

Streszczenie

Tematem niniejszej rozprawy jest wykorzystanie technologii śledzenia ruchu gałek ocznych (okulografii) w celu udoskonalenia głosowego systemu dialogowego do celów terapeutycznych. System ten został stworzony na potrzeby pacjentów z rozpoznaniem schizofrenii. Podczas rozmów pacjentów z Terabotem, wystąpił szereg problemów, głównie związanych z płynnością dialogu. W celu znalezienia odpowiednich rozwiązań, wykorzystano dane okulograficzne.

Obecnie okulografia jest szeroko wykorzystywana zarówno w badaniach naukowych jak i zastosowaniach komercyjnych obejmujących różnorodne dziedziny, np. medycynę, edukację, psychologię czy marketing. Przykładem zastosowania w dziedzinie "komunikacji człowiek – komputer" są tzw. agenty konwersacyjne lub systemy dialogowe, którym poświęca się ostatnio wiele uwagi. Jednym z zastosowań jest np. wykorzystanie zachowań niewerbalnych do nawiązania kontaktu z użytkownikiem. Kluczową koncepcją w projektowaniu tych systemów jest wykorzystanie zachowań niewerbalnych do zbierania informacji o użytkownikach, zamiast polegania wyłącznie na mowie. Aby dialog między agentem a człowiekiem stał się bardziej realistyczny, agenci mogą być wyposażeni w zdolności komunikacyjne i ekspresyjne podobne do tych obserwowanych w interakcji z ludźmi, w czasie której mamy do czynienia z ich mową, gestami mimiką twarzy czy wzrokiem). Dzięki temu dialog może przebiegać bardziej naturalnie, a systemy mogą zapewnić bardziej realistyczną rozmowę z użytkownikami. Dane ze śledzenia ruchu gałek ocznych mogą służyć jako dodatkowe źródło informacji, szczególnie podczas rozmów z agentami dialogowymi w aspekcie zdrowia psychicznego.

Terapeutyczny system dialogowy Terabot został zaprojektowany przez zespół badawczy z Politechniki Warszawskiej (PW), Wydziału Elektroniki i Technik Informacyjnych. Jego utworzenie miało na celu wzbogacenie terapii pacjentów z zaburzeniami takimi jak np. depresja, zaburzenia lękowe itp. W celu zweryfikowania skuteczności systemu przeprowadzono pilotażowe randomizowane badanie kliniczne w Instytucie Psychiatrii i Neurologii w Warszawie. 38 pacjentów miało możliwość przeprowadzania rozmów terapeutycznych z systemem dialogowym Terabot (oprócz bieżącego leczenia, w tym farmakologicznego). Podczas rozmowy z Terabotem pacjent mógł wybrać jedną z trzech emocji: lęk, strach lub wstyd. Pod koniec rozmowy Terabot zachęcał do wykonania ćwiczenia relaksacyjnego, które miało pomóc pacjentom w uspokojeniu się i opanowaniu emocji. Zbierane były dane okulograficzne oraz nagrania audio z dialogów pomiędzy pacjentami a Terabotem. Bezpośredni kontakt z pacjentami pozwolił na znacznie efektywniejszą optymalizację systemu dialogowego.

Na podstawie analiz zweryfikowano dotychczasowe działanie istniejącego systemu dialogowego. Przeprowadzono analizy interfejsu graficznego Terabota, w szczególności obszarów zainteresowania pacjentów (ang. area of interest, AOI). Analizowano gdzie pacjenci patrzyli najdłużej i najczęściej podczas gdy kończyli swoją wypowiedź. Wyniki analiz stały się kluczowym źródłem informacji, które posłużyły do zaprojektowania kolejnej wersji systemu. Nowy system jest systemem multimodalnym, wspomaganym przez sygnał okulograficzny. Jest on

w stanie reagować na zachowanie pacjenta (jego spojrzenie) podczas rozmowy. Ponadto system ten opiera się na dużych modelach językowych (ang. Large Language Models, LLM), w szczególności na jednym z modeli GPT firmy OpenAI. Takie podejście sprawia, że nowy system jest bardziej zaawansowany nie tylko naukowo, ale i technologicznie. Dzięki zastosowaniu okulografu (w przeciwieństwie do systemu dialogowego bez dodatkowych narzędzi) można było dostosować konwersacje do zachowania pacjenta. Pozwalało to na bardziej naturalną interakcję między pacjentem a systemem dialogowym. Dane okulograficzne posłużyły do rozwiązania problemów, które powodowały przerwy w dialogu między systemem a pacjentem. Ponadto technologię śledzenia ruchu gałek ocznych wykorzystano również do automatycznej oceny zaangażowania pacjentów w dialog. Proponowane rozwiązania mogą mieć pozytywny wpływ na interakcję pomiędzy Terabotem a pacjentem. Może się to również przyczynić do zwiększenia dobrego samopoczucia pacjentów i komfortu podczas sesji terapeutycznych.

Słowa kluczowe: okulografia, system dialogowy, interakcja człowiek – komputer, systemy multimodalne, psychiatria, komputerowe wspomaganie medycyny, terapia poznawczo - behawioralna, fiksacje, duże modele językowe, GPT, automatyczna ocena zaangażowania, kalibracja