

Interaktywna sonifikacja obrazów dla niewidomych – badania pilotażowe

Streszczenie. Artykuł dotyczy problemu prezentacji informacji o obrazach z wykorzystaniem techniki sonifikacyjnej polegającej na dźwiękowym, niewerbalnym opisie obrazów często określanej również mianem "wyświetlania dźwiękowego" (ang. auditory display). Omówiono krótko podstawowe metody sonifikacji stosowane w prezentacji obrazów. W referacie skupiono się na tzw. interaktywnej sonifikacji obrazów, w której użytkownik systemu wskazuje dotykiem obszar obrazu do udźwiękowania. W pracy opisano aplikację mobilną, służącą do nauki matematyki dla dzieci niewidomych, którą poddano testom użytkowym w szkole dla dzieci niewidomych. Program służy jako pomoc w nauce przebiegów wykresów funkcji wyświetlanych na ekranie dotykowym. Przemieszczenie palca wzdłuż wykresu funkcji powoduje wygenerowanie tonu o określonej częstotliwości, w miarę oddalania się od wykresu częstotliwość tonu jest zmniejszana. Oceniono przydatność takiej pomocy dydaktycznej w testach z docelowymi użytkownikami.

Abstract. The paper concerns the problem of presenting visual information using image sonification – nonverbal auditory description often referred to as „auditory display”. Common methods used in sonification of images are described. The paper focuses on so called interactive sonification, which allows the user to select an image area for auditory presentation. The paper describes a mobile application created for teaching math to blind children and tested in a school for the blind. The software aids in teaching the shape of function plots displayed on a touchscreen (moving a finger across the display generates a tone dependant on the proximity to the plotted line, with frequency increasing as the distance decreases). The usefulness of such an application in the teaching process was tested with blind users. (**Interactive sonification of images – a pilot study**).

Słowa kluczowe: substytucja sensoryczna, interaktywna sonifikacja, Android, osoby niewidome, tyflopädagogika.

Keywords: sensory substitution, interactive sonification, Android, visually impaired, education of the blind.

Wprowadzenie

Jedną z podstawowych barier napotykaną przez osoby niewidome jest dostęp do informacji graficznej, m.in. zdjęć, map, rysunków i obrazów. Istnieje potrzeba opracowania skutecznych metod przetwarzania informacji wizualnej na sygnały sensoryczne dostępne dla osób niewidomych. Prace nad sonifikacją danych i obrazów zyskują ostatnio na popularności, głównie dzięki postępowi technologicznemu i łatwej implementacji syntezy dźwięku w czasie rzeczywistym [1]. Specjalność naukowa dotycząca zastosowania dźwięku do obrazowania informacji wyłoniła się ok. 20 lat temu. Termin sonifikacja został zdefiniowany przez Kramera [2] jako: *wykorzystanie dźwięku nie będącego mową do przekazywania informacji*. W roku 1992 w Instytucie Santa Fe, USA utworzono International Community for Auditory Display (ICAD). Od tego też roku regularnie odbywają się coroczne konferencje poświęcone tematyce obrazowania dźwiękowego. W roku 2013 po raz pierwszy w Polsce konferencję ICAD zorganizował Instytut Elektroniki Politechniki Łódzkiej (www.icad2013.com).

W roku 2005 wyłoniła się wąska, ale ważna specjalność badawcza tzw. interaktywnej sonifikacji (ang. *Interactive Sonification*) [3]. U podstaw rozwoju tej specjalności leży stwierdzenie, że interakcja użytkownika z systemem sonifikującym ma kluczowe znaczenie dla procesu nauki i nabywania biegłości przez użytkowników w rozpoznawaniu schematów dźwiękowych stosowanych w sonifikacji. Temu aspektowi sonifikacji poświęcono specjalne wydanie czasopisma IEEE Multimedia w 2005 roku [4]. Zagadnienie interaktywnej sonifikacji obrazów dla osób niewidomych było podejmowane przez wielu badaczy, zarówno w Polsce jak i na świecie. Istotnym przykładem krajowym jest dźwiękowy digitizer [5].

Najbardziej znanym algorytmem sonifikacyjnym dla niewidomych jest the vOICe [6]. Algorytm polega na automatycznym odczytywaniu kolejnych kolumn pikseli obrazu monochromatycznego (od lewej do prawej) i ich przekształcania na chwilowe widmo dźwięku. Obraz jest skanowany w jednosekundowych cyklach, przez co niewidomy słyszy krótkie dźwięki o zmiennej

charakterystyce widmowej widmie.

Zauważalny jest jednak wyraźny brak badań podstawowych w dziedzinie sonifikacji dla niewidomych [7]. Publikacje dotyczą zazwyczaj badań stosowanych – przedstawiane są jednostkowe rozwiązania lub schematy sonifikacji [8, 9].

W roku 2016 rozpoczęto w Instytucie Elektroniki Politechniki Łódzkiej dwuletni grant badawczy, pt. *Interaktywna sonifikacja obrazów z przeznaczeniem dla osób niewidomych* finansowany przez Narodowe Centrum Nauki. Hipoteza badawcza projektu jest następująca: możliwe jest opracowanie sterowanych gestami dotykowymi schematów udźwiękowania pozwalających niewidomemu użytkownikowi na poprawną interpretację obrazu wyświetlanego na ekranie dotykowym.

W niniejszym komunikacie przedstawiono prace pilotażowe przeprowadzone w ramach pracy inżynierskiej, pt. *Mobilny system wspomagania nauki matematyki dla osób słabowidzących* wykonanej w Instytucie Elektroniki Politechniki Łódzkiej [10].

Substytucja sensoryczna

Technika substytucji sensorycznej (TSS) polega na reprezentacji lub przekształceniu informacji wizualnej w taki sposób aby była ona dostępna dla zmysłów osoby niewidomej, tj. dla zmysłu dotyku lub słuchu [11]. Przykładem TSS zastosowanej dla zmysłu słuchu jest wspomniana sonifikacja lub audiodeskrypcja (tj. słowny opis obrazu), zaś przykładem TSS dla zmysłu dotyku jest zapis tekstu za pomocą pisma Braille'a oraz reprezentacja obrazów m.in. za pomocą plansz z drukiem wypukłym obrazującym np. mapy geograficzne, plany budynków i ulic.

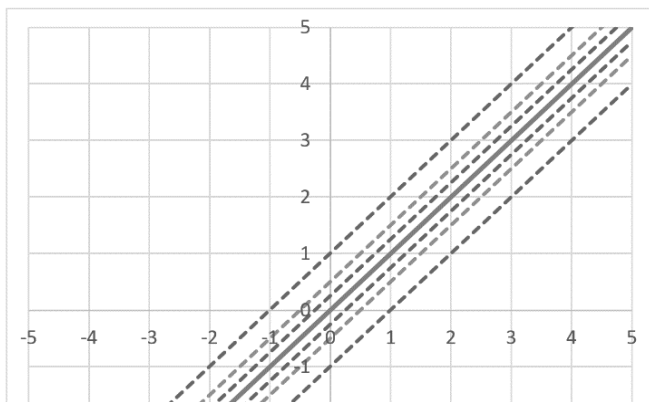
Techniki substytucji sensorycznej wykorzystuje się m.in. w nauczaniu geografii, plastyki i matematyki dla dzieci niewidomych. Przykładem przyrządu o takich właściwościach jest Graphics Aid for Mathematics, tj. tablica wykonana z gumowej maty umożliwiającej wtykanie zatyczek, na których rozpinana jest gumka recepturka przyjmująca kształt zadanej figury geometrycznej. Innym znanym produktem jest Wiki Stix, tj. sznurek pokryty

naturalnym woskiem, za pomocą którego można tworzyć dowolne kształty np. wykresy funkcji i przyklejać do dowolnej powierzchni. Pomoce takie są jednak dość kosztowne i niezbyt trwałe.

Celem podjętych badań pilotażowych było opracowanie metody i aplikacji mobilnej do udźwiękowania wykresów funkcji liniowej wyświetlanej w układzie kartezjańskim na ekranie dotykowym. Program został opracowany jako pomoc w nauczaniu matematyki dla osób słabowidzących i niewidomych. Użytkownik może zadawać współczynniki funkcji, którą zamierza przeanalizować a następnie poprzez dotyk i odpowiednią sonifikację może zbadać przebieg wyświetlanej na ekranie funkcji. Dodatkowo istnieje możliwość zlokalizowania miejsc przecięcia wykresów funkcji z osiami współrzędnych. Aplikacja jest przeznaczona dla uczniów szkoły gimnazjalnej i została napisana na urządzeniu z systemem operacyjnym Android. Z przeprowadzonego wywiadu w Ośrodku Szkolno – Wychowawczym dla Dzieci Słabowidzących i Niewidomych w Łodzi wynika, że dzieci niewidome i słabowidzące do zapoznania się z wykresami funkcji wykorzystują specjalną książkę, w której za pomocą wytłoczeń o różnej wielkości zamieszczono wybrane wykresy funkcji z wytłoczonymi równaniami oraz z układem współrzędnym i samym przebiegiem funkcji.

Metoda udźwiękowania wykresów funkcji

Do udźwiękowania wykresów funkcji zaproponowano interaktywną metodę polegającą na generowaniu tonu o określonej częstotliwości w zależności od odległości miejsca dotknięcia ekranu od wykresu funkcji. Na rysunku 1 pokazano zakresy zmian częstotliwości tonu w zależności od miejsca dotknięcia pola wykresu. Linie przerywane rozgraniczają obszary ekranu, dla których następuje zmiana częstotliwości dźwięku.



Rys. 1. Obszary zmian częstotliwości tonu w miarę oddalania się od wykresu, linią ciągłą pokazano przykładowy przebieg funkcji, liniami przerywanymi pokazano miejsca, w których następuje zmiana częstotliwości generowanego tonu

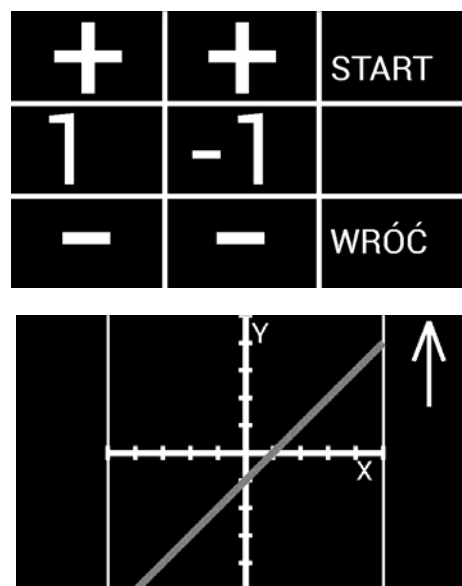
W pierwszym obszarze (w którym znajduje się linia ciągła) jest generowany ton o częstotliwości 800 Hz, w miarę oddalania się od wykresu częstotliwości tonu zmniejszają się z odstępami 100 Hz. Przy zbyt dużej odległości miejsca dotyku od wykresu dźwięk cichnie. Dla dźwięku o częstotliwości 800 Hz margines błędu wynosi $\pm 0,25$ jednostki układu kartezjańskiego dla badanej współrzędnej. Dodatkowo w obszarze najbliższym wykresowi funkcji generowana jest vibracja. Analogicznie dla częstotliwości 700 Hz różnica w wartości współrzędnej może wynosić maksymalnie $\pm 0,5$ jednostki, a dla 600 Hz ± 1 jednostki. Podane zakresy są zdefiniowane w aplikacji w postaci stałych i zostały dobrane na podstawie

przeprowadzonych badań z użytkownikami. Dodatkowo udźwiękowane są osie współrzędnych, po wskazaniu na nie, wytwarzany jest ton o małej częstotliwości. Do generowania dźwięków wykorzystano klasę *AudioTrack*, ponieważ pozwala ona na przesyłanie zbioru próbek sygnałów do bufora, z którego pobierane są dane potrzebne do wygenerowania dźwięku. W aplikacji wykorzystano kodowanie PCM (ang. *Pulse Code Modulation*) z rozdzielczością 16 bitów na próbkę i szybkością próbkowania 9000 próbek/s. Zastosowano dźwięk monofoniczny.

Użytkownik wskazuje palcem fragment obrazu do sonifikacji. Dodatkowo też, dotykając ekranu drugim palcem, może uzyskać informację o punkcie ekranu, na który aktualnie wskazuje pierwszym palcem. Odczytywane są wtedy współrzędne punktu w układzie kartezjańskim oraz informacja, czy wskazywany punkt należy do wykresu. Wszystkie komunikaty w aplikacji są udźwiękowane – do odczytu ich treści wykorzystano standardowe metody zaimplementowane w systemie operacyjnym (klasa *TextToSpeech*) pozwalające na konwersję tekstu na mowę. Dzięki temu użytkownik może korzystać ze swojego ulubionego syntezatora mowy.

Budowa i interfejs graficzny aplikacji

Ponieważ aplikacja jest przeznaczona nie tylko dla osób niewidomych, ale i słabowidzących, szczególną uwagę zwrócono na interfejs graficzny aplikacji. Wszystkie komunikaty są wyświetlane z zastosowaniem dużej, wyraźnej czcionki, a do wyświetlania wykresów funkcji oraz osi współrzędnych wykorzystano linie o odpowiedniej grubości (rys. 2).



Rys. 2. Widok ekranu do wyboru współczynników oraz ekranu do rysowania wykresów funkcji

Dodatkowo, wprowadzono możliwość zmiany zestawu kolorów wykorzystywanych w aplikacji. Dzięki temu użytkownik ma możliwość dopasowania kolorystyki do swoich preferencji.

Aplikacja może zostać uruchomiona na dowolnym urządzeniu z systemem Android, elementy graficznego interfejsu użytkownika skalują się automatycznie, niezależnie od rozdzielczości i proporcji ekranu.

Testy aplikacji udźwiękującej wykresy funkcji

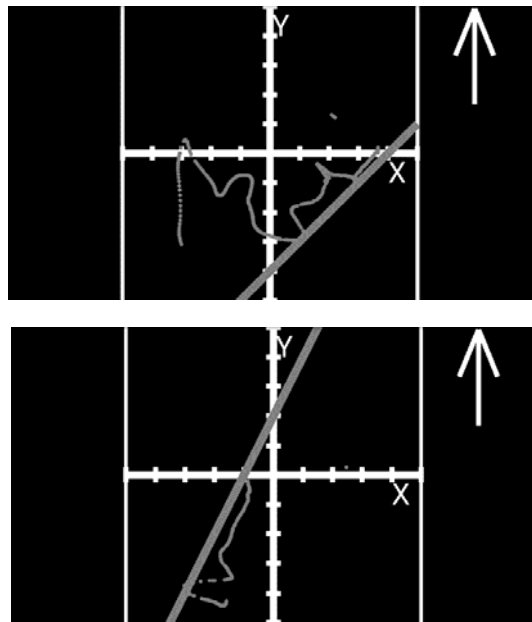
Badania wstępne zostały przeprowadzone z udziałem pięciu osób widzących, które podczas testów miały

zasłonięte oczy. Pozwoliło to na sprawdzenie, czy korzystanie z programu bez możliwości wizualnej interakcji z graficznym interfejsem użytkownika nie sprawia problemów i czy sposób udźwiękowania wyświetlanego wykresu pozwala na wnioskowanie o jego przebiegu. Do testów wykorzystano tablet o przekątnej 10 cali z ekranem dotykowym wspierającym wielokrotny dotyk. Użytkownicy mieli za zadanie wybrać wskazane współczynniki funkcji, sprawdzić jak został zaprezentowany wykres, wskazać miejsca przecięcia wykresu z osiami współrzędnych oraz zlokalizować wskazane punkty w układzie współrzędnych. Jednym z zadań było także zbadanie i określenie równania wykresu funkcji. W programie dodano możliwość rejestrowania trajektorii położenia palca podczas wykonywania zadań, co umożliwia przeprowadzenie analiz skuteczności wybranych metod prezentacji dźwiękowej oraz analizę, w jaki sposób użytkownik nabywa doświadczenie w obsłudze programu. Na rysunku 3 pokazano przykładowe trajektorie podczas poszukiwania miejsc zerowych funkcji.

Każde badanie trwało około 15 minut i poprzedzone było krótką instrukcją dotyczącą wykorzystania aplikacji oraz poruszania się po niej. Dzięki współpracy nawiązanej z Ośrodkiem Szkolno-Wychowawczym dla Dzieci Słabowidzących i Niewidomych w Łodzi, możliwe było także przeprowadzenie testu aplikacji także przez niewidomego ucznia trzeciej klasy gimnazjum. Testy zostały wykonane w obecności nauczyciela informatyki. Uczeń, w pierwszym etapie testów miał możliwość sprawdzenia w książce z wydrukiem wypukłym, jaki przebieg ma dana funkcja, a następnie przeprowadzał analizę tego samego wykresu z wykorzystaniem aplikacji sonifikującej obraz wyświetlany na ekranie tabletu.

Aplikacja została przyjęta pozytywnie i z zainteresowaniem. Interfejs użytkownika oceniono jako przyjazny i niepowodujący trudności w obsłudze programu. Również udźwiękowanie wykresu spełniło swoje zadanie. Zostało także zasugerowanych kilka rozwiązań, które w przyszłości mogłyby ulepszyć i rozszerzyć działanie programu, m.in. zasugerowano dodanie trybu sprawdzania wiedzy, co umożliwiłoby na samodzielną naukę. Przedstawiciele Ośrodka Szkolno-Wychowawczego dla Dzieci Słabowidzących i Niewidomych wyrazili chęć korzystania z tego typu aplikacji, która mogłaby stanowić

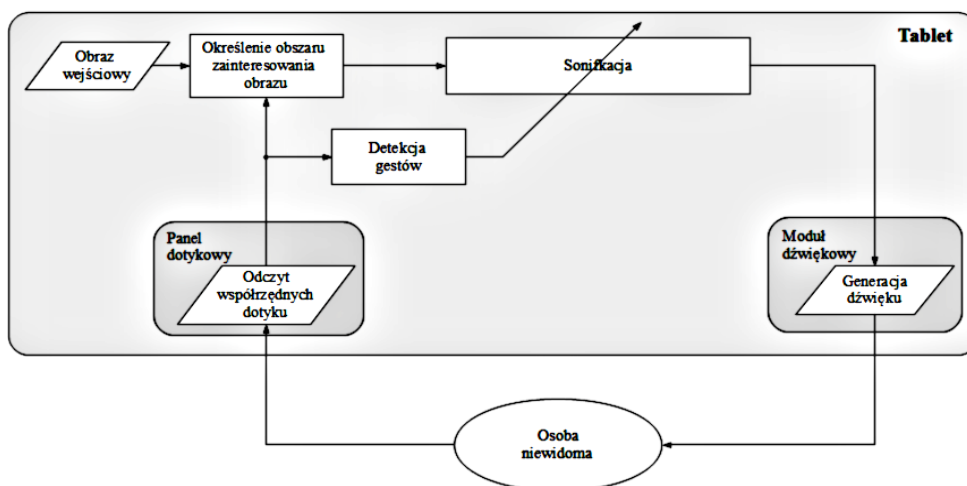
alternatywę dla dotychczasowych rozwiązań.



Rys. 3. Przykładowa droga, po jakiej użytkownik zmieniał miejsce dotyku ekranu, aby odnaleźć miejsce zerowe funkcji

Sonifikacja interaktywna

Przedstawiona w artykule koncepcja udźwiękowania obrazu stanowi punkt wyjściowy do zagadnienia sonifikacji interaktywnej, w której użytkownik otrzymuje pełną kontrolę nad tym jakie wybrane cechy obrazu, i w jakim zakresie chce udźwiękować. Kontrola ta jest realizowana w czasie rzeczywistym poprzez wykonywanie odpowiednich gestów dotykowych wskazujących tzw. obszar zainteresowania obrazu przeznaczony do sonifikacji. W tego typu sonifikacji, osoba niewidoma może być interpretowana jako sprzężenie zwrotne w procesie regulacji dźwięku, który będzie zawierał istotne informacje o analizowanym obrazie (rys. 4).



Rys. 4. Schemat blokowy procesu sonifikacji interaktywnej

Wprowadzenie mechanizmu sonifikacji interaktywnej daje możliwość analizy wyłącznie wybranych cech zadanego obszaru obrazu. Kluczowym aspektem tego mechanizmu jest umożliwienie osobie niewidomej

parametryzacji procesu sonifikacji, dzięki czemu może się ona skupić na mniej zrozumiałych aspektach analizowanych obszarów.

Podsumowanie

W artykule opisano budowę oraz testy aplikacji dla urządzenia z systemem operacyjnym Android przeznaczoną dla niewidomych i słabowidzących uczniów gimnazjum. W opracowanej wersji pilotażowej aplikacja umożliwia interaktywną sonifikację wykresów funkcji. Zaproponowana metoda udźwiękowienia została przyjęta pozytywnie przez pierwszych użytkowników. Oceniono, że aplikacja może być cenną pomocą dydaktyczną w nauczaniu wybranych działów matematyki. Na podstawie przeprowadzonych prac i testów aplikacji można stwierdzić, że bardzo ważnym elementem podczas projektowania interfejsu użytkownika jest uwzględnienie potrzeb zarówno osób niewidomych jak i osób słabowidzących. Kluczowym elementem jest zaproponowanych schemat udźwiękowienia obrazu, który nie będzie męczący dla użytkownika, a jednocześnie będzie w łatwy do przyswojenia sposób prezentował informacje graficzne.

Dalszym kierunkiem badań będzie opracowanie metod sonifikacji interaktywnej. Badania te, będą dotyczyły metod sonifikacji obrazów bardziej złożonych (m.in. map, planów i zdjęć kolorowych) dając jednocześnie użytkownikowi pełną kontrolę zarówno nad parametrami odwzorowań wizyjno-sonicznych, jak i parametrami wstępnego cyfrowego przetwarzania obrazów. Badania weryfikacyjne, rozwijanej sonifikacji interaktywnej, zostaną przeprowadzone na drodze testów metodycznych z większą grupą użytkowników.

Pracę dofinansowano ze środków projektu badawczego finansowanego przez Narodowe Centrum Nauki nr 2015/17/B/ST7/03884 realizowanego w Instytucie Elektroniki oraz w Instytucie Automatyki Politechniki Łódzkiej w latach 2016-2018. Autorzy pragną także podziękować pracownikom i uczniom Ośrodka Szkolno-Wychowawczego dla Dzieci Słabowidzących i Niewidomych w Łodzi za szereg cennych uwag i sugestii dalszych prac.

Autorzy: inż. Izabela Borowiecka, Politechnika Łódzka, Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki, Instytut Elektroniki, ul. Wólczańska 211/215, 90-924 Łódź; dr inż. Piotr Skulimowski, Politechnika Łódzka, Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki, Instytut Elektroniki, ul. Wólczańska 211/215, 90-924 Łódź, E-mail: pskul@p.lodz.pl; dr inż. Michał Bujacz, Politechnika Łódzka, Wydział Elektrotechniki,

Elektroniki, Informatyki i Automatyki, Instytut Elektroniki, ul. Wólczańska 211/215, 90-924 Łódź, dr inż. Andrzej Radecki, Politechnika Łódzka, Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki, Instytut Automatyki, ul. B. Stefanowskiego 18/22, 90-924 Łódź, prof. dr hab. inż. Paweł Strumiłło, Politechnika Łódzka, Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki, Instytut Elektroniki, ul. Wólczańska 211/215, 90-924 Łódź, E-mail: pawel.strumillo@p.lodz.pl.

LITERATURA

- [1] Hermann T., Hunt A., Neuhoff J., Eds. Sonification Handbook, Logos Publishing House, Berlin, 2011
- [2] Kramer G., Sonification of Financial Data: An Overview of Spreadsheet and Database Sonification, *The Proceedings of Virtual Reality Systems '93*, SIG Advanced Applications, New York, NY, 1993
- [3] Hermann T., Hunt A., An Introduction to Interactive Sonification, *IEEE Multimedia*, April-June, vol. 12, (2005), no. 2, 20-24
- [4] Special Issue on Interactive Sonification, *IEEE Multimedia*, vol. 12, (2005), no. 2
- [5] Krzywicka-Blum E., Kuchmister J., Modelowanie sonoryczne w edukacji niewidomych, Katedra Geodezji i Fotogrametrii Akademii Rolniczej, Wrocław, 1999
- [6] Meijer P.B.L., An Experimental System for Auditory Image Representations, *IEEE Transactions on Biomedical Engineering*, vol. 39, (1992), no. 2, 112-121
- [7] Sarkar R., Bakshi S., K'Sa P., Review on Image Sonification: A Non-visual Scene Representation, *Int. Conf. on Recent Advances in Information Technology RAIT2012*, 2012
- [8] Alonso M., Shelley S., Hermes D., Kohlrausch A., Evaluating geometrical properties of virtual shapes using interactive sonification, In: *IEEE International Workshop on Haptic Audio visual Environments and Games 2008*, 154-159, 2008
- [9] Tian Y.A., Vision Substitution Method for the Blind Based on Image Edge Detection and Sound Mapping, *Third International Conference on Computational Intelligence, Communication Systems and Networks*, 208-212, 2011
- [10] Borowiecka I., Mobilny system wspomaganie nauki matematyki dla osób słabowidzących, praca inżynierska, Instytut Elektroniki Politechniki Łódzkiej, 2015
- [11] Maidenbaum S., Buchs G., Abboud S., Lavi-Rotbain O., Amedi A., Perception of Graphical Virtual Environments by Blind Users via Sensory Substitution, *PLOS ONE*, doi:10.1371/journal.pone.0147501, 2016