

POLITECHNIKA WARSZAWSKA

Rozprawa doktorska

Dyscyplina naukowa:

Inżynieria Lądowa, Geodezja i Transport /

Dziedzina nauk: Nauki Inżynieryjno-Techniczne

mgr inż.

Artur Karol Karwel

**Orientacja wieloczasowych zdjęć lotniczych
metodami widzenia maszynowego**

Promotor

prof. dr hab. inż. **Zdzisław Kurczyński**

Promotor pomocniczy

dr hab. inż. **Jakub Markiewicz**, prof. uczelni

Warszawa 2026

Jedynemu...

Streszczenie

Archiwalne zdjęcia lotnicze stanowią unikalne źródło danych przestrzennych, umożliwiające retrospektywną analizę zmian zachodzących w środowisku naturalnym i antropogenicznym. Ich wartość informacyjna, obejmująca zarówno aspekty ilościowe, jak i jakościowe, co czyni je nieocenionym materiałem w badaniach m.in. z zakresu geoinformacji, historii urbanizacji, ochrony dziedzictwa kulturowego, a także w analizach prawnych dotyczących granic własności. Jednak skuteczne wykorzystanie tych danych wymaga ich dokładnej orientacji przestrzennej, co w przypadku materiałów archiwalnych napotyka liczne trudności – od braku metadanych, przez niską jakość radiometryczną i geometryczną po znaczne zmiany pokrycia terenu.

Tradycyjne metody orientacji, oparte na aerotriangulacji i wykorzystaniu osnowy fotogrametrycznej, okazują się często niewystarczające lub niemożliwe do zastosowania zwłaszcza w przypadku zdjęć wykonanych przed 2000 rokiem, dla których nie ma dostępnej informacji odnośnie elementów orientacji zewnętrznej. W odpowiedzi na te wyzwania, dynamiczny rozwój metod widzenia komputerowego (*Computer Vision*) oraz technik *Structure-from-Motion* (SfM) otwiera nowe możliwości automatycznego przetwarzania archiwalnych zdjęć lotniczych, bez konieczności stosowania klasycznych metod wykorzystujących manualny pomiar punktów polowej osnowy fotogrametrycznej.

W związku z tym przeprowadzono przegląd rozwiązań dotyczących metod i algorytmów umożliwiających orientację archiwalnych zdjęć lotniczych. Kluczowym problemem takiej orientacji jest automatyczna identyfikacja punktów przejściowych między zdjęciami tworzącymi blok i między blokami zdjęć pozyskanych w różnym czasie. W ramach badań eksperymentalnych przetestowano trzy główne podejścia do wyznaczania punktów wykorzystywanych w procesie orientacji serii zdjęć wieloczasowych. Pierwsze z nich, *Detect-then-Describe*, zakłada rozdzielenie etapów detekcji i opisu punktów charakterystycznych na podstawie, których wyznaczane są wspólne punkty przejściowe, drugie podejście, *Detect-and-Describe*, integruje proces detekcji i deskrypcji w jednym modelu, trzecie podejście, *Describe-then-Detect*, odwraca klasyczną kolejność działań, poprzedzając detekcję etapem deskrypcji. W ramach każdego z tych podejść przetestowano różne detektory i deskryptory.

Celem niniejszej pracy jest opracowanie i weryfikacja metodyki sekwencyjnej orientacji archiwalnych zdjęć lotniczych z wykorzystaniem algorytmów widzenia maszynowego, umożliwiającej ich dokładne georeferencjonowanie w warunkach braku osnowy fotogrametrycznej. Proponowane podejście zakłada pracę z pokrywającymi się blokami zdjęć z różnych epok, co pozwala na skuteczne przenoszenie orientacji przestrzennej ze zdjęć, posiadających znane elementy orientacji, na zdjęcia archiwalne o elementach orientacji nieznanych. Badania wykazały, że metodyka ta zachowuje skuteczność nawet w przypadku znaczących różnic czasowych pomiędzy analizowanymi materiałami, a także przy zróżnicowanej rozdzielczości przestrzennej i jakości radiometrycznej obrazów.

W ramach niniejszej pracy sformułowano szereg pytań badawczych ukierunkowanych na ocenę skuteczności i dokładności proponowanej metodyki orientacji archiwalnych zdjęć lotniczych. W pierwszej kolejności podjęto analizę wpływu rozdzielczości przestrzennej i jakości radiometrycznej zdjęć pochodzącymi z różnych epok na efektywność detekcji punktów wiążących. Następnie zbadano, które algorytmy detekcji i dopasowania punktów charakterystycznych umożliwiają uzyskanie najwyższej dokładności budowy bloków fotogrametrycznych. Ostatni aspekt badań obejmował identyfikację efektywnej/skutecznej strategii sekwencyjnego przenoszenia orientacji pomiędzy blokami zdjęć w sytuacjach charakteryzujących się znacznymi odstępami czasowymi między materiałami źródłowymi.

Przeprowadzone badania obejmowały dane lotnicze pozyskane w różnych okresach – od roku 2022, poprzez 2014, 1994 i 1986, aż do roku 1976 – co odpowiada przedziałowi czasowemu sięgającemu około 40 lat. Zastosowane zdjęcia charakteryzowały się zróżnicowaną rozdzielczością przestrzenną, różną jakością radiometryczną oraz odmiennymi parametrami rejestracji. Metodyka obejmowała dwuetapowy proces sekwencyjnej orientacji bloków zdjęć lotniczych. W pierwszym etapie wyznaczono elementy orientacji zdjęć dla każdego bloku w układzie lokalnym, a następnie wygenerowano true-ortofotomapy oraz Numeryczne Modele Pokrycia Terenu (NMPT). W drugim etapie, z wykorzystaniem wcześniej opisanych algorytmów, automatycznie określono współrzędne x , y punktów przejściowych pomiędzy true-ortofotomapami z różnych okresów oraz odpowiadające im współrzędne wysokościowe (z) na podstawie utworzonego NMPT. Uzyskany w ten sposób zbiór trójwymiarowych punktów przejściowych (x , y , z) posłużył do wyznaczenia parametrów transformacji przestrzennej pomiędzy odpowiadającymi sobie blokami czasowymi. W przeprowadzonym eksperymencie

analizie poddano czas pozyskania punktów przejściowych, dokładność ich wyznaczenia, liczbę i rozmieszczenie, a także dokładność transformacji pomiędzy orientowanymi blokami na podstawie wyznaczonych punktów przejściowych.

W ramach badań zaproponowano autorską metodykę sekwencyjnej orientacji archiwalnych bloków zdjęć lotniczych, opartą na zmodyfikowanej metodzie *Structure-from-Motion*. Podejście to polega na chronologicznym przenoszeniu orientacji pomiędzy blokami czasowymi, od najnowszych do najstarszych, przy wykorzystaniu wcześniej zorientowanego bloku referencyjnego o znanych elementach orientacji zewnętrznej. Takie podejście jest szczególnie istotne w przypadku dużych różnic czasowych między zdjęciami, gdy bezpośrednie dopasowanie najnowszych i najstarszych obrazów jest niemożliwe z powodu braku wspólnych szczegółów terenowych, wynikającego np. z intensywnych przemian urbanizacyjnych.

Eksperymenty badawcze wykazały, że zastosowane metody pozwalają na skuteczne dopasowanie i orientację zdjęć nawet w warunkach znacznych zmian pokrycia terenu, szczególnie na obszarach silnie zurbanizowanych. Zaproponowana metodyka okazała się skuteczna w procesie georeferencji archiwalnych zdjęć lotniczych, w sytuacjach, w których tradycyjne podejścia okazują się niewystarczające. Uzyskane wyniki dowodzą słuszności postawionej tezy badawczej, iż zastosowanie metod widzenia maszynowego pozwala z wysoką niezawodnością na automatyczne odtworzenie orientacji archiwalnych zdjęć lotniczych pozyskanym w znacznych odstępach czasowych (sięgających nawet 40 lat) z dokładnością odpowiadającą rozdzielczości zdjęć (wartości piksela terenowego), a w niektórych przypadkach dokładnością podpikselową.

Słowa kluczowe: automatyczna orientacja archiwalnych zdjęć lotniczych, bloki fotogrametryczne, NMPT, true-ortofotomapa, detektory punktowe, detektory powierzchniowe, deskryptory, Structure-from-Motion, Computer Vision

Abstract

Archival aerial images represent a unique source of spatial data, enabling retrospective analyses of changes in both natural and anthropogenic environments. Their informational value – encompassing both quantitative and qualitative aspects – makes them an invaluable material in studies, such as geoinformation, urbanisation history, cultural heritage preservation, and legal analyses concerning property boundaries. However, the effective use of such data requires precise spatial orientation, which in the case of archival materials poses numerous challenges, ranging from the lack of metadata and poor radiometric and geometric quality to substantial changes in land cover.

Classical orientation methods based on aerial triangulation and photogrammetric control networks often prove insufficient or impossible to apply, particularly for images acquired before the year 2000, for which no information about exterior orientation parameters is available. In response to these challenges, the rapid development of computer vision and Structure-from-Motion (SfM) techniques presents new possibilities for the automatic processing of archival aerial imagery, eliminating the need for classical approaches that rely on manual measurement of ground control points.

A review of methods and algorithms for orienting archival aerial images was conducted. The key issue in such orientation lies in the automatic identification of tie points between images within a single block and across image blocks acquired at different times. In the experimental phase, three main approaches to determining points used in the orientation of multitemporal image series were tested, along with their corresponding algorithms. The first approach, Detect-then-Describe, separates the detection and description stages used for keypoint determination, on the basis of which common tie points are determined. The second, Detect-and-Describe, integrates both processes within a single model. The third, Describe-then-Detect, reverses the traditional order by performing description before detection.

This study aims to develop and validate a methodology for sequential orientation of archival aerial images using computer vision algorithms, enabling accurate georeferencing in the absence of a photogrammetric control network. The proposed approach relies on overlapping image blocks from different epochs, allowing for effective spatial orientation transfer from images with known orientation parameters to archival images with unknown parameters. The

research demonstrated that this methodology remains effective even when there are significant temporal gaps between datasets, as well as variations in spatial resolution and radiometric quality.

A series of research questions were formulated to assess the effectiveness and accuracy of the proposed methodology for orienting archival aerial images. First, the influence of spatial resolution and radiometric quality of images from different epochs on the efficiency of tie-point detection was analysed. Next, the performance of various feature detection and matching algorithms was evaluated in terms of their ability to achieve the highest accuracy in photogrammetric block reconstruction. The final aspect concerned identifying the most effective strategies for sequential orientation transfer between image blocks acquired at significantly different times.

The conducted experiments utilised aerial data acquired across several decades, from 1976 to 1986, 1994, 2014, and 2022, covering a temporal span of approximately 40 years. The employed images varied in spatial resolution, radiometric quality, and acquisition parameters. The methodology comprised a two-stage sequential orientation process for aerial image blocks. In the first stage, image orientation parameters were determined for each block within a local coordinate system, followed by the generation of true orthophotos and Digital Surface Models (DSMs). In the second stage, using the previously described algorithms, the x and y coordinates of tie points between orthophotos from different periods were automatically identified, with corresponding elevation (z) values derived from the generated DSMs. The resulting set of three-dimensional tie points (x, y, z) was used to compute the spatial transformation parameters between corresponding temporal blocks. The experiment assessed the time required to obtain tie points, their accuracy, number and spatial distribution, as well as the accuracy of spatial transformation between oriented blocks.

As part of this research, a novel methodology for sequential orientation of archival aerial image blocks was proposed, based on a modified Structure-from-Motion approach. This method involves the chronological transfer of orientation between temporal blocks, from the most recent to the oldest, using a previously oriented reference block with known exterior orientation parameters. Such an approach is particularly relevant in cases of significant temporal gaps between images, where direct matching of the newest and oldest data is

infeasible due to the lack of common ground features, often resulting from extensive urban development.

Experimental results demonstrated that the proposed methods enable effective image matching and orientation even under conditions of significant land cover change, particularly in highly urbanised areas. The proposed methodology proved effective for georeferencing archival aerial images in situations where traditional methods are ineffective. The results confirm the research thesis that the application of computer vision methods enables reliable automatic reconstruction of the orientation of archival aerial images acquired over long temporal intervals (up to 40 years), achieving accuracy comparable to image resolution (ground pixel size), and in some cases, even sub-pixel precision.

Keywords: automatic orientation of archival aerial images, photogrammetric blocks, DSM, true-orthophotomap, keypoint detectors, blob detectors, descriptors, Structure-from-Motion, Computer Vision

Spis treści

Streszczenie.....	3
Abstract	8
Spis treści	11
Wykaz wykorzystanych skrótów	14
1. Wstęp	17
2. Motywacja i problemy badawcze podjęte w rozprawie doktorskiej	22
3. Budowa bloku fotogrametrycznego.....	31
3.1 Geometria archiwalnych zdjęć analogowych	32
3.1.1 Deformacja podłoża filmu lotniczego.....	33
3.1.2 Stabilność geometryczna filmów	34
3.1.3 Przegląd metryk kamer analogowych i generowanie metryki wirtualnej.....	36
3.2 Dokładność orientacji zdjęć lotniczych	38
3.3 Schemat pomiarowego opracowania bloku zdjęć lotniczych.....	39
3.3.1 Podejście klasyczne do orientacji zdjęć lotniczych.....	41
3.3.2 Podejście do orientacji zdjęć metodą SfM	42
3.3.3 Różnice między metodą klasyczną i SfM orientacji zdjęć lotniczych.....	44
4. Algorytmy przetwarzania obrazów cyfrowych stosowane w metodzie SfM	46
4.1 Opis wykorzystanych algorytmów CV pomiaru punktów charakterystycznych	49
4.2 Detektory punktowe i powierzchniowe (Hand-Crafted)	54
4.2.1 Detektory punktowe – FAST i BRISK	56
4.2.2 Detektory powierzchniowe – SIFT, SURF, CenSurE, ASIFT	58
4.3 Deskryptory	69

4.3.1	Deskryptor <i>Hand-Crafted</i> - SIFT	69
4.3.2	Deskryptory <i>Learned-Based</i> (HardNet, HardNet8, SOSNet, HyNet, MKDDescriptor, TFeat.....	71
4.4	Metody łączone detekcji i opisu punktów Learned-Based	87
5.	Metodyka opracowania sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych.....	94
5.1	Budowa bloków fotogrametrycznych.....	97
5.2	Generowanie NMPT i true-ortofotomap.....	106
5.3	Metodyka sekwencyjnej orientacji bloków fotogrametrycznych	108
5.4	Miary jakości wykorzystywane w ocenie dokładności orientacji zdjęć lotniczych	114
6.	Charakterystyka danych testowych wykorzystywanych w pracach eksperymentalnych	119
7.	Budowa bloków fotogrametrycznych	127
7.1	Ocena czasu detekcji punktów wiążących w bloku fotogrametrycznym	129
7.2	Ocena liczby wyznaczonych punktów wiążących i ich rozkładu w bloku fotogrametrycznym	133
7.3	Ocena dokładności pomiaru punktów wiążących na parach zdjęć w bloku	145
7.4	Ocena dokładności budowy bloku fotogrametrycznego.....	150
7.5	Wnioski i podsumowanie	160
8.	Produkty fotogrametryczne wykorzystane do sekwencyjnej orientacji bloków zdjęć	166
9.	Sekwencyjna orientacja archiwalnych bloków fotogrametrycznych	177
9.1	Ocena poprawności procesu orientacji bloków fotogrametrycznych – podejście Detect-then-Describe (algorytmy Hand-Crafted)	183
9.1.1	Ocena czasu automatycznego pomiaru punktów przejściowych w sekwencyjnej orientacji bloków fotogrametrycznych.....	184

9.1.2	Ocena liczby wykrytych punktów przejściowych i ich rozkładu w sekwencyjnej orientacji bloków fotogrametrycznych.....	186
9.1.3	Ocena dokładności wyznaczenia punktów przejściowych	202
9.1.4	Ocena dokładności orientacji bloków fotogrametrycznych	205
9.1.5	Podsumowanie wyników uzyskanych przy wykorzystaniu metody <i>Detect-then-Describe</i> i algorytmów <i>Hand-Crafted</i>	209
9.2	Ocena poprawności procesu orientacji bloków fotogrametrycznych – podejście <i>Detect-then-Describe</i> (algorytmy <i>Hand-Crafted</i> / <i>Learned-Based</i>)	213
9.3	Ocena poprawności procesu orientacji bloków fotogrametrycznych – podejście <i>Detect-and-Describe</i> (algorytmy <i>Learned-Based</i>).....	233
9.4	Ocena poprawności procesu orientacji bloków fotogrametrycznych – podejście <i>Describe-then-Detect</i> (algorytmy <i>Learned-Based</i>)	244
9.5	Wnioski i podsumowanie	256
9.5.1	Podsumowanie	262
10.	Weryfikacja przyjętej koncepcji - ocena przeniesienia orientacji bloku fotogrametrycznego z roku 2022 na blok z roku 1976 metodą sekwencyjną.....	264
11.	Podsumowanie.....	276
12.	Bibliografia	282
13.	Załączniki	298

Wykaz wykorzystanych skrótów

ABM – *Area-Based Matching*

AM – *Agisoft Metashape*

AI – *Artificial Intelligence*

API – *Application Programming Interface*

ASIFT – *Affine Scale-Invariant Feature Transform*

ASURF – *Affine Speeded Up Robust Features*

BERT – *Bidirectional Encoder Representations from Transformers*

BN – *Batch Normalization*

BRISK – *Binary Robust Invariant Scalable Keypoints*

CCD – *Charge-Coupled Device*

CenSurE – *Center Surround Extremas*

CHOP- *Check Orientation Point*

CHTP – *Check Tie Point*

CMOS – *Complementary Metal-Oxide-Semiconductor*

CNN – *Convolutional Neural Network*

CODGiK – *Centralny Ośrodek Dokumentacji Geodezyjnej i Kartograficznej*

COP – *Control Orientation Points*

CPU – *Central Processing Unit*

CS – *Central-Surround*

CTP – *Control Tie Point*

CV – *Computer Vision*

DBSCAN – *Density-Based Spatial Clustering of Applications with Noise*

DELF - *Deep Local Features*

DNN – *Deep Neural Network*

DoB – *Difference of Boxes*

DoG – *Difference of Gaussian*

DoH – *Determinant of Hessian*

ELU – *Exponential Linear Unit*

EMK – *Efficient Match Kernels*

EOZ – *Elementy Orientacji Zewnętrznej*

EuroSDR – *European Spatial Data Research*

FAST – *Features from Accelerated Segment Test*

FBM – *Feature-Based Matching*

FLANN – *Fast Library for Approximate Nearest Neighbors*

FRN – *Filter Response Normalization*

GCPs – *Ground Control Points*

GNSS-RTK – *Global Navigation Satellite Systems – Real-Time Kinematic*

GPT – *Generative Pre-trained Transformer*

GSD – *Ground Sampling Distance*

GUKiK – *Główny Urząd Geodezji i Kartografii*

HardNet – *Hard Descriptor Network*

HC – *Hand-Crafted*

HyNet – *Hybrid Network*

IDW – *Inverse Distance Weighted*

IPD – *Interest Point Decoder*

IQR – *Inter Quartile Range*

ISPRS – *International Society for Photogrammetry and Remote Sensing*

k-NN – *k-Nearest Neighbors*

LB – *Learned-Based*

LBFE – *Learned-Based - feature extraction*

LIFT – *Learning Invariant Feature Transform*

LMeds – *Least Median of Squares*

LoFTR – *Local Feature Transformer*

LoG – *Laplacian of the Gaussian*

LSD – *Line segment detector*

LSM – *Least Squares Matching*

LSM – *Least Squares Measurement*

MC – *Matching Candidates*

MKD – *Multiple-Kernel local-patch Descriptor*
MS-COCO – *Microsoft Common Objects in Context*
MVS – *Multi-View Stereo*
NC – *Number of Cross-Attention Layers*
NMPT – *Numeryczny Model Pokrycia Terenu*
NMT – *Numeryczne Modele Terenu*
PC – *Pairwise Confidence*
PCA – *Principal Component Analysis*
PROSAC – *Progressive Sample Consensus*
PZGiK – *Państwowy zasób geodezyjny i kartograficzny*
RAM – *Random Access Memory*
RANSAC – *Random Sample Consensus*
ReLU – *Rectified Linear Unit*
RGB – *Red Green Blue*
RMSD – *Root Mean Square Deviation*
RMSE – *Root Mean Square Error*
SfM – *Structure-from-Motion*
SIFT – *Scale-Invariant Feature Transform*
SLAM – *Simultaneous Localization and Mapping*
SMAD – *Sigma Median Absolute Deviation*
SOSNet – *Scale-Oriented Signature Network*
SURF – *Speeded Up Robust Features*
SVD – *Singular Value Decomposition*
TFeat – *Triplet-based Feature Extraction*
TLU – *Thresholded Linear Unit*
UAS – *Unmanned Aerial Systems*
VGG – *Visual Geometry Group*

1. Wstęp

Wiarygodne odtwarzanie informacji ilościowej i jakościowej dotyczącej danego obszaru lub obiektów już nie istniejących jest możliwe wyłącznie na podstawie archiwalnych danych obrazowych, zwłaszcza w tych przypadkach, gdzie informacje archiwalne na ich temat oparte są na znikomej dokumentacji. Z tego też powodu, niezbędne jest wieloetapowe przetworzenie archiwalnych zdjęć, w którym kluczowym aspektem jest orientacja przestrzenna danych (tak zwana georeferencja). Standardowo, wykonywane jest to w procesie aerotriangulacji, polegającej na orientacji zewnętrznej i wpasowaniu bloku zdjęć w układ terenowy w oparciu o punkty osnowy fotogrametrycznej. Dzięki temu pierwszemu, a zarazem najważniejszemu etapowi jakim jest orientacja zdjęć, możliwe jest wykonanie końcowych produktów fotogrametrycznych tj. modele 3D, Numeryczne Modele Terenu (NMT), Numeryczne Modele Pokrycia Terenu (NMPT) bądź ortofotomapy. Produkty te (oprócz ortofotomapy) są dodatkowym źródłem informacji na temat wysokości elementów terenowych, a tym samym stają się doskonałym materiałem do analiz przestrzennych z uwzględnieniem skali czasu.

Ważnym aspektem pomiarowym zdjęć historycznych jest mapowanie długoterminowych zmian pokrycia terenu odnoszących się zarówno do terenów rolniczych jak i leśnych (Kupidura, i.in., 2019; Osińska, i.in., 2019). Nie można zapominać o ich wartości informacyjnej dla wykonywania wieloczasowych analiz urbanistycznych (Markiewicz, Turek, 2014). Wykonanie orientacji przestrzennej zdjęć zgodnie z przyjętymi standardami umożliwia poprawne odtworzenie dowolnego punktu lub linii, dzięki czemu archiwalne zdjęcia lotnicze stanowią doskonałe źródło danych m.in. na temat historycznego przebiegu granic działek, a także sposobu ich zagospodarowania (Ewiak, Lulkowska, 2016). Zdjęcia lotnicze mogą również stanowić doskonałe źródło danych do modelowania hydrologicznego dolin rzecznych (Gołuch, 2003). Innym kluczowym wykorzystaniem zdjęć lotniczych, ale tym razem z niskiego pułapu jest zastosowanie ich w kontroli upraw przy użyciu bezzałogowych systemów latających UAS (*Unmanned Aerial Systems*) (Sanecki J., i.in. 2014).

Klasyczne podejście do orientacji archiwalnych zdjęć lotniczych związane jest z następującymi etapami:

- 1) orientacją wewnętrzną, która polega na odtworzeniu parametrów kamery użytej podczas wykonywania zdjęć na podstawie metryki kalibracji kamery;
- 2) orientacją wzajemną polegającą na automatycznym lub półautomatycznym pomiarze punktów wiążących, mająca na celu połączenie zdjęć w jeden spójny blok i wyznaczenie relacji geometrycznych pomiędzy poszczególnymi zdjęciami,
- 3) orientacją zewnętrzną w oparciu o fotopunkty wyszukiwane manualnie przez obserwatora i identyfikowane na obrazach archiwalnych (zeskanowanych zdjęciach analogowych) dla przyjętego układu odniesienia (państwowy geodezyjny układ odniesienia przestrzennego lub globalny układ odniesienia) (Sevara i in., 2018; Fieber i in., 2018).

W większości przypadków orientację obrazu archiwalnego wykonuje się metodami FBM (*Feature-Based Matching*) i ABM (*Area-Based Matching*) (Shen., i in., 2021). Orientacja obrazów archiwalnych z wykorzystaniem metod FBM i ABM pozwala na uzyskanie dokładności orientacji wzajemnej na poziomie około 0.5 piksela (Nocerino i in., 2012; Kapnias i in., 2008; Sauerbier i in., 2004; Micheletti i in., 2015).

Znaczącym ograniczeniem w przeprowadzeniu orientacji zdjęć jest brak lub niekompletność metadanych dotyczących parametrów orientacji wewnętrznej kamery, orientacji kamery w przestrzeni (elementy orientacji zewnętrznej kamery) oraz osnowy fotogrametrycznej (Nocerino i in., 2012, Nurminen i in., 2015, Salach, 2017, Poli i in., 2020). Podanie przybliżonych wartości tych parametrów może skutkować obniżeniem dokładności opracowania fotogrametrycznego. Do czynników determinujących jakość produktów fotogrametrycznych tworzonych na podstawie archiwalnych zdjęć lotniczych zaliczamy również niekompletność zdjęć w bloku fotogrametrycznym. Brak tych danych powoduje braki w opracowaniu produktów takich jak aerotriangulacja, NMPT, czy ortofotomapa. W dostępnych zbiorach archiwalnych zdjęć lotniczych zdarzają się również bloki, w których pokrycie podłużne czy poprzeczne zdjęć spada poniżej dopuszczalnego standardu (60% dla pokrycia podłużnego, 30% dla poprzecznego), co znacznie pogarsza lub uniemożliwia wykonanie orientacji bloku fotogrametrycznego.

Kolejnym czynnikiem determinującym dokładność orientacji zdjęć lotniczych jest ich niska jakość geometryczna i radiometryczna, spowodowana m.in. procesem skanowania

archiwalnych danych. Zdjęcia wykonane przeszło dekadę temu na emulsji światłoczułej, ze względu na długi okres¹ ich przechowywania ulegają skurczom podłoża, co przyczynia się do lokalnych błędów położenia elementów sytuacyjnych na zdjęciu. Złe przechowywanie zdjęć archiwalnych może spowodować również ich wyblaknięcie, co zasadniczo obniża ich jakość radiometryczną. Nie należy zapominać, że sam proces digitalizacji archiwalnych zdjęć ma duży wpływ na jakość opracowania. Złe zdefiniowane parametry skanowania czy użycie niefotogrametrycznych skanerów obniżają jakość geometryczną zdjęć, co przekłada się bezpośrednio na proces ich orientacji.

W przypadku zdjęć obarczonych wymienionymi czynnikami, zastosowanie klasycznych metod orientacji zdjęć, wykorzystujących pomiar oparty na algorytmach ABM (Babbar i in., 2010; Poli i in., 2020), jest znacznie utrudnione lub całkowicie niemożliwe do realizacji.

W klasycznym podejściu orientacja przestrzenna archiwalnych zdjęć lotniczych wymaga pomiaru terenowego punktów osnowy fotogrametrycznej (przy wykorzystaniu pomiaru geodezyjnego GNSS-RTK (*Global Navigation Satellite Systems – Real-Time Kinematic*), tachimetrii, lub pomiaru kameralnego (bazę stanowią współczesne ortofotomapy, mapy topograficzne i NMT). W przypadku materiałów archiwalnych odtworzenie osnowy terenowej w sposób bezpośredni (ze względu na długi odstęp czasu od daty ich wykonania) jest ograniczone lub często niemożliwe do wykonania. Znaczące zmiany urbanizacyjne zachodzące na danym obszarze są na tyle duże, że odnalezienie wystarczającej liczby punktów osnowy fotogrametrycznej staje się problematyczne, a w wielu przypadkach wręcz niewykonalne. Należy dodać, że dla większości danych archiwalnych dostępnych w państwowym zasobie geodezyjnym GUKiK (Główny Urząd Geodezji i Kartografii) odnoszących się do zdjęć lotniczych brak jest polowych osnów fotogrametrycznych, które umożliwiałyby ich prostą i bezpośrednią orientację. Dopiero wraz z początkiem okresu transformacji technologicznej (po roku 2000) wykonawcy zaczęli przekazywać wraz ze zdjęciami lotniczymi polową osnowę fotogrametryczną, jednak i ona nie zawsze była kompletna.

Mając na uwadze wszystkie wspomniane powyżej aspekty determinujące dokładność orientacji oraz braki osnowy fotogrametrycznej, można postawić pytanie czy istnieje metoda

¹ Znaczenie słowa okres związane z czasem pozyskania zdjęć lotniczych używanego w rozprawie doktorskiej odnosi się również do określenia epoki.

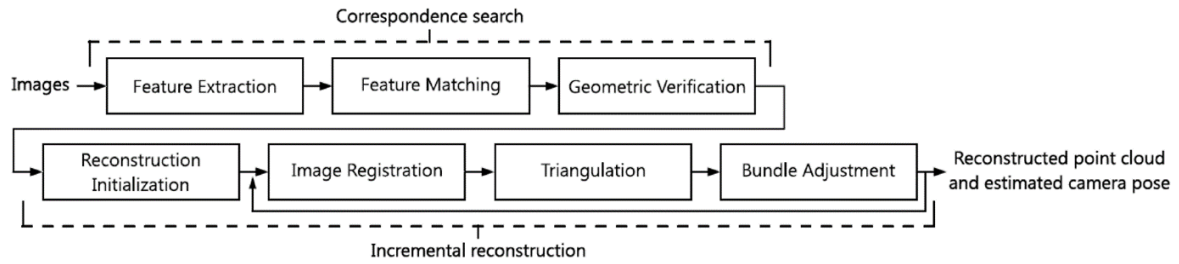
pomiaru pozwalająca na poprawne odtworzenie elementów orientacji zewnętrznej archiwalnych zdjęć lotniczych bez udziału fotogrametrycznej osnowy pomiarowej rozumianej jako osnowa oparta na klasycznych fotopunktach?

Analiza literaturowa dotycząca orientacji zdjęć pozyskiwanych z pułapu lotniczego i bliskiego zasięgu wskazuje, że fotogrametria lotnicza opiera się na jednym powielanym schemacie (wynika to z powtarzalnego układu relacji zdjęcia-obiekt), a w fotogrametrii bliskiego zasięgu mamy do czynienia z mnogością przypadków (zdjęcia wykonane z różnych pozycji, w różnych skalach, pod różnymi kątami itd.) (Habib i Morgan, 2003; Luhmann et al., 2013; Remondino i Fraser, 2006). Można przyjąć, że zdjęcia lotnicze są uogólnieniem przypadku zdjęć naziemnych. Ta mnogość przypadków geometrii zdjęć naziemnych sprzyjała rozwojowi metod ich orientacji bazujących na algorytmach widzenia maszynowego CV (*Computer Vision*), a w szczególności na podejściu SfM (*Structure-from-Motion*), głównie z wykorzystaniem algorytmów FBM. Zdjęcia lotnicze, ze względu na prostą i powtarzalną geometrię, przetwarzane są według ustandaryzowanego schematu, z wykorzystaniem głównie algorytmu ABM. Algorytm dopasowania FBM w przypadku zdjęć lotniczych służył wyłącznie do wyznaczenia ich przybliżonego położenia.

Jedną z metod widzenia maszynowego odnoszącą się do grup metod dopasowania obrazów jest metoda *Structure-from-Motion* polegająca na automatycznej orientacji wzajemnej² grupy zdjęć (metoda *Incremental* – patrz podrozdział 5.1) z jednoczesną samokalibracją w oparciu o punkty wykryte niezależnie na każdym zdjęciu i dopasowane wzajemnie (Schönberger i Frahm, 2016). Podstawą SfM jest rekonstrukcja trójwymiarowej struktury na podstawie serii obrazów uzyskanych z różnych pozycji oraz wyznaczenie relacji geometrycznych pomiędzy nimi. Metoda ta bazuje na detekcji punktów wiążących w oparciu o punkty wykryte na każdym zdjęciu charakteryzujące się niezależnością od skali, obrotu bądź zniekształceń geometrycznych. W tym celu wykorzystywane są algorytmy analizujące zmiany gradientów stopni szarości. Takie podejście pozwala na wykrycie inwariantnych punktów charakterystycznych, odpornych na zmianę skali, skręcenia i pochylenia obrazów oraz częściowo niezmiennie na zmiany oświetlenia i odporne na lokalne zniekształcenia

² Orientacja wzajemna zdjęć w bloku fotogrametrycznym w podejściu *Structure-from-Motion* jest realizowana w ramach podejścia *Incremental* (por. podrozdział 5.1)

geometryczne. Na podstawie ww. punktów wiążących przeprowadzany jest proces orientacji wzajemnej grupy zdjęć oraz wyrównanie metodą niezależnych wiązek. Ogólnie metodę SfM można podzielić na dwa główne etapy tj. detekcja wraz z opisem punktów charakterystycznych na pojedynczych zdjęciach i wyznaczenie zależności między punktami (kroki od 1 do 3) oraz iteracyjne wyznaczenie elementów orientacji zdjęć (kroki od 4 do 6) (rys.1.1).



Rys. 1.1 Schemat orientacji zdjęć metodą SfM (Bianco i in., 2018)

2. Motywacja i problemy badawcze podjęte w rozprawie doktorskiej

Motywacja podjęcia tematyki badawczej

Archiwalne zdjęcia lotnicze ze względu na swój bogaty potencjał informacyjny mają szeroki wachlarz zastosowań naukowych i praktycznych do których zalicza się m.in.:

- odtwarzanie danych topograficznych (generowanie historycznych modeli terenu),
- analizy zmian użytkowania ziemi (śledzenie zmian praktyk rolniczych i wpływu na jakość gleby),
- monitorowanie zmian środowiskowych (śledzenie zmian użytkowania ziemi, ocena degradacji ekosystemów, monitorowanie obszarów chronionych),
- śledzenie zmian w środowisku naturalnym i antropogenicznym (rekonstrukcja zasięgu lodowców i pokrywy śnieżnej, analiza przebiegu linii brzegowej),
- lokalizacja potencjalnych stanowisk archeologicznych (wykrywanie śladów dawnych osad),
- badania historyczne (badanie ewolucji miast i ich rozwoju przestrzennego, tworzeniu historycznych baz danych i map),
- planowanie przestrzenne (rekonstrukcja historii urbanistycznej miast, planowanie rozwoju miast),
- zastosowania prawne (rozwiązywanie sporów dotyczących granic własności działek).

Analiza literatury dotyczącej wykorzystania archiwalnych zdjęć lotniczych wskazuje na szerokie spektrum badań obejmujących różnorodne aspekty ich zastosowania. W badaniach tych wykorzystywane są zarówno metody klasyczne wykonywania aerotriangulacji, generowania modeli NMPT i procesu tworzenia ortoobrazów, jak również zintegrowane metody SfM i MVS (*Multi-View Stereo*) odnoszących się do rekonstrukcji kształtu czy odtwarzaniu trójwymiarowych obiektów dziedzictwa kulturowego oraz jako uzupełnienie danych naziemnych (Wiedemann i.in., 2000; Sauerbier i.in., 2004; Sonnemann i.in., 2006; Verhoeven i.in., 2013; Risbol i.in., 2014; Peppas i.in., 2018; Bolognesi i in., 2014).

Ponadto zdjęcia lotnicze mają zastosowanie w analizach wieloczasowych, detekcji i monitoringu zmian krajobrazu (Walstra i.in., 2004; Jao i.in., 2014; Nebiker i.in., 2014; Micheletti

i.in., 2015; Popelková i Mulková, 2016; Lydersen i.in., 2018; Fieber i.in., 2018; Sevara i.in., 2018; Osińska i.in., 2019; Kupidura i.in., 2019; Pinto i.in., 2019; Poli i.in., 2020; Grottoli i.in., 2020).

Inne badania odnosiły się do wyznaczania parametrów orientacji zewnętrznej kamery (EOZ) oraz na automatycznych procedurach georeferencjonowania archiwalnych zdjęć wykorzystując podejście klasyczne (Kim i.in., 2010; Redweik i.in., 2010; Verhoeven i.in., 2012; Karel i.in., 2013; Nurminen i.in., 2015; Giordano i.in., 2018) lub podejście odnoszące się do transformacji ich do referencyjnych zdjęć satelitarnych (Zambanini i Sablatnig, 2017).

Archiwalne zdjęcia lotnicze często były wykorzystywane do interpretacji terenu i odszukiwania utraconych informacji (Sevara, 2013), ale również do użycia procesów symulacyjnych wspierających interpretację danych (Siok i Ewiak, 2020). Zdjęcia lotnicze służyły również do identyfikacji kraterów bombowych związanych z II wojną światową (Meixner i Eckstein, 2016; Valjavec i.in., 2018; Clermont i.in., 2019; Dolejš i.in., 2020).

Inne badania z wykorzystaniem źródeł archiwalnych odnosiły się do produkcji historycznych map pokrycia terenu i użytkowania ziemi, wykorzystując również algorytmy sztucznej inteligencji AI (*Artificial Intelligence*) (Ratajczak i.in., 2019; Mboga i.in., 2020).

Wieloczasowe zbiory danych lotniczych mogą być orientowane z wykorzystaniem podejścia opartego na klasycznie projektowanych cechach (metoda *Hand-Crafted* -HC). W tym celu stosuje się liniowe operatory detekcji, takie jak LSD (*Line Segment Detector*), które umożliwiają ekstrakcję cech liniowych z obrazów. Następnie wyznaczone linie dopasowywane są do referencyjnych danych topograficznych, co pozwala na georeferencjonowanie zdjęć (Cléry i in., 2014).

Przedstawione powyżej podejścia zazwyczaj zawodzą w ekstrakcji wiarygodnych odpowiadających sobie cech, gdy istnieją duże różnice czasowe między danymi lub pozyskane są w różnych skalach. Z tych powodów ostatnie prace badawcze opierały się na podejściu z wykorzystaniem sieci neuronowych (*Learned-Based* – LB), w celu wyodrębnienia wiarygodnych odpowiadających sobie punktów między wieloczasowymi obrazami oraz poprawy ich dopasowania (Ressl i.in., 2020; Zhang i.in., 2021).

Z powodu tak licznych zastosowań archiwalnych zdjęć lotniczych niezbędne są metadane odnoszące się do ich georeferencji i orientacji wewnętrznej, dzięki którym możliwe jest

wyznaczenie położenia obszaru lub elementu sytuacyjnego związanego z danym zjawiskiem w przyjętym układzie odniesienia. Odnosząc się do swojego doświadczenia zawodowego zdobytego podczas realizacji prac fotogrametrycznych takich jak ekspertyzy sądowe zlecane przez Sądy Cywilne, a także Prokuraturę Krajową dotyczące m.in. wyznaczenia stanu zagospodarowania działek ewidencyjnych w określonej przestrzeni czasowej, czy też wyznaczenie przebiegu granicy użytkowania pomiędzy działkami na przestrzeni lat w oparciu o dane archiwalne, natknąłem się na trudności związane z wyznaczeniem poprawnego ich przebiegu na skutek niewystarczającej dokładności orientacji zdjęć spowodowanej słabą jakością punktów osnowy fotogrametrycznej.

Na problem jednoznacznej identyfikacji osnowy fotogrametrycznej wpływają zarówno intensywne procesy urbanizacyjne i suburbanizacyjne zachodzące na danym obszarze, co przekłada się na małą liczbę potencjalnych punktów wiążących, jak również niższa dokładność pomiaru manualnego punktów. W bazie danych Państwowego Zasobu Geodezyjnego i Kartograficznego (PZGiK), obejmującej zobrazowania lotnicze, satelitarne, ortofotomapy oraz numeryczny model terenu (NMT), dla większości zdjęć pochodzących z nalotów fotogrametrycznych wykonanych przed 2000 rokiem brakowało istotnych danych. W szczególności nie zawierała ona danych dotyczących projektów aerotriangulacji, parametrów wykorzystywanych kamer, elementów orientacji zewnętrznej ani współrzędnych punktów osnowy fotogrametrycznej wraz ze szkicami topograficznymi. Brak tych danych uniemożliwia pełne odtworzenie orientacji przestrzennej zdjęć, co znacząco ogranicza możliwość ich ponownego pomiarowego wykorzystania. Dlatego największym wyzwaniem dla tego typu prac była identyfikacja punktów osnowy, które ze względu na zmiany urbanizacyjne i suburbanizacyjne danego terenu w funkcji czasu były pracochłonne do odtworzenia, albo ich pomiar wiązał się z dużo niższą dokładnością opracowania.

Wobec braku wystarczającej liczby punktów osnowy fotogrametrycznej potrzebnych do orientacji archiwalnych zdjęć lotniczych wykorzystywano zdjęcia z pośrednimi datami ich pozyskania, dla których określona była orientacja (zdjęcia referencyjne). Proces ten polegał na manualnym odnajdywaniu punktów wspólnych pomiędzy zdjęciami referencyjnymi, a zdjęciami poddawanymi orientacji. Następnie, punktom tym przypisywano współrzędne pochodzące ze zdjęć referencyjnych, co umożliwiało wyznaczenie orientacji przestrzennej

zdjęć archiwalnych (Ewiak, Brodowska, 2012). Były to początki hierarchicznej orientacji archiwalnych zdjęć lotniczych. Oczywiście zdarzały się zdjęcia pośrednie o małej skali w stosunku do zdjęć orientowanych i często uzyskiwana dokładność orientacji była niezadawalająca i nie spełniała kryteriów związanych np. z dokładnością identyfikacji przebiegu granic działek ewidencyjnych.

W związku z zaistniałymi problemami powstało pytanie czy istnieje metoda automatycznego pomiaru, która umożliwi dokładną orientację zdjęć lotniczych. Sięgając ponad 20 lat wstecz, przy ówczesnym stanie wiedzy technologicznej, nie można było znaleźć jednoznacznej odpowiedzi na to pytanie. Badania prowadzone na Politechnice Warszawskiej przez dr. inż. Ryszarda Preussa związane z aerotriangulacją wieloczasową (Preuss, 2002) odnosiły się do zdjęć o podobnej skali, krótkim odstępie czasu ich pozyskania (małe zmiany pokrycia terenu i podobna radiometria zdjęć). Badania te, nie uwzględniały długich odstępów czasowych pomiędzy nalotami, ze względu na duże zmiany radiometryczne danych i ograniczeń stosowanych algorytmów. Również proces orientacji zdjęć nie był w pełni automatyczny. Wraz z upływem lat rozwijały się badania, szczególnie w dziedzinie widzenia komputerowego, dotyczące automatycznego wyznaczania punktów wiążących na podstawie analizy cech obrazowych (Tuytelaars i Mikołajczyk, 2007). Rozwój algorytmów w tej dziedzinie przyczynił się do opracowania metod, które umożliwiają wyznaczanie punktów o wysokiej stabilności, jednoznaczności i niezmienności. Dzięki temu punkty te mogą być wykrywane nawet na obrazach różniących się znacznie radiometrią. Takie osiągnięcia pozwoliły sformułować wniosek, że metody pomiaru punktów, pierwotnie stosowane głównie w fotogrametrii bliskiego zasięgu, mogą być z powodzeniem zaadaptowane do zastosowań w fotogrametrii lotniczej.

Biorąc pod uwagę metody i algorytmy wykorzystywane w procesie orientacji zdjęć naziemnych, w szczególności na etapie orientacji wzajemnej realizowanej poprzez dopasowanie cech między zdjęciami (Byrne, 2016; Morgan i Brogan, 2016; Friedt, 2014), można stwierdzić, że stopień ich rozwoju znacznie przewyższył rozwój metod orientacji zdjęć lotniczych (Forstner, 1986). Współczesne podejścia wykorzystują automatyczne wyznaczanie punktów wiążących między zdjęciami za pomocą algorytmów typu HC oraz LB, będących kluczowym elementem metod *Structure-from-Motion* (SfM) (rozdział 5). Pierwsze krajowe

badania dotyczące orientacji archiwalnych zdjęć lotniczych przeprowadził zespół badawczy Pawlik i Mikrut w 2007 r. w Akademii Górniczo-Hutniczej w Krakowie (AGH). Orientacja zdjęć obejmowała materiały pozyskane w różnych okresach, z różnych obszarów oraz wykonane w różnych skalach. Badania koncentrowały się głównie na wykorzystaniu metody FBM do wyznaczania punktów wiążących w danym bloku fotogrametrycznym, ze szczególnym uwzględnieniem algorytmów SIFT (*Scale-Invariant Feature Transform*), Harris oraz metody Beaudeta (Pawlik i Mikrut, 2007).

Analizując dostępną światową literaturę z ostatnich lat można zauważyć, że pojawiły się badania z zakresu orientacji archiwalnych zdjęć lotniczych wykorzystujące metody ich automatycznego dopasowania. W artykule (Zhang i.in., 2021) zespołu badawczego Lulin Zhang, Ewelina Rupnik, Marc Pierrot-Deseilligny (Francja) pod tytułem „Feature matching for multi-epoch historical aerial images” autorzy przedstawiają podejście automatycznej orientacji archiwalnych zdjęć lotniczych. Wykorzystują oni względne orientacje międzyepokowe poprzez tworzenie numerycznych modeli pokrycia terenu NMPT (Numeryczny Model Pokrycia Terenu) każdej epoki i włączenie ich w proces dopasowywania od ogólnego do dokładnego. Metoda składa się z dopasowania międzyepokowych NMPT w celu przybliżonej orientacji (tj. przekształcenie Helmerta 3D), a następnie końcowego dopasowania cech międzyepokowych za pomocą oryginalnych obrazów RGB (*Red Green Blue*). Do procesu wyznaczania punktów wiążących (punktów charakterystycznych) wykorzystano dwa algorytmy: SuperGlue i SIFT. Uzyskano wyniki orientacji danych mieszczące się od 1m do kilku metrów w zależności od zastosowanego podejścia detekcji wspólnych punktów między epokami. Drugi zespół badawczy: Friedrich Knuth, David Shean, Shashank Bhushan, Eli Schwat, Oleg Alexandrov, Christopher McNeil, Amaury Dehecq, Caitlyn Florentine, Shad O’Neel (USA, Francja) również skupił się na wyznaczeniu punktów wiążących między poszczególnymi modelami NMPT badanych epok, z tym, że wspólne dopasowania modeli wykonano w oprogramowaniu Agisoft (Knuth i.in., 2023). Trzeci zespół badawczy: Camillo Ressler, Wilfried Karel, Livia Piermattei, Gerhard Puercher, Markus Hollaus i Norbert Pfeifer (Austria, Niemcy) zaproponował metodę wspólnej orientacji zdjęć ze wszystkich epok polegającą na odnalezieniu wspólnej korespondencji punktów charakterystycznych (wiązących) pomiędzy zdjęciami z każdej epoki za pomocą algorytmu SIFT (Ressler i.in., 2020).

Odnosząc się do wspomnianych podejść orientacji archiwalnych zdjęć lotniczych i wykonując analizę literaturową odnośnie wykorzystania metod widzenia maszynowego typu klasycznego *Hand-Crafted* (HC) czy sieciowego *Learned-Based* (LB) w detekcji punktów wiążących pomiędzy obrazami zaproponowałem własną metodę orientacji danych wieloczasowych. Przedstawiona metoda opisana w rozprawie doktorskiej różni się od przytoczonych powyżej metod. Zasadniczą różnicą proponowanej metody jest to, że odnosi się ona do całego bloku, a nie pojedynczych zdjęć i ma charakter globalny. Podejście takie eliminuje problem słabej radiometrii pojedynczych zdjęć, zwiększa liczbę wykrytych punktów wiążących pomiędzy blokami oraz zapewnia ich poprawne, równomierne rozmieszczenie w bloku. Istotną różnicą metody jest również podział bloków na mniejsze części w celu detekcji większej liczby punktów przejściowych (patrz rozdział 5). Zaproponowana metodyka orientacji archiwalnych zdjęć lotniczych składa się z dwóch części: podejścia klasycznego SfM (*Incremental* – patrz podrozdział 5.1) przy wykorzystaniu metody niezależnych wiązek i SfM globalnego (orientacja bezwzględna). W pierwszym przypadku na podstawie zdjęć i wyznaczonych punktów wiążących (charakterystycznych) pozyskanych przy wykorzystaniu metod typu *Hand-Crafted* tworzone są bloki zdjęć dla poszczególnych ich epok, z których następnie otrzymuje się modele NMPT oraz true-ortofotomapy. W drugiej części przy wykorzystaniu metody *Hand-Crafted* i *Learned-Based* odnajdywane są wspólne punkty charakterystyczne między true-ortofotomapaami poszczególnych okresów na podstawie ich dopasowań. Mając wyznaczone punkty (x,y) pomiędzy true-ortofotomapą referencyjną, a orientowaną i dysponując modelem NMPT można wyznaczyć współrzędną (z) punktów przejściowych stanowiące podstawę transformacji 3D bloku orientowanego do bloku referencyjnego w układzie lokalnym (xyz) lub terenowym (XYZ).

Automatyczne przetwarzanie archiwalnych zdjęć lotniczych wciąż stanowi istotne wyzwanie, szczególnie w przypadku braku informacji o ich georeferencji, niskiej jakości radiometrycznej obrazu, ograniczonego pokrycia stereoskopowego lub znaczących zmian pokrycia terenu. Jednak dzięki najnowszym osiągnięciom w dziedzinie fotogrametrii oraz algorytmów sztucznej inteligencji (AI) stosowanych do przetwarzania i analizy obrazów, możliwe staje się realizowanie coraz bardziej zaawansowanych rozwiązań. Potencjał badawczy archiwalnych zdjęć lotniczych jest na tyle duży, że europejska organizacja EuroSDR (*European*

Spatial Data Research) utworzyła specjalny *Benchmark Time* (<https://www.eurosdrr.net/node/1657>), w którym zgromadzono bloki zdjęć lotniczych (serie czasowe) wykonanych od lat 50 XX wieku pochodzących z różnych państw wspólnoty europejskiej, zarejestrowane nad historycznymi obszarami miejskimi lub krajobrazowymi. W celu wsparcia fotogrametrycznego przetwarzania zdigitalizowanych zdjęć, dostarczane są dodatkowe dane zawierające informacje o misjach lotniczych, używanych kamerach i punktach osnowy fotogrametrycznej (GCPs – *Ground Control Points*). Celem utworzonego benchmarku jest wykonywanie różnorodnych badań i analiz przestrzennych. Główne zastosowania *Benchmark Time* to: detekcja zmian w środowisku przestrzennym, ocena dokładności i efektywności algorytmów, wsparcie dla modernizacji procesów kartograficznych, walidacja algorytmów sztucznej inteligencji, testowanie formatów danych i standardów. Część zdjęć zawartych w zbiorze *Benchmark Time* dotyczy obszaru Warszawy i została wykorzystana w niniejszych badaniach.

W oparciu o analizę literaturową odnoszącą się do archiwalnych zdjęć lotniczych i związanych z nimi licznymi trudnościami odnośnie ich orientacji przestrzennej sformułowano następujące problemy badawcze:

1. Brak jest w zasobie GUGiK punktów osnowy czy elementów orientacji zewnętrznej odnoszących się do zdjęć pozyskanych przed 2000 rokiem, co wymusza zastosowanie metod orientacji bazujących na punktach ręcznie wykrywanych na danych kartograficznych, charakteryzujących się różną skalą i dokładnością, np. mapach topograficznych, ortofotomapach.
2. Duże zmiany pokrycia terenu na danym obszarze opracowania skutkują wykryciem przez obserwatora małej liczby punktów osnowy fotogrametrycznej lub całkowitym ich brakiem, co obniża dokładność orientacji zdjęć lub uniemożliwia przeprowadzenie tego procesu.
3. Niska jakość geometryczna i radiometryczna zdjęcia spowodowana m.in. błędnym skanowaniem, deformacją podłoża i złym przechowywaniem zdjęć, co może stanowić problem kartometrycznego opracowania zdjęć archiwalnych.
4. Odstąpienie od klasycznego podejścia (fotopunkty, metryki, dobra jakość radiometryczna itp.) na rzecz metod widzenia maszynowego (CV) dla znajdowania

punktów przejściowych (punktów wspólnych) w obrębie danego kompletu zdjęć oraz znajdowania punktów wspólnych między epokami.

Problemy badawcze podjęte w pracy

1. Jak na proces pomiaru punktów pomiędzy zbiorami wielu epok wpływa różna radiometria zdjęć, różna ich skala?
2. Jak zmiany czasowe pomiędzy danymi wpływają na liczbę i rozkład punktów wspólnych na zdjęciach?
3. Jaki jest wpływ pomierzonych punktów przejściowych w oparciu o zastosowane różne metody pomiaru na dokładność orientacji danych?
4. Jakie należy zastosować algorytmy detekcji punktów charakterystycznych (bazujące na detektorach punktowych czy powierzchniowych)? Ma to kluczowe znaczenie w orientacji archiwalnych zdjęć wieloczasowych charakteryzujących się różną jakością radiometryczną, geometryczną oraz posiadających zniekształcenia powstałe w wyniku starzenia się filmu lotniczego?
5. Podejście *Hand-Crafted*, a *Learned-Based* – czy i kiedy zasadne jest stosowanie klasycznych metod identyfikacji i dopasowania punktów wiążących, a kiedy zalecane jest wykorzystanie metod opartych na sieciach neuronowych (wykorzystujących modele sztucznej inteligencji)?

Teza rozprawy

Na podstawie stwierdzonych problemów, można postawić pytanie czy istnieje metoda pomiaru pozwalająca na poprawne odtworzenie elementów orientacji zewnętrznej archiwalnych zdjęć lotniczych bez udziału fotogrametrycznej osnowy pomiarowej, rozumianej jako osnowy opartej na klasycznych fotopunktach?

W odpowiedzi na powyższe pytanie zostaje postawiona następująca teza:

Metody widzenia maszynowego pozwalają z wysoką niezawodnością na automatyczne odtworzenie orientacji archiwalnych zdjęć lotniczych z dokładnością na poziomie rozdzielczości zdjęć (piksela terenowego).

Z tego powodu krytycznym etapem staje się przeprowadzenie procesu dopasowania zdjęć z różnych okresów czasowych w oparciu o zmodyfikowaną metodę SfM wykorzystującą algorytmy widzenia maszynowego do wykrywania punktów dopasowania, co przekłada się na poprawne wykonanie orientacji bloków zdjęć.

Celem prac podjętych w ramach rozprawy doktorskiej jest wypracowanie automatycznej metody sekwencyjnej orientacji archiwalnych zdjęć lotniczych, pozwalającej na uczynienie z nich wartościowego, kartometrycznego źródła danych przestrzennych. W celu wypracowania tej metody wykorzystane zostały algorytmy detekcji punktów charakterystycznych (*keypoints*) przy zastosowaniu algorytmów widzenia maszynowego. W podejściu SfM na podstawie wykrytych punktów charakterystycznych wyznaczane są punkty wiążące wykorzystywane do orientacji wzajemnej zdjęć w bloku. Istnieje jednak możliwość wykorzystania metod widzenia komputerowego do znalezienia punktów wspólnych pomiędzy zdjęciami (lub blokami zdjęć) pochodzącymi z różnych okresów czasowych, a następnie wykonania transformacji jednego bloku fotogrametrycznego względem drugiego. W związku z tym, celem pracy jest znalezienie wzajemnych relacji pomiędzy zdjęciami archiwalnymi z różnych okresów. Dobór odpowiednich algorytmów widzenia maszynowego pozwala na w pełni automatyczną orientację zdjęć.

Przedstawione w rozprawie doktorskiej rozwiązanie orientacji zdjęć lotniczych zmodyfikowaną metodą SfM z wykorzystaniem algorytmów widzenia maszynowego charakteryzuje się:

- niezmiennością wykrytych punktów od zniekształceń radiometrycznych i geometrycznych;
- stabilnością wykrytych punktów, niezależnie od skali czy różnej radiometrii;
- unikalnością wykrytych punktów charakterystycznych jednoznacznie definiowanych na podstawie lokalnych zmian gradientu stopni szarości obrazu.

Zapewnia to prawidłową i wiarygodną orientację bloku obrazów, wysoką „sztywność geometrii bloku”, a także odporność na pomyłki (obserwacje odstające).

3. Budowa bloku fotogrametrycznego

Budowa bloku fotogrametrycznego to proces przetwarzania obrazów mający na celu określenie geometrycznych zależności zarówno pomiędzy poszczególnymi zdjęciami w bloku, jak i pomiędzy zdjęciami a referencyjnym układem współrzędnych. W proponowanej metodyce budowy bloku proces ten obejmuje zastosowanie zaawansowanych metod i algorytmów umożliwiających ekstrakcję informacji zawartych w danych obrazowych. Kluczowym elementem tego procesu jest identyfikacja wspólnych cech, którymi mogą być punkty, linie czy obszary występujące na odpowiadających sobie zdjęciach. W przypadku zdjęć lotniczych szczególną rolę odgrywają wspólne punkty wiążące, które umożliwiają wzajemne powiązanie zdjęć oraz określenie ich orientacji przestrzennej w układzie lokalnym. Powiązane wzajemnie ze sobą zdjęcia tworzą spójny geometrycznie blok spełniający założenia tzw. bryły sztywnej. Następnie, z wykorzystaniem siedmioparametrowej transformacji konforemnej, przeprowadza się wpasowanie utworzonego bloku zdjęć do układu przestrzennego, opierając się na punktach osnowy fotogrametrycznej wyznaczonych w referencyjnym układzie odniesienia.

Obecnie na rynku geoinformacyjnym do generowania produktów fotogrametrycznych takich jak aerotriangulacja, numeryczny model terenu, czy ortofotomapa w oparciu o zdjęcia lotnicze wykorzystywane są głównie programy komercyjne takie jak ImageStation (Hexagon) bądź Inpho (Trimble). Do niedawna, na etapie wyznaczenia punktów wiążących w bloku zdjęć, powyższe programy bazowały na algorytmach ABM (*Area-Based Matching*), lub na połączonym podejściu ABM-FBM (*Area-Based - Feature-Based Matching*). Podejście ABM-FBM jest podejściem dwuetapowym, w którym na niższych poziomach piramidy obrazu punkty wiążące są wykrywane naprzemiennie z wykorzystaniem algorytmów FBM oraz ABM z zastosowaniem metody najmniejszych kwadratów LSM (*Least Squares Matching*). Końcowe dopasowanie punktów wykonywane jest na zdjęciu w pełnej rozdzielczości przy użyciu podejścia ABM-LSM, co zapewnia wysoką dokładność dopasowania i zgodność geometryczną. Wraz z rozwojem algorytmów FBM opartych na analizie zmian gradientów stopni szarości w obrębie całego zdjęcia, na rynku pojawiły się również konkurencyjne zarówno cenowo jak i jakościowo rozwiązania w stosunku do wyżej wymienionych programów fotogrametrycznych takie jak Agisoft Metashape (Agisoft LLC), Pix4D (Pix4D S.A.) czy InphoUAVMaster (Trimble), jak również darmowy odpowiedniki tj. np. COLMAP (ETH Zurich, University of North Carolina at Chapel Hill).

Programy te odmiennie od wcześniej wspomnianych wykorzystują do pomiaru punktów wiążących podejście *Structure-from-Motion* (SfM) bazujące jedynie na algorytmach FBM, a które pozwalają na:

- uzyskiwanie wyższych dokładności dopasowania obrazów w porównaniu do metody ABM,
- możliwość łączenia zdjęć pozyskiwanych z różnych odległości, pod różnymi kątami,
- orientację zdjęć (budowa bloku) charakteryzujących się różną jakością radiometryczną.

3.1 Geometria archiwalnych zdjęć analogowych

Głównym czynnikiem wpływającym na jakość geometryczną i radiometryczną archiwalnych analogowych metrycznych (fotogrametrycznych) zdjęć lotniczych jest niewłaściwy sposób ich przechowywania (brak odpowiednich zabezpieczeń) w długim okresie. Do tych czynników należy również zaliczyć czynnik atmosferyczny taki jak światło słoneczne lub wilgoć, co może przełożyć się min. na wyblaknięcia, uszkodzenia mechaniczne (pęknięcia materiału fotograficznego), bądź zmiany kształtu na skutek deformacji podłoża emulsji fotograficznej. Deformacje podłoża (skurcze) przyczyniają się do występowania tzw. afinizmu zdjęć, czyli błędów geometrycznych, powodujących naruszenie fundamentalnego założenia, że zdjęcie jest pod względem geometrycznym rzutem środkowym.

Rozważając zastosowanie archiwalnych zdjęć lotniczych do celów pomiarowych, istotne jest uwzględnienie metadanych przypisanych do tych zdjęć. Metadane te zawierają informacje o parametrach metrycznych kamery, czasie wykonania zdjęć oraz elementach orientacji zewnętrznej. W kontekście parametrów metrycznych kamery, czyli elementów orientacji wewnętrznej zdjęć lotniczych, zalicza się do nich skalibrowaną wartość ogniskowej kamery, położenie punktu głównego obrazu, położenie znaczków tłowych (charakterystyczne dla zdjęć analogowych), dystorsję obiektywu (również odnoszącą się głównie do zdjęć analogowych). Te parametry są zazwyczaj zawarte w dokumencie określanym jako metryka kamery. W sytuacji, gdy metryka kamery nie jest dostępna, niektóre informacje, takie jak wartość ogniskowej czy wysokość lotu, można często odczytać z elementów pozaramkowych zdjęcia. Na dokładność pomiarową zdjęć znacząco wpływa również proces skanowania, który obejmuje: wybór odpowiedniego urządzenia skanującego oraz prawidłowy dobór parametrów skanowania,

takich jak rozdzielczość czy balans barw, poprawne ułożenie zdjęć w skanerze. Zmiana któregokolwiek z parametrów skanowania może prowadzić do istotnego obniżenia jakości danych pomiarowych. Wszystkie wymienione powyżej czynniki (niedoskonałości) negatywnie wpływają na wartość pomiarową zdjęć i możliwość ich dalszego przetwarzania.

3.1.1 Deformacja podłoża filmu lotniczego

Zdjęcia archiwalne były zwykle wykonywane przez kamery analogowe, obraz był tworzony na filmie zwojowym. Film taki był poddawany różnorodnym czynnikom deformującym podłoże. Miało to miejsce na różnych etapach:

- w trakcie wykonywania zdjęć (naprężenia mechaniczne w procesie przewijania i wyplaszczania filmu),
- w trakcie obróbki fotochemicznej (naprężenia mechaniczne, proces suszenia filmu),
- w trakcie składowania filmów (wpływ temperatury, wilgotności i tzw. starzenia się filmu),
- na etapie skanowania zdjęć.

W sposób oczywisty degradowało to jakość geometryczną zdjęć. Różny jest charakter tych deformacji (odwracalne, nieodwracalne, regularne, nieregularne). Takie problemy nie występują w przypadku zdjęć wykonywanych kamerami cyfrowymi. Nie występują tu czynniki degradujące geometrię zdjęć, obecne w przypadku zdjęć na filmie.

Niniejszy podrozdział podejmuje problem deformacji zdjęć analogowych (wykonanych na materiale światłoczułym), identyfikuje czynniki powodujące deformacje, omawia ich charakter oraz zakres, a także analizuje zasadność ich uwzględnienia w procesie pomiarowego opracowania zdjęć. Klasyczne podejście do problemu pomiarowego opracowania zdjęć lotniczych zakłada znajomość geometrii zdjęć. W praktyce przekłada się to na:

1. Znajomość elementów orientacji wewnętrznej kamery, uwzględniających, oprócz podstawowych elementów (położenie punktu głównego x_0 , y_0 oraz odległość obrazowa c_k), również dystorsje obiektywu (dystorsja radialna i tangencjalna), skutkujące odstępstwem wiązki promieni przechodzących przez obiektyw od zasady rzutu środkowego, na której zbudowana jest teoria pomiarowego opracowania zdjęć. Elementy te dla danego egzemplarza kamery są wyznaczone w procesie kalibracji kamery i zapisane w metryce kalibracji;

2. Zagwarantowanie braku czynników deformujących geometrię zdjęcia naświetlonego na filmie. Mowa tu o deformacjach podłoża filmu spowodowanych naprężeniami mechanicznymi w trakcie wykonywania zdjęcia, procesem obróbki fotochemicznej filmów (wywoływanie, suszenie), składowaniem naświetlonych filmów (tzw. proces starzenia się zdjęć). Deformacje takie powinny być minimalizowane, a ewentualne szczątkowe ich wartości eliminowane w procesie opracowania zdjęć;
3. Zagwarantowanie na etapie skanowania zdjęć na filmie zachowania geometrii skanowanych zdjęć. Dotyczy to szczególnie skanowania z użyciem skanerów innych niż specjalistyczne skanery dedykowane skanowaniu zdjęć lotniczych.

3.1.2 Stabilność geometryczna filmów

Stabilność wymiarów geometrycznych zdjęć ma decydujące znaczenie dla ich przydatności pomiarowej (American Society of Photogrammetry, 1980). O stabilności tej decyduje przede wszystkim stabilność podłoża emulsji. Podłożem filmów lotniczych były folie poliestrowe, znacznie bardziej stabilne niż używane wcześniej podłoża z folii trójoctanowej. Podłoże poliestrowe jednak nie chroniło całkowicie zdjęć od deformacji.

Wymiary zdjęć zmieniają się pod wpływem:

- zmian temperatury,
- zmian wilgotności,
- procesu obróbki fotochemicznej,
- mechanicznych naprężeń,
- „starzenia” się filmów (warunków przechowywania filmów po obróbce fotochemicznej).

Wielkość tych deformacji w znacznym stopniu zależy od grubości folii podłożowej.

W praktyce fotolotniczej spotyka się następujące grubości:

- 0.06 mm - używane w filmach dla celów interpretacyjnych (nie zalecane dla filmów przeznaczonych dla opracowań mapowych),
- 0.10 mm – zalecane dla filmów przeznaczonych dla opracowań mapowych,
- 0.18 mm – używane dla filmów duplikacyjnych.

Z punktu widzenia właściwości pomiarowych ważna jest nie tylko wielkość deformacji, ale i jej charakter. Zmiany te mogą być odwracalne lub nieodwracalne. Dla pomiarowych

właściwości zdjęć szczególnie niepożądane są zmiany nieregularne (tj. powodujące zmiany kształtu: różne w kierunku podłużnym i poprzecznym filmu), w okresie między naświetleniem filmu, a opracowaniem pomiarowym zdjęć. Współczesne filmy na podłożu poliestrowym o grubości 0.10 mm charakteryzują się odpornością na deformacje jak w tabeli 3.1.

Tab. 3.1 Stabilność wymiarów filmów lotniczych na podłożu poliestrowym o grubości 0.10 mm
(American Society of Photogrammetry, 1980)

Źródło deformacji	Deformacje		Komentarz
	Typ emulsji		
	cz - b	kolorowa	
Współczynnik rozszerzalności termalnej [%/°C]	0.002	0.002	Opracowanie zdjęć w temperaturze 5°C wyżej niż temperatura ekspozycji spowoduje różnicę odległości znaczków tłowych rzędu 21-31 μm
Współczynnik rozszerzalności wilgotnościowej [%/1%RH]	0.002	0.003	Zmiana wilgotności względnej o 10% spowoduje różnicę odległości znaczków tłowych rzędu 42-63 μm
Obróbka chemiczna. Zakres zmian: rozszerzenie [%] skurcz [%]	-0.03 +0.02	-0.04 +0.03	Obróbka fotochemiczna filmu pozostawia nieodwracalną zmianę odległości znaczków tłowych rzędu od -84 μm dla skurczu, do +61 μm dla rozszerzenia
Obróbka i starzenie się (skurcz starzejącego się wywołanego filmu [%]): 1 tydzień przy 50°C, 20% RH 1 rok przy 25°C, 60% RH	0.03 0.03	0.06 0.03	-

Dla porównania, stosowana wcześniej folia trójoctanowa charakteryzowała się współczynnikiem rozszerzalności termalnej około 4 razy większym niż stosowana później folia poliestrowa.

Zdarza się, że dostępne zdjęcia mają formę odbitek stykowych na papierowym podłożu (papier fotograficzny). Ma to miejsce na przykład w przypadku archiwalnych zdjęć rekonesansowych, wykonywanych podczas II Wojny Światowej. Skurcze papierowego podłoża są bardzo znaczne, kilkadziesiąt razy większe niż podłoża poliestrowego. Dodatkowo, ich charakterystyczną cechą jest różny skurcz wzdłuż obu osi zdjęcia, wynikający z włóknistej struktury papieru.

Dodatkowe deformacje wnosi proces skanowania zdjęć. Dotyczy to skanowania na skanerach płaskich typu *desktop*, nieprzystosowanych do skanowania zdjęć lotniczych. Cechą charakterystyczną takich skanerów jest wprowadzanie deformacji typu afinicznego, tj. różnej skali wzdłuż obu osi skanera. Deformacje takie mogą być bardzo znaczne.

W klasycznym podejściu pomiarowego opracowania zdjęć analogowych (zwykle skanów takich zdjęć) dysponujemy metryką kamery, którą wykonano zdjęcia. Metryka taka zawiera m.in. skalibrowane współrzędne znaczków tłowych, definiujących układ współrzędnych tłowych. Pierwszym etapem opracowania jest rekonstrukcja elementów orientacji wewnętrznej kamery, polegająca na pomiarze znaczków tłowych widzianych na zdjęciach (zwykle czterech lub ośmiu) i ich transformacji na współrzędne nominalne (z metryki kamery). Wykorzystywana jest przy tym transformacja konforemna, afiniczna lub rzutowa. Pozwala to na skuteczną eliminację deformacji podłoża zdjęć, bez względu na ich źródło. Szczególnie przydatna jest transformacja afiniczna lub rzutowa, uwzględniająca najgroźniejsze z tych deformacji, te o charakterze afinicznym (różny skurcz w obu kierunkach). Na etapie rekonstrukcji elementów orientacji wewnętrznej uwzględnia się również dystorsje obiektywu, o parametrach wziętych z metryki.

Sytuacja komplikuje się w przypadku braku metryki kamery, co ma dość często miejsce w przypadku opracowania zdjęć archiwalnych.

Powyższe problemy deformacji podłoża nie występują praktycznie w przypadku opracowania zdjęć pozyskiwanych kamerami cyfrowymi. Sam proces produkcji przetworników obrazowych CCD (*Charge-Coupled Device*) czy CMOS (*Complementary Metal-Oxide-Semiconductor*) i ich stabilność wymiarów w czasie gwarantują stałość elementów orientacji wewnętrznej.

3.1.3 Przegląd metryk kamer analogowych i generowanie metryki wirtualnej

Dokonano przeglądu kilkudziesięciu metryk fotogrametrycznych kamer analogowych, używanych w Polsce w okresie powojennym. Dla celów pomiarowych używano głównie kamer RC 10 (format kadru 23x23 cm) firmy Wild, również nowsze modele RC 20 i RC 30 (Leica). Rzadziej występowały kamery RMK A i RMK TOP (Zeiss) i LMK (Zeiss Jena). Przegląd metryk kamer analogowych znajduje się w rozdziale 13 - Załączniki.

Przegląd tych kamer pozwala sformułować kilka istotnych wniosków dotyczących ich jakości geometrycznej:

1. Już na etapie produkcji i montażu kamer dbano o bardzo precyzyjny montaż obiektywu kamery (stożka obiektywowego) względem ramki tłowej ze znaczkami tłowymi. W efekcie narożne znaczkii tłowe tworzą kwadrat z odstępstwami pojedynczych mikrometrów, a punkt główny obiektywu wypada w początku układu współrzędnych tłowych z dokładnością kilka mikrometrów.
2. Dystorsja (radialna i tangencjalna) obiektywów kamer pomiarowych jest nieznaczna, praktycznie nie przekracza w polu widzenia kamery dziesięciu mikrometrów, a dla nowszych kamer jest poniżej 2 μm .

Powyższe spostrzeżenia pozwalają w przypadku braku metryki kamery, wygenerować „wirtualną metrykę kamery”. Metryka taka:

- ma znaczkii tłowe o współrzędnych nominalnych dla danego typu kamery, tworzące kwadrat,
- punkt główny kamery pokrywa się z początkiem układu współrzędnych tłowych (współrzędne tłowe PPA są równe zero),
- zniekształcenia obiektywu (dystorsja) jest pomijalna, tj. przyjmuje się ich wartość równą zero,
- odległość obrazowa (stała kamery) (c_k) odpowiada wartości odczytanej z marginesu zdjęcia.

Najważniejszą zaletą tak wygenerowanej metryki jest możliwość skutecznej eliminacji w procesie opracowania znacznych deformacji podłoża zdjęć, w tym szczególnie groźnych deformacji o charakterze afinicznym. Powyższa analiza pokazuje, że błędy wprowadzone przez taką „wirtualną” metrykę (zamiast niedostępnej metryki oryginalnej) będą na poziomie kilku (rzadziej kilkunastu) mikrometrów, tj. na poziomie znacznie mniejszym niż wartość samych deformacji zdjęć. Dodatkowo, ewentualne błędy w wygenerowanych sztucznie współrzędnych znaczków tłowych spowodują bardzo nieznaczne deformacje typu konforemne, tj. z zachowaniem podobieństwa, skutkujące bardzo niewielką zmianą skali zdjęć, dające się łatwo wyeliminować na dalszych etapach opracowania. Dokładnie taki sam efekt spowoduje

przyjęcie stałej odległości obrazowej kamery nieznacznie odstającej od nieznanej wartości kalibrowanej odległości obrazowej.

3.2 Dokładność orientacji zdjęć lotniczych

Opisując orientację zdjęć lotniczych nie sposób nie wspomnieć o wymaganiach dokładnościowych. W celu poprawnego wykonania orientacji bloku zdjęć lotniczych niezbędne jest ustalenie dokładności dotyczących wyznaczenia punktów wiążących, pomiaru punktów osnowy fotogrametrycznej na zdjęciach i w terenie oraz dokładności ostatecznego wyniku wyznaczenia elementów orientacji zewnętrznej zdjęć.

W tabeli 3.2 przedstawiono minimalne wymagania dotyczące poszczególnych grup punktów pomiarowych – współrzędnych fotopunktów, punktów wiążących, tłowych oraz środka rzutów – stosowane w krajowych opracowaniach fotogrametrycznych. Na przykład maksymalny błąd średni pomiarów współrzędnej X i Y fotopunktu dla skali zdjęć 1:6 000 wynosi 0.05 m, zaś współrzędnej Z 0.07 m. Dla skali 1:14 000 maksymalny błąd średni pomiarów współrzędnej X, Y i Z fotopunktu wynosi 0.25 m. Dla skali 1:26 000 maksymalny błąd średni pomiarów współrzędnej X i Y fotopunktu wynosi 0.50 m, zaś współrzędnej Z 0.40 m (Dziennik Ustaw Nr 263, 2011 r.).

Tab. 3.2 Maksymalne błędy średnie pomiaru współrzędnych fotopunktów, środka rzutów i współrzędnych tłowych dla zdjęć analogowych (Dziennik Ustaw Nr 263, 2011 r.)

Skala zdjęć	Maksymalny błąd średni pomiarów				Maksymalny przeciętny błąd średni współrzędnej punktu wiążącego	
	Współrzędna fotopunktu [m]		Współrzędna środka rzutów [m]	Współrzędna tłowa [μ m]	Współrzędna punktu wiążącego [m]	
	M_x i M_y	M_z	M_{x0}, M_{y0}, M_{z0}	m_x i m_y	M_x i M_y	M_z
1: 6 000 (GSD $\approx$ 10 cm)	0.05	0.07	0.15	6.5	0.05	0.08
1:14 000 (GSD $\approx$ 25 cm)	0.25	0.25	0.20	6.5	0.12	0.20
1:26 000 (GSD $\approx$ 50 cm)	0.50	0.40	0.20	6.5	0.25	0.40

Błąd średni (Tab. 3.2) pomiaru współrzędnej tłowej na zdjęciu nie może przekroczyć wartości 6.5 μ m w przypadku zdjęć analogowych, a w przypadku zdjęć cyfrowych 3.0 μ m (Dziennik Ustaw Nr 263, 2011 r.). W tabeli 3.2 zawarte są również dopuszczalne błędy średnie

współrzędnej punktu wiążącego. Maksymalny błąd średni współrzędnych X i Y punktu wiążącego dla skali 1:6000 wynosi 0.05 m, natomiast dla współrzędnej Z 0.08 m, dla skali 1:14 000 wynosi odpowiednio 0.12 m i 0.20 m, a dla skali 1:26 000 wynosi odpowiednio 0.25 m i 0.40 m. Z punktu widzenia orientacji zdjęć punkty wiążące i fotopunkty odgrywają kluczową rolę w zapewnieniu dokładności i spójności procesu orientacji. Dlatego przedstawione wartości błędów ich pomiaru powinny być uwzględnione na etapie orientacji archiwalnych zdjęć lotniczych.

Dopuszczalne różnice współrzędnych D_x , D_y , D_z pomiędzy pomiarem fotogrametrycznym, a pomiarem terenowym na każdym z punktów kontrolnych (stanowią niezależną od aerotriangulacji kontrolę wyników w zakresie ochrony przed omyłkami, błędami grubymi i błędami systematycznymi) wynoszą jak w tabeli 3.3.

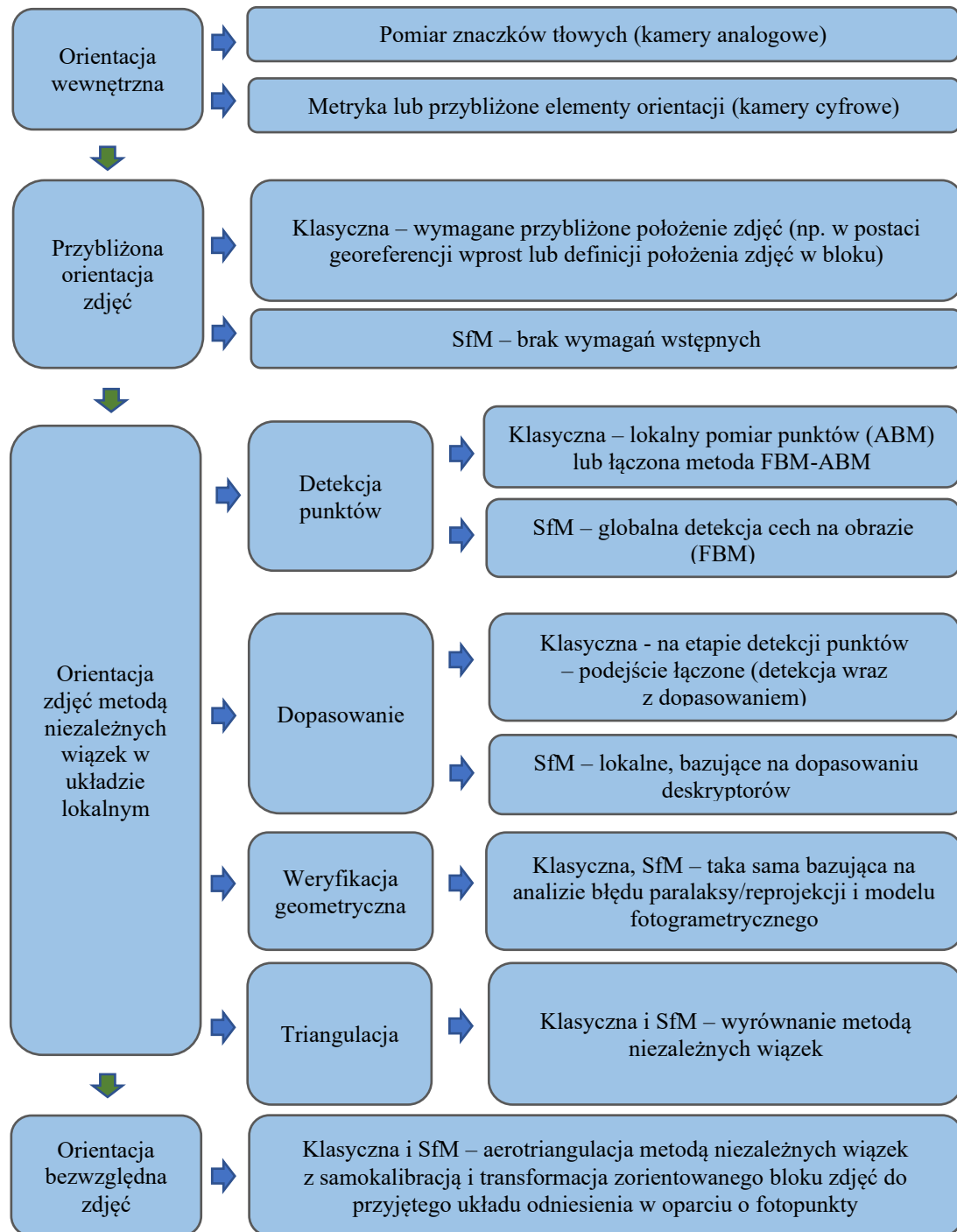
Tab. 3.3 Dopuszczalne różnice współrzędnej punktu kontrolnego między pomiarem fotogrametrycznym i terenowym, w aerotriangulacji bloków zdjęć (Dziennik Ustaw Nr 263, 2011 r.)

Skala zdjęć lotniczych	Dopuszczalna różnica współrzędnej punktu kontrolnego [m]			
	Kontrola podczas aerotriangulacji		Kontrola przez niezależny podmiot	
	D_x i D_y	D_z	D_x i D_y	D_z
1:6 000 (GSD≈10 cm)	0.15	0.21	0.30	0.30
1:14 000 (GSD≈25 cm)	0.75	0.75	1.00	1.20
1:26 000 (GSD≈50 cm)	1.50	1.20	2.00	1.60

3.3 Schemat pomiarowego opracowania bloku zdjęć lotniczych

Rozwój nowoczesnych metod orientacji zdjęć cyfrowych przyczynił się do zmiany podejścia do orientacji zdjęć lotniczych i odejścia od klasycznego podejścia stosowanego w fotogrametrii na rzecz metody *Structure-from-Motion*, powszechnie wykorzystywanej w widzeniu maszynowym i w fotogrametrii niskopułapowej oraz bliskiego zasięgu. Klasyczne podejście stosowane w fotogrametrii bazuje na punktowych algorytmach FBM i podejściu ABM, podczas gdy podejście oparte na rozwiązywaniu SfM wykorzystuje punktowe oraz powierzchniowe (typu *blob*) algorytmy FBM. Należy jednak podkreślić, że schemat przetwarzania zdjęć obiema metodami jest zbliżony, a różnice zauważalne są na etapie orientacji wzajemnej zdjęć w bloku

fotogrametrycznym. Pomimo tego wynikiem końcowym pomiarowego opracowania bloku zdjęć lotniczych dwiema metodami jest uzyskanie georeferencji każdego zdjęcia w bloku. Oba podejścia zostały przedstawione na rysunku 3.1, a ich szczegółowy opis zawarto w podrozdziałach 3.3.1 i 3.3.2.



Rys.3.1 Schemat ogólny procesu orientacji zdjęć lotniczych w ujęciu klasycznym i SfM

3.3.1 Podejście klasyczne do orientacji zdjęć lotniczych

Głównym celem orientacji archiwalnych zdjęć lotniczych jest wpasowanie bloku wzajemnie zorientowanych zdjęć w przyjęty układ odniesienia przestrzennego. Proces ten, określany jako orientacja bezwzględna, realizowany jest na podstawie punktów osnowy fotogrametrycznej (fotopunktów), których współrzędne są znane w referencyjnym układzie geodezyjnym. Do orientacji zdjęć lotniczych wykonanych przy wykorzystaniu kamer cyfrowych zazwyczaj na etapie projektu nalotu projektowana jest sygnalizowana polowa osnowa fotogrametryczna rozmieszczona i pomierzona w obszarze opracowania, natomiast dla zdjęć lotniczych wykonanych kamerami analogowymi w większości przypadków osnowę polową stanowiły punkty naturalne w postaci osi lub krawędzi skrzyżowań dróg, chodników, narożników parkingów, a także środki studzienek kanalizacyjnych i telekomunikacyjnych oraz inne elementy płaskich elementów infrastruktury technicznej mierzonych w terenie i identyfikowalnych na zdjęciu po wykonaniu nalotu fotogrametrycznego. Ze względu na brak informacji dotyczących rozmieszczenia punktów osnowy fotogrametrycznej lub współrzędnych sygnalizowanych punktów niezbędne jest „zaprojektowanie jej od nowa” poprzez odszukanie na archiwalnych zdjęciach i współczesnych materiałach kartograficznych tych samych elementów, które nie uległy zmianie w czasie.

Klasyczne podejście orientacji archiwalnych zdjęć lotniczych (analogowych i cyfrowych) stosowane w takich programach jak ImageStation Intergraph i Inpho Trimble wykonywane jest w następujących etapach (rysunek 3.1):

1. Orientacja wewnętrzna (odtworzenie parametrów kamery na podstawie metryki kamery, pomiar znaczków tłowych (kamery analogowe) celem transformacji z układu pikselowego do układu tłowego. W kamerach cyfrowych orientacja wewnętrzna zdefiniowana jest na podstawie geometrii matrycy obrazowej;
2. Przybliżona orientacja zdjęć – wymagane jest podanie przybliżonego położenie zdjęć (np. w postaci georeferencji wprost lub definicji położenia zdjęć w bloku);
3. Orientacja zdjęć metodą niezależnych wiązek (*bundle adjustment*) bazująca na punktach wykrytych przy wykorzystaniu automatycznych lub półautomatycznych metod detekcji opartych na algorytmie ABM lub zintegrowanej metody FBM-ABM. W ramach tego etapu wykonywane jest dopasowanie wykrytych punktów charakterystycznych i weryfikacja

geometryczna wraz z triangulacją, mająca na celu połączenie zdjęć wzajemnie w jeden spójny blok (na bazie warunku kolinearności);

4. Orientacja bezwzględna: wyrównanie metodą niezależnych wiązek (aerotriangulacja) wraz z transformacją zorientowanego bloku zdjęć do przyjętego układu odniesienia w oparciu o pomierzone fotopunkty (na bazie kolinearności).

3.3.2 Podejście do orientacji zdjęć metodą SfM

Ze względu na wspomnianą niską jakość obrazów archiwalnych (obrazy analogowe o słabej radiometrii i zniekształceniach geometrycznych) jak również brak dostępnych parametrów kamery, zastosowanie klasycznego podejścia orientacji zdjęć stosowanego w programach komercyjnych takich jak Inpho, czy Intergraph jest znacznie utrudnione, a niekiedy wręcz niemożliwe. W większości przypadków archiwalne zdjęcia lotnicze nie posiadają pełnej dokumentacji kalibracyjnej, umożliwiającej bezpośrednie wyznaczenie elementów orientacji wewnętrznej. Znane są jednak dane dotyczące stałych parametrów geometrycznych danej kamery, takich jak położenie znaczków tłowych, długość ogniskowej itp., na podstawie których możliwe jest oszacowanie tzw. metryki wirtualnej (por. podrozdział 3.1.3). Błędy wynikające z zastosowania takiej metryki są rzędu kilku, rzadziej kilkunastu mikrometrów, co jest wartością znacznie mniejszą niż deformacje samych zdjęć.

Podejściem umożliwiającym kompensację powstałych szacunkowych błędów – jak również błędów wynikających z afinizmu, zmian geometrycznych i radiometrycznych zdjęcia, a także niedoskonałości procesu skanowania – jest zastosowanie metod pomiarowych wykorzystywanych w widzeniu maszynowym, w szczególności metody *Structure-from-Motion* (SfM) (Markiewicz i in., 2019).

Dzięki temu, że metoda SfM pozwala na wykrycie bardzo dużej liczby równomiernie rozłożonych punktów na całej powierzchni zdjęcia, możliwe jest przeprowadzenie samokalibracji. Nie jest to jednak klasyczna kalibracja w rozumieniu fotogrametrycznym, w której wyznaczane są takie parametry, jak stała ogniskowa, współrzędne punktu głównego czy zniekształcenia radialne i tangencjalne. W przypadku zdjęć analogowych w metodzie SfM mówimy raczej o samokalibracji umożliwiającej oszacowanie zniekształceń geometrycznych i afinicznych. Dodatkową zaletą metody SfM w porównaniu do klasycznych technik kalibracyjnych jest jej odporność na afinizm, wynikająca z zastosowanych algorytmów detekcji

i dopasowania punktów. Metoda ta polega na automatycznej orientacji wzajemnej grupy zdjęć z jednoczesną samokalibracją w oparciu o punkty wykryte i dopasowane wzajemnie niezależnie od skali, obrotu bądź zniekształceń geometrycznych niezależnie na każdym zdjęciu.

Takie podejście pozwala na wykrycie inwariantnych punktów charakterystycznych odpornych na zmianę skali, skręcenia i pochylenia obrazów oraz jest odporne na lokalne zniekształcenia geometryczne (m.in. afinizm, dystorsja) (szczegółowy opis jest zawarty w rozdziale 4) (Karwel i Markiewicz, 2022; Zawieska i.in, 2019).

W przeciwieństwie do „klasycznego” podejścia stosowanego do orientacji zdjęć lotniczych, metoda SfM nie wymaga żadnych informacji początkowych odnoszących się do orientacji wewnętrznej zdjęć oraz nie jest wymagana przybliżona ich orientacja. W ogólnym procesie orientacji zdjęć lotniczych z wykorzystaniem metody SfM wyróżnia się pięć etapów przetwarzania zdjęć:

1. Na etapie orientacji wzajemnej w pierwszym etapie za pomocą detektorów wykrywane są pojedyncze cechy charakterystyczne zdjęcia w postaci punktów, linii lub całych regionów. Tworzą one tak zwane cechy lokalne, te natomiast przy większej ich liczbie stanowią strukturę globalną. Następnie za pomocą deskryptorów opisywane są cechy charakterystyczne wykrytych elementów w oparciu o gradienty stopni szarości w otoczeniu wykrytego punktu na poszczególnych poziomach piramidy obrazu.
2. Drugi etap dotyczy statystycznego dopasowania deskryptorów poprzez analizę podobieństwa między nimi.
3. W trzecim etapie przeprowadzana jest weryfikacja geometryczna w oparciu o wpasowany model fotogrametryczny i analizę wartości błędu paralaksy nazywanego również błędem reprojekcji. Na tym etapie wyznaczane są przybliżone elementy orientacji wzajemnej par zdjęć.
4. W czwartym etapie wykonywana jest triangulacja i wyrównanie metodą niezależnych wiązek celem wyznaczenia elementów orientacji wzajemnej grupy zdjęć.
5. Ostatnim etapem jest wyznaczenie elementów orientacji zewnętrznej - aerotriangulacja metodą niezależnych wiązek z samokalibracją i transformacja zdjęć na podstawie fotopunktów do przyjętego układu odniesienia.

Schemat ogólny działania podejścia SfM do orientacji zdjęć przedstawiono na rysunku 3.1.

3.3.3 Różnice między metodą klasyczną i SfM orientacji zdjęć lotniczych

Porównując klasyczną metodę orientacji zdjęć lotniczych z metodami automatycznymi opartymi na podejściu *Structure-from-Motion* (SfM), można wskazać dwie zasadnicze różnice.

Pierwsza dotyczy orientacji wewnętrznej zdjęć, która jest niezbędnym etapem w metodzie klasycznej, natomiast w podejściu SfM nie jest wymagana w sposób bezpośredni. Parametry metryczne kamery mogą być wykorzystywane jedynie jako pierwsze przybliżenie w procesie samokalibracji. W metodzie SfM przybliżone wartości elementów orientacji wewnętrznej odczytywane są z metadanych EXIF zapisanych w plikach zdjęć, co umożliwia automatyczne pozyskanie informacji o ogniskowej użytej podczas wykonywania zdjęć oraz wielkości piksela matrycy. Za współrzędne punktu głównego przyjmuje się zwykle połowę szerokości i wysokości obrazu.

Druga dotyczy tego, że w metodzie SfM nie są wymagane przybliżone elementy orientacji zewnętrznej, lecz podobnie jak w przypadku orientacji wewnętrznej ich znajomość (np. pozyskane z pomiaru GNSS/INS) może być traktowana jako przybliżona w procesie wyrównania metodą niezależnych wiązek co przekłada się na zmniejszenie złożoności obliczeniowej i skrócenie czasu orientacji zdjęć. Zaletą stosowania metody SfM jest nadliczbowość wykrywanych punktów wiążących, co przekłada się na większą spójność geometryczną całego bloku fotogrametrycznego. Dzięki temu możliwe jest zmniejszenie liczby punktów osnowy fotogrametrycznej wymaganych do przeprowadzenia orientacji bezwzględnej bloku, co czyni proces bardziej efektywnym i mniej czasochłonnym. W klasycznym podejściu niezbędne jest wstępne przybliżenie położenia zdjęć, na przykład w formie georeferencji wprost lub określenia ich położenia w bloku fotogrametrycznym. Wynika to ze specyfiki działania algorytmów detekcji punktów wiążących przy zastosowaniu metody ABM. Pozostałe etapy przetwarzania w obu metodach są zbliżone pod względem struktury i przebiegu.

Główną i zasadniczą różnicą metody orientacji zdjęć lotniczych jest podejście do pomiaru punktów wiążących (rys. 3.1). W metodzie klasycznej pomiar punktów wiążących i fotopunktów wykonywany jest przy użyciu algorytmu ABM i/lub FBM-ABM, natomiast metoda SfM bazuje na algorytmie FBM. Metoda ABM odnosi się do lokalnego wykrywania punktów, które są wyznaczane w oparciu o analizę wartości stopni szarości pikseli zdjęcia w oknie

poszukiwawczym. Metoda ta dobrze działa dla obrazów o zbliżonej teksturze i intensywności oraz zbliżonej skali. Dla zdjęć lotniczych pozyskanych z jednego okresu (okres kilku dni) zarówno jedna jak i druga metoda zapewnia wystarczającą liczbę i dokładność detekcji punktów służących do ich orientacji.

W przypadku występowania różnic w radiometrii zdjęć spowodowanej np. inną datą pozyskania dopasowanie ich do pozostałych zdjęć przy użyciu podejście ABM jest utrudnione a wręcz niemożliwe do zrealizowania, zależy to od stopnia różnic tonalnych zdjęć. Im większe są różnice (dłuższy odstęp czasowy pomiędzy wykonanymi zdjęciami, albo inna pora roku) tym algorytm ma większe trudności z dopasowaniem pikseli o zmiennych wartościach radiometrycznych. Z tego powodu w proponowanym autorskim podejściu orientacji zdjęć wieloczasowych, dla których ze względu na różną datę ich pozyskania (różną radiometrią), a także różną ich skalę obrazową zastosowanie metody ABM do wyszukiwania wspólnych punktów stanowiących podstawę ich orientacji jest niewykonalne. Różna radiometria zdjęć (słaba jakość) odnosi się do ich nieostrości, niskiego kontrastu i dużego rozmycia. Z tego względu zastosowanie podejścia SfM z uwzględnieniem algorytmów FBM stanowi odporne i inwariantne rozwiązanie tego problemu.

Znaczącą różnicą między orientacją wzajemną zdjęć lotniczych w obu podejściach jest również sposób realizacji procesu orientacji. W metodzie klasycznej orientacja zdjęć wykonywana jest jednocześnie dla wszystkich zdjęć w bloku, natomiast w podejściu SfM proces rozpoczyna się od utworzenia modelu początkowego na podstawie pary zdjęć, dla której wyznaczono największą liczbę punktów wiążących. Dla tej pary przeprowadza się triangulację oraz wyrównanie metodą niezależnych wiązek. Następnie do utworzonego modelu sukcesywnie dołączane są kolejne zdjęcia, przy czym po każdym etapie wykonywane jest ponowne wyrównanie. W końcowej fazie, po dołączeniu wszystkich zdjęć, przeprowadza się globalne wyrównanie metodą niezależnych wiązek dla całego bloku zdjęć. Opisany sposób orientacji bloku określany jest jako podejście inkrementalne (*Incremental SfM*). Podejście to, charakterystyczne dla metody SfM, umożliwia zwiększenie dokładności orientacji bloków zdjęć w porównaniu z metodami klasycznymi, zwłaszcza w przypadku zdjęć pozyskanych kamerami o nieznanej metryce.

4. Algorytmy przetwarzania obrazów cyfrowych stosowane w metodzie SfM

Współczesne metody przetwarzania obrazowego, w tym orientacja zdjęć lotniczych czy naziemnych, opierają się na zaawansowanych metodach oraz algorytmach stosowanych w widzeniu maszynowym (*Computer Vision*, CV) (Byrne, 2016; Morgan i Brogan, 2016; Friedt, 2014; Ressler i in., 2020; Zhang i in., 2021). Zastosowanie algorytmów widzenia maszynowego w procesie orientacji zdjęć wynika z ich zdolności do pełnej automatyzacji przetwarzania dużych zbiorów obrazów oraz możliwości pracy z sensorami o zróżnicowanej rozdzielczości przestrzennej. Z tego też powodu, do detekcji punktów wiążących na zdjęciach fotogrametrycznych stosuje się dwie główne metody pomiarowe tj. *Area-Based Matching* (ABM) oraz *Feature-Based Matching* (FBM).

Metoda ABM polega na dopasowaniu odpowiadających sobie fragmentów obrazów w oparciu o analizę wartości stopni szarości w oknie poszukiwawczym o wymiarach $n \times n$ (Forstner, 1986). Algorytm ABM (Mousa, 2014) jest wykorzystywany na etapie aerotriangulacji przede wszystkim do wyznaczania punktów wiążących na podstawie pary zdjęć, jak również pomiaru osnowy fotogrametrycznej, a także podczas generowania gęstych chmur punktów. Identyfikacja tych punktów opiera się na zastosowaniu tzw. okien korelacyjnych, w których porównywane są lokalne fragmenty obrazów. Analizowane jest okno korelacyjne (obszar) wokół piksela referencyjnego na jednym zdjęciu, a następnie ten sam wzorzec porównywany jest z fragmentami obrazu w drugim zdjęciu. Dopasowanie bazuje na analizie podobieństwa radiometrycznego, np. wartości jasności, pomiędzy odpowiadającymi sobie obszarami na różnych zdjęciach.

Algorytm FBM bazuje na globalnej lub lokalnej analizie gradientów stopni szarości lub wartości stopni szarości niezależnie dla każdego zdjęcia (Tuytelaars i Mikolajczyk, 2007). Analizowany jest obszar całego zdjęcia, na którym dopasowywane są względem siebie jego fragmenty charakteryzujące się zbliżonymi cechami.

Na podstawie analizy dostępnej literatury (Forstner, 1986, Rosenfeld, 1986, Tuytelaars i Mikolajczyk, 2007) odnoszącej się do metodyki dopasowania obrazów metodami ABM i FBM, można jednoznacznie stwierdzić przewagę metody FBM nad metodą ABM w kontekście

automatycznej orientacji zdjęć. Przewaga ta obejmuje zarówno strategię dopasowania, zbieżność wyników, szybkość pomiaru, jak i mniejszą wrażliwość metody na błędy. Potencjał algorytmu FBM, wynikający z możliwości wyznaczania punktów wiążących na podstawie analizy zmian gradientu obrazu, jest znacznie większy niż potencjał podejścia ABM, które bazuje na analizie zmian jasności pikseli obrazu. Z tego względu w badaniach do wyznaczenia punktów wiążących czy dopasowania (patrz rozdział 7 i 9) niezbędnych do procesu orientacji zdjęć lotniczych zdecydowano się na wykorzystanie algorytmów FBM. W tabeli 4.1 przedstawione zostały wady i zalety obu metod. Należy podkreślić, że metoda ABM przedstawiona w tabeli 4.1 ściśle jest związana z metodą LSM (*Least Squares Matching*), która znajduje zastosowanie w końcowym jej etapie dopasowania na podstawie minimalizacji różnic intensywności odpowiadających sobie pikseli na dwóch obrazach.

Tab. 4.1 Metody dopasowania ABM i FBM - porównanie (Babbar i in., 2010)

Parametry	Area-Based Matching (ABM)	Feature-Based Matching (FBM)
Strategia dopasowania obrazów	każdy z dopasowywanych punktów stanowi środek małego wzorcowego okna pikseli na pierwszym obrazie, do którego dopasowywane jest okno poszukiwawcze na drugim obrazie (o wymiarach równych oknu wzorcowemu)	analiza całości zdjęcia, z którego wybierane są jedynie punkty charakterystyczne
Konwergencja szybkość i wszechstronność	działanie ABM z mniejszą konwergencją, szybkością i wszechstronnością w porównaniu z FBM	FBM jest lepszy do korelacji obrazu z uwzględnieniem szybkości i wszechstronności oraz jest lepszy od LSM pod względem zakresu zbieżności, szybkości i wszechstronności
Uniwersalność	porównywalna	porównywalna
Dokładność	niższa (dla dobrej jakości obrazów podobna, dla słabej jakości niższa)	większa (dla dobrej jakości obrazów podobna, dla słabej jakości wyższa)
Zbieżność dopasowania	nie zawsze osiąga wysokie wartości nawet przy zastosowaniu metody <i>Cross – Corellation</i>	wysoka
Parametry wejściowe	ABM zazwyczaj wymaga bardzo dobrych wartości początkowych przy nieznanach parametrach	brak konieczności wyznaczenia parametrów wejściowych
Czułość na błędy	większa, LSM ma bardzo mały promień zbieżności. Choć jest on większy dla krzyżowej korelacji, to warunek podobieństwa musi być spełniony	mniejsza, techniki dopasowywania oparte na cechach są mniej wrażliwe na jakość przybliżonych wartości.
Czułość na występowanie martwych pól (ang. <i>Oclusion</i>)	ABM jest wrażliwy na okluzje	mniejsza, wynikająca z zaproponowanego modelu matematycznego

W metodzie FBM lokalne cechy odnoszą się do punktów, krawędzi i linii oraz regionów, zaś większe globalne cechy nazywane są strukturami. Na cechy globalne składają się cechy lokalne. Oprócz opisu cech globalnych za pomocą atrybutów lokalnych do ich opisu wprowadza się relacje między cechami lokalnymi. Według teorii Mikołajczyka, dobrze wykryte i opisane punkty charakterystyczne powinny mieć następujące właściwości (Tuytelaars i Mikołajczyk, 2007):

- jednoznaczność – wykryte punkty powinny być wyróżnialne i wysoce kontrastowe tzn. muszą odróżniać się od najbliższych sąsiadów,

- niezmiennosc – wyznaczone punkty powinny być niezmiennie w stosunku do zniekształceń geometrycznych i radiometrycznych co bezpośrednio wpływa na dokładność i wiarygodność dopasowania. W celu wyróżnienia punktów w lokalnym otoczeniu powinny one być również kontrastowe,
- stabilność – bez względu na szumy czy warunki otoczenia punkty powinny być wykrywane zawsze jednakowo,
- interpretowalność – zasada doboru punktów (krawędzi, narożników) powinna być wykonana według typu (narożnik, krawędź),
- unikalność – punkty wykrywane na zdjęciach powinny być unikalne, pozwalając tym samym na ich jednoznaczne wyodrębnienie na podstawie pewnych cech charakterystycznych, co jest istotne dla obrazów z powtarzalną teksturą.

W celu doboru odpowiedniego algorytmu dopasowania obrazów należy uwzględnić strukturę powierzchni zdjęcia lotniczego, w tym teksturę obrazu występujących na nim elementów terenowych, ponieważ ma to wpływ na dokładność i poprawność pomiaru.

Do istotnych czynników warunkujących dobór algorytmu zaliczamy:

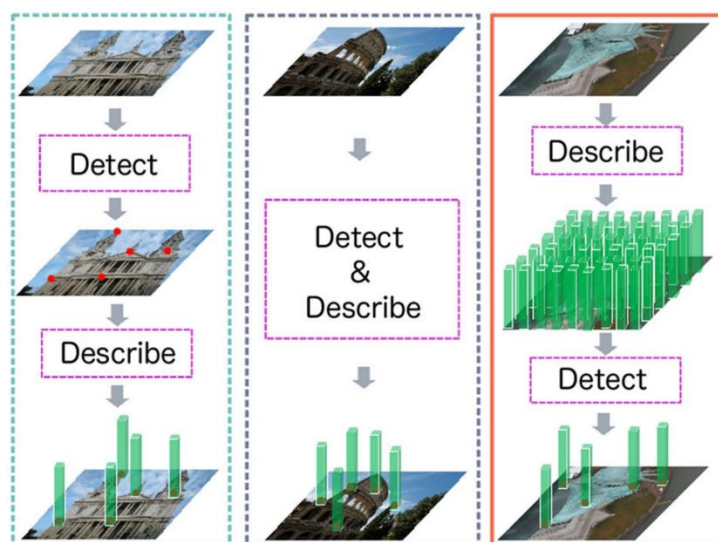
- niejednoznaczność (powtarzalność) struktury – wyznaczenie na obrazie kilku różnych punktów charakterystycznych odpowiadających jednemu punktowi na zdjęciu sąsiednim, na skutek niejednoznaczności struktury zdjęcia (powtarzalna struktura zdjęcia – np. przęsła w moście),
- uboga tekstura analizowanego obszaru z uwzględnieniem możliwości występowania szumów obrazu powodująca niepewność pomiaru i błędy detekcji punktów,
- zasłonięcie punktu na zdjęciu homologicznym w stosunku do jego położenia na zdjęciu odniesienia na skutek występowania elementów terenowych i kąta jego odwzorowania.

4.1 Opis wykorzystanych algorytmów CV pomiaru punktów charakterystycznych

Biorąc pod uwagę charakterystykę archiwalnych zdjęć lotniczych tj. występowanie zniekształceń geometrycznych (skurcze, afinizm, błędy związane z konwersją do postaci cyfrowej) i radiometrycznych (zmiana kontrastu, rozmycia, błędy konwersji barwnej podczas skanowania) do pomiaru punktów charakterystycznych (wiązących – patrz rozdział 7,

przejęciowych – patrz rozdział 9) niezbędne jest wykorzystanie algorytmów spełniających wymogi teorii Mikołajczyka (Tuytelaars i Mikołajczyk, 2007). W związku z tym zdecydowano się na implementację zarówno klasycznych algorytmów punktowych i powierzchniowych typu *Hand-Crafted* (HC), jak i najnowszych podejść opartych na wytrenowanych sieciach neuronowych - *Learned-Based* (LB). W tym celu wybrano trzy podejścia stosowane do ekstrakcji i opisu punktów charakterystycznych wykorzystujące kombinacje metod HC i LB tj. metody *Detect-then-Describe*, *Detect-and-Describe* i *Describe-then-Detect*. Działanie powyższych metod zostało zilustrowane w sposób poglądowy na rysunku 4.2, prezentując ich różnice pod względem kolejności operacji i sposobu przetwarzania danych obrazowych.

Wspominając o metodach wykrywania, dopasowania lub wykrywania i dopasowania cech charakterystycznych na zdjęciach przy wykorzystaniu metod widzenia maszynowego należy również wspomnieć o tych opartych na Głębokich Sieciach Neuronowych - DNN (*Deep Neural Network*). We współczesnym podziale metod dopasowania obrazów można wydzielić dwie grupy tj. grupa algorytmów typu HC, do których zaliczamy m.in. detektory np. SIFT (Lowe, 2004) lub SURF (Bay i Ess, 2008), oraz drugie bazujące na zastosowaniach DNN nazywane *Learned-Based-Feature Extraction* (LBFE), oparte m.in. na takich rozwiązaniach jak SuperGlue lub LoFTR (*Local Feature Transformer*). Podejście HC umożliwia wykrywanie punktów kluczowych na podstawie wartości gradientów skali szarości w najbliższym sąsiedztwie (*blob detectors*, tj. SIFT, SURF lub CenSurE) lub poprzez porównanie różnic wartości stopni szarości w stosunku do analizowanego piksela (*corner detectors*, tj. FAST (Rosten i Drummond, 2006), BRISK (Leutenegger i in., 2011)). Podejście to opiera się wyłącznie na zależnościach matematycznych, które można jednoznacznie opisać za pomocą wzorów analitycznych, co zapewnia jego przewidywalność i spójność działania w różnych warunkach.



Rys. 4.2 Porównanie metod pomiaru punktów charakterystycznych (Yan N. i in., 2023)

W ostatnich latach opracowano rozwiązania bazujące na automatycznym podejściu opartym na uczeniu maszynowym, mające na celu pokonanie istniejących ograniczeń metod HC (Verdie i in., 2015) i wykorzystujące rozwiązania bazujące na sieciach neuronowych. Podobnie jak metody HC, również metody LB oparte na głębokich sieciach neuronowych charakteryzują się wysoką odpornością na różnorodność danych wejściowych, takich jak zmienności oświetlenia, szumy, zniekształcenia geometryczne czy skalowanie obrazów. Dzięki zdolności do przetwarzania dużych zbiorów danych modele te są w stanie wyekstrahować punkty charakterystyczne, które nie można w jednoznaczny sposób opisać za pomocą znanych reguł matematycznych. To umożliwia wykrywanie i opis tych cech w sposób optymalny dla konkretnego zadania, niezależnie od złożoności danych.

Szczegółowy opis metod wykrywania punktów kluczowych na zdjęciach przedstawiono poniżej.

Wyróżnia się trzy główne rozwiązania, a mianowicie:

1. Metoda *Detect-then-Describe*, w której proces analizy rozpoczyna się od detekcji punktów charakterystycznych, a następnie przeprowadzane jest ich opisanie za pomocą deskryptora. Zarówno detektor (Barroso-Laguna i in., 2019; Verdie i in., 2015), jak i deskryptor (Ebel i in., 2019; Mishchuk i in., 2017) mogą być oparte na klasycznych metodach HC lub stanowić kombinację rozwiązań klasycznych HC i opartych sieciach neuronowych LB. Metoda ta najpierw wykonuje detekcję cech, które następnie

opisywane są za pomocą deskryptora. W tym celu wykorzystuje się tzw. cechy niezmiennicze, którymi mogą być m.in. kierunki zmiany gradientu stopnia szarości lub wartości stopni szarości dla najbliższego otoczenia. Wraz z rozwojem widzenia komputerowego i rosnącym zapotrzebowaniem na metody automatycznego dopasowywania obrazów, zaproponowano wiele algorytmów HC (m.in.: SIFT – *Scale-Invariant Feature Transform*, SURF – *Speeded Up Robust Features*, FAST – *Features from Accelerated Segment Test* czy BRIEF – *Binary Robust Independent Elementary Features*) do wykrywania cech, mających na celu wyodrębnienie stabilnych i łatwo rozróżnialnych cech z obrazów. Podstawową zasadą na tym etapie jest zapewnienie stabilności oraz powtarzalności wykrywanych punktów charakterystycznych. W drugim etapie do opisu (*Describe*) wykrytych punktów kluczowych można stosować albo algorytmy klasyczne HC (m.in.: SIFT, SURF, BRIEF) albo algorytmy głębokich sieci neuronowych LB (m.in.: SuperPoint, HardNet czy D2-Net). W przypadku tych drugich sieci neuronowe są trenowane w sposób nadzorowany, aby generować deskryptory punktów kluczowych, które są bardziej odporne na zakłócenia. Następnie za pomocą sieci neuronowych stosuje się regresję w celu przewidywania lokalizacji nowych punktów kluczowych (Tyszkiewicz et al., 2020).

2. Metoda *Detect-and-Describe*, wspólnie optymalizuje cały ciąg przetwarzania danych w celu wyodrębnienia odpornych punktów wiążących, np. SuperPoint (DeTone i in., 2017), SuperGlue (Sarlin i in., 2019), DISK (Tyszkiewicz i in., 2020). Metoda *Detect-and-Describe* jest stosowana w celu zwiększenia zarówno liczby wykrywanych punktów kluczowych, ich powtarzalności i niezawodności, a w konsekwencji dokładności ich położenia (Remondino i in., 2021). Kompleksowa strategia dopasowywania *Detect-and-Describe* łączy dwa różne etapy ekstrakcji cech obrazu, opisu i dopasowywania w jeden ciąg procesowy. Takie rozwiązanie umożliwia globalne optymalizowanie parametrów modelu oraz adaptacyjne poprawianie wydajności każdego etapu. Yi i in. w 2016 zaproponowali metodę *Learning Invariant Feature Transform* (LIFT). Ta pierwsza kompleksowa sieć głębokiego uczenia oparta jest na algorytmie SIFT. LIFT odwzorowuje proces SIFT poprzez wykrywanie punktów kluczowych, szacowanie ich orientacji oraz obliczanie deskryptorów, zapewniając niezmienniczość względem transformacji

obrazowych. LIFT uczy się całego procesu (detekcji, orientacji, deskryptora) jako sieć neuronowa, która zastępuje klasycznie projektowane etapy SIFT. Podejście to jako dane wejściowe wykorzystuje obrazy wieloskalowe, następnie za pomocą przesuwanego okna obrazowego znajduje punkty charakterystyczne. W dalszej kolejności wyodrębnia lokalne bloki obrazu w celu przypisania orientacji punktów, a na koniec wyznacza deskryptory na podstawie znormalizowanego otoczenia punktów. W rezultacie LIFT generuje punkty odporne na zmiany perspektywy czy radiometrii obrazu (Yi i in., 2016; Yan i in., 2023).

3. Kompleksowa metoda typu *Describe-then-Detect* w ekstrakcji cech lokalnych najpierw generują gęste mapy deskryptorów, a dopiero następnie wyznaczają punkty kluczowe na podstawie lokalnych maksimów lub mechanizmów uwagi (Choy i in., 2016; Rocco i in., 2018; Li i in., 2020). Do tej kategorii należą między innymi algorytmy takie jak D2-Net (Dusmanu i in., 2019) oraz DELF (*Deep Local Features*, Noh i in. 2017). W niektórych opracowaniach do tej grupy metod zaliczany jest również model LoFTR (Sun i in., 2021), ponieważ dopasowania wyprowadzane są z gęstych reprezentacji cech generowanych przez sieć neuronową. Jednocześnie w literaturze LoFTR jest klasyfikowany jako metoda typu *detector-free*, w której dopasowania pomiędzy obrazami wyznaczane są bez jawnego etapu detekcji punktów kluczowych. Model wykorzystuje architekturę Transformer (Vaswani i in., 2017) i realizuje dopasowanie w schemacie *coarse-to-fine*: najpierw wyznacza przybliżone dopasowania na poziomie gęstej mapy cech, a następnie dokonuje ich precyzyjnego dopasowania na poziomie szczegółowym. Niemniej jednak w niniejszej rozprawie doktorskiej przyjęto algorytm LoFTR w kontekście podejścia *Describe-then-Detect*.

W przypadku metod opartych na deskryptorach lokalnych, obejmujących podejścia *Detect-then-Describe* oraz *Detect-and-Describe*, dopasowanie punktów pomiędzy obrazami jest zwykle realizowane przy użyciu algorytmu k-NN (*k-Nearest Neighbors*) (Cover i Hart, 1967) lub metod wyszukiwania najbliższych sąsiadów, takich jak Brute-Force Matching czy FLANN (*Fast Library for Approximate Nearest Neighbors*) (Muja i Lowe, 2009). W metodach opartych na głębokich sieciach neuronowych, takich jak LoFTR (podejście *Describe-then-Detect*), dopasowania wyznaczane są bezpośrednio przez model, bez konieczności wyszukiwania najbliższych

deskryptorów. Natomiast na etapie weryfikacji geometrycznej, niezależnie od zastosowanego podejścia, często stosuje się algorytm RANSAC (*Random Sample Consensus*) (Fischler i Bolles, 1981), który umożliwia odrzucanie błędnych dopasowań oraz estymację transformacji pomiędzy obrazami, jak również LMeds (*Least Median of Squares*) (Rousseeuw, 1984) czy PROSAC (*Progressive Sample Consensus*) (Chum i Matas, 2005).

4.2 Detektory punktowe i powierzchniowe (Hand-Crafted)

W ostatnich latach zauważalny jest rozwój metod wykrywania i opisu punktów charakterystycznych opartych na podejściu *Learned-Based*. Jest to bezpośrednio związane z rozwojem metod wykorzystywanych w sztucznej inteligencji (AI – *Artificial Intelligence*) oraz rozwojem sprzętu komputerowego, który umożliwia przeprowadzanie złożonych operacji obliczeniowych. Aby jednak możliwe było zastosowanie sieci neuronowych, niezbędne jest wykorzystanie tzw. zbiorów uczących, będących zestawami danych przeznaczonymi do trenowania sieci neuronowych. Zbiory uczące składają się z obiektów przedstawionych za pomocą cech opisujących pewne przypisane zależności. Z tego też powodu przed rozpoczęciem przetwarzania danych przy wykorzystaniu algorytmów AI niezbędne jest posiadanie reprezentatywnego zbioru uczącego pozwalającego na poprawną predykcję punktów wiążących.

Metody bazujące na algorytmach *Hand-Crafted* nie wymagają zbiorów uczących, a wykrycie, opis i dopasowanie punktów wykonywane jest jedynie w oparciu o zależności, które można opisać wzorami matematycznymi. Takie podejście pozwala na w pełni kontrolowalne wykrywanie punktów przy mniejszej złożoności algorytmicznej i mniejszych wymaganiach sprzętowych, co przekłada się również na krótszy czas przetwarzania danych.

W metodach *Hand-Crafted* do wykrycia punktów charakterystycznych lub grup pikseli tworzących obszary tzw. *blobs* wykorzystuje się detektor, który ma na celu nie tylko wykrycie w przestrzeni obrazu ww. punktów lub grup pikseli, ale również pozwala na dodatkowe przypisanie cech, tj. kierunek zmian gradientów stopni szarości lub skali odnoszącej się do wielkości obszaru, w którym dany punkt jest jednoznacznie rozpoznawalny. W zależności od wykorzystywanego detektora wyznaczane są różne parametry dodatkowe (np. wartość gradientu, kąt rotacji czy rozmiar wykrytego obszaru).

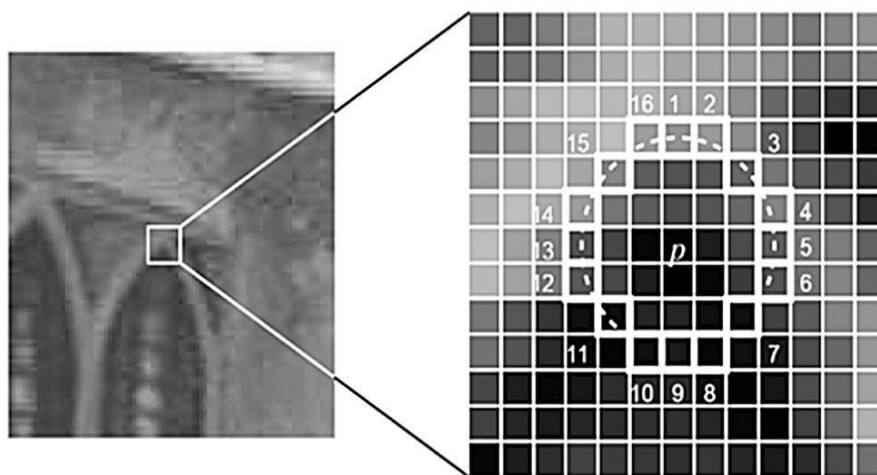
Dla efektywnego działania detektorów ważne jest zastosowanie oryginalnych, surowych i nieskompresowanych danych obrazowych. Takie podejście zapewnia większą liczbę poprawnie wykrytych punktów, a także ich lepsze rozmieszczenie. Niemniej jednak w przypadku wykorzystania detektorów punktowych możliwe jest zastosowanie wstępnego przetwarzania obrazów celem poprawy jakości wykrywanych narożników. Analizując dostępne algorytmy CV pod względem liczby wykrytych punktów, czasu i jakości (jednoznaczność, niezmienność, stabilność, interpretowalność i unikalność) ich wykrycia w pracy doktorskiej wykorzystano algorytmy z grupy FBM. Do analiz zastosowano dwie grupy detektorów, tj. detektory punktowe FAST (*Features from Accelerated Segment Test*) i BRISK (*Binary Robust Invariant Scalable Keypoints*) oraz detektory powierzchniowe (typu blob) SIFT (*Scale-Invariant Feature Transform*), SURF (*Speeded Up Robust Features*) i CenSurE (*Center Surround Extremas*). W badaniach ze względu na rodzaj, geometrię oraz sposób pozyskania zdjęć zastosowano również detektory z grupy detektorów afinicznych (A-detektor). Specyfiką A-detektorów jest wykrycie większej liczby punktów charakterystycznych w porównaniu do “normalnego podejścia” (Markiewicz, 2020). Detekcja ma miejsce dla zdjęć wykonanych pod różnymi kątami widzenia (ω , ϕ) i różnym kątem skręceniem, uwzględniając również afinizmy spowodowane różnicą skali odwzorowania wzdłuż osi ox względem osi oy . Dlatego też, A-detektory powinno stosować się dla zdjęć wykonanych z różnych kierunków i o słabej teksturze. Podejście A-detektorów daje dużą nadliczbowość punktów, co pozwala na zwiększenie dokładności opracowania. W przypadku zdjęć lotniczych pomimo usystematyzowanego ich sposób pozyskiwania (zdjęcia wykonane w jednym kierunku – pionowy) oraz dużej odległości fotografowania (małe różnice wysokości dla obiektów powierzchniowych) zastosowanie A-detektorów również wpływa na wzrost liczby wyznaczonych punktów, np. na obszarach z wysoką zabudową. Z tego względu A-detektory zostały zastosowane na etapie sekwencyjnego dopasowania bloków fotogrametrycznych (rozdział 9). Potwierdza to uzyskana różna liczba wykrytych punktów pomiędzy detektorami SURF i ASURF (*Affine Speeded Up Robust Features*) oraz FAST i AFAST (*Affine Scale-Invariant Feature Transform*) (podrozdział 9.1.2).

4.2.1 Detektory punktowe – FAST i BRISK

Detektory punktowe pozwalają na wykrycie punktów charakterystycznych w oparciu o porównanie wartości stopni szarości pikseli sąsiednich względem analizowanego piksela. Operują na lokalnych strukturach obrazu, które wykazują charakterystyczne cechy, takie jak narożniki, punkty przecięcia krawędzi oraz obszary o wysokiej zmienności gradientu w ich bezpośrednim otoczeniu. Do grupy detektorów punktowych zaliczamy m.in. algorytm FAST oraz BRISK.

FAST

Koncepcja działania detektora FAST (*Features from Accelerated Segment Test*) została przedstawiona w roku 2006 przez Rostena oraz Drummonda. Dotyczyła ona przygotowania algorytmu, który byłby w stanie w czasie rzeczywistym wykryć punkty na klatkach obrazu video. Z tego też powodu, ten detektor znalazł zastosowanie min. w orientacji zdjęć sekwencyjnych przy użyciu algorytmu SLAM (*Simultaneous Localisation and Mapping*). W celu oceny sprawdzenia czy piksel jest potencjalnym narożnikiem, porównywana jest jego wartość stopnia szarości z pikselami opartymi na okręgu *Bresenham* o promieniu 3 pikseli. Z tego powodu działanie algorytmu odnosi się do badania intensywności obrazu o cechach lokalnych w pełnej jego rozdzielczości.



Rys. 4.3 Zasada działania detektora FAST detekcji punktu charakterystycznego w oparciu o 16 pikselowy okrąg (Rosten i Drummond, 2006)

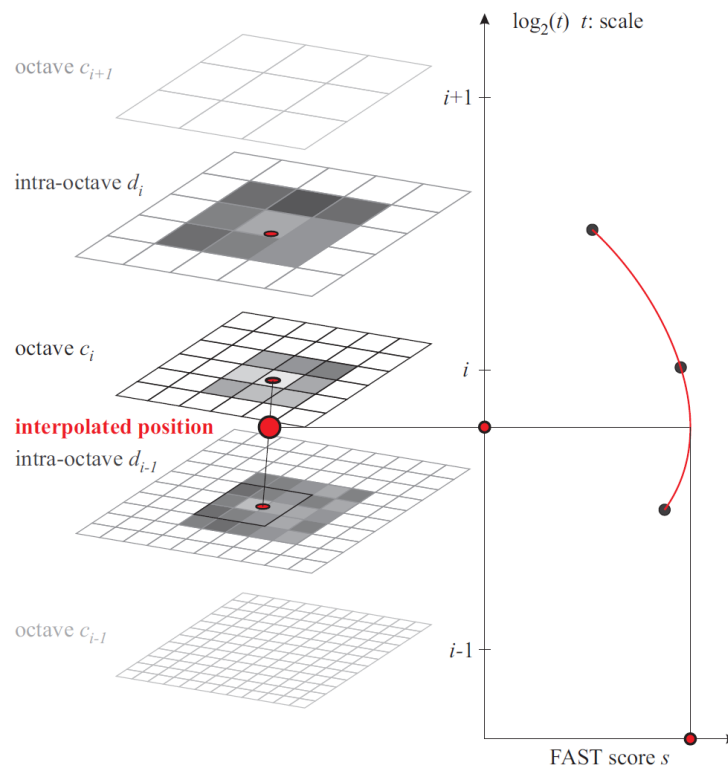
Zasada działania algorytmu FAST składa się z następujących etapów (rysunek 4.3):

- wybór na obrazie piksela w celu weryfikacji jego przynależności do punktów charakterystycznych. Wyznaczana jest wartość intensywności I_p ,
- określenie granicznej wartości progu, poniżej którego punkty są wykorzystywane w dalszych analizach,
- w celu analizy wartości pikseli wokół punktu charakterystycznego definiowany jest okrąg o obwodzie 16 pikseli (okrąg *Bresenham* o promieniu 3 pikseli),
- w przypadku gdy w otoczeniu punktu znajduje się n pikseli o intensywności większej lub równej, to do zdefiniowania go wybierane jest 12 z nich,
- dla przyspieszenia działania algorytmu na początku etapu dokonuje się selekcji i porównania pikseli o numerach nieparzystych 1, 3, 5, 13. Jeżeli wystąpi przypadek, że dwa pierwsze piksele są jaśniejsze lub ciemniejsze w stosunku do porównywalnego piksela, algorytm przechodzi do porównania dwóch kolejnych. Dla zatwierdzenia wyboru piksela jako punktu charakterystycznego niezbędna jest weryfikacja, czy trzy piksele otaczające punkt są jaśniejsze lub ciemniejsze. W przypadku, gdy przedstawione warunki są spełnione następuje porównanie kolejnych pikseli obrazu.

Ze względu na dużą niestabilność działania algorytmu wynikającą z szybkości jego działania autorzy detektora zaproponowali podejście oparte na uczeniu maszynowym umożliwiające zminimalizowanie błędów związanych z jego działaniem (Rosten i Drummond, 2006).

BRISK

Kolejnym detektorem punktowych jest detektor BRISK (*Binary Robust Invariant Scalable Keypoints*), którego sposób działania zbliżony jest do algorytmu FAST, a różnicę stanowi sposób wyszukiwania punktów charakterystycznych (Leutenegger i.in., 2011). Punkty są wyszukiwane nie na pełnej rozdzielczości obrazu, ale na różnych poziomach jego piramidy. Pozostałe kryteria decydujące o jego klasyfikacji jako punkt charakterystyczny są takie same jak dla algorytmu FAST.



Rys. 4.4 Detekcja punktów charakterystycznych na różnych poziomach piramidy obrazu – detektor BRISK (Leutenegger i.in., 2011)

W algorytmie BRISK zaproponowana została metoda generowania piramidy obrazu składającej się z n poziomów c_i , w skład których wchodzi n podpoziomy d_i . Począwszy od obrazu oryginalnego c_0 w wyniku przepróbowania co drugi piksel powstają kolejne poziomy (rysunek 4.4). Podpoziomy są tworzone na podstawie generalizacji o współczynnik 1.5 kolejnych poziomów zaczynając od poziomem 0 (obraz oryginalny). Wartość środkowa punktu charakterystycznego wyznaczana jest w oparciu o 8 wartości pikseli jego otoczenia (podobszar), na różnych poziomach piramidy. Dalej obliczane są wskaźniki na poszczególnych podobszarach dzięki którym wpasowywana jest parabola. Za pomocą paraboli wyznaczana jest wartość punktu, co umożliwia wyznaczenie jego położenia na odpowiednim poziomie piramidy obrazu (Leutenegger i.in., 2011).

4.2.2 Detektory powierzchniowe – SIFT, SURF, CenSurE, ASIFT

Ważną grupę detektorów do wykrywania punktów charakterystycznych stanowią detektory powierzchniowe zwane również detektorami *blob*. Odnoszą się one do wyszukiwania punktów na zdjęciach w granicach charakterystycznych obszarów wyznaczanych na podstawie właściwości otoczenia (zmian jasności czy zmian barw). Badany obszar charakteryzuje się

dużym kontrastem w stosunku do najbliższego otoczenia, inwariantnością niezależnie od warunków i zmian oświetlenia i zniekształceń geometrycznych (skali czy skręcenia), wyraźną odróżnialnością od otoczenia na zdjęciu i odpornością na szумы (Urban i Weinmann, 2015). Do tej grupy detektorów najczęściej zaliczamy detektor SIFT (*Scale-Invariant Feature Transform*), SURF (*Speeded Up Robust Features*) oraz CenSurE (*Center Surround Extremas*) (Tab. 4.2).

Tab. 4.2 Zasady działania algorytmów powierzchniowych CenSurE, SIFT i SURF (Agrawal i.in., 2008)

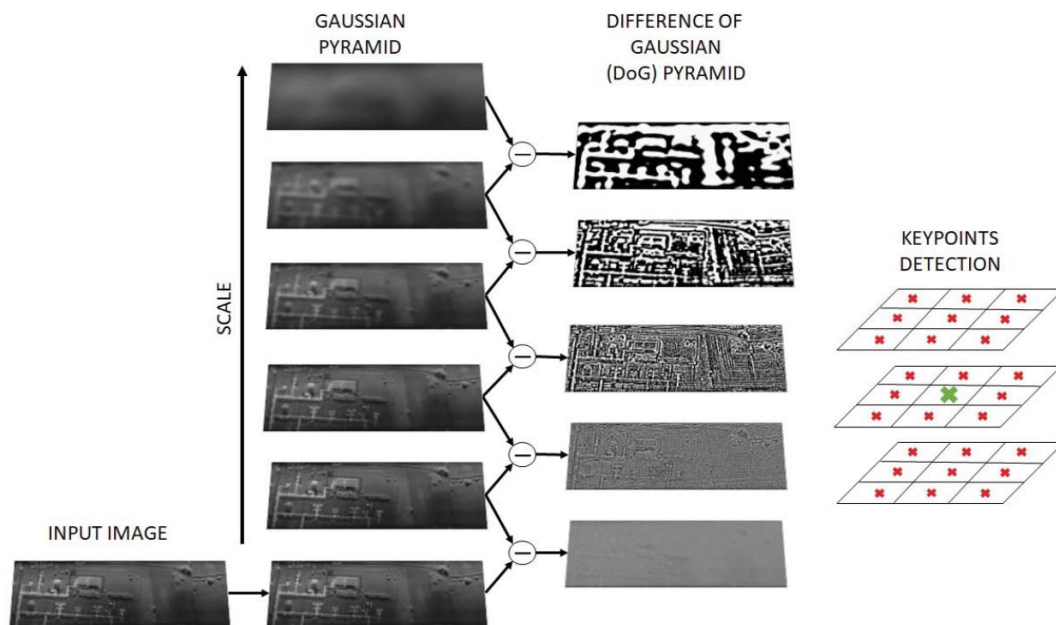
Działanie	Detektor		
	SIFT	SURF	CenSurE
Rozdzielczość przetwarzanego zdjęcia	Praca na różnych piramidach obrazu	Praca na różnych piramidach obrazu	Pełna
Filtracja obrazu	Laplasjan (DoG)	Hesjan (DoB)	Laplasjan (Centersurround)
Filtr krawędziowy	Hesjan	Hesjan	Harris
Wpływ obrotu	Tak	Nie	Przybliżony

Detektor SIFT wykorzystuje różnicę rozmytych obrazów (DoG – *Difference of Gaussians*) do lokalizacji punktów charakterystycznych w różnych skalach. SURF opiera się na aproksymacji Laplasjanu Gaussa przy użyciu filtrów pudełkowych oraz wykorzystuje wyznacznik macierzy Hesjanu. Z kolei detektor CenSurE stosuje przybliżone operatory typu *center-surround* w celu szybkiego wykrywania lokalnych ekstremów jasności, co pozwala na zwiększenie wydajności oraz zapewnienie odporności na rotację obrazu. Algorytm CenSurE odwrotnie niż dwa pozostałe działa na pełnej rozdzielczości zdjęcia. Do wyznaczania wartości maksymalnych w algorytmach SIFT i SURF wykorzystywana jest maska filtru oparta odpowiednio na operatorze DoG oraz DoB (*Difference of Boxes*). Jako filtr krawędziowy w detektorach SIFT i SURF zastosowano macierz Hessego, natomiast w detektorze CenSurE wykorzystano filtr Harrisa.

Detektor SIFT

Algorytm SIFT (*Scale Invariant Feature Transform*) został pierwotnie zaproponowany przez (Lowe, 1999) do rozpoznawania obiektowego na podstawie lokalnych cech obrazu. Cechy te są niezmiennie od skali obrazu (piramidy obrazu), translacji, obrotu oraz częściowo od zmiany oświetlenia i afinizmu. Główną ideą tego algorytmu jest przetwarzanie obrazów w następującej kolejności (rysunek 4.5):

1. wykrywanie ekstremów (wartości maksymalnych i minimalnych) funkcji obrazu różnic gaussowskich (DoG) na różnych poziomach piramidy;
2. lokalizacja punktu kluczowego (określenie jego współrzędnych i otoczenia), odrzucanie punktów o niskim kontraście oraz leżących na krawędziach;
3. przypisanie orientacji punktom na podstawie lokalnych zmian gradientu stopni szarości;
4. opis punktów kluczowych na podstawie wyznaczenia gradientów obrazu w najbliższym otoczeniu obszaru za pomocą deskryptora.



Rys. 4.5 Schemat detekcji punktów za pomocą detektora SIFT (Markiewicz i in., 2019)

Potencjalne regiony cech lokalnych (niezależne od skali i orientacji) odpowiadają lokalnym ekstremom w różnicowych obrazach (4.3) splotu ($L(x,y,\sigma)$) (4.1) z filtrem Gaussa ($G(x,y,\sigma)$) (4.2) przy różnych poziomach piramid (w poniższych wzorach odnoszących się do detektora SIFT symbolika została zaczerpnięta z Lowe, 2004):

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad (4.1)$$

gdzie:

$L(x,y,\sigma)$ – Obraz wygładzony (skalowany) dla pozycji (x,y) i skali σ ,

$I(x,y)$ – zdjęcie wejściowe poddane konwolucji,

$G(x,y,\sigma)$ – dwuwymiarowa funkcja Gaussa (rozmycie Gaussowskie),

(x,y) – położenie punktu na obrazie,

$*$ – splot (konwolucja) – operacja rozmycia obrazu przez filtr Gaussowski,

σ – wartość współczynnika filtru gaussowskiego (4.1).

Dwuwymiarową funkcję Gaussa, wykorzystywaną w algorytmie SIFT do rozmycia obrazu w celu analizy w różnych skalach zapisujemy według wzoru 4.2:

$$G(x,y,\sigma) = \left(\frac{1}{2\pi\sigma^2}\right) e^{-\frac{(x^2+y^2)}{2\sigma^2}}. \quad (4.2)$$

Punkt kluczowy jest traktowane jako maksima/minima różnicy Gaussa DoG (*Difference of Gaussian*). Różnica filtru gaussowskiego (DoG) jest obliczana dla sąsiednich obrazów (4.2):

$$\text{DoG}(x,y,\sigma) = G(x,y,k\sigma) - G(x,y,\sigma), \quad (4.3)$$

gdzie: $\text{DoG}(x,y,\sigma)$ – różnica dwóch rozmytych obrazów (*Difference of Gaussian*),

k – to współczynnik skalujący.

W kolejnym kroku obraz/obrazy są poddawane konwolucji filtrami gaussowskimi (DoG) (4.4, 4.5):

$$D(x,y,\sigma) = [G(x,y,k\sigma) - G(x,y,\sigma)] * I(x,y) \quad (4.4)$$

$$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma). \quad (4.5)$$

gdzie: $D(x,y,\sigma)$ – obraz różnicy Gaussowskiej DoG w skali σ .

Zastosowanie powyższych wzorów można opisać w następujący sposób – Obraz wejściowy jest przetwarzany za pomocą funkcji Gaussa, tak aby pozyskać obrazy oddzielone w przestrzeni piramidy stałą wartością współczynnika k ($k = 2^{1/s}$). Każda oktawa w przestrzeni piramidy (powstała przez usunięcie z każdego wiersza i kolumny co drugiego piksela) dzielona jest na stałą liczbę interwałów. Dla każdego interwału oktawy tworzone są obrazy rozmyte. Następnie, aby stworzyć obraz różnicowy DoG sąsiadujące interwały są od siebie odejmowane. W celu wykrycia punktów charakterystycznych (ekstremów) (*keypoints detection*) każdy piksel jest porównywany z ośmioma sąsiednimi w obrazie różnicowym i z dziewięcioma w piramidzie powyżej i poniżej w tzw. oknie poszukiwawczym o wymiarze 3x3 (kwadrat 3D) (rysunek 4.5).

Potencjalny punkt przypisany jest jako kandydat na punkt kluczowy wtedy, gdy ma wartość większą lub mniejszą niż sąsiadujące z nim punkty.

Lokalizacja punktów kluczowych odbywa się w laplasianie obrazu na podstawie rozwinięcia Taylora (do drugiej pochodnej) funkcji przestrzeni skali $D(x,y,\sigma)$ wokół punktu kandydującego na punkt kluczowy. Pochodne funkcji DoG są obliczane

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (4.6)$$

w tym punkcie w celu wyznaczenia dokładnego położenia ekstremum, gdzie: D to wartość funkcji DoG w punkcie, $x=(x,y,\sigma)^T$ to współrzędna w przestrzeni skali, $D(x)$ (4.6) to wartość funkcji DoG w pobliżu punktu, $\frac{\partial D}{\partial x}$ to gradient funkcji DoG, a $\frac{\partial^2 D}{\partial x^2}$ druga pochodna cząstkowa funkcja DoG. Lokalizacja punktu charakterystycznego z dokładnością do podpiksli/podskali jest przyjmowana jako ekstremum w kwadracie 3D (4.7):

$$(\hat{x}) = - \left(\frac{\partial^2 D}{\partial x^2} \right)^{-1} \frac{\partial D}{\partial x} \quad (4.7)$$

gdzie: $\hat{x}$ to dokładne położenie punktu charakterystycznego.

Wartość funkcji w ekstremum $D(\hat{x})$ jest przydatna do odrzucania niestabilnych ekstremów o niskim kontraście (4.8):

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x}. \quad (4.8)$$

gdzie: $D(\hat{x})$ - wartość funkcji DoG w dokładnym położeniu ekstremum.

Wzór ten pozwala na obliczenie przybliżonej wartości funkcji DoG w jej ekstremum, tzn. po interpolacji.

Wykrywanie ekstremów w przestrzeni skali powoduje powstanie zbyt wielu punktów kluczowych, z których część jest niestabilna.

Punkty kluczowe są filtrowane na podstawie miar ich stabilności – odrzucane są punkty, które mają niski kontrast (co odnosi się do ich wrażliwości na szum) lub są zlokalizowane wzdłuż krawędzi. Dlatego też, przewidziana jest trzyetapowa procedura filtracji:

- interpolacja pobliskich danych w celu uzyskania dokładnej pozycji,

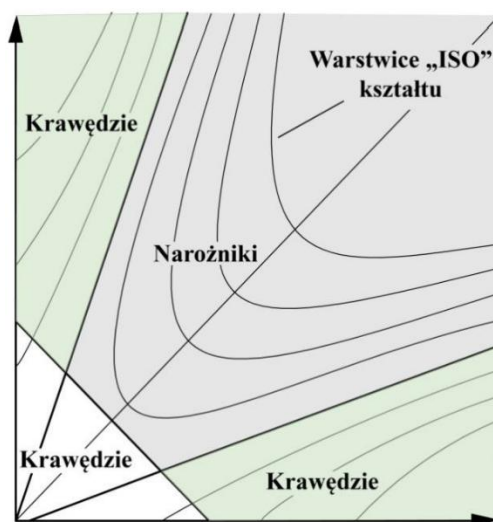
- odrzucenie punktów kluczowych o niskim kontraście,
- eliminacja odpowiedzi na krawędzi.

W celu wykrycia punktów stabilnych i odpornych na szum, nie wystarczy odrzucić punkty kluczowe o niskim kontraście. DoG będzie miał duże oddziaływanie wzdłuż krawędzi, nawet jeśli lokalizacja wzdłuż krawędzi jest słabo określona, a zatem niestabilna w przypadku niewielkich ilości szumu. Słabo zdefiniowane maksima w funkcji DoG będą miały duże wektory własne w poprzek krawędzi, ale małe wzdłuż krawędzi. Wektory własne można obliczyć z macierzy hesjanu 2x2, obliczonej w miejscu i skali punktu kluczowego (4.9):

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}, \quad (4.9)$$

gdzie: D_{xx} , D_{xy} , D_{yy} są to pochodne funkcji DoG.

Powstała wskutek przekształceń macierz hesjanu H jest macierzą wartości lokalnych otoczenia badanego punktu (rys. 4.6). W zależności od przyjętych wartości własnych macierzy definiuje się przynależność punktu do jednej z trzech klas: narożników, krawędzi i ogólnych kształtów według teorii zaproponowanej w detektorze Harris'a (Harris i Stephens, 1988).



Rys. 4.6 Przykład interpretacji wartości własnych macierzy hesjanu. Punkt zostaje uznany za narożnik w przypadku, gdy obie wartości są duże (Harris i Stephens, 1988)

Na podstawie wartości zawartych w macierzy hesjanu można ustalić kryterium, dla którego wykryte punkty można uznać za poprawne i nie znajdujące się na krawędziach. Niech α będzie

wartością własną o największej wielkości, a β najmniejszą. Następnie możemy obliczyć sumę wartości własnych z macierzy $\mathbf{H}$ (4.9) i ich iloczyn z wyznacznika (4.11):

$$\text{Tr}(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta \quad (4.10)$$

$$\text{Det}(\mathbf{H}) = D_{xx} D_{yy} - (D_{xy})^2 = \alpha\beta, \quad (4.11)$$

gdzie: α i β to wartości własne macierzy Hessego.

W mało prawdopodobnym przypadku, gdy wyznacznik jest ujemny, wartości mają różne znaki, więc punkt jest odrzucany jako niebędący ekstremum. Niech $\mathbf{r}$ jest stosunkiem największej wartości własnej do mniejszej, a $\alpha = \mathbf{r}\beta$, uzyskujemy miarę umożliwiającą eliminację punktów nie należących do krawędzi (4.12):

$$\frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\mathbf{r}\beta + \beta)^2}{(\mathbf{r}\beta)^2} = \frac{(\mathbf{r} + 1)^2}{\mathbf{r}}. \quad (4.12)$$

gdzie: $\mathbf{r}$ - stosunek wartości własnych macierzy Hessego.

Należy wziąć pod uwagę tylko stosunek wartości własnych, a nie ich indywidualne wartości. Wielkość $(\mathbf{r} + 1)^2 / 2\mathbf{r}$ jest minimalna, gdy dwie wartości własne są równe i rosnące wraz z $\mathbf{r}$. Dlatego, aby sprawdzić, czy stosunek własności własnych macierzy hesjanu jest poniżej pewnego progu $\mathbf{r}$, wystarczy tylko sprawdzić zależność (4.13):

$$\frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} < \frac{(\mathbf{r} + 1)^2}{\mathbf{r}}. \quad (4.13)$$

Powyższe podejście jest wyjątkowo efektywne obliczeniowo, wymagając mniej niż 20 operacji zmiennoprzecinkowych w celu weryfikacji każdego punktu kluczowego, oferując wysoką wydajność przy minimalnym nakładzie obliczeniowym.

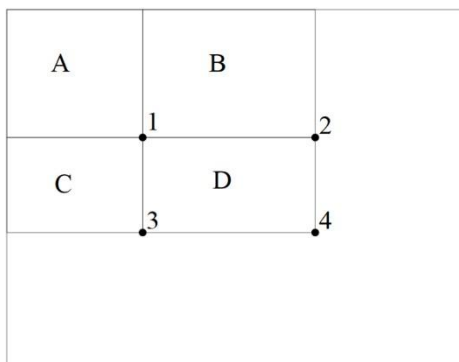
Detektor SURF

Detektor SURF (*Speeded Up Robust Features*) jest to algorytm powierzchniowy powstały na bazie algorytmu SIFT i przedstawiony przez Bay i in. w 2006 roku. Założeniem jego miało być szybsze działanie w porównaniu do detektora SIFT i wykorzystanie do jego działania macierzy Hessego.

Działanie detektora składa się z dwóch części:

1. Wyznaczenie obrazów całkowych (*integral*).

Przyspieszenie wykrywania punktów charakterystycznych poprzez zastosowanie obrazów *integral* (Viola i Jones, 2001) przekształconych z postaci pikselowej do postaci cech. Obraz całkowy powstaje w wyniku sumowania pikseli znajdujących się powyżej i z lewej strony punktu wyznaczanego. Zasada obrazu *integral* przedstawiona jest na rysunku 4.7. Wartość obrazu *integral* w punkcie 1 jest sumą pikseli w kwadracie A, w punkcie 2 jest równa sumie wartości w kwadratach A i B, w punkcie 3 równa się sumie A i C, a w punkcie 4 wynosi $A+B+C+D$ (Bay i Ess, 2008).



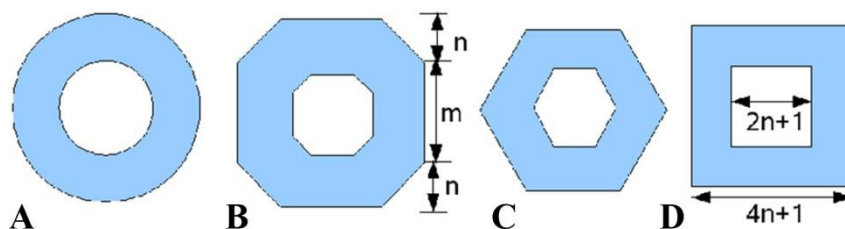
Rys.4.7 Zasada przekształcenia obrazu pikselowego w obraz *integral* (Bay i Ess, 2008)

2. Lokalizacja punktów charakterystycznych.

W celu przyspieszenia czasu wyszukiwania punktów oraz poprawę dokładności ich detekcji autorzy wykorzystali macierz Hessianu. Wykryte punkty charakterystyczne odpowiadają maksimum wyznacznika macierzy. Punkty wyznaczane są przy użyciu okna poszukiwawczego o wymiarze 9x9 pikseli. W oknie tym badane są wartości σ w najniższym poziomie piramidy. Przetwarzanie zdjęć detektorem SURF wykonywane jest filtrami prostokątnymi, a dzięki zastosowaniu obrazów *integral* nie ma potrzeby skalowania filtrów na różnych poziomach jego piramid. W pierwszym etapie przyjmuje się wartość maski 9x9 oraz wartość skali $s=1.2$, które odpowiadają parametrowi filtracji Gaussowskiej $\sigma = 1.2$. Wraz ze wzrostem poziomu piramidy obrazu zwiększana jest wielkość maski, kolejno do rozmiarów 9x9, 15x15, 21x21, 27x27 np. Również wartość współczynnika σ – np. dla filtru 27x27 ulega zmianom i będzie zwiększona $\sigma=3 \times 1.2 = 3.6$ (Bay i Ess, 2008).

Detektor CenSurE

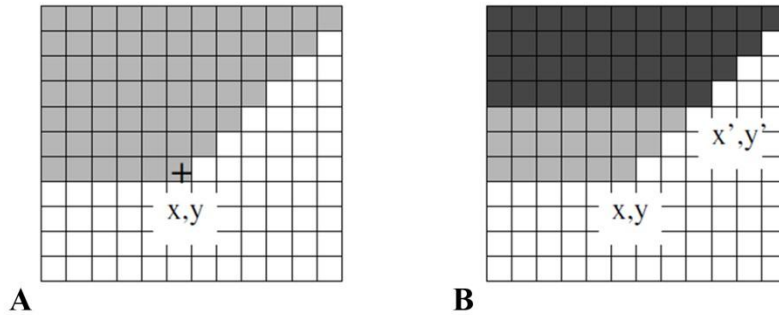
W roku 2008 autorzy Agrawal, Konolige i Blas przedstawili zasadę działania detektora powierzchniowego CenSurE (*Center Surround Extremas*). W przeciwieństwie do poprzednich detektorów detektor ten wykonuje przetworzenia obrazowe każdego piksela przy pełnej rozdzielczości zdjęcia, ze względu na to, że maksymalna wartość laplasjanu na różnych poziomach piramid jest stała. Do filtracji obrazowej zastosowano filtry klasy *Center – Surround*. Detektor CenSurE opiera się na algorytmie detekcji narożników Harrisa dając w ten sposób lepsze wyniki niż wyznacznik hesjanu. Algorytm CenSurE do wyznaczania wartości maksymalnych na zdjęciu wykorzystuje przybliżenie laplasjanu w postaci filtru *Center – Surround* (rysunek 4.8). Takie podejście pozwala na przyspieszenie procesu obliczeniowego oraz uniezależnienie od wpływu obrotu.



Rys.4.8 Rodzaje filtrów dwupoziomowych *Center – Surrounded*: A) kołowy filtr BloG (Bi – level LoG), B) filtr ośmiokątny, C) filtr sześciokątny i D) filtr prostokątny (Agrawal i.in., 2008)

Działanie detektora składa się z pięciu etapów:

- detekcja ekstremów funkcji,
- wyznaczenie punktów charakterystycznych przy zastosowaniu filtrów dwupoziomowych (tzw. Bi – *level filters*),
- porównaniu pikseli w obrębie najbliższego otoczenia,
- filtracja punktów położonych na krawędziach i liniach,
- przetworzenie zdjęcia pikselowego do zdjęcia *integral* z wykorzystaniem maski prostokątnej (rysunek 4.9).



Rys. 4.9 Przykład wygenerowanych obrazów trapezoidalnych typu integral: A) Skośne zdjęcie *integral*, gdzie piksel x,y jest sumą zaciemnionych obszarów, B) Wynik różnicy dwóch trapezoidalnych obrazów typu integral (Agrawal i.in., 2008)

Detektor ASIFT

Główną ideą algorytmu ASIFT (*Affine Scale-Invariant Feature Transform*) jest wygenerowanie zestawu symulowanych (wirtualnych) obrazów przejściowych na podstawie oryginalnych obrazów wejściowych (Yu i Morel, 2011; Markiewicz i in., 2019). Obrazy te powstają poprzez modyfikację kątów skręcenia, pochylenia i powiększenia oryginalnego zdjęcia. Następnie algorytm SIFT jest stosowany do wszystkich wygenerowanych obrazów wirtualnych. Dzięki temu ASIFT uwzględnia wszystkie sześć parametrów przekształcenia afinicznego i gwarantuje pełną niezależność względem tego przekształcenia. W algorytmie ASIFT przyjęty jest następujący model kamery oraz model afiniczny (4.14) (w poniższych wzorach odnoszących się do detektora ASIFT symbolika została zaczerpnięta z Yu i Morel, 2011):

$$\mathbf{u} = \mathbf{S}_1 \mathbf{G}_1 \mathbf{A} \mathbf{T} \mathbf{u}_0 \quad (4.14)$$

gdzie:

- $\mathbf{u}$ to obraz cyfrowy;
- $\mathbf{u}_0$ to (idealny) nieskończony obraz z widokiem frontalnym płaskiego obiektu;
- $\mathbf{T}$ i $\mathbf{A}$ to odpowiednio translacja płaszczyzny i mapowanie projektowe płaskie związane z ruchem kamery;
- $\mathbf{G}_1$ to konwolucja Gaussowska modelująca rozmycie optyczne;
- $\mathbf{S}_1$ to standardowy operator próbkowania na regularnej siatce o oczku 1.

Poniżej podany jest zapis ogólny przekształcenia afinicznego obrazu wraz z translacją (transformacja punktów obrazu). Wzór ten symuluje efekt zmiany punktu widzenia kamery (inaczej mówiąc: nachylenie płaszczyzny obrazu względem obserwatora).

$$U(x, y) \rightarrow u(ax + by + e, cx + dy + f) \quad (4.15)$$

gdzie: x, y – położenie punktu na obrazie; a, b, c, d – współczynniki macierzy afinicznej; e, f parametry przesunięcia w kierunku osi x i y .

W celu wyznaczenia wartości afinizmu stosowany jest wzór 4.16:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = H_\lambda R_1(\psi) T_t R_2(\phi) = \lambda \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \quad (4.16)$$

- gdzie: $\lambda > 0$ to wyznacznik macierzy A ,
- R_i to macierze obrotu, $\phi \in [0, \pi)$,
- a T_t to nachylenie kamery,
- (ψ) to obrót kamery, a (λ) współczynnik przybliżenia (*zoom factor*).

Możliwe jest rozłożenie parametrów ruchu kamery na dwa kąty punktu widzenia ϕ (długość geograficzna) i θ (szerokość geograficzna). Kąt punktu widzenia θ oblicza się według wzoru (4.17)

$$\theta = \arccos(1/t), \quad (4.17)$$

gdzie: t to parametr nachylenia kamery względem płaszczyzny obrazu.

W algorytmie ASIFT obrazy są poddawane rotacji o kąt θ , który jest reprezentowany przez parametr nachylenia t (4.18)

$$t = 1/\cos(\theta). \quad (4.18)$$

W algorytmie ASIFT wpływ nachylenia (rotacji szerokości geograficznej) jest wykonywany poprzez próbkowanie t i konwolucję Gaussowską ze standardowymi odchyleniami σ według wzoru 4.19:

$$\sigma = c(t^2 - 1)^{-2} \quad (4.19)$$

gdzie: c to stały próg kontrolny używany w eliminacji punktów leżących na krawędziach i przyjmuje on wartość 0.8.

Przyjmuje się, że szerokości geograficzne θ są próbkowane jako seria geometryczna $1, a, a^2, \dots, a^n$, $a > 1$, a wybór $a = (2)^{-2}$ stanowi dobry kompromis między dokładnością a liczbą iteracji. Autorzy algorytmu ASIFT zaproponowali wartość n równą 5, co skutkuje symulacją nachylenia 32 razy. W przypadku długości geograficznych ϕ , stosuje się serię arytmetyczną $0, b/t, \dots, kb/t$ z $b \cong 2\pi/5$ i $kb/t < \pi$. Po procesie generowania obrazów wirtualnych (który obejmuje pochylenie, nachylenie i obrót), można użyć dowolnego detektora, np. takiego jak SIFT lub SURF. W tym przypadku był to detektor SIFT.

Zaproponowane tworzenie symulowanych (wirtualnych) obrazów w przypadku detektora ASIFT można zaproponować dla pozostałych algorytmów tworząc ich afiniczne odpowiedniki: AFAST, ASURF itp.

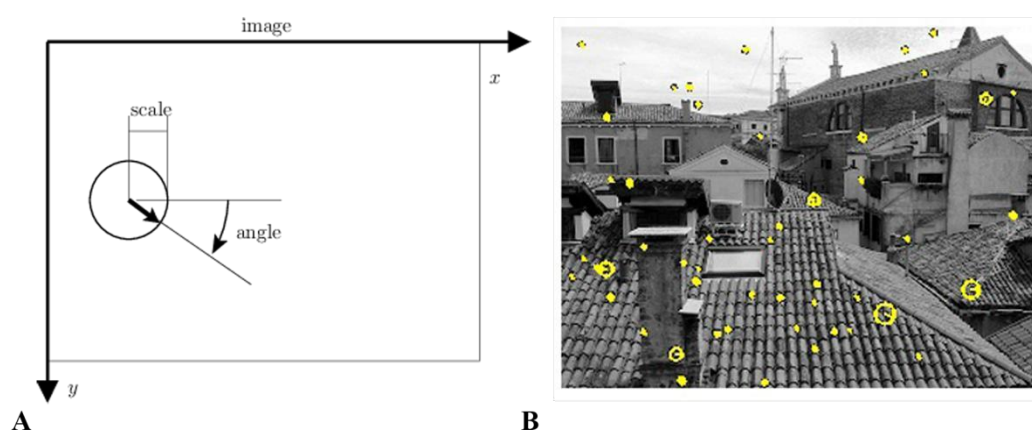
4.3 Deskryptory

Dopasowanie punktów charakterystycznych na kilku zdjęciach wymaga zastosowania opisu cech punktu na podstawie jego najbliższego otoczenia wyznaczanych niezależnie dla każdego zdjęcia (Brown i Lowe, 2002). Opis ten wykonuje się w oparciu o deskryptory pozwalające na wyznaczenie cech niezmienniczych szukanych punktów. W tym celu wykorzystuje się algorytmy bazujące na zmianach gradientów stopni szarości w otoczeniu punktu charakterystycznego bazujące jedynie na zdjęciu wejściowym (deskryptory typu *Hand-Crafted*) oraz w oparciu o charakterystyczne zależności wyznaczone w oparciu o wstępnie wytrenowane sieci neuronowe (deskryptory typu *Learned-Based*). W poniższym rozdziale opisane zostanie działanie deskryptora SIFT (*Hand-Crafted*) oraz deskryptorów HardNet, HardNet8, SOSNet, HyNet, MKDDescriptor, TFeat zaliczanych do grupy algorytmów *Learned-Based*.

4.3.1 Deskryptor *Hand-Crafted* - SIFT

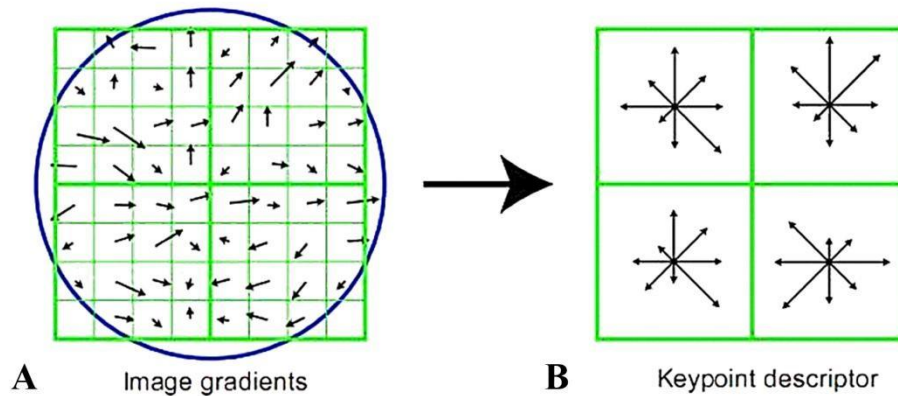
Główną ideą działania deskryptora SIFT (Brown i Lowe, 2002) jest obliczanie lokalnych gradientów obrazu na wybranym poziomie piramidy skalowej w otoczeniu kluczowego punktu badanego regionu (rysunek 4.10). Działanie deskryptora opiera się na analizie histogramów orientacji w sąsiedztwie podzielonym na siatkę 4×4 komórek, z których każda zawiera 8 przedziałów orientacji. Histogramy te są tworzone na podstawie próbkowania wartości

gradientów (wielkości i kierunku) w obszarze 16×16 pikseli wokół kluczowego punktu, przy czym dla każdego histogramu analizowany jest podobszar 4×4 pierwotnego sąsiedztwa. Takie podejście umożliwia eliminację punktów charakterystycznych, które nie spełniają kryteriów „dobrego detektora”. Oznacza to, że choć dany punkt może być lokalnie inwariantny, to globalnie może wykazywać podatność na zmiany kontrastu, przesunięcia wzdłuż linii czy drobne zmiany gradientów jasności, co mogłoby prowadzić do niejednoznaczności w późniejszych etapach dopasowywania par punktów.



Rys. 4.10 Przykłady opisu wykrytych punktów za pomocą deskryptora: A) wyznaczenie gradientu (skali) i orientacji dla pojedynczego punktu kluczowego, B) wyznaczenie parametrów wektora cech punktów charakterystycznych obrazu, zależnych od najbliższego otoczenia badanego punktu (vflfeat.org, 2024)

W wyniku opisu cech w oparciu o ww. sąsiedztwo generowany jest 128-parametrowy ($4 \times 4 \times 8 = 128$ elementów) wektor cech (rysunek 4.11) wykorzystywany na etapie wyszukiwania potencjalnych punktów wiążących (*descriptor matching*), w oparciu o analizę funkcji kosztu miar podobieństwa porównywanych deskryptorów. W celu zmniejszenia wpływu oświetlenia wektor ten jest normalizowany. Detekcja i opis cech punktów charakterystycznych za pomocą deskryptora stanowi podstawę do wyznaczenia punktów wiążących na zdjęciach. Dopasowanie wykrytych punktów poszczególnych detektorów na danych testowych można wykonać np. metodą *Brute – Force Matching* (OpenCV, 2024).

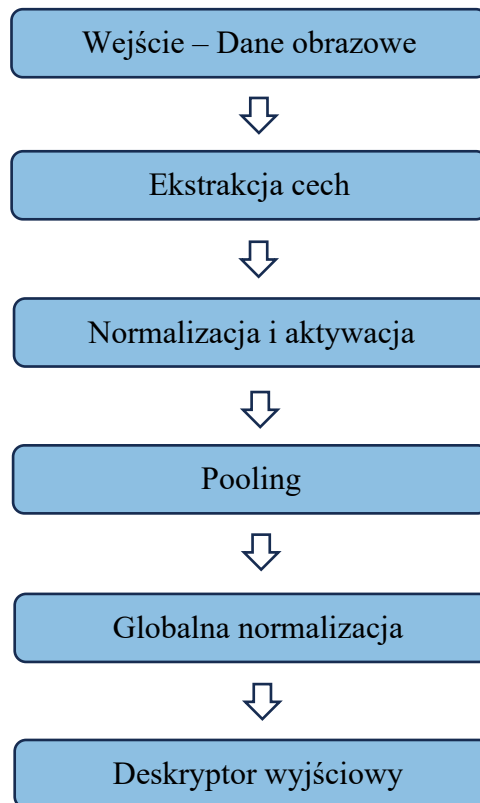


Rys. 4.11 Opis punktów charakterystycznych za pomocą deskryptor SIFT: A) zorientowane gradienty z otoczenia punktu kluczowego połączone wewnątrz regionów 4x4, B) opis wektorowy powstały w wyniku połączenia (Lowe, 2004)

4.3.2 Deskryptory *Learned-Based* (HardNet, HardNet8, SOSNet, HyNet, MKDDescriptor, TFeat)

Zasada działania sieci neuronowych w kontekście Deskryptorów

W kontekście deskryptorów, sieci neuronowe odnoszą się do ekstrakcji cech z danych obrazowych, które mogą być następnie używane do identyfikacji wzorców, rozpoznawania obiektów czy innych zadań związanych z widzeniem maszynowym. Deskryptory to wektory cech opisujące istotne informacje zawarte w danych obrazowych. Deskryptory w kontekście sieci neuronowych bazują na sieciach konwolucyjnych (rys. 4.12).

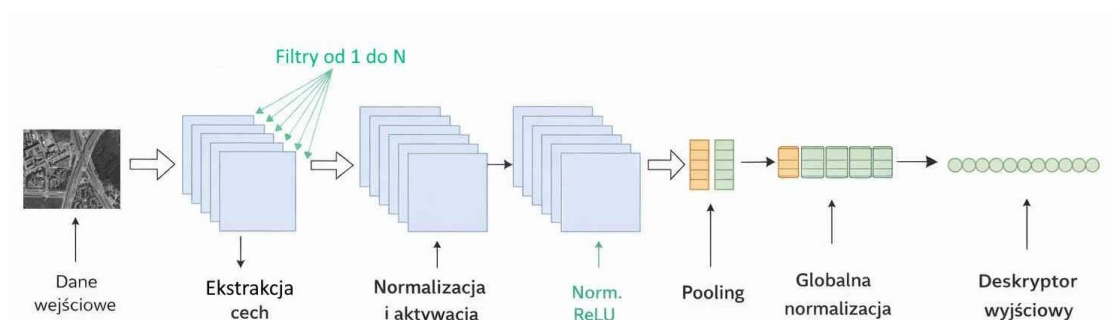


Rys. 4.12 Schemat zasady działania sieci neuronowych w kontekście deskryptorów

Rozważając przypadek ogólny wykorzystania sieci neuronowych w opisie cech wykorzystuje się następujące etapy zwane popularnie warstwami (Rys. 4.13):

- Wejście*: na wejściu sieć neuronowa przyjmuje dane w postaci fragmentów obrazów reprezentowanych przez określoną wielkość macierzy pikseli;
- Ekstrakcja cech*: Następnie wejściowe fragmenty obrazów analizowane są w oparciu o warstwy konwolucyjne sieci neuronowej, wykrywając w nich różne cechy i wzorce;
- Normalizacja i aktywacja*: Po warstwach konwolucyjnych często stosuje się normalizację i funkcję aktywacji, takie jak ReLU (*Rectified Linear Unit*), które pomagają w regulacji sygnału i wprowadzają nieliniowość;
- Pooling*: Wykonanie redukcji wymiarowości obrazu w celu zmniejszenia rozmiaru przestrzennego map cech z jednoczesnym zachowaniem najważniejszych wartości. Ułatwia to porównywanie i dopasowywanie obiektów czy obszarów na obrazie, a zarazem przyspiesza proces uczenia;

- e. *Globalna normalizacja*: Dane są przekształcane w skali globalnej, tj. na podstawie wartości całego zbioru danych. Pomaga to w utworzeniu deskryptorów znormalizowanych;
- f. *Deskryptor wyjściowy*: Na wyjściu sieci powstaje deskryptor opisujący wykryte punkty charakterystyczne, czyli wektor cech o stałej długości, umożliwiając porównywanie i dopasowywanie tych samych cech na dwóch i więcej obrazach.



Rys. 4.13 Przykład wielowarstwowej sieci konwolucyjnej

Sieć L2-Net

Współczesne deskryptory wykorzystywane do opisu cech bazują na architekturze konwolucyjnej sieci neuronowej L2-Net (Tian i in., 2017). Sieć L2-Net (rys. 4.14), oparta na architekturze CNN (Convolutional Neural Network), została zaprojektowana w celu wykrywania i opisywania lokalnych cech obrazu (punktów kluczowych), które stanowią podstawę rozpoznawania obiektów 3D, ich dopasowania oraz rekonstrukcji.

Model L2-Net wyróżnia się kilkoma istotnymi cechami:

- posiada progresywną strategię próbkowania, pozwalającą na dostęp do wielu próbek treningowych w zaledwie kilku epokach,
- kładzie nacisk na względną odległość między deskryptorami (konceptcja dopasowywania lokalnych fragmentów),
- posiada dodatkowy nadzór nad mapami cech pośrednich,
- uwzględnia zwartości deskryptora.

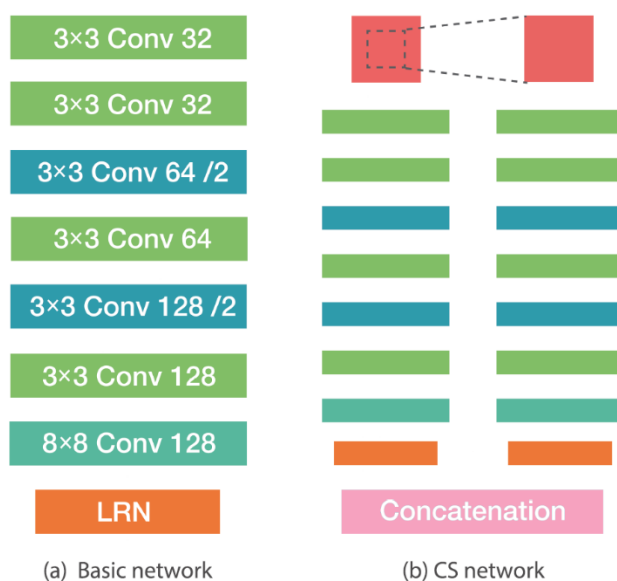
Deskryptor L2-NET wziął nazwę od odległości euklidesowej **L2** dopasowywanych cech w przestrzeni euklidesowej. Odległość euklidesowa **L2**, to miara odległości między dwoma wektorami cech (czyli linii prostej między nimi). Używana jest do mierzenia podobieństwa między wektorami cech lub do minimalizacji błędu prognozy.

Dla dwóch punktów $A(a_1, a_2 \dots a_n)$ i $B(b_1, b_2 \dots b_n)$ w n -wymiarowej przestrzeni euklidesowej, odległość $L2$ jest obliczana za pomocą następującego wzoru (4.20) (w poniższych wzorach odnoszących się do sieci neuronowej L2-Net symbolika została zaczerpnięta z Tian i in. 2017):

$$L2(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \quad (4.20)$$

gdzie: A, B – to dwa wektory deskryptorów o długości n ,

a_i, b_i – i -te współrzędne (składowe) wektorów A i B .



Rys. 4.14 Architektura sieci L2-Net, (a) – podstawowa sieć L2-Net (*Basic network*), (b) – rozszerzona wersja sieci L2-net (*Central-Surround network*) (Tian i in., 2017)

Architektura L2-Net jest przedstawiona na rysunku 4.14 a. Model ten składa się wyłącznie z warstw konwolucyjnych, a próbkowanie w dół osiągnięte jest poprzez konwolucję z krokiem równym 2. Po każdej warstwie konwolucyjnej wykonywana jest normalizacja wsadowa BN (*Batch Normalization*), jednak z niewielkimi modyfikacjami – parametry wagowe i przesunięcia warstw BN nie są aktualizowane, lecz ustalone odpowiednio na wartości 1 i 0. Ponieważ normalizacja jest istotnym krokiem w projektowaniu deskryptorów, jako warstwa wyjściowa wykorzystywana jest warstwa Normalizacji Lokalnej Odpowiedzi (LRN – *Local Response Normalization*), w celu uzyskania deskryptorów o jednostkowej długości.

L2-Net przekształca wejściowe fragmenty obrazu o rozmiarach 32×32 pikseli na deskryptory o wymiarze 128 (takim samym wymiarze jak stosowanym w deskrypcie SIFT). Implementowana jest również część L2-Net o strukturze centralno-obwodowej (CS – *Central-*

Surround). Struktura ta stanowi połączenie dwóch odrębnych L2-Netów ułożonych w układzie dwukolumnowym, jak przedstawiono na rysunku 4.14 b. Wejście (początek procesu) dla kolumny po lewej stronie jest takie same jak dla pojedynczego L2-Net, natomiast wejście do prawej kolumny generowane jest poprzez wycięcie centralnej części oryginalnego fragmentu obrazu (czerwony kwadrat na rys. 4.14b) i przeskalowanie jej do rozmiaru 32×32 . Lewa kolumna otrzymuje oryginalny *patch* o wymiarach 32×32 , natomiast prawa kolumna – środkową część tego *patch'a*, powiększoną z powrotem do rozmiaru 32×32 . Wyniki obu kolumn są następnie łączone (*concatenation*) w jeden większy deskryptor. Wykorzystując architekturę L2Net stosuje się następujące przetworzenia na różnych etapach działania sieci neuronowej:

- Normalizacja – normalizacja jest techniką stosowaną w sieciach neuronowych w celu poprawy procesu uczenia;
- BN (*Batch Normalization*) i LN (*Layer Normalization*) – te warstwy normalizacyjne pomagają zniwelować problem niestabilności numerycznej i przyspieszyć uczenie się sieci neuronowych;
- FRN – (*Filter Response Normalization*) – technika normalizacji FRN została zaprojektowana w celu eliminacji zależności od *batch'a* (zbiór próbek danych wejściowych równocześnie przetwarzanych przez sieć neuronową w jednym kroku) procesie treningu sieci neuronowych. Tradycyjne techniki normalizacji, takie jak BN, polegają na normalizacji aktywacji wewnątrz wsadu podzbioru danych treningowych (*batch'a*) podczas treningu natomiast FRN normalizuje aktywacje na poziomie filtra (kanału) bez stosowania statystyk wsadu, co sprawia, że model jest bardziej stabilny i łatwiejszy do szkolenia na różnych danych;
- Patch – w analizie obrazu *patch* odnosi się do małego fragmentu obrazu np. 3×3 . Fragmenty są wejściami dla konwolucyjnych warstw, co umożliwia rozpoznanie lokalnych cech obrazu;
- Konwolucyjne warstwy – konwolucyjne warstwy przekształcają wejściowe obrazy poprzez stosowanie filtra na jego lokalnych fragmentach. Fragmenty (*Patches*) te pozwalają wydobyć różne cechy, takie jak krawędzie, tekstury czy punkty kluczowe;
- TLU (*Threshold Logic Unit*) – prosta jednostka logiczna, która przyjmuje różne sygnały wejściowe i na podstawie przyjętego progu decyduje, czy wyjście powinno być

aktywowane (czy w danym obszarze obrazu występują punkty kluczowe w oparciu o analizy zmian gradientu intensywności pikseli z pewnym ustalonym progiem);

- Nieliniowość ReLU (*Rectified Linear Unit*) to funkcja aktywacji używana w sieciach neuronowych. Jest to prosta, nieliniowa funkcja, która przekształca wartość wejściową zgodnie z poniższym równaniem (4.21):

$$f(x) = \max(0, x), \quad (4.21)$$

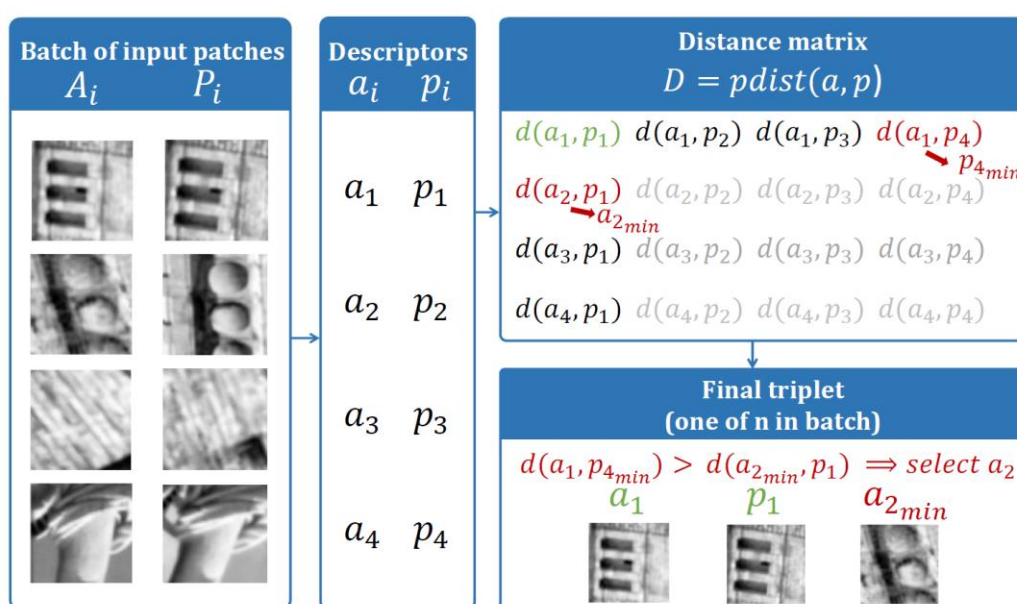
gdzie: x wartość funkcji;

- Oznacza to, że dla wartości wejściowej x , funkcja ReLU zwraca x , jeżeli x jest dodatnie, a dla x ujemnego zwraca 0. Warto zauważyć, że ReLU jest nieliniową funkcją aktywacji, co pozwala sieciom neuronowym modelować bardziej złożone zależności między danymi. Zastosowanie ReLU jako funkcji aktywacji ma kilka zalet, takich jak łatwość obliczeń (względem innych funkcji aktywacji), przyspieszanie procesu uczenia oraz rozwiązywanie problemu zanikającego gradientu. Niemniej jednak, może istnieć sytuacja, zwłaszcza w głębokich sieciach, w której neurony z funkcją ReLU mogą stać się „martwe” (nieaktywne) dla pewnych danych, co prowadzi do utraty informacji. W związku z tym, istnieją również modyfikacje funkcji ReLU, takie jak *Leaky ReLU*, *Parametric ReLU* (PreLU), czy *Exponential Linear Unit* (ELU), które próbują rozwiązać pewne wady funkcji ReLU;
- Regularyzacja *Dropout* to technika regularyzacji stosowana w sztucznych sieciach neuronowych w celu zapobiegania przeuczeniu się modelu. Metoda ta polega na losowym wyłączaniu (zerowaniu) pewnych neuronów w warstwach sieci podczas treningu. Innymi słowy, podczas każdej iteracji treningowej pewne neurony są tymczasowo ignorowane, co uniemożliwia im uczestniczenie w procesie aktualizacji wag.

Warstwy poolingu (*pooling layers*) są często używane do zmniejszenia wymiarowości danych, zachowując jednocześnie ważne cechy. Pomagają w redukcji liczby parametrów i obliczeń, a jednocześnie utrzymują ważne informacje przestrzenne. *Pooling* pomaga modelowi skoncentrować się na istotnych informacjach, co przyczynia się do niezależności od translacji, skalowania i wybranych deformacji. W detekcji obiektów czy klasyfikacji obrazów, warstwy poolingu pomagają w tworzeniu bardziej zgeneralizowanych i odpornych na zmienność reprezentacji cech.

Deskryptor HardNet

Sieć HardNet (*Hard Descriptor Network*) zaproponowany przez A. Mishchuka (2017), to deskryptor oparty na sieciach neuronowych, zaprojektowany do rozpoznawania i dopasowywania odpowiadających sobie punktów charakterystycznych na obrazach. Jego celem jest efektywne przekształcanie lokalnych punktów charakterystycznych, co umożliwia dokładne dopasowywanie obrazów pomimo znaczących zmian geometrycznych (np. zniekształceń perspektywicznych lub afinizmu) i radiometrycznych (np. różnic w oświetleniu). HardNet znajduje zastosowanie np. w rozpoznawaniu obrazów w dynamicznych środowiskach oraz w rekonstrukcji trójwymiarowych obiektów.



Rys. 4.15. Schemat próbkowania sieci HardNet (Mishchuk i in., 2017)

HardNet opiera się na architekturze L2-Net (Tian i in., 2017) oraz zbiorze uczącym *Phototour* (2024), który zawiera zdjęcia naziemne wykorzystywane do treningu modeli (zdjęcia wykonane z różnych perspektyw, przy zmiennym oświetleniu, skali i orientacji). Deskryptor ten generuje wektory cech (rys. 4.15) o długości 128, które opisują lokalne obszary obrazu o rozmiarach 32×32 piksele (tzw. *patches*). Wyodrębnione cechy są odporne na zmiany skali, rotację oraz różnice w oświetleniu, co czyni HardNet wysoce skutecznym narzędziem w systemach wizyjnych i robotyce autonomicznej. Jego dodatkową zaletą jest odporność na zniekształcenia i szumy, dzięki czemu sprawdza się także w analizie obrazów niskiej jakości. Został on również zoptymalizowany pod kątem szybkości przetwarzania obrazów w czasie rzeczywistym.

HardNet wykorzystuje architekturę syjamskiej sieci neuronowej (Vijay i.in. 2016), która uczy się optymalizować reprezentację cech w celu poprawy dokładności dopasowania. W trakcie treningu wykorzystuje się pary pasujących i niepasujących fragmentów obrazu, co umożliwia naukę określonych reprezentacji cech.

Szczegóły architektury deskryptora HardNet są następujące:

1. Dopełnianie zerami (*zero-padding*): Stosowane w każdej warstwie konwolucyjnej, z wyjątkiem ostatniej, w celu zachowania wymiarów przestrzennych danych wejściowych;
2. Brak warstw *pooling'u*: Warstwy *pooling'u* zostały pominięte, ponieważ ich użycie obniża wydajność deskryptora. W związku z tym redukcja rozmiaru przestrzennego danych wejściowych odbywa się wyłącznie poprzez kolejne operacje konwolucyjne;
3. *Batch Normalization* i nieliniowość ReLU: Po każdej warstwie konwolucyjnej, z wyjątkiem ostatniej, stosuje się normalizację wsadową oraz funkcję aktywacji ReLU;
4. Regularyzacja *dropout*: Wprowadzono *dropout* z prawdopodobieństwem 0.1 przed ostatnią warstwą konwolucyjną w celu redukcji ryzyka przeuczenia modelu;
5. Normalizacja wyjścia: Wyjście sieci jest normalizowane zgodnie z normą L2, co pozwala uzyskać deskryptor o stałej długości 128 jednostek.

Główne założenie architektury sieci HardNet (rys. 4.16) bazuje na kryterium dopasowania, wykorzystywanym w algorytmie SIFT, Ekstrakcja cech odbywa się na podstawie następującej procedury próbkowania, w której wykorzystywane są pary pozytywne oraz najtrudniejsze przykłady negatywne.

Architektura działania sieci HardNet przedstawiona jest na rysunku 4.16 i składa się z następujących etapów:

1. Generowane są fragmenty obrazu (*batch X*), zawierające lokalne dopasowania (*anchor*, A) i ich pozytywne odpowiedniki (*positive*, P). Fragmenty odpowiadają temu samemu punktowi w przestrzeni 3D;
2. Wszystkie fragmenty przechodzą przez sieć, co prowadzi do obliczenia macierzy odległości L2 między deskryptorami oznaczana jako (4.22) (w poniższych wzorach odnoszących się do deskryptor HardNet symbolika została zaczerpnięta z Mishchuk i in., 2017).

Proces ten można zapisać formalnie w postaci następujących zależności:

$$\mathbf{D} = \text{cdist}(\mathbf{a}, \mathbf{p}), \quad (4.22)$$

gdzie: $\mathbf{D}$ – macierz odległości między wszystkimi parami wektorów (deskryptorów) z dwóch zbiorów $\mathbf{a}$ i $\mathbf{p}$, $\text{cdist}(\mathbf{a}, \mathbf{p})$ – funkcja licząca odległość (zwykle euklidesową lub cosinusową) między każdym wektorem z macierzy $\mathbf{a}$ a każdym wektorem z macierzy $\mathbf{p}$. Następnie wyznaczana jest odległość euklidesowa $d(\mathbf{a}_i, \mathbf{p}_j)$ między deskryptorami $\mathbf{a}_i$ a $\mathbf{p}_j$ o rozmiarze $n \times n$ według wzoru (4.23):

$$d(\mathbf{a}_i, \mathbf{p}_j) = \sqrt{2 - 2\mathbf{a}_i \mathbf{p}_j}, \quad (4.23)$$

gdzie: $d(\mathbf{a}_i, \mathbf{p}_j)$ – to odległość euklidesowa między deskryptorami, a $\mathbf{a}_i$ i $\mathbf{p}_j$ oznaczają deskryptory fragmentów $\mathbf{A}_i$ i $\mathbf{P}_j$,

$\mathbf{a}_i$ – oznacza deskryptor zakotwiczenia (*anchor descriptor*), $\mathbf{p}_j$ – oznacza deskryptor pozytywny (*positive descriptor*);

3. Dla każdego deskryptora zakotwiczenia ($\mathbf{a}_i$) wyznaczany jest najbliższy niepasujący deskryptor (*negative descriptor*):

$\mathbf{p}_{j_{\min}}$ – najbliższy niepasujący deskryptor do $\mathbf{a}_i$, gdzie

$$j_{\min} = \text{argmin}_{j=1 \dots n, j \neq i} d(\mathbf{a}_i, \mathbf{p}_j)$$

$\mathbf{a}_{k_{\min}}$ – najbliższy niepasujący deskryptor do $\mathbf{p}_i$, gdzie

$$k_{\min} = \text{argmin}_{k=1 \dots n, k \neq i} d(\mathbf{a}_k, \mathbf{p}_i);$$

4. Na podstawie tych wyników konstruowane są trójkowe zestawy (*triplets*) do optymalizacji (4.24):

$$\begin{aligned} & \text{triplet}(\mathbf{a}_i, \mathbf{p}_i, \mathbf{p}_{j_{\min}}), \\ & \text{a jeśli } d(\mathbf{a}_i, \mathbf{p}_{j_{\min}}) < d(\mathbf{a}_{k_{\min}}, \mathbf{p}_i) \text{ to tworzony jest} \\ & \text{triplet}(\mathbf{p}_i, \mathbf{a}_i, \mathbf{a}_{k_{\min}}); \end{aligned} \quad (4.24)$$

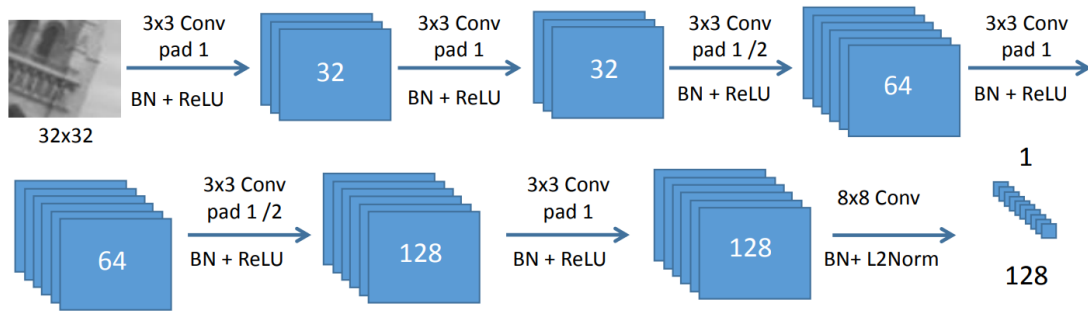
5. Funkcja straty oparta na marginesie *tripletowym* minimalizuje odległość między dopasowanymi deskryptorami, jednocześnie maksymalizując odległość do najbliższego niepasującego deskryptora (4.25):

$$L = \frac{1}{n} \sum_{i=1}^n \max(0.1 + d(a_i, p_i) - \min(d(a_i, p_{j_{\min}}), d(p_i, a_{k_{\min}}))) \quad (4.25)$$

gdzie: L – łączna strata dla batch'a, 0.1 – margines, ustalona stała wartość, która wymusza separację między pozytywną a negatywną parą, natomiast $\min(d(a_i, p_{j_{\min}}), d(p_i, a_{k_{\min}}))$ jest obliczane przed konstrukcją tripletu.

Obliczenia macierzy odległości są wykonywane na GPU, a jedynym dodatkowym kosztem w porównaniu do losowego próbkowania tripletów jest obliczanie macierzy odległości i minimalizowanie wartości wierszy i kolumn. Ponadto, w porównaniu do standardowego uczenia z tripletami, metoda ta wymaga tylko dwóch strumieni CNN, a nie trzech, co prowadzi do 30% mniejszego zużycia pamięci i obliczeń.

W przeciwieństwie do L2-Net, w metodzie tej nie korzysta się ani z głębokiego nadzoru dla warstw pośrednich, ani z ograniczenia korelacji wymiarów deskryptora. W trakcie przetwarzania nie zaobserwowano zjawiska nadmiernego dopasowywania.



Rys. 4.16. Architektura deskryptorów HardNet (Mishchuk i in. 2017)

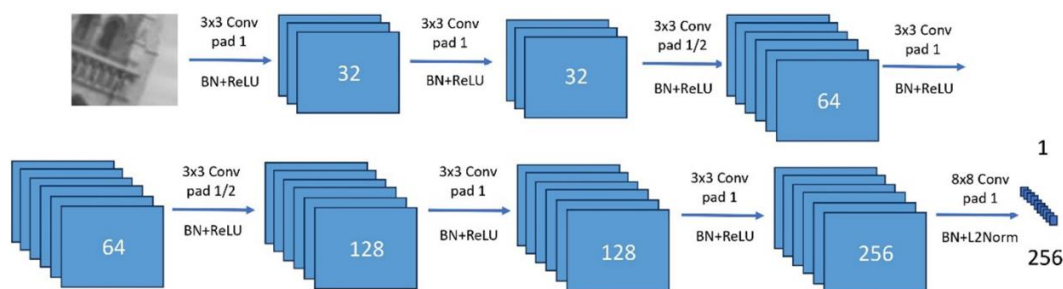
Deskryptor SOSNet

Deskryptor SOSNet (*Scale-Oriented Signature Network*) został opracowany w celu zwiększenia skuteczności opisu wykrywania punktów charakterystycznych w porównaniu z innymi metodami i jest modyfikacją sieci HardNet (Tian i in., 2019). SOSNet został zaprojektowany z myślą o poprawie dopasowywania punktów charakterystycznych. Poprzez uwzględnianie zmian skali (rozmiaru fragmentów obrazu), orientacji i informacjach przestrzennych, sieć ta ma potencjał dostarczania bardziej dokładnych i niezawodnych wyników w porównaniu do tradycyjnych metod.

W początkowej fazie algorytm na podstawie cech geometrycznych, takich jak zmiany intensywności obrazu (gradienty) oraz lokalnych zmiany struktury obrazu (np. narożniki i tekstury), identyfikuje punkty charakterystyczne na obrazie. Odbyna się to na podstawie ekstrakcji lokalnego otoczenia punktu zawierającego istotne informacje kontekstowe. Następnie na podstawie sieci neuronowej następuje przetworzenie lokalnego otoczenia każdego punktu w celu wyodrębnienia cech kluczowych. Cechy te reprezentują punkt w przestrzeni. Na podstawie wyekstrahowanych cech tworzony jest deskryptor każdego punktu za pomocą wektora liczbowego. Dalej deskryptory są normalizowane lub standaryzowane w celu poprawy ich niezawodności.

Deskryptor HardNet8

Zaprojektowana sieć HardNet8 (*Hard Descriptor Network 8*) (rys. 4.17) jest rozszerzeniem sieci konwolucyjnych L2Net, HardNet i SOSNet (Pultar, 2020). W przeciwieństwie do sieci HardNet do ekstrakcji cech na wyższym poziomie wykorzystuje aż osiem warstw konwolucyjnych. Dane wejściowe architektury HardNet mają rozmiar 32×32 piksele. Różni się on od HardNet tym, że po kroku 6 została dodana kolejna warstwa konwolucyjna o wymiarze $\text{Conv } 3 \times 3 \times 256$, aby uzyskać 256 elementowy deskryptor.



Rys. 4.17. Architektura deskryptorów HardNet8 (Pultar, 2020)

HardNet8 osiąga lepszą wydajność w stosunku do deskryptora HardNet i jest bardziej dokładny w skomplikowanych scenariuszach dopasowania obrazów, np. przy dużych różnicach oświetlenia, rotacji lub ich deformacjach. Architektury te jednak generują również dłuższy wektor wyjściowy, co ma wady, takie jak większe zużycie pamięci i wolniejsze wyszukiwanie najbliższego sąsiada.

Deskryptor HyNet

W przypadku HyNet (*Hybrid Network*) (rys 4.18), deskryptor ten jest poddawany regularyzacji, która normalizuje długość jego wektora cech do stałej wartości, co ułatwia proces uczenia i poprawia stabilność porównań między deskryptorami. Ponadto, architektura sieci HyNet wykonuje normalizację L2 na wszystkich pośrednich mapach cech (*patches*) oraz na wynikowych deskryptorach (Tian i.in., 2020). Natomiast normalizacja L2 jest techniką normalizacji danych, która sprowadza długość wektora do stałej wartości, co znacznie ułatwia proces uczenia się i poprawia ogólną wydajność sieci. HyNet poprzez dodanie warstwy ukrytej i konwertowanie zwykłego podobieństwa na podobieństwo hybrydowe ulepsza architekturę L2Net (Tian i.in., 2020).



Rys. 4.18 Architektura deskryptora HyNet

Sieć L2-Net pokazuje, że pośrednie mapy cech mogą być optymalizowane w taki sam sposób jak ostateczne deskryptory. Ponieważ mapy cech są również wektorami cech w przestrzeniach wielowymiarowych w sieci HyNet zamiast budować dodatkowe funkcje strat zaproponowano optymalne zarządzanie gradientami dla różnych warstw. Dlatego wygenerowano ortogonalne gradienty dla map cech wszystkich warstw poprzez normalizację L2 tak, aby można je było lepiej zoptymalizować pod względem kierunków naśladujących deskryptory.

W tym celu bezpośrednio zastosowano gotową normalizację odpowiedzi filtra FRN (*Filter Response Normalization*) w zadaniu klasyfikacji. Podstawowa idea FRN polega na L2 normalizacji pośrednich map cech w oparciu o możliwości nauczania parametrów afinicznych. W szczególności, FRN normalizuje każdą warstwę map cech według wzoru (4.26) (w poniższych wzorach odnoszących się do deskryptora HyNet symbolika została zaczerpnięta z Tian i.in., 2020):

$$\hat{f}_i = \gamma \sqrt{N} \frac{f_i}{\|f_i\|} + \beta \quad (4.26)$$

gdzie:

γ i β – parametry regularyzacji,

f_i – spłaszczona mapa cech,

N – liczba pikseli.

Architektura HyNet składa się z siedmiu warstw konwolucyjnych i generuje deskryptory o wymiarach 128. Jak pokazano na rysunku 4.18, wszystkie warstwy BN (*Batch Normalization*) z wyjątkiem ostatniej (przed ostateczną normalizacją L2 w oryginalnym L2-Net), zostały zastąpione warstwami FRN. Dodatkowo, po każdej warstwie FRN zastosowano funkcję aktywacji TLU (Thresholded Linear Unit) zamiast tradycyjnej funkcji ReLU (Rectified Linear Unit). Dzięki temu HyNet zachowuje tę samą liczbę wag konwolucyjnych, co HardNet i SOSNet.

Deskryptor MKD

MKD (*Multiple-Kernel local-patch Descriptor*) (Mukundan i.in., 2018) opiera się na masce deskryptora (*Hand-Crafted kernel descriptor*) zaproponowanego przez Bursuca i.in. (2015). Dzięki zastosowaniu różnych kombinacji parametrów deskryptora oraz różnych masek filtrów (jąder), MKD wykazuje wysoką odporność na różne rodzaje błędów w dopasowywaniu cech (Tolias i.in., 2015).

Patch, oznaczony symbolem P , reprezentowany jest jako zbiór pikseli $p \in P$ w jednym z obrazów. W celu porównania dwóch map cech P i Q , wykorzystywana jest funkcja podobieństwa zdefiniowana wzorem (4.27) (w poniższych wzorach odnoszących się do deskryptor HardNet symbolika została zaczerpnięta z Mukundan i.in., 2018):

$$M(P, Q) = \sum_{p \in P} \sum_{q \in Q} k(p, q), \quad (4.27)$$

gdzie: kernel k to funkcja podobieństwa zwykle nieliniowa, porównująca pary pikseli, P i Q – zbiory lokalnych punktów (cech) wyodrębnionych z dwóch fragmentów obrazu (patchy), $k(p, q)$ – funkcja kernela, która mierzy podobieństwo między dwoma punktami p i q .

W celu uproszczenia obliczeń, *Efficient Match Kernels* (EMK) wykorzystuje jawną mapę cech $\psi: \mathbb{R}^n \rightarrow \mathbb{R}^d$ do przybliżenia powyższego wyniku jako (4.28):

$$M(P, Q) = \sum_{p \in P} \sum_{q \in Q} k(p, q) \approx \sum_{p \in P} \sum_{q \in Q} \psi(p)^T \psi(q) = \sum_{p \in P} \psi(p)^T \sum_{q \in Q} \psi(q), \quad (4.28)$$

gdzie: $\psi(p)$ – funkcja mapująca punkt p do przestrzeni cech, w której można zamiast kernela $k(p,q)$ obliczyć zwykły iloczyn skalarny.

Dzięki temu każda mapa cech może być przekształcona w deskryptor oparty na funkcji jądra (*Kernel Descriptor*, KD) $V(P) = \sum_{p \in P} \psi(p)$.

Przybliżenie podobieństwa między patchami $M(P, Q)$ wyrażane jest w formie iloczynu skalarnego (4.29):

$$M(P, Q) \approx V(P)^T V(Q), \quad (4.29)$$

gdzie $V(P) \in \mathbb{R}^d$ – to tzw. globalny wektor cech (*embedding*) reprezentujący cały patch P , a $V(Q)$ – analogicznie, globalna reprezentacja patcha Q . W celu zapewnienia zgodności jednostkowego podobieństwa, deskryptor jest normalizowany zgodnie z normą L2 (4.30):

$$\bar{V}(P) = \gamma(P)V(P), \text{ gdzie } \gamma(P) = (V(P)^T V(P))^{-1/2}, \quad (4.30)$$

gdzie: $\bar{V}(P)$ – to znormalizowany wektor cech.

Kernel $k(p,q)$ (maska filtru) to sposób, w jaki oceniamy podobieństwo między pikselami. Funkcja ta bierze pod uwagę różne właściwości pikseli, takie jak ich położenie czy kierunek gradientu (zmiany jasności). Kernel można zapisać jako (4.31):

$$k(p, q) = k_1(p_1, q_1) k_2(p_2, q_2) \dots k_n(p_n, q_n), \quad (4.31)$$

gdzie każda funkcja k_n jest parametryzowaną funkcją podobieństwa dla skalarów p_n i q_n . Mapę cech ψ buduje się za pomocą iloczynu Kroneckera indywidualnych map cech ψ_n (4.32):

$$\Psi(p) = \psi_1(p_1) \otimes \psi_2(p_2) \dots \otimes \psi_n(p_n), \quad (4.32)$$

gdzie: p – punkt w lokalnym patchu, $\Psi(p)$ to całkowite odwzorowanie punktu p do przestrzeni cech deskryptora, $\otimes$ – operator łączenia (najczęściej iloczyn tensorowy), $\psi_i(p_i)$ – mapowanie cechy p do przestrzeni cech za pomocą funkcji kernela.

Iloczyn Kroneckera pozwala na reprezentację wielowymiarowych cech obrazu w formie tensora³, które zachowują informację o lokalnych różnicach w pozycji, kierunku i innych atrybutach (Shen, i.in., 2018).

³ Tensory są rozszerzeniem pojęcia wektora. Używa się ich do opisu pewnych zmian w danym ośrodku. W uproszczeniu tensor możemy utożsamić jako operator działający na wektor, w wyniku którego powstaje z niego

Do porównywania atrybutów pikseli, takich jak kierunek gradientu, wykorzystuje się funkcję matematyczną zwaną kernelem Von Misesa (Mardia i Jupp, 2000). Jest to funkcja podobna do rozkładu normalnego, która dobrze radzi sobie z danymi kątowymi, np. kierunkiem. Dzięki zastosowaniu tej funkcji z parametrem kappa κ , można regulować dokładność dopasowania – mniejsze κ oznacza większą tolerancję na różnice w kierunkach. Funkcja ta jest przybliżana za pomocą rozwinięcia Fouriera, co pozwala na szybkie obliczenia. Funkcja ta jest używana zarówno dla obrazów, jak i patchów, a jej przybliżenie oblicza się z użyciem rozwinięcia Fouriera (4.33):

$$k_{VM}(p_n, q_n) \approx \psi_{VM}(p_n)^T \psi_{VM}(q_n), \quad (4.33)$$

gdzie $\psi_{VM}: \mathbb{R} \rightarrow \mathbb{R}^{2N+1}$ jest mapą cech zależną od różnicy $\Delta_n = p_n - q_n$

Po skonstruowaniu deskryptora $V(P)$, jest on dodatkowo „centrowany” i rzutowany na mniejszy wymiar za pomocą metody PCA (*Principal Component Analysis*) lub metod nadzorowanych (Jégou i Chum, 2012). Umożliwia to usunięcie nadmiarowych informacji i skupienie się na najważniejszych cechach. Proces ten opisuje równanie (4.34):

$$\hat{V}(P) = A^T(\bar{V}(P) - \mu), \quad (4.34)$$

gdzie: μ to średnia wartość cech, a A to macierz zmieniająca wymiar danych. Na końcu deskryptor znów jest normalizowany (norma L2), aby wszystkie miały porównywalną długość.

Każda z parametryzacji wielokernelowego deskryptora lokalnego *patch* obrazu odnosząca się do położenia i kierunku gradientu zaproponowana przez autorów Mukundan i in., w 2018 roku, zapewnia odporność na inne rodzaje błędów jego orientacji. Parametryzacja polarna dla zakłóceń w dominującej orientacji, a parametryzacja kartezjańska dla niedokładnej lokalizacji punktu charakterystycznego. Te właściwości sprawiają, że MKD jest wszechstronnym narzędziem do analizy obrazów, szczególnie w kontekście błędów rejestracji cech.

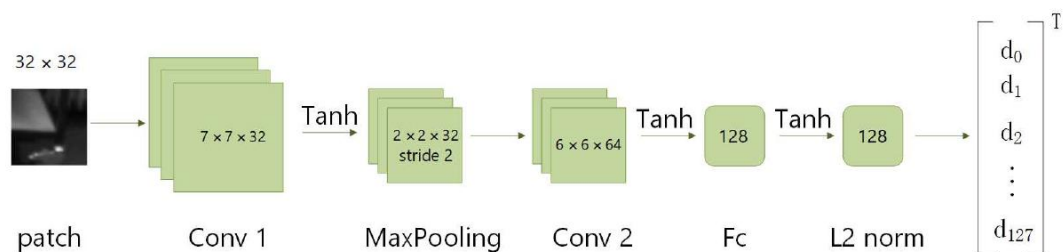
Deskryptor TFeat

Działanie algorytmu TFeat (*Triplet-based Feature Extraction*) koncentruje się na opracowaniu efektywnego i lekkiego podejścia do generowania deskryptorów cech lokalnych

nowy wektor o innym zwrocie, kierunku i wartości. W przetwarzaniu obrazowym tensor utożsamiany jest z obrazem jako wielowymiarowej tablicy liczb, gdzie dwa pierwsze wymiary odnoszą się do wymiaru macierzy obrazu, a trzeci do wartości kanałów RGB. Za pomocą tensorów wykonywane są takie operacje jak: konwolucja, normalizacja czy *pooling*.

z wykorzystaniem płytkiej sieci konwolucyjnej (Balntas i.in. 2016). TFeat został zaprojektowany z myślą o wysokiej wydajności i niskich wymaganiach obliczeniowych, co czyni go szczególnie użytecznym w zastosowaniach takich jak systemy SLAM, czy analiza obrazów wideo.

Główną właściwością algorytmu jest szybkie i dokładne dopasowywanie cech w dużych zbiorach danych wizualnych. Dla każdego wykrytego punktu kluczowego (np. przy użyciu detektora SIFT, FAST lub Harris), algorytm wyodrębnia niewielki fragment obrazu, który następnie jest przetwarzany przez konwolucyjną sieć neuronową. Wynikiem działania sieci jest znormalizowany wektor deskryptora, który numerycznie reprezentuje lokalną strukturę wokół danego punktu. Proces uczenia sieci oparty jest na funkcji straty typu *triplet loss*, co umożliwia skuteczne rozróżnianie cech należących do różnych punktów charakterystycznych i zwiększa odporność deskryptora na zakłócenia związanymi ze zmianą punktu widzenia kamery, zmianą oświetlenia elementów terenowych na zdjęciu oraz zmianą skali zdjęcia.



Rys. 4.19 Architektura deskryptora TFeat (Rong i in. 2019)

Architektura deskryptora TFeat (rys. 4.19) oparta jest na stosunkowo prostej konwolucyjnej sieci neuronowej, zaprojektowanej z myślą o wysokiej wydajności obliczeniowej. Struktura sieci przedstawia się następująco:

1. Wejście: patch 32×32.
2. Conv1 + ReLU + MaxPooling:
 - 32 filtrów o rozmiarze 7×7,
 - ReLU jako funkcja aktywacji,
 - MaxPooling 2×2.
3. Conv2 + ReLU + MaxPooling:
 - 64 filtry 6×6,
 - ReLU,

- MaxPooling 2×2.
4. Conv3 + ReLU:
- 128 filtrów 5×5,
 - ReLU.
5. Flatten + Normalizacja L2:
- Wynikowy wektor (rozmiar np. 128) jest normalizowany do długości 1 (L2).

Algorytm TFeat oparty na sieci neuronowej do dopasowywania cech obrazu wykorzystuje nowoczesne podejście do trenowania sieci z trójkami obrazów i zastosowanie specjalnych funkcji kosztu, co pozwala na uzyskanie szybkich, efektywnych i odpornych na przeuczenie deskryptorów cech. Działa to szczególnie dobrze w zadaniach dopasowywania punktów charakterystycznych między obrazami. Dzięki dużej szybkości ekstrakcji cech algorytm jest odpowiedni do pracy z dużymi zbiorami danych, a także do zastosowań wymagających szybkich wyników, takich jak systemy rozpoznawania obrazów czy śledzenia obiektów.

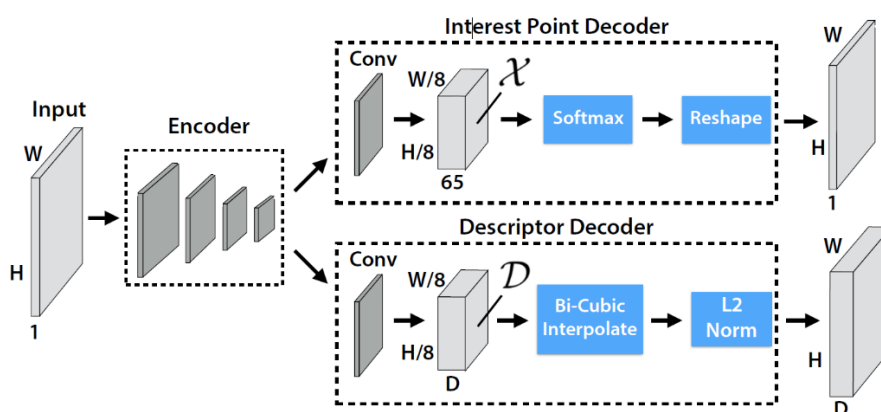
4.4 Metody łączone detekcji i opisu punktów Learned-Based

Detect-and-Describe: SuperPoint+SuperGlue

SuperPoint (De Tone i.in., 2018) to sieć w pełni konwolucyjna działająca na pełnowymiarowych obrazach, służąca zarówno do wykrywania punktów charakterystycznych, jak i do generowania ich deskryptorów w postaci wektorów o stałej długości. Sieć trenowana była zarówno na syntetycznym zbiorze danych oraz na rzeczywistych zdjęciach naziemnych ze zbioru MS-COCO (*Microsoft Common Objects in Context*) (De Tone i.in., 2018). Model sieci (rys. 4.20) posiada jeden dekodery mający zadanie przetwarzanie oraz redukcje wymiarów obrazu wejściowego. Następnie architektura sieci przechodzi w dwa dekodery specjalizujące się w różnych zadaniach: pierwszy odpowiada za wykrywanie punktów charakterystycznych IPD (*Interest Point Decoder*), a drugi za generowanie ich deskryptorów (*Descriptor Decoder*). W przeciwieństwie do innych sieci, gdzie na początku wykrywane są punkty, a w drugiej kolejności następuje ich opis, SuperPoint wykonuje te operacje jednocześnie dzięki zastosowaniu dwóch dekodery.

Dekoder używany w architekturze sieci wzorowany jest na architekturze enkodera VGG (*Visual Geometry Group*) w celu zmniejszenia wymiaru obrazu (Simonyan, i.in., 2014). Enkoder

w tej architekturze składa się z szeregu warstw konwolucyjnych, które odpowiadają za ekstrakcję kluczowych cech obrazu, oraz operacji *pooling*, które redukują wymiary przestrzenne map cech, umożliwiając hierarchiczne przetwarzanie danych wejściowych na różnych poziomach szczegółowości. Dodatkowo, zastosowanie nieliniowych funkcji aktywacji, takich jak ReLU (*Rectified Linear Unit*), pozwala sieci na modelowanie złożonych zależności i relacji między cechami w danych, co znacząco poprawia zdolność sieci w detekcji punktów charakterystycznych oraz efektywnej analizie obrazów o skomplikowanej strukturze.



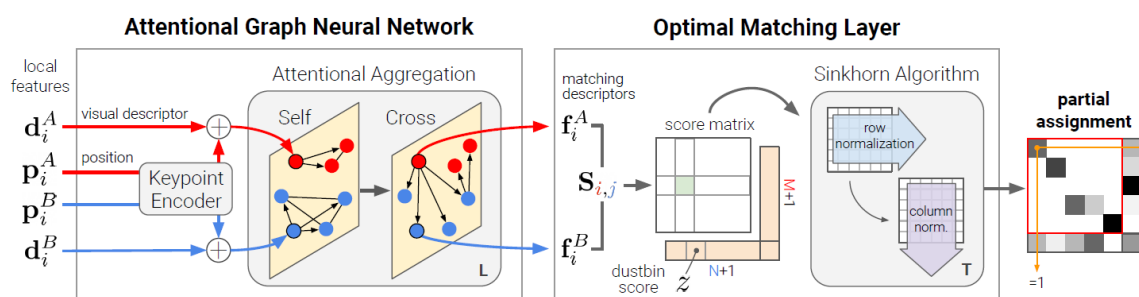
Rys. 4.20. Architektura Detektora SuperPoint (De Tone i.in., 2018)

SuperGlue (Sarlin i.in. 2019) wprowadza nowatorskie podejście do dopasowywania cech punktów charakterystycznych. W przeciwieństwie do metod opartych na metodach klasycznych (*Hand-Crafted Features*), SuperGlue wyznacza prawdopodobieństwa dopasowania cech i wykonuje ich dopasowanie bezpośrednio na podstawie danych wejściowych. Proces ten eliminuje potrzebę posiadania specjalistycznej wiedzy na temat projektowania deskryptorów i opiera się na głębokich sieciach neuronowych. Dzięki temu możliwe jest automatyczne i efektywne modelowanie zależności między cechami, co znacząco poprawia dokładność i wydajność dopasowywania punktów charakterystycznych. Detektor ten składa się z dwóch głównych części:

1. Sieć osadzająca (*Embedding Network*), odpowiedzialna przekształcanie punktów charakterystycznych z obu obrazów do wspólnej przestrzeni cech, tak aby odpowiadające sobie punkty miały podobne reprezentacje i możliwe było wyznaczenie miary podobieństwa pomiędzy nimi;

2. Moduł weryfikacji geometrycznej, który estymuje transformacje geometryczne pomiędzy punktami i ocenia dokładność ich dopasowania.

SuperGlue może korzystać z bogatych informacji kontekstowych i radzić sobie z trudnymi scenariuszami, takimi jak przesłonięcia widoczności obiektów na obrazie, zmiany kierunku widzenia czy różnice w oświetleniu, dzięki jednoczesnemu uczeniu reprezentacji cech i procesu ich dopasowywania.



Rys. 4.21. Architektura SuperGlue (Sarlin i.in. 2019)

Architektura SuperGlue (rys. 4.21) zawiera dwa kluczowe komponenty:

1. Sieć grafowa oparta na mechanizmach uwagi (*Attentional Graph Neural Network*), która obejmuje enkoder punktów kluczowych (*Keypoint Encoder*), przekształcający pozycje punktów p oraz ich wizualne deskryptory d w pojedynczy wektor reprezentacji f . Sieć wykorzystuje naprzemienne warstwy *Self-Attention* oraz *Cross-Attention*, powtarzane N -krotnie, aby uzyskać bardziej wyraziste reprezentacje cech;
2. Warstwa optymalnego dopasowania (*Optimal Matching Layer*), która generuje macierz wyników dopasowania $M \times N$, rozszerza macierz o dodatkowe elementy w celu stabilizacji dopasowania i optymalizuje dopasowanie za pomocą algorytmu Sinkhorna (Cuturi, 2013) przez T iteracji.

Warstwy *Self-Attention* pozwalają modelowi skoncentrować się na różnych częściach obrazu w zależności od kontekstu. Obraz jest podzielony na obszary (najczęściej nazywane "płaskimi obszarami" lub "płaskimi pudełkami"), a warstwa ta wykorzystywana jest w celu przypisania wag różnym obszarom obrazu na podstawie ich znaczenia w kontekście reszty obrazu. Dzięki temu model może skupić się na kluczowych cechach obrazu, które są istotne dla danego zadania, takiego jak klasyfikacja obiektów czy wyszukiwanie punktów kluczowych.

W przypadku warstwy *Cross-Attention* model koncentruje się na różnych źródłach informacji w jednym obrazie w zależności od informacji zawartej w innym obrazie lub innym źródle danych. Zależność tą wykorzystuje się w detekcji obiektów, gdzie model musi skoncentrować się na różnych częściach obrazu w zależności od tego, jakie obiekty są w nim obecne. Dzięki takim rozwiązaniom SuperPoint i SuperGlue stanowią zaawansowane narzędzia w dziedzinie detekcji oraz dopasowywania punktów charakterystycznych na obrazach.

Describe-then-Detect: LoFTR

LoFTR (*Local Feature Transformer*) to kompleksowa metoda opisu i detekcji lokalnego dopasowywania cech wprowadzona przez Sun J. i in. (2021). W przeciwieństwie do tradycyjnych podejść, które wykonują wykrywanie cech, ich opis i dopasowywanie w sposób sekwencyjny, LoFTR w pierwszym etapie tworzy wstępne dopasowania dla wszystkich pikseli obrazu, które następnie są poprawiane na poziomie szczegółowym. Architektura LoFTR opiera się na technologii Transformatorów, które pomagają w zrozumieniu globalnych zależności i informacji kontekstowych w obrazach. To podejście pozwala osiągnąć skuteczność nawet w obszarach o niskiej teksturze, gdzie tradycyjne detektory często zawodzą.

LoFTR różni się od klasycznych metod tym, że eliminuje potrzebę wyodrębniania i opisywania cech w oddzielnych krokach. Zamiast tego przewiduje bezpośrednie dopasowania pikselowe między obrazami, co znacząco upraszcza i przyspiesza cały proces. Dzięki zastosowaniu Transformatora LoFTR jest w stanie skutecznie wykrywać powiązania między punktami w różnych częściach obrazu (tzw. zależności dalekiego zasięgu) i wykorzystywać globalne informacje kontekstowe, co poprawia dokładność i niezawodność wyników.

LoFTR wykorzystuje funkcję uważalności na etapie zgrubnego dopasowywania obrazów, aby każdy piksel odnosił się do swoich sąsiadów i przechwytywał ich informacje kontekstowe w celu wygenerowania gęstego dopasowania pikselowego. Etap dopasowania zgrubnego służy do wstępnego oszacowania zgodności. LoFTR wykorzystuje sieć hierarchiczną do udoskonalania początkowych dopasowań. Sieć ta pobiera początkowe odpowiedniki i iteracyjnie je udoskonala, biorąc pod uwagę lokalne relacje przestrzenne i kontekst. Architektura oparta na transformatorach w LoFTR umożliwia przechwytywanie zależności dalekiego zasięgu, czyli relacji między punktami lub regionami obrazu, które są od siebie

znacząco oddalone w przestrzeni. Dodatkowo model integruje globalne informacje kontekstowe, co pozwala na bardziej kompleksową analizę całego obrazu. Takie podejście znacząco poprawia jakość generowanych gęstych powiązań pikselowych. Eliminuje również konieczność jawnego wykrywania cech, generując gęste deskryptory bezpośrednio na podstawie danych wejściowych. Dzięki temu LoFTR osiąga wyższą dokładność oraz lepszą wydajność procesu dopasowywania, szczególnie w przypadku złożonych obrazów o słabo wyrazistych lub niewyraźnych cechach.

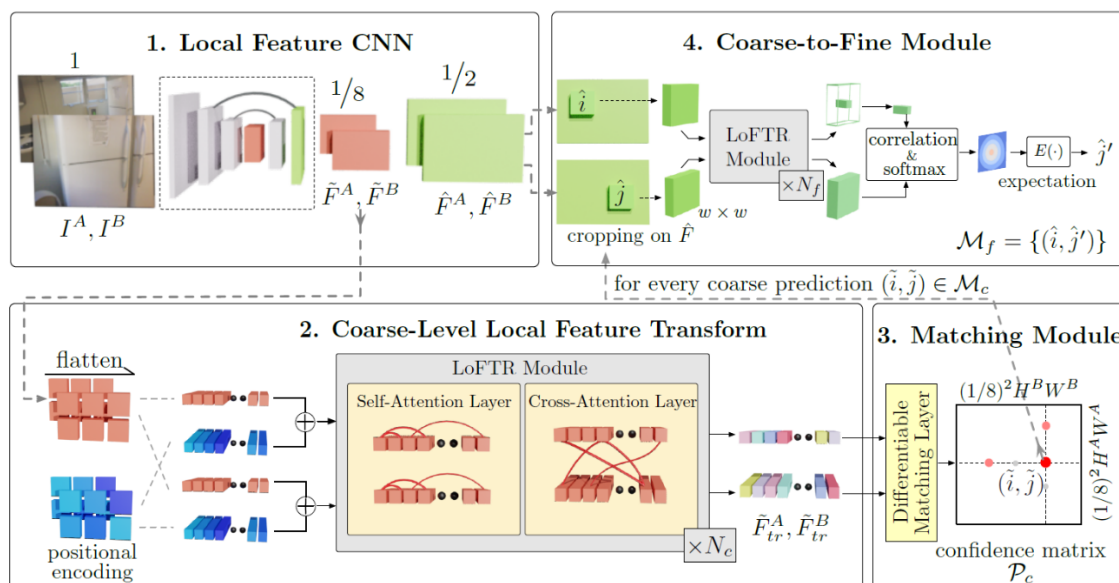
Kluczowe elementy działania LoFTR

1. Gęste dopasowania pikselowe: LoFTR nie korzysta z tradycyjnych detektorów cech, takich jak SIFT czy SURF, lecz na podstawie funkcji uwag⁴ przewiduje gęste dopasowania pikselowe bezpośrednio na obrazie wejściowym. To podejście pozwala na dokładne odwzorowanie nawet w trudnych warunkach, takich jak słaba tekstura obrazu (rys. 4.21);
2. Etap zgrubnego dopasowywania: LoFTR wykorzystuje mechanizm uwagi (*attention*), który pozwala modelowi analizować zależności między pikselami, uwzględniając ich kontekst lokalny. Dzięki temu możliwe jest wstępne oszacowanie zgodności pikseli między obrazami (rys. 4.21);
3. Hierarchiczne udoskonalanie dopasowań: Sieć hierarchiczna iteracyjnie udoskonala początkowe dopasowania, uwzględniając lokalne i globalne relacje przestrzenne między punktami kluczowymi. Proces ten poprawia zarówno dokładność, jak i wiarygodność końcowych wyników (rys. 4.21);
4. Architektura oparta na Transformatorze: Dzięki zastosowaniu Transformatora LoFTR może przechwytywać zależności dalekiego zasięgu i globalne informacje kontekstowe, co znacząco poprawia jakość gęstych powiązań pikselowych. Taka konstrukcja eliminuje konieczność jawnego wykrywania cech i bezpośrednio tworzy gęste deskryptory (rys. 4.21).

⁴ Funkcja uwagi służy do oceny przewidywania wyników modelu sieci neuronowej na podstawie dostępnych danych treningowych. Jest kluczowym elementem uczenia maszynowego umożliwiając porównanie różnych modeli i ich skuteczną ocenę (np. na podstawie błędu MSE (*Mean Squared Error*). Dostosowując odpowiednie parametry modelu minimalizuje się funkcję uwagi. Im niższa wartość funkcji uwagi, tym dokładniejszy model.

LoFTR (Sun i.in 2021) składa się z czterech komponentów (rys. 4.21):

- a. Lokalna funkcja CNN (*Local feature CNN*) – Komponent ten wyodrębnia z pary obrazów wejściowych I^A i I^B mapy cech na poziomie ogólnym $\tilde{F}^A$ i $\tilde{F}^B$ oraz szczegółowym $\hat{F}^A$ i $\hat{F}^B$;



Rys. 4.21. Architektura LoFTR (Sun i.in 2021)

- b. Transformer cech zgrubnych (*Coarse-Level Local Feature Transform*) – Zgrubne mapy cech są konwertowane do wektorów 1-D i łączone za pomocą kodowania pozycyjnego;
- c. Dodane funkcje są następnie przetwarzane przez moduł *Local Feature Transformer* (LoFTR), który zawiera warstwy samouwagi⁵ i krzyżowej⁶ NC (*Number of Cross-Attention Layers*);

⁵ Warstwa samouwagi (*self-attention layer*) jest podstawowym elementem takich modeli uczenia jak model GPT (*Generative Pre-trained Transformer*) czy BERT (*Bidirectional Encoder Representations from Transformers*). Warstwa samouwagi pozwala modelowi na efektywne analizowanie zależności między elementami wejściowymi odnoszącymi się do całej sekwencji, a nie jak tradycyjnych sieci rekurencyjnych (RNN) lub konwolucyjnych (CNN), które przetwarzają dane lokalnie lub sekwencyjnie. Mechanizm samouwagi uwzględnia wszystkich pozostałe elementy w sekwencji podczas przetwarzania. Pozwala to modelowi podejmować decyzje na podstawie kontekstu globalnego, co jest szczególnie przydatne w zadaniach związanych z przetwarzaniem języka naturalnego.

⁶ Warstwa krzyżowej uwagi (*cross-attention layer*), w odróżnieniu od warstwy samouwagi, analizuje relacje nie tylko w obrębie jednej sekwencji wejściowej, ale także między różnymi sekwencjami. Jest to szczególnie przydatne w zadaniach, gdzie konieczne jest powiązanie różnych źródeł danych, na przykład w przetwarzaniu obrazów i tekstów. W takich zastosowaniach warstwa krzyżowej uwagi pozwala modelowi analizować związki pomiędzy opisem tekstowym a odpowiednimi regionami obrazu, co prowadzi do dokładniejszego i bardziej efektywnego przetwarzania informacji.

- d. Moduł dopasowania (*Matching module*) – Na podstawie różniczkowej warstwy dopasowania generowana jest macierz zaufania P_c (*Pairwise Confidence*), która następnie służy do wyznaczania wstępnych dopasowań przybliżonych M_c (*Matching Candidates*). Dopasowania są wybierane zgodnie z progiem zaufania i kryteriami najbliższego wspólnego sąsiada;
- e. Moduł przejścia (*Coarse-to-Fine Module*) – Każde przybliżone dopasowanie $(\tilde{\mathbf{I}}, \tilde{\mathbf{J}} \in \mathbf{M}_c)$ jest udoskonalane w oknie lokalnym wyciętym z mapy cech na poziomie szczegółowym. Proces ten prowadzi do uzyskania dopasowań o dokładności subpiksela (M_f).

LoFTR jest jednym z najbardziej zaawansowanych podejść w dziedzinie dopasowywania cech obrazowych, wykorzystuje mechanizm uwagi, który umożliwia uwzględnianie zarówno lokalnego, jak i globalnego kontekstu. Dzięki eliminacji potrzeby jawnego wykrywania punktów kluczowych oraz bezpośredniemu tworzeniu gęstych dopasowań, metoda ta wykazuje wysoką odporność na trudne warunki, takie jak zmiany perspektywy, przestonięcia czy brak wyraźnych tekstur.

5. Metodyka opracowania sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych

Dotychczas stosowane metody orientacji archiwalnych zdjęć lotniczych opierały się na aerotriangulacji opartej na „klasycznym pomiarze” - bazującym na początkowych założeniach, tj. znajomości elementów orientacji wewnętrznej uzyskiwanej z metryki kamery lub z informacji poza ramkowej zdjęć, przybliżonych elementów orientacji zewnętrznej oraz punktach osnowy fotogrametrycznej. Zdjęcia lotnicze wykonane podczas misji fotolotniczych na obszarze Polski zgodnie z wymaganiami ustawy „Prawo geodezyjne kartograficzne” (Dz. U. z 2021 r.) przekazywane są do archiwizacji przez wykonawców (firmy fotolotnicze) do zasobu Głównego Urzędu Geodezji i Kartografii (dawniej CODGiK – Centralny Ośrodek Dokumentacji Geodezyjnej i Kartograficznej).

Ze względu na brak uregulowań prawnych do roku 2000 zdjęcia przekazywane były praktycznie bez żadnych dodatkowych informacji dotyczących projektu aerotriangulacji lub informacji odnoszących się do orientacji zewnętrznej, czy parametrów wykorzystywanych kamer (metryk kamery). Po 2000 roku weszła regulacja, która obligowała wykonawców do przekazania danych fotogrametrycznych do GUGiK wraz z wynikami ich orientacji. W konsekwencji, w przypadku wykorzystania archiwalnych zdjęć lotniczych wykonanych przed 2000 rokiem do różnego rodzaju opracowań fotogrametrycznych (i nie tylko), niezbędne jest ponowne przeprowadzenie pełnego procesu orientacji bloku fotogrametrycznego. Obejmuje to w szczególności wyznaczenie współrzędnych środków rzutów zdjęć oraz elementów orientacji kątowej w celu wpasowania bloku zdjęć w przyjęty przestrzenny układ odniesienia.

Największym wyzwaniem związanym z wpasowaniem zdjęć w układ odniesienia jest identyfikacja i dokładny pomiar punktów osnowy fotogrametrycznej. Trudności te wynikają przede wszystkim z różnej jakości danych archiwalnych, co wpływa na dokładność wyznaczania elementów terenowych (punktów naturalnych) na zdjęciach. Dodatkową trudność stanowi wykorzystanie współczesnych danych do identyfikacji tych punktów ze względu na zmiany urbanizacyjne i suburbanizacyjne na danym obszarze, zmiany w zagospodarowaniu przestrzennym oraz przekształcenia sposobu użytkowania terenu na przestrzeni lat.

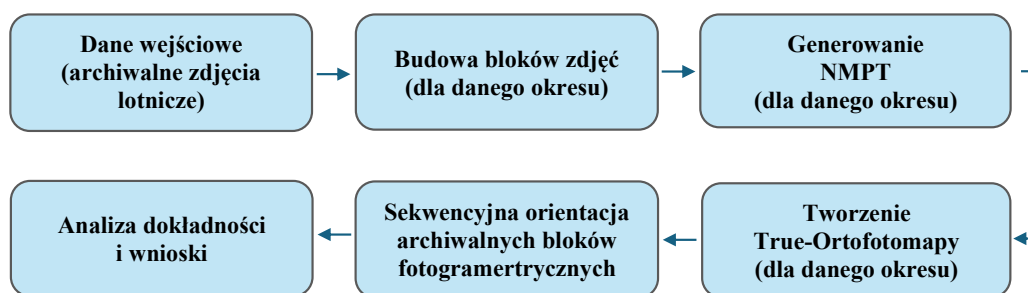
Szczególne znaczenie mają zmiany suburbanizacyjne związane z przekształcaniem terenów rolnych i leśnych w obszary miejskie, rozbudową infrastruktury czy budową dróg, powodując trudności w identyfikacji tych samych punktów na zdjęciach archiwalnych oraz na współczesnych ortofotomapach i w bazach danych obiektów topograficznych. W konsekwencji liczba zidentyfikowanych i pomierzonych punktów osnowy, niezbędnych do przeprowadzenia aerotriangulacji, okazuje się niewystarczająca, a ich niekorzystne rozmieszczenie w bloku zdjęć prowadzi do pogorszenia dokładności wyznaczania orientacji zdjęć (Kurczyński i in. 2020).

W niektórych opracowaniach dodatkowym problemem jest brak metryki kamery, a podanie przybliżonego położenia znaczków tłowych, ogniskowej kamery czy punktu głównego może doprowadzić do obniżenia dokładności opracowania. Problem ten został szczegółowo przedstawiony w podrozdziale 3.3.

Proponowane podejścia rozwiązania tego problemu na krajowym polu badawczym odnoszą się do metody manualnej wyznaczenia orientacji zdjęć wieloczasowych (Ewiak, Brodowska, 2012) lub metody automatycznej (Preuss, 2002), ale z ograniczonym udziałem działania algorytmów FBM (*Feature-Based Matching*), ABM (*Area-Based Matching*), LSM (*Least Squares Matching*). Problem ten jest również poruszany od wielu lat na polu międzynarodowym, gdzie organizacja EuroSDR (*European Spatial Data Research*) utworzyła specjalny *Benchmark Time* (Benchmark Time, 2025) zajmujący się tym zagadnieniem.

W rozprawie doktorskiej zaproponowano w pełni automatyczne podejście do sekwencyjnej orientacji bloków zdjęć lotniczych metodami CV (*Computer Vision*) na podstawie punktów wykrytych na true-ortofotomapach (wyznaczenie współrzędnych x, y), wraz z wyznaczoną ich wartością współrzędnej (z), wyinterpolowaną z NMPT (Numeryczny Model Pokrycia Terenu), w układzie lokalnym (xyz) lub terenowym (XYZ).

W celu realizacji automatycznego procesu sekwencyjnej orientacji bloków zdjęć wieloczasowych, stanowiącego kluczowy element proponowanej metody przetwarzania danych, zastosowano autorskie podejście oparte na zmodyfikowanej metodyce SfM (*Structure-from-Motion*) (rozdział 3.3.2). Na rysunku 5.1 przedstawiono ogólny schemat sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych (Rys. 5.1). Opis szczegółowy schematu znajduje się w rozdziale 5.3.



Rys. 5.1 Schemat ogólny metodyki opracowania sekwencyjnej orientacji bloków fotogrametrycznych (zmodyfikowane podejście SfM)

Pierwszy element modyfikacji metody SfM odnosi się do etapów detekcji i opisu punktów charakterystycznych. W przeciwieństwie do zamkniętych implementacji SfM (tzw. „czarnej skrzynki”), stosowanych w oprogramowaniu takim jak Pix4Dmapper czy Agisoft Metashape, zastosowano podejście umożliwiające wykorzystanie wybranych detektorów i deskryptorów — zarówno klasycznych (hand-crafted), jak i opartych na sieciach neuronowych (learned-based). Dzięki temu możliwe było dostosowanie metody do określonego scenariusza, w zależności od rodzaju danych wejściowych (zdjęcia cyfrowe lub analogowe), ich struktury (np. elementów pokrycia terenu) oraz skali obrazów. Drugim elementem modyfikacji był dobór poziomu piramidy obrazu, co umożliwiło skuteczniejsze dopasowanie true-ortofotomap oraz wyznaczenie punktów homologicznych dla zdjęć pozyskanych z różnych wysokości, charakteryzujących się odmienną skalą. Trzeci, a zarazem najistotniejszy element modyfikacji dotyczył sekwencyjnego, chronologicznego przenoszenia orientacji pomiędzy blokami zdjęć (od najnowszych do najstarszych), z wykorzystaniem uprzednio zorientowanego bloku referencyjnego w oparciu o transformację 3D. W podejściu standardowym SfM elementy orientacji zdjęć są zazwyczaj wyznaczane z wykorzystaniem procesu aerotriangulacji.

Metodyka opracowania sekwencyjnej orientacji bloków fotogrametrycznych składała się z trzech głównych etapów, tj.:

1. Budowa bloku fotogrametrycznego (poszczególne okresy czasowe niezależnie).

Pierwszym etapem wykorzystywanej metodyki orientacji archiwalnych zdjęć lotniczych (zdjęć wieloczasowych) jest przeprowadzenie orientacji wzajemnej tych zdjęć w celu stworzenia spójnych geometrycznie bloków fotogrametrycznych w lokalnych układach odniesienia. Szczegółowy opis tego etapu został przedstawiony w rozdziale 5.1;

2. Generowanie NMPT i true-ortofotomapy (na poszczególne okresy czasowe niezależnie).
W drugim etapie dysponując zorientowanym wzajemnie blokiem zdjęć wykonywany jest NMPT, na podstawie którego generowana jest true-ortofotomapa. Pierwsza część drugiego etapu realizowana jest w autorskim oprogramowaniu, natomiast druga w środowisku Agisoft Metashape. Opis tego etapu przedstawiono w podrozdziale 5.2.
3. Sekwencyjna orientacja wieloczasowych bloków fotogrametrycznych;
Kolejny etap metodyki związany jest z detekcją punktów dopasowania (przejściowych - patrz podrozdział 5.3) pomiędzy parami true-ortofotomap dla poszczególnych okresów czasowych, jest to tak zwane dopasowanie sekwencyjne. Wypracowana metoda uwzględnia wieloczasową orientację utworzonych bloków fotogrametrycznych z archiwalnych zdjęć lotniczych, rozumianą jako przeniesienie orientacji ze zdjęć współczesnych do najstarszych archiwalnych na podstawie odnalezienia wspólnych cech (szczegółów terenowych) pomiędzy zdjęciami. Szczegółowy opis orientacji bloków zdjęć znajduje się w podrozdziale 5.3.

Przedstawione w rozprawie doktorskiej rozwiązanie orientacji zdjęć lotniczych (rys. 5.1) zmodyfikowaną metodą SfM charakteryzuje się:

- niezmiennością wykrytych punktów przejściowych między parami true-ortofotomap od zniekształceń radiometrycznych i geometrycznych;
- stabilnością punktów niezależną od zmiany skali czy warunków oświetleniowych;
- unikalnością wykrytych punktów na obrazach, co zapewnia prawidłową i wiarygodną orientację bloku zdjęć, wysoką sztywność geometrii bloku z określonej epoki, a także odporność na błędy grube (rozdział 4).

W ramach pracy doktorskiej powstało autorskie oprogramowanie bazujące na językach programowania C++ i Python oraz interfejsie programowania aplikacji (API) Agisoft Metashape. Opis funkcjonalności oprogramowania przedstawiono w rozdziale 5.3.

5.1 Budowa bloków fotogrametrycznych

W celu przeprowadzenia orientacji wzajemnej zdjęć tworzących blok fotogrametryczny (budowy bloku) dokonano modyfikacji podejścia stosowanego w oprogramowaniu Agisoft Metashape. Modyfikacja ta polegała na zastąpieniu modułu odpowiedzialnego za detekcję,

opis i dopasowanie punktów wiążących zewnętrznym modulem zaimplementowanym w języku C++. Pozwoliło to na wykorzystanie powszechnie dostępnych algorytmów punktowych i powierzchniowych typu *Hand-Crafted*, odmiennych od tych stosowanych w komercyjnym oprogramowaniu.

Zastosowanie tych algorytmów skutkowało zwiększoną liczbą wykrytych punktów na parach zdjęć w stosunku do pomiaru klasycznego, co w konsekwencji poprawiło ich orientację wzajemną, umożliwiając utworzenie geometrycznie spójnego bloku. Po wyznaczeniu punktów w zewnętrznym module wyniki pomiarów algorytmów zostały zaimportowane do programu Agisoft Metashape i wykorzystane do przeprowadzenia orientacji wzajemnej zdjęć w poszczególnych blokach fotogrametrycznych.

Wybór języka C++ był podyktowany jego wysoką wydajnością obliczeniową, co ma kluczowe znaczenie dla efektywności procesu wyszukiwania, opisu i dopasowywania punktów wiążących. Zastosowanie detektorów *Hand-Crafted* w orientacji zdjęć dla poszczególnych okresów wynikało z faktu, że obrazy w obrębie danego bloku charakteryzowały się zbliżoną jakością radiometryczną i geometryczną oraz stałą wysokością lotu, co przekładało się na jednolitą skalę i wielkość piksela terenowego. Wiązało się to również z ograniczonymi zasobami obliczeniowymi, gdzie do wyznaczenia punktów wiążących przy użyciu algorytmów *Learned-Based* potrzebna była pamięć operacyjna komputera RAM (*Random Access Memory*) dużo powyżej 128 GB. W związku z tym, na tym etapie badań, nie było konieczności stosowania bardziej zaawansowanych metod, takich jak podejścia oparte na głębokich sieciach neuronowych do orientacji analizowanych bloków fotogrametrycznych.

Orientację archiwalnych zdjęć lotniczych (analogowych) w programie Agisoft Metashape można wykonać w dwóch podejściach. W pierwszym z nich, na podstawie znajomości metryki kamery, a w szczególności skalibrowanej wartości ogniskowej i współrzędnych znaczków tłowych, wykonywana jest orientacja wewnętrzna w klasycznym ujęciu. Drugie podejście (zastosowane w badaniach) uwzględnia etap samokalibracji (nieznana metryka kamery), który może być stosowany dzięki dużej nadliczbowości punktów wiążących. Należy jednak podkreślić, że ze względu na geometrię zdjęć lotniczych, ich pokrycia wzajemne i ukształtowanie terenu, samokalibrację postrzegamy tu nie jako proces, w wyniku którego wyznaczamy skalibrowaną wartość ogniskowej, współrzędne punktu głównego oraz parametry

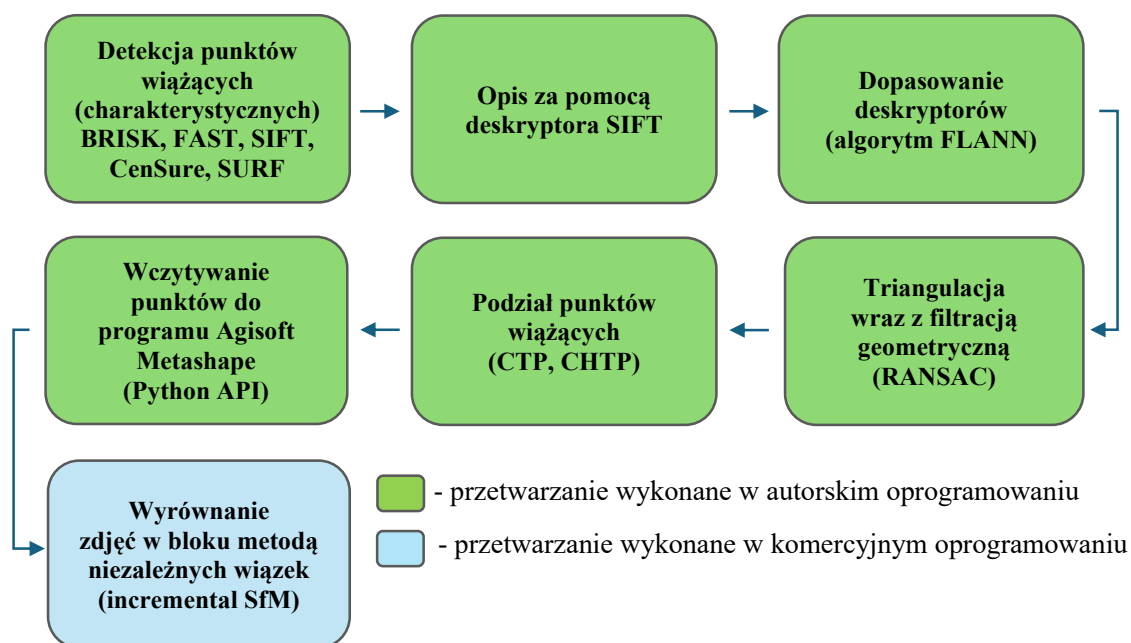
dystorsji radialnej i tangencjalnej. Zamiast tego, pełni ona funkcję etapu nakierowanego na opisanie, wyznaczenie i uwzględnienie afinizmu zdjęć spowodowanego groźnymi skurczami i deformacjami podłoża (patrz rozdział 3.1). W przypadku zdjęć cyfrowych ze względu na znane elementy orientacji wewnętrznej i metrykę kamery, etap samokalibracji nie jest wykonywany.

Jak wspomniano wcześniej, w ramach prowadzonych badań do budowy bloków fotogrametrycznych na podstawie archiwalnych analogowych zdjęć lotniczych zastosowano metodę SfM bazującą na algorytmach *Hand-Crafted*, wykorzystującą etap samokalibracji wraz z utworzeniem metryki wirtualnej dla danej kamery. Rozwiązanie to (por. podrozdział 3.3.2) zostało przyjęte ze względu na brak dostępnych danych kalibracyjnych dla wykorzystanych kamer. Zastosowane podejście umożliwiło uzyskanie spójności geometrycznej zdjęć, co w konsekwencji pozwoliło na eliminację zniekształceń afinicznych.

Proces orientacji zdjęć (rys.5.2) był procesem wieloetapowym w ramach, którego w pierwszym etapie wykonano detekcję punktów charakterystycznych w oparciu o detektory punktowe BRISK i FAST (patrz rozdział 4.2.1) oraz powierzchniowe SIFT, SURF CenSurE (patrz rozdział 4.2.2). W kolejnym etapie wykryte punkty zostały opisane za pomocą deskryptora SIFT (patrz rozdział 4.3), a przy wykorzystaniu algorytmu FLANN (*Fast Library for Approximate Nearest Neighbors*) wykonano dopasowanie deskryptorów (punktów opisanych za pomocą cech). Dla każdej par zdjęć za pomocą macierzy homografii (transformacja perspektywiczna między dwoma płaszczyznami zachowującymi współliniowość) przy wykorzystaniu algorytmu RANSAC (*Random Sample Consensus*) przeprowadzono filtrację geometryczną potencjalnych par punktów wiążących, mająca na celu usunięcie punktów, których błąd reprojektacji przekracza ustalony próg (*threshold*) wraz z jednoczesną triangulacją (patrz rozdział 5.1 - (4) Triangulacja wraz z filtracją geometryczną). Po przejściu wszystkich wymienionych etapów, uzyskane punkty zostały zakwalifikowane jako punkty wiążące.

Zanim punkty wiążące zostały zaimportowane do oprogramowania Agisoft Metashape zostały one podzielone na punkty wiążące wykorzystywane do orientacji (CTP – *Control Tie Point*) (nie są to klasyczne fotopunkty, ale punkty wiążące wykorzystane do orientacji zdjęć w układzie lokalnym) i kontrolne (CHTP – *Check Tie Point*) (nie są to klasyczne punkty kontrolne używane w pomiarach fotogrametrycznych, ale część punktów wiążących nie biorących udziału w orientacji zdjęć w układzie lokalnym, służących do niezależnej oceny dokładności

procesu orientacji). Ostatni etap przetwarzania danych obejmował import punktów do programu Agisoft Metashape z wykorzystaniem interfejsu programowania aplikacji (API) oraz języka Python, a następnie iteracyjne wyrównanie zdjęć w bloku do uzyskania wyrównania końcowego przy użyciu metody niezależnych wiązek (*bundle adjustment*). Dokładność orientacji zdjęć w bloku zweryfikowano na podstawie ww. punktów kontrolnych.



Rys. 5.2 Schemat budowy bloku zdjęć lotniczych metodą SfM

Szczegółowy schemat przetwarzania danych celem budowy bloku fotogrametrycznego można przedstawić w następujących krokach:

1. Detekcja punktów charakterystycznych

Pierwszy etap (rys. 5.2) w zmodyfikowanej metodzie Structure-from-Motion polegał na wyznaczeniu punktów charakterystycznych. W badaniach przeanalizowano dwa rodzaje algorytmów detekcji: detektory narożników/punktów (takie jak FAST, BRISK itp., patrz rozdział 4.2.1) oraz detektory blob (tj. SIFT, SURF, CenSurE, patrz rozdział 4.2.2). Etap detekcji cech był wykonywany na każdym zdjęciu z osobna w oparciu o algorytmy, które wykrywają cechy niezmiennie w stosunku do translacji obrazu, jego skalowania i obrotu, częściowo inwariantne w stosunku do zmian oświetlenia oraz odporne na lokalne zniekształcenia geometryczne. W celu implementacji ww. detektorów zdecydowano się na wykorzystanie biblioteki OpenCV;

2. Opis punktów za pomocą deskryptora SIFT

Kolejnym krokiem, po detekcji punktów charakterystycznych, był ich opis w oparciu o deskryptor SIFT (patrz rozdział 4.3). Wielkość i orientacja gradientów obrazu są badane wokół lokalizacji punktu kluczowego. W wyniku tego przetwarzania uzyskano 128 parametrowy wektor cech wykorzystywany na etapie dopasowywania odpowiadających sobie cech (potencjalnych punktów wiążących) na etapie *descriptor matching*. Dodatkowo zastosowanie deskryptora pozwala na usunięcie nieinwariantnych punktów w globalnym (a nie lokalnym) otoczeniu (patrz rozdział 4.3);

3. Dopasowanie deskryptorów przy wykorzystaniu algorytmu FLANN (dopasowanie deskryptorów)

Po przypisaniu punktom kluczowym cechy deskryptora, możliwe było określenie zależności między poszczególnymi punktami kluczowymi i dopasowanie deskryptorów na dwóch zdjęciach w pokryciu podłużnym i poprzecznym za pomocą algorytmu FLANN. Na początku zadaniem algorytmu jest zindeksowanie zbioru punktów w celu przeprowadzenia szybkiego wyszukiwania. Następuje przybliżone wyszukiwanie za pomocą takich technik jak np.: *k-d trees*, *randomized kd-trees*, *hierarchical k-means trees* (Jo J., i.in. 2017; Tavallali P.; i.in. 2021; Redmond SJ., Heneghan C., 2007) umożliwiające wybór najbardziej obiecujących obszarów zbioru, co pozwala na szybkie ograniczenie obszaru poszukiwawczego. Następnie na podstawie odległości euklidesowej (normy L2) wyznaczana jest miara podobieństwa dopasowania wektorów 128-parametrowych. Algorytm FLANN z całego zbioru na podstawie miar podobieństwa dopasowania wektorów wyznacza potencjalne punkty wiążące;

4. Triangulacja wraz z filtracją geometryczną

Wykrycie punktów wiążących jedynie na podstawie dopasowania deskryptorów nie jest zawsze wystarczające i może skutkować błędnym wyznaczeniem położenia punktu w przestrzeni 3D. Spowodowane to jest tym, że dopasowanie odbywa się jedynie w oparciu o wartości zmian gradientów stopni szarości i nie bierze pod uwagę położenia geometrycznego analizowanych punktów. W związku z tym konieczna jest weryfikacja geometryczna eliminująca wartości odstające i poprawiająca jakość punktów wiążących, a tym samym jakość ostatecznej orientacji zdjęć. W proponowanej metodyce generowania bloku fotogrametrycznego, wykonano weryfikację geometryczną punktów wiążących w oparciu

o podejście iteracyjne oparte na metodzie RANSAC przy wykorzystaniu funkcji modelu homografii i jednym z jej przypadków transformacji 2D DLT.

Metoda 2D DLT jest jednym z najczęściej stosowanych podejść do wyznaczania macierzy homografii między dwoma obrazami (np. lewym i prawym). Za pomocą tej metody możemy uzyskać macierz, która przekształca współrzędne punktów z jednej płaszczyzny obrazu na drugą (Ghosh, 2005).

Załóżmy, że mamy n -par punktów $(\mathbf{x}_i, \mathbf{y}_i)$ na obrazie 1 i $\mathbf{x}'_i, \mathbf{y}'_i$ na obrazie 2). Wyznaczamy macierz homografii $\mathbf{H}$, która przekształci punkty z obrazu 1 na odpowiadające im punkty na obrazie 2. Proces ten można opisać za pomocą następującego równania (5.1):

$$\begin{bmatrix} \mathbf{x}'_i \\ \mathbf{y}'_i \\ 1 \end{bmatrix} \sim \mathbf{H} \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \\ 1 \end{bmatrix} \quad 5.1$$

gdzie:

- $(\mathbf{x}_i, \mathbf{y}_i)$ to współrzędne punktu na obrazie 1 (lewy obraz),
- $(\mathbf{x}'_i, \mathbf{y}'_i)$ to współrzędne punktu na obrazie 2 (prawy obraz),
- $\mathbf{H}$ to macierz homografii (o wymiarach 3×3).

Macierz $\mathbf{H}$ ma postać macierzy 3×3 i zawiera 9 elementów, które chcemy wyznaczyć. Jej postać przedstawia równanie (5.2).

$$\begin{bmatrix} \mathbf{x}'_i \\ \mathbf{y}'_i \\ 1 \end{bmatrix} \sim \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \\ \mathbf{h}_4 & \mathbf{h}_5 & \mathbf{h}_6 \\ \mathbf{h}_7 & \mathbf{h}_8 & \mathbf{h}_9 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x}_i \\ \mathbf{y}_i \\ 1 \end{bmatrix} \quad 5.2$$

gdzie $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_9$ to elementy macierzy $\mathbf{H}$, które chcemy wyznaczyć i nie mają one bezpośredniej, niezależnej interpretacji geometrycznej. Homografia jest przekształceniem projektowym, w którym wszystkie elementy macierzy działają współzależnie. Wartości elementów $\mathbf{h}_7$ i $\mathbf{h}_8$ odpowiadają za nieliniowy (projektowy) charakter przekształcenia i decydują o występowaniu efektów perspektywicznych. Ich wartości decydują, czy transformacja jest afiniczna, czy projektowa. Parametr $\mathbf{h}_9$ pełni funkcję skali, dlatego macierz $\mathbf{H}$ jest określona z dokładnością do mnożnika skalarnego i może być dowolnie normalizowana (np. przyjmując $\mathbf{h}_9 = 1$).

Aby wyznaczyć elementy macierzy $\mathbf{H}$, należy wyjść od równania 5.1, a następnie rozwinąć je do postaci równania dla $\mathbf{x}'_i, \mathbf{y}'_i$ (5.3):

$$\begin{aligned}x_i'(h_7x_i + h_8y_i + h_9) &= h_1x_i + h_2y_i + h_3 \\y_i'(h_7x_i + h_8y_i + h_9) &= h_4x_i + h_5y_i + h_6\end{aligned}\tag{5.3}$$

Ponieważ każda para punktów dostarcza dwóch równań liniowych, a homografia ma 8 niezależnych parametrów (jest określona do skali), potrzebujemy co najmniej 4 par punktów (czyli $n \geq 4$) do wyznaczenia wszystkich 8 niezależnych parametrów (wektora h).

Te równania (5.3) zapisujemy w postaci macierzowej (5.4):

$$\begin{bmatrix} -x_i & -y_i & -1 & 0 & 0 & 0 & x_i x_i' & y_i x_i' & x_i' \\ 0 & 0 & 0 & -x_i & -y_i & -1 & x_i y_i' & y_i y_i' & y_i' \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{bmatrix} = 0\tag{5.4}$$

Dla n punktów wiążących tworzymy macierz A o wymiarach $2n \times 9$ (gdzie n to liczba punktów), a cały układ równań zapisujemy jako $A \cdot h = 0$. W celu rozwiązania układu równań i znalezienia macierzy H , należy zebrać informacje o punktach wiążących w jedną dużą macierz A (5.5). Każdy wiersz macierzy odpowiada jednej parze punktów, a kolumny zawierają współczynniki w układzie równań. Dzięki tej konstrukcji, macierz A (5.5) przyjmuje postać:

$$A = \begin{bmatrix} -x_1 & -y_1 & -1 & 0 & 0 & 0 & x_1 x_1' & y_1 x_1' & x_1' \\ 0 & 0 & 0 & -x_1 & -y_1 & -1 & x_1 y_1' & y_1 y_1' & y_1' \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -x_n & -y_n & -1 & 0 & 0 & 0 & x_n x_n' & y_n x_n' & x_n' \\ 0 & 0 & 0 & -x_n & -y_n & -1 & x_n y_n' & y_n y_n' & y_n' \end{bmatrix}\tag{5.5}$$

Aby rozwiązać układ równań i znaleźć macierz H , stosujemy metodę wartości osobliwych SVD (*Singular Value Decomposition*).

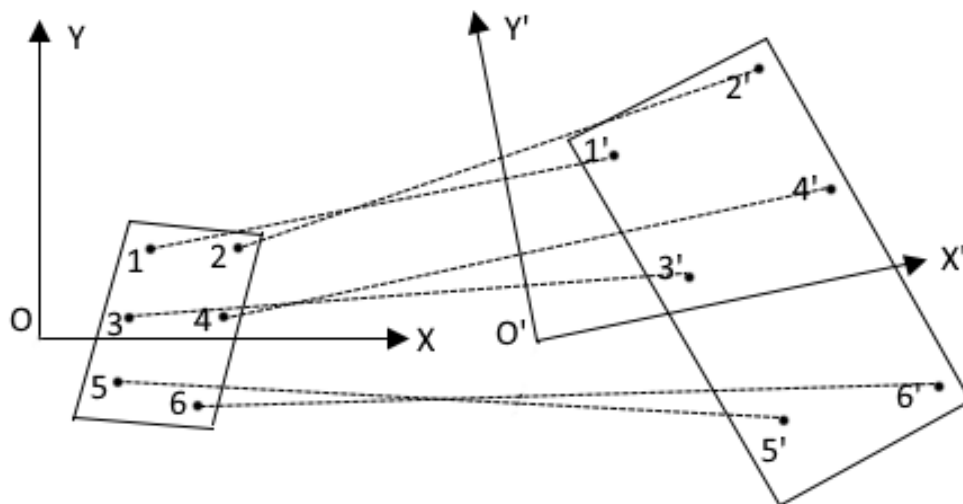
Po rozwiązaniu układu $A \cdot h = 0$ uzyskujemy wektor h , który zawiera elementy macierzy homografii H .

Po rozwiązaniu układu równań, uzyskujemy wartości współczynników $h_1, h_2, \dots, h_9$, które tworzą macierz homografii H w postaci (5.6):

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \quad 5.6$$

Ta macierz H pozwala na przekształcenie punktów z obrazu 1 do obrazu 2. Metoda 2D DLT (rys. 5.3) pozwala na wyznaczenie macierzy homografii, która opisuje transformację punktów z jednej płaszczyzny obrazu na drugą, przy pomocy układu równań liniowych. Za pomocą tej metody możemy określić powiązanie jednego zdjęcia z drugim w przestrzeni 2D.

Przed wyznaczeniem macierzy H stosowana jest normalizacja punktów wejściowych, co poprawia stabilność numeryczną i dokładność estymacji homografii (punkty są bardziej odporne na błędy wynikające z różnic w skali lub położeniu obiektów). Po wyznaczeniu macierzy H przeprowadza się denormalizację.



Rys. 5.3 Transformacja metodą 2D DLT

Metoda RANSAC przy wykorzystaniu funkcji homografii składa się z następujących etapów:

- Losowy wybór minimalnej liczby punktów (4 pary punktów) na dwóch zdjęciach;
- Przybliżone wyznaczenie homografii;
- Na podstawie punktów wykrytych (patrz podpunkt 1) wyznaczany jest model transformacji oparty na transformacji przez podobieństwo;
- Sprawdzenie dopasowań między punktami obu zdjęć. Wyznaczenie punktów, których odległość między punktem przewidywanym a faktycznym w drugim obrazie jest mniejsza od ustalonego progu;

- Iteracyjne powtórzenie kroku 1-3, aż do osiągnięcia określonej dokładności transformacji (homografii);
- Po wyznaczeniu par punktów spełniających kryterium ustalonego progu, obliczana jest ostateczna homografia. Wykonywane jest to na podstawie metody najmniejszych kwadratów i minimalizacji wartości błędów reprojektacji. Wartość progu filtracji punktów wiążących ustalono na 1 piksel, co odpowiada trzykrotności zakładanego błędu pomiarowego $\sigma \approx 0.3$ piksel, zgodnie z założeniem rozkładu normalnego. Ma to na celu usunięcie z pomiaru punktów odstających od wartości oczekiwanej.

Przefiltrowane punkty stanowią spójny geometrycznie zbiór wykorzystywany do orientacji zdjęć celem budowy bloku zdjęć fotogrametrycznych.

5. Podział punktów wiążących na CTP (*Control Tie Point*) i CHTP (*Check Tie Point*).

W klasycznym podejściu do wzajemnej orientacji grupy obrazów ocena dokładności dokonywana jest wyłącznie na podstawie wartości błędu reprojektacji (Tab. 5.1) na wszystkich punktach wykorzystywanych w wyrównaniu metodą niezależnych wiązek. Wadą tego podejścia jest brak pełnej kontroli wyrównania, ponieważ walidacja odbywa się na punktach użytych do budowy modelu matematycznego. Dlatego w przyjętej metodyce pomierzone automatycznie punkty wiążące zostały podzielone na dwie grupy: punkty wykorzystywane do orientacji zdjęć (CTP) oraz punkty kontrolne (CHTP), które nie brały udziału w procesie wyrównania i pełniły funkcję niezależnej kontroli. Punkty kontrolne zostały podzielone przy wykorzystaniu funkcji DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) (Sander J. i.in. 1998), celem równomiernego rozkładu w bloku zdjęć. Stanowiły one 10% ogólnej liczby wszystkich punktów. Warto dodać, że taki podział był możliwy ze względu na bardzo dużą nadliczbowość wykrytych punktów wiążących.

6. Import punktów do Agisoft Metashape.

W celu importu CTP oraz CHTP wykorzystano skrypt napisany w języku Python oraz API programu Agisoft Metashape.

7. Wyrównanie zdjęć w bloku metodą niezależnych wiązek (*Incremental SfM*) w Agisoft Metashape.

Proces orientacji bloku był procesem iteracyjnym (*Incremental SfM*). W pierwszym etapie algorytm wybiera parę zdjęć, dla której uzyskano największą liczbę dopasowanych punktów

wiązących CTP. Na ich podstawie konstruowany jest początkowy model trójwymiarowy. Dla wybranej pary zdjęć obliczana jest relacja geometryczna, a następnie przeprowadzana jest triangulacja punktów wiążących, co umożliwia utworzenie wstępnego modelu 3D. W kolejnym etapie proces przebiega iteracyjnie – do już zorientowanego modelu sukcesywnie dołączane są kolejne zdjęcia. Dla każdego nowego zdjęcia wyznaczana jest jego pozycja oraz orientacja w przestrzeni (wcięcie wstecz), po czym wykonywana jest triangulacja punktów wiążących. Po każdej iteracji przeprowadzane jest wyrównanie wiązek (*bundle adjustment*), którego celem jest optymalizacja zarówno parametrów orientacji wszystkich zdjęć, jak i przestrzennych współrzędnych punktów 3D. Proces ten powtarzany jest do momentu, aż wszystkie zdjęcia w bloku zostaną zorientowane. W ostatnim etapie przeprowadza się wyrównanie końcowe całego bloku zdjęć, zapewniające spójność geometryczną i minimalizację błędów rekonstrukcji, w wyniku którego otrzymano ich elementy orientacji wzajemnej w układzie lokalnym w skali zdjęcia.

Następnie przeprowadzono weryfikację w oparciu o liniowe wartości błędu reprojektacji (RMSE) i przeprowadzono filtrację przyjmując graniczny błąd reprojektacji nie większy niż 1 piksel. Badania empiryczne pokazują, że błąd graniczny reprojektacji jest na poziomie +/-1 piksela, pojawienie się większego należy traktować jako obserwację odstającą. Równocześnie przeprowadzono ocenę dokładności orientacji bloku fotogrametrycznego na punktach wiążących CTP oraz niebiorących udziału w wyrównaniu punktach kontrolnych CHTP.

5.2 Generowanie NMPT i true-ortofotomap

Generowanie NMPT i true-ortofotomap był etapem poprzedzającym etap sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych. Każdy blok zbudowany był w układzie lokalnym, gdzie jego początek związany jest parą zdjęć, dla której na etapie metody SfM wyznaczono największą liczbę punktów wiążących. Wykonano to w środowisku oprogramowania Agisoft Metashape. Po zbudowaniu bloków fotogrametrycznych dla każdej epoki program za pomocą podejścia MVS (*Multi-View Stereo*) wygenerował gęste chmury punktów stanowiące podstawę utworzenia NMPT (wykorzystany został on do wyznaczenia współrzędnej (z) punktów przejściowych) oraz true-ortofotomap.

Proces generowania NMPT i true-ortofotomap składał się z następujących etapów:

1. Generowanie gęstej chmury punktów przy wykorzystaniu algorytmu MVS z parametrem - $\frac{1}{2}$ rozdzielczości zdjęcia (powoduje to zwiększenie liczby generowanych punktów w porównaniu do ich pełnej rozdzielczości).
2. W chmurze punktów – wyszukiwanie 6 najbliższych sąsiadów i wyliczenie średniej odległości.
3. Tworzenie histogramów i wybór maksymalnej wartości częstotliwości (powtarzalność) – dobór GSD (*Ground Sampling Distance*).
4. Generowanie NMPT (w oparciu o funkcje biblioteki ArcPy). Generowanie w strukturze *grid* z GSD wyznaczonym w punkcie (3) – metoda interpolacji IDW (*Inverse Distance Weighted*).
5. Generowanie true-ortofotomap o wielkości GSD tak jak w punkcie 4 przy wykorzystaniu oprogramowania Agisoft Metashape.

Dzięki iteracyjnej metodzie orientacji zdjęć (*Incremental SfM*) zbudowane zostały bloki fotogrametryczne w układzie lokalnym dla poszczególnych okresów czasowych, które w rozumieniu matematycznym należy traktować jako tzw. „bryły sztywne”. W pojęciu fotogrametrycznym „bryła sztywna” rozumiana jest jako idealne połączenie zdjęć, które poddawane transformacji tj. skalowaniu, obrotowi i translacji nie ulegają deformacji (wzajemne położenie zdjęć nie ulega zmianie), przez co zastosowanie ww. transformacji przez podobieństwo przyczyni się jedynie do przeskalowania, obrotów i translacji całego bloku zdjęć. Dzięki zastosowaniu tej metody transformacji całego bloku możliwe jest wyznaczenie nowych elementów orientacji zewnętrznej zdjęć, a sam proces wyznaczenia nowych wartości polega na przeliczeniu macierzy orientacji zdjęcia przez wyznaczoną macierz transformacji bloku. Te kluczowe założenie stanowi podstawę zaproponowanej metodyki sekwencyjnej orientacji wieloczasowych bloków zdjęć opartych na bazie true-ortofotomap.

Wprowadzenie true-ortofotomap w metodyce

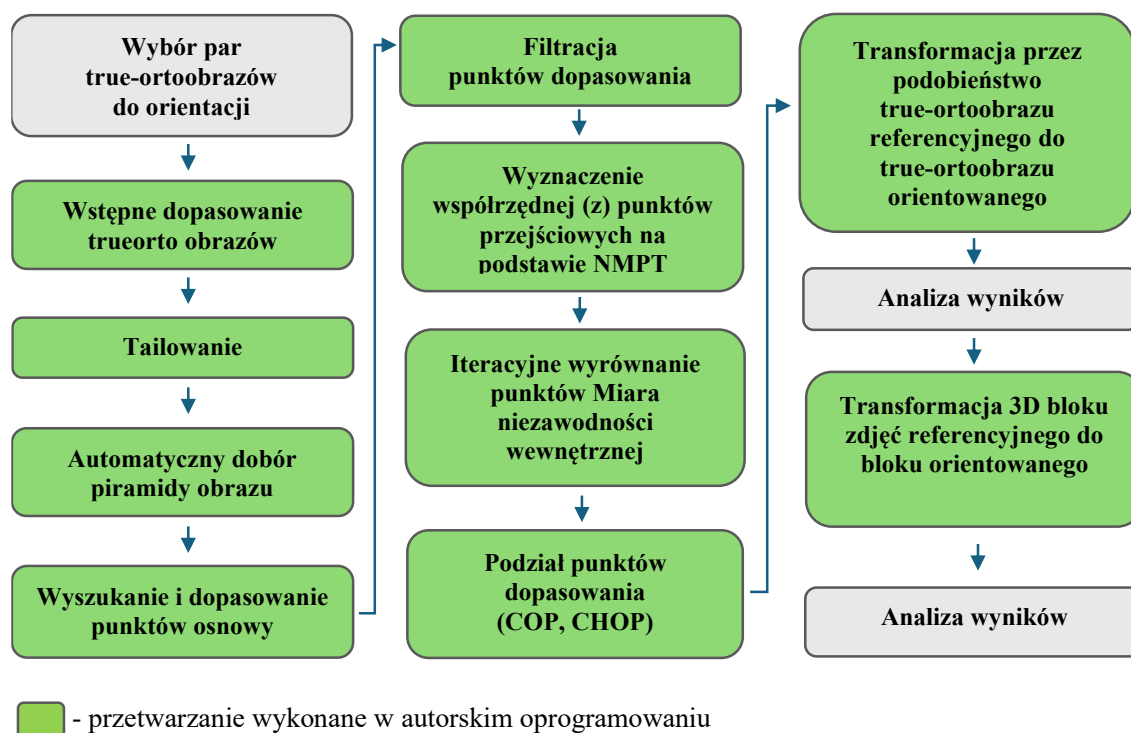
W proponowanej metodyce zrezygnowano z wykorzystania wyłącznie pojedynczych zdjęć lotniczych do wyznaczenia punktów, ze względu na ograniczenia wynikające z ich pokrycia podłużnego i poprzecznego oraz różnice w pokryciu pomiędzy poszczególnymi epokami. Zastosowano natomiast własne podejście globalne, wykorzystujące true-ortofotomapy wygenerowane na podstawie bloków fotogrametrycznych. Połączenie zdjęć w mozaikę true-

ortofotomapy znacząco rozszerza obszar możliwego dopasowania par obrazów. Dzięki temu liczba wykrywanych punktów wzrasta, co redukuje ryzyko ich nierównomiernego rozkładu. Ma to bezpośredni wpływ na zwiększenie dokładności orientacji bloków fotogrametrycznych. Wykorzystanie true-ortofotomap minimalizuje także wpływ słabej radiometrii pojedynczych zdjęć w bloku. W rezultacie, nawet w obszarach o niskiej jakości radiometrycznej, nie dochodzi do istotnych przerw w detekcji punktów, co pozytywnie wpływa na dokładność orientacji. Pomimo znaczących zmian w użytkowaniu terenu (w tym procesów urbanizacyjnych) pomiędzy różnymi epokami zdjęć lotniczych, zastosowanie bloków fotogrametrycznych oraz wynikających z nich true-ortofotomap umożliwia identyfikację wystarczającej liczby punktów przejściowych, zapewniając ich poprawną orientację. Efekt ten jest szczególnie widoczny w obszarach zurbanizowanych.

5.3 Metodyka sekwencyjnej orientacji bloków fotogrametrycznych

W oparciu o wstępnie zorientowane bloki zdjęć i wygenerowane true-ortofotomapy i NMPT możliwe było wykonanie sekwencyjnego dopasowania bloków zdjęć i przeniesienie orientacji na poszczególne zdjęcia.

Ze względu na występujące różnice w pokryciu pomiędzy utworzonymi blokami zdjęć, w celu wykonania sekwencyjnej orientacji pomiędzy nimi, utworzone na ich podstawie true-ortofotomapy zostały przycięte do wspólnego obszaru pokrycia. Umożliwiło to skuteczne dopasowanie true-ortofotomap oraz identyfikację punktów przejściowych (Abratkiewicz i in., 2019), niezbędnych do orientacji bloków. Zabieg ten nie wpływa na sekwencyjne przeniesienie orientacji pomiędzy blokami fotogrametrycznymi. Proces orientacji wieloczasowych danych był również wieloetapowym procesem iteracyjnym, którego schemat przedstawiony został na rysunku 5.4 i składał się z następujących etapów:



Rys. 5.4 Schemat metodyki sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych

1. Wybór par true-ortofotomap.

W celu automatycznej orientacji par true-ortofotomap niezbędne było określenie liczby wszystkich możliwych kombinacji ich par. W tym celu wykorzystano metodę permutacji bez powtórzeń według wzoru 5.7:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (5.7)$$

gdzie:

$k = 2$ – liczba true-ortofotomap tworzących parę,

n = liczba wszystkich true-ortofotomap

Pozwoliło to na automatyzację procesu przetwarzania danych;

2. Wstępne dopasowanie true-ortofotomap.

Proces wstępnego dopasowania true-ortofotomap polegał na wskazaniu w sposób manualny w przybliżeniu (np. zgrubnie na skrzyżowaniu dróg lub niezmienniczych obszarach) dwóch punktów na obrazie referencyjnym (względem, której będzie orientowany kolejny obraz) i orientowanym. W tym przypadku jako obraz referencyjny rozumiana jest true-ortofotomapa posiadająca georeferencję pozyskaną na podstawie przeprowadzonej

aerotriangulacji w oparciu o pomierzone fotopunkty lub na podstawie georeferencji wprost (elementów orientacji zewnętrznej pozyskanych w trakcie wykonywania zdjęć). Takie podejście pozwoliło na zastosowanie transformacji 2D Hallerta z jednoczesnym wyznaczeniem przybliżonej skali. W ramach przeprowadzonych prac testowano wykorzystanie detektora FAST i SIFT do automatycznego wstępnego dopasowania true-ortofotomap, lecz ostatecznie zdecydowano się na wykorzystanie manualnego podejścia;

3. Tailowanie.

Ze względu na rozdzielczość i wymiary generowanych true-ortofotomap nie było możliwe ich bezpośrednie wykorzystanie na etapie detekcji punktów charakterystycznych. Związane to było z działaniem detektorów *Hand-Crafted* (i ograniczeniami wykorzystywanego sprzętu) oraz architekturą wykorzystywanych sieci neuronowych (*Learned-Based*). Rozwiązaniem tego problemu mogłoby być zmniejszenie rozdzielczości true-ortofotomap (co jednocześnie wpływałoby na wykrycie mniejszej liczby punktów wiążących lub komplikacje z poprawnym ich dopasowaniem) lub ich tailowanie. W ramach rozprawy doktorskiej zdecydowano się na wykorzystanie tego drugiego podejścia, a true-ortofotomapy podzielono na 16 sekcji (2x8). W tym celu przygotowano skrypt w języku Python;

4. Automatyczny dobór piramidy obrazów.

W celu poprawnego wyszukiwania punktów wiążących przy wykorzystaniu metod *Hand-Crafted* i *Learned-Based* niezbędne jest wykorzystanie zbliżonych rozdzielczości zdjęć. Należy jednak zauważyć, że te rozwiązania dedykowane były dla zdjęć naziemnych, dla których zmiany skali nie są aż tak znaczące jak zdjęć lotniczych pozyskiwanych z różnych wysokości lotu i o różnej skali/rozdzielczości, np. zdjęć wykonywanych z GSD 10 cm i 25 cm. Z tego też powodu, w celu poprawnego zastosowania tych algorytmów dla zdjęć lotniczych zaproponowano metodę automatycznego doboru poziomu piramid na etapie wyszukiwania i dopasowywania true-ortofotomap wieloczasowych. Pierwszym założeniem działania algorytmu doboru poziomu piramidy była znajomość przybliżonej skali wyznaczonej na etapie wstępnego dopasowania true-ortofotomap. Dzięki temu możliwe było dopasowanie poziomu piramidy, dla każdej pary oddzielnie;

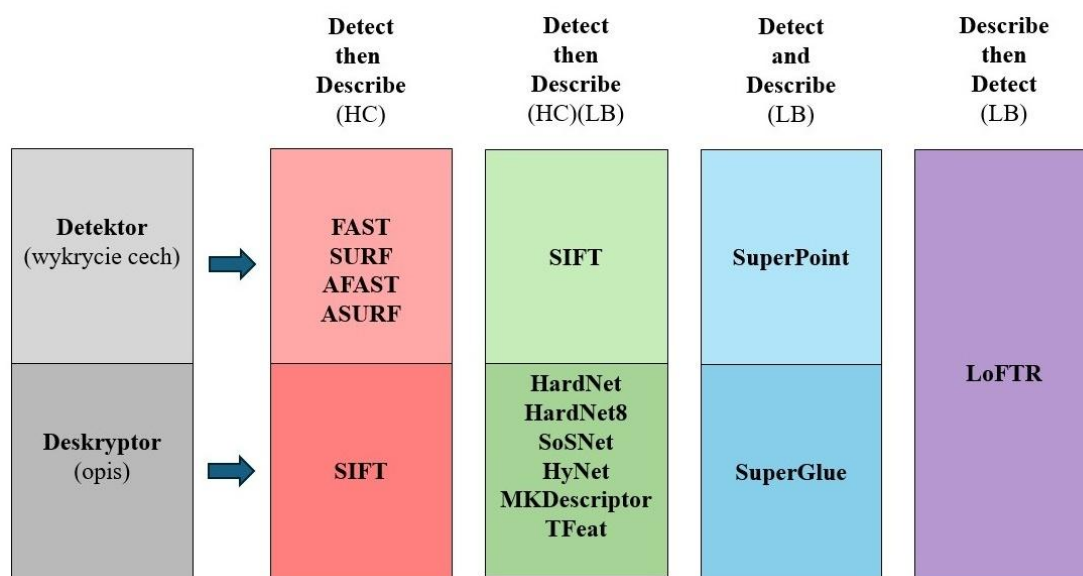
5. Wyszukiwanie i wyznaczenie punktów przejściowych (punktów charakterystycznych) pomiędzy sekwencjami (parami) true-ortofotomap.

Na etapie pomiaru punktów przejściowych dla sekwencji true-ortofotomap zastosowano trzy podejścia do ich wyznaczania (rys. 5.5): *Detect-then-Describe* (wykryj, a następnie opisz), *Detect-and-Describe* (wykryj i jednocześnie opisz) oraz *Describe-then-Detect* (opisz, a następnie wykryj) opisanych w rozdziale 4.

Pierwsze podejście, *Detect-then-Describe*, związane jest z metodami opartymi na algorytmach typu *Hand Crafted* (gdzie algorytmy detekcji to: FAST, SURF, AFAST, ASURF, a deskryptorem jest SIFT), a także na algorytmach związanych z sieciami neuronowymi typu *Learned-Based* (gdzie algorytmem detekcji jest SIFT, a algorytmami opisu są: HardNet, HardNet8, SOSNet, HyNet, MKDDescriptor, TFeat) (patrz rozdział 4).

Drugie podejście związane jest z metodą typu *Detect-and-Describe*, w której algorytm detekcji to SuperPoint, a algorytmem opisu SuperGlue (patrz rozdział 4.4).

Ostatnie podejście, *Describe-then-Detect*, polega na odwróceniu kolejności w stosunku do metody *Detect-and-Describe* – proces rozpoczyna się od opisu, a kończy na detekcji punktów. Reprezentatywnym algorytmem tej klasy jest LoFTR (patrz rozdział 4.4);



Rys. 5.5 Metody wyszukiwania i opisu punktów wiążących

6. Filtracja geometryczna.

W celu eliminacji błędów odstających wykonano filtrację geometryczną z ustalonym progiem 1 piksela. W proponowanej metodyce sekwencyjnej orientacji bloków fotogrametrycznych w oparciu o wygenerowane true-ortofotomapy, wykonano triangulację wraz z filtracją geometryczną według założeń metody opisanej w rozdziale 5.1 w punkcie 4;

7. Wyznaczenie współrzędnej (z) punktów przejściowych na podstawie modelu NMPT.

Przeniesienie orientacji z jednego bloku fotogrametrycznego na drugi wymaga znajomości parametrów transformacji w przestrzeni 3D. Parametry te zawarte są w macierzy translacyjno-rotacyjnej wyznaczanej za pomocą wykrytych punktów charakterystycznych na true-ortofotomapach (współrzędne x, y – 2D). Na podstawie współrzędnych tych punktów w oparciu o NMPT (generowany dla każdej z epok oddzielnie) wyznaczana jest współrzędna (z) (przy wykorzystaniu skryptu opartego na języku Python i bibliotece ArcPy). Dzięki temu wyznaczone zostały współrzędnych 3D punktów, które w kolejnym etapie wykorzystano do wyznaczenia parametrów transformacji między blokami zdjęć wieloczasowych;

8. Iteracyjne wyrównanie sekwencji par true-ortofotomap.

Przeniesienie orientacji (wzajemne dopasowanie) z true-ortofotomapy referencyjnej na true-ortofotomapę orientowaną w oparciu o pomierzone punkty przejściowe 3D składa się z następujących etapów:

- wykonanie wstępnej orientacji par true-ortofotomap w oparciu o pomierzone punkty przejściowe 3D;
- wyszukanie odpowiadających sobie punktów na parach true-ortofotomap z wykorzystaniem metody k -NN (*k-Nearest Neighbors*) (Fix i Hodges, 1951) o promieniu najbliższego sąsiedztwa 3 piksele, dla metod opartych na deskryptorach lokalnych (*Detect-then-Describe* oraz *Detect-and-Describe*), a w przypadku podejścia *Describe-then-Detect* dopasowania wyznaczane bezpośrednio przez sieć neuronową;
- iteracyjne wyrównanie obserwacji (punktów przejściowych) wraz z analizą błędów RMS współrzędnych x, y, z (3D) z uwzględnieniem miary niezawodności wewnętrznej;
- ostateczne wyrównanie obserwacji wraz z minimalizacją błędów orientacji par true-ortofotomap (środowisko Python).

Punkty przejściowe zostały podzielone na dwie grupy: punkty wykorzystywane do orientacji zdjęć (COP – *Control Orientation Points*) oraz punkty weryfikujące orientację (CHOP- *Check Orientation Point*), które nie brały udziału w wyrównaniu areotriangulacji metodą niezależnych wiązek i pełniły funkcję niezależnej kontroli. Punkty kontrolne zostały wydzielone przy wykorzystaniu funkcji DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*)

poprzez ich losowy wybór z grupy wszystkich punktów. Stanowiły one 10% ogólnej liczby wszystkich punktów;

9. Transformacja 3D bloku referencyjnego do bloku orientowanego.

Transformacja 3D przez podobieństwo związana jest z wyznaczeniem macierzy translacyjno-rotacyjnej pozwalającej na przeniesienie orientacji z bloku referencyjnego (true-ortofotomapy referencyjnej) do bloku orientowanego (true-ortofotomapy orientowanej) (środowisko Python). Transformacja może być wykonana w układzie lokalnym bloku (pikselowy – x,y,z) jak i zewnętrznym (terenowy – X,Y,Z) (w zależności jak zdefiniowany jest układ referencyjny). Uwzględniając właściwości „bryły sztywnej”, dzięki zastosowaniu parametrów transformacji możliwe jest przeniesienie orientacji z bloku referencyjnego na poszczególne zdjęcia i wyznaczenie elementów orientacji zewnętrznej dla każdego zdjęcia oddzielnie.

Transformacja 3D: Model siedmioparametrowy

Transformacja trójwymiarowa w postaci siedmioparametrowej (Ghosh, 2005; Strang, 2023) składa się z trzech podstawowych etapów:

1. Translacja (przesunięcie) – trzy parametry t_x, t_y, t_z ;
2. Rotacja – trzy kąty obrotu α, β, γ wokół osi x, y, z ;
3. Jednolite skalowanie – pojedynczy współczynnik s , stosowany jednakowo we wszystkich osiach.

Całość operacji wyrażana jest za pomocą macierzy jednorodnej T o wymiarach 4×4 . Współrzędne punktu (x, y, z) po zastosowaniu transformacji T opisuje równanie (5.8):

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = T \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \quad (5.8)$$

gdzie (x', y', z') oznaczają nowe współrzędne punktu.

Macierz rotacji

Rotacja w przestrzeni 3D przedstawiana jest jako produkt trzech macierzy obrotu:

$$R = R_z(\gamma) \cdot R_y(\beta) \cdot R_x(\alpha), \quad 5.9$$

gdzie:

$$R_z(\gamma) = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad R_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix},$$

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}$$

Skalowanie

Jednolite skalowanie zmienia rozmiar obiektu proporcjonalnie we wszystkich trzech osiach, przy zachowaniu kształtu (5.10):

$$R' = s \cdot R, \quad 5.10$$

gdzie s mnoży każdy element macierzy R .

Następnie tworzona jest macierz transformacji T zawierająca macierz obrotu R , skalowanie s i translację t (5.11):

$$T = \begin{bmatrix} R' & \vec{t} \\ \vec{0} & 1 \end{bmatrix} = \begin{bmatrix} s \cdot r_{11} & s \cdot r_{12} & s \cdot r_{13} & t_x \\ s \cdot r_{21} & s \cdot r_{22} & s \cdot r_{23} & t_y \\ s \cdot r_{31} & s \cdot r_{32} & s \cdot r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad 5.11$$

gdzie: r_{ij} to elementy macierzy R , a $\vec{t}$ – translacja (wektor przesunięcia $[t_x, t_y, t_z]$).

Dla punktu $\vec{p} = [x, y, z, 1]^T$ w układzie jednorodnym po transformacji otrzymujemy nowe współrzędne punktu po obrocie, skalowaniu i przesunięciu (5.12):

$$\vec{p}' = T \cdot \vec{p} = \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} s(r_{11}x + r_{12}y + r_{13}z) + t_x \\ s(r_{21}x + r_{22}y + r_{23}z) + t_y \\ s(r_{31}x + r_{32}y + r_{33}z) + t_z \\ 1 \end{bmatrix} \quad 5.12$$

10. Analiza uzyskanych wyników

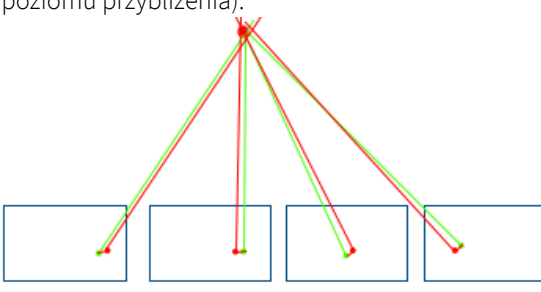
Szczegółowy opis miar jakości wykorzystywanych do oceny dokładności orientacji zdjęć wieloczasowych został zawarty w tabeli 5.1.

5.4 Miary jakości wykorzystywane w ocenie dokładności orientacji zdjęć lotniczych

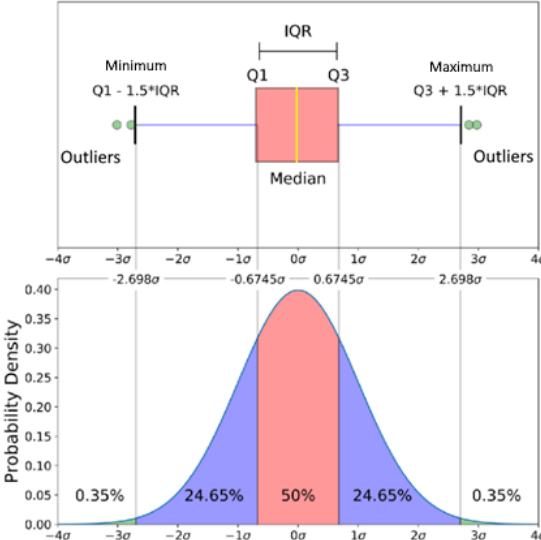
W celu wielowariantowej oceny dokładności orientacji bloków zdjęć zaproponowano szereg wskaźników umożliwiających jednoznaczną ocenę poprawności zaproponowanej metodyki.

W tabeli 5.1 przedstawiono zbiorczy ich wykaz ze szczegółowym opisem znaczenia ważności poszczególnych wskaźników oraz wspomniano, na którym etapie przetwarzania danych były one wykorzystane. W tabeli 5.1 przedstawiono również wielkości związane z oceną niezawodności orientacji zdjęć lotniczych (bloków zdjęć) i par true-ortofotomap. Zastosowanie takiego podejścia pozwoliło na szczegółową ocenę poszczególnych etapów orientacji danych, a w szczególności doboru odpowiedniego algorytmu *Hand-Crafted* lub *Learned-Based*.

Tab. 5.1 Kryteria oceny zaproponowanej metodyki orientacji danych wieloczasowych

Miary jakości	Opis	Wykorzystanie
Czas wyszukiwania i dopasowania punktów wiążących	Im krótszy czas wyszukiwania lub dopasowania punktów przez algorytm tym bardziej efektywne działanie algorytmu i krótszy czas opracowania. Czas związany jest również z liczbą wykrytych punktów.	Zdjęcia w bloku i true-ortofotomapy wieloczasowe
Procentowy udział poprawnie wykrytych punktów wiążących	Duża liczba wykrytych punktów prowadzi do stabilności powiązań między parami zdjęć i „sztywności bloku” oraz jego odporności na błędy na obserwacje odstające. Natomiast dla true-ortofotomap wieloczasowych - do ich poprawności dopasowania. Świadczy również o efektywności algorytmu detekcji i dopasowania punktów wiążących.	Zdjęcia w bloku i true-ortofotomapy wieloczasowe
Rozkład punktów wiążących	Optymalne położenie punktów wiążących umożliwia poprawną orientację danych, natomiast słaby rozkład spowoduje ich niestabilność. Rozkład należy oceniać zgodnie z teorią niezawodności wewnętrznej sieci geodezyjnej.	Zdjęcia w bloku i true-ortofotomapy wieloczasowe
Liczba dopasowanych par true-ortofotomap	Weryfikacja możliwości dopasowania true-ortofotomap wieloczasowych pomiędzy poszczególnymi okresami i kombinacja wszystkich dostępnych par.	True-ortofotomapy wieloczasowe
Błąd reprojektacji	Odtworzenie trójwymiarowych współrzędnych punktu następuje na podstawie jego położenia na wszystkich obrazach, na których się znajduje. Błąd położenia tego punktu (reprojekcji) zależy od pozycji i orientacji kamery oraz jakości punktu na obrazach (pozycji i poziomu przybliżenia).  ● Obliczony punkt 3D ● Punkt wyznaczony na zdjęciu ● Punkt rzutowany na zdjęcie — Błąd reprojektacji	Zdjęcia w bloku

	Odległość pomiędzy zaznaczonym i ponownie odwzorowanym punktem na jednym obrazie to błąd ponownej projekcji (Pix4D Support, 2025).	
Mediana paralaksy (ocena niezawodności)	W zbiorze wartości mediana, to wartość, która dzieli uporządkowany zbiór danych na dwie równe części. Mediana charakteryzuje się trzema właściwościami: prostota wyliczenia tej miary, odporność na wpływ obserwacji z wartościami skrajnymi (odstającymi) oraz uwzględnia cały rozkład, a nie tylko fragment. $M_e = \frac{x_{n+1}}{2},$ gdy n jest nie parzyste $M_e = \frac{1}{2} \left(\frac{x_n}{2} + \frac{x_{n+1}}{2} \right),$ gdy n jest parzyste.	
RMSE błędu paralaksy (Root Mean Square Error) (ocena niezawodności)	Odchylenie średniokwadratowe (RMSD) lub błąd średniokwadratowy (RMSE) jest jedną z dwóch blisko powiązanych i często używanych miar różnic między wartościami rzeczywistymi lub przewidywanymi z jednej strony a wartościami obserwowanymi lub estymatorem z drugiej strony. Błąd średni reprojektacji liczony według wzoru: $RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}$ gdzie: N – liczba obserwacji, y_i – położenie punktów na zdjęciu referencyjnym, $\bar{y}$ – położenie punktów na zdjęciu dopasowywanym.	Zdjęcia w bloku i true-ortofotomapy wieloczasowe
SIGMA – odchylenie standardowe	Jest to wskaźnik jakości pomiaru pozwalający na ocenę dokładności i wiarygodności uzyskanych wyników. SIGMA – odchylenie standardowe liczone według wzoru $\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$ gdzie: $\{x_1, x_2, \dots, x_N\}$ są zaobserwowanymi wartościami położenia punktów, N oznacza wielkość próby, a $\bar{x}$ jest wartością oczekiwaną.	
SMAD błędu paralaksy (ocena niezawodności)	SMAD (Sigma Median Absolute Deviation) $SMAD = 1,4826 * \text{mediana}(x_i - \delta)$ gdzie: x_i jest wartością w zbiorze danych, a δ jest medianą zbioru danych	
Min - Max błędu paralaksy (ocena niezawodności)	Max, min (a1, a2,...,an). Wartości skrajne w odchyłkach odnoszące się do wartości maksymalnej i minimalnej.	
Wykres pudełkowy (box plot)	Wykresy pudełkowe są rysowane jako prostokąty z pionową linią pośrodku i poziomymi liniami dołączonymi po obu stronach (tzw. „wąsami”). Ramka służy do przedstawienia rozstępu międzykwartylowego (IQR - <i>Inter Quartile Range</i>) — czyli 50 procent punktów	Zdjęcia w bloku i true-ortofotomapy wieloczasowe

	danych leżących powyżej pierwszego kwartyla i poniżej trzeciego kwartyla — w danym zbiorze danych. Wąsy służą do przedstawienia zmienności minimalnych, maksymalnych i wszelkich punktów danych odstających w porównaniu z IQR (im dłuższy wąs, tym szersza zmienność punktów danych dołączonych do IQR). Rozkład wartości błędów został przedstawiony w formie wykresu pudełkowego, gdzie: - mediana (Q2 percentyl 50): środkowa wartość zbioru danych zaznaczona linią. - pierwszy kwartyl (Q1 - percentyl 25): środkowa liczba pomiędzy najmniejszą liczbą, a medianą zbioru danych. - trzeci kwartyl (Q3 - percentyl 75): środkowa wartość między medianą a najwyższą wartością zbioru danych. - rozstęp ćwiartkowy (IQR): od percentyla 25 do 75. - wąsy wykresu: (zaznaczone na fioletowo), maksymalna i minimalna wartość danych i jej zmienność w porównaniu do IQR. - wartości odstające (błędy grube) (zaznaczone w zielonych kółkach) - maximum: $Q3 + 1.5 \cdot IQR$ - minimum: $Q1 - 1.5 \cdot IQR$  Porównanie wykresu pudełkowego o rozkładzie bliskim normalnemu i funkcji gęstości prawdopodobieństwa dla rozkładu normalnego (Galarnyk, 2025).	
Średnia wartość błędu reprojektacji	Dla bloku wartości błędu reprojektacji. Dla true-ortofotomap wartość błędu w jednostkach terenowych. Ocena niezawodności pozwala ocenić poprawność rozkładu geometrycznego punktów wiążących pozyskanych w sposób w pełni automatyczny. Spełniając minimalne wymagania dotyczące wartości na podstawie wartości błędu reprojektacji można ocenić, czy powstała sieć punktów jest odporna na występowanie błędów w obserwacjach (obserwacje odstające).	Zdjęcia w bloku i true-ortofotomapy wieloczasowe

	Wartość średnią oblicza się jako sumę wartości błędów reprojekcji, podzielonej przez liczbę obserwacji. $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$ gdzie: $x_1, x_2, \dots$ - poszczególne wartości błędów reprojekcji, n - liczba obserwacji. Średnia jest najczęściej stosowaną miarą typowości. Średnia ze względu na swoją konstrukcję (bazuje na wszystkich wartościach zmiennej) jest wrażliwa na wartości skrajne.	
--	--	--

6. Charakterystyka danych testowych wykorzystywanych w pracach eksperymentalnych

Przedstawiona w rozdziale 2 teza rozprawy doktorskiej „Metody widzenia maszynowego pozwalają z wysoką niezawodnością na automatyczne odtworzenie orientacji archiwalnych zdjęć lotniczych z dokładnością na poziomie rozdzielczości zdjęć (piksela terenowego)”. W celu wykonania orientacji takich zdjęć niezbędne jest odtworzenie punktów osnowy fotogrametrycznej poprzez odnalezienie i pomiar tych samych szczegółów terenowych na archiwalnych zdjęciach i aktualnych danych kartograficznych (ortofotomapy, bazy danych topograficznych, pomiar terenowy punktów, itp.). Proces pomiaru (odtworzenia) punktów osnowy ze względu na jego manualny charakter jest bardzo czasochłonny, co przekłada się na wysoki koszt opracowania. Trudno jest również uzyskać wymagane dokładności orientacji zdjęć (patrz podrozdział 3.2), ze względu na ich słabą radiometrię i geometrię. Należy do tego dodać, że występują zdjęcia, dla których ze względu na duże zmiany urbanizacyjne na obszarze opracowania, nie ma możliwości identyfikacji punktów osnowy. Tak scharakteryzowane zbiory zdjęć lotniczych znajdują się w centralnym zasobie geodezyjnym i kartograficznym GUGiK (patrz rozdział 5). Jeżeli chodzi o Warszawę (obszar analiz badawczych) do roku 2000 zdjęć (wykonanych w ramach nalotów fotogrametrycznych) o tej specyfice zarchiwizowano w około 40 zbiorach. Zbiory te są liczne i należy je utożsamiać z dużymi blokami fotogrametrycznymi, dla których nadanie referencji stanowi spore wyzwanie technologiczne i metodyczne. Szeroki zakres dostępnych archiwalnych zdjęć lotniczych pod względem ich skali i jakości stanowił podstawę do wyboru danych testowych, pozwalających w części eksperymentalnej zrealizować postawione zadania badawcze i udowodnić tezę rozprawy. Dobór danych testowych, oprócz daty ich pozyskania, związany był następującymi parametrami:

- skalą zdjęć,
- liczbą i pokryciem zdjęć w szeregu i pomiędzy szeregami,
- rodzajem zdjęć (cyfrowe, analogowe),
- jakością geometrycznej i radiometrycznej zdjęć,
- zróżnicowanym pokryciem terenu,
- zmianami urbanizacyjnymi zachodzącymi w czasie.

W związku z tym jako dane testowe wybrano zdjęcia lotnicze z lat: 1976, 1986, 1994, 2014 oraz 2022 (rys. 6.2 - 6.7). Różnią się one m.in. następującymi parametrami (tabela 6.1): datą wykonania, skalą zdjęć, liczbą szeregów i liczebnością zdjęć w bloku, wysokością lotu, ogniskową obiektywu, wielkością piksela obrazu, a także rodzajem i formatem zdjęć.

Problematyka opracowania zdjęć archiwalnych pozostaje aktualna, czego dowodem jest utworzony w ramach europejskiej organizacji EuroSDR (*European Spatial Data Research*) zestaw testowy Benchmark Time, zawierający m.in. zbiór zdjęć lotniczych z 1986 roku obejmujących obszar Warszawy, które zostały również wykorzystane w badaniach przedstawionych w niniejszej rozprawie doktorskiej (Benchmark Time, 2025).

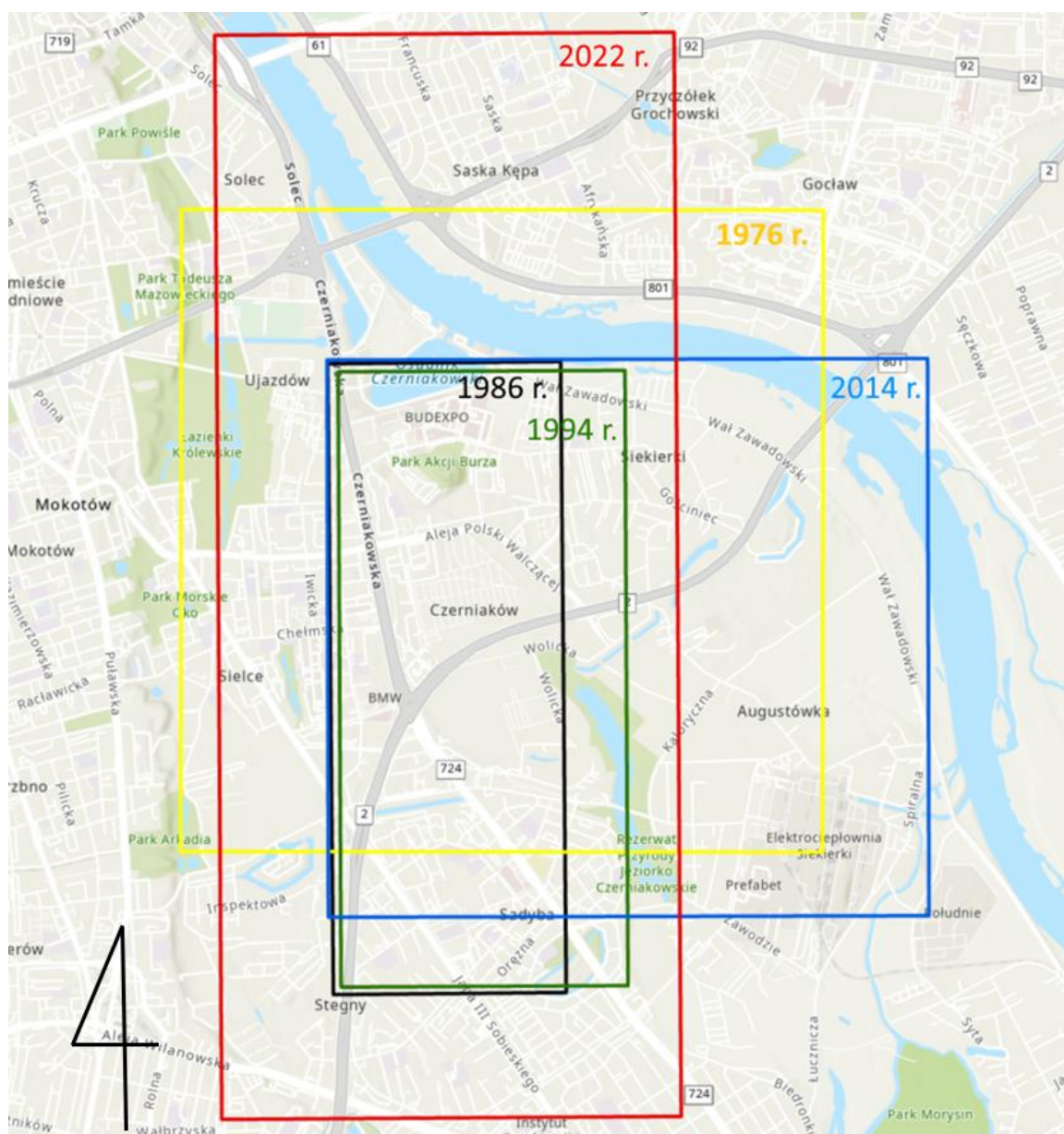
Tabela 6.1 Zestawienie podstawowych parametrów danych testowych (zdjęć lotniczych z roku 2022, 2014, 1994, 1986, 1976)

	2022 r.	2014 r.	1994 r.	1986 r.	1976 r.
Nazwa kamery	UltraCam Eagle M3	UltraCam Xp	NAGII 7154	NAGII 7154	UAgII 3074
Data wykonania	05.06.2022	11.06.2014	22.04.1994	07.05.1986	01.10.1976
Skala/GSD	5 cm	24 cm	1:5000 (7 cm)	1:3000 (5 cm)	1:18 500 (25 cm)
Liczba szeregów / Liczba wszystkich zdjęć	2/6	1/4	2/29	3/44	1/3
Wysokość lotu	1250 m	4020 m	1040 m	660 m	2770 m
Metryka kamery	tak	tak	tak	tak	tak
Elementy orientacji	Tak	Tak	brak	brak	brak
c_k	100.500 mm	100.500 mm	213.75 mm	213.75	153.45 mm
Wielkość piksela	4 μ m (matrycy)	6 μ m (matrycy)	14 μ m (skanowania)	14 μ m (skanowania)	14 μ m (skanowania)
Rodzaj zdjęcia	cyfrowe	cyfrowe	analogowe	analogowe	analogowe
Format zdjęcia	17004x26460 Pikseli	11310x17310 pikseli	18471x17165 Pikseli (24x24 cm)	17140x18569 Pikseli (24x24 cm)	18569x17140 Pikseli (24x24 cm)
Jakość radio metryczna	Bardzo dobra	Bardzo dobra	Dobra	Bardzo dobra	Słaba

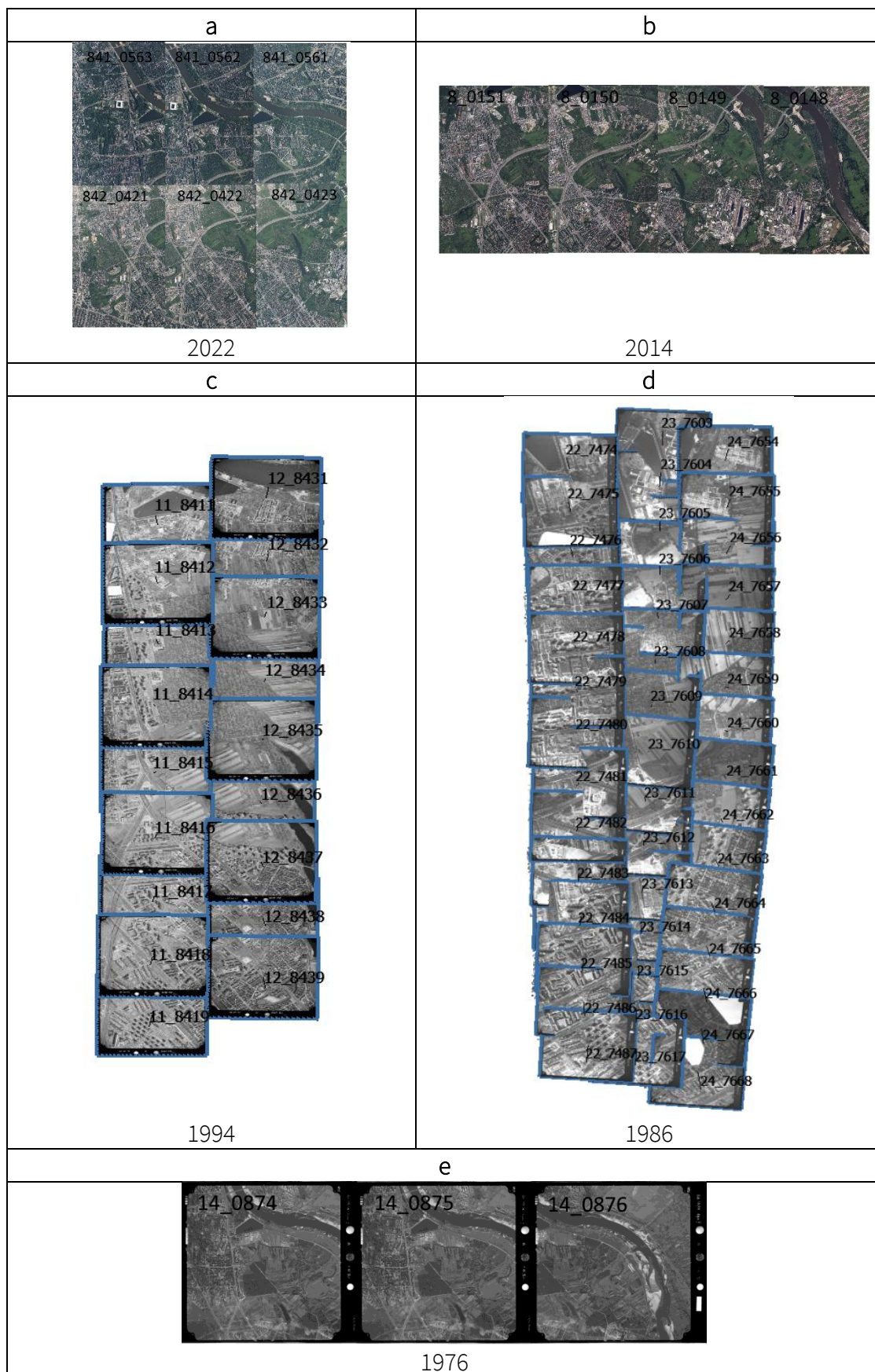
Zdjęcia lotnicze dla poszczególnych lat 1976, 1986, 1994, 2014 i 2022 obejmowały swym zasięgiem (rys. 6.1) fragment obszaru miasta Warszawy (dzielnica Śródmieście – okolice

Osadnika Czerniakowskiego). Zdjęcia te stanowiły wybrane fragmenty bloków fotogrametrycznych pozyskanych w trakcie nalotów we wskazanych latach. W większości obszar, na którym przeprowadzono badania, ma charakter miejskiej zabudowy zwartej, częściowo pokryty ogródkami działkowymi oraz użytkami rolnymi. Dotyczy to lat 1986 i 1994. Dla zdjęć z roku 2014 i 2022 obszar gruntów rolnych przechodzi w formę nieużytku, pokrytego roślinnością niską (krzaki i zadrzewienia).

Do orientacji zdjęć analogowych z lat 1976, 1986 i 1994 na podstawie dostępnej dokumentacji wykorzystano ogniskową kamery (odczytaną z informacji pozaramkowej umieszczonej na marginesie zdjęcia) oraz położenie nominalne znaczków tłowych. Oprogramowanie Agisoft Metashape, służące do przetwarzania zdjęć, posiada moduł umożliwiający pomiar tych znaczków. W przypadku zdjęć cyfrowych z lat 2014 i 2022, dzięki dostępności metryki kamery, możliwe było uwzględnienie pełnego zestawu parametrów orientacji wewnętrznej. Ponadto zdjęcia cyfrowe z roku 2022 posiadały orientację zewnętrzną i stanowiły referencję do pozostałych zdjęć lotniczych.



Rys. 6.1 Zasięgi pokryć archiwalnych zdjęć lotniczych fragmentu miasta Warszawy (okolice Czerniakowa) dla poszczególnych lat wykorzystanych w badaniach. Zasięg dla roku 1976 – kolor żółty, zasięg dla roku 1986 – kolor czarny, zasięg dla roku 1994 – kolor zielony, zasięg dla roku 2014 – kolor niebieski, zasięg dla roku 2022 – kolor czerwony



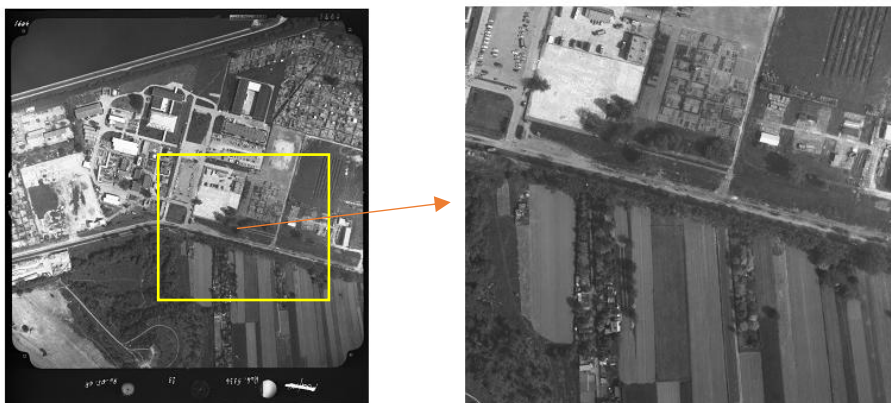
Rys. 6.2 Dane testowe (a - zdjęcia z roku 2022, b - zdjęcia z roku 2014, c - zdjęcia z roku 1994, d - zdjęcia z roku 1986, e - zdjęcia z roku 1976)

14_0875



Rys. 6.3 Przykład zdjęcia lotniczego wybranego ze zbioru pozyskanego w roku 1976 dla miasta Warszawy oraz wybrane fragmenty jego powiększenia (prawa strona rysunku)

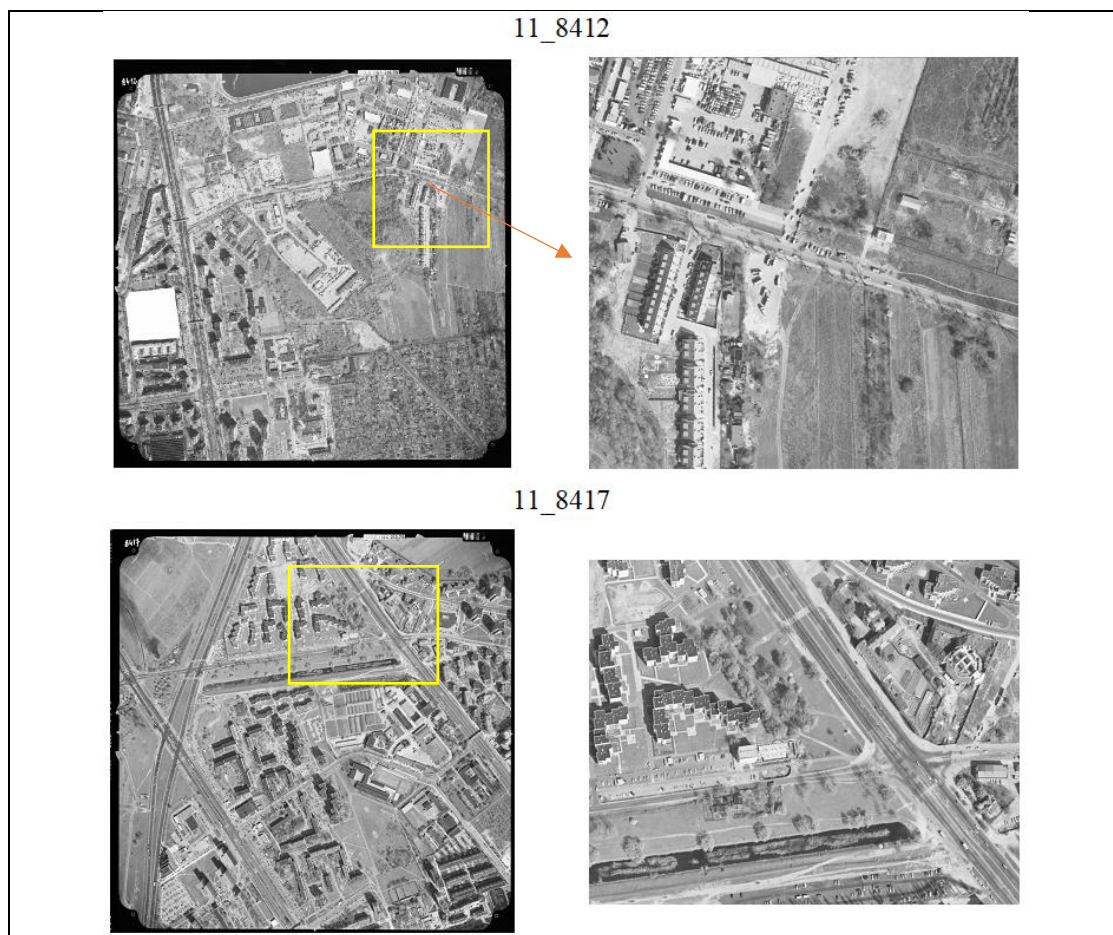
23_7604



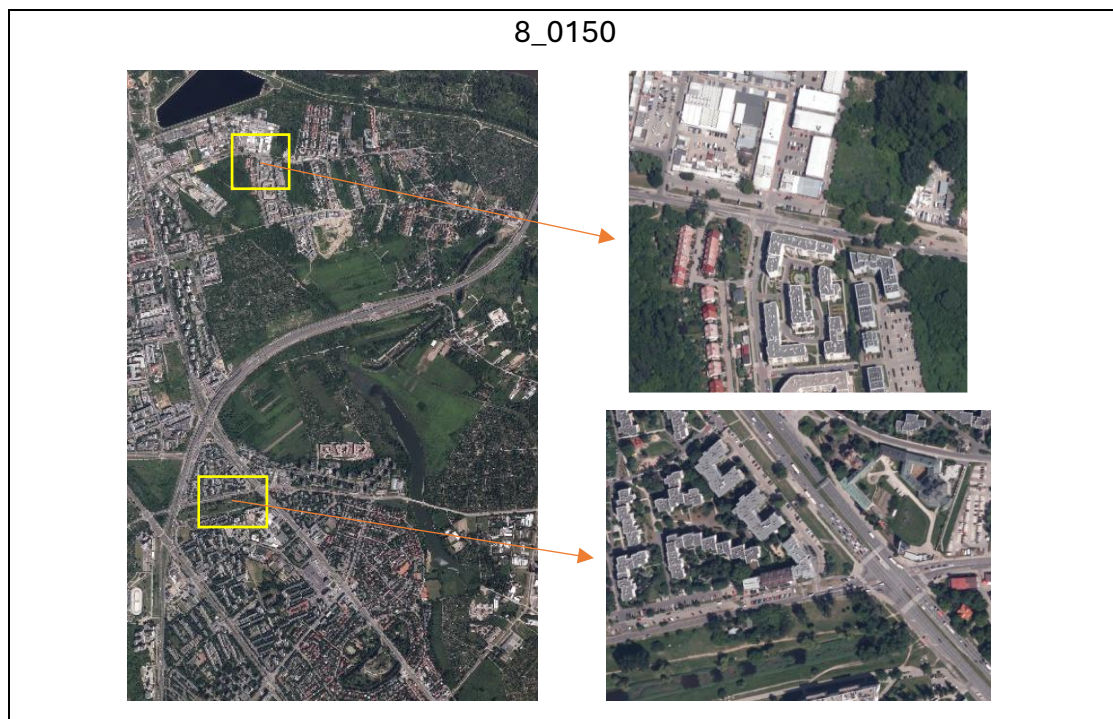
23_7612



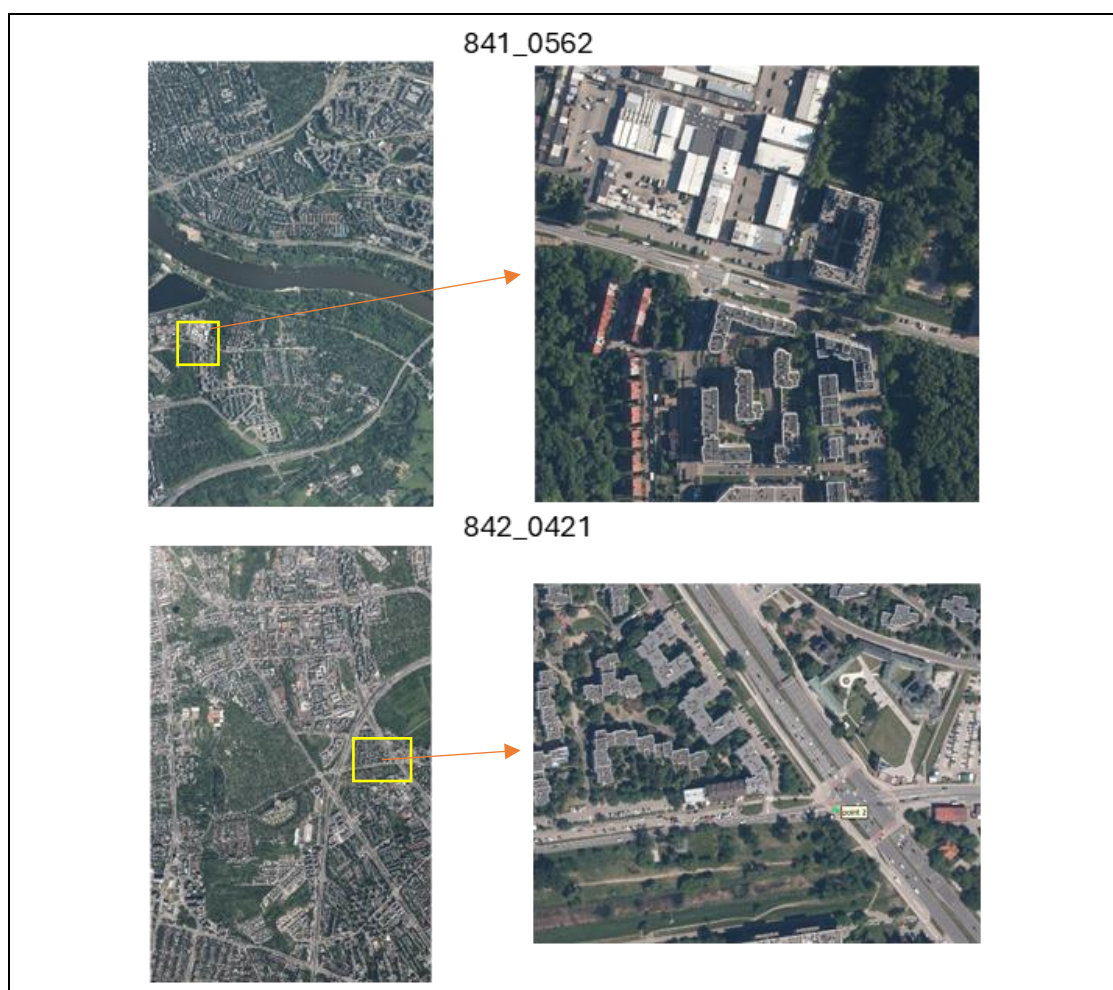
Rys. 6.4 Przykład zdjęć lotniczych wybranych ze zbioru zdjęć pozyskanych w roku 1986 dla miasta Warszawy oraz wybrane fragmenty ich powiększenia (prawa strona rysunku)



Rys. 6.5 Przykład zdjęć lotniczych wybranych ze zbioru zdjęć pozyskanych w roku 1994 dla miasta Warszawy oraz wybrane fragmenty ich powiększenia (prawa strona rysunku)



Rys. 6.6 Przykład zdjęcia lotniczego wybranego ze zbioru pozyskanego w roku 2014 dla miasta Warszawy oraz wybrane fragmenty jego powiększenia (prawa strona rysunku)



Rys. 6.7 Przykład zdjęć lotniczych wybranych ze zbioru zdjęć pozyskanych w roku 2022 dla miasta Warszawy oraz wybrane fragmenty ich powiększenia (prawa strona rysunku)

7. Budowa bloków fotogrametrycznych

Proces budowy bloków fotogrametrycznych (opisany w podrozdziale 5.1) na podstawie archiwalnych zdjęć lotniczych odbywał się w układzie lokalnym. Wykorzystano do tego zdjęcia lotnicze z roku 2022, 2014, 1994, 1986, 1976 (opisane w rozdziale 6). Do detekcji punktów charakterystycznych (Chen i in., 2020; Chen i Heipke, 2022) zastosowano powszechnie używane detektory typu *Hand-Crafted* (HC) (patrz rozdział 4), obejmujące dwa detektory punktowe: BRISK (*Binary Robust Invariant Scalable Keypoints*) i FAST (*Features from Accelerated Segment Test*), a także trzy detektorów powierzchniowe: SIFT (*Scale-Invariant Feature Transform*), CenSurE (*Center Surround Extremas*), SURF (*Speeded Up Robust Features*). Wykryte punkty zostały opisane za pomocą deskryptora SIFT i dopasowane przy wykorzystaniu algorytmu FLANN (*Fast Library for Approximate Nearest Neighbors*). W kolejnym etapie potencjalne punkty wiążące poddano filtracji geometrycznej, której celem było usunięcie wartości odstających. Na podstawie analizy literaturowej wskazującej, że typowy błąd wyznaczenia położenia punktów charakterystycznych mieści się w zakresie 0.2–0.5 piksela (Mikołajczyk i Schmid, 2005; Morel i Yu, 2009), przyjęto wartość progową błędu reprojektacji równą 1 piksel. Odchylenia przekraczające ten próg (odpowiadający 3σ w rozkładzie normalnym) uznano za statystycznie nieistotne i sklasyfikowano jako wartości odstające. W rezultacie uzyskano zbiory punktów oczyszczone z wartości odstających, które wykorzystano w dalszym procesie budowy bloków fotogrametrycznych.

Punkty charakterystyczne wykryte na zdjęciach, po przejściu etapów deskrypcji, dopasowania oraz filtracji geometrycznej, uzyskały status punktów wiążących. W dalszym etapie podzielono je na dwie grupy tj.: *Control Tie Point* (CTP) wykorzystywane do orientacji i *Check Tie Point* (CHTP) służące weryfikacji orientacji. W ostatnim etapie tworzenia bloków fotogrametrycznych zaimportowano dwie grupy punktów do programu Agisoft Metashape (AM), korzystając z interfejsu programowania aplikacji API (*Application Programming Interface*) oraz języka Python. Proces budowy bloku fotogrametrycznego oparto o metodę *Incremental SfM* (*Structure-from-Motion*) (patrz rozdział 5). Zdjęcia w bloku wyrównano przy użyciu metody niezależnych wiązek, opierając się na punktach CTP, co stanowiło modyfikację standardowego podejścia SfM zaimplementowanego w tym programie. W wyniku tego procesu powstała grupa bloków fotogrametrycznych odnoszących się do różnych okresów czasowych. Utworzone bloki

fotogrametryczne poddano statystycznej ocenie dokładności orientacji, wykorzystując błędy obliczone na punktach kontrolnych CHTP (patrz podrozdział 5.1 i 5.4).

Ponadto, w celu zweryfikowania dokładności budowy bloków fotogrametrycznych przy użyciu wcześniej wspomnianych detektorów, przeprowadzono dodatkową orientację tych bloków z wykorzystaniem algorytmu zaimplementowanego w oprogramowaniu AM (wersje 1.8 oraz 2.0). Szczegóły funkcjonowania tego detektora są zastrzeżone przez producenta. Jednak z informacji umieszczonych na forach i informacji udostępnionych przez twórców tego oprogramowania można stwierdzić, że jego mechanizm jest zbliżony do dobrze znanego algorytmu SIFT. Zastosowane w nim modyfikacje pozwalają jednak na osiągnięcie nieco wyższej dokładności wyrównania (Forum Agisoft, 2011).

Proces budowy bloku fotogrametrycznego w oprogramowaniu AM (*Incremental SfM*) przebiega w następujących etapach. Na początku dla każdego zaimportowanego zdjęcia przeprowadzana jest detekcja oraz opis punktów charakterystycznych. Następnie algorytm wybiera dwa zdjęcia spośród wszystkich o największej liczbie dopasowanych punktów charakterystycznych, na podstawie których tworzony jest początkowy model 3D. Dla tej pary zdjęć obliczana jest relacja geometryczna i przeprowadzana jest triangulacja punktów wiążących, co pozwala utworzyć pierwszy, wstępny model przestrzenny (model fotogrametryczny). Wykryte i opisane punkty poddawane są również filtracji w celu usunięcia wartości odstających. Szczegółowe kryteria filtracji stosowane przez program nie są znane. W dalszej kolejności proces przebiega iteracyjnie – do już zorientowanego modelu stopniowo dołączane są kolejne zdjęcia. Dla każdego nowego zdjęcia wyznaczana jest jego pozycja i orientacja w przestrzeni (wcięcie wstecz), a następnie wykonywana jest triangulacja nowych punktów. Po każdym takim kroku następuje wyrównanie wiązek (*bundle adjustment*), które optymalizuje zarówno orientacje wewnętrzne wszystkich zdjęć, jak i pozycje punktów 3D, a opcjonalnie również parametry wewnętrzne kamery. Proces ten jest powtarzany aż do momentu, gdy wszystkie zdjęcia zostaną zorientowane. Na końcu wykonywane jest wyrównanie końcowe całego bloku zdjęć.

Po zakończeniu budowy bloku fotogrametrycznego konieczna jest weryfikacja poprawności wyrównania. Należy zwrócić uwagę na wielkość uzyskanych błędów reprojekcji oraz na występowanie błędów grubych (obserwacji odstających), które mogą zaburzyć geometrię

bloku. Filtracja punktów w programie AM realizowana jest za pomocą funkcji *Gradual Selection*, która umożliwia usunięcie punktów na podstawie określonego progu błędu reprojektacji.

W prowadzonych badaniach filtracja punktów wiążących wyznaczonych przez algorytm programu AM dotyczyła bloków utworzonych ze zdjęć analogowych (z lat 1994, 1986 i 1976), dla których otrzymano błędy reprojektacji powyżej 3σ (podrozdział 7.4), a także zaobserwowano dużą liczbę błędów grubych, co negatywnie wpływało na ich geometrię. W celu poprawy relacji geometrycznych między zdjęciami w blokach zastosowano filtrację punktów wiążących z progiem 1 piksela. Zastosowanie filtracji znacząco zmniejszyło liczbę punktów wiążących, jednak ich liczba pozostała wystarczająca, aby zapewnić poprawną geometrię bloków. W oparciu o punkty kontrolne CHTP utworzone bloki fotogrametryczne poddano ocenie dokładności orientacji.

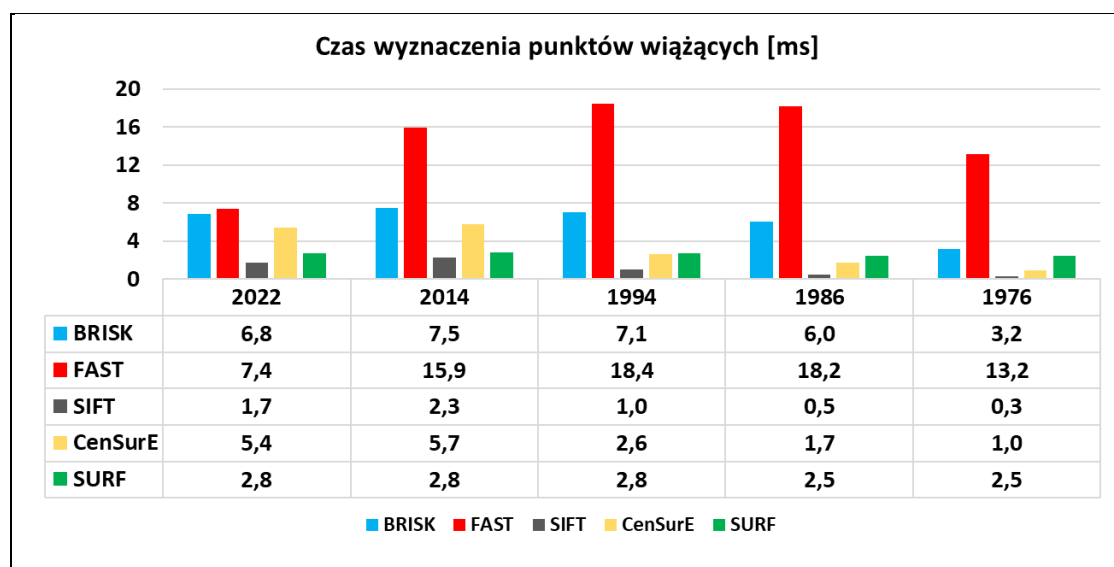
W rozdziale przedstawiono statystyczne informacje dotyczące zastosowania algorytmów widzenia maszynowego (CV – *Computer Vision*) w celu wyznaczenia punktów wiążących, które mają istotny wpływ na dokładność budowy bloków fotogrametrycznych. Analiza obejmuje takie aspekty, jak liczba wykrytych punktów, ich rozmieszczenie, wielkość błędów dopasowania par zdjęć w bloku fotogrametrycznym, błędy orientacji wzajemnej bloków fotogrametrycznych oraz czas działania poszczególnych detektorów odnoszący się do aspektu ekonomicznego pomiaru.

Przy ocenie przydatności zastosowanych algorytmów w procesie budowy bloków fotogrametrycznych uwzględniono następujące kryteria: rodzaj detektora, wielkość GSD (*Ground Sampling Distance*), rodzaj zdjęcia (cyfrowe lub analogowe), jakość zdjęć (parametry radiometryczne), format zdjęć, datę wykonania zdjęć oraz obszar pokrycia terenu.

7.1 Ocena czasu detekcji punktów wiążących w bloku fotogrametrycznym

Pierwsza statystyczna ocena budowy bloków fotogrametrycznych w podejściu HC, z wykorzystaniem algorytmów BRISK, FAST, SIFT, CenSurE i SURF, obejmowała analizę czasu pomiaru (łącznie detekcji, deskrypcji i dopasowania) pojedynczego punktu wiążącego w blokach fotogrametrycznych z lat 2022, 2014, 1994, 1986 i 1976 (Rys. 7.1). Czas ten określono jako iloraz całkowitego czasu detekcji, opisu i dopasowania wszystkich punktów oraz liczby wykrytych punktów. Jednostka określająca czas pomiaru punktu odnosiła się do milisekund

(1ms – 0.001 s). Analizując czas wyznaczenia punktu przez poszczególne algorytmy należało również odnieść się do liczby wyznaczonych punktów.



Rys. 7.1 Czas pomiaru automatycznego pojedynczego punktu wiążącego na archiwalnych zdjęciach lotniczych z roku 2022, 2014, 1994, 1986 i 1976 w oparciu o algorytmy HC

Analiza wykresu kolumnowego (rys. 7.1) wskazuje, że różnice czasowe pomiaru pojedynczego punktu przez poszczególne detektory HC w każdym bloku są na podobnym poziomie (podobny przebieg wykresu kolumnowego w każdym bloku). Najkrótszy czas pomiaru uzyskano dla detektora SIFT, a najdłuższy dla FAST. Porównując czasy pomiaru detektora SIFT w stosunku do detektora FAST w bloku 2022 był on 4-krotnie krótszy, w bloku 2014 7-krotnie, w bloku 1994 20-krotnie, w bloku 1986 37-krotnie, a w bloku 1976 40-krotnie. Algorytm FAST, zoptymalizowany pod kątem szybkiego przetwarzania, działa szczególnie efektywnie na obrazach o niskiej rozdzielczości dzięki mniejszej liczbie operacji. W przypadku obrazów o wysokiej rozdzielczości (czyli większej liczbie pikseli), ze względu na swój sposób działania, wykrywa więcej punktów kluczowych, co zwiększa zapotrzebowanie na zasoby obliczeniowe (patrz podrozdział 4.2.1).

Analizując wpływ wielkości GSD na czas pomiaru, nie zaobserwowano znaczących różnic pomiędzy danymi z 2022 roku (GSD 5 cm) i 2014 roku (GSD 24 cm), z wyjątkiem detektora FAST, dla którego czas detekcji pojedynczego punktu w 2022 roku był dwukrotnie krótszy. W odniesieniu do zdjęć analogowych z lat 1994 (GSD 7 cm) i 1986 (GSD 5 cm), gdzie różnica

w wielkości GSD była niewielka, stwierdzono, że parametr ten nie wpływa znacząco na czas wyznaczenia punktów wiążących.

Odnosząc się do parametru rodzaj zdjęć (cyfrowe vs. analogowe), parametr ten istotnie wpływa na czas pomiaru punktów wiążących tylko dla dwóch algorytmów powierzchniowych SIFT i CenSurE. W przypadku detektora SIFT różnica czasu pomiaru pomiędzy zdjęciami cyfrowymi a analogowymi była od 3 do 5 razy dłuższa dla zdjęć cyfrowych, w zależności od badanego bloku. Dla detektora CenSurE różnice te wynosiły od 2 do 6 razy. Może to wynikać z faktu, że zdjęcia cyfrowe mają dużo lepszą jakość radiometryczną w stosunku do zdjęć analogowych. Należy zaznaczyć, że oprogramowanie zastosowane w pracy przetwarzało zarówno zdjęcia cyfrowe i analogowe do postaci 8-bitowej. Wyższa jakość radiometryczna zdjęć cyfrowych jest konsekwencją ich początkowej rozdzielczości 11-bitowej, która została przekształcona do 8 bitów. Natomiast rzeczywista rozdzielczość radiometryczna zdjęć analogowych, które poddano skanowaniu z 8-bitową głębią zapisu, jest mniejsza od tej wartości. W rezultacie obrazy cyfrowe, charakteryzujące się lepszym odwzorowaniem tonalnym (większą rozpiętością i lepszym odwzorowaniem ciągłości tonalnej), umożliwiają skuteczniejsze wykrywanie punktów charakterystycznych przez algorytmy takie jak SIFT czy CenSurE, które opierają się na analizie lokalnych zmian gradientów stopni szarości. Większa liczba wykrywanych punktów przekłada się jednak na wydłużenie czasu przetwarzania i pomiaru.

W przypadku detektorów BRISK, FAST i SURF rodzaj zdjęcia (cyfrowe, analogowe) nie miał istotnego wpływu na czas pomiaru punktów, co pozwala uznać te algorytmy za bardziej powtarzalne pod względem wydajności, niezależnie od typu obrazu. Algorytmy FAST i BRISK opierają swoje działanie na analizie lokalnych zmian jasności pikseli w obrazie. W detektorach FAST i BRISK intensywności pikseli są bezpośrednio porównywane w określonych sąsiedztwach, natomiast SURF wykorzystuje obrazy całkowite do efektywnego obliczania gradientów i filtracji splotowej.

Detektorem, dla którego czas wykrywania punktów pozostawał na stałym poziomie (około 3 ms), niezależnie od wartości GSD, rodzaju zdjęcia (cyfrowe lub analogowe) oraz jakości radiometrycznej zdjęć (blok 1976), był detektor SURF. Odmienną tendencję zaobserwowano w przypadku bloku 1976, w którym czasy pomiaru pojedynczego punktu dla detektorów BRISK,

FAST, SIFT i CenSurE były o 50–70% krótsze w porównaniu z blokiem 1986. Zależność tą można przypisać niskiej jakości radiometrycznej zdjęć z tego okresu, co wpłynęło na specyfikę działania wspomnianych algorytmów.

W przypadku budowy dużych bloków fotogrametrycznych, w których liczba punktów wiążących może sięgać kilkunastu milionów, wybór odpowiedniego detektora ma kluczowe znaczenie ekonomiczne, zarówno pod względem czasu obliczeń, jak i zasobów obliczeniowych. Wskazuje to na konieczność uwzględnienia nie tylko czasu detekcji pojedynczego punktu, ale również i całkowitego czasu obliczeniowego związanego z wyznaczeniem wszystkich punktów wiążących, co w znaczący sposób wpływa na ogólną wydajność procesu fotogrametrycznego.

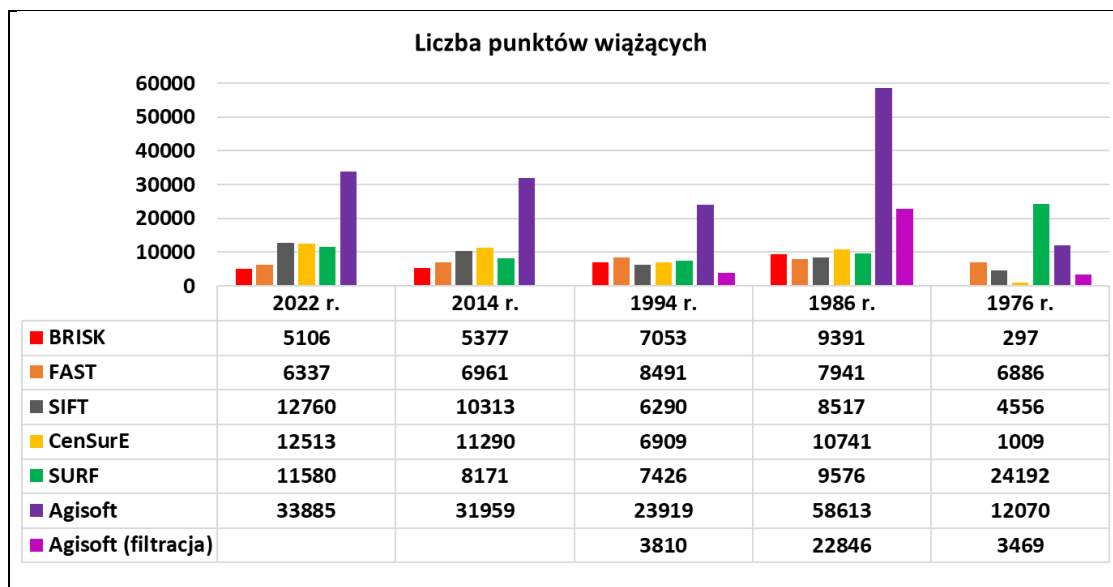
Wnioski:

1. Detektor SIFT zapewniał najkrótszy czas pomiaru pojedynczego punktu wiążącego we wszystkich analizowanych blokach, a detektor FAST – najdłuższy. Różnice w czasie pomiaru między SIFT a FAST stawały się bardziej wyraźne w blokach zdjęć analogowych (np. blok 1976: SIFT 40 razy szybszy od FAST).
2. Różnice w wielkości GSD pomiędzy blokami z lat 2022 (GSD 5 cm) i 2014 (GSD 24 cm) nie miały istotnego wpływu na czas pomiaru. Wyjątek stanowi detektor FAST, dla którego czas pomiaru w 2022 roku był dwukrotnie krótszy, co potwierdza założenia teoretyczne działania algorytmu.
3. Analogiczne wnioski dotyczą zdjęć analogowych z lat 1994 (GSD 7 cm) i 1986 (GSD: 5 cm), gdzie różnica w GSD również nie wpłynęła znacząco na czas pomiaru.
4. Detektory SIFT i CenSurE działały wolniej dla zdjęć cyfrowych (3-6 razy dłuższy czas w porównaniu do zdjęć analogowych) ze względu na wyższą rozdzielczość radiometryczną obrazów cyfrowych.
5. Działanie detektorów BRISK, FAST i SURF nie były zależne od rodzaju zdjęcia, co czyni je bardziej powtarzalnymi w kontekście wydajności.
6. Najbardziej stabilnym algorytmem okazał się SURF, którego czas detekcji punktów był stały (ok. 3 ms), niezależnie od: wielkości GSD, rodzaju zdjęcia (cyfrowe czy analogowe), jakości radiometrycznej zdjęć (np. blok 1976).

7. Przy tworzeniu bloków o dużym zasięgu, obejmujących dużą liczbę punktów wiążących wybór odpowiedniego detektora jest kluczowy z ekonomicznego punktu widzenia, wpływając na czas obliczeń i wykorzystanie zasobów.
8. Ważne jest uwzględnienie nie tylko czasu pomiaru pojedynczego punktu (detekcja, deskrypcja, dopasowanie), ale także całkowitego czasu obliczeniowego dla wszystkich punktów, co bezpośrednio przekłada się na wydajność i koszty procesu przetwarzania.

7.2 Ocena liczby wyznaczonych punktów wiążących i ich rozkładu w bloku fotogrametrycznym

Kolejna analiza budowy bloku obejmowała ocenę liczby wyznaczonych punktów wiążących oraz ich rozmieszczenia w obszarze wspólnego pokrycia zdjęć, odnoszącym się zarówno do par zdjęć w szeregu o pokryciu podłużnym, jak i do par zdjęć z sąsiednich szeregów o pokryciu poprzecznym. Przy ocenie liczby wyznaczonych punktów wiążących w utworzonych blokach fotogrametrycznych uwzględniono różne czynniki, takie jak rodzaj zastosowanego detektora (szczegółowo opisanego w rozdziale 4), liczba zdjęć w bloku i ich wzajemne pokrycie, skala zdjęć, rozdzielczość zdjęć oraz ich jakość radiometryczną. Zróznicowanie bloków odnoszące się do wspomnianych czynników wymagało przeprowadzenia niezależnej analizy liczby punktów i ich rozmieszczenia dla każdego bloku, z uwzględnieniem poszczególnych algorytmów HC.



Rys. 7.2 Liczba punktów wiążących (łącznie CTP i CHTP) wykrytych przez detektory HC w blokach fotogrametrycznych z roku 2022 r., 2014 r., 1994 r., 1986 r., 1976 r.

Liczba punktów wiążących (łącznie punktów CTP i CHTP) wykrytych przy zastosowaniu detektorów punktowych i powierzchniowych została przedstawiona na rysunku 7.2. Wyniki zaprezentowano w formie wykresu kolumnowego, obejmując zarówno wyniki dla detektorów zaimplementowane w środowisku C++ (BRISK, FAST, SIFT, CenSurE i SURF), jak i te dostępne w środowisku komercyjnego oprogramowania AM.

W przypadku punktów wyznaczonych za pomocą algorytmu oprogramowania AM (rys. 7.2) uwzględniono zarówno punkty przed etapem filtracji, jak i po jej wykonaniu (próg filtracji 1 piksel).

W bloku utworzonym ze zdjęć z roku 2022 najmniej punktów wykryto za pomocą detektorów punktowych BRISK i FAST, osiągając około 40-50% liczb punktów wykrytych przez detektory powierzchniowe SIFT, CenSurE i SURF. Taka różnica jest zgodna z zasadami działania tych algorytmów (patrz rozdziały 4.2.1 i 4.2.2). W bloku utworzonym ze zdjęć z roku 2014 zaobserwowano podobny trend, liczba punktów wykrytych przez detektory punktowe była o około 50-60% mniejsza niż dla detektorów powierzchniowych. Jedynie detektor SURF wykrył o 21-28% mniej punktów w porównaniu do pozostałych detektorów powierzchniowych. Przyczyną tych różnic może być odmienne pokrycie terenu, ponieważ w bloku z 2014 roku występowało mniej elementów zabudowy, a więcej obszarów niezabudowanych w porównaniu do bloku z 2022 roku, co mogło wpłynąć na mniejszą liczbę punktów wykrywanych przez algorytmy.

W bloku utworzonym ze zdjęć z roku 1994 liczba punktów wykrytych przez detektory BRISK, FAST, SIFT, CenSurE i SURF mieściła się w przedziale od 6290 do 8491. Zbliżona do siebie liczba wykrytych punktów może wynikać z tego, że większość obszaru analizowanego w tym przypadku stanowiła zabudowa. W bloku utworzonym ze zdjęć z roku 1986 liczba punktów wykrytych przez detektory punktowe była o około 12-25% mniejsza w porównaniu z detektorami powierzchniowymi. Różnica ta nie jest znacząca. Zróźnicowanie liczby punktów wykrytych przez detektory punktowe i powierzchniowe w bloku 1986 wynika częściowo z większego zasięgu obszaru tego bloku w porównaniu do bloku z 1994 roku oraz nieco odmiennego charakteru pokrycia terenu (rys. 6.1).

Porównując liczbę punktów wykrytych w blokach zdjęć cyfrowych i analogowych, zauważalna jest mniejsza liczba punktów na zdjęciach analogowych, szczególnie w przypadku detektorów powierzchniowych. Przyczyną tego jest słabsza radiometria zdjęć analogowych w porównaniu ze zdjęciami cyfrowymi.

Analizując blok utworzony ze zdjęć z roku 1976, które charakteryzują się słabą radiometrią, widoczny jest spadek liczby wykrytych punktów dla detektora punktowego BRISK (23-krotny w stosunku do FAST) oraz dla detektora powierzchniowego CenSurE (4,5-krotny w stosunku do SIFT). Spadek liczby wykrywanych punktów w detektorze BRISK może wynikać z tego, że analizuje on otoczenie punktów kluczowych przy użyciu wzorca pierścieniowego oraz wykorzystuje cechy oparte na lokalnych różnicach jasności, identyfikując stabilne punkty zgodnie z teorią Lowe'a (1999) oraz Mikołajczyka i Tuytelaarsa (2007). W przypadku zdjęć o słabej radiometrii, czyli niskim kontraście, detektor – ze względu na zastosowane podejście piramidalne (gdzie wykrywanie punktów odbywa się na różnych poziomach piramidy z odmiennymi parametrami *octree*) – eliminuje większą liczbę niestabilnych punktów, co prowadzi do ograniczenia liczby wykrytych cech. Z kolei FAST, który opiera się na prostszej analizie różnic jasności wokół piksela przy zastosowaniu wzorca kołowego, jest mniej wrażliwy na słabą radiometrię. Dzięki swojej prostocie potrafi wykryć więcej punktów nawet przy słabej jakości obrazu.

Mniejsza liczba punktów wykrywanych przez detektor CenSurE w porównaniu do SIFT jest przypuszczalnie efektem różnic w sposobie analizy struktury obrazu. CenSurE opiera się na detekcji ekstremów różnicy wartości jasności na pełnej rozdzielczości zdjęcia, przy czym

wykorzystuje aproksymację filtrów Laplace'a (LoG). Jego działanie jest ściśle uzależnione od lokalnych zmian gradientów stopni szarości, co czyni go bardziej podatnym na spadek efektywności w warunkach słabej radiometrii.

Z kolei SIFT, który wykorzystuje różnicę Gaussów (DoG) jako przybliżenie operatora Laplace'a (LoG), działa z większą dokładnością dzięki gaussowskiej aproksymacji w różnych skalach oraz z wykorzystaniem informacji o gradientach. Te zaawansowane metody pozwalają mu lepiej radzić sobie w sytuacjach niskiej jakości radiometrycznej obrazu, co przekłada się na większą liczbą wykrywanych punktów.

Detektor SURF w tym bloku wykrył 5 razy więcej punktów niż detektor SIFT. Wyższa liczba punktów wykrytych przez SURF może wynikać z tego, że algorytm wykorzystuje szybką aproksymację macierzy Hessego do wykrywania punktów w różnych skalach (podrozdział 4.4.2), co pozwala mu skutecznie znajdować cechy nawet przy niskim kontraście. Dodatkowo dzięki zastosowaniu filtrów pudełkowych i obrazów całkowych jest bardziej odporny na szумы i rozmycia, które często występują a zdjęciach lotniczych analogowych o słabej radiometrii. Ponadto, wykorzystanie falek Haara pozwala na dokładniejsze wyznaczanie kierunku gradientów (Bay i in., 2008).

Odnosząc się do rozdzielczości terenowej GSD zdjęć lotniczych, dla małych wartości GSD (np. 5–7 cm) algorytmy wykrywają większą liczbę punktów kluczowych, co wynika z wyższej szczegółowości obrazu. Natomiast przy dużych wartościach GSD (np. 20 – 30 cm) ze względu na ograniczoną czytelność detali na obrazach liczba wykrytych punktów jest zdecydowanie mniejsza.

Zastosowanie algorytmu oprogramowania AM wykazało, że w każdym bloku (2022, 2014, 1994, 1986), poza blokiem 1976 (ze względu na słabą radiometrię), liczba wyznaczonych punktów wiążących była znacznie większa (około 3–6 razy) w porównaniu z wynikami uzyskanymi przy użyciu innych algorytmów (rys. 7.2). Dla bloków 1994, 1986 i 1976 (zdjęcia analogowe) przeprowadzono filtrację punktów w oparciu o algorytm AM z filtracją, co było konieczne z uwagi na duże błędy pomiarowe. Natomiast dla bloków 2022 i 2014 (zdjęcia cyfrowe), ze względu na niewielkie błędy pomiarowe, etap filtracji nie był wymagany (rys. 7.2). Porównując algorytm pomiaru punktów w oprogramowaniu AM bez filtracji z wersją z filtracją (próg filtracji: 1 piksel), widać znaczący spadek liczby punktów (około 3–6 krotny). Liczba

punktów po filtracji pozostaje jednak wystarczająco duża, by nie obniżyć dokładności orientacji bloku. Usunięcie odstających punktów wiążących z pomiaru wzmocniło relacje geometryczne bloku, co wpłynęło na poprawę dokładności jego orientacji. Ze względu na brak szczegółowych informacji o zasadach działania algorytmu AM, pełna analiza jego działania (wykrywalności) nie jest możliwa.

Wnioski odnoszące się do liczby wyznaczonych punktów wiążących w blokach fotogrametrycznych:

1. Liczba wyznaczonych punktów wiążących przez algorytmy HC bez względu czy były to detektory punktowe (BRISK, FAST), czy powierzchniowe (SIFT, CenSurE i SURF) jest wystarczająca do budowy bloków fotogrametrycznych z roku 2022, 2014, 1994, 1986, 1976.
2. Detektory powierzchniowe są bardziej efektywne i wykrywają więcej punktów niż detektory punktowe szczególnie w terenach zurbanizowanych o dużej liczbie szczegółów. Detektory punktowe (np. BRISK, FAST) wykrywają istotnie mniejszą liczbę punktów wiążących w porównaniu do detektorów powierzchniowych (np. SIFT, CenSurE, SURF). Różnica ta wynosi średnio 40–60%, co wynika z odmiennych zasad działania algorytmów.
3. Rodzaj pokrycia terenu (zabudowa vs. obszary niezabudowane) wpływa na liczbę wykrywanych punktów. Zdjęcia przedstawiających obszary o dużej gęstości zabudowy generują więcej punktów wiążących, co wzmacnia relacje geometryczne bloków.
4. Detektor SURF wykazuje wyraźną przewagę w analizie zdjęć o słabej jakości radiometrycznej (np. blok 1976), co jest związane z jego odpornością na szумы i rozmycia, dzięki zastosowaniu szybkiej aproksymacji macierzy Hessego oraz filtrów pudełkowych. Obszary o dużej liczbie detali (np. zabudowa miejska) sprzyjają detekcji większej liczby punktów wiążących.
5. Gorsza radiometria zdjęć analogowych wpływa negatywnie na liczbę wykrytych punktów wiążących, co zaobserwowano szczególnie w bloku z roku 1976. Jednak detektor SURF, dzięki swojej odporności na szумы i niską jakość radiometryczną, okazał się w tych warunkach bardziej skuteczny niż inne detektory.

6. Zdjęcia cyfrowe wykazują większą liczbę punktów wiążących niż analogowe, co jest wynikiem lepszej jakości radiometrycznej zdjęć cyfrowych. Różnica ta była szczególnie widoczna w przypadku detektorów powierzchniowych.
7. Przy mniejszych wartościach GSD algorytmy wykrywają więcej punktów ze względu na większą szczegółowość obrazu. Dla większych wartości GSD liczba punktów spadała ze względu na ograniczoną czytelność.
8. Algorytm AM w stosunku do algorytmów BRISK, FAST, SIFT, CenSurE i SURF, mimo braku szczegółowych informacji o jego działaniu, wykazał największą skuteczność w liczbie wykrywanych punktów bez względu na rodzaj zdjęcia.
9. Dla zdjęć analogowych punkty wyznaczone przez algorytm AM ze względu na słabszą radiometrię tych zdjęć powinny przejść proces filtracji.

Na rysunkach 7.4 i 7.5 przedstawiono przestrzenny rozkład punktów wiążących w poszczególnych blokach fotogrametrycznych (2022 r., 2014 r., 1994 r., 1986 r., 1976 r.) dla wykorzystanych detektorów. Rozmieszczenie punktów zależało zarówno od charakterystyki działania algorytmów HC (BRISK, FAST, SIFT, CenSurE i SURF), jak i od innych czynników, takich jak:

- data wykonania zdjęć – odzwierciedlająca zmiany pokrycia terenu w czasie;
- miesiąc wykonania zdjęć – związany z okresem wegetacyjnym roślin i jego wpływem na pokrycie terenu;
- zasięg bloków fotogrametrycznych – uzależniony od skali zdjęć i ich rozmiaru.

W analizowanych blokach fotogrametrycznych punkty wiążące, w zależności od rodzaju detektora, zostały wykryte na różnych elementach pokrycia terenu:

1. Detektory punktowe.

Zgodnie z ich charakterystyką działania (wykrywanie narożników i lokalnych maksimów kontrastu) punkty były identyfikowane na elementach o wyraźnych krawędziach, takich jak:

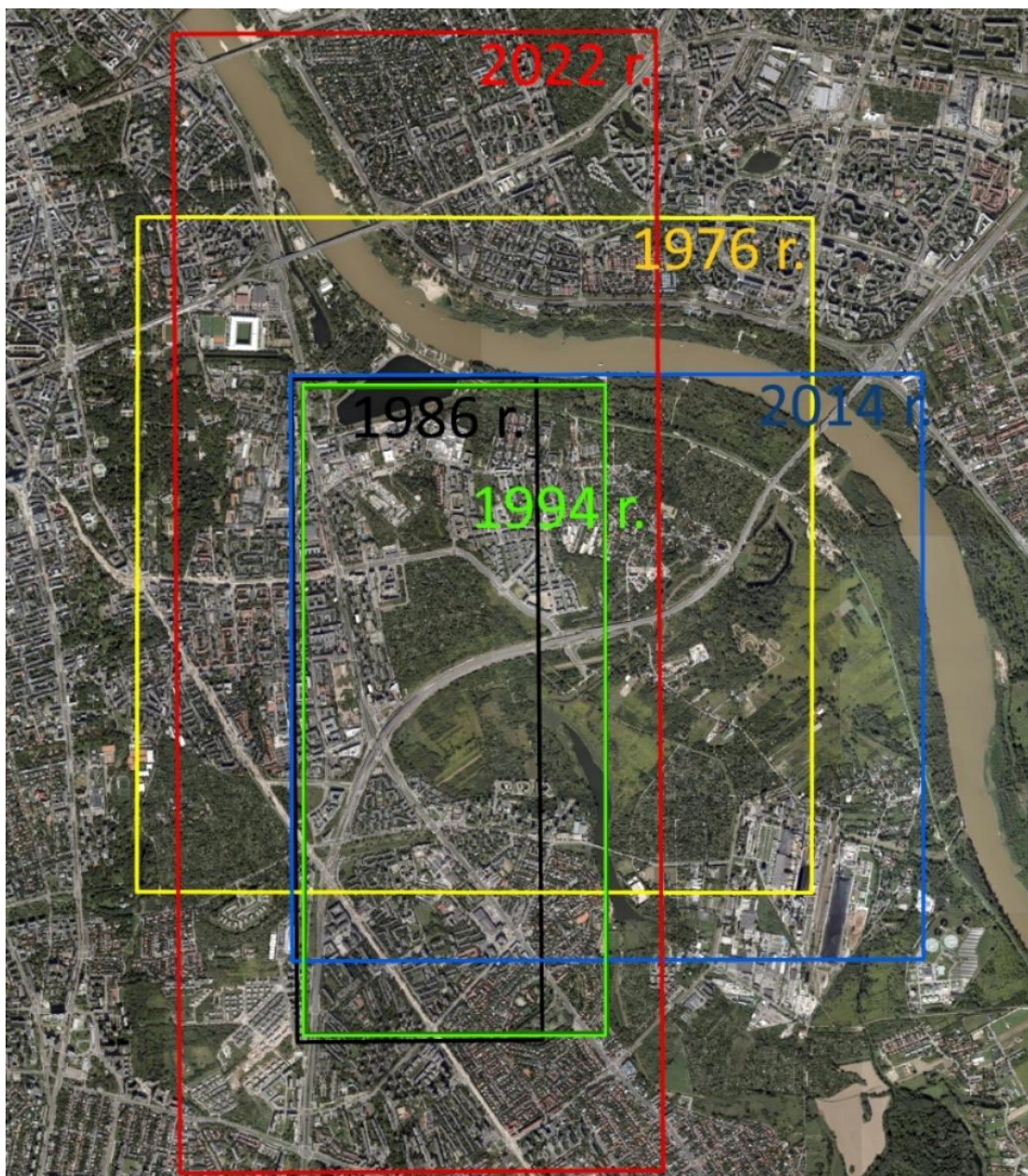
- narożniki budynków,
- krawędzie dróg i chodników,
- narożniki poziomych znaków drogowych,
- drzewa z wyraźnymi krawędziami cienia.

2. Detektory powierzchniowe.

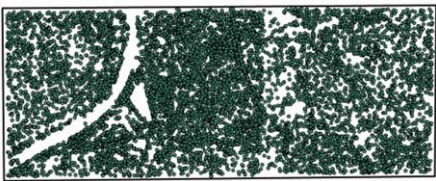
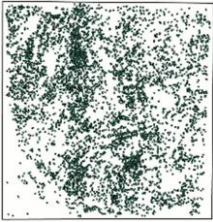
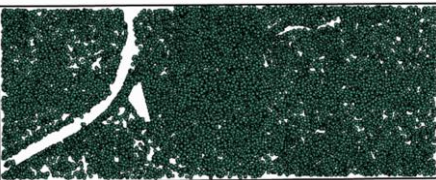
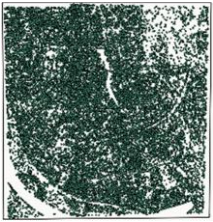
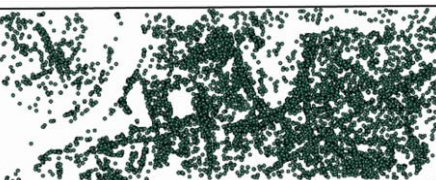
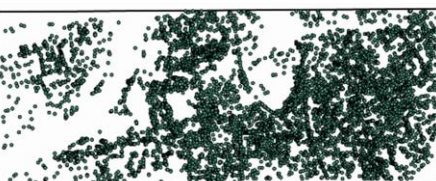
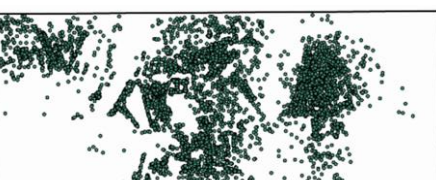
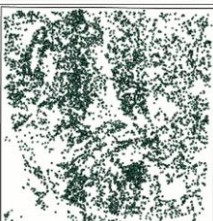
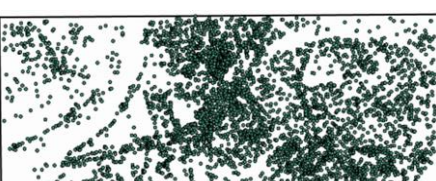
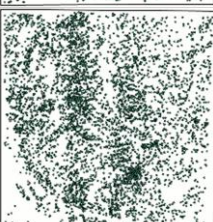
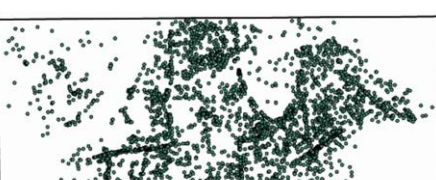
Zgodnie z ich charakterystyką działania (identyfikują punkty kluczowe na podstawie lokalnych zmian gradientu stopni szarości w obrazie w różnych skalach i kierunkach poprzez analizę większych obszarów) punkty wiążące były identyfikowane na takich elementach jak:

- powierzchnie budynków, dachów i chodników,
- powierzchnie dróg i poziomych znaków drogowych,
- powierzchnie ogródków działkowych,
- obszary porośnięte niską roślinnością (np. krzakami),
powierzchnie zadrzewione.

W celu dokładniejszej oceny rozmieszczenia punktów wiążących w poszczególnych blokach fotogrametrycznych, różniących się skalą, rozdzielczością oraz stopniem pokrycia zdjęć, a także dla porównania tych rozkładów między blokami, na rysunku 7.3 przedstawiono zasięgi pokryć tych bloków na obszarze miasta Warszawy.



Rys. 7.3 Zasięgi pokrycia utworzonych bloków fotogrametrycznych na obszarze miasta Warszawy (okolice Czerniakowa) na podkładzie ortofotomapy. Zasięg bloku 1976 – kolor żółty, zasięg bloku 1986 – kolor czarny, zasięg bloku 1994 – kolor zielony, zasięg bloku 2014 – kolor niebieski, zasięg bloku 2022 – kolor czerwony

	2022 r.	2014 r.
AM (filtracja)		
AM		
SURF		
CenSurE		
SIFT		
FAST		
BRISK		

Rys. 7.4 Rozkład punktów wiążących wyznaczonych przez detektory HC dla bloków 2022 i 2014 (zdjęcia cyfrowe)

	1994 r.	1986 r.	1976 r.
AM (filtracja)			
AM			
SURF			
CenSurE			
SIFT			
FAST			
BRISK			

Rys. 7.5 Rozkład punktów wiążących wyznaczonych przez detektory HC dla bloków 1994, 1986, 1976 (zdjęcia analogowe)

Ze względu na podobną radiometrię zdjęć w tworzonych blokach fotogrametrycznych, a także odpowiednią liczbę i optymalny rozkład pomierzonych punktów wiążących nie było potrzeby szczegółowej analizy miejsc położenia tych punktów na elementach terenowych. Taka analiza została przedstawiona dla detekcji punktów charakterystycznych pomiędzy zdjęciami o odmiennej radiometrii (ze względu m.in. różnej daty ich wykonania) w rozdziale 9 (*Sekwencyjna orientacja archiwalnych bloków fotogrametrycznych*).

Niskie wartości błędów orientacji bloków ze zdjęć cyfrowych (patrz podrozdział 7.4), w oparciu o punkty wiążące wyznaczone przez algorytm AM, nie wymagały dodatkowej filtracji. Dlatego na rysunku 7.4 oznaczono to jako miejsce puste. Natomiast w przypadku bloków utworzonych ze zdjęć analogowych, wielkości błędów orientacji były na tyle duże, że punkty wyznaczone przez algorytm AM musiały zostać poddane dodatkowej filtracji. Najlepszy rozkład punktów wiążących pod względem ich rozmieszczenia, liczby oraz dokładności wyznaczenia spośród wszystkich zastosowanych algorytmów uzyskano przy wykorzystaniu algorytmu AM w blokach utworzonych ze zdjęć cyfrowych. Dla bloków utworzonych ze zdjęć analogowych również algorytm AM zapewnił najlepszy rozkład punktów, jednak dotyczyło to punktów poddanych filtracji (rys. 7.5).

Dla wykorzystanych detektorów BRISK, FAST, SIFT, CenSurE i SURF rozkład pomierzonych punktów wiążących w blokach był następujący:

Blok 2022 – Najlepszy rozkład punktów uzyskano dla detektorów FAST, CenSurE i SURF, natomiast gorszy ich rozkład dla detektorów BRISK i SIFT. Ze względu na zastosowaną filtrację w południowej i północnej części bloku nie zidentyfikowano punktów. Ponadto, na obszarach związanych z wodą (charakteryzujących się stałym kontrastem) żaden z detektorów nie wykrył punktów charakterystycznych;

Blok 2014 – Najlepsze wyniki uzyskano dla detektorów BRISK, FAST, SIFT, CenSurE i SURF. Na powierzchniach wodnych (stały kontrast) detektory nie wykryły żadnych punktów, a w obszarach nieużytków liczba wykrytych punktów była znacząco niższa;

Blok 1994 – Zarówno w przypadku detektorów punktowych, jak i powierzchniowych liczba wykrytych punktów była duża, a ich rozmieszczenie – równomierne. Podobnie jak w innych analizowanych przypadkach, detektory nie wykryły punktów na obszarach wodnych (stały kontrast), a liczba punktów wykrytych na terenach nieużytków była ograniczona;

Blok 1986 – Najlepszy rozkład punktów uzyskano zarówno przy pomiarze przez detektory punktowe, jak i powierzchniowe. Na powierzchniach wodnych (stały kontrast) nie zidentyfikowano punktów, a na terenach nieużytków liczba wykrytych punktów była niewielka;

Blok 1976 – Bardzo dobre wyniki uzyskano dla detektorów FAST, SIFT i SURF, natomiast nieco gorsze wyniki odnotowano dla detektorów BRISK i CenSurE. Detektory nie wykryły punktów na obszarach wodnych (stały kontrast), a na terenach nieużytków liczba punktów była ograniczona. Detektor BRISK w dużym stopniu opiera się na lokalnych różnicach jasności, co sprawia, że jego wydajność spada w warunkach słabej radiometrii. Z kolei detektor CenSurE opiera się na analizie ekstremów aproksymowanego Laplasjanu (LoG - *Laplacian of the Gaussian*) w wielu skalach, dlatego przy słabej jakości radiometrycznej — gdy gradienty są mało wyraźne — jego skuteczność również ulega pogorszeniu.

Wnioski odnoszące się do rozkładu wyznaczonych punktów wiążących w blokach fotogrametrycznych:

1. Analiza rozkładu punktów we wszystkich utworzonych blokach wskazuje, że niezależnie od zastosowanego detektora (FAST, BRISK, SIFT, CenSurE, SIFT, SURF) rozmieszczenie punktów jest poprawne i zapewnia prawidłową budowę wszystkich bloków.
2. Na obszarach wodnych, charakteryzujących się stałym kontrastem, żaden z testowanych detektorów nie wykrył punktów charakterystycznych. Wynika to z trudności w identyfikacji cech na jednolitych powierzchniach o niewielkim zróżnicowaniu tekstury.
3. Na nieużytkach wykrywano mniej punktów, co może być spowodowane niewielką różnorodnością jego pokrycia.
4. Najlepsze wyniki uzyskano przy wykorzystaniu detektorów FAST, CenSurE i SURF – Pomiar za pomocą tych detektorów pozwalał wielokrotnie uzyskać najlepszy rozkład punktów w analizowanych blokach czasowych. Są one skuteczne w zróżnicowanym pokryciu terenowym, co czyni je bardziej uniwersalnymi w porównaniu do innych detektorów.
5. Brak wyraźnych różnic w skuteczności detektorów punktowych i powierzchniowych. W blokach 1994 i 1986 zarówno detektory punktowe, jak i powierzchniowe osiągnęły najlepszy rozkład punktów, co sugeruje, że oba podejścia mogą być efektywne, niezależnie od pokrycia terenu.

6. Dla detektora zastosowanego w oprogramowaniu AM rozkład punktów, zarówno przed, jak i po filtracji, wykazuje najwyższą jakość pod względem równomiernego rozmieszczenia. Punkty te zostały zidentyfikowane zarówno na elementach infrastruktury tj. drogach, chodnikach, budynkach jak i na obszarach porośniętych roślinnością, co świadczy o wszechstronności tego algorytmu. Ze względu na brak szczegółowych informacji o działaniu zastosowanego detektora, nie można jednoznacznie określić przyczyn tak wysokiej skuteczności w wykrywaniu punktów charakterystycznych.
7. Algorytm AM zapewnił najlepszy rozkład punktów wiążących w blokach utworzonych ze zdjęć cyfrowych, co wynika z jego wysokiej skuteczności w przetwarzaniu obrazów o dobrej radiometrii.
8. Z punktu widzenia jakości budowy bloku fotogrametrycznego, rozkład punktów wyznaczonych zarówno przez algorytmy FAST, BRISK, SIFT, CenSurE, SURF, jak i algorytm oprogramowania AM, we wszystkich blokach, z wyjątkiem bloku 1976, był bardzo dobry. W przypadku bloku 1976, ze względu na słabą radiometrię zdjęć, rozkład punktów wyznaczonych przez detektory BRISK i CenSurE można ocenić jako dobry.

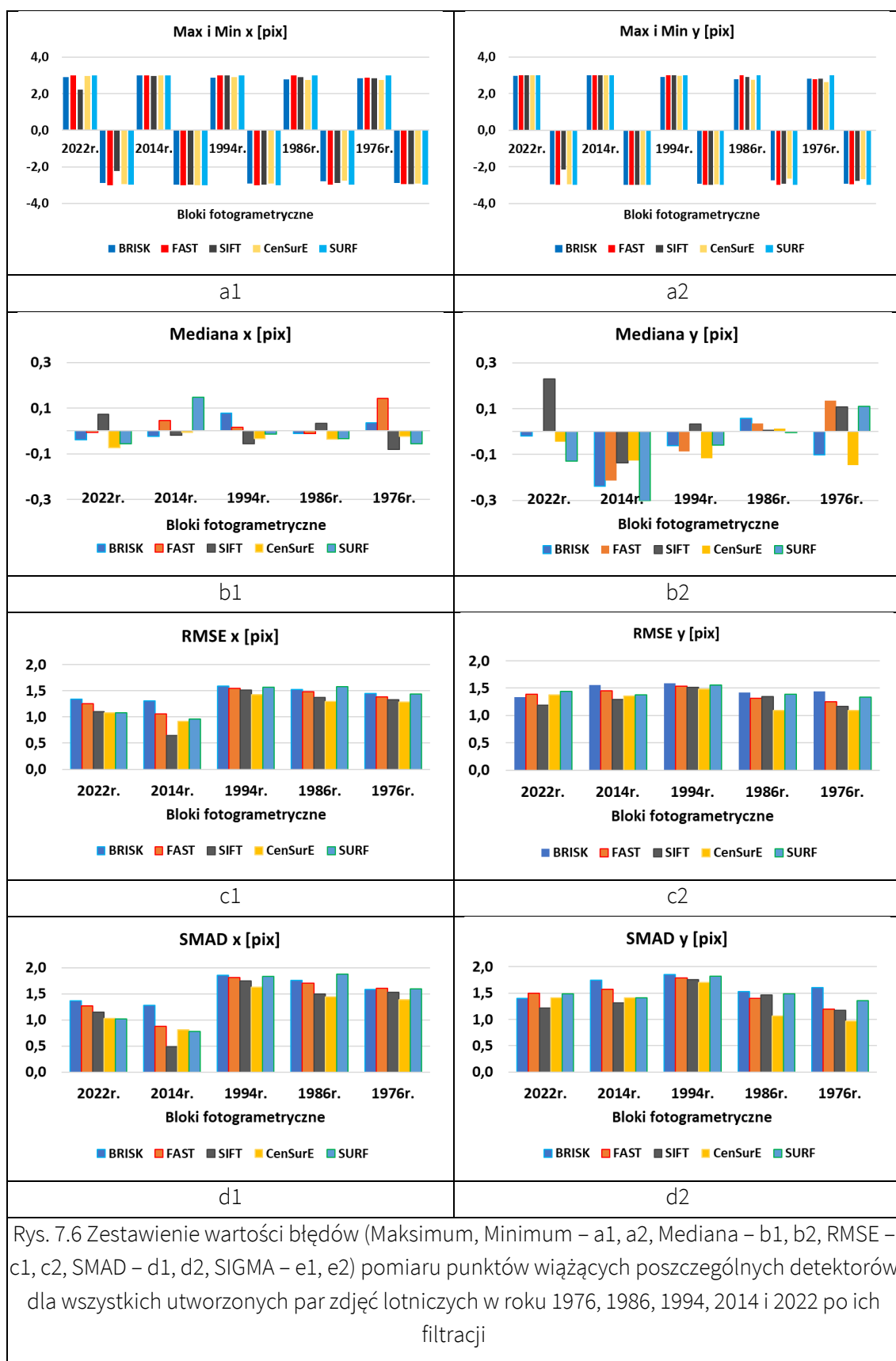
7.3 Ocena dokładności pomiaru punktów wiążących na parach zdjęć w bloku

Pierwszym i zarazem kluczowym etapem w procesie SfM jest wyznaczenie punktów wiążących na parach zdjęć w oparciu o model orientacji wzajemnej zdjęć. Dzięki temu możliwe jest określenie parametrów orientacji wzajemnej na podstawie punktów wiążących, które w kolejnym etapie przetwarzania danych służą jako dane wejściowe do całościowego wyrównania obserwacji w procesie budowy bloków bazującym na metodzie niezależnych wiązek (patrz podrozdział 5.1). Z tego też powodu istotna jest ocena wykrytych punktów wiążących poprzez analizę wartości błędu reprojekcji (współrzędnej x i y punktów) na zdjęciach (pary zdjęć) o wspólnym pokryciu podłużnym i poprzecznym w całym bloku fotogrametrycznym (rys. 7.6). Współrzędne x , y odnoszą się do układu obrazowego, gdzie oś x oznacza kierunek lotu, a oś y kierunek prostopadły do kierunku lotu.

Utworzone zbiory punktów wiążących, wygenerowane z wykorzystaniem algorytmów HC, stosowane są zarówno na etapie orientacji zdjęć w bloku fotogrametrycznym, jak i w procesie

generowania gęstych chmur punktów z zastosowaniem metod *Multi-View Stereo* (MVS), stanowiących podstawę tworzenia NMPT i ortofotomapy. Ze względu na konieczność zachowania kompromisu między dokładnością wyznaczenia punktów wiążących (kluczową dla prawidłowej orientacji zdjęć w bloku) a ich liczebnością (istotną dla efektywności metod MVS), wartość progu filtracji ustalono na poziomie 3 pikseli, co odpowiada trzykrotnemu odchyleniu standardowemu (3σ) w rozkładzie normalnym.

Poniższa analiza statystyczna odnosi się do punktów wiążących przed ich podziałem na CTP i CHTP.



W celu przeprowadzenia takiej oceny wykonano analizę błędów reprojektacji punktów wiążących dla par zdjęć w blokach z lat 1976, 1986, 1994, 2014 i 2022, opierając się na wybranych miarach statystycznych, które zestawiono w tabeli 5.1 (rozdział 5.4).

Na rysunku 7.6 przedstawiono wykresy ze zbiorczymi wartościami analizowanych wielkości błędów pomiaru punktów wiążących przy użyciu detektorów punktowych BRISK i FAST oraz powierzchniowe SIFT, SURF CenSurE dla zdjęć z roku 1976, 1986, 1994, 2011 i 2022.

Analizując maksymalną i minimalną wartość błędu (rys.7.6 a1, a2) pomiaru punktów wiążących (x, y) w poszczególnych blokach i dla wszystkich detektorów, stwierdzono, że mieszczą się one w granicach ± 3 pikseli, zgodnie z przyjętym progiem. Dzięki filtracji możliwe było odrzucenie wszystkich punktów, które mogłyby negatywnie wpłynąć na dokładność wyznaczenia orientacji bloku i jakość generowanych gęstych chmur punktów.

Wartość mediany błędu reprojektacji punktów wiążących przy użyciu algorytmów HC dla współrzędnych x i y (rys. 7.6 b1, b2) we wszystkich blokach jest bliska zero (od -0.3 do 0.3 piksela), co wskazuje na minimalny wpływ błędów systematycznych na pomiar.

W odniesieniu do wartości RMSE reprojektacji (rys. 7.6 c1, c2) stwierdzono, że we wszystkich blokach nie przekraczały one 1.6 piksela. Szczegółowa analiza wykazała, że różnice w wartościach RMSE dla współrzędnych x i y pomiędzy poszczególnymi algorytmami były niewielkie, z wyjątkiem bloku 2014. W tym przypadku różnica RMSE dla współrzędnej x pomiędzy detektorami BRISK i SIFT wyniosła 0.6 piksela.

W blokach utworzonych ze zdjęć analogowych (1994, 1986 i 1976) wartość RMSE reprojektacji dla pomiaru punktów wiążących w przypadku współrzędnej x mieściła się w przedziale od 1.3 do 1.5 piksela, natomiast dla współrzędnej y od 1.1 do 1.5 piksela. W blokach utworzonych ze zdjęć cyfrowych (2022 i 2014) wartość RMSE dla współrzędnej x również wynosiła od 1.3 do 1.5 piksela, podczas gdy dla współrzędnej y mieściła się w przedziale od 0.6 do 1.3 piksela. Niższe wartości RMSE dla współrzędnej x i y w blokach ze zdjęć cyfrowych w porównaniu do bloków ze zdjęć analogowych można przypisać lepszej jakości radiometrycznej zdjęć cyfrowych. Dla zdjęć analogowych, różnice w wartościach RMSE reprojektacji między współrzędnymi x i y są praktycznie niezauważalne.

Najmniejsze wartości RMSE w blokach utworzonych ze zdjęć cyfrowych uzyskano dla detektora SIFT, natomiast w blokach utworzonych ze zdjęć analogowych dla detektora

CenSurE. Największe wartości RMSE zarówno dla bloków cyfrowych i analogowych uzyskano dla detektora BRISK. Największa różnica wartości RMSE powstała dla bloku z roku 2014 pomiędzy detektorami BRISK i SIFT i wyniosła około 0.5 piksela. Niskie wartości błędów pomiaru punktów wiążących uzyskane przez analizowane detektory wskazują na poprawność budowy spójnych geometrycznie bloków fotogrametrycznych.

Wartości błędów SMAD (rys. 7.6 d1, d2) uzyskanych w wyniku pomiarów punktów wiążących przez algorytmy HC we wszystkich blokach fotogrametrycznych dla współrzędnych x i y był zbliżony do wartości RMSE, co wskazuje na brak błędów grubych.

Uzyskane wielkości błędów pomiarowych przez algorytmy HC w poszczególnych blokach jednoznacznie potwierdzają brak występowania błędów grubych w danych pomiarowych.

Analizując wartości błędów wyrażonych w skalach zdjęć zawarte na rysunku 7.6 odnoszących się do pomiaru punktów wiążących w kontekście wielkości GSD zdjęć (2022 – 5 cm, 2014 – 24 cm, 1994 – 7 cm, 1986 – 5 cm, 1976 – 25 cm) nie jest widoczny wpływ tego parametru na dokładność pomiaru.

Wnioski:

1. Wyznaczenie znaczącej równomiernie rozłożonej liczby punktów wiążących w oparciu o algorytmy HC jest podstawowym zadaniem procesu SfM. Jakość tych punktów, mierzona wartością błędu reprojekcji, istotnie wpływa na dokładność dalszych etapów przetwarzania, w tym na proces wyrównania przy użyciu metody niezależnych wiązek.
2. W żadnym z analizowanych bloków, poza blokiem z 2014 roku, zastosowane detektory nie wyróżniają się jednoznacznie jako najlepsze lub najgorsze pod względem generowanych błędów RMS. Zaobserwowane różnice są marginalne i mogą wynikać ze specyficznych cech danego bloku (różne rodzaj zdjęć, różne pokrycie). Dla zwiększenia dokładności budowy bloku fotogrametrycznego zaleca się stosowanie detektora powierzchniowego CenSurE w przypadku zdjęć analogowych oraz detektora SIFT w przypadku zdjęć cyfrowych.
3. W bloku z 2014 roku wystąpiła wyraźna różnica wartości RMSE pomiędzy detektorami BRISK i SIFT, która wyniosła około 0.5 piksela.

4. W blokach ze zdjęć cyfrowych (2022, 2014) wartości RMSE pomiarów punktów wiążących przez algorytmy HC były niższe niż w blokach ze zdjęć analogowych (1994, 1986, 1976), co można przypisać lepszej jakości radiometrycznej zdjęć cyfrowych.
5. Wartości błędów SMAD był zbliżony do wartości błędów RMS, co wskazuje na brak błędów grubych w pomiarze.
6. Pomimo różnic technologicznych (zdjęcia cyfrowe i analogowe) oraz warunków nalogu, uzyskane wyniki wskazują na stabilność pomiarów we wszystkich analizowanych blokach. Wartość mediany dla współrzędnych x i y oscylowała wokół 0 ± 0.3 piksela, co wskazuje na minimalny wpływ błędów systematycznych i wysoką jakość analizowanych danych.
7. Filtracja geometryczna z progiem 3 pikseli skutecznie wyeliminowała błędy grube, zapewniając poprawność danych wejściowych do procesu aerotriangulacji i generowania gęstych chmur punktów.
8. Brak wyraźnej zależności między wielkością GSD zdjęć a wartościami błędów pomiaru punktów wiążących wyrażonych w skali zdjęć (pikselach) wskazuje, że dokładność pomiaru punktów pozostaje na podobnym poziomie niezależnie od rozdzielczości przestrzennej zdjęć. Uzyskane błędy pomiarowe dla wszystkich bloków fotogrametrycznych mieszczą się w przyjętych granicach dokładnościowych.

7.4 Ocena dokładności budowy bloku fotogrametrycznego

Na podstawie wyznaczonych punktów wiążących (CTP) w oparciu o wykorzystane algorytmy, przeprowadzono proces aerotriangulacji każdego bloku (dla lat 2022, 2014, 1994, 1986 oraz 1976) w układzie lokalnym w skali zdjęcia przy użyciu oprogramowania AM. Przed przystąpieniem do procesu aerotriangulacji przeprowadzono filtrację punktów wiążących w celu eliminacji wartości odstających, przyjmując próg 1 piksela. Wartość ta została określona na podstawie doświadczeń praktycznych oraz danych literaturowych (Mikolajczyk i Schmid, 2005; Morel i Yu, 2009). Zastosowanie filtracji umożliwiło usunięcie błędnych dopasowań i poprawę jakości orientacji zdjęć, z uwzględnieniem rzeczywistej dokładności detekcji punktów charakterystycznych. Proóg ten został przyjęty zarówno dla punktów wyznaczonych przez wykorzystane detektory jak i detektor AM.

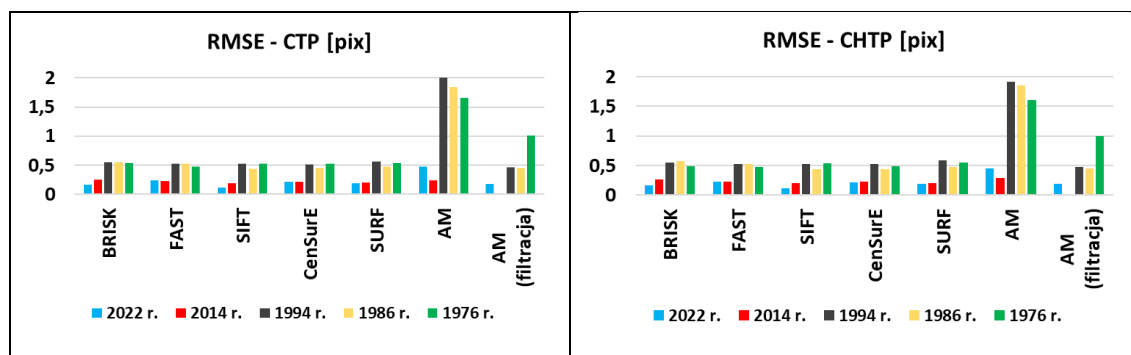
Proces aerotriangulacji realizowany w AM zarówno dla bloków tworzonych za pomocą punktów pomierzonych przez wskazane algorytmy jak i punktów pomierzonych przez algorytm AM odpowiada tzw. metodzie *Incremental SfM*. Rozpoczyna się ona od rekonstrukcji początkowego modelu na podstawie najlepiej dopasowanej pary zdjęć o największym pokryciu i liczbie punktów wiążących, a następnie na sukcesywnym dodawaniu kolejnych zdjęć do już zrekonstruowanego modelu. W procesie tym uwzględniano standardowe parametry, takie jak ogniskowa obiektywu, rozdzielczość zdjęć oraz wagi obserwacji współrzędnych obrazowych.

Ze względu na brak dostępnych danych kalibracyjnych dla kamer analogowych z lat 1994, 1986 i 1976, proces budowy bloków fotogrametrycznych w tych przypadkach obejmował etap samokalibracji kamer, realizowany w oparciu o wirtualne modele kamer. Parametry geometryczne niezbędne do utworzenia wirtualnej kamery, takie jak położenie znaczków tłowych, długość ogniskowej, odczytywano bezpośrednio ze zdjęć oraz z informacji zawartych w ich elementach pozaramkowych. W tym przypadku samokalibrację utożsamiana jest jako proces ukierunkowany na identyfikację, modelowanie i kompensację afinizmu zdjęć, wynikającego z istotnych skurczów i deformacji materiału nośnego, a nie jako procedurę prowadzącą do wyznaczenia skalibrowanej wartości ogniskowej, współrzędnych głównego punktu obrazu oraz parametrów dystorsji radialnej i tangencjalnej. Ze względu na dostępność metryk kalibracyjnych dla kamer cyfrowych, którymi wykonano zdjęcia lotnicze z lat 2022 i 2014, zastosowanie samokalibracji w tym przypadku nie było konieczne.

Przeprowadzono również analizę błędów orientacji bloków fotogrametrycznych uzyskanych przy użyciu algorytmów HC z wynikami generowanymi przez komercyjne oprogramowanie AM. Ocena dokładności orientacji wzajemnej bloków zdjęć została wykonana w oparciu o uzyskane wielkości błędów na punktach wiążących (CTP) oraz niezależnie na punktach kontrolnych (CHTP), które stanowiły 10% ogólnej ich liczby. Takie podejście umożliwiło weryfikację działania algorytmów pod kątem dokładności budowy bloku zdjęć w układzie lokalnym. Jako miarę dokładności przyjęto błąd średniokwadratowy (RMSE - *Root Mean Square Error*) reprojektacji (Tab. 5.1).

Na rysunku 7.7 przedstawiono wyniki orientacji dla ww. detektorów oraz algorytmów zastosowanych w oprogramowaniu AM. Analiza wartości błędów orientacji wykazała, że

średnia dokładność orientacji bloków wynosiła około 0.3 piksela. W związku z tym w celu eliminacji wartości odstających zdecydowano się przeprowadzić filtrację w oparciu o progową wartość błędu reprojektacji wynoszącej 1 piksel. Proces filtracji dla algorytmów BRISK, FAST, SIFT, CenSurE i SURF zrealizowano za pomocą autorskiego rozwiązania (programu) wykorzystującego API programu AM.



Rys. 7.7 Wartości RMSE (wypadkowa x y) orientacji wzajemnej zdjęć w bloków (2022 r., 2014 r., 1994 r., 1986 r., 1976 r.) uzyskanych na punktach CTP i CHTP dla zastosowanych algorytmów HC oraz algorytmu programu AM

Analizując wyniki przedstawione na rysunku 7.7 można zauważyć, że dokładności orientacji wzajemnej bloków zdjęć na podstawie zaimplementowanych detektorów punktowych i powierzchniowych (2022 r., 2014 r., 1994 r., 1986 r., 1976 r.) zarówno na punktach CTP jak i punktach CHTP są zbliżone. Różnice w wynikach pomiędzy nimi są zauważalne jedynie w dziesiątych częściach piksela, co można uznać za pomijalne, a jednocześnie świadczy to o poprawności orientacji wzajemnej sieci zdjęć (bloków fotogrametrycznych).

Interpretując RMSE można stwierdzić, że dla zdjęć analogowych z lat 1976, 1986 i 1994 jego wartości nie przekraczają 0.6 piksela, a dla zdjęć cyfrowych 0.3 piksela.

Analizując wyniki z poszczególnych lat można zauważyć:

- w roku 2022 największą wartość RMSE (0.2 piksela) uzyskano dla detektora FAST, a najmniejszy (0.1 piksela) dla detektora SIFT;
- w roku 2014 największą wartość RMSE (0.3 piksela) zanotowano dla detektora BRISK, a najmniejszy (0.2 piksela) dla detektorów SIFT i SURF;
- w roku 1994 największą wartość RMSE (0.6 piksela) wystąpił przy zastosowaniu detektora SURF, a najmniejszy (0.4 piksela) dla detektora SIFT;
- w roku 1986 największą wartość RMSE (0.6 piksela) uzyskano dla detektora BRISK, a najmniejszy (0.4 piksela) dla detektora SIFT;

- w roku 1976 wartość błędu reprojekcji na punktach wpasowania wyniósł 0.5 piksela zarówno dla detektorów punktowych (BRISK, FAST), jak i powierzchniowych (SIFT, CenSurE, SURF).

Odnosząc się do przedstawionych wyników, można stwierdzić, że algorytm SIFT charakteryzuje się najwyższą dokładnością we wszystkich analizowanych latach, uzyskując najniższe wartości RMSE. Największe błędy najczęściej generują detektory BRISK i FAST, natomiast detektor SURF wykazuje większą zmienność dokładności. W roku 1976 dla wszystkich zastosowanych detektorów osiągnięto porównywalny poziom błędu.

Analiza wyników wskazuje także, że dokładności orientacji dla zdjęć cyfrowych są średnio dwukrotnie wyższe niż dla zdjęć analogowych. Należy jednak podkreślić, różnice w wartościach RMSE dla zdjęć analogowych były niewielkie (rzędu 0.2 piksela), co pozwala stwierdzić, że wszystkie detektory zapewniają wykrywanie punktów wiążących z podobną dokładnością.

Na rysunku 7.7 przedstawiono również wartości RMSE orientacji wzajemnej bloków zdjęć uzyskane na punktach wykrytych przez algorytm AM przed filtracją i po filtracji w oparciu o progową wartość błędu reprojekcji (1 piksel). Filtracja punktów pomierzonych za pomocą algorytmu AM jest realizowana przy użyciu wbudowanej funkcji dostępnej w graficznym interfejsie programu. Rozwiązanie to, mimo swojej użyteczności, nie zapewnia pełnej eliminacji wartości odstających, co może negatywnie wpływać na stabilność i poprawność geometrii bloku fotogrametrycznego. W przeciwieństwie do tego, autorski program filtracyjny (środowisko Python) działa w sposób jednoznaczny, odrzucając z całego zbioru wszystkie punkty przekraczające zdefiniowany próg wartości błędu reprojekcji. Takie podejście pozwala na skuteczne eliminowanie punktów odstających i zwiększenie dokładności orientacji wzajemnej bloków zdjęć. Przed zastosowaniem filtracji geometrycznej punktów odstających wyznaczonych za pomocą algorytmu AM, wartości RMSE reprojekcji dla bloków utworzonych ze zdjęć analogowych wynosiły odpowiednio:

- 1.9 piksela dla roku 1994,
- 1.9 piksela dla roku 1986,
- 1.7 piksela dla roku 1976.

Zastosowanie filtracji pozwoliło na średnią 4-krotną redukcję błędu reprojekcji dla zdjęć analogowych, co umożliwiło uzyskanie wartości RMSE porównywalnych do wyników

uzyskanych w ramach zaproponowanej metody orientacji wzajemnej opartej na powszechnie wykorzystywanych detektorach HC. Dla bloków zdjęć cyfrowych błędy RMS reprojektacji wynosiły przed filtracją odpowiednio:

- 0.3 piksela dla roku 2014,
- 0.5 piksela dla roku 2022.

Ze względu na stosunkowo małe wartości błędów reprojektacji, dla bloków zdjęć cyfrowych dodatkowa filtracja punktów odstających nie była konieczna.

Wnioski oceny dokładności orientacji wzajemnej zdjęć w bloku (wykres kolumnowy)

1. Zarówno detektory punktowe, jak i powierzchniowe zapewniają porównywalną dokładność orientacji wzajemnej bloków zdjęć niezależnie od ich rodzaju (analogowe lub cyfrowe).
2. Wielkości błędów reprojektacji zdjęć analogowych i cyfrowych są niewielkie (0.2 – 0.5 piksela), co pozwala uznać, że wszystkie detektory zapewniają dobrą dokładność w wykrywaniu punktów wiążących zapewniając tym samym wysoką poprawność budowy bloków fotogrametrycznych.
3. Zdjęcia cyfrowe wykazują średnio dwukrotnie niższe wartości błędów reprojektacji w porównaniu do zdjęć analogowych.
4. Najmniejsze wartości błędów RMS orientacji wzajemnej zdjęć analogowych i cyfrowych uzyskano dla detektora SIFT. Największe błędy reprojektacji uzyskano dla detektorom BRISK (dla 1986, 2014) oraz SURF (dla 1994).
5. W podejściu AM bez filtracji wartości RMSE reprojektacji dla zdjęć analogowych wynosiły 1.7–2.1 piksela, co znacząco obniżało jakość orientacji.
6. Dla zdjęć analogowych zastosowanie filtracji punktów odstających z użyciem funkcji dostępnej w AM umożliwiło redukcję błędów reprojektacji średnio czterokrotnie. Wyniki RMSE po filtracji były zbliżone do tych uzyskanych dla zdjęć cyfrowych.
7. W przypadku zdjęć cyfrowych w podejściu AM, błędy RMS reprojektacji były relatywnie niskie (0.3–0.5 piksela), dlatego dodatkowa filtracja nie była konieczna.
8. Zastosowanie autorskiego programu filtracyjnego, który w sposób jednoznaczny usuwa wszystkie punkty przekraczające ustalony próg błędu reprojektacji, pozwala na znaczącą poprawę dokładności orientacji wzajemnej zdjęć w bloków. W przeciwieństwie do

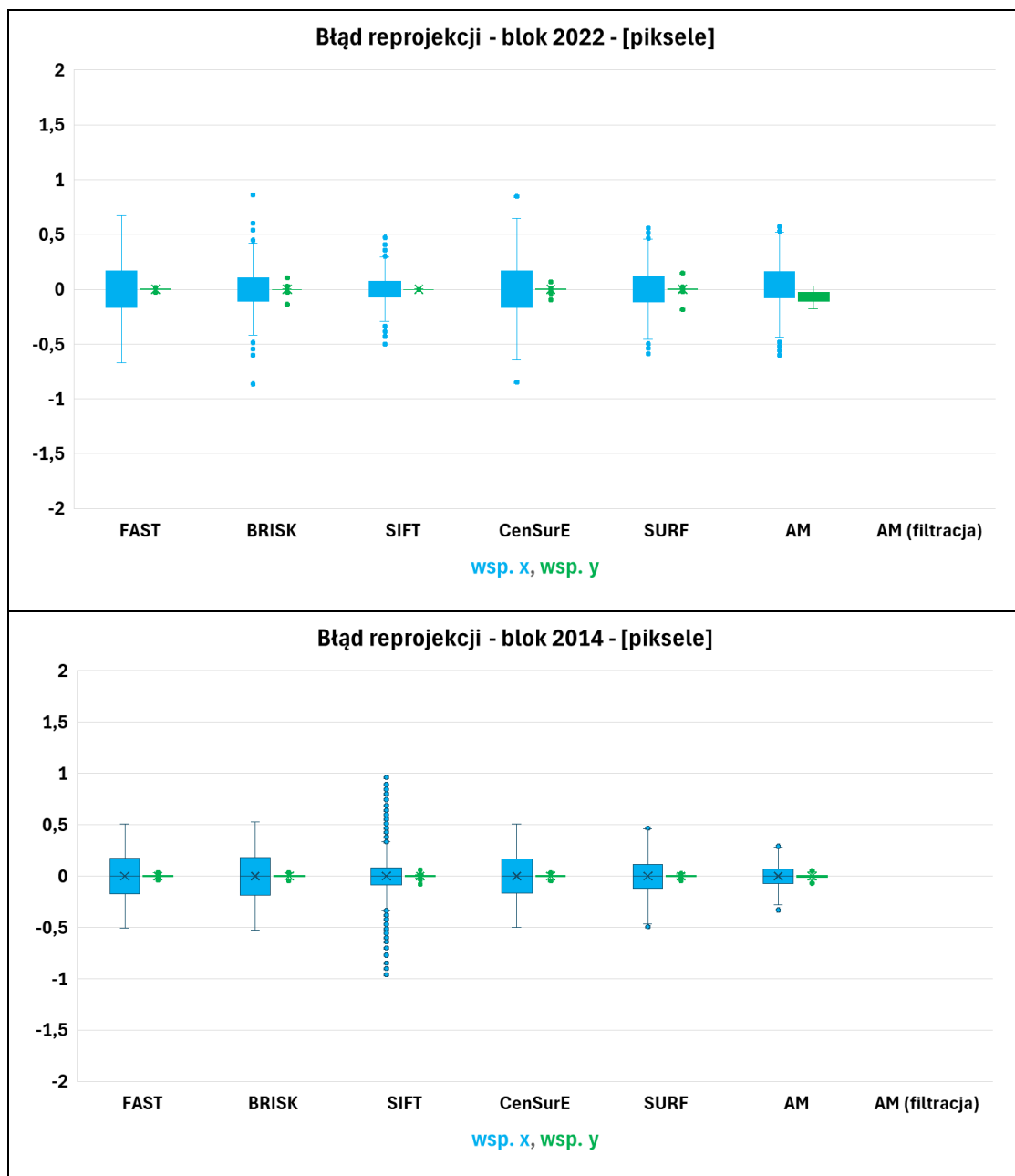
standardowej funkcji dostępnej w programie AM, takie podejście minimalizuje ryzyko wpływu punktów odstających na geometrię bloków fotogrametrycznych.

W celu kompletnej analizy statystycznej orientacji wzajemnej zdjęć w poszczególnych blokach wyniki błędów orientacji zostały przedstawione w formie wykresów pudełkowych (tab. 5.1) na rysunkach 7.8 i 7.9. Z uwagi na różną jakość zdjęć analogowych i cyfrowych rozkłady te należy rozpatrywać oddzielnie, tj. zdjęcia cyfrowe z 2022 i 2014 roku oraz zdjęcia analogowe z roku 1976, 1986 i 1994. Ze względu na istotnie niższe wartości błędu reprojekcji współrzędnej y w porównaniu ze współrzędną x (rys. 7.8, 7.9), dalsza analiza błędów orientacji zdjęć odnosi się głównie do składowej x . Zaobserwowaną różnicę pomiędzy wartościami błędów współrzędnych x i y można wyjaśnić geometrią rejestracji zdjęć lotniczych. W klasycznym nalocie baza zdjęć przebiega niemal w całości wzdłuż kierunku lotu samolotu, który odpowiada osi x w układzie obrazowym. Oznacza to, że przesunięcie stanowiska obrazowania między kolejnymi zdjęciami – a zatem obserwowana paralaksa – ma dominującą składową poziomą. W rezultacie w kierunku x występuje znacznie większa paralaksa niż w kierunku y , gdzie jej wartość jest bliska zeru. Większa paralaksa w osi x oznacza, że punkty na dwóch zdjęciach są względem siebie bardziej przesunięte w poziomie (Fraser, C. S., 1997; Mikhail, E. M., Bethel, J. S. i McGlone, J. C., 2001). Przez to algorytm ma większy zakres możliwych pozycji, w których może szukać dopasowania danego punktu – dlatego łatwiej o większy błąd. Nawet niewielka pomyłka w dopasowaniu mocniej wpływa na wynik właśnie w kierunku x . W efekcie podczas orientacji wzajemnej otrzymuje się większy błąd RMS dla współrzędnej x niż dla współrzędnej y .

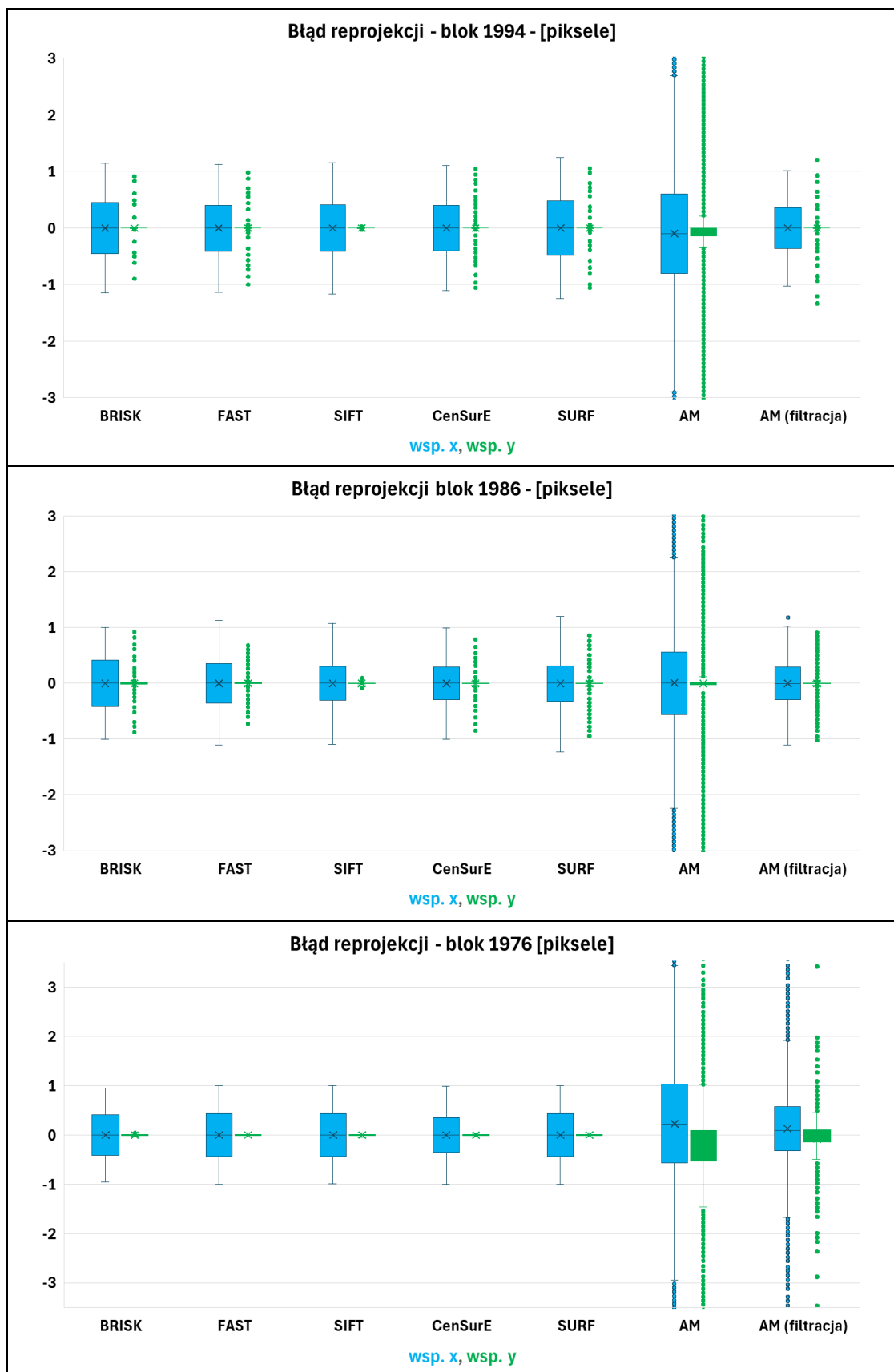
Przeprowadzona analiza statystyczna (rys. 7.8 i 7.9) rozkładu błędów reprojekcji współrzędnych x i y (obejmująca pięć okresów: 2022, 2014, 1994, 1986 i 1976), wykazała, że najmniejsze błędy oraz najlepszą spójność geometryczną bloku uzyskano dla punktów wykrytych za pomocą detektora SIFT na zdjęciach cyfrowych z roku 2022. Dla tego bloku dolny i górny kwartył mieścił się w granicy ± 0.1 pix, a dolny i górny wąs wynosił ± 0.2 pix dla współrzędnej x .

W bloku z 2014 r. najmniejsze błędy uzyskał detektor SURF, z dolnym i górnym kwartylem wynoszącym $\pm 0,1$ pix oraz dolnym i górnym wąsem ± 0.3 pix dla współrzędnej x . W przypadku bloku z 1994 r. najlepsze wyniki osiągnęły detektory BRISK, FAST, SIFT oraz CenSurE, dla

których dolny i górny kwartył mieścił się w granicy ± 0.4 pix, a dolny i górny wąs wynosił ± 1.2 pix. Dla bloku z 1986 r. najmniejsze błędy odnotowano dla detektora CenSurE, przy dolnym i górnym kwartylu wynoszącym ± 0.3 pix oraz dolnym i górnym wąsie ± 1.0 pix. W bloku z 1976 r. najmniejsze błędy reprojektacji uzyskały detektory SIFT, CenSurE oraz SURF, gdzie dolny i górny kwartył wynosiły ± 0.4 pix, a dolny i górny wąs ± 1.0 pix dla współrzędnej x. Analiza błędów reprojektacji (x, y) orientacji wzajemnej w blokach zdjęć lotniczych wykonanych w różnych latach (rys. 7.8, 7.9) wskazuje, że jakość detekcji punktów kluczowych zależy zarówno od zastosowanego algorytmu, jak i od rodzaju zdjęć (cyfrowe, analogowe). Najmniejsze błędy uzyskano dla Detektorów SURF, SIFT oraz CenSurE, co świadczy o ich wysokiej skuteczności. Dla bloku 2014 r. najlepszą dokładnością orientacji uzyskał detektor SURF, natomiast w blokach z lat 1976–1994 dobre wyniki orientacji uzyskano również dla detektorów CenSurE i SIFT. Sugeruje to, że detektory te można wykorzystać do orientacji zdjęć lotniczych archiwalnych zarówno cyfrowych jak analogowych.



Rys. 7.8 Rozkład błędów reprojektacji (x, y) orientacji bloku z roku 2022 i 2014 utworzonych ze zdjęć cyfrowych w zależności od użytego detektora (wykresy pudełkowe)



Rys. 7.9 Rozkład błędów reprojektacji (x, y) orientacji bloku z roku 1994, 1986 i 1976 utworzonych ze zdjęć analogowych w zależności od użytego detektora (wykresy pudełkowe)

Wartości błędów uzyskanych dla algorytmu AM bez filtracji w przypadku zdjęć analogowych (rys. 7.8) były istotnie większe w porównaniu z innymi detektorami, co wyklucza jego zastosowanie w końcowej orientacji bloków. Największe odchylenia odnotowano w bloku z 1976 roku, gdzie błędy dla współrzędnej x były dwukrotnie (kwartyle) i trzykrotnie (wąsy) większe niż w przypadku pozostałych metod, a dla współrzędnej y nawet kilkudziesięciokrotnie większe; dodatkowo przesunięcie kwartyłu wskazuje na występowanie błędów systematycznych, prawdopodobnie spowodowanych niską jakością radiometryczną zdjęć. W przypadku bloków utworzonych ze zdjęć cyfrowych (2014 i 2022) rozkład błędów był znacznie korzystniejszy, a pojedyncze błędy grube, stanowiące poniżej 0.5% zbioru, nie wpływały istotnie na geometrię bloków.

Odnosząc się do algorytmu AM po filtracji stosowanego do bloków utworzonych ze zdjęć analogowych (1994, 1986 i 1976) uzyskano podobne wartości błędów jak dla algorytmów HC wykorzystane w badaniach. Niemniej jednak, zaproponowana metoda, uwzględniająca wybrane detektory HC, charakteryzuje się większą dokładnością, ponieważ wykryte punkty wiążące są pozbawione wartości odstających, co znacząco poprawia dokładność orientacji bloków.

Wnioski oceny dokładności orientacji wzajemnej zdjęć w bloku (wykres pudełkowy)

1. Dokładność wyznaczania punktów wiążących przez zastosowane detektory nie zależy od wielkości GSD zdjęć lotniczych, a wybór detektora nie jest uwarunkowany rozdzielczością terenową obrazu.
2. Zdjęcia cyfrowe (np. z lat 2022 i 2014) wykazują wysoką jakość radiometryczną oraz najmniejsze błędy reprojekcji. Zaleca się ich wykorzystanie jako podstawę do sekwencyjnej orientacji bloków fotogrametrycznych (rozdział 9). Detektory takie jak SIFT czy SURF dla tych zdjęć zapewniają wysoką dokładność wyznaczenia punktów wiążących.
3. Najmniejsze błędy reprojekcji (najlepszą spójność geometryczną bloku) uzyskano dla punktów wykrytych detektorem SIFT na zdjęciach cyfrowych z 2022 roku.
4. Algorytm AM bez filtracji generuje istotnie większe błędy dla zdjęć analogowych. Takie błędy są zbyt duże, by traktować algorytm AM bez filtracji jako rozwiązanie końcowe.

5. Orientację zdjęć lotniczych cyfrowych w środowisku programowym Agisoft Metashape można przeprowadzić bez konieczności stosowania dodatkowej filtracji punktów wiążących, przy zachowaniu weryfikacji błędów grubych w pomiarze.
6. Ze względu na lepszą radiometrię zdjęć cyfrowych blok utworzony z tych zdjęć charakteryzuje się około 2-3 krotnie mniejszym błędem reprojekcji.
7. Algorytm AM bez filtracji nie jest zalecany, zwłaszcza dla zdjęć o słabej radiometrii (blok 1976), ze względu na wysokie wartości błędów.
8. Algorytm AM po filtracji (AM_filtracja) uzyskał wyniki porównywalne z wynikami detektorów HC (BRISK, FAST, SIFT, CenSurE, SURF) dla bloków opartych na zdjęciach analogowych.
9. Zaproponowana autorska metoda pomiaru i filtracji punktów wiążących z wykorzystaniem wybranych detektorów HC charakteryzuje się wyższą niezawodnością, ponieważ eliminuje ona wartości odstające, co istotnie poprawia dokładność orientacji bloków.

7.5 Wnioski i podsumowanie

W niniejszym rozdziale przedstawiono ogólne wnioski dotyczące budowy bloków fotogrametrycznych z zastosowaniem metody SfM (*Structure-from-Motion*). Omówiono czas detekcji punktów wiążących, liczbę i rozkład punktów w bloku, dokładność pomiaru punktów na parach zdjęć oraz dokładność orientacji wzajemnej zdjęć w bloku.

Czas detekcji punktów wiążących

1. Wydajność detektorów: Detektor SIFT charakteryzuje się najkrótszym czasem pomiaru punktów wiążących, podczas gdy detektor FAST działa najwolniej. Efektywność SIFT była szczególnie widoczna dla zdjęć analogowych.
2. Różnice w GSD zdjęć w większości przypadków nie wpływały na czas pomiaru, co podkreśla stabilność badanych algorytmów.
3. Różnice w rodzajach zdjęć: Detektory SIFT i CenSurE działały wolniej w przypadku zdjęć cyfrowych niż analogowych, co wynikało z większej liczby wykrywanych punktów w obrazach cyfrowych, charakteryzujących się wyższą rozdzielczością radiometryczną.

4. Stabilność detektorów: Algorytmy BRISK, FAST i SURF wykazały wysoką niezmienną wydajności, ponieważ ich czas detekcji punktów był niezależny od rodzaju zdjęć i ich radiometrii. SURF okazał się najbardziej niezmiennym algorytmem.
5. Ekonomia procesu: Wybór detektora ma kluczowe znaczenie przy pracy z dużymi blokami danych, ponieważ wpływa na czas obliczeń oraz koszty procesu fotogrametrycznego.
6. Kompleksowe podejście: Ważne jest uwzględnienie całkowitego czasu obliczeniowego dla wszystkich punktów, a nie tylko pojedynczego punktu, co ma kluczowe znaczenie dla efektywności procesów fotogrametrycznych.

Detektor SIFT, choć bardziej złożony, okazał się najefektywniejszy czasowo, szczególnie w analizach dużych zbiorów danych. Detektor SURF wyróżniał się przewidywalnością działania, co czyni go odpowiednim dla różnorodnych parametrów zdjęć. Z kolei detektor FAST, choć szybki w niskiej rozdzielczości, tracił na wydajności przy większej szczegółowości obrazu. Wybór detektora powinien być oparty na specyfice zadania, uwzględniając zarówno liczbę punktów, jak i wymagania obliczeniowe procesu.

Liczba i rozmieszczenie punktów wiążących w bloku fotogrametrycznym

1. Wszystkie analizowane algorytmy detekcji HC (BRISK, FAST, SIFT, CenSurE, SURF) wykrywają wystarczającą liczbę punktów wiążących do budowy bloków fotogrametrycznych w latach 2022, 2014, 1994, 1986 i 1976, niezależnie od rodzaju zdjęć (cyfrowe vs. analogowe).
2. Algorytmy powierzchniowe są preferowane w terenach o urozmaiconym pokryciu oraz w przypadku zdjęć cyfrowych.
3. Detektor SURF jest najskuteczniejszy pod względem liczby wykrytych punktów dla zdjęć analogowych o słabej radiometrii.
4. Detektory punktowe są mniej efektywne w terenach niezurbanizowanych, co ogranicza ich zastosowanie w takich obszarach.
5. Detektor w oprogramowaniu AM wykazał najwyższą jakość pod względem równomierności rozmieszczenia punktów zarówno przed, jak i po filtracji, jednak jakość tych punktów jest niska, co wymaga zastosowania dodatkowego procesu filtracji przed ich wykorzystaniem, zwłaszcza dla zdjęć analogowych.

6. Najlepsze rozmieszczenie punktów wyznaczonych przez algorytmy HC w analizowanych blokach uzyskano dla algorytmów takich jak: FAST, CenSurE i SURF.
7. Zdjęcia cyfrowe wyraźnie przewyższają analogowe pod względem liczby wyznaczonych punktów wiążących, co wynika z lepszej ich jakości radiometrycznej.

W przypadku analiz zdjęć o słabej jakości radiometrycznej (np. archiwalne zdjęcia analogowe), zaleca się stosowanie detektora SURF. Na terenach nieurbanizowanych rekomendowane są detektory powierzchniowe (SIFT, CenSurE, SURF) z uwagi na większą liczbę wykrywanych punktów. W mniej zróżnicowanych pokryciach terenu można rozważyć także algorytmy punktowe (FAST, BRISK). Przy przetwarzaniu zdjęć można korzystać z algorytmu AM ze względu na jego wysoką skuteczność zarówno w liczbie wyznaczanych punktów, jak i ich rozmieszczeniu. Należy przy tym pamiętać, że stosowanie tego algorytmu dla zdjęć analogowych wymaga filtracji punktów. Obszary wodne oraz tereny nieużytków stanowią największe wyzwanie dla wszystkich detektorów.

Dokładność pomiaru punktów wiążących na parach zdjęć

1. Wyznaczenie punktów wiążących w oparciu o algorytmy HC jest kluczowe dla dokładności dalszych etapów przetwarzania, takich jak proces wyrównania zdjęć w bloku metodą niezależnych wiązek, a także jakość generowanych chmur punktów. Jakość tych punktów, mierzona błędem reprojekcji, ma istotny wpływ na końcowy wynik. Mimo różnic technologicznych (zdjęcia cyfrowe i analogowe) oraz warunków nalotu, uzyskane wyniki wskazują na stabilność pomiarów.
2. Różnice w wartościach RMSE między detektorami (FAST, SURF, BRISK) są marginalne i mogą być wynikiem budowy bloku (pokrycie podłużne, poprzeczne), a nie samego działania algorytmu.
3. Dla zdjęć analogowych najlepiej sprawdza się detektor CenSurE, a dla zdjęć cyfrowych – SIFT.
4. Wartości RMSE wyznaczenia punktów wiążących przez detektory na zdjęciach cyfrowych były 2 razy mniejsze niż dla zdjęć analogowych.
5. Brak wyraźnej zależności między wielkością GSD a dokładnością pomiaru punktów wskazuje, że poziom dokładności pozostaje stabilny niezależnie od rozdzielczości przestrzennej zdjęć.

6. Analiza błędów w procesie wyrównania zdjęć w bloku metodą niezależnych wiązek potwierdza brak błędów grubych w danych, co świadczy o poprawności wyznaczenia punktów wiążących.

Odnosząc się do dokładności pomiaru punktów na parach zdjęć w bloku, detektor SIFT okazał się najbardziej efektywny w przypadku zdjęć cyfrowych, podczas gdy detektor CenSurE sprawdził się lepiej w przypadku zdjęć analogowych. Niskie wartości błędów SMAD i SIGMA potwierdzają stabilność i wiarygodność pomiarów, co jest kluczowe dla budowy dokładnych modeli fotogrametrycznych. Przeprowadzona analiza potwierdziła zasadność przyjętej metodologii oraz skuteczność zastosowanych detektorów, umożliwiając uzyskanie dokładnych wyników pomiaru punktów wiążących niezależnie od charakteru materiału fotograficznego (analogowego czy cyfrowego).

Budowa bloków fotogrametrycznych w lokalnym układzie odniesienia

1. Zarówno detektory punktowe, jak i powierzchniowe zapewniają dobrą dokładność orientacji zdjęć w bloku fotogrametrycznym, niezależnie od rodzaju zdjęć (analogowe lub cyfrowe).
2. Dokładność orientacji bloków budowanych na podstawie zdjęć cyfrowych była dwukrotnie wyższa od dokładności orientacji bloków budowanych ze zdjęć analogowych, co jest wynikiem ich lepszej jakości radiometrycznej.
3. Najmniejsze wartości RMSE orientacji wzajemnej uzyskano dla detektora SIFT zarówno dla zdjęć cyfrowych, jak i analogowych.
4. Największe błędy reprojektacji stwierdzono w przypadku detektorów BRISK (dla bloków z 1986 i 2014) oraz SURF (dla bloku z 1994 roku), co jest zgodne z zasadą ich działania.
5. Zastosowanie filtracji punktów odstających wyznaczonych przez algorytm programu AM umożliwiło redukcję błędów reprojektacji zdjęć analogowych średnio czterokrotnie, do poziomu zbliżonego do wyników uzyskiwanych dla algorytmów HC na zdjęciach analogowych.
6. Zaproponowana autorska metoda pomiaru i filtracji punktów przy użyciu detektorów HC (BRISK, FAST, SIFT, CenSurE, SURF) charakteryzuje się wyższą niezawodnością niż metoda dostępna w oprogramowaniu AM z własną funkcją filtracji.

Dla zdjęć cyfrowych zaleca się wykorzystanie detektorów powierzchniowych SIFT i SURF, które zapewniają najlepszą dokładność w procesie wykrywania punktów wiążących oraz orientacji bloków. W przypadku zdjęć analogowych, dla których błędy reprojekcji są większe, korzystne jest zastosowanie detektora SIFT oraz dodatkowej filtracji punktów odstających. Zaleca się stosowanie autorskiej metody filtracji punktów wiążących, która efektywnie eliminuje błędy wynikające z wartości odstających, zwiększając dokładność orientacji bloków fotogrametrycznych.

Na podstawie przeprowadzonej oceny (tab.7.1) przydatności algorytmów HC detekcji punktów wiążących dla wszystkich okresów czasowych (zdjęcia z roku: 2022, 2014, 1994, 1986, 1976) najlepszy wynik (I miejsce) uzyskał detektor powierzchniowy SIFT z sumą punktów 38/40. Detektor punktowy CenSurE uzyskały równie wysoki wynik zajmując II miejsce. Pozostałe detektory uplasowały się na trzeciej pozycji. Każde z 4 kryterium było oceniane w skali czterostopniowej, gdzie: 2 oznaczało ocenę negatywną, 3 – dostateczną, 4 – dobrą, a 5 – bardzo dobrą. Do tworzenia bloków fotogrametrycznych rekomendowane są wszystkie oceniane detektory, jednak z wyraźnym wskazaniem na detektor SIFT jako najbardziej efektywny.

Tab. 7.1 Podsumowanie oceny kryteriów (skala oceny: 2 oznaczało ocenę negatywną, 3 - dostateczną, 4 - dobrą, a 5 - bardzo dobrą) przydatności algorytmów detekcji punktów dla wszystkich okresów czasowych

Kryteria oceny	Kryteria oceny	BRISK	FAST	SIFT	CenSurE	SURF
Czas detekcji	Zdjęcia cyfrowe	4	3	5	4	4
	Zdjęcia analogowe	4	3	5	4	4
Liczba punktów i ich rozmieszczenie	Zdjęcia cyfrowe	5	5	4	4	4
	Zdjęcia analogowe	4	4	5	5	5
Dokładność pomiaru punktu	Zdjęcia cyfrowe	4	4	5	5	5
	Zdjęcia analogowe	4	4	4	5	4
Orientacja bloku	Zdjęcia cyfrowe	3	4	5	4	3
	Zdjęcia analogowe	3	4	5	4	3
Suma		31	31	38	35	32
Ocena końcowa		III	III	I	II	III

Ze względu na lepszą rozdzielczość i jakość radiometryczną cyfrowych zdjęć lotniczych, w porównaniu ze zdjęciami analogowymi, efektywność detekcji algorytmów HC na obrazach cyfrowych jest istotnie wyższa zarówno pod względem liczby wykrywanych punktów, jak i dokładności ich wyznaczenia. Jednak dzięki zastosowaniu odpowiednich technik filtracji i zaawansowanych detektorów (np. SIFT, CenSurE), zdjęcia analogowe mogą również zapewnić

dobrą jakość wyników, choć kosztem większego nakładu pracy. Algorytmy punktowe i powierzchniowe wykazują dobrą skuteczność, co daje elastyczność wyboru narzędzi w zależności od warunków i potrzeb projektu.

Zastosowanie metody SfM w orientacji bloków fotogrametrycznych okazało się skuteczniejsze od tradycyjnych technik stosowanych w fotogrametrii lotniczej. Podejście SfM oferuje wyższą dokładność, większą automatyzację oraz lepszą adaptację do zdjęć o zróżnicowanej jakości, co czyni je praktyczną i nowoczesną alternatywą dla klasycznych metod aerotriangulacji. Uzyskane błędy reprojektacji oraz stabilność detekcji punktów kluczowych potwierdzają, że podejście SfM może stanowić efektywną i konkurencyjną alternatywę dla klasycznych metod aerotriangulacji w kontekście cyfrowej fotogrametrii lotniczej.

8. Produkty fotogrametryczne wykorzystane do sekwencyjnej orientacji bloków zdjęć

W proponowanej metodyce sekwencyjnej orientacji bloków fotogrametrycznych, podstawę ich orientacji stanowiły dwa pośrednie produkty fotogrametryczne tj. true-ortofotomapy i NMPT (Numeryczne Modele Pokrycia Terenu). Zasięgi tych produktów odpowiadały zasięgom utworzonych bloków. Wygenerowane produkty fotogrametryczne zostały utworzone w oparciu o zbudowane wcześniej bloki fotogrametryczne (patrz rozdział 5.1).

W celu wykonania sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych niezbędna jest identyfikacja punktów homologicznych (przejęciowych, tj. tych samych elementów terenowych) pomiędzy blokami, które stanowią podstawę transformacji 3D (patrz podrozdział 5.3). Umożliwia to, wzajemną orientację jednego bloku względem drugiego.

Zazwyczaj pomiar punktów przejęciowych pomiędzy dwoma blokami fotogrametrycznymi można wykonać na parach zdjęć pochodzących z tych bloków, pod warunkiem występowania między nimi wspólnego pokrycia. W takim przypadku mowa jest o pomiarze lokalnym, który odnosi się do konkretnej pary zdjęć. Ze względu na różnice w skalach zdjęć, formatach obrazów oraz stopnia wzajemnego pokrycia zarówno w obrębie pojedynczych bloków, jak i między nimi, obszar wspólnego dopasowania zdjęć jest zmienny. Przykładem tego jest sekwencja bloków z lat 1994 i 2022, gdzie pokrycie pomiędzy dopasowywanymi parami zdjęć wynosi 100%. W takim przypadku możliwe jest wyznaczenie dużej liczby punktów homologicznych, pozwalających na przeprowadzenie sekwencyjnej orientacji bloków. Natomiast w przypadku sekwencji bloków 1994 i 1986, w której pokrycie wynosi jedynie 20–30%, liczba możliwych do zidentyfikowania punktów homologicznych jest niewystarczająca do realizacji tego procesu.

Dodatkowym wyzwaniem jest duża liczba par zdjęć utworzonych pomiędzy blokiem referencyjnym i orientowanym, co znacząco zwiększa liczbę obliczeń wykonywanych przez algorytm w procesie wyszukiwania punktów homologicznych, prowadząc do wydłużenia czasu przetwarzania.

Ze względu na ww. trudności w wyznaczaniu punktów przejęciowych na podstawie pojedynczych par zdjęć, zastosowano alternatywne podejście metodyczne w sekwencyjnej orientacji bloków. Zamiast klasycznych par zdjęć wykorzystano pary utworzonych true-

ortofotomap – wykorzystanie podejścia globalnego na rzecz lokalnego dopasowania parami poszczególnych zdjęć w bloku. Takie rozwiązanie pozwoliło na eliminację wielu problemów związanych z pomiarem punktów na parach zdjęć (patrz wyżej), co doprowadziło do globalnego zwiększenia obszaru wspólnego dopasowania bloków, wzrostu liczby punktów przejściowych oraz poprawy ich rozmieszczenia.

W podejściu globalnym wykorzystywane są jedynie dwie true-ortofotomapy (po jednej dla każdego bloku), w przeciwieństwie do podejścia lokalnego, które wymaga analizy dużej liczby par zdjęć. W konsekwencji wymagana liczba punktów przejściowych do przeprowadzenia transformacji 3D i orientacji bloków w podejściu globalnym jest znacznie mniejsza niż w podejściu lokalnym, co przekłada się na wyższą efektywność i skuteczność tej metody. Ponadto zastosowanie true-ortofotomap eliminuje problemy związane z trudnościami w detekcji punktów w obszarach o zmiennym pokryciu (leśne, rolnicze, nieużytki, urbanizacyjne), a także na zdjęciach charakteryzujących się niską jakością radiometryczną. W podejściu lokalnym opartym na analizie pojedynczych zdjęć (dwie pary zdjęć) często można natrafić na przypadek, że dla danej pary zdjęć brak jest punktów (lub jest ich za mało). Nie ma tego problemu w podejściu globalnym, gdzie true-ortofotomapa pokrywa cały obszar opracowania.

W zaproponowanym rozwiązaniu orientacja bloku zdjęć referencyjnych przenoszona jest do bloku zdjęć orientowanych (zob. rozdział 9). Zastosowanie metody globalnej, obejmującej szeroki obszar opracowania, umożliwiło uzyskanie wystarczającej liczby punktów wiążących oraz zapewniło ich równomierne rozmieszczenie. Pozwoliło to na przeprowadzenie transformacji 3D pomiędzy blokami, a tym samym zapewniło poprawną orientacji zdjęć w blokach.

Sekwencyjna orientacja bloków fotogrametrycznych (rozdział 5.3) została wykonana na podstawie relacji geometrycznych punktów przejściowych, wykrytych między true-ortofotomapami poszczególnych epok. Po wyznaczeniu punktów przejściowych (x , y) pomiędzy true-ortofotomapą referencyjną, a orientowaną na podstawie NMPT wyznaczono współrzędne (z). Stanowiło to podstawę transformacji 3D bloku orientowanego do bloku referencyjnego.

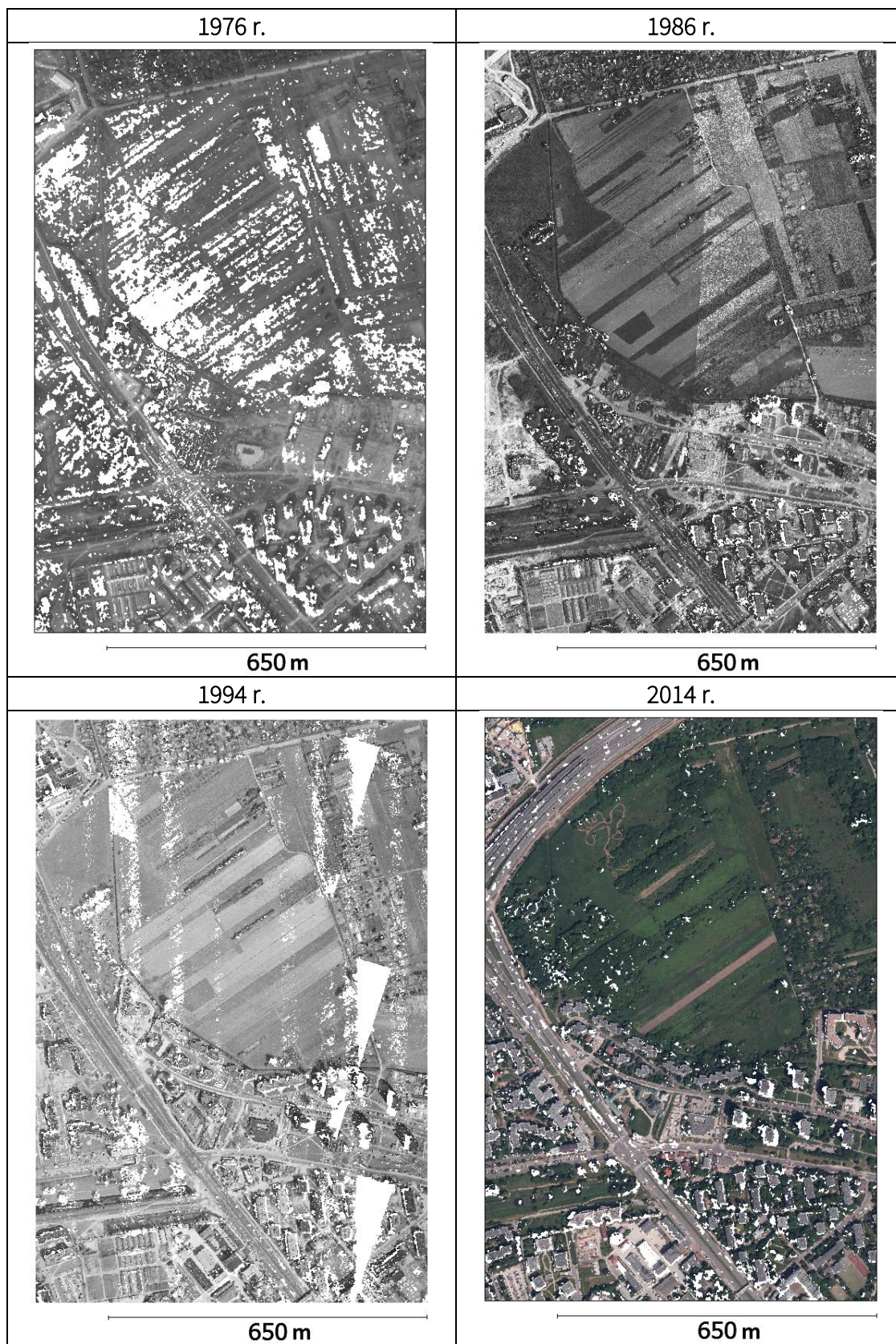
Zgodnie z rekomendacjami (podrozdział 7.5) podstawą tworzenia produktów fotogrametrycznych takich jak true-ortofotomapy i NMPT, były bloki fotogrametryczne budowane w oparciu o punkty wiążące wyznaczone przez detektor SIFT dla wszystkich epok. Utworzone true-ortofotomapy i NMPT odnosiły się do układów lokalnych poszczególnych bloków wyrażonych w skali zdjęcia.

NMPT

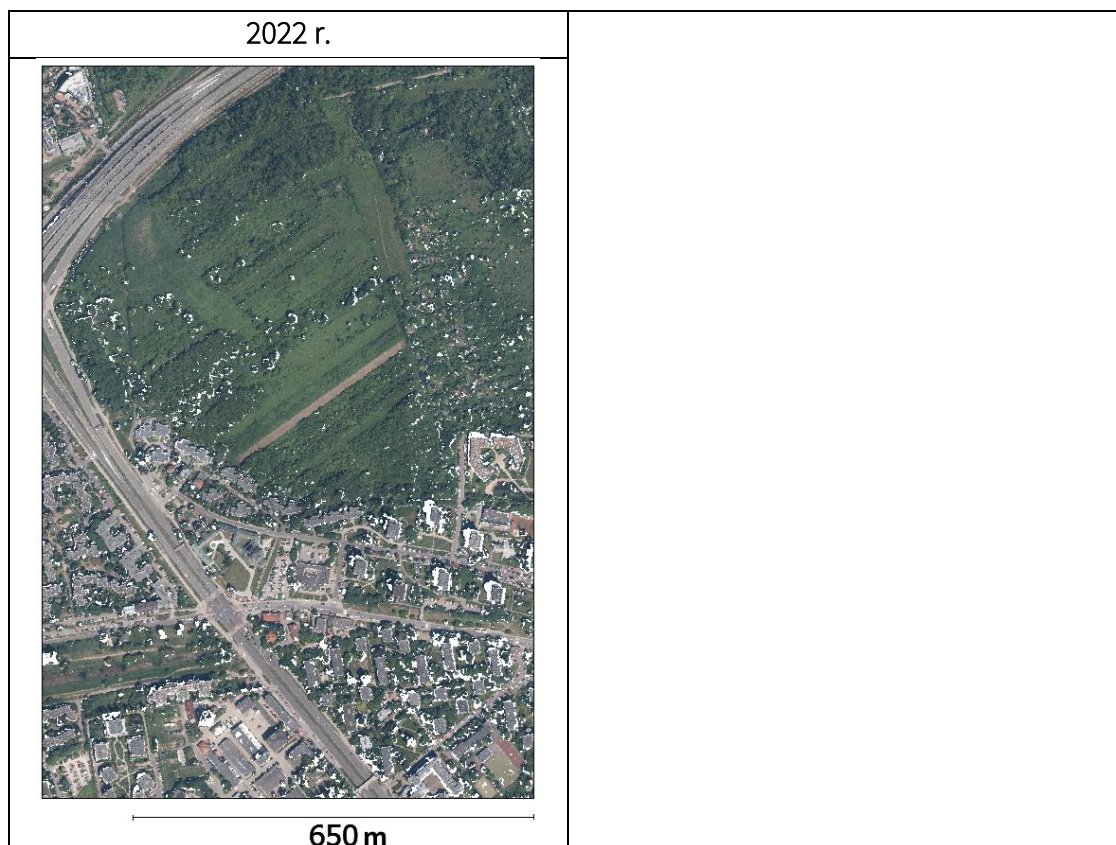
Dla bloku fotogrametrycznego z każdej epoki wygenerowano gęste chmury punktów w oprogramowaniu Agisoft Metashape (AM), przy czym parametr jakościowy (rozumiany jako poziom piramidy; *Quality*) spośród pięciu dostępnych poziomów (*Ultra-High*, *High*, *Medium*, *Low*, *Lowest*) został ustawiony na *High*. Wartość *Ultra-High* oznacza generowanie gęstej chmury punktów na podstawie zdjęć w pełnej rozdzielczości (najniższy poziom piramidy obrazu), natomiast każdy kolejny poziom oznacza czterokrotne zmniejszenie rozdzielczości obrazu poprzez usunięcie co drugiego piksela w rzędzie i kolumnie, z jednoczesnym zastosowaniem filtru Gaussa. Poziom *High* wybrano, ponieważ wprowadzenie parametru *Ultra-High* nie wpływało na zwiększenie dokładności wyznaczenia punktów w chmurze, pomimo jej większej gęstości, a czas ich wyznaczenia był znacznie dłuższy. Dodatkowo zauważalne były braki w chmurach punktów (tzw. „dziury”) co wpływało na poprawność generowania NMPT.

Celem uniknięcia błędnego klasyfikowania niektórych elementów pokrycia terenu jako wartości odstających, w procesie generowania chmury punktów zastosowano filtrację. Z dostępnych czterech poziomów filtracji (*Disabled*, *Mild*, *Moderate*, *Aggressive*) wybrano parametr *Mild* dla generowanej chmury punktów. Poziom filtrowania *Mild* jest wykorzystywany, jeżeli na zdjęciach istnieją przestrzennie wyróżnione drobne szczegóły, a także elementy terenowe o słabej teksturze. Jest on najczęściej stosowanym filtrem dla zdjęć lotniczych (Agisoft Helpdesk Portal, 2025). Z tego względu, że przetwarzane archiwalne zdjęcia charakteryzują się drobnymi szczegółami przestrzennymi (pokrycia terenu roślinnością i zabudową) jako tryb filtrowania zastosowano *Mild*. Z kolei filtracja na poziomie *Aggressive* stosowana jest w sytuacjach, gdy obszar pomiarowy nie zawiera istotnych drobnych szczegółów. W tym trybie usuwane są punkty, które nie pasują do ogólnej geometrii terenu lub obiektów (Agisoft Helpdesk Portal, 2025). Tryb filtrowania *Moderate* o umiarkowanym poziomie filtrowania zapewnia wyniki pomiędzy podejściem łagodnym a agresywnym. Do generowania

chmur punktów wykorzystano mapy głębi, które w fotogrametrii powiązane są z pojęciem paralaksy, przy czym wysokość punktów chmury wyznaczana była na podstawie dopasowania homologicznych punktów na zdjęciach z zastosowaniem metod *Multi-View Stereo* (MVS) oraz triangulacji. Dla wszystkich utworzonych chmur punktów (2022, 2014, 1994, 1986, 1976) zastosowano następujące parametry: jakość generowania chmur punktów ustawioną na poziom *High*, filtrację w trybie *Mild* oraz tworzenie mapy głębi z parametrem jw.



Rys. 8.1 Zestawienie wygenerowanych chmur punktów dla bloku 1976, 1986, 1994, 2014
(fragment obszaru testowego)



Rys. 8.2 Zestawienie wygenerowanych chmur punktów dla bloku 2022 (fragment obszaru testowego)

Wygenerowane dla każdej epoki za pomocą podejścia MVS (*Multi-View Stereo*) gęste chmury punktów stanowiły podstawę utworzenia NMPT, które wykorzystane zostały do wyznaczenia współrzędnej (z) punktów przejściowych stanowiących podstawę sekwencyjnej orientacji bloków fotogrametrycznych (rys. 8.1 i 8.2). W celu ułatwienia interpolacji wysokości punktów przejściowych modele NMPT tworzone były w środowisku oprogramowania ArcGIS w formacie *grid* za pomocą funkcji „*las to raster*”, gdzie wielkość oczka siatki modelu odpowiadała połowie wartości GSD zdjęć.

W celu szczegółowej analizy pod względem jakościowym wygenerowanych chmur punktów na rysunku 8.1 przedstawiono powiększony fragment obszaru testowego dla wszystkich chmur. Zakres wysokości punktów w każdej z chmur dla poszczególnych okresów mieścił się w przedziale od 40 do 130 m. Analiza jakości chmur punktów, przeprowadzona pod kątem ich kompletności, określonej na podstawie występowania tzw. luk (obszarów bez pomierzonych punktów) oraz regionów o niskiej gęstości, wykazała, że największa liczba tych obszarów wystąpiła w chmurze z 1976 roku. Chmura ta została wygenerowana na podstawie zdjęć

o niskiej jakości radiometrycznej, co wpłynęło na ograniczoną skuteczność pomiaru punktów w tej chmurze (rys. 8.1). Regiony o niskiej gęstości koncentrowały się głównie na powierzchniach użytków gruntowych natomiast luki w bezpośrednim sąsiedztwie wysokich budynków, co było efektem występowania stref cienia. Podobne regiony o niskiej gęstości punktów oraz obszary bez pomierzonych punktów w sąsiedztwie zabudowy zaobserwowano również w chmurze punktów wygenerowanej na podstawie zdjęć z 1986 roku, jednak ich liczba była znacznie mniejsza w porównaniu do chmury z 1976 roku. Dodatkowo w chmurze z roku 1986, z uwagi na ograniczone pokrycie poprzeczne między zdjęciami w sąsiednich szeregach bloku, wynikające z odchylenia trajektorii lotu samolotu od osi szeregu (rys. 6.2, rozdział 6), w niektórych obszarach modelu nie było możliwe dokonanie pomiaru punktów, co skutkowało powstaniem powierzchni o kształcie trójkąta pozbawionych punktów (rys. 8.1). W chmurach punktów utworzonych ze zdjęć z lat 2014 i 2022 liczba obszarów bez pomierzonych punktów była znacznie mniejsza i nie miała istotnego wpływu na jakość chmur.

Tab. 8.1 Zestawienie gęstości generowanych chmur punktów dla bloku 1976, 1986, 1994, 2014, 2022

Chmura punktów	1976 r.	1986 r.	1994 r.	2014 r.	2022 r.
Gęstość Punkt/1 m ²	4	160	50	6	14

W tabeli 8.1 przedstawiono wartości gęstości poszczególnych chmur punktów, wyrażone jako liczba punktów przypadających na jednostkę powierzchni 1 m². Najmniejszą gęstość punktów odnotowano w chmurach z lat 1976 i 2014, natomiast największą w chmurze z roku 1986. Wysoka gęstość chmury punktów z 1986 roku (160 punktów/m²) wynika z dużej liczby zdjęć o wspólnym pokryciu podłużnym i poprzecznym, co wynikało ze skali tych zdjęć w stosunku do obszaru opracowania. Z kolei czterdziestokrotnie i trzydziestokrotnie mniejsza gęstość chmur z lat 1976 i 2014 jest konsekwencją mniejszej liczby zdjęć o wspólnym pokryciu podłużnym, co również wynikało ze skali tych zdjęć w stosunku do obszaru opracowania (patrz rozdział 6).

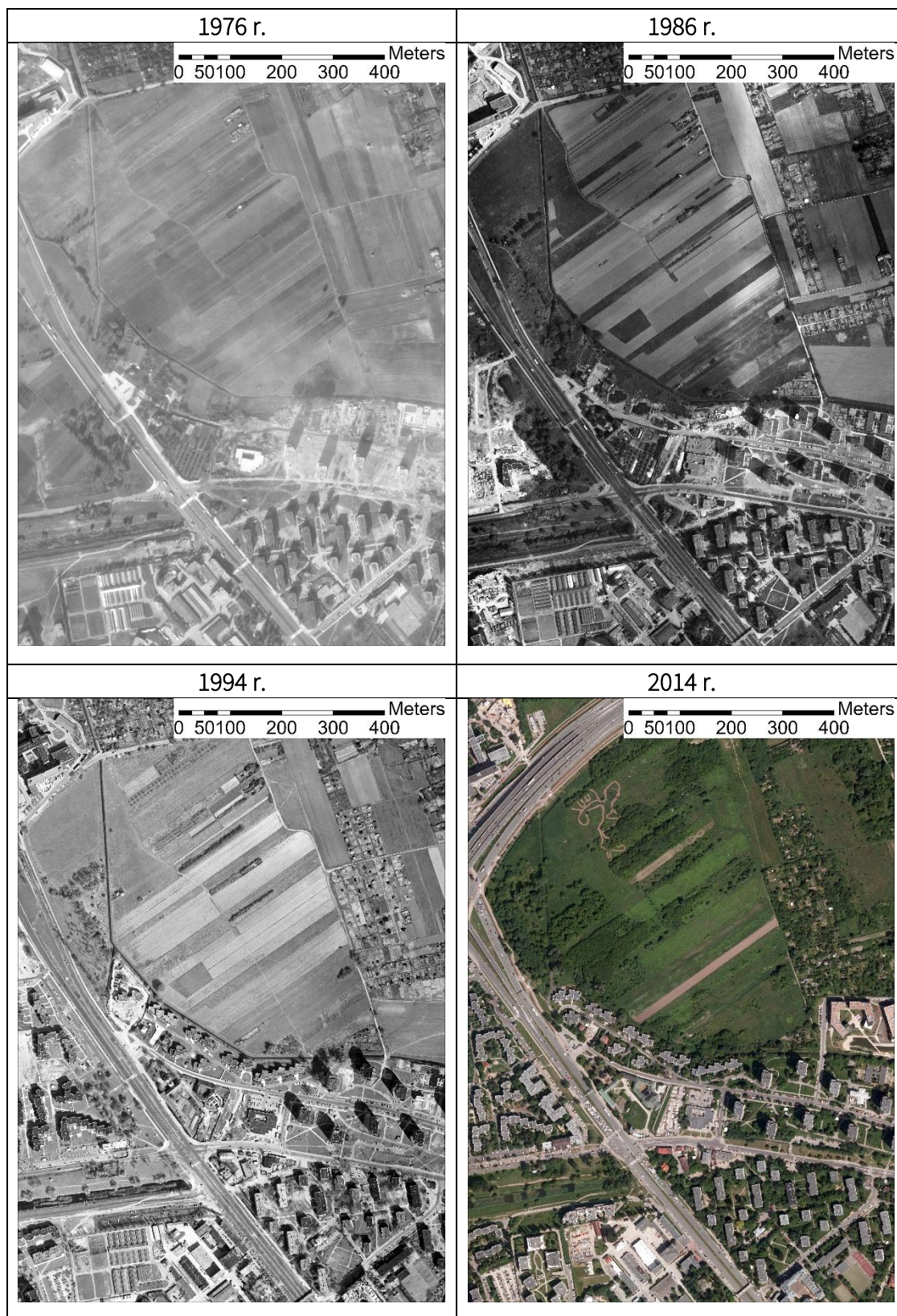
Jakość chmur punktów oceniono w skali czterostopniowej na podstawie dwóch parametrów: kompletności chmury punktów oraz gęstości punktów w chmurze. W przyjętej skali ocena 2 oznaczała jakość niską, 3 – dostateczną, 4 – dobrą, a 5 – bardzo dobrą. Chmura punktów z 1976 roku otrzymała ocenę 3, z 1986 roku – 4, natomiast chmury z lat 1994, 2014 i

2022 uzyskały ocenę 5. Wszystkie wygenerowane chmury punktów spełniają kryterium jakościowe i mogą być wykorzystane do tworzenia z nich true-ortofotomap oraz NMPT (w strukturze *grid*) do wyznaczania współrzędnej (z) tych punktów przejściowych stanowiących podstawę sekwencyjnej orientacji bloków fotogrametrycznych (rozdział 9). Należy nadmienić, że gorsza jakość chmury punktów z roku 1976, skutkuje ograniczoną liczbą wyznaczonych punktów przejściowych, co z kolei może obniżyć dokładność sekwencyjnej orientacji tego bloku (rozdział 9).

Jakość radiometryczna zdjęć lotniczych tworzących poszczególne bloki fotogrametryczne ma istotny wpływ na jakość numerycznych modeli powierzchni terenu (NMPT) generowanych z wykorzystaniem metod MVS. Dokładność tych modeli bezpośrednio determinuje dokładność wyznaczenia współrzędnej wysokościowej (z) punktów wiążących, które są niezbędne do przeprowadzenia transformacji między blokami. Lepsza dokładność wyznaczenia współrzędnych wysokościowych przekłada się na wyższą dokładność wzajemnej orientacji bloków fotogrametrycznych.

True-ortofotomapy

W oparciu o zbudowane bloki fotogrametryczne i wygenerowane dla nich gęste chmury punktów utworzono true-ortofotomapy (rys. 8.4 i 8.5) w środowisku oprogramowania AM (Rys. 8.2). Do generowania true-ortofotomap w programie wykorzystano dwa parametry. Pierwszy parametr odnosi się do wyboru modelu terenu z dwóch dostępnych *grid* lub *mesh* (siatka trójkątów). W tym przypadku na podstawie wygenerowanych chmur punktów utworzono modele *mesh*, które użyto do procesu ortorektyfikacji. Wybór modelu *mesh* związany był z wierniejszym odwzorowaniem obiektów 3D w stosunku do prostszego geometrycznie modelu NMPT (struktura *grid*). *Mesh* uwzględnia także kształt i położenie obiektów 3D. Dzięki temu piksele na zdjęciach są dokładnie przypisane do swoich rzeczywistych pozycji, co minimalizuje błędy geometryczne. Drugim istotnym parametrem wpływającym na proces generowania true-ortofotomapy był parametr mozaikowania *Blending Mode*. Spośród dostępnych opcji: *Mosaic*, *Average* oraz *Disabled*, wybrano tryb *Mosaic*, który określa metodę integracji sąsiadujących zdjęć. W tym trybie piksele z nakładających się obszarów zdjęć są łączone w sposób umożliwiający uzyskanie płynnych przejść między zdjęciami, co obejmuje m.in. korekcję jasności, kontrastu oraz barw.



Rys. 8.4 Zestawienie true-ortofotomap dla każdej epoki 1976, 1986, 1994, 2014 (wykorzystane parametry mozaikowania dla wszystkich bloków fotogrametrycznych: *Blending mode – Mosaic*, *Surface – Mesh*)



Rys. 8.6 Zestawienie true-ortofotomap dla każdej epoki 2022 (wykorzystane parametry mozaikowania dla wszystkich bloków fotogrametrycznych: *Blending mode – Mosaic, Surface – Mesh*)

Dzięki temu możliwe jest stworzenie jednolitej, spójnej mozaiki (true-ortofotomapy) w procesie przetwarzania w oprogramowaniu AM. Drugi parametr *Average* wykorzystuje średnią ważoną wartość koloru (zdjęcia barwne) lub fototonu (zdjęcia panchromatyczne) wszystkich pikseli z poszczególnych obrazów. Przy zastosowaniu trzeciego parametru *Disabled* oprogramowanie wybiera wartość koloru lub fototonu piksela z jednego zdjęcia, które zostało wykonane pod kątem najbardziej zbliżonym do normalnego (prostopadłego) do powierzchni modelu. Zazwyczaj analizowane zdjęcia lotnicze w danym bloku charakteryzują się porównywalną radiometrią. Jednak mogą zdarzać się odstępstwa od tych cech, szczególnie w przypadku archiwalnych zdjęć lotniczych, takich jak te z 1976 roku, które często wyróżniają się niską jakością radiometryczną. Dlatego w celu zwiększenia poprawności mozaikowania jako parametr mozaikowania dla wszystkich zdjęć w poszczególnych blokach użyto parametr *Mosaic*.

Ocena jakości uzyskanych true-ortofotomap dla poszczególnych okresów czasowych, przeprowadzona w skali czterostopniowej, uwzględniała dwa kluczowe aspekty: jakość

geometryczną (poprawność odwzorowania kształtów obiektów terenowych), zależną od jakości chmury punktów, oraz jakość radiometryczną, wynikającą z właściwości radiometrycznych zdjęć składających się na dany blok fotogrametryczny. W przyjętej skali ocena 2 oznaczała jakość niską, 3 – dostateczną, 4 – dobrą, a 5 – bardzo dobrą. True-ortofotomapa z 1976 roku otrzymała ocenę 3, z 1986 roku – 4, natomiast true-ortofotomapy z lat 1994, 2014 i 2022 uzyskały ocenę 5.

Wszystkie wygenerowane true-ortofotomapy spełniają przyjęte kryteria dokładnościowe i mogą zostać wykorzystane w dalszych etapach analizy, w szczególności przy wyszukiwaniu punktów przejściowych niezbędnych do przeprowadzenia sekwencyjnej orientacji bloków fotogrametrycznych (rozdział 9). Należy jednak podkreślić, że niższa jakość true-ortofotomapy z 1976 roku, wynikająca ze słabej jakości radiometrycznej zdjęć, może wpłynąć na dokładność sekwencyjnej orientacji bloku.

9. Sekwencyjna orientacja archiwalnych bloków fotogrametrycznych

Sekwencyjna orientacja archiwalnych bloków fotogrametrycznych zmodyfikowaną metodą *Structure-from-Motion* (SfM) zbudowanych w oparciu o zdjęcia lotnicze pozyskane dla roku 2022, 2014, 1994, 1986 i 1976 polega na wyznaczeniu elementów orientacji jednego bloku (orientowanego) w oparciu o blok referencyjny. Zaproponowany w pracy termin sekwencja odnosi się do chronologicznego procesu orientacji bloków zdjęć, poczynając od najnowszych do najstarszych. Zastosowanie podejścia sekwencyjnego ma kluczowe znaczenie w przypadku dużych różnic czasowych pomiędzy zbiorami danych, gdy bezpośrednie przeniesienie orientacji z najnowszego na najstarszy blok jest niemożliwe ze względu na brak wspólnych cech obrazowych (tj. szczegółów terenowych) pomiędzy zdjęciami. Brak ten może wynikać m.in. z intensywnych przemian urbanizacyjnych oraz zmian w zagospodarowaniu terenu na analizowanym obszarze. W takich przypadkach wykorzystanie bloków pośrednich, obejmujących zdjęcia wykonane w okresach przejściowych, umożliwia identyfikację wspólnych cech obrazowych i stopniowe przeniesienie orientacji z danych współczesnych na najstarsze archiwalne.

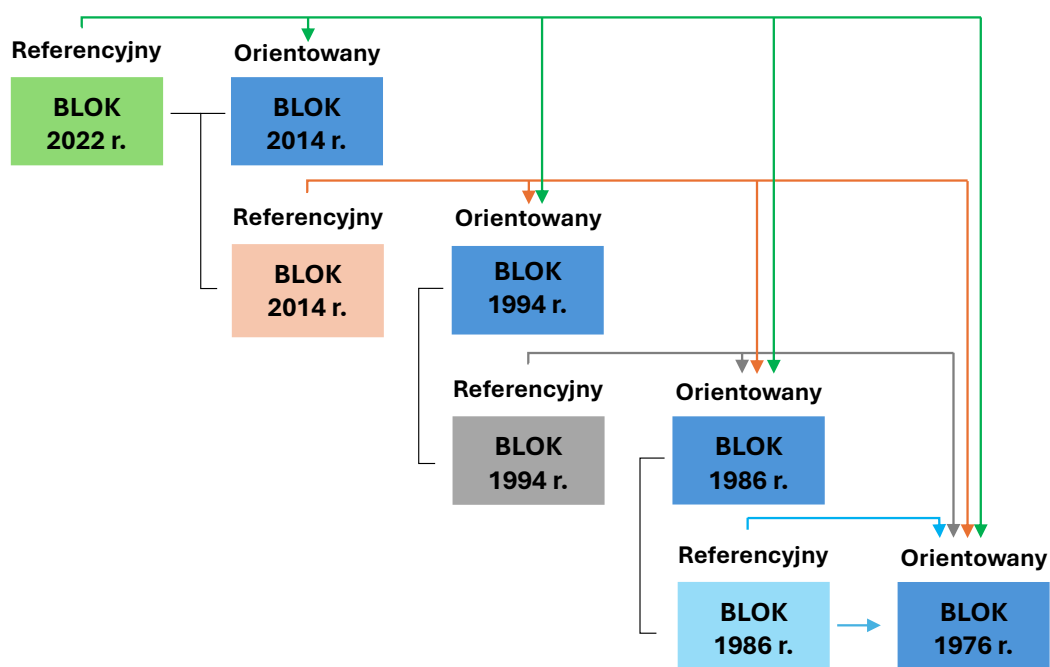
W celu realizacji tego zadania na podstawie pozyskanych zbiorów zdjęć (patrz rozdział 6) zbudowano spójne geometrycznie bloki fotogrametryczne (patrz rozdział 7). Proces budowy bloków odbywał się w układzie lokalnym odnoszącym się do skali zdjęć. Następnie dla każdego bloku wygenerowano true-ortofotomapy oraz Numeryczne Modele Pokrycia Terenu (NMPT). Pomiedzy parami true-ortofotomap z poszczególnych okresów wyznaczano automatycznie punkty przejściowe, dla których wyznaczono współrzędne 2D (x,y) (podrozdział 5.3). Dla każdego z wyznaczonych punktów, na podstawie Numerycznego Modelu Powierzchni Terenu (NMPT), wyinterpolowano wartość współrzędnej (z), co umożliwiło wyznaczenie pełnych współrzędnych trójwymiarowych (xyz). Następnie, uzyskane punkty 3D wykorzystano do obliczenia parametrów transformacji przestrzennej pomiędzy parami bloków. W rezultacie, dla każdej z utworzonych par bloków wyznaczono macierz transformacji (4×4) (patrz podrozdział 5.3). Dzięki zastosowaniu parametrów transformacji możliwe było przeniesienie orientacji ze zdjęć w bloku referencyjnym na zdjęcia w bloku orientowanym, zapewniając zachowanie jednolitości odniesienia przestrzennego dla danych z różnych okresów.

W wyniku tego procesu utworzono sekwencyjne pary true-ortofotomap obejmujące lata 2022, 2014, 1994, 1986 i 1976. Jako referencję przyjęto true-ortofotomapę utworzoną z bloku zdjęć z 2022 r. dla którego dostępne były Elementy Orientacji Zewnętrznej (EOZ) udostępnione przez Główny Urząd Geodezji i Kartografii (GUGiK). W przypadku braku EOZ w głównym zasobie geodezyjnym i kartograficznym dla zdjęć tworzących blok referencyjny, orientację można wykonać samodzielnie poprzez pomiar polowej osnowy fotogrametrycznej oraz przeprowadzenie procesu aerotriangulacji.

Pierwsze sekwencyjne dopasowanie true-ortofotomap przeprowadzono pomiędzy rokiem 2022, a rokiem 2014, wykorzystując wyznaczone wspólne homologiczne punkty.

Analogiczne dopasowanie wykonano dla kolejnych lat (rys. 9.1):

- 2014→1994 (referencją była true-ortofotomapa z 2014 r., a orientowaną – z 1994 r.),
- 1994→1986 (referencją była true-ortofotomapa z 1994 r., a orientowaną – z 1986 r.),
- 1986→1976 (referencją była true-ortofotomapa z 1986 r., a orientowaną – z 1976 r.).



Rys. 9.1 Schemat sekwencyjnego przenoszenia orientacji pomiędzy blokami fotogrametrycznymi

Rozwiązanie to ma szczególne znaczenie w przypadku, gdy nie ma możliwości wyznaczenia orientacji bezpośrednio z sekwencji 2022–1976 ze względu na brak wspólnych wykrytych punktów.

Takie pośrednie sekwencje dopasowania przenoszące orientacje bloków od daty najnowszej do najstarszej można również zastosować do pozostałych danych wieloczasowych w przypadku braku detekcji punktów w bezpośrednich sekwencjach.

W celu znalezienia najlepszego pośredniego rozwiązania orientacji bloków utworzono następujące sekwencje zawarte w tabeli 9.1. Należy pamiętać przy tym, że blokami referencyjnymi w tym przypadku są na przemian bloki: 2022, 2014, 1994 i 1986.

Tabela 9.1 Sekwencje dopasowania bloków fotogrametrycznych

Sekwencje	2022	2014
		1994
		1986
		1976
	2014	1994
		1986
		1976
	1994	1986
		1976
	1986	1976

Dla zapewnienia kompleksowej analizy możliwości orientacji danych wieloczasowych przeprowadzono badania dotyczące automatycznego wyznaczania punktów przejściowych, opierając się na trzech głównych podejściach stosowanych w widzeniu maszynowym (rozdział 5.3):

- 1) *Detect-then-Describe*,
- 2) *Detect-and-Describe*,
- 3) *Describe-then-Detect*.

W pierwszym podejściu *Detect-then-Describe* (wykryj, a następnie opisz) dla algorytmów typu *Hand-Crafted* do detekcji punktów zastosowano te same algorytmy (rozdział 4), które wykorzystano przy budowie bloków fotogrametrycznych (rozdział 5.1). Ponieważ w prawie pionowych zdjęciach lotniczych nie występują znaczące wpływy zniekształceń afinicznych, w badaniach zastosowano detektory punktowe FAST (*Features from Accelerated Segment Test*) i BRISK (*Binary Robust Invariant Scalable Keypoints*) oraz powierzchniowe SIFT (*Scale-Invariant*

Feature Transform), SURF (*Speeded Up Robust Features*) i CenSurE (*Center Surround Extremas*). Do opisu cech wykrytych przez ww. detektory, wykorzystano deskryptor SIFT.

Pierwsze wyniki badań wykazały, że tylko przy wykorzystaniu detektorów FAST i SURF możliwe było wykrycie równomiernie rozmieszczonych punktów wspólnych (przejściowe) pomiędzy sekwencjami true-ortofotomap. Dla detektorów BRISK, CenSurE i SIFT nie było możliwe wyznaczenie punktów wiążących, albo ich liczba była niższa niż przyjęta wartość progowa 15. Za minimalną liczbę niezbędną do przeprowadzenia poprawnej transformacji przyjęto 5 punktów, natomiast pozostałe punkty traktowano jako element kontrolny weryfikujący poprawność transformacji. Dodatkowo, ich rozmieszczenie w obszarze wspólnego pokrycia nie zapewniało równomiernego rozkładu, co obniżało dokładność orientacji między blokami fotogrametrycznymi. Spowodowane to było sposobem działania detektorów FAST i SURF, które – w przeciwieństwie do pozostałych algorytmów – generują znacznie większy zbiór punktów charakterystycznych (*keypoints*, wykorzystywanych w późniejszych etapach do wyznaczenia punktów przejściowych), co pozwala na poprawne wykonanie orientacji (rozdział 4.2). W związku z tym algorytmy BRISK, SIFT i CenSurE nie zostały uwzględnione w dalszych badaniach.

Dodatkowo ze względu na zbyt małą liczbę punktów wyznaczonych przez detektory FAST i SURF pomiędzy niektórymi parami true-ortofotomap zdecydowano się na wykorzystanie ich afinicznych odpowiedników - AFAST (*ang. Affine Features from Accelerated Segment Test*) oraz ASURF (*Affine Speeded Up Robust Features*). Celem tej analizy była ocena czy wykorzystanie detektorów afinicznych zwiększy liczbę wykrytych punktów. Algorytmy te, dzięki uwzględnianiu transformacji afinicznej poprzez wirtualne dodanie kątów obrotu wokół osi kamery, skrócenia oraz pochylenia wirtualnej kamery (rozdział 4), umożliwiają wykrycie większej liczby punktów o bardziej równomiernym rozmieszczeniu w obszarze wspólnego pokrycia. W konsekwencji przyczyniają się do wykrycia większej liczby punktów wiążących co jednoznacznie wpływa na poprawę transformacji między blokami fotogrametrycznymi w danej sekwencji (rozdział 4.2.2).

W podejściu *Detect-then-Describe* zdecydowano się również na wykorzystanie algorytmu typu *Hand-Crafted* SIFT na etapie detekcji punktów, natomiast do ich opisu zastosowano deskryptory oparte na sieci neuronowej (rozwiązane nazwane w tej pracy jako *Learned-Based*) - HardNet, HardNet8, SOSNet, HyNet, MKDDescriptor oraz TFeat (podrozdział 4.4). Wybór tych

deskryptorów był podyktowany ich wysoką skutecznością, potwierdzoną w licznych badaniach naukowych (Pultar, 2020; Tian i in., 2020, 2019; Munkundan i in., 2018; Miles i Mikołajczyk, 2020; Balntas i in., 2016).

W drugim podejściu *Detect-and-Describe* (wykryj i opisz) jako detektor punktów charakterystycznych zastosowano algorytm SuperPoint, natomiast do ich opisu skorzystano z powiązanego z nim deskryptora SuperGlue.

W trzecim podejściu *Describe-then-Detect* (opisz, następnie wykryj) zarówno na etapie opisu i wykrycia wykorzystano algorytm typu *Learned-Based* bazujący na zastosowaniu głębokich sieci neuronowych zaimplementowanych w rozwiązaniu LoFTR.

Warto podkreślić, że w przeprowadzonych badaniach wykorzystano wszystkie analizowane sieci neuronowe, które zostały wcześniej przetrenowane przez ich twórców na zbiorach zdjęć naziemnych. Oznacza to, że parametry tych modeli były dostosowane do specyfiki obrazów naziemnych, a ich zastosowanie w kontekście danych lotniczych wiązało się z koniecznością oceny ich skuteczności dla zdjęć lotniczych, a w szczególności archiwalnych. Taki dobór sieci umożliwił przeprowadzenie analizy ich wydajności oraz dokładności w zakresie detekcji i opisu punktów przejściowych na obrazach lotniczych, pomimo że dane tego typu nie były uwzględnione w procesie pierwotnego treningu.

Dodatkowo w celu zwiększenia liczby wykrywanych punktów dokonano podziału (*tailowania*) wszystkich utworzonych true-ortofotomap (każda z osobna) na 16 (2x8) części (podobszarów). Każda z utworzonych części true-ortofotomapy każdego okresu czasowego odpowiadała w przybliżeniu temu samemu obszarowi (rysunek: 9.6, 9.9, 9.14, 9.15, 9.27, 9.33). Podział taki znacząco podniósł liczbę wyznaczonych punktów, co wpłynęło na równomierne ich rozmieszczenie w obszarze wspólnego pokrycia pomiędzy blokami. Należy podkreślić, że bez zastosowania podziału obrazu na *taile* większość algorytmów ze względu na ograniczenia pamięci operacyjnej RAM (*Random Access Memory*) w komputerze (do 128 GB) nie jest w stanie wykonać pomiaru punktów przejściowych. Wynika to z dużej rozdzielczości zdjęć lotniczych (patrz rozdział 6).

Na rysunkach 9.6, 9.9, 9.14, 9.15, 9.27, 9.33 przedstawiono wyniki dopasowania bloków w poszczególnych sekwencjach, a podobszary, w których nie wyznaczono punktów przejściowych lub ich liczba jest niewystarczająca, oznaczono kolorem czerwonym.

Podobszary charakteryzujące się odpowiednią liczbą punktów przejściowych oznaczono kolorem zielonym.

W celu oceny przydatności algorytmów zastosowanych w trzech podejściach do wyznaczania położenia punktów przejściowych wykorzystywanych w procesie automatycznej orientacji archiwalnych bloków fotogrametrycznych przeprowadzono analizę obejmującą:

- 1) czas detekcji pojedynczego punktu przejściowego,
- 2) liczbę wykrytych punktów przejściowych,
- 3) dokładność wyznaczenia punktów w sekwencjach true-ortofotomap,
- 4) dokładność orientacji bloków fotogrametrycznych na punktach COP – (*Control Orientation Points*).

Całkowity czas wyznaczania punktów przejściowych przez algorytmy w danej sekwencji bloków był determinowany przez proces ich detekcji i opisu. Wszystkie obliczenia związane z pomiarem punktów przejściowych przeprowadzono niezależnie dla każdego algorytmu, wykorzystując pojedynczy procesor obliczeniowy CPU (*Central Processing Unit*), co zapewniało porównanie czasu działania detektorów HC i LB.

Analiza wartości błędów pomiaru punktów przejściowych w oparciu o metody HC i LB wykazała, że średnia wartość RMSE wynosiła około 0.3 piksela. W związku z tym, w celu zachowania poprawnej geometrii bloków we wszystkich algorytmach, przeprowadzono filtrację geometryczną punktów przejściowych. Miała ona na celu eliminację błędnie zidentyfikowanych punktów homologicznych, uzyskanych w procesie automatycznego pomiaru, dla których wartość błędu przekraczała przyjęty próg 1 piksela, co odpowiada około trzykrotności średniego błędu RMS.

Przy ocenie czasu detekcji wszystkich punktów przejściowych (COP i CHOP - *Check Orientation Points*) brany był pod uwagę rodzaj detektora (jego charakterystyka), rodzaj zdjęcia (analogowe, cyfrowe), jakość zdjęcia (radiometria).

W ocenie liczby wyznaczonych punktów przejściowych w poszczególnych blokach fotogrametrycznych kluczowym parametrem pozostawał rodzaj detektora, który w zależności od jego specyfiki (opisanej w rozdziale 4) wykrywa większą bądź mniejszą ich liczbę. Każdy blok (true-ortofotomapę) analizowano indywidualnie, z uwzględnieniem zróżnicowanej liczby zdjęć składowych, ich odmiennej skali oraz różnego stopnia ich wzajemnego pokrycia. Pierwsza

część oceny obejmowała analizę wartości błędów punktów wykrytych przez algorytm, opisanych i dopasowanych o współrzędnych (x,y) . Natomiast dalsza część uwzględniała także wymiar wysokościowy (współrzedną z), wyinterpolowany z Numerycznego Modelu Pokrycia Terenu (NMPT), niezbędną do przeprowadzenia transformacji bloków.

Należy jednak podkreślić, że ze względu na ograniczenia budowy NMPT, wynikające z dyskretnej struktury siatki grid, wykryte przez algorytm punkty (detektory punktowe – patrz rozdział 4.2) na krawędziach zabudowy mogą znajdować się pomiędzy węzłami siatki. W takich przypadkach interpolacja wysokości może być obciążona znacznym błędem, co może wpłynąć na dokładność orientacji. W przypadku detektorów typu *blob* (patrz rozdział 4.2), dzięki ich powierzchniowemu charakterowi, wpływ tej niedokładności jest znacznie ograniczony.

Ocena dokładności sekwencyjnej orientacji bloków, obejmująca dwa aspekty: ocenę wyznaczonych punktów 2D, wykrytych na true-ortofotomapach na podstawie błędu transformacji 2D, oraz ocenę orientacji bloków na podstawie punktów 3D (xyz) , opisaną średnim błędem kwadratowym (RMSE). Miary oceny liczone były według wzoru zawartego w tabeli 5.1.

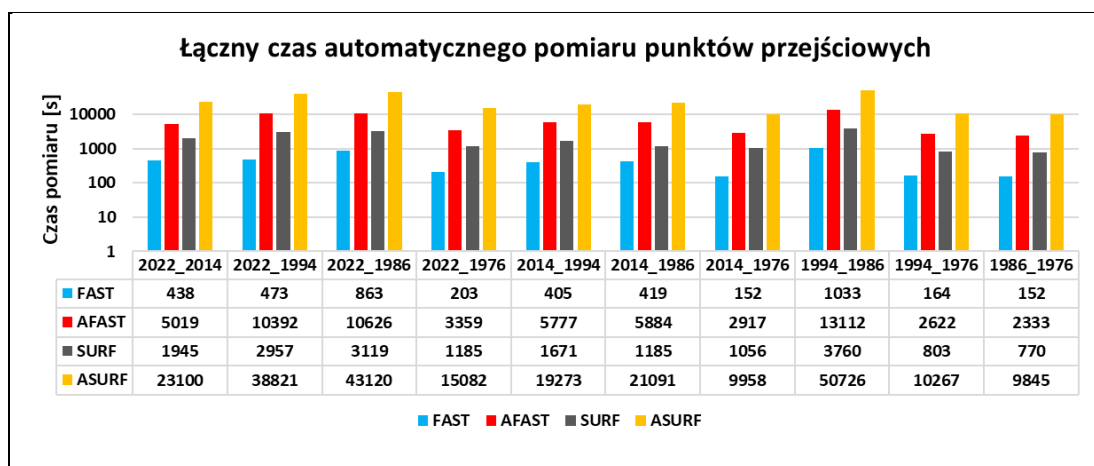
Końcowa ocena odnosiła się do weryfikacji dokładności sekwencyjnego dopasowania bloków fotogrametrycznych w oparciu o niezależne punkty kontrolne CHOP wyznaczone przez dany algorytm CV. Punkty te nie należy utożsamiać z punktami polowej osnowy fotogrametrycznej mierzonymi w terenie, ale z punktami przejściowymi wyznaczonymi przez grupy zastosowanych algorytmów.

9.1 Ocena poprawności procesu orientacji bloków fotogrametrycznych – podejście Detect-then-Describe (algorytmy Hand-Crafted)

Po etapie dwuwymiarowej filtracji geometrycznej punktów wiążących, przeprowadzonej w procesie transformacji homograficznej, przystąpiono do całościowej oceny dopasowania bloków fotogrametrycznych w podejściu *Detect-then-Describe*. Obejmowała ona analizę czasu obliczeń, liczby i rozmieszczenia punktów przejściowych, a także dokładności orientacji bloków fotogrametrycznych.

9.1.1 Ocena czasu automatycznego pomiaru punktów przejściowych w sekwencyjnej orientacji bloków fotogrametrycznych

Pierwsza analiza dopasowania sekwencyjnego bloków fotogrametrycznych w podejściu *Detect-then-Describe* z wykorzystaniem algorytmów HC/HC polegała na ocenie czasu wykrycia, opisu i dopasowania punktów przejściowych w danej sekwencji bloków fotogrametrycznych utworzonych dla lat 2022-2014, 2014-1994, 1994-1986 i 1986-1976 (rys. 9.2). W każdej sekwencji punkty przejściowe wyznaczone były dla pary bloków tworzących tę sekwencję. Całkowity czas automatycznego pomiaru punktów przejściowych dla każdej sekwencji obliczano jako sumę czasu ich detekcji, opisu i dopasowania.



Rys. 9.2 Łączny czas automatycznego wyznaczenia punktów przejściowych przez algorytmy HC/HC w podejściu *Detect-then-Describe* w sekwencjach bloków fotogrametrycznych z roku 2022, 2014, 1994, 1986 i 1976

Na rysunku 9.2 przedstawiono czasy całkowitego wyznaczenia punktów przejściowych w danej sekwencji bloków związany z ich wykryciem i opisem przez algorytmy HC/HC. Wszystkie obliczenia związane z pomiarem punktów wykonano niezależnie dla każdego algorytmu na pojedynczym procesorze obliczeniowym, co zapewniało jednakowe warunki wyznaczenia czasu.

Najkrótsze czasy wyznaczenia punktów przejściowych dla algorytmów HC/HC (rys. 9.2) odnotowano dla detektora FAST we wszystkich sekwencjach bloków i mieściły się w przedziale 2–3 min. Najdłuższe czasy wyznaczenia punktów przejściowych uzyskano przy wykorzystaniu detektora ASURF we wszystkich sekwencjach czasowych i mieściły się w granicy 164–171 min. Czas wyznaczenia punktów przejściowych we wszystkich sekwencjach przez detektor FAST

w stosunku do detektora AFAST był od 11 do 21 razy krótszy, w stosunku do detektora SURF od 3 do 6 razy krótszy, a w stosunku do detektora ASURF od 50 do 70 razy krótszy.

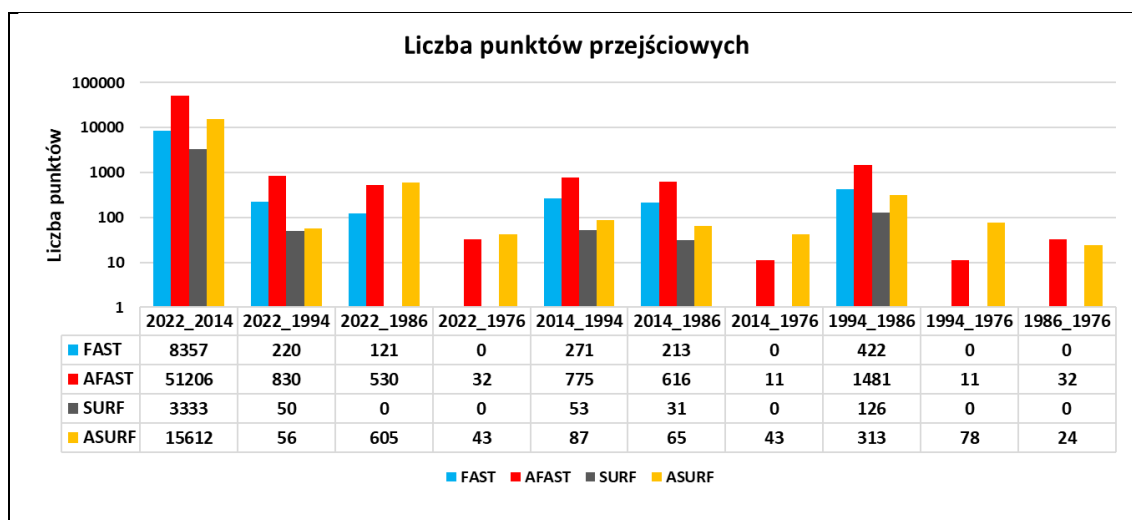
Najdłuższy czas wyznaczenia punktów przejściowych algorytmów HC odnotowano dla bloków tworzących sekwencję 1994–1986, gdzie wyznaczenie punktów przy udziale algorytmu ASURF trwało 14 godzin. Zbliżone wartości pomiaru czasu wyznaczenia punktów przejściowych przez algorytmy HC (FAST około 400 punktów, AFAST około 6000 punktów, SURF około 1500 punktów i ASURF około 20000 punktów) uzyskano dla bloków utworzonych ze zdjęć o podobnej rozdzielczości przestrzennej w sekwencjach 2014–1994 i 2014–1986. Analizując wpływ rozdzielczości przestrzennej true-ortofotomapy na czas wyznaczenia punktów przejściowych widać, że im rozdzielczość przestrzenna zdjęcia jest większa (mniejszy piksel terenowy) tym czas wyznaczenia punktów jest dłuższy. Zależność ta widoczna jest w sekwencjach 2022–1986 i 1994–1986.

Warto również zauważyć, że w sekwencjach czasowych utworzonych ze zdjęć z roku 1976 (o mniejszej rozdzielczości przestrzennej) czasy pomiaru punktów przez poszczególne algorytmy HC były kilkakrotnie krótsze od pozostałych sekwencji. Prawdopodobnie wynika to z wyznaczenia mniejszej liczby punktów przejściowych, co było konsekwencją niskiej rozdzielczości przestrzennej zdjęć (duża wartość GSD; rozdział 6) oraz ich słabej jakości radiometrycznej. W związku z tym do wyznaczenia punktów przejściowych algorytmy potrzebowały kilkakrotnie mniej czasu.

Podsumowując można stwierdzić, że czas automatycznego wyznaczania punktów przejściowych jest determinowany przede wszystkim przez rozdzielczość przestrzenną obrazu – im większa rozdzielczość, tym dłuższy czas obliczeń. Rodzaj zdjęcia (cyfrowe czy analogowe) nie ma istotnego wpływu na czas przetwarzania danych, natomiast istotnym czynnikiem jest jakość radiometryczna obrazu. Ponadto czas obliczeń zależy od charakterystyki działania poszczególnych algorytmów, co potwierdzają uzyskane wyniki: najdłuższy czas wyznaczania punktów przejściowych charakteryzował detektor ASURF, natomiast najkrótszy – detektor FAST, co jest zgodne z założeniami metodologicznymi opisanymi w rozdziale 4.

9.1.2 Ocena liczby wykrytych punktów przejściowych i ich rozkładu w sekwencyjnej orientacji bloków fotogrametrycznych

Druga analiza dopasowania sekwencyjnego bloków fotogrametrycznych w podejściu *Detect-then-Describe* z wykorzystaniem algorytmów HC/HC polegała na ocenie liczby wyznaczonych punktów przejściowych i ich rozkładu na obszarze dopasowania.



Rys. 9.3. Liczba pomierzonych automatycznie punktów przejściowych COP i CHOP w poszczególnych sekwencjach bloków fotogrametrycznych (1976 r., 1986 r., 1994 r., 2011 r., 2022 r.) dla algorytmów HC (FAST, AFAST, SURF i ASURF) w podejściu *Detect-then-Describe*

Na rysunku 9.3 przedstawiono liczbę wyznaczonych punktów przejściowych (łącznie COP i CHOP) w sekwencjach czasowych pomiędzy utworzonymi blokami, obejmujących okres od 2022 do 1976 roku, w podejściu *Detect-then-Describe*. Analiza wyników wskazuje, że największą liczbę punktów przejściowych uzyskano dla sekwencji 2022–2014 (zdjęcia cyfrowe), a najwyższy wynik osiągnięto przy wykorzystaniu detektora AFAST (51 206 punktów).

W większości analizowanych sekwencji czasowych (2022–2014, 2022–1994, 2022–1986, 2014–1986, 1994–1986 oraz 1986–1976) największą liczbę wykrytych punktów uzyskano również przy zastosowaniu detektora afinicznego AFAST.

Najmniejsza liczba punktów została wyznaczona w sekwencjach 2022–1976, 2014–1976, 1994–1976 i 1986–1976. W tym przypadku należy zauważyć, że wszystkie te sekwencje odnoszą się do bloku utworzonego ze zdjęć z roku 1976, a wpływ na liczbę wyznaczanych punktów ma zarówno mała skala tych zdjęć (1:18 500 ~ GSD 25cm), jak i niższa jakość radiometryczna. Również dla tych sekwencji algorytmy FAST i SURF nie były w stanie wyznaczyć punktów

przejściowych, co potwierdza wniosek znaczącego wpływu jakości radiometrycznej tych zdjęć na pomiar. Zasada działania algorytmu FAST polega na tym, że wykrywa punkty na podstawie lokalnych różnic jasności (narożniki i krawędzie), a SURF bazuje na analizie gradientów zmian stopni szarości. Oba algorytmy cechują się wysoką wrażliwością na niski kontrast i słabą jakość radiometryczną.

Analizując sekwencje obejmujące blok z 2022 roku oraz bloki z lat 1994, 1986 i 1976, można zaobserwować systematyczny spadek liczby wykrywanych punktów wraz ze wzrostem różnicy czasowej między porównywanymi danymi. Jest to wynikiem istotnych zmian urbanizacyjnych zachodzących na danym obszarze w okresie 46 lat (od 1976 do 2022 roku), takich jak budowa dróg, osiedli mieszkaniowych czy przekształcanie terenów rolnych na obszary zabudowane. Im większy odstęp czasowy pomiędzy pozyskanymi danymi, tym mniejsza liczba wspólnych cech charakterystycznych, co przekłada się na systematyczne obniżenie liczby wykrywanych punktów przejściowych.

Analogiczny trend spadkowy obserwuje się w sekwencjach obejmujących rok 2014 (tj. 2014–1994, 2014–1986, 2014–1976), co wynika z podobnych czynników – zmiany pokrycia terenu spowodowane procesami urbanizacyjnymi nasilają się wraz ze wzrostem różnicy czasowej między pozyskanymi zdjęciami lotniczymi.

Drugi co do wielkości zbiór punktów przejściowych uzyskano dla wszystkich zastosowanych detektorów w sekwencji 1994–1986 (zdjęcia analogowe). Wpłynął na to krótki odstęp czasowy pomiędzy pozyskanymi danymi (8 lat), podobna jakość radiometryczna oraz zbliżona rozdzielczość zdjęć w blokach. Biorąc pod uwagę liczbę wykrytych punktów najslabszym detektorem HC okazał się detektor powierzchniowy SURF (brak wyznaczonych punktów w sekwencjach 2022–1986, 2022–1976, 2014–1976, 1994–1976, 1986–1976). Ocena wpływu rodzaju zdjęć (analogowych i cyfrowych) na liczbę wykrytych punktów przejściowych jest utrudniona ze względu na znaczną dynamikę zmian pokrycia terenu w okresie od pozyskania najstarszych zdjęć (1976 r.) do najnowszych obrazów (2022 r.). Zmiany te, głównie wynikające z procesów urbanizacyjnych, prowadzą do istotnych różnic w liczbie wykrywanych punktów w zależności od analizowanej sekwencji bloków fotogrametrycznych oraz stosowanego algorytmu. Niemniej jednak porównując sekwencję 2022–2014, utworzoną na podstawie zdjęć cyfrowych o różnych rozdzielczościach przestrzennych, z sekwencją 1994–1986, składającą się

ze zdjęć analogowych o zbliżonej rozdzielczości przestrzennej, stwierdzono, że liczba wykrytych punktów przejściowych w przypadku sekwencji 2022–2014 była od kilku do kilkunastu razy większa niż w przypadku sekwencji 1994–1986, w zależności od zastosowanego algorytmu. Prawdopodobną przyczyną tych różnic jest lepsza jakość radiometryczna zdjęć cyfrowych w porównaniu do zdjęć analogowych.

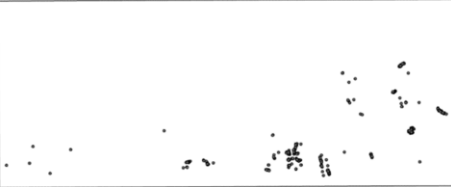
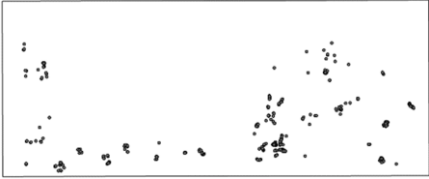
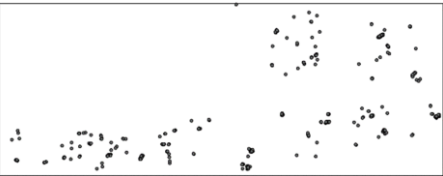
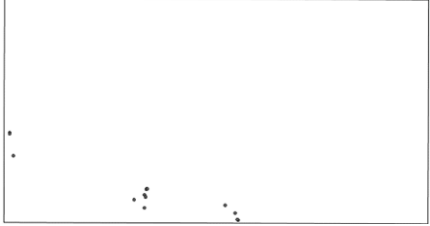
W podejściu *Detect-then-Describe* dla kombinacji HC/HC do wyznaczenia punktów zastosowano detektor FAST, AFAST, SURF i ASURF, a do ich opisu deskryptory SIFT. W związku z tym zbiory punktów wykrytych przez detektory HC stanowiły podstawę do ich opisu przez deskryptor SIFT, a opis wyznaczanych punktów odnosi się bezpośrednio do detektorów.

Podsumowując wyniki uzyskanych analiz dotyczących liczby wykrytych punktów przejściowych i ich rozkładu w sekwencjach bloków fotogrametrycznych można stwierdzić, że:

1. Wykorzystanie zdjęć cyfrowych skutkuje znacząco większą liczbą wykrytych punktów przejściowych w porównaniu do zdjęć analogowych.
2. Największą liczbę punktów przejściowych uzyskano dla sekwencji 2022–2014, a najskuteczniejszym detektorem okazał się AFAST, co wynika z jego zdolności do generowania dodatkowych obrazów.
3. Liczba wykrywanych punktów przejściowych maleje wraz ze wzrostem różnicy czasowej między zdjęciami, co wynika z intensywnych zmian urbanizacyjnych.
4. Wpływ jakości zdjęć przekłada się bezpośrednio na liczbę wyznaczonych punktów. Najmniejszą liczbę punktów wykryto w sekwencjach zawierających zdjęcia z 1976 roku, co jest efektem ich małej skali (1:18 500) a przede wszystkim gorszej jakości radiometrycznej. Za pomocą algorytmów FAST i SURF nie było można wyznaczyć punktów w tych przypadkach.
5. Stabilność detektorów zależna jest od różnic czasowych pomiędzy poszczególnymi sekwencjami. W sekwencjach o krótkim odstępie czasowym (np. 1994–1986) uzyskano więcej punktów ze względu na podobną jakość radiometryczną i rozdzielczość zdjęć.
6. Detektor SURF okazał się najłagodniejszy, nie identyfikując punktów w wielu sekwencjach zwłaszcza dla sekwencji obejmujących duże odstępy czasowe pomiędzy blokami.

Algorytm	2022-2014	2022-1994	2022-1986	2022-1976	2014-1994
	FAST				
Algorytm	2014-1986	2014-1976	1994-1986	1994-1976	1986-1976
	FAST				

Rys. 9.4 Rozkład punktów przejściowych wyznaczonych przez algorytm HC (FAST) w podejściu *Detect-then-Describe* sekwencji bloków fotogrametrycznych (2022, 2014, 1994, 1986, 1976)

Algorytm	2022-2014	2022-1994	2022-1986	2022-1976	2014-1994
					
	AFAST	2014-1986	2014-1976	1994-1986	1986-1976
					

Rys. 9.5 Rozkład punktów przejściowych wyznaczonych przez algorytm HC (AFAST) w podejściu *Detect-then-Describe* sekwencji bloków fotogrametrycznych (2022, 2014, 1994, 1986, 1976)

Z analizy rozkładu punktów wiążących (rys. 9.4 i 9.5) wynika, że punkty wykryte przez detektor FAST znajdowały się głównie na elementach zabudowy, w szczególności na dachach budynków, w miejscach o dużej różnicy kontrastu w otoczeniu punktu, co wynika ze sposobu działania algorytmu (detektor punktowy). Były to narożniki dachów, załamania krawędzi wewnętrznych elementów dachowych (attyki) oraz narożniki kominów wentylacyjnych. Dodatkowo, część punktów została zidentyfikowana na drogach, wzdłuż krawędzi poziomych znaków drogowych, a także na przecięciach krawędzi chodników i jezdni. Niewielką liczbę punktów wykryto na przecięciach linii boisk sportowych. Na obszarze częściowo rolniczym wykorzystanie detektora FAST nie umożliwiło wykrycia punktów przejściowych, a w rejonie ogródków działkowych wykryto jedynie nieliczne punkty, głównie na krawędziach dachów altan i domków działkowych oraz na przecięciach wewnętrznych dróg dojazdowych i chodników.

Analiza rozkładu i liczby punktów wykrytych przez detektor FAST w podejściu *Detect-then-Describe* (rys. 9.4) wskazuje, że najbardziej równomierne pod względem liczby i rozmieszczenia punktów wyniki uzyskano dla sekwencji bloków 2022–2014. Wpływ na to miały dwa czynniki. Po pierwsze, stosunkowo krótki odstęp czasowy (8 lat) między porównywanymi danymi, który sprawił, że zmiany w pokryciu terenu były minimalne. Po drugie, wysoka jakość zdjęć lotniczych z lat 2022 i 2014, uzyskanych przy użyciu kamery cyfrowej o bardzo dobrej radiometrii i geometrii. Te dwa czynniki sprawiły, że algorytmy były w stanie wyznaczyć duże zbiory punktów przejściowych, co znacząco wpłynęło na poprawne wykonanie orientacji bloków fotogrametrycznych. W analizowanej sekwencji większość punktów przejściowych została zidentyfikowana w obszarach zabudowy oraz na drogach, natomiast najmniej na terenach ogródków działkowych i obszarach nieużytków.

Analizując rok 2022 i kolejne tworzone z nim sekwencje czasowe (2014, 1994, 1986, 1976) widoczny jest spadek liczby punktów (rys. 9.4) oraz ich mniej równomierne rozmieszczenie. W przypadku sekwencji 2022–1976 nie udało się wyznaczyć żadnych punktów przejściowych. Główną przyczyną tego trendu są intensywne zmiany urbanizacyjne zachodzące na danym obszarze. Im dłuższy okres między dopasowywanymi blokami, tym większe różnice w pokryciu terenu, co prowadzi do zmniejszenia liczby wykrywanych wspólnych cech. W przypadku sekwencji 2022–1976, algorytm FAST nie był w stanie wykryć punktów nie tylko z powodu

dużego odstępu czasowego między danymi, ale także ze względu na niską jakość radiometryczną zdjęć z 1976 roku tworzącego blok.

Ta sama zależność została zauważona dla sekwencji czasowych obejmujących rok 2014 (1994, 1986, 1976), gdzie również w sekwencji 2014–1976 algorytm nie wyznaczył automatycznie punktów przejściowych. W przypadku sekwencji bloków fotogrametrycznych utworzonych ze zdjęć panchromatycznych z roku 1994, 1986, 1976 punkty zostały wyznaczone wyłącznie w sekwencji 1994–1986. Krótki okres między datami pozyskania zdjęć (8 lat) sprzyjał skutecznej detekcji. W sekwencjach 1994–1976 i 1986–1976 algorytm nie był w stanie pomierzyć automatycznie punktów ze względu na słabą jakość radiometryczną zdjęć z roku 1976.

W celu szerszej analizy przestrzennego rozkładu punktów przejściowych wyniki dopasowania (liczba i rozmieszczenie) dla sekwencji bloków fotogrametrycznych, z uwzględnieniem algorytmów HC podejścia *Detect-then-Describe*, zostały przedstawione w formie graficznej na rysunkach 9.6 i 9.9. W badanych sekwencjach kolor czerwony wskazuje na brak wyznaczonych punktów przejściowych lub ich niewystarczającą liczbę (przyjęta minimalna liczba 15 punktów), natomiast kolor zielony oznacza ich odpowiednią liczbę.

FAST									
2022-2014		2022-1994		2022-1986		2022-1976		2014-1994	
1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16
FAST									
2014-1986		2014-1976		1994-1986		1994-1976		1986-1976	
1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16
AFAST									
2022-2014		2022-1994		2022-1986		2022-1976		2014-1994	
1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16
AFAST									
2014-1986		2014-1976		1994-1986		1994-1976		1986-1976	
1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16

Rys. 9.6 Rozkład punktów przejściowych (w podziale tailowym – 16) sekwencji (2022–2014, 2022–1994, 2022–1986, 2022–1976, 2014–1994, 2014–1986, 2014–1976, 1994–1986, 1994–1976, 1986–1976) bloków fotogrametrycznych (true–ortofotomap) z uwzględnieniem algorytmów HC (FAST, AFAST) podejścia *Detect-then-Describe* (kolor zielony oznacza wystarczającą liczbę punktów w danym podobszarze, natomiast kolor czerwony oznacza ich brak lub niewystarczającą liczbę)

Na podstawie zestawienia wyników dopasowania sekwencji bloków fotogrametrycznych (rys. 9.6) w podejściu *Detect-then-Describe* algorytmu FAST stwierdzono, że możliwe jest wykonanie orientacji czasowej (przeniesienie orientacji z jednego bloku na drugi) dla sekwencji 2022–2014, 2022–1994, 2022–1986, 2014–1994, 2014–1986 i 1994–1986. W pozostałych sekwencjach 2022–1976, 2014–1976, 1994–1976, 1986–1976 z powodu braku wyznaczonych punktów przejściowych przez algorytm FAST nie ma możliwości wykonania orientacji bloków fotogrametrycznych.

Punkty wyznaczone za pomocą algorytmu AFAST (rys. 9.5) były zlokalizowane na tych samych elementach, co w przypadku algorytmu FAST, tj. na narożnikach dachów, załamaniach krawędzi wewnętrznych elementów dachowych, krawędziach linii znaków drogowych poziomych na drogach, przecięciach krawędzi chodników i dróg, a także na krawędziach dachów altan i domków działkowych. Różnica polegała na liczbie wykrytych punktów, która w zależności od sekwencji czasowej była od trzech do sześciu razy większa na korzyść algorytmu AFAST.

Największą liczbę punktów oraz ich równomierny rozkład uzyskano dla sekwencji czasowej 2022–2014 (odstęp czasowy między danymi wynosił 8 lat). Analizując sekwencje utworzone na podstawie bloku 2022 (zdjęć cyfrowych) z blokami 1994, 1986, 1976 (zdjęcia analogowe) oraz bloku 2014 (zdjęć cyfrowych) z tymi samymi blokami 1994, 1986, 1976 (zdjęcia analogowe), zauważalny jest spadek liczby punktów przejściowych oraz mniej korzystny ich rozkład wraz z odstępem czasowym pomiędzy poszczególnymi blokami (rys. 9.5). Główną przyczyną tego zjawiska są zmiany urbanizacyjne zachodzące na analizowanym obszarze. Im dłuższy okres czasowy dzieli dopasowywane bloki, tym większe różnice w pokryciu terenu, co prowadzi do zmniejszenia liczby wykrywanych punktów przejściowych oraz ich mniej równomiernego rozmieszczenia.

W sekwencji 2022–1994 w części wspólnego obszaru dopasowania obejmującej teren ogródków działkowych i użytków rolnych nie wykryto punktów przejściowych. Porównując sekwencję 2022–1994 (odstęp czasowy 28 lat) z sekwencją 2022–1986 (odstęp czasowy 36 lat), stwierdzono, że pomimo dłuższego odstępu czasowego algorytm wykrył punkty w tym obszarze. Może to wynikać z większej skali zdjęć z 1986 roku (1:3 000), co przełożyło się na wyższą szczegółowość obrazu, a tym samym umożliwiło detektorowi identyfikację większej liczby punktów charakterystycznych.

Ze względu na dużą różnicę czasową między pozyskanymi danymi oraz słabą jakość radiometryczną bloku fotogrametrycznego z 1976 roku, dla sekwencji 2022–1976 oraz 2014–1976 uzyskano najmniejsze zbiory punktów przejściowych. W przypadku sekwencji utworzonych na podstawie zdjęć analogowych największą liczbę punktów wyznaczono w sekwencji 1994–1986, co wynikało z krótkiego odstępu czasowego pomiędzy danymi oraz zbliżonej skali zdjęć. Natomiast najmniejszą liczbę punktów algorytm AFAST wyznaczył

w sekwencjach 1994–1976 oraz 1986–1976 ze względu na niską jakość radiometryczną zdjęć z roku 1976 oraz ich małą skalę.

Na podstawie analizy wyników dopasowania sekwencji bloków fotogrametrycznych (rys. 9.6) w podejściu *Detect-then-Describe* z wykorzystaniem algorytmu AFAST stwierdzono, że możliwe jest wykonanie orientacji czasowej (przeniesienia orientacji z jednego bloku na drugi) dla sekwencji 2022–2014, 2022–1994, 2022–1986, 2014–1994, 2014–1986 oraz 1994–1986. W przypadku sekwencji 2022–1976, 2014–1976, 1994–1976 oraz 1986–1976, z uwagi na niewielką liczbę wykrytych punktów przejściowych oraz ich nierównomierne rozmieszczenie, orientacja bloków fotogrametrycznych nie jest możliwa.

Większość wykrytych punktów przez detektor SURF położona była na powierzchniach dachów budynków. Część punktów została pomierzona przez algorytm na drogach, głównie na powierzchniach znaków drogowych poziomych, a także powierzchniach chodników i dróg. Na obszarze częściowo rolniczym stwierdzono brak punktów. W przypadku terenów ogródków działkowych wykryto niewielką liczbę punktów położonych głównie na powierzchni dachów domków działkowych oraz dróg i chodników.

Analizując rozkład i liczbę punktów wykrytych przez detektor SURF (rys. 9.7) najlepsze rozwiązanie pod względem liczby i rozmieszczenia punktów przypadło na sekwencje 2022–2014. Wpływają na to dwa czynniki. Pierwszy odnosi się do krótkiego odstępu czasu (8 lat) pozyskania danych w sekwencji bloków (małe zmiany pokrycia terenu powodują możliwość detekcji dużej liczby punktów na całej jego powierzchni). Drugi odnosi się do jakości zdjęć lotniczych, gdzie zdjęcia z roku 2022 i 2014 charakteryzowały się bardzo dobrą jakością radiometryczną (wykonane kamerą cyfrową). W sekwencji tej najwięcej punktów wyznaczono na powierzchniach zabudowy i dróg, natomiast najmniej na ogródkach działkowych i obszarach nieużytków.

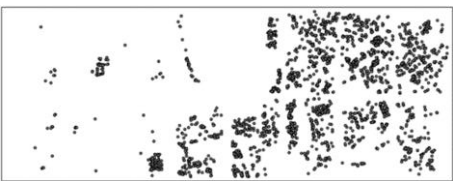
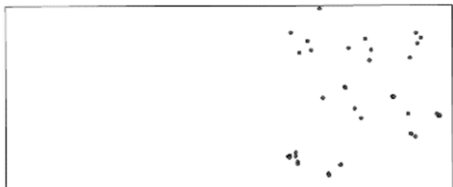
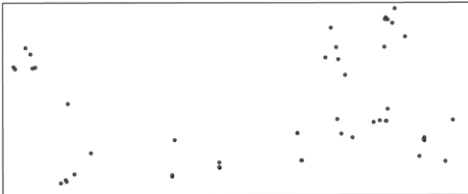
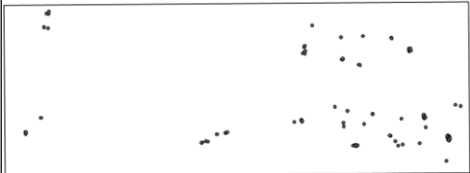
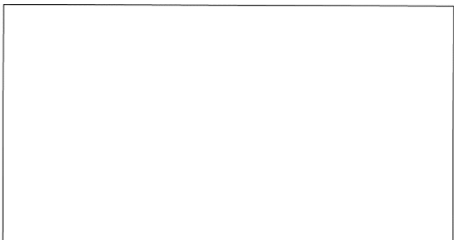
Analizując kolejne sekwencje czasowe utworzone w odniesieniu do roku 2022 (2014, 1994, 1986, 1976), zauważalny jest znaczący spadek liczby wykrytych punktów – od 25 do nawet 100 razy (rys. 9.7). W sekwencji 2022–1994 rozkład punktów nie jest równomierny, natomiast w sekwencjach 2022–1986 i 2022–1976 algorytm nie wykrył żadnych punktów przejściowych. Wynika to z rosnącego wpływu zmian w pokryciu terenu wraz z wydłużaniem się odstępu

czasowego między dopasowywanymi blokami. Brak punktów w sekwencjach 2022–1986 oraz 2022–1976 jest również konsekwencją niskiej jakości radiometrycznej zdjęć z 1976 roku.

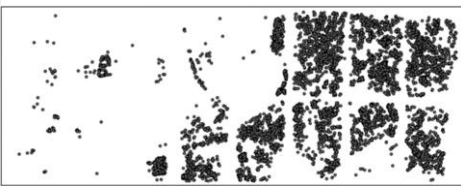
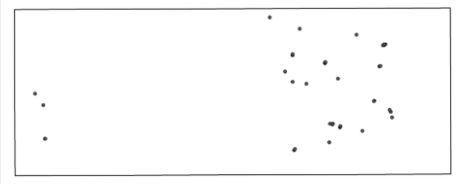
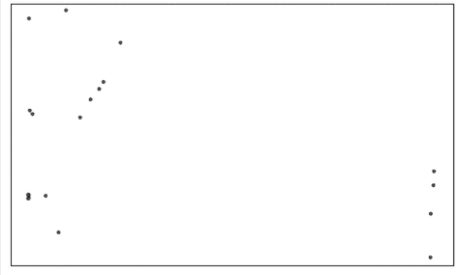
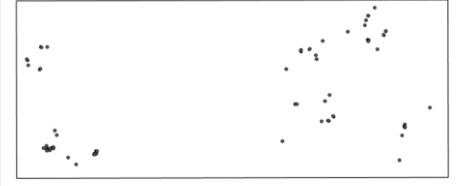
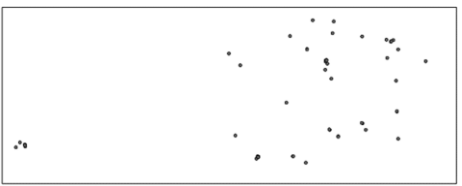
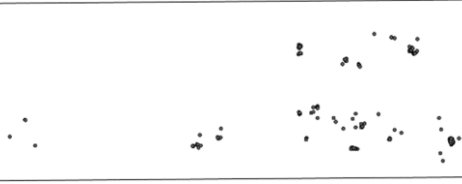
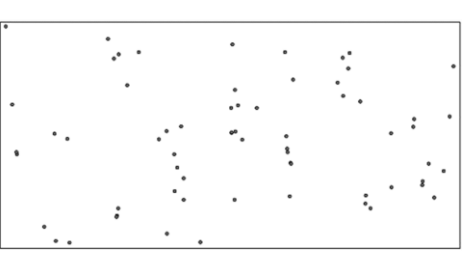
Podobne zależności zaobserwowano w przypadku sekwencji tworzonych z rokiem 2014 (1994, 1986, 1976). W sekwencji 2014–1976 algorytm nie był w stanie automatycznie wyznaczyć punktów przejściowych, natomiast w sekwencji 2014–1986 punkty zostały wykryte. W przeciwieństwie do sekwencji 2022–1986, w której punktów nie wykryto, w przypadku 2014–1986 ich obecność może wynikać z krótszego odstępu czasowego (28 lat w porównaniu do 36 lat dla 2022–1986), co skutkowało mniejszymi zmianami w pokryciu terenu i umożliwiło detekcję.

Dla sekwencji czasowych bloków fotogrametrycznych utworzonych ze zdjęć analogowych z lat 1994, 1986, 1976, punkty przejściowe zostały wykryte jedynie w sekwencji 1994–1986. Krótki odstęp czasowy między danymi ograniczył wpływ zmian w pokryciu terenu, co sprzyjało detekcji punktów. Natomiast w sekwencjach 1994–1976 oraz 1986–1976 algorytm nie wykrył punktów przejściowych, co było konsekwencją niskiej jakości radiometrycznej zdjęć z 1976 roku.

Na podstawie zestawienia wyników dopasowania sekwencji bloków fotogrametrycznych (rys. 9.9) w podejściu *Detect-then-Describe* algorytmu SURF stwierdzono, że możliwe jest wykonanie orientacji czasowej (przeniesienie orientacji z jednego bloku na drugi) dla sekwencji 2022–2014, 2022–1994, 2022–1986, 2014–1994, 2014–1986 i 1994–1986. W przypadku sekwencji 2022–1976, 2014–1976, 1994–1976 oraz 1986–1976, ze względu na brak wykrytych punktów przejściowych, orientacja bloków fotogrametrycznych nie jest możliwa.

Algorytm	2022-2014	2022-1994	2022-1986	2022-1976	2014-1994
					
	SURF	2014-1986	2014-1976	1994-1986	1986-1976
					

Rys. 9.7 Rozkład punktów przejściowych wyznaczonych przez algorytm HC (SURF) w podejściu *Detect-then-Describe* sekwencji bloków fotogrametrycznych (2022, 2014, 1994, 1986, 1976)

Algorithm	ASURF	2022-2014		2022-1994		2022-1986		2022-1976		2014-1994	
		2014-1986		2014-1976		1994-1986		1994-1976		1986-1976	

Rys. 9.8 Rozkład punktów przejściowych wyznaczonych przez algorytm HC (ASURF) w podejściu *Detect-then-Describe* sekwencji bloków fotogrametrycznych (2022, 2014, 1994, 1986, 1976)

SURF									
2022-2014		2022-1994		2022-1986		2022-1976		2014-1994	
1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16
SURF									
2014-1986		2014-1976		1994-1986		1994-1976		1986-1976	
1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16
ASURF									
2022-2014		2022-1994		2022-1986		2022-1976		2014-1994	
1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16
ASURF									
2014-1986		2014-1976		1994-1986		1994-1976		1986-1976	
1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16

Rys. 9.9 Rozkład punktów przejściowych (w podziale tailowym – 16) sekwencji (2022–2014, 2022–1994, 2022–1986, 2022–1976, 2014–1994, 2014–1986, 2014–1976, 1994–1986, 1994–1976, 1986–1976) bloków fotogrametrycznych (true-ortofotomap) z uwzględnieniem algorytmów HC (SURF i ASURF) podejścia *Detect-then-Describe* (kolor zielony oznacza wystarczającą liczbę punktów w danym podobszarze, natomiast kolor czerwony oznacza ich brak)

Punkty wykryte za pomocą algorytmu ASURF (rys. 9.8) położone były na tych samych elementach co w przypadku algorytmu SURF, z tą różnicą, że liczba wykrytych punktów była pięciokrotnie większa. Największa liczba punktów wyznaczonych przez detektor AFAST oraz najbardziej równomierny ich rozkład wystąpiły w sekwencji czasowej 2022–2014 (zdjęcia cyfrowe). W sekwencjach utworzonych na podstawie bloku 2022 i bloków 1994, 1986, 1976, a także bloku 2014 z blokami 1994, 1986, 1976, widoczna jest tendencja spadkowa liczby punktów przejściowych oraz mniej równomierny ich rozkład (rys. 9.8).

Podobnie jak w przypadku wcześniejszych algorytmów, kluczowy wpływ na ten wynik miały zmiany urbanizacyjne zachodzące na danym obszarze. Im większy odstęp czasowy między dopasowywanymi blokami, tym większe zmiany miały miejsce w tym czasie, co bezpośrednio prowadzi do obniżenia liczby wyznaczanych punktów przejściowych przez algorytm oraz pogorszenia ich rozmieszczenia. W sekwencji 2022–1994 w części wspólnego obszaru dopasowania obejmującej teren ogródków działkowych i użytków rolnych brak jest wykrytych punktów.

Porównanie sekwencji czasowych 2022–1994 (odstęp czasowy między pozyskaniem danych: 28 lat) oraz 2022–1986 (odstęp: 36 lat) wykazało, że w przypadku sekwencji 2022–1986 algorytm wykrył punkty przejściowe w tej samej części obszaru dopasowania, pomimo większego odstępu czasowego pomiędzy zestawami danych. Prawdopodobną przyczyną tego jest większa skala zdjęć z 1986 roku (1:3 000) w porównaniu do skali zdjęć z 1994 roku (1:5 000), co zwiększyło poziom szczegółowości obrazów i ułatwiło detektorowi identyfikację większej liczby punktów w tym obszarze.

Dla sekwencji czasowych utworzonych na podstawie zdjęć cyfrowych i analogowych najmniej punktów wyznaczono w sekwencjach 2022–1976, 2014–1976 ze względu na dużą różnicę czasową pozyskanych danych jak również ze względu na słabą jakość radiometryczną zdjęć w bloku z 1976 r. W przypadku sekwencji czasowych utworzonych na podstawie zdjęć analogowych najwięcej punktów wyznaczono w sekwencjach 1994–1986, co jest efektem krótkiego odstępu czasowego między pozyskaniem danych oraz dużej skali zdjęć. Natomiast najmniejszą liczbę punktów algorytm AFAST wyznaczył dla sekwencji czasowych 1994–1976, a także sekwencji 1986–1976 utworzonych zdjęć analogowych, co jest konsekwencją niskiej jakości radiometrycznej zdjęć z 1976 roku.

Na podstawie analizy wyników dopasowania sekwencji bloków fotogrametrycznych (rys. 9.9) w podejściu *Detect-then-Describe* z wykorzystaniem algorytmu AFAST stwierdzono, że możliwe jest wykonanie orientacji czasowej (przeniesienie orientacji z jednego bloku na drugi) dla sekwencji 2022–2014, 2022–1994, 2022–1986, 2014–1994, 2014–1986 oraz 1994–1986. W przypadku pozostałych sekwencji 2022–1976, 2014–1976, 1994–1976, 1986–1976, z powodu małej liczby wyznaczonych punktów przejściowych przez algorytm AFAST oraz braku ich równomiernego rozkładu, nie było możliwe wykonanie orientacji bloków fotogrametrycznych.

Podsumowując wyniki analiz rozkładu wyznaczonych punktów przejściowych w sekwencjach bloków fotogrametrycznych, można stwierdzić, że:

1. Największa liczba punktów przejściowych została wyznaczona w sekwencji 2022–2014, głównie dzięki krótkiemu odstępowi czasowemu (8 lat) i minimalnym zmianom urbanizacyjnym oraz wysokiej jakości zdjęć lotniczych wykonanych kamerą cyfrową o dobrej radiometrii.
2. Detektor AFAST osiągnął najlepsze wyniki, wykrywając najwięcej punktów w większości analizowanych sekwencji czasowych. Jego skuteczność wynika z zastosowania afinizmu, który zwiększa liczbę wykrywanych punktów.
3. Detektor SURF okazał się najmniej efektywny w detekcji punktów przejściowych, szczególnie dla zdjęć starszych i o słabszej jakości radiometrycznej, co wynikało z jego ograniczonej odporności na zmiany radiometryczne.
4. Liczba wykrywanych punktów spada wraz ze wzrostem różnicy czasowej między porównywanymi danymi. Jest to wynikiem zmian urbanizacyjnych (np. budowa dróg, osiedli, przekształcenie terenów rolnych) oraz różnic w jakości radiometrycznej zdjęć archiwalnych tworzących bloki.
5. Najmniejsza liczba punktów przejściowych została wyznaczona w sekwencjach obejmujących rok 1976, co wynika z małej skali zdjęć (1:18 500) i słabej jakości radiometrycznej oraz znacznych różnic w pokryciu terenu w stosunku do nowszych zdjęć.
6. Brak punktów w niektórych obszarach (np. ogródki działkowe, tereny rolnicze) wynikał z niskiego kontrastu (jednolita powierzchnia) i braku charakterystycznych cech (brak unikalnych i dobrze zdefiniowanych struktur, takich jak krawędzie budynków) wykrywanych przez algorytmy.
7. Rozkład punktów zależał od algorytmu detekcji. Detektory FAST i AFAST wykrywały punkty głównie na dachach budynków, znakach drogowych i chodnikach, natomiast SURF i ASURF znajdowały punkty głównie na powierzchniach dachów i dróg. Detektor AFAST wykrywał od 3 do 6 razy więcej punktów niż FAST, a detektor ASURF wykrywał 5 razy więcej punktów niż SURF.

8. Najlepsze wyniki uzyskano dla sekwencji bloków 2022–2014, a skuteczność detekcji punktów przejściowych była silnie uzależniona od małej różnicy czasowej między zdjęciami i ich jakości radiometrycznej.

9.1.3 Ocena dokładności wyznaczenia punktów przejściowych

Po etapie filtracji geometrycznej 2D punktów wiążących w procesie transformacji homograficznej przystąpiono do wykonania trzeciej analizy odnoszącej się do

Trzecia analiza dopasowania sekwencyjnego bloków fotogrametrycznych w podejściu *Detect-then-Describe* z wykorzystaniem algorytmów HC/HC polegała na ocenie geometrycznej dokładności pomiaru punktów przejściowych pomiędzy sekwencjami czasowymi bloków fotogrametrycznych. Ocena ta przedstawiona została w formie rozkładu błędów pomiaru (dopasowania) punktów przejściowych w sekwencjach bloków fotogrametrycznych z roku 1976, 1986, 1994, 2014 i 2022.

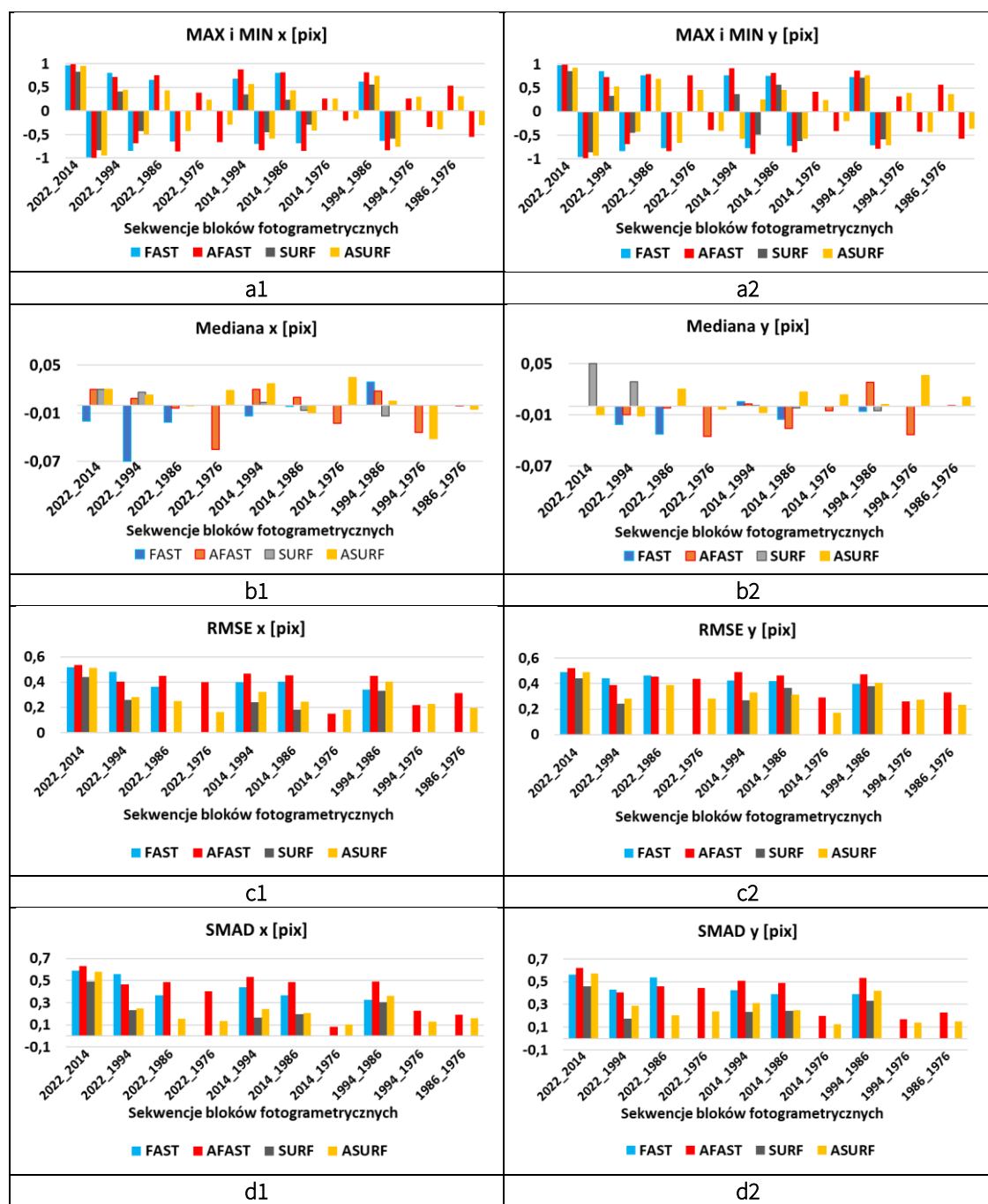
W celu zapewnienia poprawnej transformacji bloków dla wszystkich detektorów przeprowadzono filtrację geometryczną punktów przejściowych, której celem było usunięcie punktów odstających. Fakt ten znajduje potwierdzenie w rozkładzie wartości maksymalnych i minimalnych błędów pomiaru, przedstawionym na rysunku 9.10 a1 i a2.

Analiza wartości maksymalnych i minimalnych błędów pomiarowych punktów przejściowych (x, y) we wszystkich sekwencjach oraz dla wszystkich detektorów wykazała, że mieszczą się one w przedziale ± 1 piksela, co pozostaje zgodne z przyjętym progiem filtracji. W przypadku sekwencji bloków obejmujących rok 1976 (tj. 2022–1976, 2014–1976, 1994–1976, 1986–1976) wartości błędów minimalnych i maksymalnych były mniejsze niż 1 piksel i oscylowały wokół 0.5 piksela.

Sekwencja 2022–2014, charakteryzująca się największą liczbą punktów przejściowych, wykazuje błędy maksymalne i minimalne zgodne z ustalonym progiem filtracji.

Wartości mediany błędu wyznaczenia punktów przejściowych przez algorytmy HC dla współrzędnych x i y (rysunek 9.10 b1, b2) we wszystkich sekwencjach jest bliska zero (od -0.07 do 0.04 piksela), co świadczy o braku składowej systematycznej błędów.

Interpretacja wartości RMSE wykazała, że dla wszystkich sekwencji wartość tego błędu nie przekraczała 0.6 piksela (rysunek 9.10 c1 i c2).



Rys. 9.10 Zestawienie wartości błędów (Maksimum, Minimum – a1, a2, Mediana – b1, b2, RMSE – c1, c2, SMAD – d1, d2) pomiaru punktów przejściowych w sekwencjach bloków fotogrametrycznych z roku 1976, 1986, 1994, 2014 i 2022 przy wykorzystaniu algorytmów HC/HC stosowanych w podejściu *Detect-then-Describe*

W sekwencjach bloków 2022–1994, 2022–1986, 2014–1994 oraz 2014–1986, algorytmy SURF i ASURF osiągnęły wartości błędu średniokwadratowego (RMSE) od 50% do 80% mniejsze w porównaniu do algorytmów punktowych FAST i AFAST, co wynikało z obecnych w tych sekwencjach istotnych różnic jakości radiometrycznej pomiędzy obrazami cyfrowymi i analogowymi. SURF oraz jego afiniczna wersja ASURF to algorytmy wykorzystujące analizę

cech powierzchniowych (*blob detection*), co pozwala im na wykrywanie bardziej stabilnych i powtarzalnych struktur, mniej podatnych na zmiany radiometryczne. W przeciwieństwie do nich algorytmy FAST i AFAST są wrażliwe na zmiany oświetlenia, kontrastu oraz degradację radiometryczną.

W sekwencji 2022–2014 (zdjęcia cyfrowe) oraz w sekwencji 1994–1986 (zdjęcia analogowe), gdzie różnice radiometryczne pomiędzy blokami były niewielkie, wartości RMSE uzyskane dla wszystkich zastosowanych algorytmów były na zbliżonym poziomie. Wynik ten potwierdza, że przy małych zmianach radiometrycznych algorytmy zachowują stabilność i wykazują porównywalną skuteczność.

Niewielka różnica w wartościach RMSE, obserwowana na korzyść sekwencji utworzonej ze zdjęć cyfrowych w porównaniu do sekwencji utworzonej ze zdjęć analogowych (1994–1986), wynika ze słabszej jakości radiometrycznej zdjęć analogowych. Wartość uzyskanych błędów RMS przez algorytmy HC w podejściu *Detect-then-Describe* świadczy o poprawności pomiaru punktów przejściowych w poszczególnych sekwencjach czasowych stanowiąc tym samym solidną podstawę do późniejszej transformacji 3D bloków fotogrametrycznych.

Wartości błędów SMAD (rys. 9.10 d1, d2) uzyskanych w wyniku pomiaru punktów przejściowych przez algorytmy HC we wszystkich sekwencjach dla obu współrzędnych były zbliżone do wartości błędów RMS, co potwierdza brak występowania błędów grubych i systematycznych w pomiarze.

Podsumowując przeprowadzone badania można wysunąć następujące wnioski:

1. Brak błędów systematycznych.

Analiza mediany wartości błędów wyznaczonych przez algorytmy HC dla współrzędnych x i y potwierdziła brak błędów systematycznych w pomiarze.

2. Porównanie skuteczności detektorów.

Wartości RMSE we wszystkich sekwencjach nie przekraczały 0.6 piksela, co potwierdza wysoką jakość ich dopasowania.

3. Wpływ różnic radiometrycznych na dokładność algorytmów.

W sekwencjach bloków 2022–1994, 2022–1986, 2014–1994 oraz 2014–1986, gdzie występowały znaczne różnice radiometryczne pomiędzy zdjęciami cyfrowymi i analogowymi, wartość RMSE uzyskana dla algorytmów SURF i ASURF była od 50% do

80% niższa niż dla algorytmów punktowych FAST i AFAST. Świadczy to o większej odporności detektorów opartych na analizie cech powierzchniowych na zmiany radiometryczne w porównaniu do algorytmów punktowych.

4. Stabilność algorytmów przy niewielkich zmianach radiometrycznych.

W sekwencjach 2022–2014 (zdjęcia cyfrowe) oraz 1994–1986 (zdjęcia analogowe), gdzie różnice radiometryczne były niewielkie, wartości RMSE dla wszystkich algorytmów były na zbliżonym poziomie, co wskazuje na ich stabilność i porównywalną skuteczność przy niewielkich zmianach radiometrycznych.

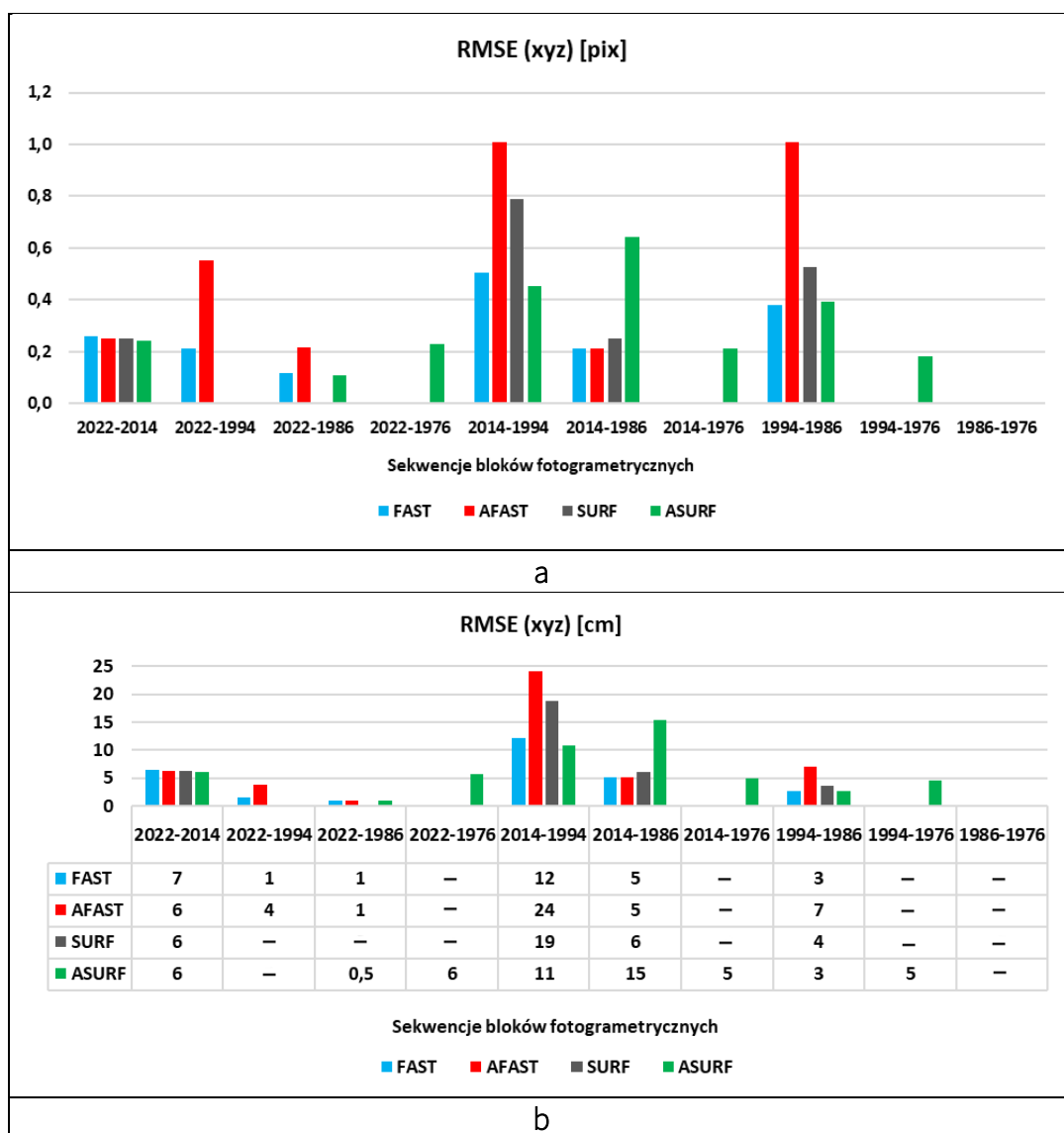
5. Wpływ jakości zdjęć analogowych na RMSE.

Niewielkie różnice w wartościach RMSE, obserwowane na korzyść sekwencji ze zdjęć cyfrowych w porównaniu do sekwencji ze zdjęć analogowych (1994–1986), wynikają z niższej jakości radiometrycznej zdjęć analogowych.

6. Otrzymane wartości błędów potwierdzają, że utworzone zbiory punktów przejściowych stanowią podstawę do transformacji 3D bloków fotogrametrycznych.

9.1.4 Ocena dokładności orientacji bloków fotogrametrycznych

Analiza sekwencyjnej orientacji bloków fotogrametrycznych z wykorzystaniem algorytmów HC/HC odnosiła się do dokładności orientacji bloków realizowanej na podstawie punktów przejściowych COP. Oceny dokładności dokonano na podstawie pomiarów RMSE (xyz) na niezależnych punktach kontrolnych CHOP, a wyniki przedstawiono zarówno w jednostkach pikselowych (rys. 9.11a), jak i terenowych (rys. 9.11b).



Rys. 9.11 Wartości RMSE sekwencyjnej orientacji bloków fotogrametrycznych (1976, 1986, 1994, 2014, 2022) wyznaczone na punktach kontrolnych CHOP, przy zastosowaniu algorytmów HC/HC w podejściu *Detect-then-Describe* w miarach pikselowych (a) oraz w miarach terenowych (b)

Na rysunku 9.11 przedstawione zostały wartości RMSE (xyz) sekwencyjnej orientacji bloków fotogrametrycznych utworzonych z archiwalnych zdjęć lotniczych z roku 2022, 2014, 1994, 1986 i 1976, uzyskanych na punktach CHOP (10% wszystkich punktów przejściowych – patrz podrozdział 5.3) wyznaczonych przez detektory HC (FAST, AFAST, SURF i ASURF). Należy zaznaczyć, że wartości GSD odpowiadająca 1 pikselowi obrazu poszczególnych true-ortofotomap wynosiła:

- rok 2022 – 5cm,
- rok 2014 – 24 cm,

- rok 1994 – 7 cm,
- rok 1986 – 4 cm,
- rok 1976 – 25 cm.

Na rysunku 9.11 symbolem „–” oznaczono przypadki braku orientacji bloku w sekwencji, co było spowodowane niewystarczającą liczbą wyznaczonych punktów oraz ich nierównomiernym rozkładem. Wpływ na to miało działanie poszczególnych algorytmów, jak i błędy pomiarowe współrzędnej (z) tych punktów na Numerycznym Modelu Powierzchni Terenu (NMPT).

Dokładność transformacji bloku 2014 do bloku 2022, w oparciu o wyznaczone punkty przejściowe dla algorytmów FAST, AFAST, SURF, ASURF wyniosła około 0.25 piksela, co odpowiada około 6 cm. Wynik ten potwierdza poprawność orientacji bloków.

W sekwencji 2022–1994 najmniejszy błąd orientacji uzyskano dla algorytmu FAST (0.2 piksela – 1.5 cm), natomiast w sekwencji 2022–1986 najmniejszy błąd osiągnięto dla algorytmu ASURF (0.1 piksela – 0.4 cm). Porównując sekwencje 2022–2014, 2022–1994 i 2022–1986 widać, że wraz ze stopniowym zmniejszeniem wielkości GSD true–ortofotomapy błąd orientacji maleje (od 0.25 piksela do 0.1 piksela).

Porównując sekwencję bloku 2022–1976 (dla 1976 GSD 25 cm) i sekwencję bloku 2022–2014 (dla 2014 GSD 24 cm) stwierdzono, że w obu przypadkach błąd orientacji wynosił 0.25 piksela. Potwierdza to zależność, że im większa wartość GSD dopasowywanego obrazu, tym wyższa wartość RMSE.

W analizie sekwencji 2014–1994 oraz 2014–1986 zaobserwowano 2-3-krotny spadek wartości błędu orientacji dla detektorów FAST, AFAST i SURF na korzyść sekwencji 2014–1986. Wyjątek stanowi detektor ASURF, gdzie błąd był większy o 0.2 piksela w przypadku sekwencji 2014–1994. Tendencja ta wynika z mniejszej wartości GSD dla true-ortofotomapy z 1986 roku (4 cm).

W większości analizowanych sekwencji wartość RMSE dla detektora FAST była wyższa niż dla detektora AFAST. Przyczyną tego zjawiska może być fakt, że algorytm AFAST generuje większą liczbę punktów charakterystycznych, jednak ich stabilność jest niższa, co skutkuje mniejszą dokładnością orientacji bloków.

W sekwencjach 2022–2014 i 2022–1976 o podobnych wielkościach GSD (2014 – 24 cm, 1976 – 25 cm) wielkości błędów były zbliżone (0.2 piksela). W sekwencji 2014–1994 najłabszym detektorem (o największym RMSE) okazał się detektor AFAST i SURF (RMSE około 1 piksel). W sekwencji 1994–1986 najłabszym detektorem (o największym RMSE) okazał się detektor AFAST. Analizując sekwencje bloku tworzonego z danych 1976 roku (2022–1976, 2014–1976, 1994–1976) punkty przejściowe wyznaczone zostały tylko dla detektora ASURF, a jego błąd orientacji wyniósł około 0.2 piksela. Najłabszym detektorem okazał się SURF, dla którego, ze względu na niewystarczającą liczbę punktów przejściowych, nie było możliwe wykonanie orientacji w sekwencjach 2022–1994, 2022–1986, 2022–1976, 2014–1976, 1994–1976 oraz 1986–1976. W przypadku sekwencji 1986–1976 żaden z zastosowanych algorytmów nie umożliwił wyznaczenia orientacji z powodu braku dostatecznej liczby punktów przejściowych. Przyczyną tego była słaba jakość radiometryczna zdjęć z bloku 1976.

Podsumowując można stwierdzić, że:

1. Na wartość błędu orientacji bloków wpływa rozdzielczość przestrzenna zdjęć (GSD). Im mniejsza wartość GSD, tym niższy błąd orientacji bloków. Przykładowo, w sekwencjach 2022–2014 (GSD 24 cm), 2022–1994 (GSD 7 cm) i 2022–1986 (GSD 4 cm) błąd orientacji zmniejszył się z 0.25 piksela do 0.1 piksela. W sekwencjach o podobnych wartościach GSD (np. 2022–2014 i 2022–1976, gdzie GSD wynosi odpowiednio 24 cm i 25 cm) błędy orientacji były zbliżone (0.2 piksela).
2. Najniższy błąd orientacji w sekwencji 2022–1994 (1.5 cm) oraz w sekwencji 2022–1986 (1.0 cm) uzyskano przy wykorzystaniu punktów wykrytych za pomocą algorytmu FAST. AFAST w większości sekwencji miał wyższe wartości RMSE niż FAST, szczególnie w sekwencji 2014–1994 (24.0 cm) i 1994–1986 (6.9 cm). Był najmniej skutecznym detektorem w tych przypadkach. Przyczyną tego zjawiska może być fakt, że algorytm AFAST generuje więcej punktów, jednak są one mniej stabilne, co prowadzi do obniżonej dokładności orientacji bloków.

Detektor SURF okazał się najmniej skuteczny — ze względu na niewystarczającą liczbę punktów charakterystycznych niemożliwe było przeprowadzenie orientacji w sekwencjach 2022–1994, 2022–1986, 2022–1976, 2014–1976, 1994–1976 oraz 1986–

1976, co wskazuje na jego ograniczoną efektywność w przypadku dużych różnic czasowych pomiędzy zdjęciami oraz w przypadkach słabej ich jakości radiometrycznej. ASURF był najbardziej uniwersalnym detektorem, radzącym sobie nawet w trudnych warunkach (np. w sekwencjach z 1976 roku). W sekwencji 2022–1986 osiągnął najmniejszy błąd (0.4 cm). Jednak w sekwencji 2014–1986 miał wyższy błąd niż inne detektory (15.4 cm).

3. Punkty przejściowe dla danych z 1976 roku (GSD 25 cm) zostały wyznaczone tylko przez detektor ASURF. W sekwencjach 2022–1976, 2014–1976 i 1994–1976 ASURF osiągnął błąd orientacji około 0.2 piksela, co potwierdza jego skuteczność w przypadku danych o niskiej rozdzielczości.
4. FAST sprawdził się dobrze w sekwencjach z mniejszym GSD, ale miał problemy z danymi o większym GSD.
5. Dla detektorów AFAST i SURF uzyskano zmienne wyniki dokładności orientacji, zależne od sekwencji, ale generalnie były mniej skuteczne niż FAST i ASURF.
6. Najłabszym detektorem okazał się SURF, który dla wszystkich sekwencji zawierających blok z 1976 roku nie wyznaczył orientacji ze względu na brak dostatecznej liczby punktów przejściowych wynikający ze słabej radiometrii tych zdjęć.
7. Skuteczność detektorów zależy od jakości danych (GSD), a ASURF okazał się najbardziej niezawodnym detektorem w przypadku danych archiwalnych o różnej rozdzielczości.

9.1.5 Podsumowanie wyników uzyskanych przy wykorzystaniu metody *Detect-then-Describe* i algorytmów *Hand-Crafted*

W niniejszym podrozdziale przedstawiono podsumowanie wniosków dotyczących orientacji bloków fotogrametrycznych w podejściu *Detect-then-Describe* i algorytmów *Hand-Crafted*.

Czas detekcji punktów przejściowych

1. Czas automatycznego wyznaczania punktów przejściowych przez algorytmy w danej sekwencji zależy od liczby wyznaczonych punktów, na który wpływ ma rozdzielczość przestrzenna obrazu – im wyższa, tym dłuższy czas obliczeń. Zależność ta jest zgodna z teoretycznymi podstawami przetwarzania obrazów cyfrowych.

2. Rodzaj zdjęcia (cyfrowe czy analogowe) nie ma istotnego wpływu, natomiast jakość radiometryczna obrazu odgrywa istotną rolę.
3. Różnice w czasie analizy wynikają także z charakterystyki algorytmów – detektor ASURF działa najwolniej, a detektor FAST najszybciej, co jest zgodne z założeniami metodologicznymi.

Liczba i rozmieszczenie punktów przejściowych w sekwencji bloków fotogrametrycznych

1. Odstęp czasowy – Krótszy odstęp czasowy bloków tworzących sekwencję (np. 2022–2014) sprzyja większej liczbie wykrytych punktów przejściowych, natomiast ich liczba maleje wraz ze wzrostem różnicy czasowej pomiędzy blokami ze względu na zmiany urbanizacyjne i jakość radiometryczną zdjęć.
2. Najlepszy detektor – AFAST wykazał najwyższą skuteczność, a zastosowanie afinizmu umożliwiającego generowanie symulowanych obrazów przejściowych poprawiło zarówno liczbę, jak i rozmieszczenie punktów przejściowych, dlatego zaleca się wykorzystanie zmodyfikowanego algorytmu FAST. Charakteryzuje się on najwyższą skutecznością w detekcji punktów przejściowych.
3. Najmniej skuteczny detektor – SURF osiągnął najłabsze wyniki, szczególnie w przypadku starszych zdjęć o gorszej jakości radiometrycznej.
4. Wpływ jakości zdjęć – Najmniej punktów wykryto dla zdjęć z 1976 roku, co wynikało z ich małej skali a zwłaszcza słabej jakości radiometrycznej.
5. Stabilność detektorów – W sekwencjach z krótkim odstępem czasowym pomiędzy blokami (np. 1994–1986) liczba punktów była większa dla zdjęć o podobnej jakości radiometrycznej i rozdzielczości.
6. Rozkład punktów – AFAST i FAST wykrywały punkty głównie na dachach budynków, znakach drogowych i chodnikach, a SURF i ASURF na powierzchniach dachów i dróg.
7. Zdjęcia cyfrowe vs. analogowe – Zdjęcia cyfrowe o dobrej radiometrii pozwalały wykryć więcej punktów przejściowych niż zdjęcia analogowe.
8. Brak punktów w niektórych obszarach – W terenach rolniczych i ogródkach działkowych liczba punktów była niska z powodu słabego kontrastu (jednolita powierzchnia) i braku wyraźnych cech wykrywanych przez algorytmy.

Dokładność pomiaru punktów przejściowych w sekwencji bloków

1. Detektory wykazały wysoką skuteczność, a wartości RMSE nie przekraczały 0.6 piksela.
2. Algorytmy oparte na analizie cech powierzchniowych (SURF, ASURF) lepiej radziły sobie z różnicami radiometrycznymi niż algorytmy punktowe (FAST, AFAST).
3. Niewielkie różnice w wartościach RMSE, obserwowane na korzyść sekwencji ze zdjęć cyfrowych w porównaniu do analogowych, wynikają z niższej jakości radiometrycznej zdjęć analogowych.

Orientacja bloków fotogrametrycznych

1. Najlepsze wyniki orientacji uzyskano na podstawie punktów wykrytych przy wykorzystaniu algorytmu FAST w sekwencjach z małym GSD - algorytm dobrze radzi sobie z obrazami wysokiej rozdzielczości. W przypadku zdjęć o większym GSD błędy orientacji były większe, co sugeruje jego ograniczoną skuteczność dla analogowych, niskorozdzielczych obrazów. AFAST był mniej skutecznym detektorem w stosunku do algorytmu FAST, pomimo większej liczby generowanych punktów, które jednak charakteryzowały się niską stabilnością, prowadząc do obniżonej dokładności orientacji.
2. SURF okazał się najmniej skuteczny spośród analizowanych detektorów. W wielu przypadkach, zwłaszcza przy dużych różnicach czasowych między zdjęciami a także o słabej jakości radiometrycznej, nie był w stanie poprawnie wyznaczyć orientacji bloków.
3. ASURF wykazał największą wszechstronność i skuteczność, radząc sobie nawet z niską rozdzielczością zdjęć. Na podstawie punktów wykrytych przez ten algorytm uzyskano najmniejszą wartość RMSE dla sekwencji 2022–1986, jednak w innych przypadkach, np. w sekwencji 2014–1986, wartość RMSE była wyższa niż innych detektorów.

REKOMENDACJA zastosowania algorytmów *Hand-Crafted* do orientacji bloków fotogrametrycznych

Tabela 9.1 przedstawia wytyczne dotyczące oceny przydatności algorytmów HC/HC do orientacji bloków fotogrametrycznych. Kryteria oceny obejmują czas detekcji, liczbę i rozmieszczenie punktów oraz wielkość błędu transformacji. Przyjęto czterostopniową, porządkową skalę ocen: 0-2 – ocena negatywna, 3 – dostateczna, 4 – dobra, 5 – bardzo dobra.

Jeśli którekolwiek z kryteriów otrzymało ocenę negatywną (0-2), oznacza to, że dany algorytm nie jest rekomendowany do orientacji analizowanej sekwencji bloków fotogrametrycznych.

Tab. 9.1 Podsumowanie oceny kryteriów przydatności algorytmów HC/HC do orientacji sekwencji bloków fotogrametrycznych w podejściu *Detect-then-Describe* dla wszystkich okresów czasowych w skali od 0 do 5 punktów

Detektor	FAST									
Sekwencje	2022_2014	2022_1994	2022_1986	2022_1976	2014_1994	2014_1986	2014_1976	1994_1986	1994_1976	1986_1976
Czas detekcji	5	5	5	5	5	5	5	5	5	5
Liczba punktów i ich rozmieszczenie	5	5	4	0	5	4	0	5	0	0
RMSE	5	5	5	0	4	5	0	4	0	0
SUMA	15	15	14	5	14	14	5	14	5	5
Detektor	AFAST									
Sekwencje	2022_2004	2022_0994	2022_0986	2022_0976	2004_0994	2004_0986	2004_0976	0994_0986	0994_0976	0986_0976
Czas detekcji	4	4	4	4	4	4	4	4	4	4
Liczba punktów i ich rozmieszczenie	5	4	5	0	4	4	0	5	0	0
RMSE	5	4	5	0	2	5	0	2	0	0
SUMA	9	8	9	4	8	8	4	9	4	4
Detektor	SURF									
Sekwencje	2022_2004	2022_0994	2022_0986	2022_0976	2004_0994	2004_0986	2004_0976	0994_0986	0994_0976	0986_0976
Czas detekcji	4	4	4	4	4	4	4	4	4	4
Liczba punktów i ich rozmieszczenie	5	3	0	0	5	3	0	4	0	0
RMSE	5	0	0	0	3	5	0	4	0	0
SUMA	14	7	4	4	12	12	4	12	4	4
Detektor	ASURF									
Sekwencje	2022_2004	2022_0994	2022_0986	2022_0976	2004_0994	2004_0986	2004_0976	0994_0986	0994_0976	0986_0976
Czas detekcji	4	4	4	4	4	4	4	4	4	4
Liczba punktów i ich rozmieszczenie	5	3	4	3	4	4	5	4	5	0
RMSE	5	0	5	5	4	4	5	4	5	0
SUMA	14	7	13	12	12	12	14	12	14	4

rekomendowane nierekomendowane

Na podstawie oceny wpływu detektora/deskryptora na dokładność orientacji bloków fotogrametrycznych stwierdzono, że detektory/deskryptory uzyskujące wynik powyżej 11 punktów są najbardziej efektywne, a najlepszy z nich został wyróżniony pogrubieniem (tab.9.2):

Tab. 9.2 Rekomendacja wyboru algorytmu HC/HC w podejściu *Detect-then-Describe* do orientacji bloków fotogrametrycznych w danej sekwencji

Sekwencja bloków	Rodzaj zdjęcia	Rozdzielczość	Jakość radiometryczna	Najlepszy algorytm
2022 2014	cyfrowe cyfrowe	5 cm 24 cm	bardzo dobra bardzo dobra	FAST, AFAST, SURF, ASURF
2022 1994	cyfrowe analogowe	5 cm 7 cm	bardzo dobra dobra	FAST, AFAST
2022 1986	cyfrowe analogowe	5 cm 5 cm	bardzo dobra dobra	FAST, AFAST, ASURF
2022 1976	cyfrowe analogowe	5 cm 25 cm	bardzo dobra słaba	ASURF
2014 1994	cyfrowe analogowe	24 cm 7 cm	bardzo dobra dobra	FAST, SURF, ASURF
2014 1986	cyfrowe analogowe	24 cm 5 cm	bardzo dobra dobra	FAST, AFAST, SURF, ASURF
2014 1976	cyfrowe analogowe)	24 cm 25 cm	bardzo dobra słaba	ASURF
1994 1986	analogowe analogowe	7 cm 25 cm	dobra dobra	ASURF
1994 1976	analogowe analogowe	7 cm 25 cm	dobra słaba	FAST, SURF, ASURF
1986 1976	analogowe analogowe	5 cm 25 cm	dobra słaba	Brak

9.2 Ocena poprawności procesu orientacji bloków fotogrametrycznych – podejście *Detect-then-Describe* (algorytmy *Hand-Crafted* / *Learned-Based*)

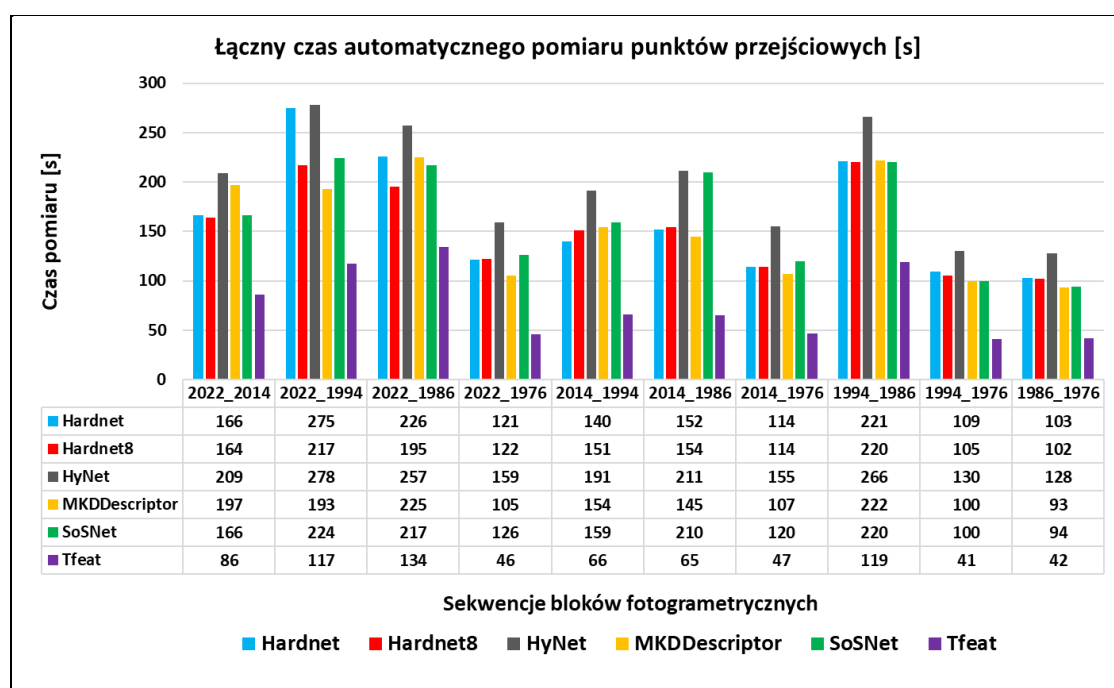
W podrozdziale 9.2 zaproponowano tę samą metodykę oceny poprawności procesu orientacji bloków fotogrametrycznych, co w podrozdziale 9.1. Obejmuje ona analizę czasu automatycznego wyznaczania punktów przejściowych, ich liczby i rozmieszczenia, dokładności ich wyznaczenia oraz dokładności orientacji bloków. Struktura opisu w niniejszym podrozdziale jest analogiczna do struktury przedstawionej w podrozdziale 9.1.

W ramach podejścia *Detect-then-Describe* zastosowano następującą kombinację algorytmów: detektor (HC): SIFT oraz deskryptory (LB): HardNet, HardNet8, SoSNet, HyNet, MKDDescriptor i TFeat.

Ocena czasu automatycznego pomiaru punktów przejściowych w sekwencyjnej orientacji bloków fotogrametrycznych

W związku z tym, że w podejściu *Detect-then-Describe* zastosowano ten sam detektor (SIFT), czas detekcji pozostawał stały, a całkowity czas wyznaczania punktów przejściowych był determinowany głównie przez zastosowany deskryptor.

Przeprowadzona analiza czasu wyznaczania punktów przejściowych wykazała, że algorytm oparty na deskrytorze TFeat był najbardziej wydajny (rys. 9.12). W każdej sekwencji osiągnięto dla niego najkrótszy czas automatycznego pomiaru, nawet dwukrotnie krótszy w porównaniu do pozostałych algorytmów.



Rys. 9.12 Łączny czas automatycznego wyznaczania punktów przejściowych przez algorytmy HC/LB w podejściu *Detect-then-Describe* w sekwencjach bloków fotogrametrycznych z roku 2022, 2014, 1994, 1986 i 1976

Z kolei algorytm HyNet charakteryzował się najdłuższym czasem wyznaczania punktów we wszystkich analizowanych sekwencjach, co czyni go najmniej korzystnym rozwiązaniem w tym aspekcie. Zauważono również, że czas potrzebny na wyznaczenie punktów przejściowych silnie zależy od rozdzielczości przestrzennej zdjęć tworzących daną sekwencję. Im wyższa

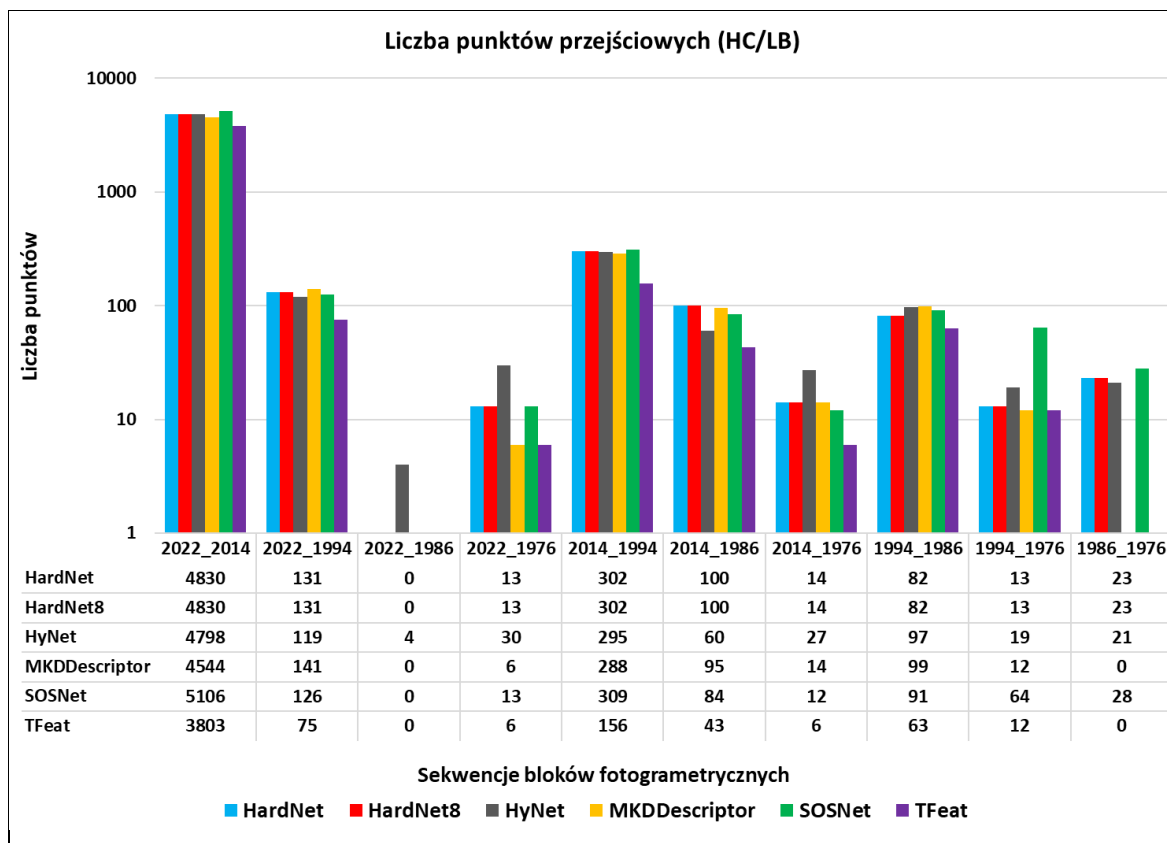
rozdzielczość, tym dłuższy czas obliczeń (np. sekwencje 2022–1994, 2022–1986, 1994–1986), natomiast przy niższej rozdzielczości czas ten ulega skróceniu (np. sekwencje 2022–1976, 2014–1976). Należy podkreślić, że do detekcji punktów charakterystycznych zastosowano ten sam algorytm, natomiast do ich opisu wykorzystano różne deskryptory. W związku z tym różnice w czasie pomiaru wynikają przede wszystkim z doboru deskryptora, który przekładał się na liczbę dopasowanych punktów, a tym samym na łączny czas potrzebny do ich dopasowania.

Dodatkowym czynnikiem wpływającym na skrócenie czasu obliczeń była niska jakość radiometryczna zdjęć z roku 1976. W przypadku algorytmów LB zaobserwowano zbliżone wartości czasu wyznaczania punktów przejściowych dla sekwencji utworzonych z obrazów o podobnej rozdzielczości przestrzennej (np. 2014–1994, 2014–1986).

Podsumowując na czas pomiaru automatycznego punktów przejściowych w utworzonych sekwencjach nie wpływa rodzaj zdjęcia (cyfrowe, analogowe), lecz przede wszystkim metoda działania algorytmu (uwzględniająca m.in. rozdzielczość wejściowego obrazu), rozdzielczość przestrzenna zdjęć (im większa tym więcej wykrywanych jest punktów, a co za tym idzie – dłuższy czas pomiaru), a także ich jakość radiometryczna. Spośród analizowanych algorytmów najdłuższy czas pomiaru odnotowano dla HyNet, natomiast najkrótszy dla TFeat.

Ocena liczby wykrytych punktów przejściowych i ich rozkładu w sekwencyjnej orientacji bloków fotogrametrycznych

Dla wszystkich zastosowanych algorytmów największą liczbę punktów przejściowych uzyskano w sekwencji 2022–2014, składającej się ze zdjęć cyfrowych (rys. 9.13). Najwyższy wynik osiągnął algorytm SOSNet, wyznaczając 5106 punktów. Różnice między algorytmami były niewielkie, co wskazuje na ich zbliżoną skuteczność.



Rys. 9.13 Liczba pomierzonych automatycznie punktów przejściowych (COP i CHOP) w poszczególnych sekwencjach bloków fotogrametrycznych (1976 r., 1986 r., 1994 r., 2011 r., 2022 r.) dla algorytmów HC/LB (SIFT/HardNet, HardNet8, SOSNet, HyNet, MKDDescriptor, TFeat) w podejściu *Detect-then-Describe*

Wraz ze wzrostem różnicy czasowej między obrazami liczba wyznaczonych punktów maleje. Dla sekwencji 2022–1986 większość algorytmów nie wyznaczyła punktów, co najprawdopodobniej wynika z intensywnych zmian urbanizacyjnych na analizowanym obszarze, prowadzących do zmniejszenia liczby wspólnych cech w obrazach i utrudniających wyznaczenie punktów.

W sekwencjach 1994–1986 oraz 1994–1976 liczba wyznaczonych punktów mieściła się w przedziale 63–99, pomimo stosunkowo krótkiego odstępu czasowego (odpowiednio 8 i 18 lat). Główną przyczyną tego zjawiska może być niższa jakość radiometryczna zdjęć analogowych w porównaniu do cyfrowych, co wpływa na skuteczność detekcji. W sekwencji 2022–2014 wyznaczono od 50 do 60 razy więcej punktów niż w sekwencji 1994–1986, co można tłumaczyć zarówno wyższą jakością zdjęć cyfrowych, jak i większymi zmianami urbanistycznymi w okresie 1994–1986, ograniczającymi liczbę wspólnych cech w obrazach. Należy również zauważyć, że różnice w liczbie wyznaczonych punktów pomiędzy

poszczególnymi algorytmami były niewielkie, co świadczy o ich podobnej skuteczności w analizowanych sekwencjach.

Wyniki dopasowania, przedstawione graficznie na rysunkach 9.14 i 9.15, obrazują rozmieszczenie punktów w poszczególnych sekwencjach bloków fotogrametrycznych z uwzględnieniem algorytmów LB w podejściu *Detect-then-Describe*. Ich celem jest szczegółowa analiza przestrzennego rozkładu punktów przejściowych. W badanej sekwencji kolor czerwony wskazuje na brak wyznaczonych punktów przejściowych lub ich niewystarczającą liczbę, natomiast kolor zielony oznacza ich odpowiednią liczbę.

2022_2014						2022_1994					
HardNet		HardNet8		HyNet		HardNet		HardNet8		HyNet	
1	2	1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16	15	16
MKDDDescriptor		SoSNet		Tfeat		MKDDDescriptor		SoSNet		Tfeat	
1	2	1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16	15	16
2022_1986						2022_1976					
HardNet		HardNet8		HyNet		HardNet		HardNet8		HyNet	
1	2	1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16	15	16
MKDDDescriptor		SoSNet		Tfeat		MKDDDescriptor		SoSNet		Tfeat	
1	2	1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16	15	16
2014_1994											
HardNet		HardNet8		HyNet							
1	2	1	2	1	2						
3	4	3	4	3	4						
5	6	5	6	5	6						
7	8	7	8	7	8						
9	10	9	10	9	10						
11	12	11	12	11	12						
13	14	13	14	13	14						
15	16	15	16	15	16						
MKDDDescriptor		SoSNet		Tfeat							
1	2	1	2	1	2						
3	4	3	4	3	4						
5	6	5	6	5	6						
7	8	7	8	7	8						
9	10	9	10	9	10						
11	12	11	12	11	12						
13	14	13	14	13	14						
15	16	15	16	15	16						

Rys. 9.14 Rozkład punktów przejściowych (w podziale tailowym – 16) sekwencji (2022–2014, 2022–1994, 2022–1986, 2022–1976, 2014–1994) bloków fotogrametrycznych (true-ortofotomap) z uwzględnieniem algorytmów LB (HardNet, HardNet8, HyNet, MKDDDescriptor, SoSNet, TFeat) podejścia *Detect-then-Describe* (kolor zielony oznacza wystarczającą liczbę punktów w danym podobzszarze, natomiast kolor czerwony oznacza ich brak)

2014_1986						2014_1976					
HardNet		HardNet8		HyNet		HardNet		HardNet8		HyNet	
1	2	1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16	15	16
MKDDDescriptor		SoSNet		Tfeat		MKDDDescriptor		SoSNet		Tfeat	
1	2	1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16	15	16
1994_1986						1994_1976					
HardNet		HardNet8		HyNet		HardNet		HardNet8		HyNet	
1	2	1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16	15	16
MKDDDescriptor		SoSNet		Tfeat		MKDDDescriptor		SoSNet		Tfeat	
1	2	1	2	1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16	15	16	15	16
1986_1976											
HardNet		HardNet8		HyNet							
1	2	1	2	1	2						
3	4	3	4	3	4						
5	6	5	6	5	6						
7	8	7	8	7	8						
9	10	9	10	9	10						
11	12	11	12	11	12						
13	14	13	14	13	14						
15	16	15	16	15	16						
MKDDDescriptor		SoSNet		Tfeat							
1	2	1	2	1	2						
3	4	3	4	3	4						
5	6	5	6	5	6						
7	8	7	8	7	8						
9	10	9	10	9	10						
11	12	11	12	11	12						
13	14	13	14	13	14						
15	16	15	16	15	16						

Rys. 9.15 Rozkład punktów przejściowych (w podziale tailowym – 16) sekwencji (2014–1986, 2014–1976, 1994–1986, 1994–1976, 1986–1976) bloków fotogrametrycznych (true-ortofotomap) z uwzględnieniem algorytmów LB (HardNet, HardNet8, HyNet, MKDDDescriptor, SoSNet, TFeat) podejścia Detect-then-Describe (kolor zielony oznacza wystarczającą liczbę punktów w danym podobszarze, natomiast kolor czerwony oznacza ich brak)

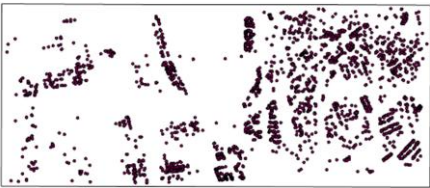
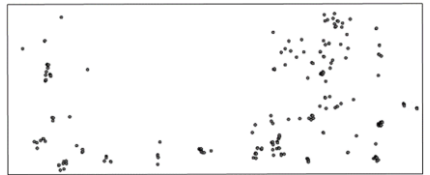
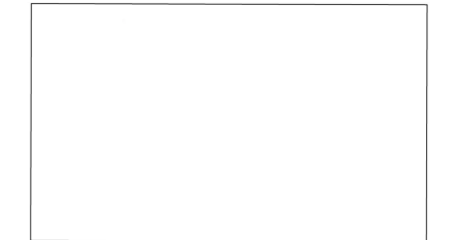
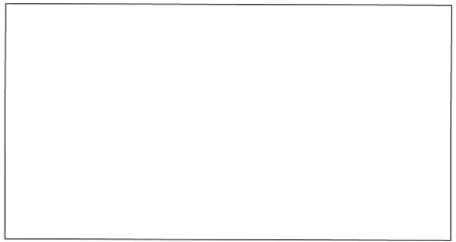
W podejściu *Detect-then-Describe* HB/LB kluczową rolę w procesie wykrywania punktów w poszczególnych sekwencjach odgrywał detektor SIFT. Istotny wpływ na wyznaczanie

punktów przejściowych miało również zastosowanie deskryptorów LB, odpowiedzialnych za opis wykrytych punktów. Efektywność wykrywania punktów przejściowych znacząco zmniejszała się wraz ze wzrostem różnicy czasowej między blokami fotogrametrycznymi (rys. 9.16 - 9.21). W starszych blokach, zwłaszcza z lat 1976 i 1986, jakość radiometryczna oraz skala zdjęć wpływały negatywnie na skuteczność algorytmów, co skutkowało brakiem wyznaczonych punktów i uniemożliwiało przeprowadzenie orientacji czasowej. Najlepsze wyniki uzyskano dla sekwencji 2022–2014, gdzie liczba punktów oraz ich rozmieszczenie były równomierne i wystarczające do przeprowadzenia orientacji czasowej.

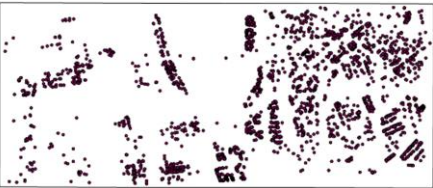
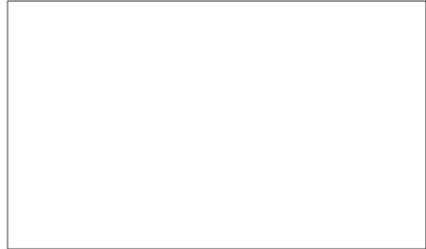
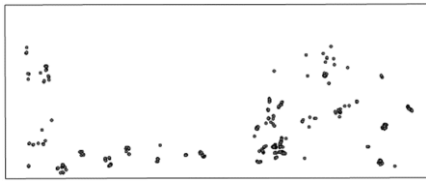
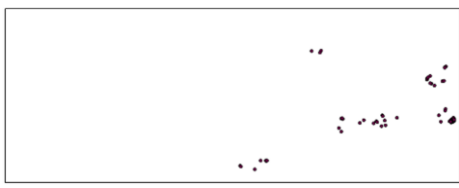
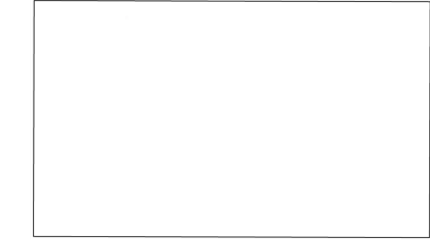
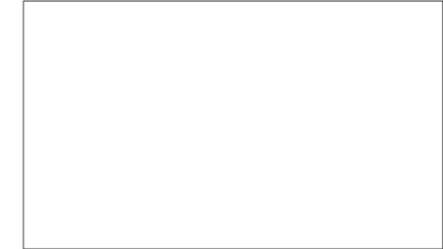
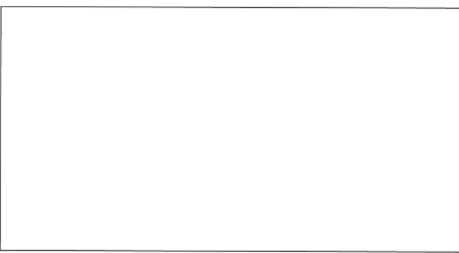
Sekwencje zawierające zdjęcia z roku 1976 charakteryzowały się niemal całkowitym brakiem punktów przejściowych, co wykluczało możliwość dopasowania bloków. W przypadku sekwencji 2014–1976 jedynie przy udziale algorytmu HyNet możliwe było wyznaczenie punktów charakterystycznych, co świadczy o jego przewadze wynikającej z mechanizmu opisu cech. W zakresie zastosowania poszczególnych algorytmów zaobserwowano, że HardNet i HardNet8 wykazywały identyczne wyniki, co oznacza, że oba algorytmy cechują się podobną skutecznością. Algorytm HyNet jako jedyny umożliwił dopasowanie sekwencji 2014–1976, świadcząc o jego wyższej odporności na zmiany jakości radiometrycznej zdjęć.

MKDDDescriptor i SOSNet różniły się w zakresie wyznaczania punktów przejściowych w sekwencjach 2022–1994 oraz 1994–1976, co wskazuje na unikalne cechy opisu punktów charakterystycznych przez te deskryptory. Wyniki uzyskane przez algorytm TFeat były zbliżone do wyników algorytmu HardNet, jednak liczba wyznaczonych punktów była nieco mniejsza. Na podstawie analizy dopasowania sekwencji bloków fotogrametrycznych możliwe było przeprowadzenie orientacji czasowej dla sekwencji 2022–2014, 2022–1994 oraz 2014–1994. Dla pozostałych sekwencji orientacja bloków była niemożliwa ze względu na brak punktów przejściowych lub ich nierównomierne rozmieszczenie.

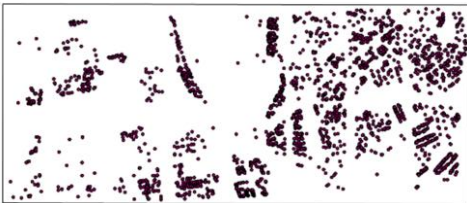
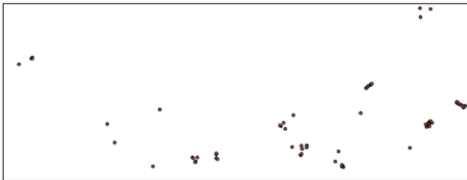
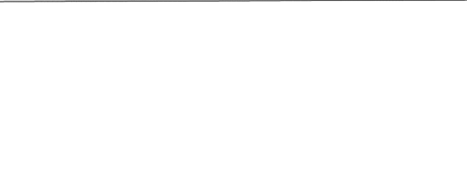
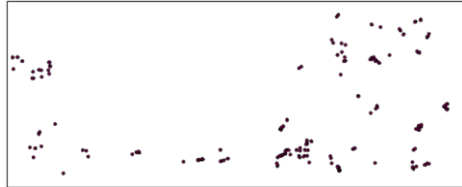
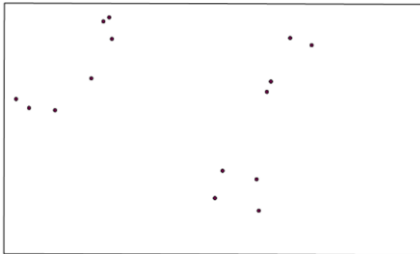
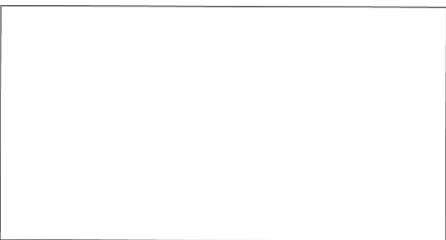
Podsumowując, skuteczność algorytmów w podejściu *Detect-then-Describe* jest silnie uzależniona od jakości radiometrycznej zdjęć, skali zmian urbanizacyjnych oraz doboru deskryptorów. Algorytm HyNet okazał się najbardziej odporny na zmiany radiometrii zdjęć, natomiast HardNet i jego warianty zapewniały stabilne wyniki dla sekwencji zdjęć utworzonych z nowszych danych.

Algorytm	HardNet	2022–2014		2022–1994		2022–1986		2022–1976		2014–1994	
		2014–1986		2014–1976		1994–1986		1994–1976		1986–1976	

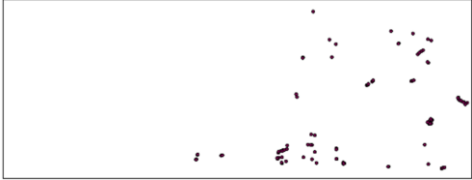
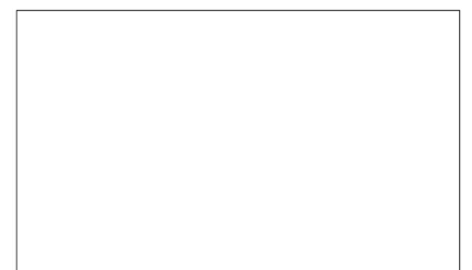
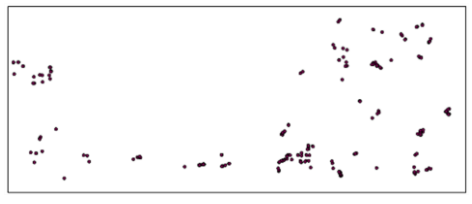
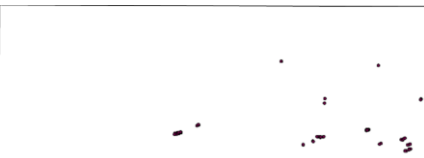
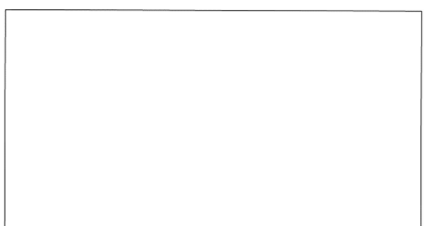
Rys. 9.16 Rozkład punktów przejściowych dla algorytmu LB (HardNet) podejścia *Detect-then-Describe* sekwencji bloków fotogrametrycznych 2022–2014, 2022–1994, 2022–1976, 2014–1994, 2014–1986, 2014–1976, 1994–1986, 1994–1976, 1986–1976

Algorytm	HardNet8	2022-2014		2022-1994		2022-1986		2022-1976		2014-1994	
		2014-1986		2014-1976		1994-1986		1994-1976		1986-1976	
		Algorytm	HardNet8	Algorytm	HardNet8	Algorytm	HardNet8	Algorytm	HardNet8	Algorytm	HardNet8

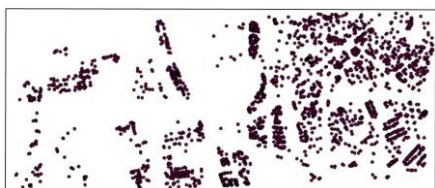
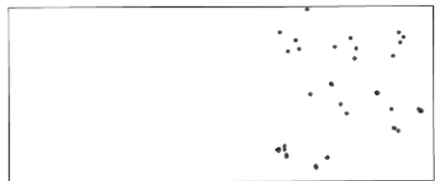
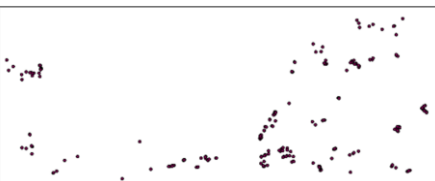
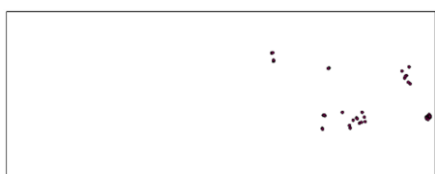
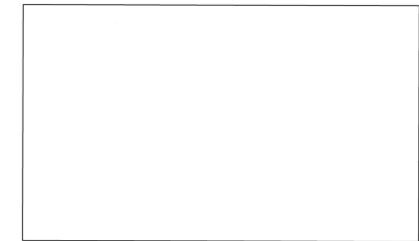
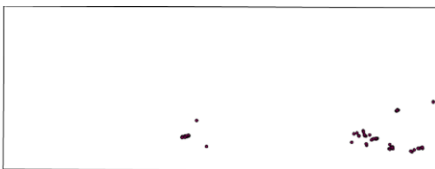
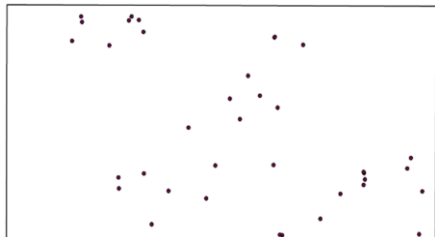
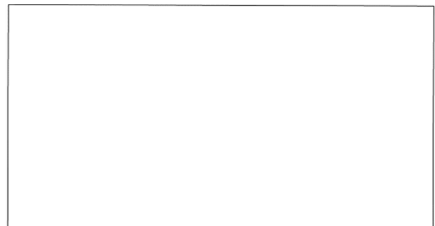
Rys. 9.17 Rozkład punktów przejściowych dla algorytmu LB (HardNet8) podejścia *Detect-then-Describe* sekwencji bloków fotogrametrycznych 2022-2014, 2022-1994, 2022-1976, 2014-1994, 2014-1986, 2014-1976, 1994-1986, 1994-1976, 1986-1976

Algorithm	2022-2014	2022-1994	2022-1986	2022-1976	2014-1994
HyNet					
	2014-1986	2014-1976	1994-1986	1994-1976	1986-1976
HyNet					

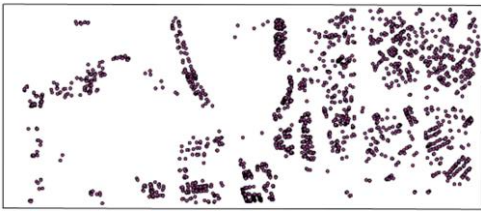
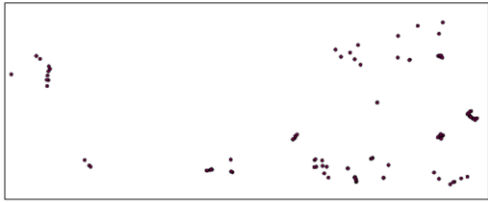
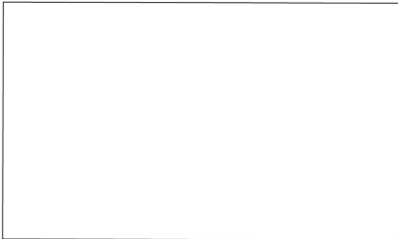
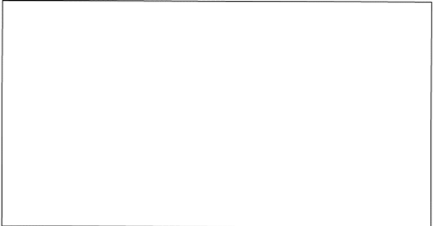
Rys. 9.18 Rozkład punktów przejściowych dla algorytmu LB (HyNet) podejścia *Detect-then-Describe* sekwencji bloków fotogrametrycznych 2022-2014, 2022-1994, 2022-1976, 2014-1994, 2014-1986, 2014-1976, 1994-1986, 1994-1976, 1986-1976

Algorytm	2022–2014	2022–1994	2022–1986	2022–1976	2014–1994
					
	MKD Descr ptor				
	2014–1986	2014–1976	1994–1986	1994–1976	1986–1976
					
Algorytm					
MKD Descr ptor					

Rys. 9.19 Rozkład punktów przejściowych dla algorytmu LB (MKDDescriptor) podejścia *Detect-then-Describe* sekwencji bloków fotogrametrycznych 2022–2014, 2022–1994, 2022–1976, 2014–1994, 2014–1986, 2014–1976, 1994–1986, 1994–1976, 1986–1976

Algorytm	SOSNet	2022-2014		2022-1994		2022-1986		2022-1976		2014-1994	
		Algorytm	SOSNet								
Algorytm	SOSNet	2014-1986		2014-1976		1994-1986		1994-1976		1986-1976	
		Algorytm	SOSNet								

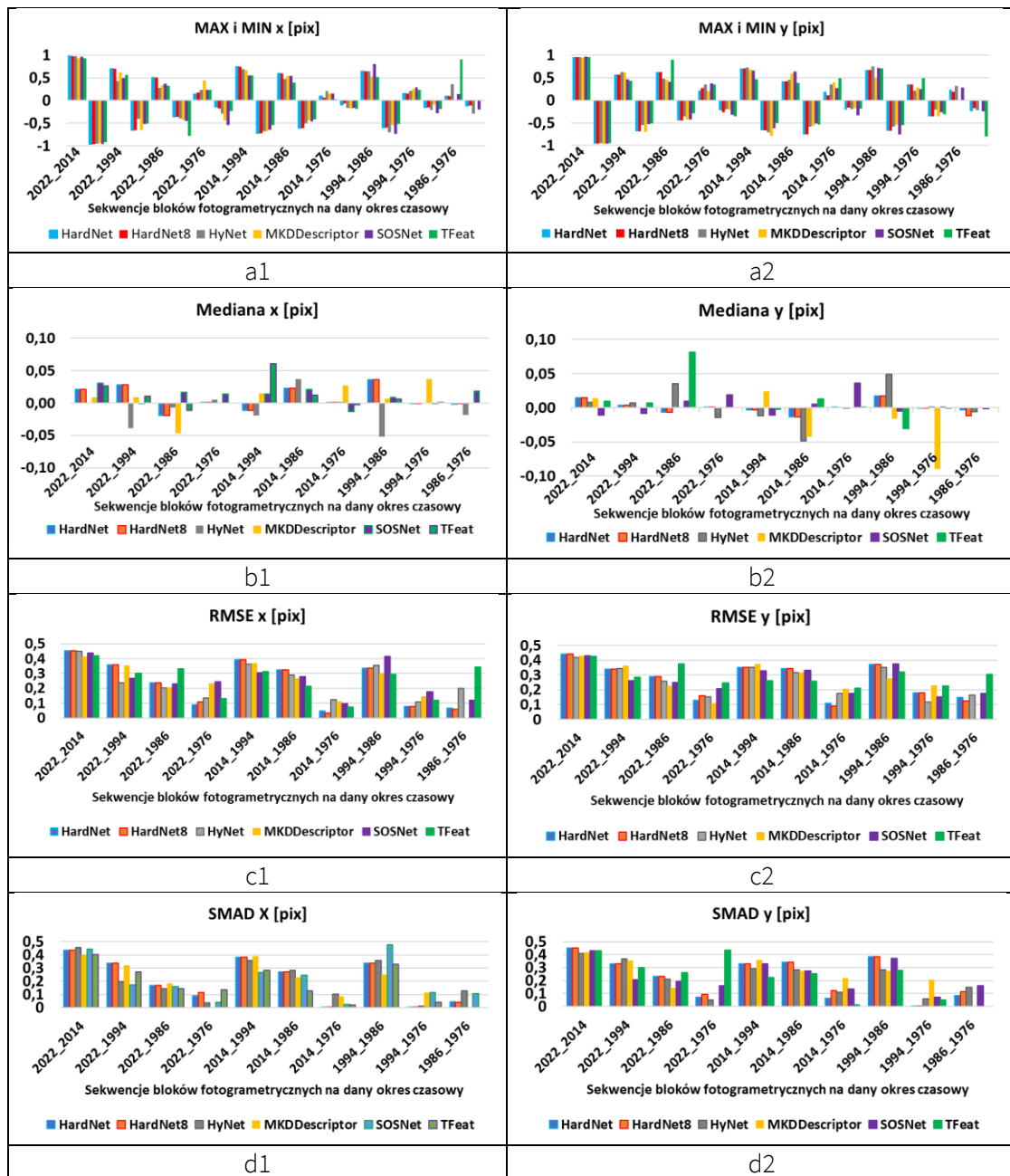
Rys. 9.20 Rozkład punktów przejściowych dla algorytmu LB (SOSNet) podejścia *Detect-then-Describe* sekwencji bloków fotogrametrycznych 2022-2014, 2022-1994, 2022-1976, 2014-1994, 2014-1986, 2014-1976, 1994-1986, 1994-1976, 1986-1976

Algorytm TFeat	2022-2014		2022-1994		2022-1986		2022-1976		2014-1994	
	2014-1986		2014-1976		1994-1986		1994-1976		1986-1976	

Rys. 9.21 Rozkład punktów przejściowych dla algorytmu LB (TFeat) podejścia *Detect-then-Describe* sekwencji bloków fotogrametrycznych 2022-2014, 2022-1994, 2022-1976, 2014-1994, 2014-1986, 2014-1976, 1994-1986, 1994-1976, 1986-1976

Ocena dokładności wyznaczenia punktów przejściowych

Na rysunku 9.22 zostały przedstawione wartości błędów pomiaru punktów przejściowych przy użyciu sieci neuronowych. Przeprowadzona analiza dopasowania sekwencyjnego bloków fotogrametrycznych w podejściu *Detect-then-Describe* wykazała, że zastosowane algorytmy HC/LB w większości sekwencji zapewniają poprawność pomiaru punktów przejściowych. Filtracja punktów przejściowych skutecznie eliminowała punkty o błędach grubych, co wpłynęło na poprawę jakości dopasowania.



Rys. 9.22 Zestawienie wartości błędów (Maksimum, Minimum, Mediana, SMAD, SIGMA, RMSE) pomiaru punktów przejściowych (podejście *Detect-then-Describe*) przy użyciu sieci neuronowych (HardNet, HardNet8, SOSNet, HyNet, MKDDescriptor, TFeat) dla par true-ortofotomap z roku 1976, 1986, 1994, 2014 i 2022

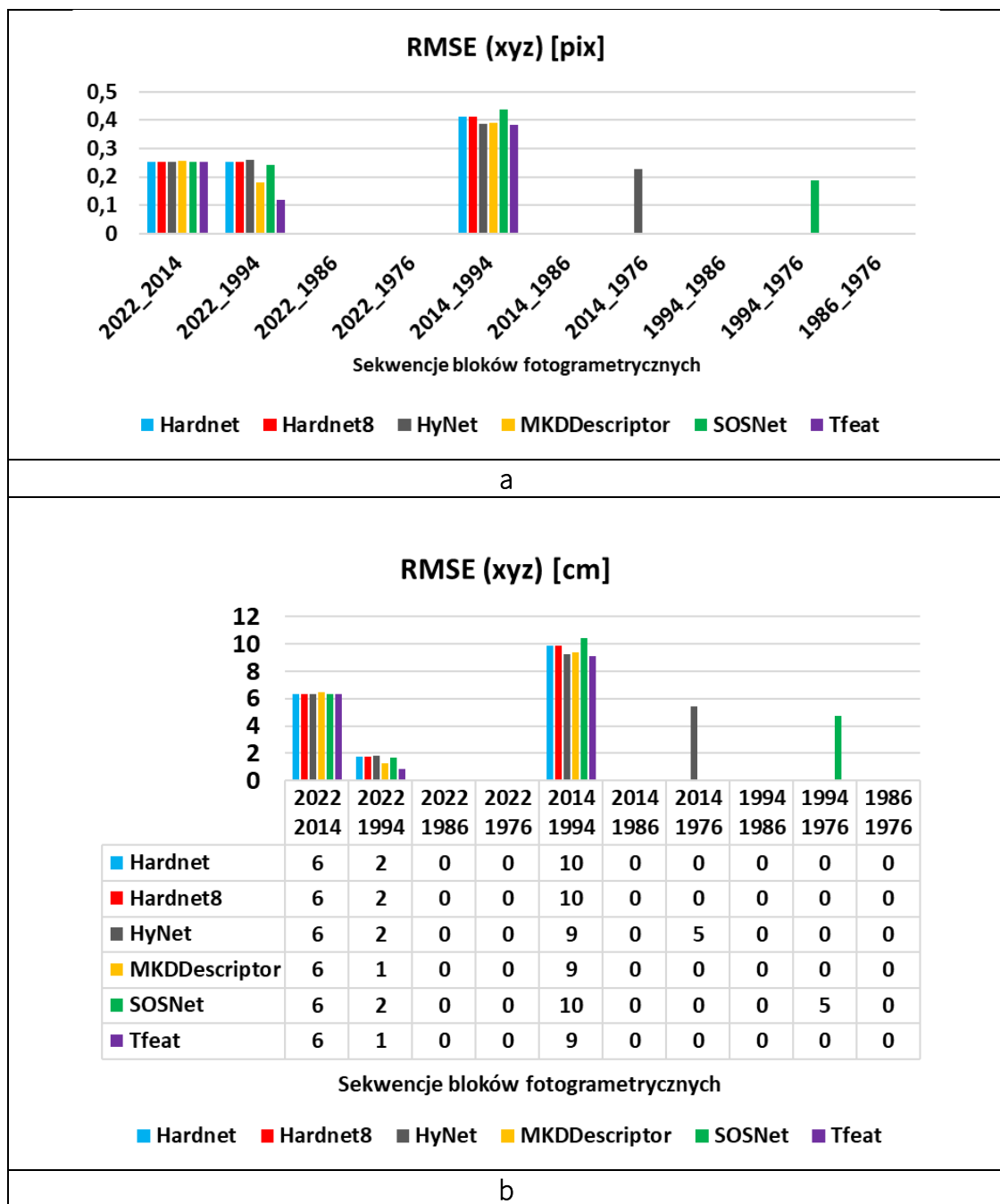
Wartości mediany błędów zbliżone do 0 (mieszczące się w zakresie od -0.07 do 0.07 piksela) potwierdzają brak systematyki w danych. Interpretacja wartości RMSE (wyrażonego w pikselach w odniesieniu do skali analizowanych zdjęć) dla poszczególnych sekwencji czasowych ujawnia, że nie przekraczały one 0.5 piksela, co świadczy o wysokiej dokładności pomiarów.

Porównanie sekwencji bloków 2022–2014 (zdjęcia cyfrowe) oraz 1994–1986 (zdjęcia analogowe) wykazało, że wartości RMSE dla poszczególnych algorytmów były zbliżone, co wskazuje na porównywalną efektywność zastosowanych deskryptorów niezależnie od rodzaju obrazów. Dodatkowo, zarówno wartości RMSE, jak i wartości błędy SMAD potwierdzają brak błędów grubych w pomiarach oraz wskazuje na wysoką dokładność uzyskanych wyników.

Podsumowując, analiza dopasowania sekwencyjnego bloków fotogrametrycznych metodą *Detect-then-Describe* potwierdza wysoką dokładność pomiaru punktów przejściowych oraz brak istotnych błędów systematycznych, co wskazuje na możliwość skutecznego wykorzystania algorytmów HC/LB w pomiarach punktów.

Ocena dokładności orientacji bloków fotogrametrycznych

Uzyskane wartości RMSE orientacji bloków fotogrametrycznych (rys. 9.23), mieszczące się w przedziale 0.2–0.4 piksela dla sekwencji 2022–2014, 2022–1994, 2014–1994, 2014–1976 oraz 1994–1976, świadczą o poprawności wykonania transformacji w tych przypadkach.



Rys. 9.23 Wartości RMSE sekwencyjnej orientacji bloków fotogrametrycznych (1976, 1986, 1994, 2014, 2022) wyznaczone na punktach kontrolnych CHOP, przy zastosowaniu algorytmów HC/LB w podejściu *Detect-then-Describe* w miarach pikselowych (a) w miarach terenowych (b)

W niektórych sekwencjach orientacja bloków nie była możliwa ze względu na niewystarczającą liczbę punktów przejściowych oraz ich nierównomierne rozmieszczenie. Dotyczy to sekwencji 2022–1986, 2022–1976, 2014–1986, 2014–1976 (z wyjątkiem algorytmu HyNet), 1994–1986, 1994–1976 (z wyjątkiem algorytmu SOSNet) oraz 1986–1976.

Analiza wpływu rozdzielczości przestrzennej zdjęć na dokładność orientacji nie wykazała jednoznacznego wpływu tego parametru na wartość RMSE, co potwierdzają porównania

sekwencji 2022–2014 (5 cm–24 cm) oraz 2022–1994 (5 cm–6 cm). Wyjątkiem była sekwencja 2022–1994 dla algorytmu TFeat, w której uzyskano dwukrotnie niższą wartość RMSE. Istotnym ograniczeniem w orientacji okazały się zdjęcia z 1986 roku – we wszystkich sekwencjach zawierających ten blok nie udało się przeprowadzić poprawnej orientacji ze względu na zbyt małą liczbę punktów przejściowych. Prawdopodobną przyczyną był niski kontrast zdjęć, który znacząco utrudnił detekcję punktów przez algorytm SIFT. Mimo tych trudności zastosowane algorytmy HC/LB w większości przypadków umożliwiły uzyskanie wysokiej dokładności orientacji bloków fotogrametrycznych.

REKOMENDACJA zastosowania algorytmów HC/LB do orientacji bloków fotogrametrycznych

Wytyczne dotyczące oceny przydatności algorytmów HC/LB do orientacji bloków fotogrametrycznych zostały zawarte w tabeli 9.3. Ocenie podlegają czas detekcji, liczba i rozmieszczenie punktów oraz wielkość błędu transformacji. Przyjęto czterostopniową, porządkową skalę ocen: 0-2 – ocena negatywna, 3 – dostateczna, 4 – dobra, 5 – bardzo dobra. Otrzymanie oceny negatywnej (0-2) w którymkolwiek z kryteriów oznacza, że dany algorytm nie jest rekomendowany do orientacji analizowanej sekwencji bloków fotogrametrycznych.

Tab. 9.3 Podsumowanie oceny przydatności algorytmów HC/LB do orientacji sekwencji bloków fotogrametrycznych w podejściu *Detect-then-Describe* dla wszystkich okresów czasowych w skali od 0 do 5 punktów

Algorytm	HardNet									
Sekwencje	2022_2014	2022_1994	2022_1986	2022_1976	2014_1994	2014_1986	2014_1976	1994_1986	1994_1976	1986_1976
Czas detekcji	4	4	4	4	4	4	4	4	4	4
Liczba punktów i ich rozmieszczenie	5	4	0	0	5	0	0	3	0	0
RMSE	5	5	0	0	4	0	0	0	0	0
SUMA	14	13	4	4	13	4	4	7	4	4
Algorytm	HardNet8									
Czas detekcji	4	4	4	4	4	4	4	4	4	4
Liczba punktów i ich rozmieszczenie	5	4	0	0	5	0	0	3	0	0
RMSE	5	5	0	0	4	0	0	0	0	0
SUMA	14	13	4	4	13	4	4	7	4	4
Algorytm	HyNet									
Czas detekcji	5	5	5	5	5	5	5	5	5	5
Liczba punktów i ich rozmieszczenie	5	4	0	0	4	0	4	3	0	0
RMSE	5	5	0	0	4	0	5	0	0	0
SUMA	15	14	5	5	13	5	14	8	5	5
Algorytm	MKDDDescriptor									
Czas detekcji	4	4	4	4	4	4	4	4	4	4
Liczba punktów i ich rozmieszczenie	5	3	0	0	5	3	0	3	0	0
RMSE	5	5	0	0	4	0	0	0	0	0
SUMA	14	12	4	4	13	7	4	7	4	4
Algorytm	SOSNet									
Czas detekcji	4	4	4	4	4	4	4	4	4	4
Liczba punktów i ich rozmieszczenie	5	2	0	0	5	2	0	2	4	0
RMSE	5	5	0	0	4	0	0	0	5	0
SUMA	14	11	4	4	13	6	4	6	13	4
Algorytm	Tfeat									
Czas detekcji	4	4	4	4	4	4	4	4	4	4
Liczba punktów i ich rozmieszczenie	5	3	0	0	5	0	0	2	0	0
RMSE	5	5	0	0	4	0	0	0	0	0
SUMA	14	12	4	4	13	4	4	6	4	4
			rekomendowane			nierekomendowane				

Na podstawie oceny wpływu detektora/deskryptora na dokładność orientacji bloków fotogrametrycznych stwierdzono, że detektory/deskryptory uzyskujące wynik powyżej 11 punktów są najbardziej efektywne. Dla każdej analizowanej sekwencji wskazano najbardziej skuteczne rozwiązanie (wyróżnione pogrubieniem), które uzyskało najlepsze rezultaty w ocenie jakości orientacji (tab. 9.4).

Tab. 9.4 Rekomendacja wyboru algorytmu HC/LB w podejściu *Detect-then-Describe* do orientacji bloków fotogrametrycznych w danej sekwencji

Sekwencja bloków	Rodzaj zdjęcia	Rozdzielczość	Jakość radiometryczna	Najlepszy algorytm
2022 2014	cyfrowe cyfrowe	5 cm 24 cm	bardzo dobra bardzo dobra	HyNet, HardNet, HardNet8, MKDDescriptor, SOSNet, TFeat
2022 1994	cyfrowe analogowe	5 cm 7 cm	bardzo dobra dobra	HyNet, HardNet, HardNet8, MKDDescriptor, TFeat
2022 1986	cyfrowe analogowe	5 cm 5 cm	bardzo dobra dobra	Brak
2022 1976	cyfrowe analogowe	5 cm 25 cm	bardzo dobra słaba	Brak
2014 1994	cyfrowe analogowe	24 cm 7 cm	bardzo dobra dobra	HardNet, HardNet8, HyNet, MKDDescriptor, SOSNet, TFeat
2014 1986	cyfrowe analogowe	24 cm 5 cm	bardzo dobra dobra	Brak
2014 1976	cyfrowe analogowe)	24 cm 25 cm	bardzo dobra słaba	HyNet
1994 1986	analogowe analogowe	7 cm 25 cm	dobra dobra	Brak
1994 1976	analogowe analogowe	7 cm 25 cm	dobra słaba	SOSNet
1986 1976	analogowe analogowe	5 cm 25 cm	dobra słaba	Brak

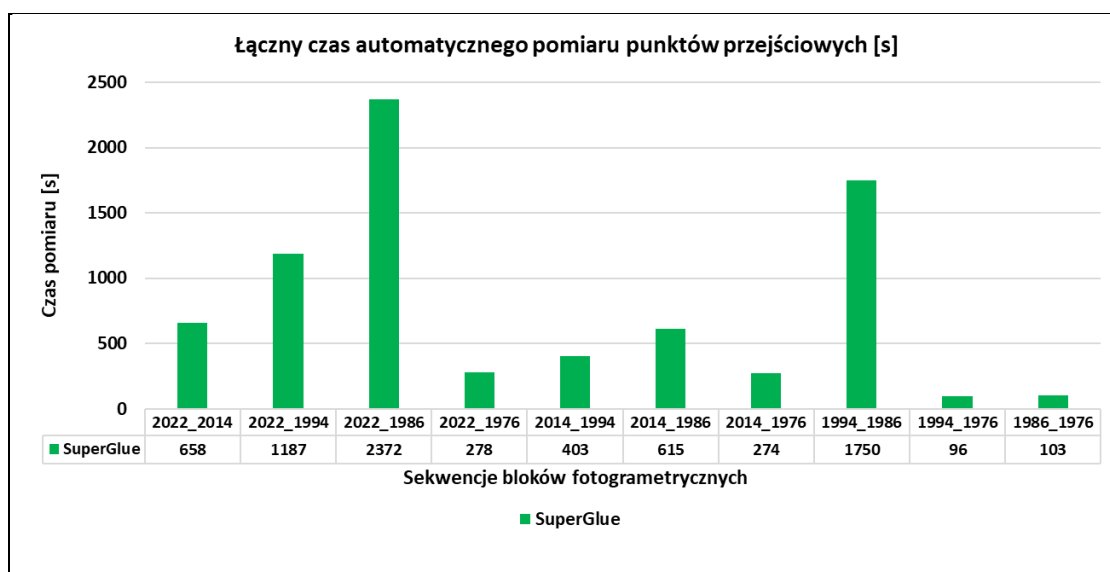
9.3 Ocena poprawności procesu orientacji bloków fotogrametrycznych – podejście Detect-and-Describe (algorytmy Learned-Based)

Kolejne badania polegały na ocenie możliwości wykorzystania podejścia *Detect-and-Describe* opartego na łącznym wykorzystaniu detektora *SuperPoint* i deskryptora *SuperGlue*. Przeprowadzono tę samą metodykę oceny jak opisaną w podrozdziałach 9.1 i 9.2.

Ocena czasu automatycznego pomiaru punktów przejściowych w sekwencyjnej orientacji bloków fotogrametrycznych

Analiza wskazuje, że czas wyznaczenia punktów przejściowych zależy (podobnie jak dla rozwiązania *Detect-then-Describe*) jest od rozdzielczości przestrzennej (GSD) true-

ortofotomapy (rys 9.24). Im mniejsze GSD (większa rozdzielczość obrazów/rastrów), tym dłuższy czas obliczeń.



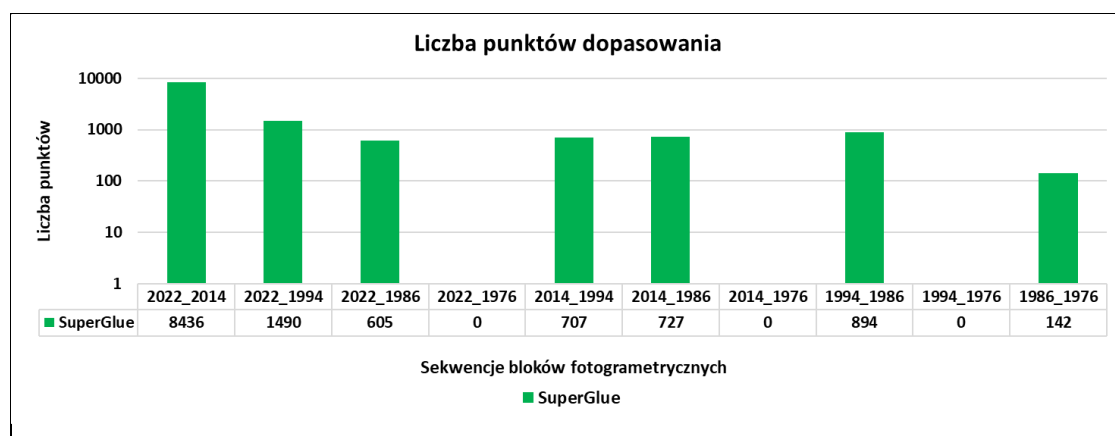
Rys. 9.24 Łączny czasu automatycznego wyznaczenia punktów przejściowych przez algorytmy SuperPoint/SuperGlue w podejściu *Detect-end-Describe* w sekwencjach bloków fotogrametrycznych z roku 2022, 2014, 1994, 1986 i 1976

Najdłuższy czas pomiaru zaobserwowano dla sekwencji obejmujących zdjęcia o najwyższej rozdzielczości (np. 2022–1986: 40 minut, 1994–1986: 30 minut). W przypadku sekwencji 2022–1976 oraz 2014–1976 czas pomiaru punktów był znacznie krótszy (ok. 5 minut), co stanowi wyjątek w dopasowywanych sekwencjach. Prawdopodobnym wyjaśnieniem tego jest niska jakość radiometryczna zdjęć z 1976 roku, która mogła ograniczyć liczbę wykrytych punktów przejściowych, a tym samym skrócić czas przetwarzania. Najkrótszy czas pomiaru, odnotowany dla sekwencji 1994–1976 oraz 1986–1976 (około 1.5 minuty), sugeruje, że w tych przypadkach na czas pomiaru, zależny od liczby wyznaczonych punktów, istotny wpływ miała niższa jakość radiometryczna zdjęć analogowych w porównaniu do zdjęć cyfrowych oraz szczególnie niska jakość zdjęć z 1976 roku.

Podsumowując, przeprowadzona analiza wskazuje, że głównymi czynnikami wpływającymi na czas pomiaru punktów przejściowych są GSD obrazów oraz ich jakość radiometryczna; wniosek ten pozostaje spójny dla wszystkich analizowanych kombinacji detektor–deskryptor.

Ocena liczby wykrytych punktów przejściowych i ich rozkładu w sekwencyjnej orientacji bloków fotogrametrycznych

Najlepsze dopasowanie (rys. 9.26, 9.27) punktów przejściowych uzyskano dla sekwencji 2022–2014 (8436 punktów), mimo różnicy w rozdzielczości zdjęć (5 cm i 24 cm). Wysoka liczba punktów wynikała z faktu, że oba bloki były cyfrowe, miały dobrą jakość radiometryczną oraz obejmowały stosunkowo krótki okres czasowy (8 lat), co oznaczało niewielkie zmiany w pokryciu terenu.

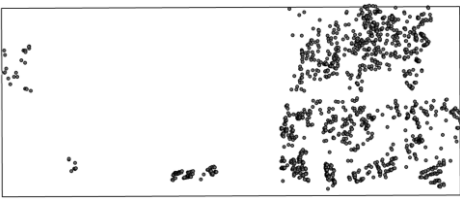
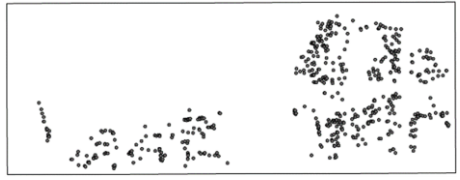
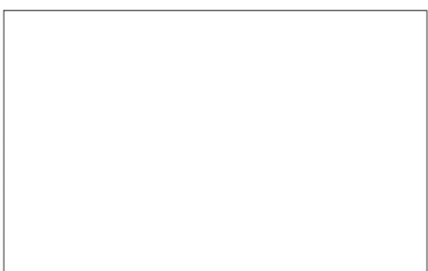
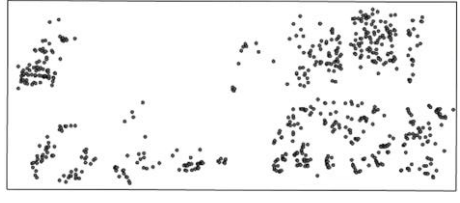
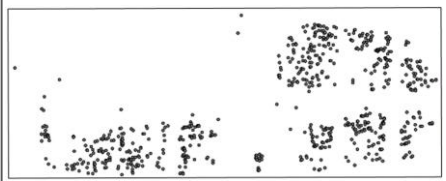
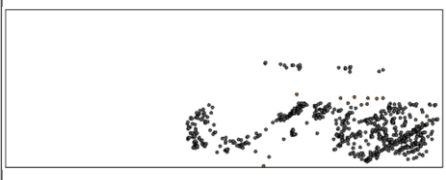
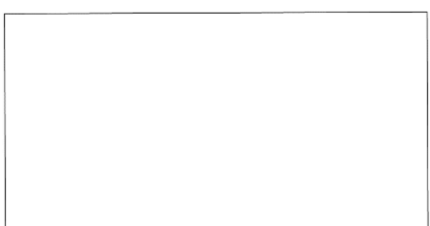
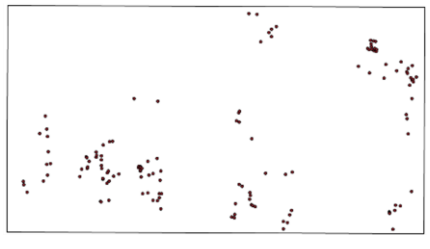


Rys. 9.25 Liczba pomierzonych automatycznie punktów przejściowych w poszczególnych sekwencjach bloków fotogrametrycznych (1976 r., 1986 r., 1994 r., 2011 r., 2022 r.) dla algorytmów SuperPoint/SuperGlue w podejściu *Detect-and-Describe*

Rozdzielczość przestrzenna wpływa na liczbę wykrytych punktów – sekwencja 2022–1994 (5 cm i 7 cm) miała ponad dwukrotnie więcej punktów (1490) niż 2014–1994 (24 cm i 7 cm, 707 punktów), co wskazuje, że wyższa rozdzielczość sprzyja lepszemu dopasowaniu (rys. 9.25).

Widoczny jest systematyczny spadek liczby punktów w starszych sekwencjach zawierających zdjęcia z 2022 roku – od 1490 punktów w sekwencji 2022–1994, przez 605 punktów w 2022–1986, aż po brak punktów w 2022–1976. Głównymi przyczynami tego zjawiska mogą być zmiany urbanistyczne oraz różnice radiometryczne między obrazami. Dopasowanie sekwencji zawierających zdjęcia z 1976 roku (np. 2022–1976, 2014–1976) okazało się szczególnie problematyczne – punktów przejściowych nie udało się wyznaczyć lub ich liczba była niewystarczająca do poprawnej orientacji bloków zdjęć. Potencjalnymi przyczynami są: niska jakość radiometryczna zdjęć z 1976 roku, ich mała skala (duże GSD) oraz niedopasowanie zbioru treningowego algorytmu SuperGlue do tego typu danych.

Najmniejszą liczbę punktów (142) uzyskano w sekwencji 1986–1976 (oba bloki analogowe, 5 cm i 25 cm). Mimo to, w przeciwieństwie do innych sekwencji z rokiem 1976, punkty zostały wyznaczone, co można tłumaczyć krótkim odstępem czasowym (10 lat), a więc mniejszymi zmianami w pokryciu terenu i lepszym dopasowaniem obrazów. W sekwencjach o podobnej rozdzielczości przestrzennej (2014–1994 i 2014–1986 – odpowiednio 24 cm, 7 cm, 5 cm), liczba punktów przejściowych była zbliżona (707 i 727 punktów). Wysoka rozdzielczość przestrzenna (5–7 cm) sprzyja lepszemu dopasowaniu, podczas gdy mała rozdzielczość (24–25 cm) utrudnia wykrywanie punktów przejściowych. Dlatego dla zdjęcia z 1976 roku, wykonane w małej skali i niskiej jakości radiometrycznej, algorytm napotkał największe trudności w procesie dopasowania. Najlepsze wyniki uzyskano dla sekwencji zawierających obrazy cyfrowe o wysokiej rozdzielczości.

Algorithm	2022-2014	2022-1994	2022-1986	2022-1976	2014-1994
Super Glue					
Algorithm	2014-1986	2014-1976	1994-1986	1994-1976	1986-1976
Super Glue					

Rys. 9.26 Rozkład punktów przejściowych wyznaczonych przez algorytm SuperGlue w podejściu *Detect-and-Describe* sekwencji bloków fotogrametrycznych z roku 2022, 2014, 1994, 1986 i 1976

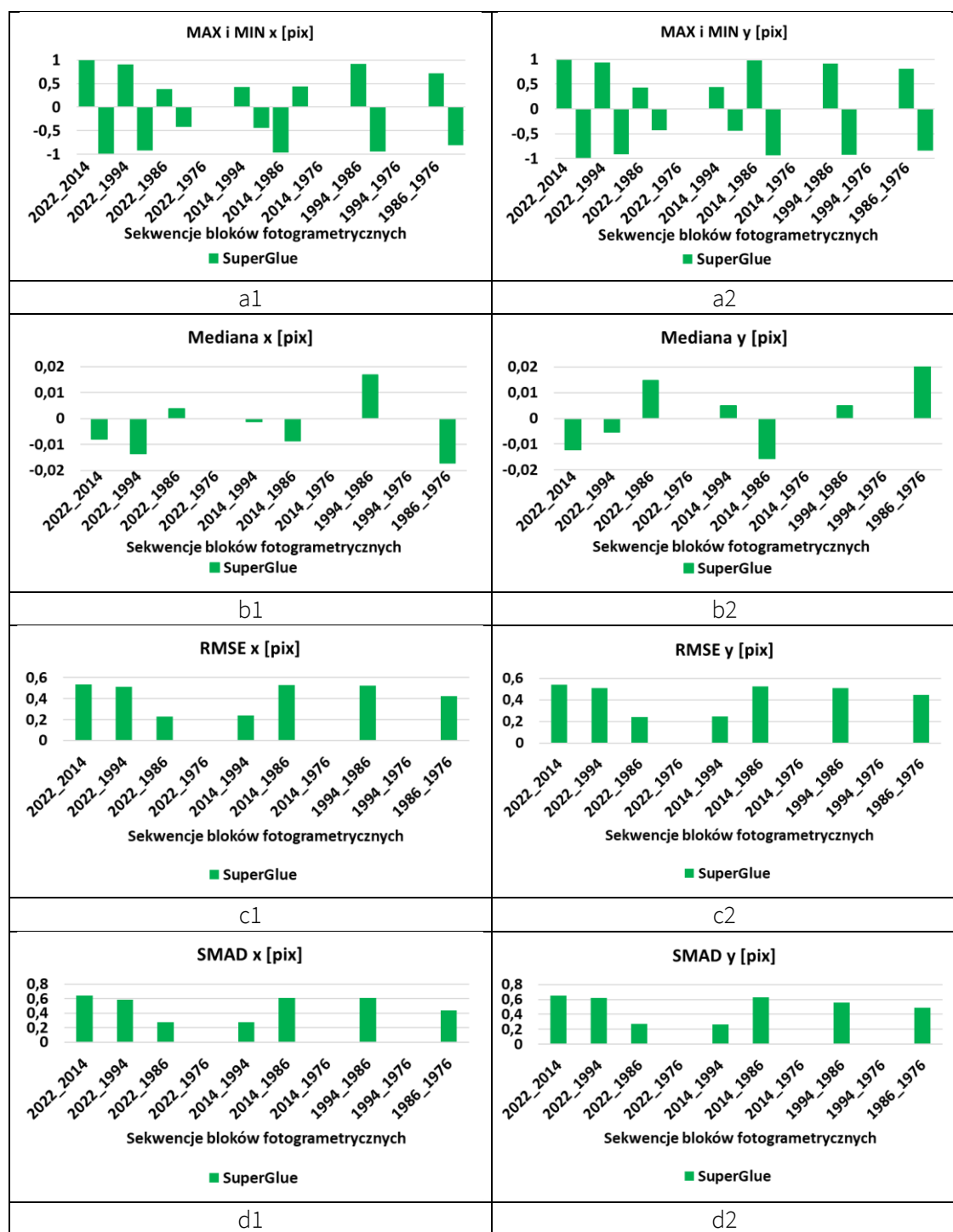
2022-2014		2022-1994		2022-1986		2022-1976	
SuperGlue		SuperGlue		SuperGlue		SuperGlue	
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16
2014-1994		2014-1986		2014-1976		1994-1986	
SuperGlue		SuperGlue		SuperGlue		SuperGlue	
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16
1994-1976		1986-1976					
SuperGlue		SuperGlue					
1	2	1	2				
3	4	3	4				
5	6	5	6				
7	8	7	8				
9	10	9	10				
11	12	11	12				
13	14	13	14				
15	16	15	16				

Rys. 9.27 Rozkład punktów przejściowych (w podziale tailowym – 16) sekwencji (2022–2014, 2022–1994, 2022–1986, 2022–1976, 2014–1994, 2014–1986, 2014–1976, 1994–1986, 1994–1976, 1986–1976) bloków fotogrametrycznych (true-ortofotomap) z uwzględnieniem algorytmów LB (SuperPoint/SuperGlue) podejścia *Detect-end-Describe* (kolor zielony oznacza wystarczającą liczbę punktów w danym podobzdarze, natomiast kolor czerwony oznacza ich brak)

Podsumowując, największą liczbę punktów uzyskano dla sekwencji zawierających zdjęcia cyfrowe o wysokiej rozdzielczości, natomiast najgorsze wyniki dotyczyły sekwencji z dużymi odstępami czasowymi oraz zdjęciami analogowymi o niskiej rozdzielczości. Rozdzielczość przestrzenna zdjęć ma istotny wpływ na skuteczność wykrywania punktów przejściowych. Im krótszy przedział czasowy między obrazami, tym większa liczba wykrytych punktów, co sprzyja lepszemu dopasowaniu sekwencji. Zdjęcia analogowe, szczególnie z 1976 roku, stanowią istotne ograniczenie dla procesu orientacji bloków fotogrametrycznych, głównie ze względu na ich niską jakość radiometryczną.

Ocena dokładności wyznaczenia punktów przejściowych

Na rysunku 9.28 przedstawione zostały wartości błędów pomiaru punktów przejściowych przy użyciu algorytmu SuperGlue. Analiza mediany błędów wyznaczonych przez algorytmy SuperPoint/SuperGlue dla współrzędnych x i y wykazała, że ich wartości były zbliżone do zera (w zakresie od $-0,02$ do $0,02$ piksela), co wskazuje na brak błędów systematycznych w pomiarach. Wartości RMSE (wyrażonego w pikselach w odniesieniu do skali analizowanych zdjęć) dla wszystkich analizowanych sekwencji nie przekraczały 0.5 piksela. Najniższe wartości, na poziomie około 0.2 piksela, odnotowano dla sekwencji 2022–1986 oraz 2014–1994.



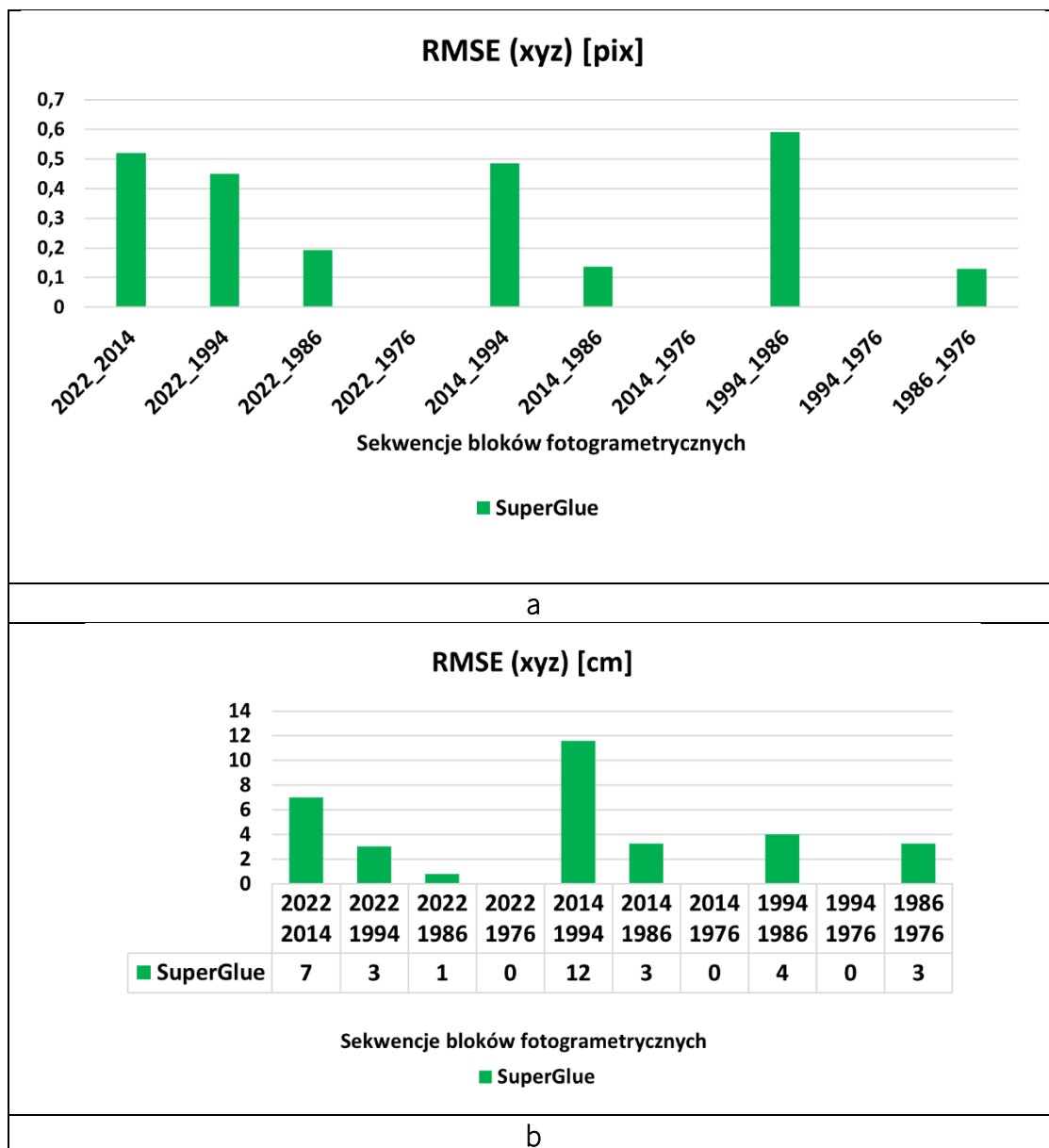
Rys. 9.28 Zestawienie wartości błędów (Maksimum, Minimum, Mediana, SMAD, SIGMA, RMSE) pomiaru punktów przejściowych (podejście *Detect-and-Describe*) przy użyciu algorytmów SuperPoint /SuperGlue dla par true–ortofotomap z roku 1976, 1986, 1994, 2014 i 2022

Analiza wpływu rodzaju zdjęć (cyfrowe vs. analogowe) oraz rozdzielczości przestrzennej (5 cm – 24 cm dla sekwencji 2022–2014 oraz 7 cm – 5 cm dla 1994–1986) na wartość RMSE nie wykazała istotnych różnic. Oznacza to, że zarówno typ zdjęć, jak i różnice w rozdzielczości przestrzennej, nie miały znaczącego wpływu na dokładność uzyskanych pomiarów.

Dodatkowo, wartości błędów SMAD uzyskane za pomocą algorytmów SuperPoint/SuperGlue były zbliżone dla obu współrzędnych i odpowiadały wartościom RMSE. Potwierdza to brak błędów grubych w pomiarach oraz wskazuje na wysoką dokładność wyników uzyskanych w procesie dopasowania.

Ocena dokładności orientacji bloków fotogrametrycznych

Analiza wyników RMSE (rys. 9.29) orientacji bloków fotogrametrycznych wskazuje, że dla większości sekwencji wartość błędu utrzymuje się na poziomie około 0.5 piksela, co potwierdza powtarzalność dokładności transformacji niezależnie od analizowanego okresu. Najniższe wartości błędów odnotowano dla sekwencji 1986–1976 (0.1 piksela), 2014–1986 (0.1 piksela) oraz 2022–1986 (0.2 piksela), co sugeruje wyższą jakość dopasowania punktów przejściowych w tych zakresach czasowych. Brak możliwości orientacji w sekwencjach 2022–1976, 2014–1976 oraz 1994–1976 wynikał z niewystarczającej liczby punktów oraz ich nierównomiernego rozmieszczenia.



Rys. 9.29 Wartości RMSE sekwencyjnej orientacji bloków fotogrametrycznych (1976, 1986, 1994, 2014, 2022) wyznaczone na punktach kontrolnych CHOP, przy zastosowaniu algorytmów LB (*SuperPoint* i *SuperGlue*) w podejściu *Detect-and-Describe* w miarach pikselowych (a) w miarach terenowych (b)

Porównanie wartości RMSE w sekwencjach 2022–2014 (0.5 piksela) oraz 1994–1986 (0.6 piksela) nie wykazało istotnych różnic wynikających z rodzaju zdjęć (cyfrowych i analogowych). Brak możliwości przeprowadzenia orientacji dla sekwencji 2022–1976, 2014–1976 oraz 1994–1976 był spowodowany niską jakością radiometryczną zdjęć z 1976 roku oraz dużą rozbieżnością tonalną i kontrastową pomiędzy epokami. Algorytmy *SuperPoint* i *SuperGlue*, działające na podstawie ekstrakcji i dopasowania cech, nie były w stanie wykryć

Dodatkowo, rozdzielczość przestrzenna zdjęć w analizowanych przedziałach czasowych nie miała istotnego wpływu na dokładność orientacji bloków wyrażoną w pikselach (tj. w skali zdjęć). Uzyskane wyniki potwierdzają, że algorytmy SuperPoint i SuperGlue, działające w podejściu *Detect-and-Describe*, skutecznie identyfikują i dopasowują punkty przejściowe, zapewniając spójną jakość transformacji bloków w analizowanych sekwencjach czasowych.

W tabeli 9.5 przedstawiono wytyczne oceny przydatności algorytmów LB SuperGlue w podejściu *Detect-and-Describe* do orientacji bloków fotogrametrycznych. Ocenie poddano takie aspekty jak czas detekcji, liczba i rozmieszczenie punktów przejściowych oraz dokładność transformacji. Do klasyfikacji wyników zastosowano czterostopniową, porządkową skalę ocen, gdzie 0-2 oznacza ocenę negatywną, 3 – dostateczną, 4 – dobrą, a 5 – bardzo dobrą. W przypadku uzyskania oceny negatywnej (0-2) dla któregośkolwiek z kryteriów, uznano, że algorytm ten nie można wykorzystać do orientacji danej sekwencji bloków fotogrametrycznych.

[illegible]

243

Tab. 9.6 Rekomendacja wyboru algorytmu LB SuperGlue w podejściu *Detect-end-Describe* do orientacji bloków fotogrametrycznych w danej sekwencji

Sekwencja bloków	Rodzaj zdjęcia	Rozdzielczość	Jakość radiometryczna	Algorytm SuperGlue
2022 2014	cyfrowe cyfrowe	5 cm 24 cm	bardzo dobra bardzo dobra	Rekomendowany
2022 1994	cyfrowe analogowe	5 cm 7 cm	bardzo dobra dobra	Rekomendowany
2022 1986	cyfrowe analogowe	5 cm 5 cm	bardzo dobra dobra	Rekomendowany
2022 1976	cyfrowe analogowe	5 cm 25 cm	bardzo dobra słaba	Nierekomendowany
2014 1994	cyfrowe analogowe	24 cm 7 cm	bardzo dobra dobra	Rekomendowany
2014 1986	cyfrowe analogowe	24 cm 5 cm	bardzo dobra dobra	Rekomendowany
2014 1976	cyfrowe analogowe)	24 cm 25 cm	bardzo dobra słaba	Nierekomendowany
1994 1986	analogowe analogowe	7 cm 25 cm	dobra dobra	Nierekomendowany
1994 1976	analogowe analogowe	7 cm 25 cm	dobra słaba	Nierekomendowany
1986 1976	analogowe analogowe	5 cm 25 cm	dobra słaba	Rekomendowany

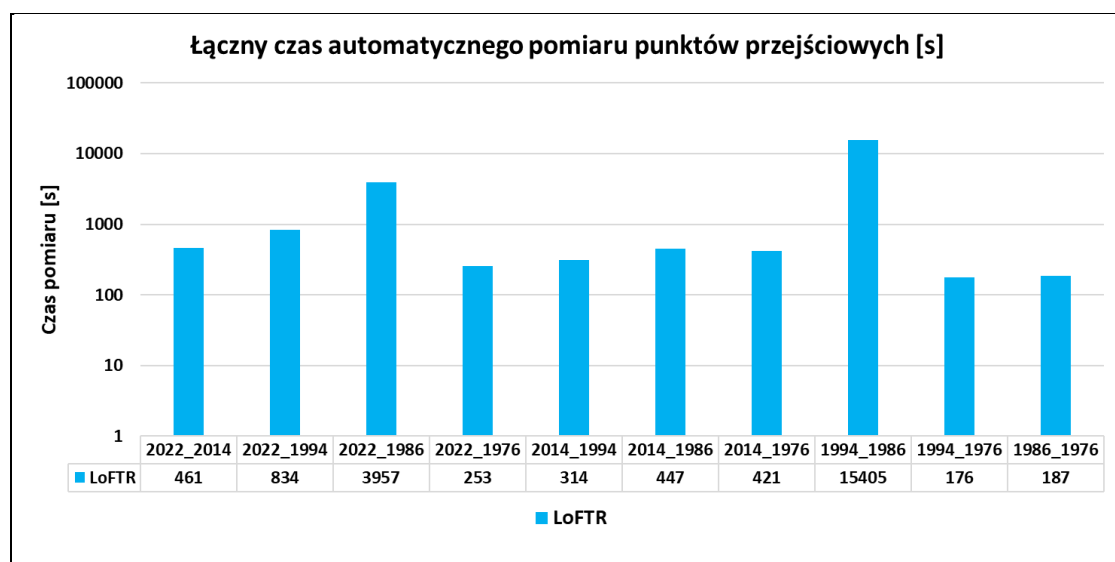
9.4 Ocena poprawności procesu orientacji bloków fotogrametrycznych – podejście Describe-then-Detect (algorytmy Learned-Based)

W podrozdziale 9.4 zaproponowano tę samą metodykę oceny poprawności procesu orientacji bloków fotogrametrycznych, co w podrozdziale 9.1. W tym celu wykorzystano podejście *Describe-then-Detect* oparte na algorytmie LoFTR.

Ocena czasu automatycznego pomiaru punktów przejściowych w sekwencyjnej orientacji bloków fotogrametrycznych

Najkrótszy czas pomiaru (rys. 9.30) punktów przejściowych przez algorytm LoFTR odnotowano dla sekwencji 2022–1976 (253 ms), 1994–1976 (176 ms) i 1986–1976 (187 ms).

Wynika to z faktu, że zdjęcia z 1976 roku cechują się niską rozdzielczością przestrzenną (GSD około 25 cm) oraz słabą jakością radiometryczną. Ograniczona liczba wykrytych punktów była konsekwencją mniejszej liczby wyraźnych elementów obrazu, takich jak krawędzie, narożniki, budynki, drogi, rzeki i kontury pól, co skróciło czas przetwarzania.



Rys. 9.30 Łączny czasu automatycznego wyznaczenia punktów przejściowych przez algorytm LoFTR w podejściu *Describe-then-Detect* w sekwencjach bloków fotogrametrycznych z roku 2022, 2014, 1994, 1986 i 1976

Najdłuższy czas pomiaru (15405 ms) wystąpił w sekwencji 1994–1986, obejmującej zdjęcia o wysokiej rozdzielczości przestrzennej wykonane w krótkim odstępie czasowym, co skutkowało dużą liczbą wspólnych cech obrazowych i znacznym wydłużeniem procesu dopasowania. Zdjęcia w obu seriach są analogowe i mają wysoką rozdzielczość przestrzenną (1994 – GSD 7 cm, 1986 – GSD 5 cm). LoFTR, który analizuje cały obraz zamiast polegać jedynie na lokalnych cechach, wymaga w tym przypadku znacznie dłuższego czasu na dopasowanie punktów przejściowych. Duże różnice tonalne i radiometryczne dodatkowo zwiększają złożoność obliczeniową.

Porównując najdłuższy czas pomiaru algorytmu LoFTR (dla sekwencji 1994–1986) z czasami działania algorytmów zastosowanych w poprzednich metodach, można stwierdzić, że był on około 8 razy dłuższy niż w przypadku algorytmu SuperGlue, około 70 razy dłuższy w porównaniu do większości sieci z rodziny SuperGlue (LB), trzykrotnie krótszy niż dla algorytmu ASURF (HC) oraz około 15 razy dłuższy niż dla algorytmu FAST (HC). Natomiast w przypadku najkrótszego czasu pomiaru (dla sekwencji 1986–1976), algorytm LoFTR działał około dwukrotnie wolniej niż

SuperGlue i sieci LB, około 66 razy szybciej niż ASURF oraz osiągnął porównywalny czas działania względem algorytmu FAST.

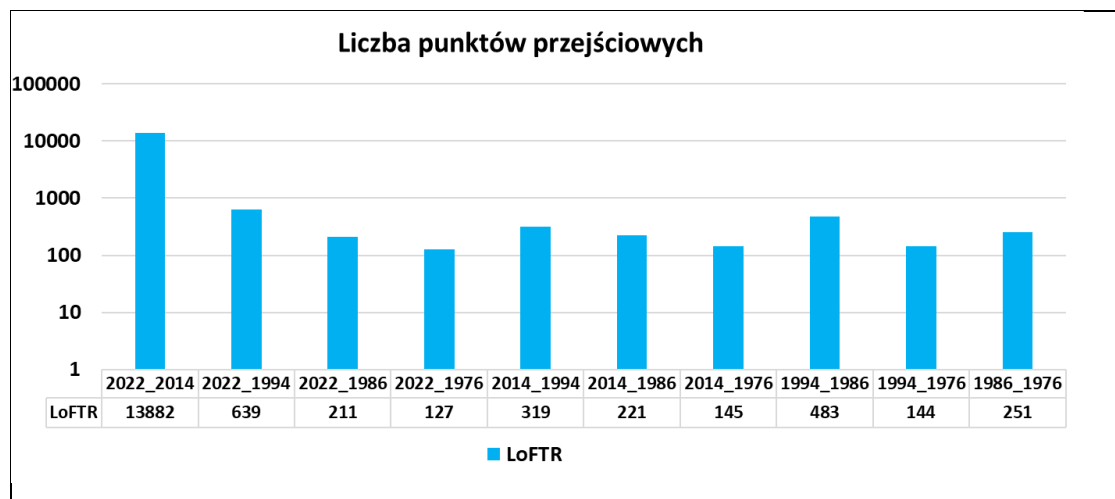
Długi czas pomiaru wystąpił również w sekwencji 2022–1994 (834 ms), ale był około czterokrotnie krótszy niż dla sekwencji 1994–1986. Można to przypisać bardzo dobrej jakości radiometrycznej zdjęć cyfrowych z 2022 roku, która wpływa na zwiększenie liczby mierzonych punktów charakterystycznych, co przekłada się na wyznaczenie większej liczby punktów przejściowych.

Sekwencje obejmujące zdjęcia z 2014 roku (np. 2014–1976: 421 ms, 2014–1986: 447 ms) wykazują relatywnie krótkie czasy przetwarzania. Jest to zgodne z wcześniejszymi obserwacjami, że zdjęcia z 2014 roku mają stosunkowo niską rozdzielczość przestrzenną (GSD 24 cm), co redukuje liczbę wykrytych punktów i przyspiesza obliczenia. Zastosowanie algorytmu LoFTR do zdjęć lotniczych wpływa na czas obliczeń ze względu na to, że algorytm został trenowany na obrazach naziemnych o znacznie wyższej rozdzielczości przestrzennej. Sekwencje obejmujące zdjęcia o wysokiej rozdzielczości i dużych różnicach radiometrycznych wymagają dłuższego czasu przetwarzania, ponieważ algorytm musi analizować globalny kontekst obrazu i odrzucać błędne dopasowania.

Podsumowując, czynniki takie jak rozdzielczość przestrzenna, jakość radiometryczna zdjęć tworzących bloki mają istotny wpływ na czas wyznaczania punktów przejściowych przez LoFTR. Najdłuższe czasy występują w sekwencjach obejmujących zdjęcia analogowe o wysokiej rozdzielczości, a najkrótsze - w przypadku zdjęć o niskiej rozdzielczości i słabej jakości radiometrycznej.

Ocena liczby wykrytych punktów przejściowych i ich rozkładu w sekwencyjnej orientacji bloków fotogrametrycznych

Największą liczbę punktów przejściowych wykryto w sekwencji 2022–2014, co wynika z wysokiej jakości radiometrycznej i przestrzennej zdjęć cyfrowych (rys. 9.31). Algorytm LoFTR, wykorzystujący głębokie sieci neuronowe do wykrywania i dopasowywania cech obrazu, działa najskuteczniej, gdy obrazy mają dobrą rozdzielczość przestrzenną i radiometryczną, co umożliwia wyznaczenie większej liczby punktów przejściowych.



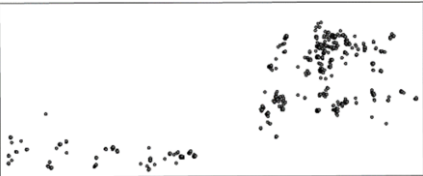
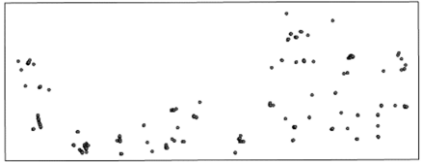
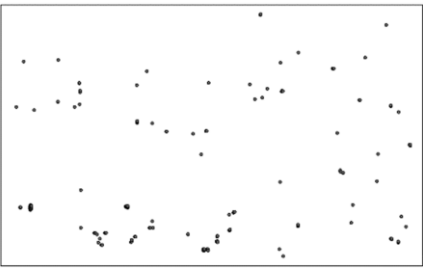
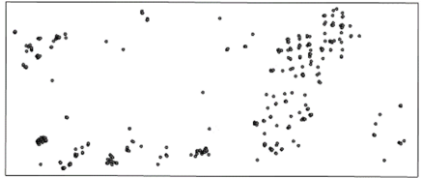
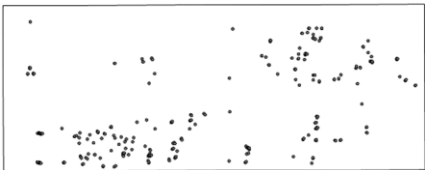
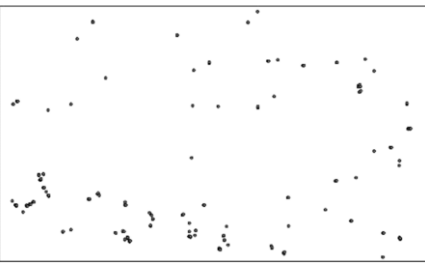
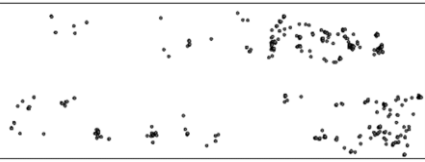
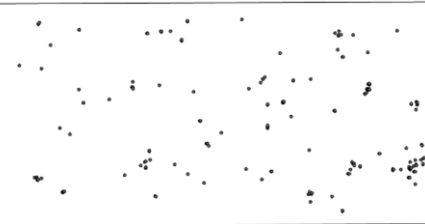
Rys. 9.31 Liczba pomierzonych automatycznie punktów przejściowych COP i CHOP w poszczególnych sekwencjach bloków fotogrametrycznych (1976 r., 1986 r., 1994 r., 2011 r., 2022 r.) dla algorytmu LB LoFTR w podejściu *Describe-then-Detect*

Liczba punktów maleje wraz ze wzrostem odstępu czasowego pomiędzy pozyskanymi zdjęciami lotniczymi, co wynika z dynamicznych zmian w pokryciu terenu (np. urbanizacja, zmiana układu dróg, przekształcenia terenów rolniczych) oraz niższej jakości radiometrycznej analogowych zdjęć lotniczych.

Algorytm LoFTR opiera się na uczeniu maszynowym i ekstrakcji głębokich cech obrazu, pomimo problemów wyszukiwania odpowiadających sobie cech między obrazami, charakteryzuje się dobrą skutecznością dopasowania (rys. 9.32, 9.31). Wraz ze wzrostem odstępu czasowego zmieniają się elementy krajobrazu (np. budynki, drogi), co utrudnia algorytmowi znalezienie punktów wiążących. Najmniejszą liczbę punktów wykryto w sekwencjach z roku 1976, co wynika nie tylko z dużych zmian krajobrazowych, lecz także z niskiej jakości radiometrycznej i przestrzennej zdjęć analogowych tworzących ten blok. Algorytm LoFTR opiera się na dopasowywaniu cech głębokich, takich jak kształty dróg, kontury zabudowy miejskiej czy granice działek, rozpoznawanych przez wyuczoną wcześniej głęboką sieć neuronową. W przypadku obrazów o niskim kontraście i słabej jakości radiometrycznej wyznaczenie punktów przejściowych jest dla algorytmu znacząco utrudnione. Wyjątkiem jest sekwencja 1986–1976, gdzie liczba punktów była większa niż w 2022–1976. Jest to efektem mniejszego odstępu czasowego (10 lat), co oznacza, że struktury (np. drogi i zabudowa) w terenie były mniej zmienione, a algorytm mógł skuteczniej znaleźć dopasowania.

Algorytm LoFTR najlepiej wykrywa punkty na elementach o wysokim kontraście, takich jak dachy budynków, krawędzie konstrukcji i elementy infrastruktury drogowej. Dzieje się tak,

ponieważ LoFTR, w przeciwieństwie do klasycznych metod wykrywania cech (np. SIFT, SURF), bazuje na globalnym kontekście obrazu i analizie przestrzennej. Wysoki kontrast pozwala na jednoznaczne przypisanie cech do konkretnej struktury zdjęcia, co ułatwia ich dopasowanie między obrazami. Natomiast na terenach o jednolitej strukturze (np. pola uprawne, nieużytki) brakuje wyraźnych, charakterystycznych cech, co prowadzi do mniejszej liczby wykrytych punktów. LoFTR grupuje punkty w lokalne struktury o zbliżonych odległościach, tworząc charakterystyczny wzór. Wynika to z architektury sieci neuronowej, która analizuje zarówno relacje lokalne, jak i globalne między cechami obrazu. Dzięki temu punkty są wykrywane w zorganizowany sposób, co przekłada się na ich równomierne rozmieszczenie w obszarach charakteryzujących się dużą szczegółowością, takich jak tereny zabudowane czy sieć dróg.

LoFTR	2022-2014	2022-1994	2022-1986	2022-1976	2014-1994
					
	2014-1986	2014-1976	1994-1986	1994-1976	1986-1976
					
	LoFTR	LoFTR			

Rys. 9.32 Rozkład punktów przejściowych wyznaczonych przez algorytm LoFTR w podejściu *Describe-then-Detect* sekwencji bloków fotogrametrycznych z roku 2022, 2014, 1994, 1986 i 1976

2022-2014		2022-1994		2022-1986		2022-1976	
LoFTR		LoFTR		LoFTR		LoFTR	
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16
2014-1994		2014-1986		2014-1976		1994-1986	
LoFTR		LoFTR		LoFTR		LoFTR	
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
5	6	5	6	5	6	5	6
7	8	7	8	7	8	7	8
9	10	9	10	9	10	9	10
11	12	11	12	11	12	11	12
13	14	13	14	13	14	13	14
15	16	15	16	15	16	15	16
1994-1976		1986-1976					
LoFTR		LoFTR					
1	2	1	2				
3	4	3	4				
5	6	5	6				
7	8	7	8				
9	10	9	10				
11	12	11	12				
13	14	13	14				
15	16	15	16				

Rys. 9.33 Rozkład punktów przejściowych (w podziale tailowym – 16) sekwencji (2014–1986, 2014–1976, 1994–1986, 1994–1976, 1986–1976) bloków fotogrametrycznych (true–ortofotomap) z uwzględnieniem algorytmu LB (LoFTR) podejścia *Describe-then-Detect* (kolor zielony oznacza wystarczającą liczbę punktów w danym podobszarze, natomiast kolor czerwony oznacza ich brak)

Podsumowując, LoFTR działa najskuteczniej na zdjęciach cyfrowych o bardzo dobrej jakości radiometrycznej oraz w sekwencjach o niewielkim odstępie czasowym pozyskania danych, ponieważ jego głębokie sieci neuronowe lepiej dopasowują punkty, gdy obrazy mają podobne cechy strukturalne i radiometryczne. Algorytm LoFTR jest w stanie skutecznie dopasowywać cechy nawet przy niskiej zgodności obrazów (wynikającej np. z ich słabej jakości radiometrycznej), ponieważ wykorzystuje analizę globalnego kontekstu całego obrazu, w przeciwieństwie do metod takich jak HC, które opierają się na lokalnym wykrywaniu i dopasowywaniu pojedynczych punktów charakterystycznych. Jego skuteczność maleje przy dużych zmianach krajobrazowych i niskiej jakości archiwalnych zdjęciach lotniczych, ponieważ trudniej jest znaleźć jednoznaczne cechy do dopasowania. Pomimo mniejszej liczby wykrytych

punktów, algorytm LoFTR nadal umożliwia pozyskanie punktów wystarczających do przeprowadzenia poprawnej orientacji sekwencji bloków fotogrametrycznych.

Ocena dokładności wyznaczenia punktów przejściowych

Na rysunku 9.34 przedstawione zostały wartości RMSE (wyrażonego w pikselach w odniesieniu do skali analizowanych zdjęć) wyznaczenia punktów przejściowych przy użyciu algorytmu LoFTR.



Rys. 9.34 Zestawienie wartości błędów (Maksimum, Minimum, Mediana, SMAD, SIGMA, RMSE) pomiaru punktów przejściowych (podejście *Describe-then-Detect*) przy użyciu algorytmu LoFTR dla par true-ortofotomap z roku 1976, 1986, 1994, 2014 i 2022

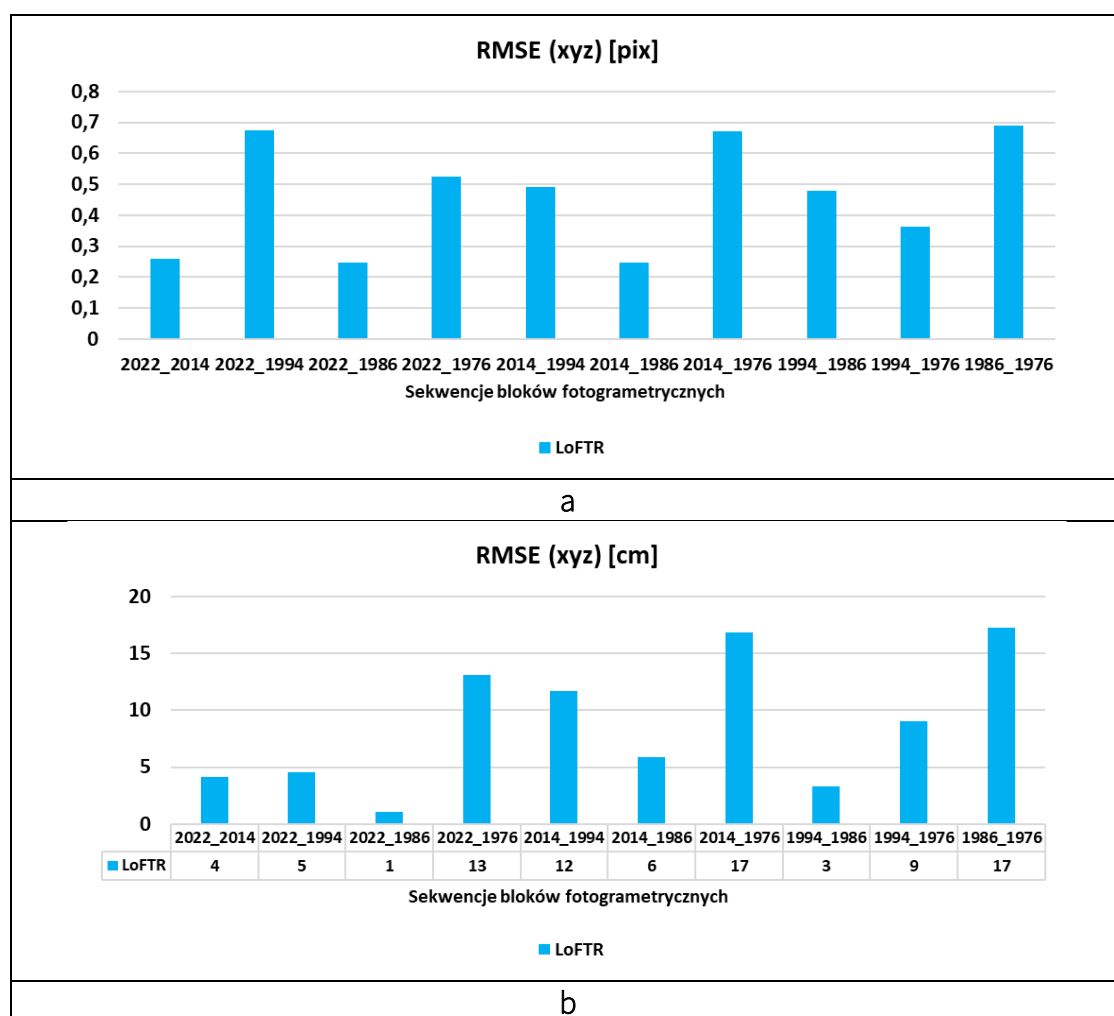
Mediana błędów pomiarowych dla współrzędnych x i y była zbliżona do zera (w zakresie od -0.02 do 0.03 piksela) wskazując na brak błędów systematycznych w pomiarach, co potwierdza wysoką jakość danych. Wartości RMSE dla wszystkich analizowanych sekwencji nie

przekraczają 0.5 piksela, co oznacza wysoką dokładność pomiarów. Wartości błędów SMAD są zbliżone do wartości RMSE, co dodatkowo potwierdza brak błędów grubych oraz stabilność pomiarów punktów przejściowych w różnych sekwencjach czasowych.

Podsumowując, zastosowanie algorytmu LoFTR w podejściu *Describe-then-Detect* pozwoliło na dokładne dopasowanie bloków fotogrametrycznych w różnych sekwencjach czasowych, przy jednoczesnym zachowaniu wysokiej dokładności i stabilności pomiarowej.

Ocena dokładności orientacji bloków fotogrametrycznych

Jakość radiometryczna zdjęć odgrywa kluczową rolę w dokładności orientacji bloków fotogrametrycznych, zwłaszcza w kontekście wyznaczania współrzędnej (z) punktów przejściowych, która opiera się na Numerycznym Modelu Pokrycia Terenu (NMPT).



Rys. 9.35 Wartości RMSE sekwencyjnej orientacji bloków fotogrametrycznych (1976, 1986, 1994, 2014, 2022) wyznaczone na punktach kontrolnych CHOP, przy zastosowaniu algorytmu LB (LoFTR) w podejściu *Describe-then-Detect* w miarach pikselowych (a) w miarach terenowych (b)

Starsze zdjęcia analogowe, takie jak te z 1976 roku, charakteryzowały się niższą jakością radiometryczną, co skutkowało większymi błędami (rys.9.35). Dla bloków utworzonych ze zdjęć cyfrowych (2022–2014) uzyskano najniższe wartości RMSE, sięgające około 0.2 piksela.

Na dokładność dopasowania punktów wpływał również kontrast – zdjęcia z 1986 roku wykazywały wyższy kontrast niż te z 1994 roku, co przełożyło się na niższe RMSE w sekwencjach 2022–1986 i 2014–1986.

Dla sekwencji zawierających wyłącznie zdjęcia analogowe, takich jak 1994–1986, wartości RMSE były wyższe, wynosząc około 0.5 piksela, czyli dwukrotnie więcej niż w przypadku zdjęć cyfrowych. Największe błędy zaobserwowano dla zdjęć z 1976 roku, których niska jakość radiometryczna znacząco wpłynęła na wzrost RMSE w sekwencjach 2022–1976, 2014–1976 i 1986–1976.

Analiza nie wykazała jednoznacznego wpływu rozdzielczości przestrzennej zdjęć na dokładność orientacji bloków fotogrametrycznych, co sugeruje brak istotnej korelacji między tymi parametrami. Pomimo różnic w jakości zdjęć i poziomie błędów, orientacja bloków fotogrametrycznych została we wszystkich analizowanych sekwencjach przeprowadzona prawidłowo.

REKOMENDACJA zastosowania algorytmu LoFTR w podejściu *Describe-then-Detect* do orientacji bloków fotogrametrycznych

W tabeli 9.7 zawarte są wytyczne odnośnie oceny przydatności algorytmów LB LoFTR do orientacji bloków fotogrametrycznych odnoszące się do takich parametrów jak czas detekcji, liczba punktów i ich rozmieszczenie oraz wartości błędu transformacji. Do klasyfikacji wyników zastosowano czterostopniową, porządkową skalę ocen, gdzie 0-2 oznacza ocenę negatywną, 3 – dostateczną, 4 – dobrą, a 5 – bardzo dobrą.

Detektor	LoFTR									
Sekwencje	2022_2014	2022_1994	2022_1986	2022_1976	2014_1994	2014_1986	2014_1976	1994_1986	1994_1976	1986_1976
Czas detekcji	5	5	4	4	4	4	4	4	4	4
Liczba punktów i ich rozmieszczenie	5	4	5	5	5	5	5	5	5	5
RMSE	5	4	5	4	4	5	4	4	5	4
SUMA	15	13	14	13	13	14	13	13	14	13
			rekomendowane			nierekomendowane				

Tab. 9.8 Rekomendacja wyboru algorytmu LB LoFTR w podejściu *Describe-then-Detect* do orientacji bloków fotogrametrycznych w danej sekwencji

255

9.5 Wnioski i podsumowanie

W podrozdziale przedstawiono ogólne wnioski dotyczące sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych, polegającej na przenoszeniu elementów orientacji z bloku referencyjnego na blok orientowany przy wykorzystaniu tzw. punktów przejściowych. Zdjęcia z lat 2022, 2014, 1994, 1986 i 1976, tworzące analizowane bloki, charakteryzowały się różną datą ich wykonania, różną skalą oraz odmiennymi właściwościami radiometrycznymi.

Automatyczna detekcja punktów przejściowych w parach true-ortofotomap pozwoliła na określenie ich współrzędnych 2D, a dzięki NMPT także wartości (z), umożliwiając wyznaczenie współrzędnych 3D (xyz). Na podstawie tych punktów obliczono parametry transformacji przestrzennej między blokami w postaci macierzy transformacji (4×4). Implementacja tych transformacji pozwala na wprowadzenie georeferencji do bloku orientowanego.

W zaproponowanej metodzie sekwencyjnej orientacji bloków fotogrametrycznych, opartej na zmodyfikowanym podejściu *Structure-from-Motion* (SfM), do wyznaczania punktów przejściowych zastosowano trzy podejścia (metody) wykorzystywane w widzeniu maszynowym: *Detect-then-Describe*, *Detect-and-Describe*, *Describe-then-Detect*. Zaprezentowane w pracy trzy podejścia obejmują wszystkie powszechnie stosowane metodyki pomiarowe orientacji zdjęć lotniczych uwzględniające metody HC i LB oraz ich kombinacje.

Wnioski uzyskane z analiz poszczególnych grup metod orientacji (rozdziały 9.1–9.4) wykazują zasadniczą zgodność, to znaczy, że poszczególne metody pomiaru są zależne m.in. od rozdzielczości zdjęć, jakości radiometrycznej zdjęć, rodzaju zdjęć (cyfrowe - analogowe), a także od przedziału czasowego danych. Należy jednak zaznaczyć, że niezawodność i dokładność każdej z tych metod zależy od tych parametrów. W niniejszym podrozdziale dokonano syntetycznego zestawienia tych metod oraz sformułowano uogólnione wnioski, które mogą być zastosowane w różnorodnych scenariuszach georeferencji archiwalnych zdjęć lotniczych.

Podobieństwa między trzema podejściami

Wszystkie trzy podejścia mają pewne wspólne cechy, które wpływają na ich skuteczność w wyznaczaniu punktów przejściowych niezbędnych do sekwencyjnej orientacji bloków fotogrametrycznych.

1. Jakość radiometryczna zdjęć.

Jednym z najważniejszych czynników wpływających na każde podejście jest jakość radiometryczna zdjęć. Algorytmy działające na zdjęciach cyfrowych uzyskują znacznie lepsze wyniki niż w przypadku zdjęć analogowych, zwłaszcza o słabej jakości radiometrycznej. Na starszych zdjęciach, które charakteryzują się niższym kontrastem i mniejszą liczbą szczegółów, liczba wykrytych punktów przejściowych jest mniejsza, co negatywnie wpływa na dokładność orientacji bloków.

2. Rozdzielczość zdjęć (GSD).

Rozdzielczość obrazu odgrywa kluczową rolę w procesie wykrywania punktów przejściowych. W przypadku zdjęć o wysokiej rozdzielczości algorytmy wykrywają więcej punktów, co prowadzi do poprawy dokładności orientacji. Jednak poszczególne podejścia są mniej lub bardziej zależne od tego parametru – niektóre algorytmy, jak FAST, są bardziej efektywne dla mniejszych wartości GSD zdjęć, podczas gdy LoFTR w podejściu *Describe-then-Detect* jest mniej zależny od rozdzielczości.

3. Odstępy czasowe między zdjęciami.

Im większa różnica czasu między wykonaniem zdjęć, tym trudniej znaleźć wspólne punkty, co wynika głównie ze zmian w krajobrazie i różnic radiometrycznych. Każde podejście wykazuje spadek skuteczności przy bardzo dużych odstępach czasowych między danymi. Jednak algorytmy LoFTR w *Describe-then-Detect* oraz SuperPoint w *Detect-and-Describe* radzą sobie lepiej niż klasyczne metody oparte na detekcji punktów kluczowych (*Detect-then-Describe*).

4. Dokładność pomiaru punktów przejściowych.

Wszystkie metody osiągają niskie wartości RMSE, czyli zapewniają wysoką dokładność. W większości przypadków wartości RMSE wynoszą poniżej 0.6 piksela, a w najlepszych przypadkach spadają poniżej 0.5 piksela. Największe błędy pojawiają się tam, gdzie zdjęcia mają niską jakość radiometryczną lub gdy występują duże odstępy czasowe.

5. Rozmieszczenie punktów przejściowych.

Niezależnie od wybranego podejścia, punkty przejściowe najczęściej są wykrywane w obszarach o dużej szczegółowości, takich jak dachy budynków, drogi i znakach powierzchniowych infrastruktury drogowej. Trudniej znaleźć stabilne punkty w terenach

rolniczych, na łąkach czy w obszarach o jednolitej strukturze. Podejście *Describe-then-Detect* z algorytmem LoFTR wykazuje największą równomierność rozmieszczenia punktów, podczas gdy *Detect-then-Describe* często wykrywa je w miejscach o wysokim kontraście.

6. Skuteczność w orientacji bloków fotogrametrycznych.

Każda metoda pozwala na skuteczną orientację bloków zdjęć, ale różnią się one pod względem liczby wykrytych punktów i ich dokładności. W przypadku zdjęć wysokiej jakości wszystkie podejścia osiągają porównywalne wyniki, ale przy trudniejszych warunkach (np. duże odstępów czasowe) niektóre metody wykazują większą odporność.

7. Różnice między trzema podejściami.

Choć wszystkie metody mają wspólne cechy, to różnią się sposobem wykrywania i opisywania punktów charakterystycznych, co wpływa na ich skuteczność w różnych warunkach. Aspekty porównawcze trzech metod pomiarowych *Detect-then-Describe*, *Detect-and-Describe*, *Describe-then-Detect* zostały przedstawione w tabeli 9.9.

Tab. 9.9 Ocena metod wykrywania punktów przejściowych

		Aspekty porównawcze				
		Metoda wykrywania punktów przejściowych	Wrażliwość na jakość radiometryczną	Wpływ rozdzielczości na skuteczność	Skuteczność przy dużych odstępach czasowych	Liczba wykrytych punktów i dokładność
Podejścia	Detect-then-Describe	Działa dwuetapowo – najpierw identyfikuje punkty kluczowe, a potem opisuje ich cechy. Pozwala to na elastyczny dobór detektora i deskryptora, ale skuteczność zależy od ich kombinacji	Charakteryzują się ograniczoną skutecznością – niektóre detektory, jak FAST, działają dobrze na zdjęciach o niskim kontraście, ale inne, jak SURF mają trudności.	Jest najbardziej zależne od rozdzielczości zdjęć – większa rozdzielczość oznacza więcej punktów przejściowych, ale różne detektory reagują na to w odmienny sposób	Może mieć problemy przy bardzo dużych odstępach czasowych pomiędzy danymi, ponieważ detektory (np. FAST, SURF) nie zawsze znajdują dopasowania	Wykrywa więcej punktów dla zdjęć cyfrowych, ale mniej dla starszych zdjęć analogowych
	Detect-and-Describe	Łączy te dwa etapy w jeden, co zapewnia stabilniejsze dopasowanie punktów, ale jest bardziej wrażliwe na jakość radiometryczną	Jest bardziej wrażliwe na różnice radiometryczne, zwłaszcza dla zdjęć analogowych.	Wykazuje ograniczoną skuteczność – dla zdjęć o wysokiej rozdzielczości wykrywa więcej punktów, ale nie jest to kluczowy czynnik	Wykazuje wyższą skuteczność niż podejście Detect-then-Describe, ale dla zdjęć o słabej jakości radiometrycznej wykrywa mało punktów przejściowych	Zapewnia stabilną liczbę punktów, ale ich liczba spada dla zdjęć o niskiej jakości radiometrycznej
	Describe-then-Detect	Odwraca kolejność – najpierw opisuje obrazy, a potem identyfikuje punkty, co pozwala wykrywać więcej punktów, zwłaszcza w trudnych warunkach	Również wymaga dobrej jakości radiometrycznej zdjęć, ale lepiej radzi sobie z dużymi odstępami czasowymi między danymi	Jest najmniej wrażliwe na rozdzielczość – liczba punktów przejściowych zależy głównie od struktury zdjęcia, a nie od samej rozdzielczości	Działa najlepiej przy dużych odstępach czasowych pomiędzy danymi, ponieważ algorytm LoFTR wykrywa dopasowania nawet tam, gdzie inne metody zawodzą	Generuje najwięcej punktów na zdjęciach o wyraźnej strukturze i są bardziej równomierne w rozmieszczeniu

Każda metoda ma swoje zalety i ograniczenia. *Detect-then-Describe* oferuje elastyczność w doborze detektora i deskryptora, lecz może napotykać trudności przy dużych odstępach czasowych pozyskanych danych. Z tego też powodu zalecana jest przy przetwarzaniu danych nie w pełni automatyczny sposób, lecz w procesie „nadzorowanym”. *Detect-and-Describe* cechuje się wysoką dokładnością, jednak słabiej radzi sobie w warunkach niskiej jakości

radiometrycznej zdjęć. *Describe-then-Detect* charakteryzuje się najwyższą odpornością na duże odstępy czasowe między zbiorami danych oraz związane z nimi zmiany urbanistyczne, nawet w przypadku zdjęć o niskiej jakości radiometrycznej. Wybór optymalnego podejścia zależy od warunków obrazowania i celu analizy.

W tabeli 9.10 zawarta jest zbiorcza ocena kryteriów przydatności algorytmów CV do orientacji bloków fotogrametrycznych trzech metod pomiaru dla której liczba uzyskanych punktów oznaczono:

- przedziałem (150-120) – miejsce I (najlepsze rozwiązanie),
- przedziałem (119-100) – miejsce II (dobre rozwiązanie),
- przedziałem (99-80) – miejsce III (słabe rozwiązanie).

Tab. 9.10 Zbiorcza ocena kryteriów przydatności algorytmu CV do orientacji bloków fotogrametrycznych dla trzech metod pomiaru: *Detect-then-Describe*, *Detect-and-Describe*, *Describe-then-Detect*

Sekuencja bloków	Parametry zdjęć			Detect-then-Describe				Detect-then-Describe						Detect-and-Describe	Describe-then-Detect
				Algorytmy Hand Crafted				Algorytmy Learned Based							
	Rodzaj zdjęcia	Rozdzielczość [cm]	Jakość radiometryczna	FAST	AFAST	SURF	ASURF	HardNet	HardNet8	HyNet	MKDDescriptor	SOSNet	TFeat	SuperGlue	LoFTR
2022 2014	cyfrowe cyfrowe	5 24	bardzo dobra bardzo dobra	15	14	14	14	14	14	15	14	14	14	15	15
2022 1994	cyfrowe analogowe	5 7	bardzo dobra dobra	15	12	8	8	13	13	14	12	11	12	14	13
2022 1986	cyfrowe analogowe	5 5	bardzo dobra dobra	14	14	6	13	6	6	6	6	6	6	13	14
2022 1976	cyfrowe analogowe	5 25	bardzo dobra słaba	7	6	6	12	6	6	11	6	6	6	7	13
2014 1994	cyfrowe analogowe	24 7	bardzo dobra dobra	14	10	12	12	13	13	10	13	13	13	14	13
2014 1986	cyfrowe analogowe	24 5	bardzo dobra dobra	14	13	12	12	6	6	7	8	7	6	15	14
2014 1976	cyfrowe analogowe	24 25	bardzo dobra słaba	7	6	6	14	6	6	10	6	6	6	7	13
1994 1986	analogowe analogowe	7 5	dobra dobra	14	11	12	12	8	8	9	8	7	7	10	13
1994 1976	analogowe analogowe	7 25	dobra słaba	7	6	6	14	6	6	7	6	13	6	7	14
1986 1976	analogowe analogowe	5 25	dobra słaba	7	6	6	6	6	6	7	6	6	6	14	13
Suma				114	98	88	117	84	84	96	85	89	82	116	135
Ocena końcowa				II	III	III	II	III	III	III	III	III	III	II	I

Odnosząc się do tabeli 9.4 najlepszym rozwiązaniem okazało się podejście *Describe-then-Detect* wykorzystujące algorytm sieciowy LoFTR. Drugie miejsce przypadło algorytmowi SuperGlue w podejściu *Describe-then-Detect* oraz FAST i ASURF w podejściu *Detect-then-Describe*. Pozostałe algorytmy (AFAST, HardNet, HardNet8, HyNet, MKDDescriptor, SOSNet) mogą być również zastosowane, ale nie we wszystkich przypadkach. W związku z tym rekomendowanymi algorytmami, które mogą być wykorzystane do sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych są: LoFTR, SuperGlue i FAST i ASURF.

9.5.1 Podsumowanie

Na podstawie przeprowadzonych badań można uznać, że algorytm LoFTR (*Describe-then-Detect*) jest najbardziej wszechstronnym i skutecznym rozwiązaniem w zakresie sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych. Do jego głównych zalet należy zaliczyć wysoką dokładność niezależnie od jakości zdjęć, odporność na zmiany radiometryczne (co umożliwia analizę materiałów archiwalnych), efektywność w detekcji punktów przejściowych w obszarach zabudowy i sieci dróg oraz stabilność wyników nawet w przypadku sekwencji o dużych odstępach czasowych. Dzięki tym cechom LoFTR stanowi optymalne narzędzie do orientacji sekwencji bloków fotogrametrycznych, zapewniając ich wysoką spójność georeferencyjną.

Algorytmy takie jak FAST, AFAST czy SURF (w podejściu *Detect-then-Describe*) dobrze sprawdzały się w przypadku zdjęć cyfrowych o wysokiej rozdzielczości, jednak wykazywały istotne trudności w analizie starszych zdjęć analogowych. Ich skuteczność była również ograniczona w przypadku bloków fotogrametrycznych o dużych odstępach czasowych. Choć algorytm HyNet charakteryzował się najwyższą odpornością na zmiany radiometryczne, jego czas działania był istotnie dłuższy w porównaniu do LoFTR. Warto zaznaczyć, że LoFTR potrafił skutecznie wykrywać punkty przejściowe w sekwencjach, w których algorytm SURF zawodził.

Z kolei w odniesieniu do podejścia *Detect-and-Describe* (SuperPoint wraz SuperGlue), zastosowanie algorytmu LoFTR zapewniło równie wysoką dokładność, jednak skuteczność tego podejścia była silnie uzależniona od liczby i rozmieszczenia punktów przejściowych. W przypadku zdjęć archiwalnych o niskiej jakości radiometrycznej pojawiały się istotne trudności. Algorytm LoFTR okazał się bardziej niezawodny, wykrywając punkty tam, gdzie SuperGlue nie był w stanie ich zidentyfikować.

Pewnym ograniczeniem LoFTR są nieco większe błędy orientacji w niektórych sekwencjach bloków fotogrametrycznych w porównaniu do innych algorytmów. Dlatego w celu uzyskania jak najwyższej dokładności orientacji zaleca się zastosowanie rozwiązania hybrydowego, łączącego podejścia *Detect-then-Describe*, *Detect-and-Describe* oraz *Describe-then-Detect* (tab. 9.4).

Uzyskane wyniki sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych charakteryzują się bardzo wysoką dokładnością, mimo istotnych trudności wynikających z charakterystyki zdjęć archiwalnych, takich jak zróżnicowana rozdzielczość, rodzaj zdjęć (analogowe-cyfrowe), jakość fotograficzna oraz znaczne odstępy czasowe między obrazami. Potwierdza to wysoką skuteczność metod opartych na technikach komputerowego rozpoznawania obrazów (CV), w szczególności podejścia SfM oraz zastosowania kombinacji metod klasycznych (HC) i uczenia maszynowego (LB). Zaproponowane rozwiązanie sprawdziło się w kontekście georeferencji zdjęć archiwalnych, osiągając dokładności przewyższające te uzyskiwane za pomocą tradycyjnych metod opracowania współczesnych, pionowych zdjęć lotniczych.

10. Weryfikacja przyjętej koncepcji - ocena przeniesienia orientacji bloku fotogrametrycznego z roku 2022 na blok z roku 1976 metodą sekwencyjną

Celem rozprawy doktorskiej było opracowanie kompleksowej koncepcji orientacji wieloczasowych zdjęć lotniczych, tworzących bloki fotogrametryczne, w taki sposób, aby możliwe było ich przekształcenie w kartometryczne źródło produktów przestrzennych. Proponowane podejście zakładało wykorzystanie algorytmów oraz metod analizy obrazów rozwijanych w ramach widzenia maszynowego, co umożliwiło skuteczne uwzględnienie różnic czasowych pozyskania zdjęć lotniczych, oraz różnic geometrycznych (tj. rozdzielczość, pokrycie, afinizm itp.) i radiometrycznych pomiędzy zobrazowaniami.

Zgodnie z przyjętą metodyką sekwencyjnej orientacji bloków wieloczasowych (por. rozdział 5), jej realizacja wymagała identyfikacji wspólnych punktów (przejściowych) – tj. charakterystycznych cech obrazowych pomiędzy blokami. Stanowiły one podstawę do przeprowadzenia transformacji jednego bloku względem drugiego i wyznaczenia macierzy transformacji (4x4). Do wyznaczenia punktów przejściowych (COP – *Control Orientation Points*) (patrz rozdział 5.3) wykorzystano trzy powszechnie stosowane metody w *Computer Vision* (CV) tj.: *Detect-then-Describe*, *Detect-and-Describe* oraz *Describe-then-Detect* (patrz rozdział 4). Współrzędne (x, y) tych punktów zostały wyznaczone na podstawie true-ortofotomap, wygenerowanych dla każdego z bloków fotogrametrycznych, a współrzędna wysokościowa (z) została wyinterpolowana na podstawie Numerycznego Modelu Pokrycia Terenu (NMPT), również utworzonego w oparciu o dany blok. Wyznaczone punkty COP (3D) posłużyły do określenia parametrów transformacji przestrzennej pomiędzy parami bloków fotogrametrycznych.

Proces przenoszenia orientacji polegał na przypisaniu współrzędnych terenowych X, Y, Z punktom przejściowym (COP) wyznaczonym pomiędzy orientowaną parą bloków, uzyskanym na etapie sekwencyjnej orientacji bloków fotogrametrycznych (zob. rozdział 9), na podstawie znanych elementów orientacji bloku referencyjnego. Następnie, dysponując współrzędnymi terenowymi tych punktów, przeprowadzono siedmioparametrową transformację bloku

orientowanego w danej sekwencji do bloku referencyjnego. Efektem była macierz transformacji w postaci macierzy jednorodnej 4×4 (zob. podrozdział 5.3).

Zastosowanie wyznaczonych macierzy transformacji umożliwiło przeniesienie orientacji z bloku referencyjnego na kolejne bloki orientowane, zapewniając spójność przestrzenną danych pozyskanych w różnych okresach czasowych. W ten sposób każdemu zdjęciu wchodzącemu w skład danego bloku orientowanego przypisano orientację wynikającą z bloku referencyjnego.

W badaniach blokiem referencyjnym był blok zdjęć z 2022 roku, dla którego znane były Elementy Orientacji Zewnętrznej (EOZ) zdjęć, udostępnione przez Główny Urząd Geodezji i Kartografii (GUGiK).

W niniejszej rozprawie doktorskiej zaproponowano autorską metodę orientacji archiwalnych zdjęć lotniczych, opartą na sekwencyjnym przenoszeniu orientacji pomiędzy blokami fotogrametrycznymi utworzonymi na podstawie zdjęć lotniczych pozyskanych w różnych latach (zob. rozdział 9). Polegało to, na wyznaczaniu elementów orientacji kolejnych bloków (bloków orientowanych) w oparciu o blok referencyjny, w układzie chronologicznym – od najnowszego do najstarszego – z wykorzystaniem bloków pośrednich. Elementy orientacji bloku referencyjnego w pierwszej sekwencji były przenoszone na blok orientowany, który następnie stawał się blokiem referencyjnym w kolejnej sekwencji. W taki sposób elementy orientacji zewnętrznej bloku z roku 2022 zostały przeniesione na pozostałe bloki. W rezultacie utworzono sekwencje par bloków obejmujące lata: 2022, 2014, 1994, 1986 oraz 1976 zgodnych z tabelą 9.1 (zob. rozdział 9).

Ze względu na wysoką rozdzielczość przestrzenną zdjęć cyfrowych w bloku referencyjnym z 2022 roku (GSD 5 cm), błąd orientacji tego bloku jest relatywnie niewielki – w przybliżeniu około 0.2–0.3 piksela, co odpowiada wartości rzędu 1–2 cm, w porównaniu do błędów występujących w pozostałych blokach. W związku z tym, w dalszych analizach przyjęto, że wpływ błędu orientacji tego bloku (2022) jest pomijalny. W pozostałych sekwencjach dokładność transformacji pomiędzy blokami wyznaczono zgodnie z teorią przenoszenia błędów średnich Gaussa jako suma błędu nawiązania (bloku referencyjnego) oraz błędu wpasowania (bloku orientowanego). (tab. 10.1).

Ocenę dokładności sekwencyjnej orientacji bloków wykonano na podstawie średniego liniowego (XYZ) błędu kwadratowego (RMSE) transformacji 3D, uzyskanego na niezależnych punktach kontrolnych CHOP (*Check Orientation Points*), stanowiących 10% wszystkich wyznaczonych punktów przejściowych. Wyniki tych błędów przedstawiono zarówno w jednostkach pikselowych, jak i terenowych w tabeli 10.1. Wartość pikselowa i terenowa uzyskanego błędu orientacji w danej sekwencji odnosiła się do bloku o niższej rozdzielczości zdjęć. Na przykład, dla sekwencji 1986 (7 cm) – 1976 (25 cm) przyjęto rozdzielczość 25 cm, natomiast dla sekwencji 1994 (7 cm) – 1986 (5 cm) przyjęto rozdzielczość 7 cm. Liczba oraz rozmieszczenie punktów COP dla danej sekwencji bloków fotogrametrycznych, na których wykonano transformację dla poszczególnych algorytmów CV, są tożsame z ich liczbą i rozmieszczeniem opisanym w rozdziale 9 (9.1, 9.2, 9.3, 9.4).

Tab. 10.1 Zbiorcze wyniki wartości liniowej RMSE (XYZ) sekwencyjnej orientacji bloków fotogrametrycznych z wykorzystaniem algorytmów CV dla trzech metod pomiaru: Detect-then-Describe, Detect-and-Describe, Describe-then-Detect uwzględniający wpływ błędów bloku referencyjnego i orientowanego. Symbol „-” oznacza brak orientacji

Sekwencja bloków	Parametry zdjęć			RMSE (XYZ) $\left[\frac{cm}{piksele}\right]$											
				Detect-then-Describe				Detect-then-Describe				Detect-and-Describe	Describe-then-Detect		
				Algoritmy Hand Crafted				Algoritmy Learned Based							
	Rodzaj zdjęcia	Rozdzielczość	Jakość radiometryczna	FAST	AFAST	SURF	ASURF	HardNet	HardNet8	HyNet	MKDDescriptor	SOSNet	TFeat	SuperGlue	LoFTR
2022 2014	Cyfrowe cyfrowe	5 24	bardzo dobra bardzo dobra	7	6	6	6	6	6	6	6	6	6	7	4
				0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.2
2022 1994	Cyfrowe analogowe	5 7	bardzo dobra dobra	1	4	-	-	2	2	2	1	-	1	3	5
				0.1	0.6			0.3	0.3	0.3	0.1		0.1	0.4	0.7
2022 1986	Cyfrowe analogowe	5 5	bardzo dobra dobra	1	1	-	1	-	-	-	-	-	-	1	1
				0.2	0.2		0.2							0.2	0.2
2022 1976	Cyfrowe analogowe	5 25	bardzo dobra słaba	-	-	-	6	-	-	-	-	-	-	-	13
							0.3								0.5
2014 1994	Cyfrowe analogowe	24 7	bardzo dobra dobra	14	-	20	13	12	12	-	11	12	11	14	13
				0.6		0.8	0.5	0.5	0.5		0.5	0.5	0.6	0.5	
2014 1986	Cyfrowe analogowe	24 5	bardzo dobra dobra	9	8	8	16	-	-	-	-	-	-	8	7
				0.4	0.3	0.3	0.7							0.3	0.3
2014 1976	Cyfrowe analogowe	24 25	bardzo dobra słaba	-	-	-	8	-	-	-	-	-	-	-	17
							0.3								0.7
1994 1986	Analogowe analogowe	7 5	dobra dobra	3	-	20	14	-	-	-	-	-	-	-	6
				0.4		3.0	2.0								0.9
1994 1976	Analogowe analogowe	7 25	dobra słaba	-	-	-	14	-	-	-	-	13	-	-	9
							0.6					0.5			0.4
1986 1976	Analogowe analogowe	5 25	dobra słaba	-	-	-	-	-	-	-	-	-	-	5	17
														0.2	0.7

Na podstawie przeprowadzonej analizy sekwencyjnej orientacji bloków fotogrametrycznych utworzonych z archiwalnych zdjęć lotniczych (zob. rozdział 9) oraz analizy danych zawartych

w tabeli 10.1, przedstawiającej wartości błędów orientacji uzyskanych przy zastosowaniu algorytmów widzenia maszynowego, można stwierdzić, że wpływ na wartość tych błędów mają: typ zdjęć (analogowe vs. cyfrowe), ich rozdzielczość przestrzenna, ich jakość geometryczna i radiometryczna, zastosowany algorytm, a także zmiany w pokryciu terenu (np. intensyfikacja procesów urbanizacyjnych).

Wymienione czynniki, a w szczególności odstęp czasowy pomiędzy pozyskanymi danymi, wpływają bezpośrednio na liczbę i rozmieszczenie wykrytych punktów przejściowych (patrz podrozdział 9.1, 9.2, 9.3, 9.4). W tabeli 10.1 miejsca oznaczone „-” wskazują na brak możliwości przeprowadzenia orientacji bloków (brak wartości RMSE), co wynika z niewystarczającej liczby punktów przejściowych pomierzonych przez zastosowane algorytmy, niezbędnych do wykonania transformacji bloków (patrz rozdział 9).

Najmniejsze wartości błędu średniokwadratowego (RMSE) odnotowano dla sekwencji bloków 2022–2014, utworzonych na podstawie zdjęć cyfrowych o bardzo dobrej jakości fotogrametrycznej, gdzie RMSE mieściły się w przedziale od 0.2 do 0.3 piksela dla wszystkich analizowanych algorytmów.

Największe błędy orientacji wystąpiły w sekwencji bloków 1994–1986 dla algorytmów SURF (RMSE równy 3.0 piksele) i ASURF (RMSE wynoszący 2.0 piksele). Wynikało to z faktu, że dla tej sekwencji orientację bloku 1986 wyznaczono na podstawie orientacji bloku 1994 pochodzącej z sekwencji 2014–1994 (GSD 24 cm). W tym przypadku do orientacji nie można było skorzystać z sekwencji 2022–1994 charakteryzującej się mniejszym błędem, ponieważ sekwencja ta nie została utworzona. W rezultacie, na skutek propagacji większego błędu z sekwencji 2014–1994 (spowodowanego m.in. niższą rozdzielczością zdjęć), błędy orientacji bloku 1986 w sekwencji 1994–1986 wyniosły 3 piksele dla detektora SURF oraz 2 piksele dla detektora ASURF. W pozostałych sekwencjach błąd orientacji bloków był poniżej 1 piksela.

Sekwencja 2022–2014 (zdjęcia cyfrowe) była jedyną, dla której uzyskano wartości RMSE we wszystkich zastosowanych algorytmach CV. Zaobserwowano wyraźną zależność (tabela 10.1): im niższa rozdzielczość zdjęć tworzących daną sekwencję, tym większy błąd orientacji. Spośród wszystkich algorytmów jedynie algorytm LoFTR umożliwił wyznaczenie wartości RMSE we wszystkich analizowanych sekwencjach.

W kontekście przenoszenia orientacji pomiędzy najbardziej odległymi czasowo blokami 2022 i 1976 pełną bezpośrednią orientację udało się uzyskać jedynie dla algorytmów ASURF (RMSE równy 0.3 piksela, podejście *Detect-then-Describe*) oraz LoFTR (RMSE równy 0.5 piksela, podejście *Describe-then-Detect*) – por. tabela 10.1.

Na podstawie analizy tabeli 10.1 stwierdzono, że sekwencyjne przeniesienie orientacji z bloku z roku 2022 na blok z roku 1976 dla wybranego algorytmu, z uwzględnieniem bloków pośrednich, udało się wyłącznie w przypadku metody *Describe-then-Detect*, przy zastosowaniu algorytmu LoFTR. Dla pozostałych dwóch metod pomiarowych nie uzyskano pełnego ciągu sekwencyjnej orientacji (2022–2014–1994–1986–1976) dla żadnego z zastosowanych algorytmów.

Sekwencyjne przeniesienie orientacji z bloku 2022 na blok 1976 z wykorzystaniem bloków pośrednich może być również wykonane jako kombinacji dedykowanych algorytmów z zastosowaniem 3 metod pomiarowych CV (tab. 10.2). Proces ten powinien przebiegać kolejno poprzez sekwencje bloków: 2022–2014, 2014–1994, 1994–1986 oraz 1986–1976 (patrz tab. 10.1). W ten sposób zewnętrzne parametry orientacji bloku z roku 2022 zostają sukcesywnie przeniesione i zastosowane dla bloku zdjęć z roku 1976. Dokładność przeniesienia orientacji z bloku zdjęć z 2022 roku na blok z 1976 roku, z wykorzystaniem bloków pośrednich, przedstawiono w tabeli 10.2 jako wartość liniowego błędu RMS (XYZ). Na końcowy błąd orientacji bloku z 1976 roku wpływa nie tylko błąd jego własnej orientacji, lecz również skumulowany błąd orientacji bloków pośrednich (2014, 1994, 1986), którego wartości zestawiono w ostatniej kolumnie tabeli 10.2. Wartość błędu w tabeli 10.2 kolumna 6 wyrażono w miarach terenowych (cm) lub pikselach przy znanej rozdzielczości przestrzennej zdjęć z roku 1976.

Tab. 10.2 Wyniki wartości liniowej RMSE (XYZ) sekwencyjnej orientacji bloku 1976 (GSD 25 cm), uzyskane dla czterech serii czasowych (2022–2014–1994–1986–1976), z wykorzystaniem pięciu przykładowych kombinacji algorytmów HC i LB

(1) Numer kombinacji	Sekwencja bloków				(6) RMSE (XYZ) Blok 1976 [cm/piksel] (GSD = 25 cm)
	(2) 2022 2014	(3) 2014 1994	(4) 1994 1986	(5) 1986 1976	
1	LoFTR	LoFTR	LoFTR	LoFTR	23/0.9
2	FAST	SURF	ASURF	Super Glue	16/0.6
3	AFast	ASURF	ASURF	Super Glue	14/0.6
4	Super Glue	Super Glue	ASURF	Super Glue	16/0.6
5	LoFTR	MKD Descriptor	FAST	Super Glue	13/0.5

W tabeli 10.2 przedstawiono wartości liniowe RMSE (XYZ) sekwencyjnej orientacji bloku 1976, uzyskanych dla czterech serii czasowych (2022–2014–1994–1986–1976). Kolumna 1 tabeli donosi się do kombinacji zastosowanej metody HC i LB. Kolumna 2, 3, 4, 5 odnosi się do wykorzystanego algorytmu CV do orientacji w danej sekwencji czasowej. W kolumnie 6 zestawiono wartości błędu orientowanego bloku 1976 odniesioną do jego rozdzielczości terenowej 25 cm.

Analiza wyników sekwencyjnej orientacji przedstawionych w tabeli 10.2 wskazuje, że najskuteczniejszą kombinacją algorytmów, umożliwiającą przeniesienie orientacji z roku 2022 i kolejne bloki, aż do bloku 1976 była kombinacja nr. 5, obejmująca wszystkie trzy metody *Detect-then-Describe*, *Detect-and-Describe* oraz *Describe-then-Detect*. Dla tego wariantu uzyskano dokładność orientacji bloku z 1976 roku na poziomie 13 cm (ok. 0.5 piksela), co należy uznać za wynik bardzo dobry, potwierdzający poprawność przeprowadzonej orientacji i poprawność zastosowanej metodologii. Zastosowanie pozostałych kombinacji 3 i 4 (*Detect-then-Describe* wraz *Detect-and-Describe*) również pozwala na osiągnięcie wysokich dokładności i może być zastosowane do orientacji tego bloku. Największy błąd orientacji uzyskano w kombinacji 1 dla podejścia *Describe-then-Detect* (LoFTR) i wyniósł on 23 cm (około 0.9 piksela).

Oprócz przeniesienia orientacji z Bloku 2022 na Blok 1976 z wykorzystaniem pełnego ciągu czasowego (obejmującego wszystkie pośrednie epoki: 2022–2014–1994–1986–1976), możliwe

było również zastosowanie jego skróconych wariantów, opartych na sekwencjach 2022–1986–1976 lub 2022–1994–1976. Tabela 10.3 zawiera wyniki orientacji bloku 1976 dla dwóch wspomnianych wariantów sekwencji, z uwzględnieniem zastosowanych algorytmów CV. Pierwsza i druga kolumna tabeli 10.3 odnoszą się do rozpatrywanych sekwencji bloków oraz zawierają listę algorytmów wykorzystanych do wyznaczenia punktów przejściowych niezbędnych do realizacji transformacji. W trzeciej kolumnie przedstawiono wartości liniowego błędu średniokwadratowego (RMSE w układzie XYZ) uzyskane dla orientacji sekwencyjnej bloku 1976, z wykorzystaniem tych algorytmów (wyróżnione pogrubieniem – kolumna 1 i 2), które w danej sekwencji zapewniły najwyższą dokładność transformacji.

W przypadku sekwencji 2022–1986–1976 najniższą wartość RMSE uzyskano przy zastosowaniu kombinacji algorytmów HC FAST, MKDDescriptor, TFeat oraz LB LoFTR; błąd orientacji wyniósł 5 cm (0.2 piksela). Dla sekwencji 2022–1994–1976 najlepsze rezultaty uzyskano przy użyciu algorytmów HC FAST oraz LB LoFTR, przy czym wartość RMSE wyniosła 9 cm (0.4 piksela) (tab.10.3).

Tab. 10.3 Wartości liniowego błędu średniokwadratowego (RMSE w układzie XYZ) dla sekwencyjnej orientacji bloku 1976 (GSD 25 cm) z wykorzystaniem kombinacji algorytmów typu HC i LB (dwie serie czasowe). Wyróżniono algorytmy, dla których uzyskano najwyższą dokładność transformacji w danej sekwencji

	Sekwencje		Blok 1976
	(1) 2022–1986	(2) 1986–1976	(3) RMSE (XYZ) [cm / piksel] (GSD = 25 cm)
Algorytmy	FAST AFAST ASURF Super Glue LoFTR	Super Glue LoFTR	5 / 0.2
	2022–1994	1994–1976	RMSE (XYZ) [cm / piksel]
Algorytmy	FAST AFAST HardNet HardNet8 HyNet MKDDescriptor TFeat Super Glue LoFTR	ASURF SOSNet LoFTR	9 / 0.4

Przeprowadzone analizy oraz uzyskane wyniki w zakresie orientacji archiwalnych zdjęć lotniczych umożliwiają dobór odpowiedniej metody komputerowego rozpoznawania obrazów (CV – *Computer Vision*), z uwzględnieniem takich czynników, jak jakość materiałów (cyfrowe vs. analogowe), ich rozdzielczość, różnice radiometryczne oraz zmienność pokrycia terenu wynikającą z procesów urbanizacyjnych. W związku z powyższym, poniżej przedstawiono zestawienie rekomendowanych algorytmów, umożliwiających poprawne przeprowadzenie sekwencyjnej orientacji archiwalnych zdjęć lotniczych.

Tab. 10.4 Rekomendacja algorytmów CV w zależności od czynników mających wpływ na dokładność sekwencyjnej orientacji archiwalnych zdjęć lotniczych

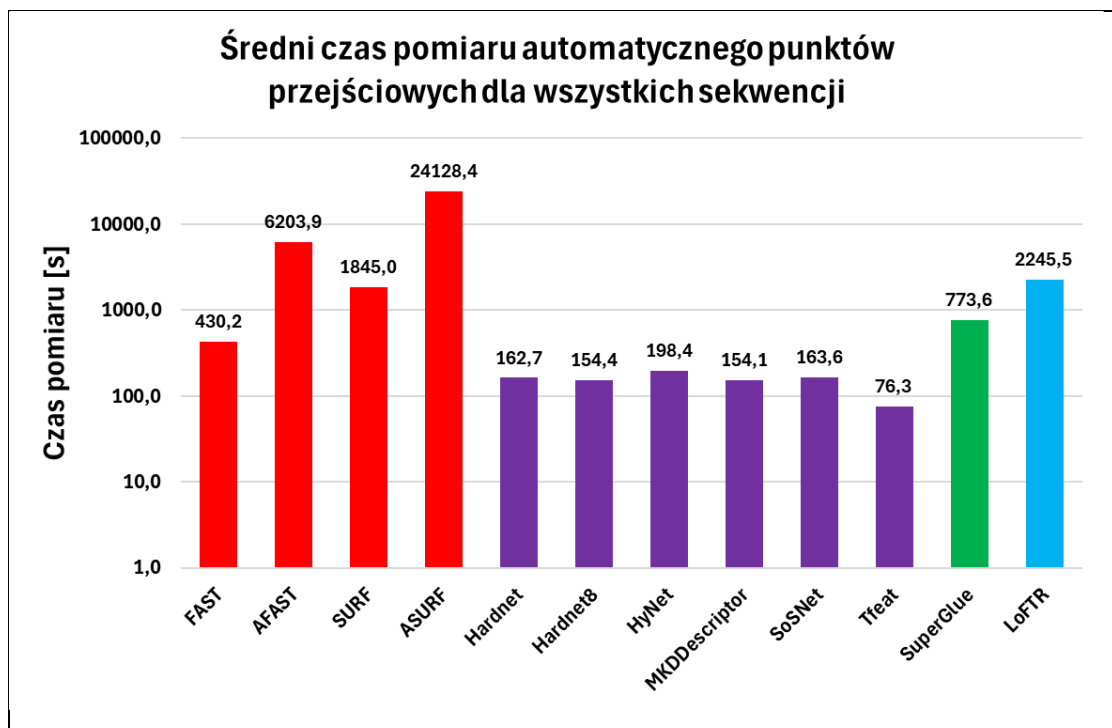
Czynniki wpływające na dokładność orientacji zdjęć lotniczych	Rekomendowane algorytmy
Zdjęcia cyfrowe (wysoka jakość)	FAST, AFAST, SURF, ASURF, HardNet, HardNet8, HyNet, MKDDescriptor, SOSNet, TFeat, SuperGlue, LoFTR
Zdjęcia analogowe dobrej jakości geometrycznej i radiometrycznej	FAST, LoFTR
Zdjęcia analogowe niskiej jakości geometrycznej i radiometrycznej	ASURF, LoFTR
Zdjęcia cyfrowe i analogowe	FAST, AFAST, ASURF, HardNet, HardNet8, MKDDescriptor, TFeat, SuperGlue, LoFTR
Duże zmiany w pokryciu terenu	LoFTR
Duże odstępy czasowe między zdjęciami	LoFTR
Największa dokładność orientacji	Orientacja jednoepokowa (2022–1976): ASURF lub LoFTR. Orientacja dwuepokowa (2022–1986–1976): (2022–1986) – FAST, AFAST, ASURF, Super Glue, LoFTR; (1986–1976) – SuperGlue. Orientacja czteroepokowa (2022–2014–1994–1986–1976): (2022–2014) – LoFTR, (2014–1994) – MKDDescriptor, (1994–1986) – FAST (1986–1976) – SuperGlue

Dodatkowym czynnikiem warunkującym proces orientacji archiwalnych zdjęć lotniczych jest czas wyznaczania punktów przejściowych (COP). Na rysunku 10.1 przedstawiono średni czas pomiaru punktów przejściowych dla algorytmów CV we wszystkich trzech podejściach,

obliczony jako średnia czasów pomiaru ze wszystkich analizowanych sekwencji (patrz tab. 9.2, tab. 9.12, tab. 9.24, tab. 9.30). Najkrótsze czasy odnotowano dla podejścia typu *Detect-and-Describe* z wykorzystaniem algorytmu TFeat, natomiast najdłuższy czas uzyskano dla algorytmu ASURF. Czas działania TFeat był około 316 razy krótszy niż czas działania ASURF, co wskazuje na istotne różnice w efektywności poszczególnych metod. W porównaniu z algorytmem SuperGlue, TFeat działał około 10 razy szybciej, a względem algorytmu LoFTR – około 30 razy szybciej.

W przypadku algorytmów typu *Hand-Crafted* (HC), takich jak ASURF czy AFAST, a także metod opartych na sieciach neuronowych (np. LoFTR, SuperGlue), należy uwzględnić ich relatywnie wysoki koszt obliczeniowy, zwłaszcza w stosunku do algorytmów HC. Czas przetwarzania zależy zarówno od charakterystyki wybranego algorytmu, jak i od wielkości orientowanego bloku, co ma istotne znaczenie w kontekście efektywności ekonomicznej i dostępnych zasobów sprzętowych. Ostateczny wybór algorytmu należy więc do wykonawcy, który powinien uwzględnić zarówno wymagania dokładnościowe, jak i ograniczenia techniczne (sprzętowe).

W przypadku bloków fotogrametrycznych tworzonych z sekwencji zdjęć o niewielkich odstępach czasowych, charakteryzujących się stabilnym pokryciem terenu oraz wysoką jakością radiometryczną, zasadne może być zastosowanie szybszego klasycznego algorytmu HC FAST, lub algorytmów opartych na sieciach neuronowych LB np. SOSNet, TFeat czy HyNet.



Rys. 10.1 Średni czas wyznaczania punktów przejściowych dla zastosowanych algorytmów widzenia maszynowego (CV) z uwzględnieniem wszystkich sekwencji czasowych. Przyjęto oznaczenia kolorystyczne odpowiadające typom algorytmów i ich podejściom pomiarowym: czerwony - HC (*Detect-then-Describe*), purpurowy - LB (*Detect-then-Describe*), zielony - LB (*Detect-and-Describe*), niebieski - LB (*Describe-then-Detect*)

Na podstawie przeprowadzonych badań potwierdzono skuteczność przyjętej koncepcji sekwencyjnej orientacji wieloczasowych zdjęć lotniczych z wykorzystaniem metod widzenia maszynowego. Opracowana koncepcja sekwencyjnej orientacji bloków wieloczasowych pozwala na skuteczne przenoszenie elementów orientacji zewnętrznej z bloku referencyjnego na pozostałe bloki fotogrametryczne, nawet w przypadku znacznych różnic czasowych, geometrycznych i radiometrycznych pomiędzy obrazowaniami. Kluczowym elementem metody jest identyfikacja i dopasowanie punktów charakterystycznych przy wykorzystaniu nowoczesnych algorytmów CV, w tym metod *Detect-then-Describe*, *Detect-and-Describe* oraz *Describe-then-Detect*.

Analiza wyników wskazuje, że skuteczność orientacji w dużej mierze zależy od jakości zdjęć (cyfrowe vs. analogowe), ich rozdzielczości oraz doboru odpowiednich algorytmów dopasowania. Najmniejsze wartości RMSE odnotowano dla sekwencji opartych na zdjęciach cyfrowych, natomiast największe dla sekwencji zawierających zdjęcia analogowe o niższej

jakości. Jedynie algorytm LoFTR pozwolił na uzyskanie pełnej orientacji we wszystkich analizowanych sekwencjach.

W przypadku dużych różnic czasowych oraz braku możliwości bezpośredniej orientacji bloków tworzących daną sekwencję, konieczne staje się wykorzystanie bloków pośrednich, pozyskanych w okresach przejściowych, umożliwiających etapowe przeniesienie orientacji. Takie podejście dowodzi elastyczności i efektywności opracowanej metody.

Całokształt przeprowadzonych analiz potwierdza, że przy odpowiednim doborze algorytmów oraz właściwym podejściu do przetwarzania danych możliwe jest skuteczne i dokładne przeniesienie orientacji pomiędzy blokami zdjęć wykonanych nawet z ponad 40-letnim odstępem czasowym, co otwiera nowe możliwości wykorzystania archiwalnych zdjęć lotniczych jako wiarygodnych źródeł danych kartometrycznych.

W oparciu o uzyskane wyniki rekomenduje się stosowanie kombinacji metod *Detect-then-Describe* oraz *Detect-and-Describe* jako najbardziej skutecznej strategii orientacji bloków zdjęć lotniczych w orientacji wieloczasowej. Pomimo większego błędu, podejście *Describe-then-Detect* z algorytmem LoFTR również stanowi wartościową alternatywę, zwłaszcza w przypadku danych o ograniczonej jakości radiometrycznej.

Na podstawie przeprowadzonych analiz można stwierdzić, że czas wyznaczania punktów przejściowych stanowi istotny czynnik wpływający na efektywność procesu orientacji archiwalnych zdjęć lotniczych. Znaczące różnice w czasie działania poszczególnych algorytmów, zwłaszcza pomiędzy podejściami klasycznymi HC, a metodami opartymi na sieciach neuronowych LB, wskazują na konieczność świadomego doboru narzędzi w zależności od charakterystyki danych oraz dostępnych zasobów obliczeniowych. W przypadku bloków o jednorodnym pokryciu terenu i dobrej jakości radiometrycznej zasadne jest stosowanie szybszych algorytmów, co pozwala na redukcję kosztów obliczeniowych bez istotnej utraty jakości orientacji. Tym samym dobór algorytmu powinien być kompromisem pomiędzy dokładnością, długością czasu przetwarzania a możliwościami technicznymi, co ma szczególne znaczenie w zadaniach związanych z przetwarzaniem dużych zbiorów archiwalnych danych obrazowych.

11. Podsumowanie

Motywacją podjęcia omawianego zagadnienia jest fakt, że w zasobach Głównego Urzędu Geodezji i Kartografii znajdują się obszerne zbiory archiwalnych zdjęć lotniczych o zróżnicowanej rozdzielczości przestrzennej oraz jakości radiometrycznej, dla których brak jest informacji o georeferencji. Jej nadanie metodami klasycznymi, polegającymi na manualnym wyszukiwaniu punktów homologicznych pomiędzy zdjęciami archiwalnymi a współczesnymi materiałami referencyjnymi (np. zdjęciami lotniczymi z przypisaną georeferencją, ortofotomapami), wymaga zazwyczaj czasochłonnej rekonstrukcji polowej osnowy fotogrametrycznej. W wielu przypadkach, ze względu na znaczny upływ czasu i istotne przekształcenia pokrycia terenu, manualne wyznaczenie osnowy okazuje się jednak niemożliwe. Równocześnie dostępne są nowoczesne metody pomiarowe, wywodzące się m.in. z dziedziny *Computer Vision*, umożliwiające przeprowadzenie orientacji zdjęć w oparciu o automatyczny pomiar punktów przejściowych przez algorytmy typu *Hand-Crafted* czy *Learned-Based* bez konieczności odtwarzania tradycyjnej osnowy. Dynamiczny rozwój metod analitycznych, w tym metody *Structure-from-Motion* oraz algorytmów opartych na sztucznej inteligencji, otwiera nowe możliwości efektywnego wykorzystania tych unikalnych zasobów.

Motywacją do podjęcia tematu rozprawy doktorskiej jest zatem potrzeba zagospodarowania rozległych zbiorów archiwalnych zdjęć lotniczych, które charakteryzują się znaczącym potencjałem pomiarowym, poprzez nadanie im georeferencji z zastosowaniem nowoczesnych metod z zakresu *Computer Vision*. Mogą one znaleźć zastosowanie m.in. w planowaniu przestrzennym i urbanistyce (np. analiza kierunków rozwoju osadnictwa), w geologii i gospodarce przestrzennej (np. ocena procesów osuwiskowych, monitorowanie zmian pokrycia terenu), a także w prawie i administracji (np. jako materiał dowodowy w sporach dotyczących przebiegu granic nieruchomości czy sposobu użytkowania gruntów). Z tego względu nadrzędnym celem niniejszej rozprawy doktorskiej jest opracowanie metod umożliwiających nadanie georeferencji archiwalnym zdjęciom lotniczym. Problem ten ma charakter uniwersalny – występuje nie tylko na rynku krajowym, lecz również w ujęciu globalnym.

Głównym problemem orientacji archiwalnych zdjęć lotniczych jest brak w dostępnych zbiorach elementów orientacji zewnętrznej tych zdjęć. Odtworzenie tych parametrów na

podstawie pomiaru naturalnych punktów osnowy fotogrametrycznej (szczegółów terenowych) jest znacznie utrudnione, a często niemożliwe, ze względu na istotne zmiany w pokryciu terenu zachodzące w czasie. Jednym z podejść pozwalających częściowo przezwyciężyć ten problem jest wykorzystanie zdjęć pośrednich (pozyskanych pomiędzy terminem wykonania zdjęć archiwalnych będących przedmiotem zainteresowania a stanem współczesnym), dla których możliwe jest wyznaczenie szczegółów terenowych pełniących rolę fotopunktów. Manualne wyznaczanie punktów osnowy jest jednak procesem pracochłonnym i czasochłonnym oraz obarczonym niską dokładnością, ze względu na różnice w rozdzielczości, rodzaju oraz jakości radiometrycznej dostępnych zdjęć. Ponadto automatyczne metody dopasowywania obrazów stosowane w klasycznych aplikacjach fotogrametrycznych, oparte m.in. na algorytmie *Area-Based Matching*, okazują się niewystarczające w przypadku obrazów pochodzących z różnych okresów, charakteryzujących się dużymi różnicami radiometrycznymi. Postęp w dziedzinie robotyki i przetwarzania obrazów (*Computer Vision*) umożliwił jednak opracowanie bardziej zaawansowanych metod analizy obrazowej, pierwotnie projektowanych dla innych zastosowań niż fotogrametria, które mogą zostać zaadaptowane do rozwiązania tego problemu. W związku z tym podjęto decyzję o wykorzystaniu metody *Structure-from-Motion* oraz jej modyfikacji w celu umożliwienia automatycznej orientacji archiwalnych zdjęć lotniczych.

W niniejszej pracy doktorskiej postawiono tezę, że metody widzenia maszynowego pozwalają z wysoką niezawodnością na automatyczne odtworzenie orientacji archiwalnych zdjęć lotniczych z dokładnością na poziomie rozdzielczości zdjęć (piksela terenowego), a nawet lepszą – podpikselową.

Problemem badawczym przedstawionym w niniejszej rozprawie doktorskiej dotyczy weryfikacji użyteczności metod widzenia maszynowego w procesie orientacji archiwalnych zdjęć lotniczych. W pracy dokonano oceny skuteczności zarówno klasycznych technik przetwarzania obrazu w kontekście pomiaru punktów przejściowych, jak i metod opartych na głębokich sieciach neuronowych. Przeprowadzone badania potwierdziły zasadność zaproponowanej autorskiej metodologii orientacji archiwalnych zdjęć lotniczych, opartej na sekwencyjnym procesie nadania im georeferencji z wykorzystaniem metody *Structure-from-Motion*.

W niniejszej pracy doktorskiej przyjęto autorskie podejście oparte na pracy z blokami zdjęć archiwalnych (serii czasowych), zamiast ograniczać się do pojedynczych zdjęć. Zastosowanie całych bloków fotogrametrycznych zwiększyło obszar dopasowania, co przełożyło się na wzrost liczby wykrywanych punktów wiążących oraz brak wymogu ich równomiernego rozkładu na całym obszarze opracowania (w tym także dopuszczenie występowania podobszarów pozbawionych punktów wspólnych, odpowiadających pojedynczym zdjęciom). W rezultacie opracowanie blokowe umożliwiło przeprowadzenie georeferencji dla wszystkich zdjęć wchodzących w skład danego bloku, co w podejściu opartym na pojedynczych zdjęciach w wielu przypadkach okazałoby się niemożliwe do zrealizowania.

W ramach opracowanej metodyki zaproponowano autorskie podejście do orientacji zdjęć lotniczych, polegające na sekwencyjnej orientacji archiwalnych bloków fotogrametrycznych z wykorzystaniem zmodyfikowanej metody *Structure-from-Motion*. Proces ten polega na chronologicznym przenoszeniu orientacji bloków zdjęć – od najnowszych do najstarszych – przy wykorzystaniu wcześniej zorientowanego bloku referencyjnego. Takie podejście pozwoliło skutecznie przezwyciężyć problemy związane z brakiem wspólnych cech obrazowych w blokach oddalonych czasowo, wynikającym ze zmian urbanistycznych i przekształceń zagospodarowania terenu. Wykorzystanie bloków pośrednich zapewniło ciągłość procesu orientacji i potwierdziło uniwersalność przyjętego podejścia.

Dane testowe wykorzystane do przeprowadzenia badań dotyczyły tego samego obszaru i charakteryzowały się znacznym zróżnicowaniem pod względem okresu ich pozyskania, skali obrazowania, rodzaju danych (zarówno w kontekście kamer, jak i formatów), a także jakości radiometrycznej. Zróżnicowanie parametrów zdjęć wskazuje na celowo przyjętą strategię doboru pól testowych, ukierunkowaną na objęcie możliwie najszerszego spektrum przypadków praktycznych, które mogą wystąpić w rzeczywistych zastosowaniach. Takie podejście stworzyło warunki do wszechstronnej oceny skuteczności i uniwersalności zaproponowanej metodyki, w tym jej zdolności do nadania georeferencji zdjęć.

Zdjęcia wykorzystane w badaniach obejmowały obszar miejski intensywnie przekształcany w wyniku procesów urbanizacyjnych, co czyniło go szczególnie wymagającym przypadkiem praktycznym ze względu na ograniczoną liczbę możliwych do wykrycia punktów charakterystycznych. Złożoność tej sytuacji wynikała z dynamicznych zmian w strukturze

przestrzennej oraz znacznego zróżnicowania pokrycia terenu w czasie. Należy zatem podkreślić, że opracowana metodyka, sprawdzona w warunkach wysokiej zmienności pokrycia terenu (obszar zurbanizowany), tym bardziej może znaleźć skuteczne zastosowanie w odniesieniu do zdjęć lotniczych obejmujących obszary rolnicze i leśne, w których warunki pokrycia terenu wykazują znacznie większą stabilność przestrzenną i czasową.

W ramach badań przeprowadzono szczegółowe analizy związane z budową bloków fotogrametrycznych oraz procesem ich sekwencyjnej orientacji. Zastosowano przy tym szeroki zakres algorytmów (detektorów, deskryptorów) z obszaru widzenia komputerowego, co pozwoliło na ocenę skuteczności metody w różnych wariantach przetwarzania oraz konfiguracjach danych.

W celu przeprowadzenia sekwencyjnej orientacji bloków zdjęć wykonano dwa zasadnicze etapy. W pierwszym etapie wyznaczono elementy orientacji zdjęć przy wykorzystaniu metody niezależnych wiązek w poszczególnych blokach fotogrametrycznych (dla poszczególnych lat) w lokalnym układzie odniesienia, które stanowiły podstawę do wygenerowania Numerycznych Modeli Pokrycia Terenu (NMPT) oraz true-ortofotomap. Drugi etap polegał na automatycznym wyznaczeniu punktów przejściowych pomiędzy true-ortofotomapami z różnych okresów czasowych oraz wyinterpolowanej wysokości (z) z NMPT. Utworzony w ten sposób zbiór punktów 3D (x, y, z) posłużył do wyznaczenia parametrów transformacji przestrzennej pomiędzy odpowiadającymi sobie blokami czasowymi. W ramach eksperymentu badany był czas pomiaru punktów przejściowych, dokładność ich wyznaczenia, ich liczba i rozkład, a także dokładność transformacji pomiędzy orientowanymi blokami na wyznaczonych punktach przejściowych.

Współrzędne (x, y) punktów przejściowych wyznaczane za pomocą algorytmów widzenia maszynowego charakteryzują się podpixselową dokładnością. Choć dokładność współrzędnej (z) zależy od jakości numerycznego modelu powierzchni terenu (NMPT), to dzięki dużej liczbie punktów generowanych w metodach CV wpływ ewentualnych błędów NMPT jest ograniczony – nadmiarowość danych pozwala na skuteczną eliminację punktów odstających bez istotnego obniżenia jakości końcowych wyników.

Jako blok referencyjny wykorzystano blok zdjęć z 2022 roku, dla którego dostępne były Elementy Orientacji Zewnętrznej (EOZ) udostępnione przez Główny Urząd Geodezji i Kartografii

(GUGiK). Proces przenoszenia orientacji polegał na nadaniu współrzędnych terenowych X, Y, Z punktom przejściowym, wyznaczonym pomiędzy parą bloków – orientującym i orientowanym – na podstawie znanych parametrów orientacji bloku referencyjnego. Przenoszenie orientacji realizowano zarówno bezpośrednio między blokiem referencyjnym a kolejnymi blokami, jak i w ciągach sekwencyjnych obejmujących kolejne epoki zdjęć: od bloku 2022, poprzez 2014, 1994 i 1986, aż do bloku 1976. W przypadku braku georeferencji dla współczesnych zdjęć lotniczych ich elementy orientacji zewnętrznej mogą zostać wyznaczone na podstawie terenowego pomiaru punktów osnowy fotogrametrycznej, opartej na istniejących punktach naturalnych.

Należy podkreślić, że uzyskana dokładność sekwencyjnej orientacji archiwalnych zdjęć lotniczych na poziomie 1 piksela (odpowiadającego terenowej rozdzielczości obrazu) jest bardzo wysoka, mimo szeregu trudności wynikających z ich specyfiki: zróżnicowanej rozdzielczości przestrzennej, odmiennych cech radiometrycznych poszczególnych bloków, zmiennej jakości zdjęć, różnej liczby i stopnia pokrycia zdjęć, niejednorodnego pokrycia terenu oraz zmian związanych z procesami urbanizacyjnymi. Wskazuje to na skuteczność zastosowanej zmodyfikowanej metody *Structure-from-Motion*, która łączy klasyczne podejście *Hand-Crafted* z metodami uczenia głębokiego *Learned-Based*. Zaproponowana metodyka okazała się efektywna w kontekście georeferencji archiwalnych zdjęć lotniczych, gdzie dla wielu przypadków klasyczne podejścia byłyby niewystarczające. Osiągnięte dokładności w wielu przypadkach były porównywalne, a nawet równorzędne z rezultatami uzyskiwanymi dla współczesnych, wysokiej jakości cyfrowych zdjęć lotniczych.

Wysoka dokładność uzyskanych wyników w zakresie sekwencyjnej orientacji zdjęć lotniczych stanowi istotny dowód trafności przyjętej koncepcji badawczej oraz skuteczności opracowanej metodyki. Przeprowadzone badania mają wyraźny wymiar aplikacyjny — umożliwiają praktyczne wykorzystanie zaproponowanych rozwiązań, poprzez dobór odpowiedniego algorytmu do konkretnych warunków projektowych oraz wymaganej dokładności opracowania.

Uzyskane wyniki jednoznacznie wskazują, że opracowana metodyka pozwala na skuteczne nadanie georeferencji archiwalnych zdjęć lotniczych pozyskanym w dużym odstępie czasowym (sięgającym nawet 40 lat), mimo występowania licznych ograniczeń geometrycznych

i radiometrycznych. Zastosowanie sekwencyjnego podejścia do orientacji archiwalnych zdjęć stanowi często jedyne możliwe rozwiązanie w przypadku braku ich bezpośredniej orientacji pomiędzy najbardziej odległymi czasowo blokami fotogrametrycznymi. Tak przygotowane dane stanowią podstawę do szerokiego spektrum zastosowań, takich jak: aktualizacja archiwalnych baz danych przestrzennych dla Głównego Urzędu Geodezji i Kartografii, odtworzenie położenia obiektów historycznych już nie istniejących, analiza zmian użytkowania i pokrycia terenu, ocena dynamiki procesów urbanizacyjnych, a także wspieranie badań w zakresie historii środowiska oraz ochrony dziedzictwa kulturowego. Osiągnięte w niniejszej rozprawie doktorskiej wyniki mogą być również bezpośrednio wykorzystane w pomiarowym opracowaniu archiwalnych zdjęć lotniczych z okresu II wojny światowej, stosowanych m.in. do lokalizacji miejsc masowego ludobójstwa.

W ten sposób zrealizowano założenia badawcze oraz osiągnięto cel główny, prowadząc tym samym do potwierdzenia tezy sformułowanej na etapie wstępnym pracy.

W oparciu o przeprowadzoną analizę oraz uzyskane wyniki, rekomenduje się kontynuację badań nad algorytmami widzenia maszynowego wykorzystującymi głębokie sieci neuronowe. W szczególności zasadne jest ukierunkowanie dalszych badań na trenowanie modeli głębokiego uczenia z wykorzystaniem archiwalnych zdjęć lotniczych, w celu opracowania metod umożliwiających automatyczną ekstrakcję punktów charakterystycznych, odporną na ograniczenia danych wejściowych, takie jak zróżnicowana jakość radiometryczna zdjęć, różna rozdzielczość radiometryczna i przestrzenna.

12. Bibliografia

Aanæs, H., Dahl, A. L., i Steenstrup Pedersen, K. (2012). Interesting interest points. *International Journal of Computer Vision*, 97(1), 18–35. <https://doi.org/10.1007/s11263-011-0473-8>.

Abratkiewicz, K., Gromek, D., Samczyński, P., Markiewicz, J. i Ostrowski, W. (2019). The concept of applying a SIFT algorithm and orthophotomaps in SAR-based augmented integrity navigation systems. *Proceedings of the International Radar Symposium*, 1–12. <https://doi.org/10.23919/IRS.2019.8768136>.

Babbar, G., Bajaj, P., Chawla, A. i Gogna, M. (2010). Comparative study of image matching algorithms. *International Journal of Information Technology and Knowledge Management*, 2(2), 337–339.

Bakuła, K., Mills, J. P. i Remondino, F. (2019). A review of benchmarking in photogrammetry and remote sensing. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-1/W2, 1–8. <https://doi.org/10.5194/isprs-archives-XLII-1-W2-1-2019>.

Balntas, V., Riba, E., Ponsa, D. i Mikolajczyk, K. (2016). Learning local feature descriptors with triplets and shallow convolutional neural networks. *Proceedings of the British Machine Vision Conference*, 2016. <https://doi.org/10.5244/C.30.119>.

Barroso-Laguna, A., Riba, E., Ponsa, D. i Mikolajczyk, K. (2019). Key.Net: Keypoint detection by handcrafted and learned CNN filters. *Proceedings of the IEEE International Conference on Computer Vision*. <https://doi.org/10.48550/arXiv.1904.00889>.

Bay, H., Ess, A., Tuytelaars, T. i Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110, 346–359. <https://doi.org/10.1016/j.cviu.2007.09.014>.

Bay, H., Tuytelaars, T. i Van Gool, L. (2006). SURF: Speeded-up robust features. *Proceedings of the European Conference on Computer Vision*, 404–417.

Bianco, S., Ciocca, G., i Marelli, D. (2018). Evaluating the performance of structure from motion pipelines. *Journal of Imaging*, 4, 1–18. <https://doi.org/10.3390/jimaging4080098>.

Bolognesi, M., Furini, A., Russo, V., Pellegrinelli, A. i Russo, P. (2014). Accuracy of cultural heritage 3D models by RPAS and terrestrial photogrammetry. *ISPRS Annals of the*

Photogrammetry, Remote Sensing and Spatial Information Sciences, II-5, 113–119. <https://doi.org/10.5194/isprsannals-II-5-113-2014>.

Brown, M. i Lowe, D. G. (2002). Invariant features from interest point groups. Proceedings of the 13th British Machine Vision Conference, 253–262.

Bursuc, A., Tolas, G. i Jégou, H. (2015). Kernel local descriptors with implicit rotation matching. Proceedings of the International Conference on Multimedia Retrieval. <https://doi.org/10.1145/2671188.2749379>.

Byrne, D. (2016). ContextCapture – oprogramowanie do automatycznego generowania szczegółowych modeli 3D na podstawie fotografii. Bentley Systems. <https://www.bentley.com>.

Chen, L. i Heipke, C. (2022). Deep learning feature representation for image matching under large viewpoint and viewing direction change. ISPRS Journal of Photogrammetry and Remote Sensing, 190, 94–112. <https://doi.org/10.1016/j.isprsjprs.2022.06.003>.

Chen, L., Rottensteiner, F. i Heipke, C. (2020). Feature detection and description for image matching: From hand-crafted design to deep learning. *Geo-spatial Information Science*. <https://doi.org/10.1080/10095020.2020.1843376>.

Chen, L., Rottensteiner, F. i Heipke, C. (2020). Deep learning based feature matching and its application in image orientation. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-2-2020*, 25–32. <https://doi.org/10.5194/isprs-annals-V-2-2020-25-2020>.

Chen, Y. i Wang, G. (2019). Bundle adjustment revisited. <https://doi.org/10.48550/arXiv.1912.03858>.

Choy, C. B., Gwak, J. Y., Savarese, S. i Chandraker, M. (2016). Universal correspondence network. Proceedings of the Conference on Neural Information Processing Systems. <https://doi.org/10.48550/arXiv.1606.03558>.

Chum, O. i Matas, J. (2005). Matching with PROSAC – Progressive sample consensus. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 220–226. <https://doi.org/10.1109/CVPR.2005.221>.

Clermont, D., Kruse, C., Rottensteiner, F. i Heipke, C. (2019). Supervised detection of bomb craters in historical aerial images using convolutional neural networks. International Archives

of Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W16, 67–74. <https://doi.org/10.5194/isprs-archives-XLII-2-W16-67-2019>.

Cléry, I., Pierrot-Deseilligny, M. i Vallet, B. (2014). Automatic georeferencing of a heritage of old analog aerial photographs. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, II-3. <https://doi.org/10.5194/isprsannals-II-3-33-2014>.

Cover, T. i Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>.

Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. Proceedings of the Conference on Neural Information Processing Systems. <https://doi.org/10.48550/arXiv.1306.0895>.

DeTone, D., Malisiewicz, T. i Rabinovich, A. (2018). SuperPoint: Self-supervised interest point detection and description. CVPR Workshop on Deep Learning for Visual SLAM. <https://doi.org/10.48550/arXiv.1712.07629>.

Dolejš, M., Pacina, J., Veselý, M. i Brétt, D. (2020). Aerial bombing crater identification: Exploitation of precise digital terrain models. ISPRS International Journal of Geo-Information, 9(12), 713. <https://doi.org/10.3390/ijgi9120713>.

Dusmanu, M., Rocco, I., Pajdla, T., Pollefeys, M., Sivic, J., Torii, A. i Sattler, T. (2019). D2-Net: A trainable CNN for joint detection and description of local features. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 8092–8101. <https://doi.org/10.1109/CVPR.2019.00828>.

Ebel, P., Mishchuk, A., Yi, K. M., Fua, P. i Trulls, E. (2019). Beyond cartesian representations for local descriptors. Proceedings of the IEEE International Conference on Computer Vision. <https://doi.org/10.48550/arXiv.1908.05547>.

Ewiak, I. i Brodowska, P. (2012). Wykorzystanie zasobu geoinformacyjnego do opracowania archiwalnych fotogrametrycznych danych obrazowych. Archiwum Fotogrametrii, Kartografii i Teledetekcji, 23, 101–110.

Ewiak, I. i Lulkowska, P. (2016). Rola pomiarów stereoskopowych w procesie identyfikacji elementów pokrycia i użytkowania terenu na archiwalnych zdjęciach lotniczych. XX Ogólnopolskie Sympozjum Naukowe „Fotogrametria dla geoinformacji”.

Farella, E. M., Morelli, L., Remondino, F., Mills, J. P., Haala, N. i Crompvoets, J. (2022). The EuroSDR time benchmark for historical aerial images. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2022, 1175–1182. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-1175-2022>.

Fieber, K. D., Mills, J. P., Miller, P. E., Clarke, L., Ireland, L. i Fox, A. J. (2018). Rigorous 3D change determination in Antarctic Peninsula glaciers from stereo WorldView-2 and archival aerial imagery. *Remote Sensing of Environment*, 205, 18–31. <https://doi.org/10.1016/j.rse.2017.10.042>.

Fischler, M. A. i Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395. <https://doi.org/10.1145/358669.358692>.

Fix, E. i Hodges, J. L. (1951). Discriminatory analysis. Nonparametric discrimination: Consistency properties. USAF School of Aviation Medicine.

Förstner, W. (1986). Feature based correspondence algorithm for image matching. *International Archives of Photogrammetry*, 26(3/3).

Fraser, C. S. (1997). Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(4), 149–159. [https://doi.org/10.1016/S0924-2716\(97\)00005-1](https://doi.org/10.1016/S0924-2716(97)00005-1)

Friedt, J. (2014). Photogrammetric 3D structure reconstruction using MicMac: Basics of SfM. 1–28.

Gabara, G. i Sawicki, P. (2023). CRBeDaSet: A benchmark dataset for high accuracy close range 3D object reconstruction. *Remote Sensing*, 15, 1116. <https://doi.org/10.3390/rs15041116>.

Giordano, S., Le Bris, A. i Mallet, C. (2018). Toward automatic georeferencing of archival aerial photogrammetric surveys. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2. <https://doi.org/10.5194/isprs-annals-IV-2-105-2018>.

Gołuch, P. (2003). Numeryczny model terenu, numeryczny model pokrycia terenu i ortofotomapa jako źródło danych dla przeprowadzenia modelowania hydrodynamicznego. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, 13B, 361–370.

Grottoli, E., Biaisque, M., Rogers, D., Jackson, D. W. i Cooper, J. A. G. (2020). Structure-from-motion-derived digital surface models from historical aerial photographs: A new 3D application for coastal dune monitoring. *Remote Sensing*, 13(1), 95. <https://doi.org/10.3390/rs13010095>.

Habib, A. i Morgan, M. (2003). Linear features in photogrammetry. *Geodetic Science Bulletin*, 9(1), 3–24.

Harris, C. i Stephens, M. (1988). A combined corner and edge detector. *Proceedings of the Alvey Vision Conference*, 23.1–23.6. <https://doi.org/10.5244/C.2.23>

Harwin, S., Lucieer, A. i Osborn, J. (2015). The impact of the calibration method on the accuracy of point clouds derived using unmanned aerial vehicle multi-view stereopsis. *Remote Sensing*, 7, 11933–11953. <https://doi.org/10.3390/rs70911933>.

Hatzopoulos, J. N., Stefanakis, D., Georgopoulos, A., Tapinaki, S., Pantelis, V. i Liritzis, I. (2017). Use of various surveying technologies for 3D digital mapping and modelling of cultural heritage structures. *Mediterranean Archaeology and Archaeometry*, 17, 311–336. <https://doi.org/10.5281/zenodo.1048937>.

Jao, F. J., Chu, H. J. i Tseng, Y. H. (2014). Historical image registration and land-use land-cover change analysis. *Environments*, 1(2), 181–189. <https://doi.org/10.3390/environments1020181>.

Jégou, H. i Chum, O. (2012). Negative evidences and co-occurrences in image retrieval: The benefit of PCA and whitening. *Proceedings of the European Conference on Computer Vision*.

Jo, J., Seo, J. i Fekete, J. D. (2017). A progressive kd-tree for approximate k-nearest neighbors. *IEEE Workshop on Data Systems for Interactive Analysis*. <https://doi.org/10.1109/DSIA.2017.8339084>.

Kahmen, O., Rofallski, R. i Luhmann, T. (2020). Impact of stereo camera calibration to object accuracy in multimedia photogrammetry. *Remote Sensing*, 12(12), 2057. <https://doi.org/10.3390/rs12122057>.

Kapnias, D., Milenov, P. i Kay, S. (2008). Guidelines for best practice and quality checking of ortho imagery. *European Commission*.

Karel, W., Doneus, M., Verhoeven, G., Briese, C., Ressel, C. i Pfeifer, N. (2013). ORIENTAL – Automatic geo-referencing and orthorectification of archaeological aerial photographs. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5/W1, 175–180.

Karwel, A. K. i Markiewicz, J. (2022). The methodology of the archival aerial image orientation based on the SfM method. *Sensors and Machine Learning Applications*, 1(2), 1–19. <https://doi.org/10.55627/smla.001.02.0015>.

Kim, J. S., Miller, C. C. i Bethel, J. (2010). Automated georeferencing of historic aerial photography. *Journal of Terrestrial Observation*, 2(1), 6.

Knuth, F., Shean, D., Bhushan, S., Schwat, E., Alexandrov, O., McNeil, C., Dehecq, A., Florentine, C. i O'Neel, S. (2023). Historical structure from motion (HSfM): Automated processing of historical aerial photographs for long-term topographic change analysis. *Remote Sensing of Environment*, 285, 113379. <https://doi.org/10.1016/j.rse.2022.113379>.

Kumar, V., Carneiro, G. i Reid, I. (2016). Learning local image descriptors with deep siamese and triplet convolutional networks by minimising global loss functions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5385–5394. <https://doi.org/10.1109/CVPR.2016.581>.

Kupidura, P., Osińska-Skotak, K., Lesisz, K. i Podkow, A. (2019). The efficacy analysis of determining the wooded and shrubbed area based on archival aerial imagery using texture analysis. *ISPRS International Journal of Geo-Information*, 8(10), 450.

Kurczyński, Z., Bakuła, K., Pilarska, M. i Ostrowski, W. (2020). The problem of using and measurement of identifiable ground control points on high resolution aerial images. *Archives of Photogrammetry, Cartography and Remote Sensing*, 31, 135–144. <https://doi.org/10.2478/apcrs-2019-0009>.

Leutenegger, S., Chli, M. i Siegwart, R. Y. (2011). BRISK: Binary robust invariant scalable keypoints. *Proceedings of the IEEE International Conference on Computer Vision*, 2548–2555. <https://doi.org/10.1109/ICCV.2011.6126542>.

Li, X., Han, K., Li, S. i Prisacariu, V. (2020). Dual resolution. <https://doi.org/10.48550/arXiv.2006.08844>.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. *Proceedings of the International Conference on Computer Vision*. <https://doi.org/10.1109/ICCV.1999.790410>.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60, 91–110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>.

Lydersen, J. M. i Collins, B. M. (2018). Change in vegetation patterns over a large forested landscape based on historical and contemporary aerial photography. *Ecosystems*, 21, 1348–1363. <https://doi.org/10.1007/s10021-018-0225-5>.

Marelli, D., Morelli, L., Farella, E. M., Bianco, S., Ciocca, G., i Remondino, F. (2023). ENRICH: Multi-purpose dataset for benchmarking in computer vision and photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 198, 84–98.

Markiewicz, J., Abratkiewicz, K., Gromek, A., Ostrowski, W., Samczyński, P., i Gromek, D. (2019). Geometrical matching of SAR and optical images utilizing ASIFT features for SAR-based navigation aided systems. *Sensors*, 19(24), 1–33. <https://doi.org/10.3390/s19245500>.

Markiewicz, J., i Turek, A. (2014). Wykorzystanie wieloczasowych zdjęć lotniczych w badaniu zmian zagospodarowania przestrzennego. *Przegląd Geodezyjny*, 7, 3–9.

Markiewicz, J. S. (2016). The use of computer vision algorithms for automatic orientation of terrestrial laser scanning data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B3, 315–322. <https://doi.org/10.5194/isprsarchives-XLI-B3-315-2016>.

Mboga, N., Grippa, T., Georganos, S., Vanhuysse, S., Smet, B., Dewitte, O., Wolff, E., i Lennert, M. (2020). Fully convolutional networks for land cover classification from historical panchromatic aerial photographs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 385–395. <https://doi.org/10.1016/j.isprsjprs.2020.07.005>.

Meixner, P., i Eckstein, M. (2016). Multi-temporal analysis of WWII reconnaissance photos. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B8, 973–2016. <https://doi.org/10.5194/isprs-archives-XLI-B8-973-2016>.

Micheletti, N., Lane, S. N., i Chandler, J. H. (2015). Application of archival aerial photogrammetry to quantify climate forcing of alpine landscapes. *Photogrammetric Record*, 30, 143–165. <https://doi.org/10.1111/phor.12099>.

Mikolajczyk, K., i Schmid, C. (2005). A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10), 1615–1630. <https://doi.org/10.1109/TPAMI.2005.188>.

Miles, R., i Mikolajczyk, K. (2020). Compression of descriptor models for mobile applications. *arXiv preprint arXiv:2001.03102*.

Mishchuk, A., Mishkin, A., Radenovic, F., i Matas, J. (2017). Working hard to know your neighbor's margins: Local descriptor learning loss. *Proceedings of the Conference on Neural Information Processing Systems*. <https://doi.org/10.48550/arXiv.1705.10872>.

Morel, J. M., i Yu, G. (2009). ASIFT: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2(2), 438–469. <https://doi.org/10.1137/080732730>.

Morgan, J. A., i Brogan, D. J. (2016). *How to VisualSFM*. Colorado State University, 1– 21.

Muja, M., i Lowe, D. G. (2009). Fast approximate nearest neighbors with automatic algorithm configuration. *Proceedings of the International Conference on Computer Vision Theory and Applications*, 331–340. <https://doi.org/10.5220/0001787803310340>.

Mukundan, A., Tolas, G., Bursuc, A., Jégou, H., i Chum, O. (2018). Understanding and improving kernel local descriptors. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.48550/arXiv.1811.11147>.

Nebiker, S., Lack, N., i Deuber, M. (2014). Building change detection from historical aerial photographs using dense image matching and object-based image analysis. *Remote Sensing*, 6, 8310–8336. <https://doi.org/10.3390/rs6098310>.

Nocerino, E., Menna, F., i Remondino, F. (2012). Multi-temporal analysis of landscapes and urban areas. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B4, 85–90. <https://doi.org/10.5194/isprsarchives-XXXIX-B4-85-2012>.

Noh, H., Araujo, A., Sim, J., Weyand, T., i Han, B. (2017). Large-scale image retrieval with attentive deep local features. *Proceedings of the IEEE International Conference on Computer Vision*, 3456–3465. <https://doi.org/10.1109/ICCV.2017.373>.

Nurminen, K., Litkey, P., Honkavaara, E., Vastaranta, M., Holopainen, M., Lyytikäinen-Saarenmaa, P., Kantola, T., i Lyytikäinen, M. (2015). Automation aspects for the georeferencing of photogrammetric aerial image archives in forested scenes. *Remote Sensing*, 7, 1565–1593. <https://doi.org/10.3390/rs70201565>.

Osińska-Skotak, K., Jełowicki, Ł., Bakuła, K., Michalska-Hejduk, D., Wylazłowska, J., i Kopeć, D. (2019). Analysis of using dense image matching techniques to study the process of secondary

succession in non-forest Natura 2000 habitats. *Remote Sensing*, 11(8), 893. <https://doi.org/10.3390/rs11080893>.

Pawlik, P., i Mikrut, S. (2007). Porównanie dokładności wybranych metod dopasowania obrazów zdjęć lotniczych. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, 17B, 603–612.

Peppas, M. V., Mills, J. P., Fieber, K. D., Haynes, I., Turner, S., Turner, A., Douglas, M., i Bryan, P. G. (2018). Archaeological feature detection from archive aerial photography with an SfM–MVS and image enhancement pipeline. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2, 869–2018. <https://doi.org/10.5194/isprs-archives-XLII-2-869-2018>.

Pinto, A. T., Gonçalves, J. A., Beja, P., i Pradinho, H. J. (2019). From archived historical aerial imagery to informative orthophotos: A framework for retrieving the past in long-term socioecological research. *Remote Sensing*, 11(11), 1388. <https://doi.org/10.3390/rs11111388>.

Poli, D., Casarotto, C., Strudl, M., Bollmann, E., Moe, K., i Legat, K. (2020). Use of historical aerial images for 3D modelling of glaciers in the province of Trento. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2020, 1151–1158. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-1151-2020>.

Popelková, R., i Mulková, M. (2016). Multitemporal aerial image analysis for the monitoring of the processes in the landscape affected by deep coal mining. *European Journal of Remote Sensing*, 49(1), 973–1009. <https://doi.org/10.5721/EuJRS20164951>

Preuss, R. (2002). Aerotriangulacja wieloczasowa. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, 12, 325–334.

Pultar, M. (2020). Improving the HardNet descriptor. *arXiv preprint arXiv:2007.09699*. <https://doi.org/10.48550/arXiv.2007.09699>.

Ratajczak, R., Crispim-Junior, C. F., Faure, E., Fervers, B., i Tougne, L. (2019). Automatic land cover reconstruction from historical aerial images: An evaluation of features extraction and classification algorithms. *IEEE Transactions on Image Processing*, 28, 3357–3371. <https://doi.org/10.1109/TIP.2019.2896492>.

Redecker, A. P. (2008). Historical aerial photographs and digital photogrammetry for impact analyses on derelict land sites in human settlement areas. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII-B8, 5–10.

Redmond, S. J., i Heneghan, C. (2007). A method for initialising the K-means clustering algorithm using kd-trees. *Pattern Recognition Letters*, 28(8), 965–973. <https://doi.org/10.1016/j.patrec.2007.01.001>.

Redweik, P., Roque, D., Marques, A., Matildes, R., i Marques, F. (2010). Triangulating the past – Recovering Portugal’s aerial images repository. *Photogrammetric Engineering and Remote Sensing*, 76, 1007–1018. <https://doi.org/10.14358/PERS.76.9.1007>.

Remondino, F., i Fraser, C. S. (2006). Digital camera calibration methods: Considerations and comparisons. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVI-5, 266–272. <https://doi.org/10.3929/ethz-b-000158067>.

Remondino, F., Menna, F., i Morelli, L. (2021). Evaluating hand-crafted and learning-based features for photogrammetric applications. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2021, 549–556. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-549-2021>.

Ressl, C., Karel, W., Piermattei, L., Puercher, G., Hollaus, M., i Pfeifer, N. (2020). Multi-epoch bundle block adjustment for processing large dataset of historical aerial images. *EGU General Assembly Conference Abstracts*, 22544.

Risbol, O., Briese, C., Doneus, M., i Nesbakken, A. (2014). Monitoring cultural heritage by comparing DEMs derived from historical aerial photographs and airborne laser scanning. *Journal of Cultural Heritage*, 16(2), 202–209. <https://doi.org/10.1016/j.culher.2014.04.002>.

Rocco, I., Cimpoi, M., Arandjelovic, R., Torii, A., Pajdla, T., i Sivic, J. (2018). Neighbourhood consensus networks. *Proceedings of the Conference on Neural Information Processing Systems*. <https://doi.org/10.48550/arXiv.1810.10510>.

Rong, K., Jieqi, S., Xueming, L., Yang, L., i Xiao, L. (2019). DF-SLAM: A deep-learning enhanced visual SLAM system based on deep local features. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.48550/arXiv.1901.07223>.

Rosten, E., i Drummond, T. (2006). Machine learning for high-speed corner detection. *Proceedings of the European Conference on Computer Vision*, 430–443. https://doi.org/10.1007/11744023_34.

Różycki, S., Karwel, A. K., i Kurczyński, Z. (2023). German extermination camps on WWII reconnaissance photographs: Orthorectification process for archival aerial images of cultural heritage sites. *Remote Sensing*, 15, 2587. <https://doi.org/10.3390/rs15102587>.

Salach, A. (2017). SAPC – Application for adapting scanned analogue photographs to use them in structure-from-motion technology. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-1/W1, 197. <https://doi.org/10.5194/isprs-archives-XLII-1-W1-197-2017>.

Sander, J., Ester, M., Kriegel, H. P., i Xu, X. (1998). Density-based clustering in spatial databases: The algorithm GDBSCAN and its applications. *Data Mining and Knowledge Discovery*, 2, 169–194. <https://doi.org/10.1023/A:1009745219419>.

Sanecki, J., Konieczny, J., Klewski, A., Stępień, G., Wołęjsza, P., Konieczny, K., i Beczkowski, K. (2014). Kontrola upraw rolnych z wykorzystaniem bezzałogowych systemów latających (UAS). *Polska Akademia Nauk*, 891–902.

Sarlin, P.-E., DeTone, D., Malisiewicz, T., i Rabinovich, A. (2019). SuperGlue: Learning feature matching with graph neural networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR42600.2020.00499>.

Sauerbier, M., Kunz, M., Fluehler, M., i Remondino, F. (2004). Photogrammetric reconstruction of adobe architecture at Túcume, Peru. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVI-5/W1. <https://doi.org/10.3929/ethz-a-005018568>.

Saurer, O., Fraundorfer, F., i Pollefeys, M. (2010). OmniTour: Semi-automatic generation of interactive virtual tours from omnidirectional video. *Proceedings of 3DPVT*, 1–8.

Schönberger, J. L., i Frahm, J. M. (2016). Structure-from-motion revisited. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2016.445>.

Seitz, S. M., Curless, B., Diebel, J., Scharstein, D., i Szeliski, R. (2006). A comparison and evaluation of multi-view stereo reconstruction algorithms. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 519–528. <https://doi.org/10.1109/CVPR.2006.19>.

Sevara, C. (2013). Top secret topographies: Recovering two- and three-dimensional archaeological information from historic reconnaissance datasets using image-based

modelling techniques. *International Journal of Heritage in the Digital Era*, 2, 395–418. <https://doi.org/10.1260/2047-4970.2.3.395>.

Sevara, C., Verhoeven, G., Doneus, M., i Draganits, E. (2018). Surfaces from the visual past: Recovering high-resolution terrain data from historic aerial imagery for multitemporal landscape analysis. *Journal of Archaeological Method and Theory*, 25, 611–642. <https://doi.org/10.1007/s10816-017-9348-9>.

Shen, Y., Xiao, T., Li, H., Yi, S., i Wang, X. (2018). End-to-end deep Kronecker-product matching for person re-identification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.48550/arXiv.1807.11182>.

Shen, J., Guo, X., Zhou, W., i Li, J. (2021). Feature point matching method for aerial image based on recursive diffusion algorithm. *Symmetry*, 13(3), 407. <https://doi.org/10.3390/sym13030407>.

Simonyan, K., i Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. <https://doi.org/10.48550/arXiv.1409.1556>.

Siok, K., i Ewiak, I. (2020). The simulation approach to the interpretation of archival aerial photographs. *Open Geosciences*, 12(1), 1–10. <https://doi.org/10.1515/geo-2020-0001>.

Sonnemann, T. M., Remondino, F., i Schoretter, G. (2006). Reality-based 3D modeling of the Angkorian temples using aerial images. *Proceedings of the 2nd International Conference on Remote Sensing in Archaeology*, 573–579.

Strecha, C., von Hansen, W., Van Gool, L., Fua, P., i Thoennessen, U. (2008). On benchmarking camera calibration and multi-view stereo for high resolution imagery. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–8. <https://doi.org/10.1109/CVPR.2008.4587706>.

Sun, J., Shen, Z., Wang, Y., Bao, H., i Zhou, X. (2021). LoFTR: Detector-free local feature matching with transformers. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 8918–8927. <https://doi.org/10.48550/arXiv.2104.00680>.

Tavallali, P., Tavallali, P., i Singhal, M. K. (2020). Means tree: An optimal clustering tree for unsupervised learning. *Journal of Supercomputing*. <https://doi.org/10.1007/s11227-020-03436-2>.

Tian, Y., Barroso-Laguna, A., Ng, T., Balntas, V., i Mikolajczyk, K. (2020). HyNet: Learning local descriptor with hybrid similarity measure and triplet loss. *Advances in Neural Information Processing Systems*, 33, 7401–7412. <https://doi.org/10.48550/arXiv.2006.10202>.

Tian, Y., Fan, B., i Wu, F. (2017). L2-Net: Deep learning of discriminative patch descriptor in Euclidean space. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2017.649>.

Tian, Y., Yu, X., Fan, B., Wu, F., Heijnen, H., i Balntas, V. (2019). SOSNet: Second order similarity regularization for local descriptor learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.48550/arXiv.1904.05019>.

Tobiasz, A., Markiewicz, J., Łapiński, S., Nikel, J., Kot, P., i Muradov, M. (2019). Review of methods for documentation, management, and sustainability of cultural heritage. *Sustainability*, 11, 7046. <https://doi.org/10.3390/su11247046>.

Tola, E., Lepetit, V., i Fua, P. (2010). DAISY: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32, 815–830. <https://doi.org/10.1109/TPAMI.2009.77>.

Tolias, G., Bursuc, A., Furon, T., i Jégou, H. (2015). Rotation and translation covariant match kernels for image retrieval. *Computer Vision and Image Understanding*, 140, 9–20. <https://doi.org/10.1016/j.cviu.2015.06.007>.

Tuytelaars, T., i Mikolajczyk, K. (2007). Local invariant feature detectors: A survey. *Foundations and Trends in Computer Graphics and Vision*, 3, 177–280. <https://doi.org/10.1561/06000000017>.

Tyszkiewicz, M., Fua, P., i Trulls, E. (2020). DISK: Learning local features with policy gradient. *Advances in Neural Information Processing Systems*. <https://doi.org/10.48550/arXiv.2006.13566>.

Valjavec, M. B., Zorn, M., i Ribeiro, D. (2018). Mapping war geoheritage: Recognising geomorphological traces of war. *Open Geosciences*, 10(1), 385–394. <https://doi.org/10.1515/GEO-2018-0030>.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., i Polosukhin, I. (2017). Attention is all you need. *Proceedings of the Conference on Neural Information Processing Systems*. <https://doi.org/10.48550/arXiv.1706.03762>.

Verdie, Y., Yi, K. M., Fua, P., i Lepetit, V. (2015). TILDE: A temporally invariant learned detector. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5279–5288. <https://doi.org/10.1109/CVPR.2015.7299165>.

Verhoeven, G., Doneus, M., Briese, C., i Vermeulen, F. (2012). Mapping by matching – A computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science*, 39, 2060–2070. <https://doi.org/10.1016/j.jas.2012.02.022>.

Verhoeven, G., Sevara, C., Karel, W., Ressel, C., Doneus, M., i Briese, C. (2013). Undistorting the past: New techniques for orthorectification of archaeological aerial frame imagery. In *Good Practice in Archaeological Diagnostics: Natural Science in Archaeology* (pp. 31–67). https://doi.org/10.1007/978-3-319-01784-6_3.

Walstra, J., Chandler, J., Dixon, N., i Dijkstra, T. (2004). Time for change – Quantifying landslide evolution using historical aerial photographs and modern photogrammetric methods. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(B4), 475–480.

Wiedemann, A., Hemmleb, M., i Albertz, J. (2000). Reconstruction of historical buildings based on images from the Meydenbauer archives. *International Archives of Photogrammetry and Remote Sensing*, 33(B5/2), 887–893.

Yan, N., Mei, Y., Xu, L., Yu, H., Sun, B., Wang, Z., i Chen, Y. (2023). Deep learning on image stitching with multi-viewpoint. *Neural Processing Letters*, 55, 3863–3898. <https://doi.org/10.1007/s11063-023-11226-z>.

Yi, K. M., Trulls, E., Lepetit, V., i Fua, P. (2016). LIFT: Learned invariant feature transform. *Proceedings of the European Conference on Computer Vision*. https://doi.org/10.1007/978-3-319-46466-4_28.

Yu, G., i Morel, J. M. (2011). ASIFT: An algorithm for fully affine invariant comparison. *Image Processing On Line*, 1, 11–38. <https://doi.org/10.5201/ipol.2011.my-asift>.

Zambanini, S., i Sablatnig, R. (2017). A Hough voting strategy for registering historical aerial images to present-day satellite imagery. *Proceedings of the International Conference on Image Analysis and Processing*, 595–605. https://doi.org/10.1007/978-3-319-68560-1_53.

Zawieska, D., Markiewicz, J., i Kopiasz, J. (2019). Development of true orthophotomaps of the fortified settlement at Biskupin, site 4, based on archival data. *Archaeological Prospection*, 26, 333–360. <https://doi.org/10.1002/arp.1748>.

Zhang, L., Rupnik, E., i Pierrot-Deseilligny, M. (2021). Feature matching for multi-epoch historical aerial images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 182, 176–189. <https://doi.org/10.1016/j.isprsjprs.2021.10.008>.

Strony internetowe:

Galarnyk, M. (2018). Boxplots are graphs that tell you how your data's values are spread out. Here's how to read a boxplot and even create your own. *BuiltIn*. <https://builtin.com/data-science/boxplot>.

Skenzo Ltd. (n.d.). AiGeekProgrammer. <https://aigeekprogrammer.com/pl>.

University of Washington. (n.d.). Photo tourism. <https://phototour.cs.washington.edu>.

Vedaldi, A., i Fulkerson, B. (n.d.). The VLFeat open-source library. <https://www.vlfeat.org>.

OpenCV.org. (n.d.). OpenCV library. <https://opencv.org>.

Weisstein, E. W. (n.d.). MathWorld — A Wolfram Web Resource. <https://mathworld.wolfram.com>

EuroSDR. (2025). Benchmark Time. <https://www.eurocdr.net/node/1657>

Fondazione Bruno Kessler. (2025). TIME Benchmark. <https://time.fbk.eu/home>.

Agisoft LLC. (2025). Orthomosaic/DEM generation without GCPs. Agisoft Helpdesk Portal. <https://agisoft.freshdesk.com/support/solutions/articles/31000157908-orthomosaic-dem-generation-without-gcps>.

Agisoft LLC. (2011). Agisoft forum discussion: GCP accuracy. <https://www.agisoft.com/forum/index.php?topic=89.0>.

Książki:

American Society of Photogrammetry. (1980). *Manual of photogrammetry* (4th ed.).

Dziennik Ustaw Nr 263. (2011). Rozporządzenie Ministra Spraw Wewnętrznych i Administracji z dnia 9 listopada 2011 r. w sprawie standardów technicznych wykonywania geodezyjnych pomiarów sytuacyjnych i wysokościowych oraz opracowywania i przekazywania wyników tych pomiarów do państwowego zasobu geodezyjnego i kartograficznego. <https://isap.sejm.gov.pl>.

Dziennik Ustaw z 2021 r. poz. 1990; z 2022 r. poz. 1846 i 2185. (2021–2022). Ustawa z dnia 17 maja 1989 r. – Prawo geodezyjne i kartograficzne. <https://isap.sejm.gov.pl>.

Ghosh, S. K. (2005). Fundamentals of computational photogrammetry. Concept Publishing Company Pvt. Ltd, ISBN-10: 818069187X.

Gilbert, S. (2016). Introduction to linear algebra. Wellesley–Cambridge Press, ISBN-10: 0980232775.

Mardia, K. V., i Jupp, P. E. (2000). Directional statistics. Wiley.

Mikhail, E. M., Bethel, J. S., i McGlone, J. C. (2001). Introduction to modern photogrammetry. Wiley.

Moussa, W. (2014). Integration of digital photogrammetry and terrestrial laser scanning for cultural heritage data recording. University of Stuttgart. ISBN 9783769651379, https://www.ifp.uni-stuttgart.de/dokumente/Dissertationen/c-725_w.moussa.pdf.

Luhmann, T., Robson, S., Kyle, S., i Boehm, J. (2020). Close-range photogrammetry and 3D imaging. De Gruyter. <https://doi.org/10.1515/9783110607253>.

13. Załączniki

Załącznik nr.1. - Przegląd metryk kamer lotniczych

Tab. Przegląd metryk kamer analogowych (rok wydania, nazwa firmy, nazwa kamery, format zdjęcia, ogniskowa kamery, wartości nominalne znaczków tłowych, położenie punktu głównego zdjęcia, dystorsja obiektywu kamery)

Rok	Firma	Kamera	Format [cm]	Ck [mm]	Znaczki tłowe [mm]	PPA [mm]	Dystorsja (średnia) ± [mm]
1963	Wild	RC5	18x18	115.29	A-B=163.993, B-C=163.997 C-D=163.992, D-A=163.998	Pp of autocollim 0.02	0.08
1995	Wild	RC5	18x18	115.34	1 X=81.999, Y=-81,997 2 X=-81.995, Y=81,996 3 X=81.998, Y=-81,997 4 X=81.996, Y=-81,998	Pp of autocollim 0.02	0.08
1963	Wild	RC5	18x18	115.26	1-2=163.999, 2-3=163.991 3-4=163.995, 4-1=163.991	0.017	0.08
1966	Wild	RC5	9x9	152.03	1-2=212.001, 2-3=212.000 3-4=212.001, 4-1=212.009	0.01	0.008
1986	Wild	RC10	23x23	153.26	1 X=106.000, Y=-106.002 2 X=-105.998, Y=-106.000 3 X=-106.006, Y=106.008 4 X=105.998, Y=106.000	X=0.002 Y=0.011	0.0063
1992	Wild	RC10	23x23	153.26	1 X=106.000, Y=-105.998 2 X=-106.007, Y=-106.010 3 X=-106.000, Y=105.998 4 X=106.007, Y=106.010	X=-0.005 Y=0.011	0.0065
1987	Wild	RC10	9x9	152.80	1 X=105.999, Y=105.999 2 X=-105.992, Y=105.984 3 X=-106.000, Y=-105.998 4 X=105.998, Y=-105.992	X=0.002 Y=-0.001	0.001
1989	Wild	RC10	9x9	213.75	1 X=105.999, Y=105.999 2 X=-105.992, Y=105.984 3 X=-106.000, Y=-105.998 4 X=105.998, Y=-105.992	X=-0.001 Y=-0.002	0.003
1986	Wild	RC10	9x9	213.75	1 X=105.999, Y=105.999 2 X=-106.008, Y=106.003 3 X=-106.000, Y=-106.004 4 X=105.998, Y=-106.006	X=-0.004 Y=-0.009	0.004
1972	Wild	RC10	9x9	152.22	1-2=212.012, 2-3=212.005 3-4=212.002, 4-1=212.010	X=0.002 Y=-0.002	0.007
1995	Leica	RC20	23x23	152.87	1 X=106.001, Y=-106.001 2 X=-106.001, Y=-106.002 3 X=-106.001, Y=106.001 4 X=106.000, Y=106.001	X=-0.006 Y=0.010	0.0016
1997	Leica	RC20	23x23	152.98	1 X=106.001, Y=-106.004	X=0.004	0.0003

					2 X=-105.999, Y=-106.000 3 X=-106.001, Y=106.004 4 X=106.003, Y=106.004 5 X=0.001, Y=-109.997 6 X=-109.997, Y=-0.001 7 X=-0.001, Y=109.998 8 X=110.002, Y=0.002	Y=0.003	
1994	Leica	RC20	23x23	152.88	1 X=105.999, Y=-106.001 2 X=-106.004, Y=-106.007 3 X=-105.992, Y=105.994 4 X=105.999, Y=106.002 5 X=-0.004, Y=-109.998 6 X=-110.006, Y=-0.008 7 X=0.005, Y=109.996 8 X=109.997, Y=0.000	X=0.003 Y=-0.001	0.0026
1992	Leica	RC20	23x23	152.36	1 X=106.002, Y=-106.005 2 X=-106.001, Y=-106.002 3 X=-106.004, Y=106.007 4 X=106.002, Y=106.002 5 X=0.002, Y=-109.996 6 X=-110.010, Y=0.001 7 X=-0.001, Y=110.013 8 X=110.004, Y=0.004	X=0.003 Y=0.013	0.0009
1995	Leica	RC20	23x23	153.17	1 X=105.995, Y=105.998 2 X=-106.004, Y=106.007 3 X=-105.995, Y=105.998 4 X=106.001, Y=106.003	X=0.003 Y=0.004	0.0018
2000	Leica	RC20	23x23	213.707	1 X=106.004, Y=-106.002 2 X=-106.002, Y=-106.001 3 X=-106.003, Y=106.002 4 X=106.003, Y=106.001 5 X=0.001, Y=-109.999 6 X=-110.001, Y=0.004 7 X=-0.004, Y=110.002 8 X=110.004, Y=0.000	X=-0.010 Y=-0.005	0.0042
2004	Leica	RC20	23x23	213.71	1 X=106.004, Y=-106.003 2 X=-106.002, Y=-106.000 3 X=-106.004, Y=106.000 4 X=106.003, Y=106.003 5 X=0.003, Y=-109.998 6 X=-110.004, Y=0.001 7 X=-0.003, Y=110.003 8 X=110.003, Y=-0.002	X=-0.010 Y=-0.005	0.0041
2005	Leica	RC20	23x23	213.55	1 X=106.001, Y=-106.000 2 X=-106.001, Y=-105.998 3 X=-106.004, Y=106.003 4 X=106.001, Y=105.998 5 X=0.001, Y=-109.998	X=-0.007 Y=0.009	0.0024

					6 X=-110.000, Y=0.003 7 X=-0.001, Y=110.007 8 X=110.004, Y=0.004		
2005	Wild	RC20	23x23	152.665	1 X=-106.012, Y=-105.993 2 X=106.004, Y=105.995 3 X=-105.989, Y=106.010 4 X=105.986, Y=-105.993 5 X=-110.005, Y=0.009 6 X=110.005, Y=-0.003 7 X=0.015, Y=110.003 8 X=-0.018, Y=-110.000	X=0 Y=0	0.002
1998	Wild	RC30	23x23	153.396	1 X=-106.007, Y=-106.000 2 X=105.993, Y=106.000 3 X=-105.996, Y=105.999 4 X=105.998, Y=-106.000 5 X=-112.004, Y=-0.003 6 X=111.996, Y=0.000 7 X=0.004, Y=112.013 8 X=-0.008, Y=-111.997	X=0 Y=0	0.002
1998	Wild	RC30	23x23	153.93	1 X=106.002, Y=-106.004 2 X=-105.995, Y=-105.995 3 X=-106.003, Y=106.005 4 X=109.999, Y=105.999 5 X=0.007, Y=-110.004 6 X=110.002, Y=0.005 7 X=-0.002, Y=110.003 8 X=109.993, Y=-0,002	X=0.000 Y=0.017	0.0019
1995	Leica	RC30	23x23	153.81	1 X=106.004, Y=-106.003 2 X=-106.000, Y=-106.000 3 X=-106.003, Y=106.002 4 X=105.998, Y=105.999 5 X=0.000, Y=-111.997 6 X=111.997, Y=0.006 7 X=-0.003, Y=112.000 8 X=111.997, Y=0,001	X=-0.003 Y=-0.006	0.0023
1996	Leica	RC30	23x23	153.12	1 X=106.002, Y=-106.001 2 X=-106.001, Y=-106.002 3 X=-106.002, Y=106.002 4 X=106.002, Y=106.003 5 X=-0.001, Y=-111.999 6 X=112.000, Y=0.003 7 X=0.000, Y=112.002 8 X=112.001, Y=0,002	X=0.002 Y=0.001	0.0009
1995	Zeiss	RMK A	23x23	305.309	1 X=112.953, Y=-0.021 2 X=-113.032, Y=-0.034 3 X=-0.025, Y=112.950 4 X=-0.038, Y=-113.028	X=-0.045 Y=-0.034	0.0023
1995	Zeiss	RMK A	23x23	153.566	1 X=112.955, Y=-0.007 2 X=-113.000, Y=-0.005	X=-0.006 Y=-0.006	0.002

					3 X=-0.004, Y=113.006 4 X=-0.002, Y=-113.005		
1991	Zeiss	RMK A	23x23	153.554	1 X=113.003, Y=-0.004 2 X=-112.998, Y=-0.007 3 X=-0.000, Y=112.999 4 X=-0.004, Y=-113.002	X=-0.004 Y=-0.005	0.002
1995	Zeiss	RMK A	23x23	304.928	1 X=112.973, Y=-0.014 2 X=-113.030, Y=-0.013 3 X=-0.029, Y=112.987 4 X=-0.028, Y=-113.011	X=-0.027 Y=-0.014	0.003
2001	Zeiss	RMK TOP	23x23	153.802	1 X=113.000, Y=0.004 2 X=-112.995, Y=0.007 3 X=0.000, Y=113.000 4 X=0.003, Y=-112.987 5 X=113.002, Y=113.000 6 X=-112.993, Y=-112.996 7 X=-113.004, Y=113.000 8 X=113.000, Y=-112.998	X=0.001 Y=0.005	0.009
2009	Zeiss	RMK TOP	23x23	153.873	1 X=112.989, Y=0.016 2 X=-113.112, Y=0.021 3 X=-0.013, Y=113.020 4 X=-0.008, Y=-112.981 5 X=112.978, Y=113.025 6 X=113.006, Y=-112.981 7 X=-113.018, Y=113.023 8 X=112.992, Y=-112.977	X=-0.014 Y=0.003	0.002
2002	Zeiss	RMK TOP	23x23	154.007	1 X=113.010, Y=0.008 2 X=-112.989, Y=0.002 3 X=0.010, Y=113.010 4 X=0.004, Y=-112.990 5 X=113.000, Y=113.006 6 X=-112.987, Y=-112.996 7 X=-112.997, Y=113.007 8 X=113.008, Y=-112.992	X=-0.005 Y=0.005	0.002
1994	Zeiss	RMK TOP	23x23	153.881	1 X=112.996, Y=-0.001 2 X=-113.001, Y=0.005 3 X=-0.004, Y=113.010 4 X=0.002, Y=-112.991 5 X=112.996, Y=113.001 6 X=-112.991, Y=-112.996 7 X=-112.999, Y=112.998 8 X=112.996, Y=-112.999	X=-0.011 Y=0.004	0.002
1994	Zeiss	RMK TOP	23x23	153.977	1 X=112.999, Y=-0.001 2 X=-112.997, Y=0.002 3 X=0.003, Y=113.000 4 X=0.006, Y=-112.994 5 X=112.991, Y=113.000 6 X=-112.999, Y=-113.012 7 X=-113.004, Y=112.990	X=-0.003 Y=0.005	0.004

					8 X=112.999, Y=-112.995		
1994	Zeiss	RMK TOP	23x23	154.009	1 X=113.009, Y=-0.003 2 X=-112.976, Y=-0.005 3 X=0.018, Y=113.006 4 X=0.016, Y=-113.002 5 X=113.009, Y=112.990 6 X=-112.974, Y=-113.005 7 X=-112.984, Y=112.988 8 X=113.007, Y=-113.003	X=0.009 Y=0.005	0.002
1995	Zeiss Jena	LMK	23x23	305.21	1 X=-109.993, Y=-109.963 2 X=110.001, Y=110.031 3 X=-109.996, Y=110.032 4 X=109.999, Y=-109.960 5 X=-109.999, Y=0.036 6 X=109.995, Y=0.036 7 X=-0.001, Y=110.022 8 X=0.000, Y=-109.968	X=0.02 Y=0.02	0.005
1995	Zeiss Jena	LMK	23x23	305.34	1 X=-110.003, Y=-110.010 2 X=110.002, Y=109.993 3 X=-109.995, Y=109.991 4 X=109.991, Y=-110.009 5 X=-111.997, Y=-0.009 6 X=112.005, Y=-0.009 7 X=0.005, Y=111.995 8 X=-0.007, Y=-112.020	X=0.01 Y=-0.01	0.003
1996	Zeiss Jena	LMK	23x23	152.188	1 X=112.008, Y=0.018 2 X=-111.993, Y=0.015 3 X=0.010, Y=112.015 4 X=0.026, Y=-112.000 5 X=110.010, Y=110.017 6 X=-109.995, Y=-109.992 7 X=-109.991, Y=110.011 8 X=110.003, Y=-109.982	X=0.07 Y=0.08	0.010
1995	Zeiss Jena	LMK	23x23	305.206	1 X=-110.021, Y=-110.000 2 X=-109.998, Y=-0.008 3 X=0.019, Y=110.000 4 X=0.012, Y=-110.014 5 X=-110.025, Y=110.000 6 X=-109.989, Y=-110.004 7 X=-109.980, Y=110.004 8 X=110.010, Y=-110.003	X=0.01 Y=0.17	0.002
1995	Zeiss Jena	LMK	23x23	304.966	1 X=111.994, Y=0.002 2 X=-111.992, Y=0.000 3 X=0.009, Y=112.002 4 X=-0.008, Y=-112.007 5 X=110.002, Y=109.998 6 X=-110.001, Y=-109.991 7 X=-109.993, Y=109.991 8 X=109.990, Y=-110.000	X=0.034 Y=0.052	0.005

1995	Zeiss Jena	LMK	23x23	152.60	1 X=-109.999, Y=-110.008 2 X=110.006, Y=109.983 3 X=-109.989, Y=109.995 4 X=109.992, Y=-110.008 5 X=-111.999, Y=-0.007 6 X=112.003, Y=-0.007 7 X=0.009, Y=111.986 8 X=-0.004, Y=-112.014	X=0.01 Y=-0.01	0.012
1998	Zeiss	LC	23x23	152.233	1 X=110.005, Y=-0.023 2 X=-110.009, Y=-0.014 3 X=-0.024, Y=109.989 4 X=0.005, Y=-110.024 5 X=109.938, Y=109.974 6 X=-110.000, Y=-110.018 7 X=-110.004, Y=109.984 8 X=110.003, Y=-110.024	X=0.004 Y=-0.002	0.007