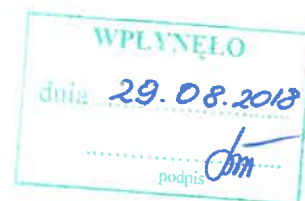


Prof. dr hab. inż. Bożena Kostek, prof. zw. PG
Politechnika Gdańska, Wydział Elektroniki, Telekomunikacji i Informatyki
Lab. Akustyki Fonicznej



14. 08. 2018 r.

**Opinia nt. rozprawy doktorskiej mgr inż. Katarzyny
Barczewskiej pt.: „Automatyczne rozpoznawanie gestów na potrzeby
wspomagania komunikacji z komputerem osób z porażeniem
mózgowym ze szczególnym uwzględnieniem dzieci”**

Cel i tezy rozprawy, zawartość rozprawy i metodyka rozwiązywania problemów

Przedmiotem recenzji jest rozprawa doktorska mgr inż. Katarzyny Barczewskiej. Recenzowana rozprawa doktorska ma charakter eksperymentalno-aplikacyjny, obejmuje 129 stron tekstu wraz z rysunkami i składa się z siedmiu rozdziałów, dwóch Załączników oraz Bibliografii liczącej 111 pozycji.

Podjęta w rozprawie problematyka dotyczy w ogólności opracowania algorytmów, pozwalających na zautomatyzowanie procesu wspomagania komunikacji z komputerem przez niepełnosprawnych użytkowników cierpiących na zaburzenia i ograniczenia ruchowe związane z dziecięcym porażeniem mózgowym (MPD). **Głównym (i ambitnym) celem pracy** było zaproponowanie metod z zakresu modelowania gestów, które pozwoliłyby na stworzenie systemu, który na podstawie niewielkiej liczby powtórzeń pojedynczych gestów wykonywanych przez nowego użytkownika byłby w stanie tworzyć spersonalizowany zestaw gestów na potrzeby interakcji z komputerem.

W pierwszej kolejności należy wskazać na istotę i społeczny charakter problemu komunikacji niepełnosprawnego użytkownika cierpiącego na zaburzenia i ograniczenia ruchowe związane z dziecięcym porażeniem mózgowym. W ogólności można powiedzieć, że charakterystyka MPD jest bardzo złożona, gdyż ma wiele postaci i odnosi się do bardzo szerokiego spektrum różnorodnych, zmieniających się wraz z wiekiem zaburzeń ruchu i postawy, często współistniejących z innymi objawami trwałego, organicznego uszkodzenia mózgu. Problematyka ta jest w sposób oczywisty istotna ze względów społecznych, ale również bardzo aktualna w kontekście możliwości technologicznych tworzenia automatycznych systemów komunikacji użytkownik-komputer (HCI). Należy zwrócić również uwagę na fakt, że zastosowane przez doktorantkę podejście ma **na celu stworzenie zindywidualizowanego, a więc personalnego systemu rozpoznawania**

pojedynczych gestów. Tego typu metodologia może z łatwością być rozszerzona na inne grupy osób z problemami zdrowotnymi, np. komunikację osób z uszkodzonym ośrodkowym układem nerwowym (OUN), osób z autyzmem, zaburzoną kontrolą aparatu mowy, itd. Proponowana metodyka wpisuje się też w pojęcie komunikacji alternatywnej i wspomagającej (AAC – *Augmentative and Alternative Communications*), która jest obecnie standardowym postępowaniem rehabilitacyjnym osób, które mają problemy z werbalną komunikacją. Jak wcześniej wspomniano – istotny jest cel społeczny, a na ten aspekt zwraca również uwagę doktorantka. Nie do przecenienia jest bowiem zapewnienie możliwości ekspresji osobom mających problemy z komunikacją werbalną, ale może jeszcze w większym stopniu na możliwości ciągłej rehabilitacji (i motywowania) osób cierpiących na zaburzenia związane z dziecięcym porażeniem mózgowym.

Ze względu na potrzebę odniesienia się w dalszej części recenzji do postawionych w rozprawie tez, zostaną one przywołane poniżej. Są one wynikiem wypracowanych przez doktorantkę założeń i ich pogłębionej dyskusji. W ten sposób sformułowana została w rozprawie jedna główna teza:

Teza 1: Z wykorzystaniem metod sztucznej inteligencji można skonstruować system, który sam uczy się nowych gestów wykonywanych przez nowego użytkownika na podstawie niewielkiej ilości materiału treningowego.

oraz pięć dodatkowych tez pomocniczych:

Teza 2: Metodologię modelowania gestów języka migowego można wykorzystać na potrzeby modelowania innych gestów. Połączenie tej metodologii z elementami automatycznej analizy danych może poprawić jakość działania systemu.

Teza 3: Każdy nieznaną gest dynamiczny, niekoniecznie będący gestem języka migowego, można zakodować przy pomocy sekwencji znanych podjednostek gestu. Dekodowanie sekwencji podjednostek może się odbyć z wykorzystaniem kryterium najwyższej wiarygodności.

Teza 4: Model nieznanego gestu można zbudować poprzez połączenie znanych modeli podjednostek gestu odpowiadających odkodowanej sekwencji. Połączenie sekwencji modeli podjednostek w model gestu może zostać zrealizowane poprzez iloczyn prawdopodobieństw.

Teza 5: Do zbudowania modelu nowego gestu pozwalającego na uzyskanie wysokiej wartości skuteczności rozpoznawania wystarczy pojedyncze powtórzenie nowego gestu.

Teza 6: Stworzenie alternatywy modeli dla jednego gestu pozwoli lepiej odzwierciedlić osobnicze zróżnicowanie w sposobie wykonywania tego samego gestu, a tym samym wpłynie na wzrost skuteczności rozpoznawania.

Zawartość rozprawy: We Wstępie autorka w pierwszej kolejności przedstawiła genezę tematu rozprawy oraz w konsekwencji potrzebę opracowania systemu rozpoznawania gestów osób z MPD. Jest to wynik braku rozwiązań automatycznego rozpoznawania gestów osób niepełnosprawnych, które w szybki i prosty dla użytkownika sposób pozwoliłyby na zdefiniowanie zindywidualizowanego zestawu gestów na potrzeby interakcji z komputerem. W rozdziale tym podane są również w sposób dość szczegółowy wymagania dotyczące modelowania i rozpoznawania gestów w proponowanym systemie, tezy oraz cel pracy, a także zakres i układ pracy.

Rozdział 2. jest rozdziałem przeglądowym i odwołuje się do stanu wiedzy, techniki i technologii związanego z automatycznym rozpoznawaniem gestów. Doktorantka odnosi się nie tylko do rozwiązań, które zostały przygotowane dla osób cierpiących z powodu dziecięcego porażenia mózgowego, ale traktuje szerszą tematykę akwizycji, modelowania oraz rozpoznawania gestów. Rozdział ten zawiera również zagadnienia komunikacji z komputerem, interakcji z wirtualną rzeczywistością, systemów wspomagania rehabilitacji za pomocą gestów przez niepełnosprawnych użytkowników czy wreszcie aplikacji umożliwiającej definiowanie gestów, szczególną uwagę zwracając na możliwość adaptacji do gestów nowego użytkownika. Doktorantka podaje też definicje gestów, które różnią się w zależności od źródeł literatury, co z kolei ma duże znaczenie w kontekście sformułowania założeń projektowych. Kolejne rozdziały - konsekwentnie w kontekście modułów tworzonego systemu – obejmują przegląd metod akwizycji danych opisujących gesty użytkowników, metody przetwarzania danych z podziałem na sposób akwizycji danych, tj. ekstrahowane z sygnałów pochodzących z czujników oraz parametry wykorzystujące obraz z kamery/kamer, jak również automatyczne rozpoznawanie gestów. Ten ostatni wymieniony podrozdział jest bardzo istotny w przytoczonym przeglądzie, gdyż prowadzi do wniosków dotyczących aktualnych wyzwań stawianych rozwiązaniom technologicznym i tworzącym systemy wspomagania osób niepełnosprawnych. Interesujący jest układ tego podrozdziału, w którym autorka rozprawy w pierwszej kolejności odnosi się do możliwych scenariuszy prowadzenia eksperymentu, co – jak wiadomo - ma zasadniczy wpływ na sposób akwizycji, przetwarzania i rozpoznawania danych, a także na metody adaptacji istniejących modeli gestów zawartych w słowniku do tych samych gestów wykonywanych przez nowego użytkownika. W końcowej części tego rozdziału znalazł się podrozdział dotyczący zastosowań automatycznego rozpoznawania gestów oraz wspomniane już wyzwania, które stanowią listę potrzeb i nierozwiązanych problemów w tematyce automatycznego rozpoznawania gestów.

Rozdział 3. jest pierwszym z rozdziałów odnoszących się do pracy własnej autorki rozprawy. We wstępie do tego rozdziału przedstawione zostało streszczenie poszczególnych etapów realizacji pracy, które następnie konsekwentnie zostały opisane w kolejnych podrozdziałach: akwizycja i przetwarzanie danych, grupowanie cech, modelowanie gestów, trening i rozpoznawanie gestów, równoległe ukryte modele Markowa, fuzja modeli równoległych, dopasowanie podjednostek do modelowania gestów spoza pierwotnej bazy gestów, metodyka ewaluacji opracowanego systemu, w tym miary oceny działania systemu. Rozdział ten został opatrzony odpowiednio: wzorami, przykładami zaimplementowanych algorytmów oraz schematami ilustrującymi działanie poszczególnych algorytmów.

Kolejny z rozdziałów autorskich (prace własne) obejmuje eksperymenty wykonane przez autorkę rozprawy. Wykorzystywane bazy zostały opisane w pierwszym z podrozdziałów i zestawione tabelarycznie (tab. 4.1). Bazy zawierają gesty wykonywane zarówno przez zdrowych, jak i niepełnosprawnych użytkowników. Bazy gestów IMU oraz IMU-SZKOŁA powstały z wykorzystaniem sensora inercyjnego (ang. *Inertial Measurement Unit*, IMU). Baza IMU zawiera gesty osób pełnosprawnych, natomiast IMU- SZKOŁA gesty osób z porażeniem mózgowym zarejestrowane w szkole specjalnej. Baza WiTKOM jest wynikiem wcześniejszych prac w ramach projektu NCBiR, u podstaw której była rejestracja danych z wykorzystaniem sensora Kinect 2.0 i wreszcie baza DR-TEST, która posłużyła do walidacji opracowanych metod. Uwarunkowania rejestracji, jak również umiejscowienie sensora zmieniały się, aby uzyskać w ten sposób bardziej zróżnicowany materiał do badań i ewaluacji systemu. Ponadto, co jest jednym z bardziej istotnych elementów charakterystyki tej bazy, potencjalni użytkownicy tworzyli własne gesty, które mogłyby zostać wykorzystane do obsługi typowych funkcjonalności komputera. W końcowej części tego rozdziału zostały przedstawione scenariusze badań. Szczegółowe wyniki badań zostały zawarte w rozdziale 5. Zostały one przedstawione w odniesieniu do poszczególnych baz. Istotne jest każdorazowe podsumowanie wyników, które zawiera szczegółowe wnioski, wynikające z przeprowadzonych eksperymentów dla poszczególnych baz, jak również zbiorcze zestawienie tych wniosków. Interesujące są hipotezy badawcze, które doktorantka zweryfikowała dla bazy DR-TEST, zawierającej gesty zaproponowane przez użytkowników, a więc najbardziej prawdopodobny scenariusz wykorzystania opracowanych algorytmów.

W rozdziale 6. zawarte zostały wnioski, które zostały przedstawione w odniesieniu do sformułowanych przez doktorantkę tez rozprawy tak, aby wykazać ich prawdziwość. W tym miejscu należy nadmienić, że przytoczone skrócone wnioski wynikające z przeprowadzonych eksperymentów w pełni uzasadniają prawdziwość zaproponowanych tez pracy. Rozdział 7. jest podsumowaniem rozprawy, wskazuje na oryginalne elementy rozprawy, napotkane w badaniach trudności oraz możliwe

kierunki dalszej pracy związanej z tematem, które niektóre z nich warto - za autorką rozprawy - wymienić:

- stworzenie uniwersalnego leksykonu podjednostek gestów przy pomocy klasteryzacji podjednostek;
- zastosowanie modeli bigramów i trigramów w modelowaniu gestykulacji przy pomocy ograniczonego leksykonu podjednostek;
- przetestowanie różnych modeli podjednostek – zwiększanie liczby stanów przypadających na pojedynczą podjednostkę, sprawdzenie różnych topologii modelu HMM;
- wprowadzenie założenia o modelowaniu każdego nowego gestu przy pomocy takiej samej liczby podjednostek, a następnie przeprowadzenie rozpoznawania na podstawie porównywania ciągów kodów przy pomocy metody DTW (ang. *Dynamic Time Warping*) lub odległości edycji;
- porównanie modelowania przy pomocy HMM z *Conditional Random Fields*;
- wykorzystanie innych metod fuzji modelu równoległego, w tym sieci neuronowe i maszyny wektorów nośnych (ang. *Support Vector Machine*);
- modelowanie sekwencji gestów przy pomocy dodania do HMM stanu odpowiadającego „ciszy”;
- poszerzenie badań o gesty wykonywane obydwoma rękami;
- poszerzenie badań o gesty statyczne;
- poszerzenie wektora cech o parametry opisujące kształt dłoni;
- przetestowanie zaproponowanych metod na bazach gestów dostępnych na stronach organizacji ChaLearn;
- sprawdzenie innego sposobu grupowania cech.

Ostatni z rozdziałów: Bibliografia został zamieszczony po dwóch Załącznikach, które zawierają odpowiednio: Eksperymenty z grami na tablet oraz Aplikację demonstracyjną.

Ocena zawartości rozprawy i metod rozwiązywania postawionych problemów

W celu weryfikacji zaproponowanych tezy doktorantka przyjęła następujący schemat postępowania:

- sformułowała pojęcie gestu oraz podała sposób, w jaki można dokonywać akwizycji danych opisujących gestykulację,
- przedstawiła metody przetwarzania danych w zależności od wybranego algorytmu ich akwizycji,

- dokonała przeglądu metod modelowania gestów wykorzystywanych w systemach automatycznego rozpoznawania oraz zastosowań automatycznego rozpoznawania gestów w kontekście aktualnych wyzwań związanych z tematyką rozpoznawania gestów,
- przedyskutowała kolejne elementy i etapy procesu tworzenia systemu komunikacji użytkownika z komputerem: urządzenia do akwizycji danych (czujnik IMU 9DoF oraz sensor Microsoft Kinect 2.0), metody przetwarzania danych, modelowanie gestów przy pomocy podjednostek i ukrytych modeli Markowa,
- przedyskutowała metodykę adaptacji wypracowanych modeli do gestów nowego użytkownika, a także parametry wykorzystane do ewaluacji systemu,
- eksperymentalnie zweryfikowała podstawione w rozprawie tezy.

Oceniając rozprawę, należy w pierwszej kolejności stwierdzić, że przytaczane treści w rozdziałach przeglądowych są trafnie dobrane. Należy podkreślić, że erudycja i prace algorytmiczne autorki w obszarze tematyki rozprawy tj. implementacji metod przetwarzania oraz rozpoznawania danych pozyskanych z akwizycji gestów zasługują na wysokie uznanie. Zostały one poprawnie zrozumiane i zaimplementowane. Można w ogólności stwierdzić, że wybór algorytmów jest właściwy, gdyż pozwala na osiągnięcie postawionego celu rozprawy i udowodnienie tez rozprawy. Należy też zwrócić uwagę na pracowitość prac badawczych, również w części przygotowania bazy gestów. W związku z trudnością pozyskania danych od docelowych niepełnosprawnych użytkowników, doktorantka zaproponowała specjalne podejście do tworzenia nagrań, w trakcie których każdy pełnosprawny użytkownik wykonywał dany zestaw gestów na trzy sposoby: naturalnie, przesadnie powoli, przesadnie szybko. Taki sposób tworzenia bazy gestów pozwolił w lepszym stopniu odzwierciedlić różnicowanie w sposobie wykonywania gestów przez osoby niepełnosprawnych. Ponadto uwzględniła możliwość tworzenia własnego zestawu gestów przez użytkowników i w różnych uwarunkowaniach (np. zmienne warunki oświetlenia).

Na uwagę zasługuje przyjęta metodologia podejścia do rozpoznawania gestów, znalazły się w niej bowiem nowatorskie rozwiązania, niektóre z nich przywołam poniżej.

Doktorantka zaproponowała nowy prosty wektor cech opisujący gesty dynamiczne wykonywane jedną ręką na podstawie danych szkieletowych z sensora Kinect. Wyniki uzyskane dla modeli wytrenowanych na tym wektorze cech porównane zostały z opisanym w literaturze sposobem parametryzacji

danych szkieletowych. Zaproponowała metodę automatycznego podziału cech do równoległych kanałów modelu PaHMM (równoległe ukryte modele Markowa (ang. *Parallel Hidden Markov Models*), w której macierz odległości korelacji pomiędzy cechami z wektora cech stanowi wejście do klasteryzacji hierarchicznej, mającej na celu zgrupowanie cech w niezależne grupy. W obrębie każdej z grup cech wykonała także analizę składowych głównych, która poza redukcją liczby wykorzystanych cech miała na celu uniezależnienie ich od siebie, co jest istotne podczas trenowania wykorzystanych modeli podjednostek (HMM-GMM). Udowodniła, że podział cech do poszczególnych kanałów modelu równoległego nie musi być zgodny ze wskazówkami lingwistycznymi wykorzystywanymi przy modelowaniu gestów języka migowego. Przeprowadziła selekcję kanałów wykorzystywanych w fuzji modelu równoległego PaHMM, które najlepiej sprawdzają się z punktu widzenia uzyskania najwyższej wartości skuteczności rozpoznawania, a także błędu zrównoważonego. Opracowała ponadto metodę automatycznej indywidualnej rekomendacji gestów na potrzeby interakcji z komputerem, przy pomocy której można wskazać zestaw gestów, które najlepiej się sprawdzą w interakcji człowiek komputer.

Metoda automatycznego grupowania cech zaproponowana przez autorkę pracy, pozwoliła osiągnąć takie same wyniki jak podejście, w którym podział cech następuje na podstawie wskazówek lingwistycznych zaczerpniętych z analizy gestów języka migowego, a więc uwzględnienie wskazówek lingwistycznych nie jest konieczne do modelowania gestów przy pomocy modelu równoległego.

W wyniku przeprowadzonych eksperymentów z wykorzystaniem przygotowanych baz gestów doktorantka wykazała, że zaproponowana metoda najlepszego dopasowania podjednostek pozwala stworzyć modele nowych gestów nowego użytkownika na podstawie niewielkiej liczby powtórzeń – najlepsze rezultaty uzyskano dla sześciu powtórzeń nieznanego gestu. Jednak nawet modele stworzone w oparciu o pojedyncze powtórzenie dają satysfakcjonujące wyniki automatycznego rozpoznawania gestów. Ponownie, warto uwypuklić, że ważnym aspektem wykorzystania metody automatycznego grupowania cech jest zredukowanie liczby cech budujących poszczególne kanały PaHMM (równoległe ukryte modele Markowa (ang. *Parallel Hidden Markov Models*), a więc uproszczenie modelu. Liczba grup, na które powinien być dzielony wektor cech w klasteryzacji hierarchicznej została wskazana na podstawie uzyskania maksymalnej separowalności przy zachowaniu najmniejszej liczby grup.

Należy też zwrócić uwagę na fakt, iż rozprawa ma wyraźnie walory praktyczne i może być w przyszłości wykorzystana przez osoby niepełnosprawne. W celu zweryfikowania opracowanych metod w rzeczywistym środowisku pracy niepełnosprawnego użytkownika, autorka rozprawy zaprojektowała dwie gry na tablet, które pozwoliły na przeprowadzenie dodatkowych eksperymentów (Załącznik 1) oraz aplikację demonstracyjną (Załącznik 2), umożliwiającą pracę z komputerem

przy pomocy spersonalizowanych gestów. Aplikacja ta umożliwi zarówno nagrywanie własnych gestów, wybór funkcjonalności komputera, które mają być obsługiwane przy pomocy gestu, jak i sterowanie komputerem przy pomocy nagranych zestawów gestów. Automatyzacja procesu rozpoznawania pojedynczych gestów jest na pewno cennym osiągnięciem, dlatego również zawartość techniczna zasługuje na wysoką ocenę.

Poniżej przedstawię uwagi i pytania szczegółowe, które się nasunęły w trakcie czytania rozprawy:

1. Sądzę, że warto by zweryfikować zestaw „własnych” gestów odpowiadających podstawowym funkcjonalnościom komputera w grupie osób, zaproponowanych przez „zdrowych” użytkowników, osobom, które nie mają doświadczenia w obsłudze komputera oraz przedstawieniu je do krytycznej oceny osobom niepełnosprawnym. Oczywiście można się zgodzić, że ponieważ system pozwala na personalizowanie gestów, to w jakimś stopniu odpowiedź na tę sugestię znalazła się w rozprawie, ale dla lepszej systematyki warto by chyba było pokusić się o takie dodatkowe badania.
2. Może warto by było, aby autorka rozprawy sięgnęła do wypracowanych metod wspomnianego w opinii procesu rehabilitacji osób z problemami komunikacji werbalnej, tj. metod komunikacji alternatywnej i wspomagającej (AAC – *Augmentative and Alternative Communications*), która z jednej strony zawiera system gestów i symboli graficznych (MAKATON), ale również wskazuje, w jaki sposób należy tworzyć stopnie trudności w opracowywanych systemach gier czy zabawy.
Dodatkowo, w proponowanych metodach rehabilitacji częstym zabiegiem jest stosowanie słownika gestów (dostosowanych do niepełnosprawnego użytkownika), który jest wykorzystywany jako podpowiedź. W badaniach tego typu badacze (i rehabilitanci) wskazują na zwiększoną motywację do ćwiczeń osób niepełnosprawnych.
3. Zaprezentowany przez doktorantkę przegląd literatury można uznać za pełny w zakresie zagadnień przedstawianych w rozprawie. Jednak warto może by sięgnąć do źródeł literatury, które pozornie nie są związane z tematem pracy p. Katarzyny Barczewskiej, ale mogłyby wnieść dodatkowe usprawnienia w tworzeniu ewentualnego słownika czy zestawu gestów, zwłaszcza w kontekście prac rozwojowych.
Jako przykład rozwiązania opartego całkowicie na przetwarzaniu obrazu wizyjnego, umożliwiającego rozpoznawanie prostych gestów dłoni i stóp, podać można system Music Maker [M. Gorman, M. Betke, E. Saltzman, A.

Lahav, Music Maker - A Camera-based Music Making Tool for Physical Rehabilitation, Computer Science Technical Report, 2005-032, 2005]. System umożliwia przetwarzanie sygnałów muzycznych za pomocą gestów wykonywanych w ramach rehabilitacji ruchowej. Użytkownik poprzez ruch dłoni lub stopy kontroluje takie cechy sygnału, jak poziom i wysokość dźwięku. Spośród interfejsów umożliwiających przetwarzanie dźwięków muzycznych za pomocą gestów rąk wymienić można również system WAVE (*Virtual Audio Environment*) autorstwa Valboma i in. [*L. Valbom, Adérito Marcos, WAVE: Sound and music in an immersive environment, Computers & Graphics, 29, 6, 871-881, 2005*]. Rozwiązanie to umożliwia wyzwalenie za pomocą gestów rąk wybranych pętli muzycznych lub odtwarzanie dźwięków skal chromatycznych. Ruch rąk odzwierciedla ruch wirtualnych różdżek na ekranie, za pośrednictwem których możliwe jest wybieranie i zmiana położenia obiektów. Berthaut i in. bazując na pracach Cadoza [*C. Cadoz, Musique, Geste Technologie, Cultures Musicales: Les Nouveaux Gestes de la Musique, H. Genevoix and R. De Vivo (red.), Marseille, Parentheses, 47-92, 1999*] dotyczących gestów muzycznych wyszczególnili trzy grupy gestów, tj. gesty selekcji – służące do wyboru obiektów reprezentujących dźwięki, gesty modulacji – odpowiedzialne za przetwarzanie dźwięków i gesty pobudzenia – służące do wyzwiania dźwięków. Prace tego typu były również zrealizowane przez dra Michała Lecha w ramach jego rozprawy doktorskiej [*M. Lech, B. Kostek, The evaluation of influence of ergonomics and multimodal perception on sound mixing results employing the novel gesture-based mixing interface, J. Audio Eng. Soc. 61, 5, 301-313, 2013*]. Zamierzeniem autora rozprawy było stworzenie systemu, który poza wyznaczeniem ergonomicznego sposobu obsługi procesów miksowania za pomocą gestów, zapewniałby warunki umożliwiające wiarygodne sprawdzenie wpływu bodźców wzrokowych na decyzje podejmowane w trakcie kreowania podstawowych elementów każdego miksu, tj. balansu (poziomy i relacje częstotliwościowe), szerokości bazy, dynamiki i głębi. W podejściu dra Lecha do obsługi operacji miksowania dźwięku opracowany został zunifikowany słownik gestów rąk. Słownik ten powstał w oparciu o badanie ankietowe oczekiwań potencjalnych użytkowników systemu względem sposobu obsługi za pomocą gestów oraz intuicyjności gestów i przypisanych im funkcji, zaproponowanych przez autora rozprawy. Oczywiście są to tylko przykłady rozwiązań dotyczące automatycznego rozpoznawania gestów do zastosowania w operowaniu dźwiękiem.

4. W rozprawie nie znalazłam dyskusji na temat problemu skalowalności opracowanej algorytmiki w kontekście tworzenia bardziej uniwersalnych zestawów gestów (złożoność obliczeniowa algorytmów, uwarunkowania

czasowe). Pomimo iż, postawiony cel w rozprawie mówi o personalizowaniu gestów i nauce systemu na podstawie pojedynczego (pojedynczych) gestów, pytanie takie wydaje się zasadne.

5. Sądzę, że tytuł rozdziału „Motywacja” obejmuje szereg dodatkowych i istotnych zagadnień, które jednak nie odnoszą się do motywacji leżącej u podstaw tematu czy genezy tematu rozprawy. Dla przykładu – sądzę, że bardzo ważna dyskusja dotycząca wymagań systemowych powinna znaleźć się w rozdziale pt. Założenia projektowe. Dodatkowo niektóre opisy zawarte w tym rozdziale odnoszą się do zakresu pracy, który stanowi osobny podrozdział Wstępu.
6. Niezbyt fortunny jest tytuł: „Wykorzystanie gestów osób z porażeniem mózgowym”. Zawarty w tym podrozdziale opis odnosi się do zagadnienia komunikacji z komputerem, interakcji z wirtualną rzeczywistością, systemów wspomagania rehabilitacji za pomocą gestów przez niepełnosprawnych użytkowników czy wreszcie aplikacji umożliwiającej definiowanie gestów.
7. Tytuł kolejnego podrozdziału (2.2) nie jest z kolei typowy dla pracy o charakterze technicznym („Czym są gesty”).
8. Pewien niedosyt pozostawia opis akwizycji danych. Dla przejrzystości przedstawione zestawienie przykładów ekstrahowanych parametrów wektora cech mogłoby być tabelaryczne z odniesieniem do źródeł, które podaje autorka rozprawy.
9. Tytuł rozdziału 2.5.3 wydaje się zbyt skrótowy. Ta sama uwaga dotyczy rozdziału 3.
10. Jak wspomniano wcześniej, Bibliografia została zamieszczona po dwóch Załącznikach, które zawierają odpowiednio Eksperymenty z grami na tablet oraz Aplikację demonstracyjną, co nie jest typowe - Bibliografia jest integralną częścią rozprawy.
11. W Bibliografii autorka wykorzystuje zapis „et al.”. Zwykle w spisie literatury w rozprawach doktorskich czy książkach zamieszcza się pełną listę nazwisk (nawet, jeśli jest kilku autorów).

Podsumowując tę część opinii, chciałabym podkreślić, przeprowadzone eksperymenty badawcze mają bardzo szeroko zakrojone prace eksperymentalne, a dyskusja wyników jest przedstawiona bardzo systematycznie, wręcz wzorcowo.

Uwagi redakcyjne

Autorka nie ustrzegła się przed drobnymi potknięciami. W tekście rozprawy można znaleźć dość liczne usterki interpunkcyjne oraz drobne błędy edytorskie, brak spacji czy tzw. „literówki”. W języku polskim odwołania do rysunków i tabel są z małej litery (rys. x.x, tab. x.y, itp.) Usterki te nie zostały przywołane w niniejszej recenzji ze względu na ich małą istotność w zrozumieniu treści przekazu rozprawy.

Podsumowanie

Tematyka rozprawy mgr inż. Katarzyny Barczewskiej wpisuje się w aktualne potrzeby dotyczące automatycznego rozpoznawania gestów w szczególności w kontekście osób niepełnosprawnych. Do osiągnięć rozprawy zaliczam w pierwszej kolejności przygotowanie i zestawienie kilku baz nagrań gestów, które zostały następnie wykorzystane w procesie automatycznego rozpoznawania gestów. Zaproponowane algorytmy automatycznego tworzenia gestów zostały następnie w sposób bardzo staranny poddane weryfikacji eksperymentalnej i pozwoliły na wykazanie prawdziwości tez rozprawy. Zaproponowana metoda najlepszego dopasowania podjednostek pozwala stworzyć modele nowych gestów nowego użytkownika na podstawie niewielkiej liczby powtórzeń. Również modele stworzone w oparciu nawet o pojedyncze powtórzenie dają satysfakcjonujące rezultaty automatycznego rozpoznawania gestów. Dodatkowym ważnym elementem rozprawy jest przygotowanie gier na tablet oraz aplikacji demonstracyjnej, które dodatkowo pozwoliły na ocenę jakości opracowanej metodologii. Warto w tym miejscu wspomnieć, że rozwiązania praktyczne wymagają uwzględnienia wielu dodatkowych praktycznych aspektów implementacyjnych, jak np. zmienne warunki otoczenia, z którymi doktorantka poradziła sobie doskonale. Można to stwierdzić poprzez analizę wyników uzyskanych przez osoby niepełnosprawne w trakcie wykonywania zadanego ćwiczenia.

Reasumując, mogę stwierdzić, że Autorka rozprawy wykazała należyty poziom wiedzy w zakresie ogólnym i w zakresie będącym przedmiotem rozprawy, biegłość w zakresie programowania, prowadzenia eksperymentu badawczego oraz umiejętność wdrożenia wyników eksperymentalnych w formie aplikacji komputerowej do sterowania grą na tablecie. Doktorantka osiągnęła postawiony sobie cel rozprawy oraz wykazała przydatność zaproponowanej metodyki, co pozwoliło na sformułowanie wniosku o udowodnieniu tezy głównej i tez pomocniczych.

W podsumowaniu stwierdzam, że przedłożona mi do recenzji rozprawa p. mgr inż. Katarzyny Barczewskiej spełnia wymagania stawiane w Ustawie rozprawom doktorskim (**dyscyplina: biocybernetyka i inżynieria biomedyczna**) i wnoszę o jej dopuszczenie do publicznej obrony.

Ze względu na twórcze podejście do analizowanego problemu oraz aplikacyjny charakter rozprawy, jak również publikację doktorantki w czasopiśmie zawartym w wykazie MNiSW (lista A), wnoszę do Wysokiej Rady o wyróżnienie rozprawy.

Kortej