

„Blind Phone” – nowy interfejs wspomagający niewidomych w obsłudze smartfonu

Streszczenie. W artykule zaprezentowano nowy interfejs wspomagający osoby niewidome i słabowidzące w obsłudze smartfonów z systemem Android. Pozwala on telefonować, wysyłać i odbierać SMS-y, korzystać z „książki telefonicznej”, a także z dodatkowych funkcji, np. pozycjonowania czy monitorowania stanu baterii, za pomocą komend głosowych. Szczegółowo omówiono budowę programu oraz działanie poszczególnych funkcji.

Abstract. This article presents a new interface supporting blind and partially sighted people in the Android smartphones use. It enables them to call, send, and receive text messages, make use of a „phone book” as well as of additional options, e.g., positioning or battery monitoring, via voice commands. The software concept and structure of this application have been discussed along with its main functionalities. („Blind Phone” – a new interface supporting the blind in smartphone use).

Słowa kluczowe: niewidomi, niedowidzący, słabowidzący, smartfon, Android.

Keywords: the blind, partially sighted, visually impaired, smartphone, Android.

Wprowadzenie

Obserwowany w ostatnich latach dynamiczny rozwój technik komunikacyjnych [1, 2, 3], informacyjnych i multimedialnych sprawił, że klasyczne telefony komórkowe zostały prawie całkowicie wyparte przez swoje nowsze odpowiedniki, tzw. „inteligentne” telefony (z jęz. ang. „smartfony”) o zupełnie innej konstrukcji. Smartfony, przeważnie z dużym ekranem dotykowym, coraz bardziej przypominają podręczny komputer, a telefonowanie niekoniecznie jest ich główną funkcją. Nawet popularne modele smartfonów oferują szerokie spektrum możliwości – od robienia zdjęć z dużą rozdzielczością, filmowania za pomocą minikamery, kończąc na aktywnym korzystaniu z internetu bądź z rozmaitych aplikacji współpracujących z różnymi czujnikami umieszczonymi w telefonie, np. z odbiornikiem sygnałów GPS, czy elektronicznym żyroskopem.

Mimo ogromnych możliwości, wbudowanego procesora i dużej mocy obliczeniowej, popularne smartfony nie odpowiadają jednak potrzebom osób niewidomych, gdyż zdolność ich obsługi przez te osoby jest znacząco ograniczona. Współczesne smartfony posiadają bowiem zaledwie parę przycisków fizycznych (które można wyczuć poprzez dotyk), zazwyczaj do regulacji głośności, blokowania ekranu, włączania i wyłączania telefonu, natomiast szczegółowe menu dostępne jest jedynie poprzez rozbudowany interfejs graficzny, prezentowany użytkownikowi na ekranie dotykowym. Rozwój telefonów komórkowych w tym kierunku okazał się więc dla osób niewidomych niekorzystny. W rezultacie, niewidomi nadal chętnie używają starszych modeli telefonów (z klawiszami), lecz nie mogą korzystać z nowoczesnych modeli, a zarazem z możliwości, jakie one oferują.

Popularne aplikacje telefoniczne

Wraz z rozwojem technologii i oprogramowania, coraz liczniej zaczęły powstawać dedykowane aplikacje na urządzenia mobilne, rozszerzające istniejące i oferujące pomocnicze funkcje, także dla osób niepełnosprawnych [4, 5]. W przypadku użytkowania smartfonów przez osoby niewidome, pomocne okazać się mogą niektóre aplikacje oraz „nakładki” na systemy operacyjne telefonów, pierwotnie zaprojektowane dla użytkowników o wzroku prawidłowym. Wśród narzędzi do głosowego wybierania numerów wymienić można: „Wybieranie Głosowe”, „Voice Dialer” czy „Speak 2 Call Free”, jednak jedyną z powyższych aplikacji, umożliwiającą zatelefonowanie do

osoby ze spisu kontaktów, jest „Voice Dialer” [6]. Pozostałe programy pozwalają tylko na głosowe wyszukanie kontaktu, natomiast dzwonienie następuje po kliknięciu bądź przesunięciu palcem danego kontaktu, co wyklucza możliwość posługiwania się nimi przez osoby z całkowitą utratą wzroku. Niestety „Voice Dialer” dzwoni zawsze pod pierwszy znaleziony kontakt pasujący do wypowiedzianej frazy, co często prowadzi do błędów w wyborze rozmówcy.

Istnieją też bardziej zaawansowane „nakładki” na systemy operacyjne smartfonów, np. „Google Talk Back” [7] dla systemu Android. Przekształca on poszczególne elementy ekranu w komunikaty słowne i w ten sposób pozwala kontrolować telefon bez konieczności patrzenia na ekran. Z powodu dużej ilości przekazywanych informacji, korzystanie z tej „nakładki” nie jest wcale wygodne. Ponadto nie ułatwia ona dostępu do konkretnych funkcji telefonu, a jedynie zmienia formę ich prezentacji. Popularnym programem do głosowej obsługi wybranych funkcji smartfonu jest również „Google Now” [8]. Nie jest on jednak dedykowany osobom niewidomym, ponieważ nie wszystkie operacje kończą się komunikatem słownym. Do zdecydowanie mniej popularnych (przynajmniej w Polsce) należą aplikacje „Ray” [9] oraz „Georgie” [10]. W obu rozwiązaniach obsługa smartfonu odbywa się poprzez naciśnięcia ekranu i przesuwanie palca w odpowiednie strony, a także gesty. Jak się okazuje, nie mały problem stanowi jednak wdrożenie języka polskiego. Z polskich aplikacji na uwagę zasługuje „Dziobaq” [11], który uruchamia niektóre funkcje telefonu za pomocą poleceń głosowych, ale wymaga sporego zaangażowania ze strony użytkownika.

Przedstawiony stan rzeczy uzasadnia potrzebę opracowania aplikacji dla aktualnie produkowanych smartfonów niższej i średniej klasy, z popularnym systemem operacyjnym, prymarnie zaprojektowanej oraz zoptymalizowanej z uwzględnieniem możliwości osób niewidomych i słabowidzących.

Założenia dla nowej aplikacji

Uwzględniając charakter niepełnosprawności osób niewidomych przyjęto, że najlepszym sposobem komunikacji niewidomego z urządzeniem (i jego obsługi) są komendy głosowe. Powinny one być jednoznaczne i zarazem krótkie, a w procesie rozpoznawania mowy bezbłędnie rozróżniane. Dla osób słabowidzących dopuszczalny jest jednak pomocniczy interfejs graficzny, oczywiście dostosowany do dysfunkcji użytkowników. Musi być on zatem bardzo prosty,

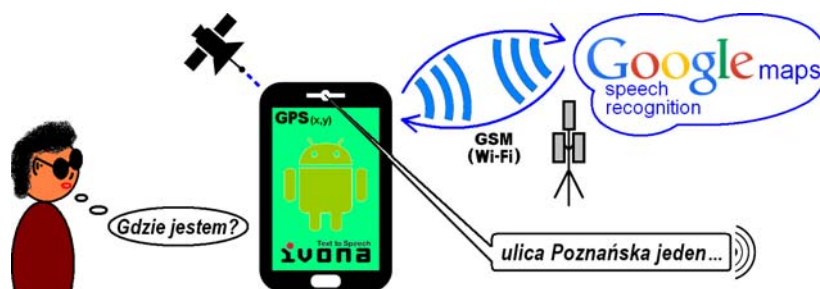
z symbolami bądź znakami w odpowiednio dużym rozmiarze, i o właściwie dobranej kolorystyce [12]. Z kolei efekty działania danej funkcji, względnie jej przebieg, winny być przekazywane zwrotnie również w formie komunikatów słownych, możliwie głośno (z opcją regulacji), ewentualnie poprzez słuchawkę douszną.

Ogromna popularność, wsparcie w postaci dedykowanego środowiska programistycznego oraz łatwy dostęp do dokumentacji przesądziły o wyborze systemu Android jako platformy docelowej. Aktualnie ponad połowa urządzeń mobilnych na świecie jest oparta na tym systemie, m.in. smartwatche, smartfony i tablety, a ponadto nowoczesne odbiorniki telewizyjne. Istotną zaletą systemu Android jest też możliwość korzystania z wielu różnorodnych serwisów internetowych, przede wszystkim firmy Google.

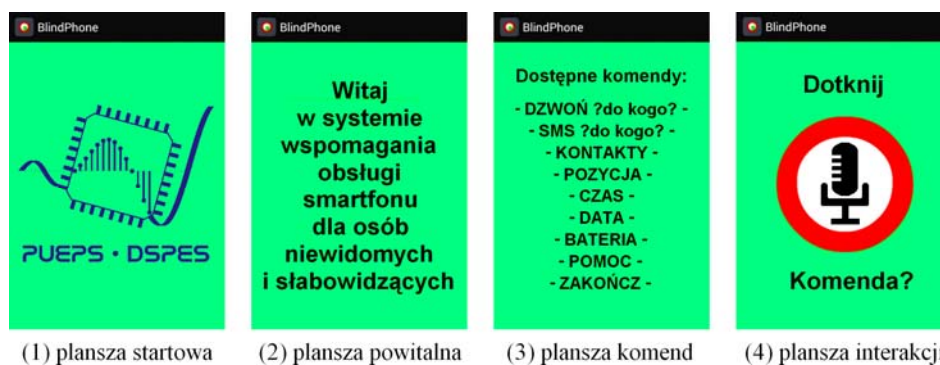
Zasada działania interfejsu „Blind Phone”

Koncepcję funkcjonowania interfejsu „Blind Phone” zilustrowano na rysunku 1. Aby uzyskać pożądaną

informację albo efekt działania konkretnej funkcji smartfonu, przepływ sygnałów pomiędzy elementami systemu może być wielotorowy, a także wielokierunkowy. Najpierw, w razie potrzeby, niewidomy wypowiada komendę głosową z listy dostępnych komend. Dzięki dużej czułości mikrofonów elektretowych nie ma konieczności zbliżania telefonu bezpośrednio do ust – może się on znajdować nawet w odległości jednego metra, jakkolwiek w takim przypadku niekorzystną rolę mogą odegrać dźwięki otoczenia. Alternatywnie można korzystać z zestawu słuchawkowego (w tym z bezprzewodowego) z mikrofonem, np. w standardzie Bluetooth [13], ale osłabi to zdolność orientacji przestrzennej osoby niewidomej, dla której słuch, w zastępstwie wzroku, staje się głównym zmysłem pozyskiwania informacji o najbliższym otoczeniu [14]. Wybór sposobu komunikacji z urządzeniem powinien pozostać w gestii samego użytkownika i zależeć wyłącznie od jego preferencji.



Rys. 1. Koncepcja działania interfejsu „Blind Phone”



Rys. 2. Plansze pomocniczego interfejsu graficznego, dedykowanego osobom słabowidzącym

Za pośrednictwem sieci GSM (względnie Wi-Fi), wypowiedziana komenda głosowa poddawana jest procedurze rozpoznawania mowy w serwisie Google Speech Recognition [15]. Moduł rozpoznawania mowy zwraca łańcuch znaków, odpowiadający rozpoznanyemu słowu, który powraca do smartfonu tą samą drogą. Następnie, program aplikacji sprawdza, czy otrzymane wyrazy składają się na jedną z dostępnych komend. Jeśli nie, użytkownik usłyszy komunikat „nie ma takiej komendy”, natomiast w przypadku odnalezienia komendy w takim brzmieniu, niezwłocznie wykonywana jest przypisana do niej funkcja.

W przykładzie z rysunku 1 obieg danych trwa dłużej, bowiem współrzędne geograficzne wyznaczone w telefonie na podstawie odebranych sygnałów GPS, są przesyłane jeszcze siecią GSM (lub Wi-Fi) do serwisu Google Maps [16], gdzie następuje tzw. odwrotne geokodowanie, tj. rzutowanie współrzędnych na mapę cyfrową i odnalezienie odpowiadającego im adresu. Znaleziony adres, w postaci

ciągu znaków, wraca do smartfonu i jest w nim odczytywany na głos przez syntezytor mowy Ivona [17]. Wyniki działania innych funkcji telefonu inicjowanych z poziomu aplikacji, jak również wszelkie inne interakcje z użytkownikiem, za każdym razem wieńczy komunikat słowny, odczytywany głośno przez syntezytor. Natężenie dźwięku można regulować standardowymi przyciskami głośności w telefonie, a w przypadku korzystania z zestawu słuchawkowego – także dodatkowymi przyciskami, o ile takie występują.

Pomocniczy interfejs graficzny

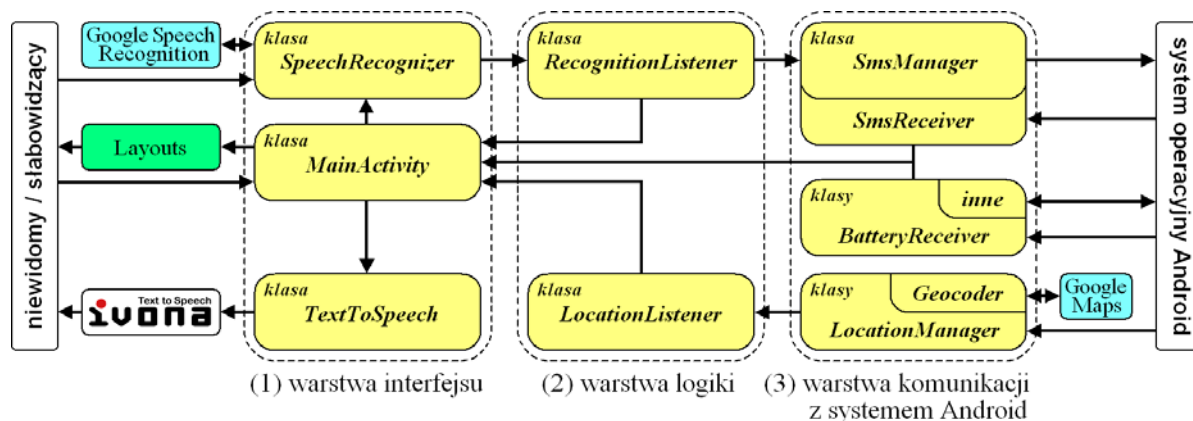
Z myślą o osobach słabowidzących przygotowano pomocniczy interfejs graficzny, wyświetlany na ekranie smartfonu. Jego forma została celowo uproszczona, aby niepotrzebnie nie utrudniać percepcji. Przewidziano cztery plansze (rys. 2): (1) uruchomieniową, z logo Pracowni Układów Elektronicznych i Przetwarzania Sygnałów, w skrócie

PUEPS (ang. DSPES), w której powstała aplikacja, (2) planszę powitalną, (3) planszę z listą najważniejszych komend, i (4) interakcyjną. Odpowiadają one fazom działania aplikacji i są z nimi zsynchronizowane.

Wybrano silnie kontrastujące kolory: jasnozielone tło i czarną, pogrubioną czcionkę [12]. Mimo, iż pożądane jest maksymalne podświetlenie ekranu, to po upływie czasu określonego w ustawieniach telefonu, docelowa plansza interakcji (4) przygasa aż do momentu ponownego dotknięcia ekranu. Zabieg ten ma na celu zmniejszenie poboru prądu i tym samym wydłużenie czasu pracy smartfonu.

Budowa programu

W programie aplikacji wyróżnić można trzy warstwy (rys. 3): (1) warstwę interfejsu, (2) logiki oraz (3) warstwę komunikacji z systemem Android. Warstwę interfejsu (1) tworzą klasy odpowiedzialne za interakcje z użytkownikiem. Warstwa logiczna (2) obejmuje klasy odpowiadające za implementację poszczególnych funkcji, a warstwa komunikacji z systemem Android (3) zawiera klasy umożliwiające pozyskiwanie i wymianę danych między aplikacją a systemem operacyjnym smartfonu.



Rys. 3. Przepływ informacji między warstwami i głównymi klasami

W trakcie uruchamiania programu inicjowane są cztery główne procesy wykonywane niezależnie. Obok procesu obsługującego interfejs użytkownika, realizowanego przez klasę aktywności *MainActivity*, równoległe toczą się jeszcze dwa procesy nasłuchiwanie powiadomień o nowych wiadomościach tekstowych (klasa *SmsReceiver*) i o zmianie poziomu naładowania akumulatora (*BatteryReceiver*), a ponadto proces związany z uzyskiwaniem współrzędnych geograficznych (funkcja klasy *LocationManager* i *LocationListener*).

Klasa *SpeechRecognizer* zapewnia dostęp do modułu rozpoznawania mowy Google [15] i zarazem odpowiada za transfer danych pomiędzy aplikacją a serwerem. Z kolei klasa *RecognitionListener* umożliwia śledzenie procesów zachodzących w klasie *SpeechRecognizer* i uzyskanie od niej wektora odpowiedzi lub ewentualnie numeru błędu. Uwzględniono łącznie kilkanaście rodzajów błędów, w tym brak lub przerwanie połączenia internetowego, brak odpowiedzi po stronie serwera, czy też brak wyników.

Za pośrednictwem klasy *TextToSpeech*, do syntezatora mowy Ivona [17] są przekazywane ciągi znaków kwalifikowane do odczytu. Klasa ta pozwala ponadto na monitorowanie procesu syntezy i, dzięki temu, na wstrzymanie wywołań kolejnych funkcji do czasu wybrzmienia bieżących treści. Konieczne okazało się również odpowiednie kolejowanie komunikatów, gdyż autonomiczne procesy aplikacji nieustannie nasłuchują powiadomień o nowych zdarzeniach (np. o nadejściu SMS-a), a natychmiastowe komunikaty głosowe mogłyby zakłócić inne procesy, w szczególności rozpoznawanie mowy.

Zaimplementowane funkcje

W prezentowanej aplikacji utworzono niezależny, dedykowany „spis kontaktów” (z numerami telefonów), znacznie efektywniejszy i szybszy w obsłudze niż systemowa „książka telefoniczna” smartfonu. Poszczególne pozycje są zapisywane w pliku tekstowym według

konwencji *nazwa_1, tel_1; nazwa_2, tel_2* itd., którego zawartość jest wczytywana do pamięci operacyjnej smartfonu przy każdym uruchomieniu programu. Od tego momentu można ją wygodnie przeszukiwać oraz edytować za pomocą komend głosowych. Oczywiście można wprowadzać nowe kontakty poprzez komendy „dodaj” lub „nowy”, bądź też likwidować istniejące wpisy komendą „usuń”. Procedury te kończą się nadpisaniem pliku z danymi. Możliwy jest także odsłuch wszystkich rekordów dzięki komendzie „kontakty”.

Komendy służące telefonowaniu są generowane dynamicznie na podstawie zawartości „listy kontaktów”, którą samodzielnie tworzy użytkownik aplikacji. Składają się one ze zbitki dwóch słów: „dzwoń + nazwa_X” (np. „dzwoń Ala”). Jeżeli wypowiedziane imię lub nazwisko, po procesie rozpoznawania mowy, zostanie odnalezione w „spisie kontaktów”, wówczas bezzwłocznie nastąpi wybranie właściwego numeru i usłyszymy komunikat „dzwońię”. Jeśli nazwa nie zostanie znaleziona, program powiadomi o tym odpowiednim komunikatem.

Do wysyłania wiadomości tekstowych (SMS-ów) służy komenda „esemes + nazwa_Y” (np. „esemes Adam”). Po pomyślnym rozpoznaniu i odnalezieniu danego adresata w „spisie kontaktów”, użytkownik jest proszony o wypowiedzenie treści SMS-a, po czym, po procedurze rozpoznania mowy, następuje kontrolny odczyt tejże wiadomości przez syntezator i możliwość ostatecznego zatwierdzenia bądź anulowania SMS-a. W przyjętym rozwiązaniu, dla zagwarantowania poprawnego odczytu przez syntezator, dopuszczono transfer polskich znaków diakrytycznych, choć skraca to długość SMS-ów.

W przypadku nadejścia SMS-a, po jego udanym zdekodowaniu, tj. wyodrębnieniu treści oraz numeru telefonu nadawcy, treść trafia najpierw do bufora wiadomości, a numer jest poszukiwany w „spisie kontaktów”. Przy braku aktywności, program powiadamia o otrzymanej wiadomości i podaje nazwę nadawcy, jeśli został on

zidentyfikowany. W celu wysłuchania wiadomości należy wypowiedzieć komendę „czytaj”.

Dla uzyskania informacji na temat aktualnego położenia (adresu), można skorzystać z jednej z synonimicznych komend: „pozycja”, „położenie”, „lokalizacja”, „adres”, „miejsce” albo zadać pytanie „gdzie (ja) jestem?”. Procedura pozycjonowania została już omówiona w punkcie 4, w akapicie opisującym rysunek 1. Warto jednak dodać, iż w przypadku zaniku sygnałów GPS, ustalenie przybliżonej pozycji odbywa się z wykorzystaniem sieci komórkowej GSM (metodą triangulacji) lub też na podstawie dostępnej sieci Wi-Fi. Ponieważ są to metody niedokładne, użytkownik zostanie uprzedzony o tym fakcie adekwatnym komunikatem.

Wypowiadając komendy: „data”, „dzień”, „czas”, „godzina”, „która (jest) godzina?”, usłyszymy natomiast precyzyjną odpowiedź, ponieważ dane te są pobierane bezpośrednio z systemu operacyjnego telefonu. Równie dokładna będzie informacja o poziomie naładowania baterii, podawana w procentach na komendę „bateria”, „stan baterii” albo „poziom baterii”. Poza kontrolą na żądanie, równoległe (w trybie ciągłym) trwa proces monitorowania stanu baterii, i w razie takiej potrzeby, zabrzmi komunikat o konieczności podłączenia ładowarki.

Wymagania sprzętowe

Smartfon współpracujący z aplikacją powinien posiadać ekran dotykowy o przekątnej nie mniejszej niż 3,27”, wbudowany odbiornik GPS oraz system operacyjny Android w wersji przynajmniej 4.0 („Ice Cream Sandwich”). Warto, aby cechował się dobrą jakością dźwięku, w tym mikrofonu, od którego zależy interakcja z niewidomym użytkownikiem, jak również dużym kontrastem i jasnością ekranu, w przypadku osoby słabowidzącej. Istotna jest też konstrukcja i kształt przycisków fizycznych, a także pojemność i żywotność akumulatora.

Po wgraniu pliku BlindPhone.apk do telefonu i zainstalowaniu aplikacji, należy jeszcze doinstalować preferowany syntezytor mowy (w razie jego braku) i ustawić go jako domyślny. Ponadto, trzeba włączyć transmisję danych, lokalizację za pomocą systemu GPS, sieci GSM oraz Wi-Fi, i zezwolić na usługę lokalizacji. Dla osoby słabowidzącej należy ustawić maksymalną jasność ekranu.

Podsumowanie

Z powodzeniem opracowano dedykowany interfejs, wspomagający osoby niewidome oraz słabowidzące w obsłudze nowoczesnego telefonu (smartfonu) z systemem Android. Za pomocą predefiniowanych komend głosowych można wykonać wiele różnych czynności, m.in. zatelefonować, wysłać i odbierać SMS-y, swobodnie korzystać z „książki telefonicznej”, a ponadto ustalić swoje położenie, uzyskać informację o aktualnej godzinie i kontrolować stan baterii, bez specjalnego zaangażowania. W niedługim czasie planuje się przeprowadzenie testów z udziałem uczniów z Ośrodka Szkolno-Wychowawczego dla Dzieci Niewidomych im. „Synów Pułku” w Owińskach koło Poznania.

Praca zrealizowana w ramach projektu PUEPS DS-2016.

Autorzy: dr inż. Piotr Kardys, prof. dr hab. inż. Adam Dąbrowski, Politechnika Poznańska, Katedra Sterowania i Inżynierii Systemów, Pracownia Układów Elektronicznych i Przetwarzania Sygnałów, ul. Jana Pawła II 24, 60-965 Poznań, E-mail: adam.dabrowski@put.poznan.pl; inż. Damian Huderek, inż. Marcin Iwanowski, Politechnika Poznańska, Wydział Informatyki.

LITERATURA

- [1] Kardys P., Dąbrowski A., Standard Bluetooth na tle bezprzewodowych systemów telekomunikacyjnych krótkiego zasięgu w aplikacjach bio- i telemedycznych, *Elektronika* 52 (5), 5/2011, 56-59
- [2] Dąbrowski A., Kardys P., System transmisji bezprzewodowej Bluetooth, Krajowa Konferencja Radio-komunikacji, Radiofonii i Telewizji, Poznań 2001, 0.4.1-0.4.8.
- [3] Kardys P., Dąbrowski A., Bluetooth--nowe możliwości nowoczesnych urządzeń elektronicznych (3), *Elektronizacja* (6), 6/2001, 14-15
- [4] Dąbrowski A., Kardys P., Marciniak T., Bluetooth technology applications dedicated to supporting blind and hearing as well as speech handicapped people, The 47th IEEE ELMAR Symp., Zadar 2005, 295-298
- [5] Dąbrowski A., Kardys P., Marciniak T., Technical aid to handicapped people by means of the Bluetooth wireless communication system, The 11th Symp. „The Art of Sound Engineering”, Kraków 2005, 70-78
- [6] Ruthruff A., How to use Voice Dialer on an Android phone, Groovy Post LLC, Port Orchard 2010
- [7] Hildenbrand J., What is Google TalkBack, AndroidCentral.com, 2014
- [8] Guha R., Gupta V., Raghunathan V., Srikant R., User modeling for a personal assistant, The 8th WSDM International Conference, Shanghai 2015, 275-284
- [9] Katz L., Ray turns Android phone into device for the blind, Cnet.com, 2012
- [10] Trenholm R., Georgie, first smartphone for the blind, Cnet.com, 2012
- [11] Masarczyk P., Obsługa telefonu bez dotykania, Tyfloświat, Fundacja IRR, Kraków 2015
- [12] Dziamska D., Portalska H., Portalski M., W harmonii z barwą, Wydawnictwo BP, Poznań 2004
- [13] Kardys P., Dąbrowski A., Przetwarzanie sygnałów audio i mowy w systemie Bluetooth, Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji, Poznań 2001, 16.2.1-16.2.4
- [14] Bujacz M., Skulimowski P., Strumiłło P., Sonification of 3D scenes using personalized spatial audio to aid visually impaired persons, The 17th ICAD International Conference, Budapeszt 2011
- [15] Aleksic P., Ghodsi M., Michaely A., Allauzen C., Hall K., Roark B., Rybach D., Moreno P., Bringing contextual information to Google speech recognition, The 16th Interspeech Conference, Drezno 2015, 468-472
- [16] Madrigal A., How Google builds its maps, Technology, TheAtlantic.com, 9/2012
- [17] Grabowski J., Korzystamy z syntezytoru mowy Ivona, *Komputer Świat* (dwutygodnik), lipiec 2009