



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

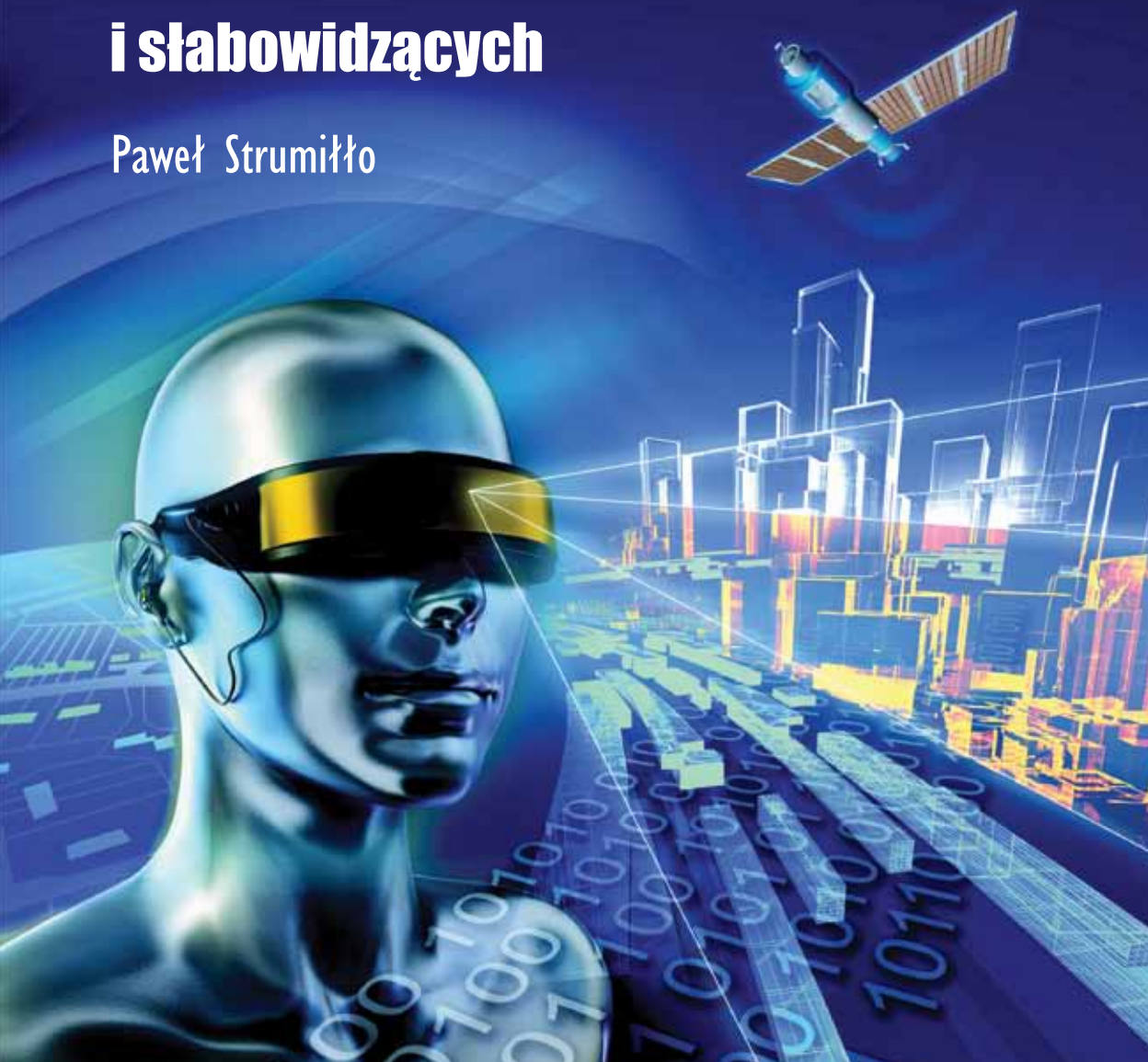


CZŁOWIEK – NAJLEPSZA INWESTYCJA

Książka jest współfinansowana przez Unię Europejską w ramach Europejskiego Funduszu Społecznego

Elektroniczne systemy nawigacji osobistej dla niewidomych i słabowidzących

Paweł Strumiłło



Elektroniczne systemy nawigacji osobistej dla niewidomych i słabowidzących

Paweł Strumiłło

pod patronatem
Komitetu Elektroniki i Telekomunikacji
Polskiej Akademii Nauk

Politechnika Łódzka
Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki

Łódź 2012



Politechnika Łódzka

Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki

Copyright by Politechnika Łódzka ©

Copyright by Paweł Strumiłło ©

Łódź 2012

Recenzenci:

prof. dr hab. inż. Jerzy Lewandowski

prof. dr hab. inż. Andrzej Materka

Redaktor:

Michał Strzelecki

Projekt okładki i rysunki:

Anna Janisz

Skład i łamanie:

Monika Popielata

Druk:

CMYK Studio sp. z o.o.

ul. H. Sienkiewicza 44, 26-300 Opoczno

cmyk@cmykstudio.pl

Wydawca:

Politechnika Łódzka

Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki

ISBN 978-83-933359-0-9

Książkę wydrukowano i wydano ze środków projektu „Innowacyjna dydaktyka bez ograniczeń – zintegrowany rozwój Politechniki Łódzkiej – zarządzanie Uczelnią, nowoczesna oferta edukacyjna i wzmacniania zdolności do zatrudnienia, także osób niepełnosprawnych” współfinansowanego przez Unię Europejską w ramach Europejskiego Funduszu Społecznego.

Spis treści

Wykaz ważniejszych skrótów	7
Wstęp	11
 1. Wzrok, jego niepełnosprawności oraz konsekwencje	
utraty wzroku	17
1.1. System wzrokowy człowieka	17
1.2. Badanie, ocena sprawności i wady wzroku	22
1.3. Konsekwencje utraty wzroku	27
1.4. Orientacja przestrzenna i poruszanie się osób niewidomych	28
 2. Techniki estymacji głębi i modelowanie sceny trójwymiarowej	41
2.1. Aktywne i pasywne techniki estymacji głębi	43
2.2. Wybrane zagadnienia stereowizji	55
2.3. Modelowanie sceny trójwymiarowej	68
 3. Elektroniczne techniki nawigacji	77
3.1. Rodzaje technik nawigacji	78
3.2. Globalne systemy pozycjonowania	81
3.3. Zastosowanie map elektronicznych w nawigacji pieszej	85

3.4. Wyznaczanie położenia z zastosowaniem nadajników naziemnych	87
3.5. Nawigacja zliczeniowa	90
3.5.1. Nawigacja inercyjna	91
3.5.2. Nawigacja wizyjna	93
3.5.3. Techniki integracji danych nawigacyjnych	100
3.5.4. Zastosowanie filtracji Kalmana i filtracji cząsteczkowej w nawigacji pieszej	102
4. Niewizualne techniki obrazowania otoczenia	117
4.1. Porównanie własności sensorycznych zmysłów wzroku, słuchu i dotyku	118
4.2. Systemy substytucji sensorycznej	126
4.2.1. Dźwiękowe techniki obrazowania	132
4.2.1.1. Techniki sonifikacji	133
4.2.1.2. Modele słyszenia przestrzennego i ich identyfikacja	135
4.2.1.3. Wytwarzanie dźwięków przestrzennych	142
4.2.2. Dotykowe i haptyczne techniki obrazowania	147
5. Wybrane elektroniczne systemy nawigacji osobistej dla niewidomych	155
5.1. Elektroniczne detektory przeszkód	157
5.2. Elektroniczne systemy niewizualnej prezentacji obrazów otoczenia	160
5.3. Systemy nawigacji osobistej o architekturze rozproszonej	177
5.4. Tele-nawigacja przy pomocy zdalnego przewodnika	185
Podsumowanie	195
Literatura	199

Słowo wstępne

Często można usłyszeć opinię, że systemy wspomagające osoby starsze i niepełnosprawne będą w najbliższych latach jednym z ważniejszych obszarów rozwoju techniki. Osobiście podzielam taki pogląd. Jest on słuszny zwłaszcza w odniesieniu do systemów wykorzystujących zaawansowane techniki przetwarzania i przesyłania informacji, w tym metody przetwarzania obrazów i nowoczesne interfejsy pomiędzy człowiekiem a komputerem. Postęp w budowie mikrokomputerów oraz urządzeń zbierających informację o głębi stereoskopowej umożliwia konstrukcję bardzo lekkich, zasilanych z baterii systemów, mogących na bieżąco prowadzić złożoną analizę sceny przestrzennej i generować odpowiednie sygnały dźwiękowe lub dotykowe. Te osiągnięcia niosą milionom ludzi o niesprawnym wzroku nadzieję na poprawę jakości życia, a badaczom i inżynierom zajmującym się tą szlachetną działalnością pozwalają marzyć o urządzeniach jeszcze doskonalszych.

W świetle powyższych uwag łatwo zrozumieć, że z prawdziwą radością przedstawiam monografię dotyczącą zaawansowanych systemów osobistej nawigacji ułatwiających poruszanie się osobom niewidomym i słabowidzącym. Przedstawiona monografia jest nie tylko cenna z powodu ważnej tematyki, ale jest godna poleceniu ze względu na nowoczesne ujęcie tematu, uwzględniające najnowsze konstrukcje urządzeń i ostatnie osiągnięcia naukowe.

Życzę Autorowi monografii, by znalazła ona licznych Czytelników zarówno w środowiskach naukowych i technicznych, jak i medycznych oraz rehabilitacyjnych, i by stała się stymulatorem postępu w tej niezwykle przydatnej dla społeczeństwa dziedzinie badań naukowych.

Marek Domański
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki
Politechniki Poznańskiej

Wykaz ważniejszych skrótów

AGPS	– usługa skracająca czas wyznaczania położenia geograficznego w globalnym systemie pozycjonowania (ang. <i>assisted Global Positioning System</i>) *
AMD	– zwyrodnienie plamki żółtej (<i>age-related macular degeneration</i>)
AOA	– kąt dotarcia (<i>angle of arrival</i>)
AR	– rzeczywistość wzbogacona (<i>augmented reality</i>)
<i>b-frame</i>	– układ współrzędnych związany z obserwatorem (<i>body frame</i>)
CCD	– układ elektroniczny ze sprzężeniem ładunkowym (<i>charge coupled device</i>)
CMOS	– para komplementarna tranzystorów polowych z izolowaną bramką (<i>complementary metal oxide semiconductor</i>)
CUDA	– uniwersalna architektura procesorów wielordzeniowych (<i>compute unified device architecture</i>)
DGPS	– różnicowy globalny system pozycjonowania (<i>differential Global Positioning System</i>)
DAISY	– standard cyfrowego zapisu informacji (<i>digital accessible information system</i>)
DoF	– stopnie swobody (<i>degrees of freedom</i>)
DSP	– cyfrowy procesor sygnałowy (<i>digital signal processor</i>)

* W wykazie skrótów i w tekście książki wszystkie nazwy obcojęzyczne są w języku angielskim; w wykazie skrótów ich tłumaczenia polskie nie zawsze są dosłowne

EKF	– rozszerzony filtr Kalmana (<i>extended Kalman filter</i>)
ETA	– elektroniczny system wspomagający poruszanie się (<i>electronic travel aid</i>)
fMRI	– czynnościowe obrazowanie rezonansu magnetycznego (<i>functional magnetic resonance imaging</i>)
FPGA	– programowalna macierz bramek logicznych (<i>field programmable gate arrays</i>)
<i>fps</i>	– ramki obrazu na sekundę (<i>frames per second</i>)
GPRS	– usługa pakietowej transmisji danych (<i>general packet radio service</i>)
GPS	– globalny system pozycjonowania (<i>Global Positioning System</i>)
GPU	– procesor graficzny (<i>graphics processor unit</i>)
GUI	– graficzny interfejs użytkownika (<i>graphical user interface</i>)
HCI	– interakcja człowiek-komputer (<i>human computer interaction</i>)
HDOP	– „rozmycie” dokładności położenia w kierunku poziomym (<i>horizontal positioning dilution of precision</i>)
HRIR	– akustyczna odpowiedź impulsowa odniesiona do głowy (<i>head related impulse response</i>)
HRTF	– widmowa charakterystyka akustyczna odniesiona do głowy (<i>head related transfer function</i>)
HSDPA	– szybka transmisja z komutacją pakietów (<i>high speed downlink packet access</i>)
ICD	– Międzynarodowa Klasyfikacja Chorób (<i>International Classification of Diseases</i>)
ICF	– Międzynarodowa Klasyfikacja Funkcjonowania, Niepełnosprawności i Zdrowia (<i>International Classification of Functioning, Disability and Health</i>)
ILD	– międzyuszna różnica poziomu (<i>interaural level difference</i>)
ITD	– międzyuszna różnica czasu (<i>interaural time difference</i>)
KF	– filtr Kalmana (<i>Kalman filter</i>)
LTE	– standard telefonii komórkowej IV generacji (<i>long term evolution</i>)
MEMS	– mikrosystemy elektromechaniczne (<i>micro electro-mechanical systems</i>)
MIDI	– standard wymiany informacji pomiędzy elektronicznymi instrumentami muzycznymi i systemami komputerowymi (<i>Musical Instrument Digital Interface</i>)

<i>n-frame</i>	– nawigacyjny układ współrzędnych (<i>navigation frame</i>)
NMEA	– Narodowe Stowarzyszenie Elektroniki Morskiej (<i>National Marine Electronics Association</i>)
PCA	– analiza składowych głównych (<i>principal component analysis</i>)
PDOP	– rozmycie dokładności pozycjonowania (<i>positioning dilution of precision</i>)
PET	– tomografia z emisją pozytronów (<i>positron emission tomography</i>)
POI	– miejsce użyteczne lub punkt zainteresowania (<i>point of interest</i>)
RFID	– układ identyfikacji radiowej (<i>radio frequency identification device</i>)
RP	– retinopatia barwnikowa (<i>retinitis pigmentosa</i>)
RSS	– odebrany poziom sygnału (<i>received signal strength</i>)
RTOF	– czas powrotu wiązki wysłanej i odbitej (<i>return time of flight</i>)
SAD	– suma modułów różnic (<i>sum of absolute differences</i>)
TC/IP	– internetowy protokół komunikacyjny (<i>Transmission Control Protocol/Internet Protocol</i>)
TDOA	– różnica czasu dotarcia (<i>time difference of arrival</i>)
TOA	– czas dotarcia (<i>time of arrival</i>)
TOF	– czas przelotu (<i>time of flight</i>)
TTF	– okres czasu do pierwszego pozycjonowania (<i>time to first fix</i>)
UMTS	– uniwersalny system telekomunikacji ruchomej (<i>Universal Mobile Telecommunications System</i>)
UTC	– uniwersalny czas koordynowany (<i>universal time clock</i>)
VDOP	– „rozmycie” dokładności pozycjonowania w kierunku pionowym (<i>vertical positioning dilution of precision</i>)
WHO	– Światowa Organizacja Zdrowia (<i>World Health Organization</i>)
WiMAX	– technologia radiowej transmisji danych (<i>Worldwide Interoperability for Microwave Access</i>)
WLAN	– bezprzewodowa sieć lokalna (<i>wireless local area network</i>)

Kształtowane przez człowieka środowisko życia, na które składają się architektura miast, środki lokomocji, metody nauczania i pracy, urządzenia i narzędzia były zaprojektowane dla cywilizacji sprawnych ludzi. Stan taki utrwał się przez wieki, pomimo tego, że niepełnosprawność była i jest nieodłącznym elementem funkcjonowania społeczeństw. Od czasów starożytnych, przez średniowiecze i wieki późniejsze niepełnosprawność była traktowana jako niechciana odmienność, którą na różne sposoby wyłączano poza społeczność. W Sparcie okrutny los spotykał ułomne fizycznie noworodki. Inkowie byli przekonani, że kalekie dzieci są karą bogów. Podobnie było w średniowiecznej Europie aż do końca XVIII wieku. Dopiero okres oświecenia, rewolucji społecznych i przemysłowych zaowocował zmianami w nastawieniu do osób o ograniczonej sprawności. Zaczęto na gruncie medycyny wyjaśniać przyczyny różnych niepełnosprawności, podejmować próby leczenia i budować pomoce rehabilitacyjne. Pomimo tych zmian niepełnosprawność była nadal traktowana jako problem jednostki napotykającej wiele barier wykluczających ją z normalnego życia społecznego i zawodowego. O prawdziwym przełomie w podejściu do problemu niepełnosprawności można mówić dopiero w drugiej połowie XX wieku (Krawczyk-Pasławska, 2010).

Lata 1982–1993 zostały ogłoszone przez ONZ dekadą osób niepełnosprawnych. W roku 2001 Światowa Organizacja Zdrowia (*World Health Organization – WHO*) opublikowała Międzynarodową Klasyfikację Funkcjonowania, Niepełnosprawności i Zdrowia (*International Classification of Functioning, Disability and Health – ICF*). Kluczowym stwierdzeniem zawartym w tym dokumencie było rozszerzenie pojęcia niepełnosprawności tradycyjnie pojmowanej jako problem

indywidualny jednostki o tzw. model społeczny niepełnosprawności¹. Zmianę tego modelu celnie określił Woźniak (2009) jako przejście od definicji związanej z tzw. pierwotnym ograniczeniem sprawności (*Co kto ma?*) do szerszej definicji obejmującej również tzw. wtórne ograniczenie sprawności (*Co z tego wynika?*).

W modelu społecznym niepełnosprawności wskazuje się na bariery po stronie środowiska życia osób niepełnosprawnych, m.in. bariery architektoniczne, edukacyjne, zawodowe (Lewandowski i Boczkowska, 2008), prawne i społeczne (przejawiające się postawą innych wobec osób niepełnosprawnych). Model społeczny niepełnosprawności jest popierany przez inicjatywy i dekrety Unii Europejskiej oraz Europejskie Forum Niepełnosprawności, którego misją jest tworzenie Europy bez barier (www.edf-feph.org) (Kabsh, 2009).

Niepełnosprawność wzroku w liczbach

Według raportów Światowej Organizacji Zdrowia (WHO, 2011) na świecie żyje prawie 300 mln osób z niepełnosprawnością wzroku w tym ok. 39 mln osób niewidomych. Blisko 90% osób z niepełnosprawnością wzroku zamieszkuje kraje rozwijające się. W Unii Europejskiej, na każdych 1000 mieszkańców cztery osoby to niewidomi lub słabowidzący. Szacuje się, że w roku 2030 co czwarty Europejczyk będzie w wieku emerytalnym, a wśród osób starszych liczba niepełnosprawnych wzrokowo potraja się z każdym dziesięcioleciem ich życia (Helal i inni, 2008). Wyniki ankiet przeprowadzonych wśród mieszkańców Unii Europejskiej pokazują, że 76% respondentów uważa osoby niewidome za grupę społeczną mającą największe trudności w dostępie do infrastruktury miasta i usług. W następnej kolejności ankietowani wskazali na niepełnosprawnych ruchowo i osoby niesłyszące (Wapiennik i Piotrowicz, 2002).

W Polsce, wg danych Polskiego Związku Niewidomych, liczba niepełnosprawnych z dysfunkcją wzroku wynosi ok. 80 tys. Potwierdzają się dane o trudnościach jakie napotkają niewidomi w pokonywaniu barier społecznych i zawodowych. Tylko bardzo bogate społeczeństwa stać na finansowanie systemów opieki zmniejszających te bariery. W Szwecji 5,5% niewidomych to bezrobotni, podczas gdy w Polsce wskaźnik ten wynosi 87% (Balcerzak-Paradowska, 2002).

¹ Terminu niepełnosprawność (*disability*) zaczęto używać dopiero w drugiej połowie XX wieku. Dzisiaj są stosowane różne definicje niepełnosprawności. W ustawie RP nt. rehabilitacji zawodowej i społecznej osób niepełnosprawnych zastosowano definicję: *Niepełnosprawnymi są osoby, których stan fizyczny, psychiczny lub umysłowy trwale lub okresowo utrudnia, ogranicza bądź uniemożliwia wypełnianie ról społecznych, a w szczególności ogranicza zdolności do wykonywania pracy zawodowej* (Ustawa RP, 1997).

Technologie wspierające niewidomych — zakres tematyczny książki

W ostatnich latach dokonał się bardzo duży postęp w technologii interfejsów wspomagających niewidomych w komunikacji z otoczeniem. Należą do nich m.in. takie urządzenia i programy komputerowe jak: wielofunkcyjne terminale Braille’a (monitory, notatniki, drukarki), syntezytory mowy, telefony komórkowe z dźwiękowym menu (Strumiłło i inni, 2019b) oraz programy komputerowe z interfejsem dźwiękowym (Strumiłło i Materka, 2011c). Jednakże przestrzenna orientacja i samodzielne poruszanie się nadal pozostają podstawowymi, nierozwiązanymi problemami osób niewidomych. Niewidomi podkreślają, że właśnie ograniczenia związane z samodzielnym, bezpiecznym poruszaniem i nawigowaniem w nieznanym terenie są podstawową barierą ograniczającą ich mobilność, a w konsekwencji wszystkie aktywności związane z funkcjonowaniem zawodowym i życiem społecznym.

Niniejsza monografia jest poświęcona systemom technicznym wspomagającym niewidomych w samodzielnym poruszaniu się i nawigacji. Prace nad tej klasy systemami są prowadzone od wielu lat (Hersh i Johnson, 2008). Wymagają interdyscyplinarnego podejścia badawczego – zastosowania specjalnych metod pomiarowych, technik obliczeniowych oraz spełnienia szczególnych wymagań ergonomicznych niewidomego użytkownika. Szeroki zakres i złożoność tej tematyki ilustruje skrócona historia badań nad urządzeniami i systemami budowanymi specjalnie dla osób niewidomych i słabowidzących.

Stosowaną od wieków pomocą nawigacyjną osoby niewidomej był kij lub laska. Wraz z postępującą urbanizacją i rozwojem motoryzacji niewidomi w coraz większym stopniu byli narażeni na niebezpieczeństwo. W latach 30. XX wieku Francuzka Guilly d’Herbement spopularyzowała swój pomysł używania przez niewidomych laski oznaczonej białym kolorem. Od tamtego czasu biała laska (*white cane*) stanowi nie tylko podstawową pomoc niewidomego w samodzielnym poruszaniu się ale również pełni ważną funkcję sygnalizacyjną dla otoczenia.

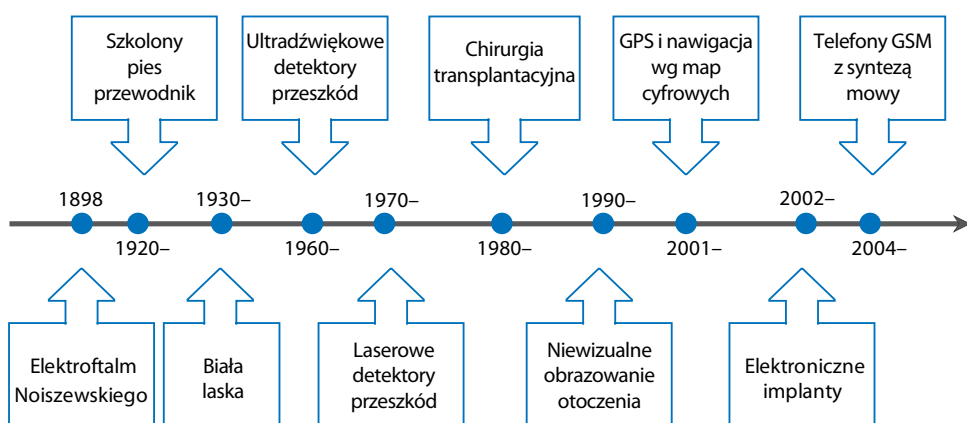
Urządzenia elektroniczne wspierające mobilność osób niewidomych (*electronic travel aid – ETA*) mają ponad stuletnią historię. Za pierwsze urządzenie tego typu uznaje się elektroftalm opracowany ok. 1898 roku przez polskiego okulistę Kazimierza Noiszewskiego. W urządzeniu tym, nazywanym sztucznym okiem, Noiszewski zastosował fotokomórkę selenową. Sygnały elektryczne odbierane z fotokomórki sterowały układami wytwarzającymi sygnały dźwiękowe lub bodźce dotykowe o amplitudzie proporcjonalnej do średniej jasności otoczenia.

Na rys. 1 pokazano zdjęcie zmodyfikowanej wersji elektroftalmu wykonanego w 1969 r. w Pomorskiej Akademii Medycznej w zespole prof. Witolda Starkiewicza. Urządzenie to miało duży ciężar i pobór mocy i nie znalazło praktycznego zastosowania.



Rys. 1. Elektroftalm mocowany na głowie pacjentki (wersja z 1969 r.), na czole są widoczne tzw. pulsatory wytwarzające bodźce dotykowe
Źródło: <http://pl.wikipedia.org/wiki/Elektroftalm>

Dopiero po II wojnie światowej zanotowano kolejny postęp w budowaniu elektronicznych urządzeń wspomagających niewidomych (rys. 2). W latach 60. XX wieku zbudowano pierwsze ultradźwiękowe czujniki odległości dla niewidomych, a w latach 70. zastosowano dalmierze laserowe jako detektory przeszkód.



Rys. 2. Historia współczesnych technik oraz procedur medycznych wspomagających osoby niewidome i słabowidzące (oś czasu w skali nieliniowej)

Źródło: opracowanie własne

Rozwój mikroprocesorów i urządzeń wizyjnych jaki dokonał się w latach 80-tych i 90-tych umożliwił kolejny postęp w technologiach wspomagających mobilność osób niewidomych. Rozpoczęto próby budowania osobistych systemów substytucji sensorycznej, urządzeń przetwarzających obrazy otoczenia na bodźce dźwiękowe lub dotykowe. Pomimo wielu opracowanych urządzeń prototypowych tej klasy z trudem znajdują one akceptację społeczności osób niewidomych. Opracowanie skutecznych technik, przekazujących osobie niewidomej informacje o otoczeniu w postaci niewizualnych „obrazów” dźwiękowych lub dotykowych okazuje się trudnym, interdyscyplinarnym problemem badawczym. Obejmuje on m.in. takie zagadnienia jak:

- techniki estymacji głębi, rekonstrukcji sceny trójwymiarowej i rozpoznawania obiektów otoczenia,
- rejestracja i analiza sekwencji obrazów w czasie rzeczywistym,
- pomiary wielkości nieelektrycznych, m.in. z wykorzystaniem czujników inercyjnych,
- prawa psychologii poznania oraz właściwości sensoryczne zmysłu wzroku oraz słuchu i dotyku, jako zmysłów wykorzystywanych do obrazowania otoczenia osobie niewidomej,
- metody orientacji przestrzennej i poruszania się osób niewidomych.

Podstawy wymienionych technik i zagadnień badawczych omówiono w pierwszym i drugim rozdziale książki. Obszar tej tematyki można określić mianem mikronawigacji, tj. technik związanych z rozpoznawaniem otoczenia w bliskiej przestrzeni, selekcją obiektów sceny (przeszkód) do prezentacji oraz automatycznym zliczaniem kroków osoby niewidomej.

Znaczący postęp dokonał się w ostatnich kilku latach w zakresie technik makronawigacji, tj. metod określania położenia geograficznego i wyznaczania ścieżek w dostępnej sieci dróg. W zastosowaniach cywilnych (w nawigacji samochodowej i pieszej) jest szeroko wykorzystywany globalny system pozycjonowania GPS (*Global Positioning System*) oraz technologie map cyfrowych. Rozwijane są metody integrujące dane mikro- i makronawigacyjne dla uzyskania lepszych estymat wyznaczanych położenia. Tematyka ta jest przedmiotem 3. rozdziału książki, w którym omówiono:

- podstawy globalnych systemów pozycjonowania,
- techniki nawigacji względnej, inercyjnej i wizyjnej,
- wyznaczanie położenia z zastosowaniem lokalnych i rozległych sieci radiowych,
- filtrację danych nawigacyjnych pozyskiwanych z różnych źródeł za pomocą filtracji optymalnej Kalmana oraz wywodzącej się z metod symu-

lacji Monte Carlo, bardzo skutecznej w tych zastosowaniach filtracji cząsteczkowej.

Osobiste systemy nawigacji, będą skuteczną pomocą dla niewidomego użytkownika, jeżeli uzyska on odpowiednią ilość i jakość informacji prezentowanej w postaci sygnałów dźwiękowych lub dotykowych. Temu niełatwemu problemowi, określanemu mianem substytucji sensorycznej obejmującemu podstawy budowy niewizualnych interfejsów użytkownika poświęcono rozdział 4. książki.

Charakter przeglądowy ma rozdział 5., w którym zawarto omówienie i porównanie różnej klasy najnowszych systemów ETA. W przeglądzie tym uwzględniono zarówno proste detektory przeszkód, bardziej złożone systemy obrazowania otoczenia, systemy globalnego pozycjonowania oraz systemy teleinformatyczne umożliwiające „prowadzenie” osoby niewidomej przez zdalnego przewodnika.

Podziękowania

Duża część materiału opisanego w niniejszej książce jest wynikiem zespołowej pracy badawczej prowadzonej w Instytucie Elektroniki Politechniki Łódzkiej. Składam podziękowania doktorantom i pracownikom Instytutu, którzy mieli wkład w te prace.

Autor wyraża podziękowanie Ministerstwu Nauki i Szkolnictwa Wyższego za dofinansowanie prac badawczych i rozwojowych poświęconych tematyce systemów wspomagania osób niewidomych. Przygotowanie książki zostało dofinansowane ze środków Narodowego Centrum Badań i Rozwoju w ramach grantu rozwojowego N R02 0083 10 realizowanego w Instytucie Elektroniki Politechniki Łódzkiej w latach 2010-2013. Wydanie książki było możliwe dzięki wsparciu finansowemu Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.

Szczególne podziękowania składam Okręgowi Łódzkiemu Polskiego Związku Niewidomych oraz dużej grupie osób niewidomych i słabowidzących wielokrotnie biorących udział w długotrwałych testach opracowywanych prototypów.

Serdeczne podziękowania kieruję również na ręce prof. dr hab. inż. Andrzeja Materki i dr hab. inż. Michała Strzeleckiego prof. PŁ za ich nieustającą zachętę do napisania tej książki i pomocne dyskusje.

Wzrok, jego niepełnosprawności oraz konsekwencje utraty wzroku

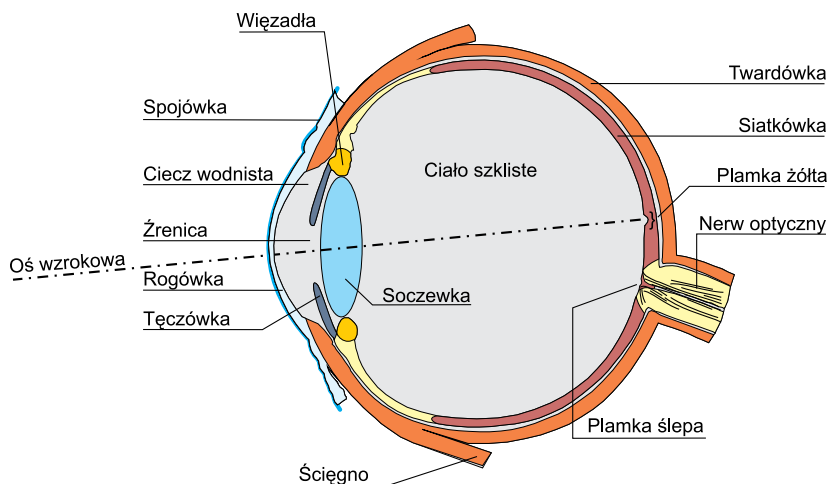
1.1. System wzrokowy człowieka

Wzrok jest najważniejszym zmysłem człowieka. Szacuje się, że dostarcza on ok. 90% informacji sensorycznej o otaczającym świecie. W przetwarzaniu i interpretacji tej bogatej informacji bierze udział ok. 30% zasobów mózgu (Kalat, 2011). Za pomocą zmysłu wzroku człowiek potrafi odbierać i analizować bodźce świetlne w szerokim zakresie kątowym nazywanym polem widzenia. Zdolność ta umożliwia szybką orientację przestrzenną i określanie własnej lokalizacji w otaczającej przestrzeni (Granlund, 1999). Zmysł wzroku pełni podstawową rolę w poruszaniu się, m.in. w utrzymaniu równowagi ciała, wykrywaniu przeszkód i wytyczaniu ścieżki marszu (McNamara i Shelton, 2002).

System wzrokowy człowieka składa się z części sensorycznej, części przewodzącej sygnały wzrokowe do wnętrza mózgu (tzw. drogi wzrokowej) i części analizującej te sygnały umiejscowionej w płacie potylicznym mózgu nazywanej korą wzrokową (Żernicki, 1983).

Funkcję sensoryczną w systemie wzrokowym pełnią ruchome gałki oczne, których ruch jest u człowieka wzajemnie zsynchronizowany. Przekrój przez gałkę oczną pokazano na rys. 1.1.

Światło widzialne (wąski zakres widma promieniowania elektromagnetycznego o długości fal od 380 do 780 nm) wnika do wnętrza oka przez źrenicę, otwór o regulowanym przez tęczęwkę promieniu odsłaniający soczewkę. Część fotoczułą oka stanowi siatkówka, cienka błona rozmieszczona na tylnej ścianie gałki ocznej. Znajduje się w niej ponad 100 mln fotoreceptorów (ok. 100 mln pręcików i ok. 6 mln czopków). Pręciki umożliwiają czarno-białe widzenie i są



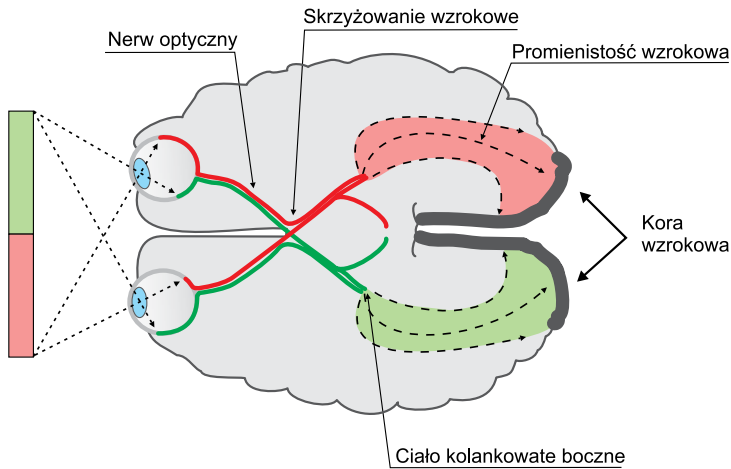
Rys. 1.1. Przekrój poziomy przez prawe oko człowieka (widok z góry)

Źródło: opracowanie własne

ok. 1000 razy czulsze niż czopki, które umożliwiają widzenie kolorów (Longstaff, 2006).

Siatkówka jest swoistym ekranem, na którym są odwzorowywane obrazy otoczenia, dzięki skupiającej światło rogówce i soczewce. Tylko 10% światła wnikałego do oka dociera do siatkówki. Fotoreceptory przetwarzają te słabe bodźce świetlne na potencjały czynnościowe o energii ok. 10^5 większej od padającego na nie światła. Największe zagęszczenie fotoreceptorów siatkówki (czopków) występuje w obrębie plamki żółtej. Ma ona średnicę ok. 1,5 mm i leży na osi wzrokowej. W osi tej człowiek ma największą ostrość widzenia.

Potencjały czynnościowe wytwarzane przez fotoreceptory są doprowadzane do nerwu optycznego liczącego ok. 1,5 mln włókien nerwowych. Włókna nerwów optycznych biegnących z lewego i prawego oka krzyżują się w tzw. skrzyżowaniu wzrokowym (rys. 1.2). Włókna przewodzące sygnały z lewej części siatkówki lewego oka i lewej części siatkówki oka prawego są kierowane do prawej półkuli mózgu – odpowiednio lewa półkula mózgu otrzymuje sygnały z prawych części siatekówek oczu. Następnie sygnały te poprzez tzw. promienistość wzrokową docierają do kory wzrokowej, cienkiej warstwy tkanki nerwowej, w której sygnały wzrokowe są analizowane i integrowane w spójne przestrzenne wrażenia wzrokowe. Należy zaznaczyć, że sygnały wzrokowe zbierane z plamki żółtej są odwzorowywane i analizowane przez obszar zajmujący prawie całą korę wzrokową. Plamka żółta leży na tzw. osi wzrokowej, linii wyznaczającej punkt skupienia wzroku (rys. 1.1).

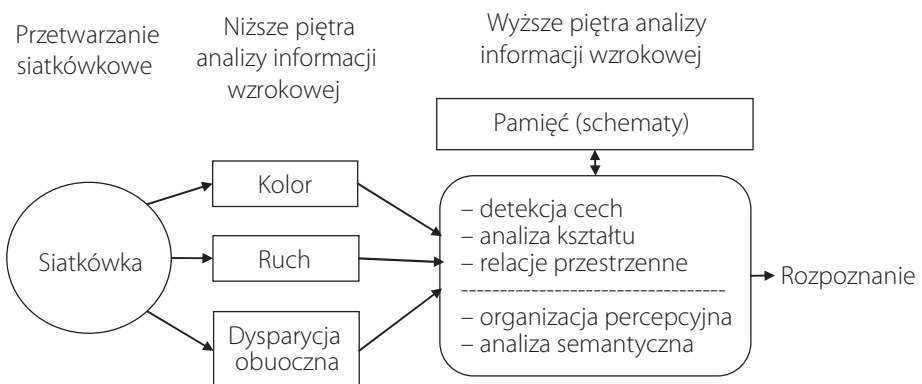


Rys. 1.2. System wzrokowy człowieka (w uproszczeniu)

Źródło: opracowanie własne

Wyniki badań systemu wzrokowego człowieka pokazują, że analiza sygnałów wzrokowych przebiega równolegle w kilku kanałach wzrokowych i charakteryzuje się hierarchiczną strukturą (Milner i Goodale, 2008). W oddzielnych kanałach zachodzą procesy związane z rozpoznawaniem kolorów, spostrzeganiem ruchu i głębi oraz relacji przestrzennych. Na wyższych piętrach systemu wzrokowego (w korze mózgowej) zachodzi organizacja percepcyjna i analiza semantyczna informacji wzrokowej (rys. 1.3).

Kluczowe znaczenie dla prawidłowej orientacji przestrzennej ma zdolność do percepcji głębi (DeAngelis, 2000). Dzięki tej zdolności, na podstawie dwu-

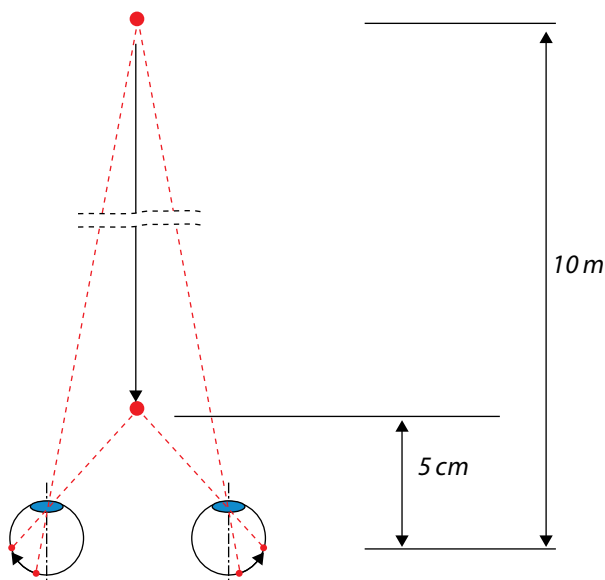


Rys. 1.3. Uproszczony schemat hierarchicznego przetwarzania informacji wzrokowej

Źródło: opracowanie własne

wymiarowych obrazów człowiek umie odtworzyć trójwymiarową strukturę otoczenia. Do percepcji głębi człowiek wykorzystuje tzw. wskazówki jednooczne i dwuoczne. Do wskazówek jednoocznych należą cechy obrazów otoczenia, takie jak: perspektywa, zmiana skali obiektów o znanych rozmiarach, okluzja (tj. zasłonięcie obiektu dalszego przez obiekt bliższy), cienie i ruch obiektów. Wskazówki jednooczne mają podstawowe znaczenie w spostrzeganiu odległych przedmiotów.

Podstawą dwuocznego spostrzegania głębi jest stereoskopia. Ten mechanizm widzenia zachodzi dla pola widzenia objętego obojgiem oczu. Ponieważ miejsca obserwacji otoczenia przez oko lewe i prawe są względem siebie przesunięte, względnemu przesunięciu (dysparycji) podlegają również obrazy siatkówkowe wytworzone na siatkówkach, tym większemu im bliżej obserwatora znajduje się obserwowany przedmiot (rys. 1.4). Przesunięcia te mają wartości ułamków milimetra. Układ wzrokowy mózgu porównuje te przesunięcia uwzględniając też punkt fiksacji (tj. skupienia wzroku) i wypracowuje informację o głębi (Wexler, 2002). Ocena stereoskopowa głębi jest skuteczna dla odległości do ok. 10 m. Człowiek wspomaga się w niej mikroruchami gałek ocznych i ruchami głowy (Wexler i Boxter, 2005). Stereoskopia jest podstawową metodą estymacji głębi w technicznych układach wizyjnych omówionych w rozdz. 2.1.2.



Rys. 1.4. Ilustracja mechanizmu widzenia stereoskopowego; obraz punktu otoczenia na siatkówce podlega odsuwaniu od osi wzrokowej wraz ze zbliżaniem się tego punktu do obserwatora (na rysunku przyjęto, że punkt fiksacji znajduje się w nieskończoności)

Źródło: opracowanie własne

Szczególnie złożony jest mechanizm związany z rozpoznawaniem kształtów obiektów, zwłaszcza twarzy (Kalat, 2011). Procesy rozpoznawania angażują obszary mózgu związane z pamięcią, w której są przechowywane schematy (prototypy) rozpoznawanych obiektów (Wallis i Bülthoff, 1999)). Człowiek skutecznie przypisuje obserwowane obiekty do schematów pomimo różnego kontekstu i perspektywy w jakich obiekty te występują. Wiele obiektów na rys. 1.5 człowiek przypisze do schematu 'fotel' pomimo, że obrazy foteli odwzorowane na siatkówce mają różne kształty i występują w różnym kontekście; m.in. skutecznie rozróżni cień fotela i obraz fotela od samego przedmiotu oraz rozpozna fotel częściowo zasłonięty przez inne przedmioty.



Rys. 1.5. Obraz ilustrujący skuteczność systemu wzrokowego człowieka w rozpoznawaniu obiektów i przypisywania ich do określonej klasy schematów

Źródło: (Wallis i Bülthoff, 1999) za pozwoleniem wydawnictwa Elsevier

Znany jest przypadek Mikea Maya, który w wieku 3 lat na skutek wypadku stracił wzrok. Dzięki postępowi jaki dokonał się w medycynie transplantacyjnej, po 40 latach przeszczepiono mu rogówkę od innego dawcy i przywrócono wzrok w jednym oku (Kurson, 2008). Mike wykrywał ruch i poznawał kolory, nie rozpoznawał natomiast twarzy osób, liter i tych samych przedmiotów widzianych z różnej perspektywy. Mylił się też co do głębi, m.in. wpadał nieoczekiwanie na przeszkody w czasie marszu. W czynnościowym badaniu tomograficznym fMRI (*functional resonance imaging*) stwierdzono, że ośrodki kory wzrokowej Mikea wykorzystywane do rozpoznawania kształtów i głębi na skutek wieloletniego braku sygnałów wzrokowych utraciły swe pierwotne funkcje i przystoso-

wały się do pełnienia innych funkcji, np. przetwarzania bodźców słuchowych lub dotykowych (Fine i inni, 2003). W neuropsychologii mechanizm ten nazywa się plastycznością mózgu. Co więcej proces ten u osoby dorosłej jest utrwalony i nieodwracalny. Stwierdzono również, że ośrodki mózgu Mikea odpowiedzialne za rozpoznawanie kolorów i ruchu zachowały swe pierwotne funkcje. Wynik ten potwierdza udokumentowany w psychologii rozwoju fakt o znacznie prostszym mechanizmie rozpoznawania tych cech otoczenia utrwalanych we wczesnym dzieciństwie. Studiowany przez neuropsychologów przypadek Mikea Maya pokazuje, że pomimo przywrócenia mu prawidłowej czynności drogi wzrokowej nie odzyska on pełnej sprawności rozpoznawania otoczenia i orientacji przestrzennej (Milner i Goodale, 2008). Zatem nie można oczekiwać, że postęp w technikach chirurgicznych (m.in. transplantologii) umożliwi w pełni skuteczne przywracanie wzroku osobom, które są niewidome od urodzenia lub pozostają niewidome przez długi okres życia (Bach-y-Rita i inni, 2003b).

Duży wkład w wyjaśnienie znaczącej roli jaką w percepcji otoczenia pełni przetwarzanie informacji wzrokowej w mózgu wnieśli już w ubiegłym wieku psychologowie postaci (Koffka, 1935). Sformułowane przez nich prawa percepcji omówiono w rozdz. 4. Tematykę podkreślającą znaczenie kognitywnego aspektu spostrzegania wzrokowego¹ rozwijał również brytyjski psycholog Richard Gregory (Gregory, 2005).

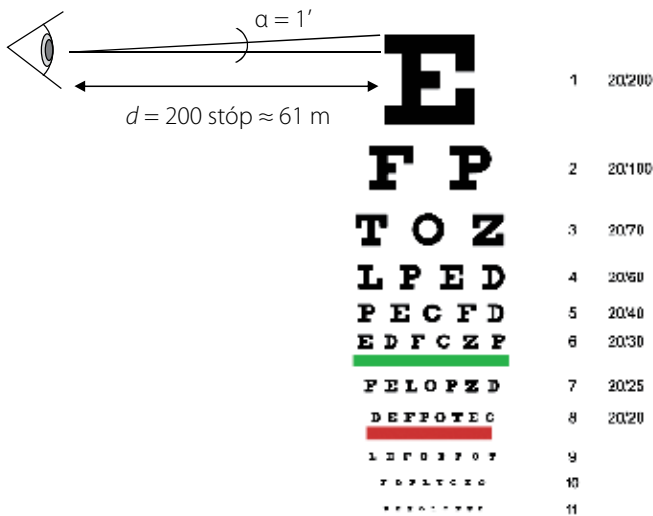
1.2. Badanie, ocena sprawności i wady wzroku

Stopień sprawności wzroku ocenia się na podstawie specjalistycznych pomiarów okulistycznych. Do podstawowych badań należą: pomiar ostrości widzenia (*visual acuity*) oraz pomiar pola widzenia (*field of vision test*). Inne badania funkcjonalne wzroku dotyczą percepcji głębi, widzenia barwnego oraz bardziej zaawansowanych badań tzw. wzrokowych potencjałów wywołanych (*visual evoked potentials – VEP*) polegających na rejestracji i analizie odpowiedzi elektrycznej kory wzrokowej na bodziec świetlny (Bronzino, 1999).

Ostrość (rozdzielczość) widzenia definiuje się jako zdolność oka do różnicowania niewielkich obiektów znajdujących się blisko siebie. Pomiar ostrości widzenia polega na badaniu zdolności do rozpoznania liter o różnej wielkości

¹ „Nie widzimy rzeczy jakimi są; widzimy rzeczy jakimi my jesteśmy” to trafne określenie spostrzegania wzrokowego podała francuska pisarka Anaïs Nin (1903–1977).

widzianych kolejno lewym i prawym okiem z zadanej standardowo odległości 6 m (20 stóp) i ustalonych warunkach oświetlenia. Do pomiaru są stosowane specjalne tablice (tzw. tablice Snellena) z wierszami liter o coraz mniejszej czcionce (rys. 1.6). Ułamki znajdujące się obok każdego wiersza tablicy są miarą ostrości widzenia, którą interpretuje się następująco. Liczba w liczniku ułamka oznacza odległość, z której osoba badana widzi tablicę, tj. standardowo z odległości 6 m (20 stóp). Liczba w mianowniku ułamka oznacza odległość (wyrażoną odpowiednio w stopach lub metrach), z której widziana grubość linii użytej w kroju czcionki odpowiada jednej minucie kątowej ($\alpha = 1' = 1/60^\circ$) dla kierunków promieni obrazujących linię czcionki i zbiegających się w oku obserwatora.



Rys. 1.6. Tablica Snellena używana do badania ostrości wzroku

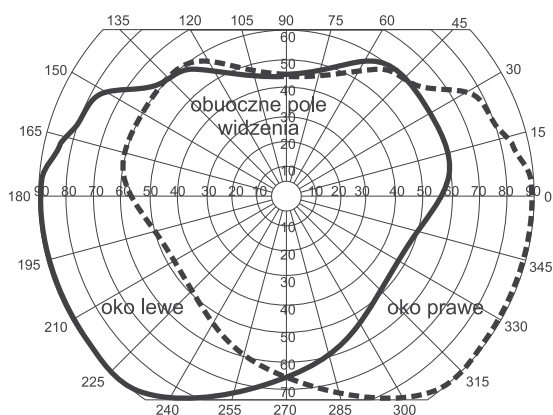
Źródło: opracowanie własne na podstawie http://en.wikipedia.org/wiki/Snellen_chart

Za normę uznaje się ostrość widzenia 20/20 (odpowiednio dla metrycznego odpowiednika 6/6), tj. gdy badana osoba rozpoznaje litery do 8-ego wiersza łącznie. Miary 20/40 i 20/10 oznaczają odpowiednio dwukrotnie gorszą i dwukrotnie lepszą ostrość widzenia w stosunku do normy.

Pomiar pola widzenia (tzw. badanie perymetryczne) jest stosowane do oceny czynności siatkówki, nerwu wzrokowego, dróg wzrokowych oraz kory wzrokowej. W badaniu perymetrycznym osoba badana koncentruje wzrok na wskazanym obiekcie (punkcie referencyjnym). Następnie (w tzw. perymetrii kinetycznej) do punktu referencyjnego z różnych kierunków jest przemieszczane punktowe źródło światła. Osoba badana sygnalizuje dla jakiego położenia ruchomy punkt świetlny został zauważony. Wynik badania w postaci obszarów określających pole wide-

nia dla oka lewego i prawego jest wykreślany na tle koncentrycznych okręgów odpowiadających kolejnym odległościom kątowym od punktu referencyjnego. Przykładowy wynik badania perymetrycznego pokazano na rys. 1.7.

Dla prawidłowego pola widzenia zakres kątowy jednoocznego pola widzenia wynosi wzdłuż kierunku pionowego ok. 120° , a dla kierunku poziomego (linia skroń–nos) ok. 150° . Obszar, w którym pola widzenia oka lewego i prawego nakładają się nazywa się obuocznym polem widzenia. Wycinek otoczenia widziany w tym obszarze może być postrzegany stereoskopowo.



Rys. 1.7. Diagram ilustrujący wynik badania pola widzenia, linia ciągła – granica pola widzenia oka lewego, linia przerywana – granica pola widzenia oka prawego
Źródło: opracowanie własne na podstawie portalu internetowego mediweb.pl

Zgodnie z zaleceniami Międzynarodowej Klasyfikacji Chorób (*International Classification of Diseases – ICD*) stopień niepełnosprawności wzrokowej (jednostka chorobowa H54) jest oceniany na podstawie miary ostrości widzenia oraz wielkości pola widzenia. W odniesieniu do ostrości widzenia wyróżniono kategorię dla widzenia normalnego (miary ostrości w zakresie 6/6–6/18) oraz pięć kategorii kwalifikowanych jako niepełnosprawność wzrokową. Kategorie te i odpowiadające im miary ostrości widzenia (po uwzględnieniu najlepszej dostępnej korekcji wady wzroku) wyszczególniono w tabeli 1.1.

Wymienione kategorie upośledzenia wzroku wg klasyfikacji ICD są używane do celów orzecznictwa medycznego i niekoniecznie odnoszą się do możliwości funkcjonalnych osoby niepełnosprawnej, np. osoba słabowidząca o ostrości widzenia 1/60 potrafi liczyć palce u dłoni z odległości 1 m, a osoba o ostrości widzenia 2/60 może poruszać się bez przewodnika. Takie „resztki” widzenia diametralnie zwiększają szanse uczestnictwa osoby słabowidzącej w życiu społecznym i zawodowym.

W świetle polskiego prawa niewidomym jest osoba, u której ostrość wzroku jest nie większa niż 0,1 lub ma pole widzenia nie większe niż 30° (Yeadon, 2000). Za niewidomą uważa się osobę, która jest niewidoma od urodzenia lub utraciła wzrok przed 5. rokiem, a ociemniałą jest osoba, która utraciła wzrok w późniejszym okresie życia i pamięta obrazy wzrokowe.

Tabela. 1.1. Stopnie niepełnosprawności wzrokowej wg WHO (2011)

Kategoria	Ostrość widzenia po korekcji	Określenia standardowe wg WHO	Określenia funkcjonalne wg WHO
0	6/6 – 6/18	Normalny	Normalne
1	6/18 – 6/60	Upośledzenie wzroku	Słabowidzenie
2	6/60 – 3/60	Poważne upośledzenie wzroku	Słabowidzenie
3	3/60 – 1/60*	Ślepotą	Słabowidzenie
4	0.02 – poczucie światła	Ślepotą	Słabowidzenie
5	Brak poczucia światła	Ślepotą	Całkowita ślepotą

* 1/60 – liczenie palców dłoni z odległości 1 m.

Źródło: opracowanie własne na podstawie WHO (2011)

Ponadto wadę wzroku, w której pole widzenia jest zawężone do 10° a jest większe niż 5° zalicza się do kategorii niepełnosprawności wzrokowej 3, a wadę, w której pole widzenia jest ograniczone do 5° zalicza się do kategorii 4 (pomimo normalnej ostrości widzenia w osi optycznej oka). Tego typu zawężenie pola widzenia nazywa się widzeniem lunetowym (lub tunelowym), rys. 1.8, jego przyczyną może być zwyrodnienie siatkówki oka.



Rys. 1.8. Prawidłowe widzenie (z lewej) oraz tzw. widzenie lunetowe (z prawej)

Źródło: opracowanie własne

W roku 2001 Światowa Organizacja Zdrowia zastąpiła stosowaną od wielu lat międzynarodową klasyfikację zdrowia ICD nową klasyfikacją. Międzynarodowa Klasyfikacja Funkcjonowania, Niepełnosprawności i Zdrowia (*International Classification of Functioning, Disability and Health – ICF*) jest obecnie obowiązującym standardem stosowanym do usystematyzowanej oceny stanu zdrowotnego człowieka i jego niepełnosprawności (Helal i inni, 2008). Elementem nowego podejścia do definicji zdrowia w klasyfikacji ICF jest wyróżnienie dwóch aspektów niepełnosprawności, tj. aspektu związanego z oceną medyczną niepełnosprawności (stosowaną w klasyfikacji ICD) oraz aspektu środowiskowego i społecznego, w których są umiejscowione bariery ograniczające zdolność osoby niepełnosprawnej w aktywnym uczestniczeniu (funkcjonowaniu) w życiu społecznym i zawodowym.

Niepełnosprawność wzroku może wynikać z wielu przyczyn, m.in. chorób, wypadków lub mieć podłoże genetyczne. Za podstawowe wady wzroku uważa się: krótkowzroczność, nadwzroczność, astygmatyzm, daltonizm, zez i zaćmę (zmętnienie soczewki zwane też kataraktą). Większość wad wzroku, z wyjątkiem ślepoty barw, można korygować odpowiednimi szklami korekcyjnymi lub leczyć operacyjnie, np. stosować przeszczep soczewki.

Chorobami oczu, które mogą prowadzić do ślepoty to choroby siatkówki:

- zwyrodnienie płamki żółtej (*age-related macular degeneration – AMD*) powodowane powolną degeneracją fotoreceptorów występującą głównie u osób w starszym wieku,
- retinopatia barwnikowa (*retinitis pigmentosa – RP*) – choroba dziedziczna powodowana odkładaniem barwnika w siatkówce i prowadząca do zmian zanikowych fotoreceptorów siatkówki (wyróżnia się również tzw. retinopatię cukrzycową, której przyczyną jest uszkodzenie naczyń krwionośnych siatkówki).

Choroby mogące powodować nieodwracalne uszkodzenie nerwu wzrokowego jest jaskra (*glaukoma*) zwykle objawiająca się podwyższonym ciśnieniem wewnątrz gałki ocznej oraz zapaleniem nerwu wzrokowego.

Utrata widzenia może być również wynikiem poważnych urazów gałki ocznej oraz uszkodzenia innych części systemu wzrokowego, np. przerwania nerwu wzrokowego oraz uszkodzenia kory wzrokowej w wyniku urazu mechanicznego lub udaru mózgu. Utrata wzroku wywołana uszkodzeniem kory wzrokowej nazywa się ślepotą korową, zależnie od miejsca uszkodzenia może ona powodować trudności w rozpoznawaniu przedmiotów i osób oraz ślepotę na ruch (Kalat, 2011).

1.3. Konsekwencje utraty wzroku

Utrata wzroku jest najpoważniejszą niepełnosprawnością sensoryczną. Człowiek niewidomy jest pozbawiony podstawowego zmysłu, który u zdrowej osoby dostarcza ok. 90% informacji o świecie zewnętrznym. Jest całkowicie pozbawiony możliwości wzrokowego poznawania świata i zjawisk. W dużym stopniu traci orientację przestrzenną i zdolność do samodzielnego poruszania się i podróżowania. Ograniczona mobilność ma negatywne konsekwencje dla sprawności psychofizycznej osoby niewidomej oraz niekorzystnie oddziałuje na sferę emocjonalną (Paplińska, 2008). Osoba niewidoma nie ma dostępu do informacji przekazywanej w postaci wizualnej, m.in. znaków informacyjnych, tekstu pisanego oraz obrazów, filmów i innych dzieł sztuki. Ma duże trudności w wykonywaniu czynności życia codziennego. Ma ograniczoną dostępność do edukacji, usług (np. medycznych, komunikacji miejskiej) i urzędów (m.in. systemów informatycznych, np. bankomatów). Brak wykształcenia skutkuje trudnościami w podjęciu pracy zawodowej. W dużym stopniu ma utrudnioną komunikację interpersonalną stanowiącą barierę w normalnym uczestnictwie w życiu społecznym.

Instytut Elektroniki Politechniki Łódzkiej (PŁ) jest jednym z kilku ośrodków badawczych w kraju podejmującym tematykę związaną z budową systemów elektronicznych i teleinformatycznych wspomagających osoby niewidome. W ramach prowadzonych prac ankietowano 20 osób niewidomych i słabowidzących zrzeszonych w Okręgu Łódzkim Polskiego Związku Niewidomych (IE PŁ, 2007). W ankietach zawarto pytania dotyczące podstawowych problemów i barier ograniczających normalne funkcjonowanie osób z niepełnosprawnością wzroku. Syntetyczne podsumowanie wyników zestawiono w trzech grupach problemów, na które przede wszystkim wskazały osoby niewidome.

1. Samodzielne poruszanie się i bezpieczeństwo:
 - omijanie przeszkód i innych użytkowników drogi,
 - lokalizacja nieciągłości nawierzchni (nierówne chodniki, wykopy, schody i rozpoznanie czy prowadzą w dół czy w górę)
 - unikanie kolizji z pojazdami w ruchu,
 - unikanie napadów i kradzieży,
 - 99% osób niewidomych nie wskazuje na potrzebę korzystania z systemów wspomagania w poruszaniu się wewnątrz budynków (z wyjątkiem bardzo dużych pomieszczeń, hal sportowych, terenów wystawowych).
2. Orientacja przestrzenna i nawigacja:
 - określanie własnej lokalizacji geograficznej i orientacji w otwartej przestrzeni (tj. na zewnątrz budynków);

- utrzymywanie kierunku (w tym informacja o stronach świata),
- zliczanie własnych kroków,
- planowanie trasy, nawigowanie wg trasy i wyznaczanie trasy powrotnej.

3. Dostęp do informacji:

- wykrywanie i rozpoznawanie tekstu,
- znaki graficzne (przystanek, przejście dla pieszych),
- identyfikacja pojazdów komunikacji miejskiej (lokalizacja wejść do pojazdu),
- obsługa telefonu komórkowego.

Bardziej szczegółową analizę przeprowadzonych ankiet zawarto w rozdz. 1.4.

Kwapiszowie (1990) zwracają uwagę na znaczenie i korzyści płynące z jak najszybszego podjęcia nauki orientacji przestrzennej i samodzielnego poruszania się przez osoby niewidome i ociemniałe. Do korzyści zaliczają: nabycie bezpiecznych technik bezpiecznego poruszania się, rozwijanie sposobów poznawania świata i zjawisk, poczucie satysfakcji z pokonania własnej bierności ruchowej oraz poprawę sprawności fizycznej, zwiększenie szans podjęcia regularnej edukacji i pracy zawodowej oraz zmniejszenie kosztów ponoszonych na opiekę i leczenie osób niewidomych.

1.4. Orientacja przestrzenna i poruszanie się osób niewidomych

Poznanie otaczającej przestrzeni i własnego położenia względem obiektów tej przestrzeni jest czynnością sprawnie wykonywaną przez osoby ze zdrowym wzrokiem bez potrzeby przemieszczania się. Umiejętność tę określa się mianem orientacji przestrzennej, którą Kwapiszowie (1990) w odniesieniu do osób niewidomych, definiują następująco:

Orientację przestrzenną można określić jako sprawność jednostki w zakresie poznawania swego otoczenia oraz zachodzących w nim stosunków czasowych i przestrzennych. Zasadniczą rolę odgrywają tu procesy poznawcze, zasób pojęć, znajomość schematu ciała, wyobrażenia przestrzenna, wiedza o otoczeniu, operowanie relacjami odległości i czasu. Nie bez znaczenia jest tutaj aspekt motoryczności jako warunku kształtowania pojęć w praktycznym działaniu.

Warto w tej definicji zwrócić uwagę na wzmiankowane relacje czasowo-przestrzenne i aspekt motoryczności w poznawaniu otoczenia. Podczas gdy osoby sprawne wzrokowo mogą szybko lokalizować obiekty otoczenia, to osoba niewidoma do ich lokalizacji musi przemieścić się kolejno do tych obiektów

i spostrzegać je w konkretnej relacji czasowej, a nie przestrzennej. Zatem orientacja przestrzenna dla osoby niewidomej jest zadaniem znacznie trudniejszym. Wzajemne relacje przestrzenne obiektów (ich wzajemne odległości i kierunki) są u osób niewidomych wypracowywane wg czasowych zależności w kolejno poznawanych obiektach przestrzeni na podstawie zapamiętanej liczby kroków i kierunku marszu (oraz zmiany tego kierunku). Taki sposób poznawania otoczenia nazywany integracją ścieżki (*path integration*) nie jest zbyt skuteczny gdyż kolejne błędy odległości i kierunku kumulują się (Loomis i inni, 1999).

Opisane problemy w orientacji przestrzennej osób niewidomych powodują, że zarówno poznawanie nowego otoczenia jak i poruszanie się w znanym otoczeniu stanowi dla wielu osób barierę trudną do pokonania. Nabycie umiejętności samodzielnego, sprawnego i bezpiecznego poruszania się wymaga odpowiedniego treningu i jest uwarunkowane wieloma czynnikami, m.in. wiekiem, sprawnością fizyczną i kognitywną osoby niewidomej, asertywną postawą do własnej niepełnosprawności oraz okresem życia, w którym osoba utraciła wzrok (Millar, 2002).

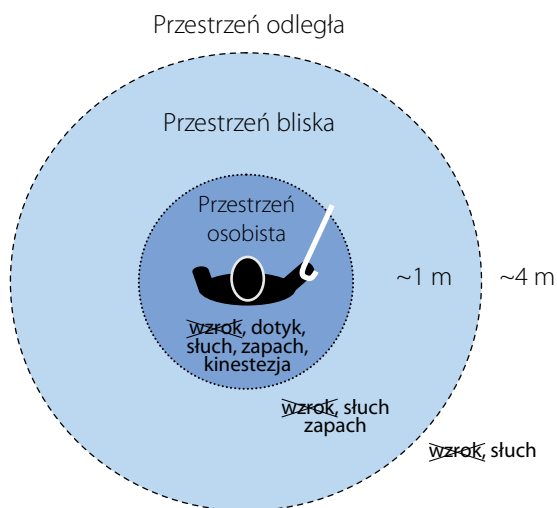
Istnieje bardzo bogata literatura poświęcona tematyce orientacji przestrzennej i mobilności (*spatial orientation and mobility – O&M*) osób niewidomych. Wśród amerykańskich badaczy należy wymienić przede wszystkim prace Roberty Klatzky (www.psy.cmu.edu/faculty/klatzky) oraz niewidomego od 1984 roku Reginalda Golledgea (1999) i Jacka Loomisa (<http://www.psych.ucsb.edu/~loomis>), a wśród badaczy europejskich Susannę Millar (2002), Marion Hersh (2008) i Edwigę Pissaloux (2007). W Polsce podstawowym źródłem wiedzy o tej tematyce jest klasyczny podręcznik Jadwigi i Jacka Kwapiśzów (1990) oraz nowsze opracowanie pod redakcją Paplińskiej (2008), które tylko w części jest poświęcone orientacji przestrzennej. Ta ostatnia pozycja jest wydana w tzw. formacie DAISY (*digital accessible information system*) dostępnym dla osób mających problemy w czytaniu, m.in. dla osób niewidomych (www.daisy.com).

W niniejszym rozdziale zebrano i krótko objaśniono podstawowe pojęcia związane z orientacją przestrzenną i mobilnością osób niewidomych. Podjęto próbę usystematyzowania wielu bliskoznacznych pojęć, które w potocznym użyciu są stosowane zamiennie i nie zawsze precyzyjnie.

Podstawowym pojęciem, definiowanym w psychologii poznawczej jest świadomość istnienia otaczającej przestrzeni i jej spostrzegania. Podstawowym wnioskiem wynikającym z prac Ungara i innych (2004) jest stwierdzenie, że brak zdolności do spostrzegania wzrokowego nie zamyka możliwości percepcji przestrzennej.

Spostrzeganie przestrzeni – to zdolność oceny geometrycznej struktury otaczającego środowiska, świadomość własnej lokalizacji w tej przestrzeni i znajomość relacji odległość–kierunek pomiędzy obiektami w otaczającej przestrzeni trójwymiarowej.

W pracach poświęconych percepcji przestrzennej osób niewidomych definiuje się tzw. przestrzeń bliską i odległą (Millar, 2002). Jeszcze bardziej szczegółowego podziału przestrzeni dokonał amerykański antropolog Edward Hall. Na rys. 1.9 zilustrowano model tzw. przestrzeni społecznej wg Halla. Współśrodkowymi kołami zakreślono obszary otaczającej przestrzeni podzielonej na przestrzeń osobistą (w zasięgu rąk, tzw. przestrzeń haptyczną), przestrzeń bliską (w promieniu ok. 4 m) oraz przestrzeń odległą. Osoba niewidoma dla poznania otaczającej przestrzeni (zależnie od jej skali) wykorzystuje zmysł dotyku (np. do manualnej oceny kształtu elementów przestrzeni), przemieszczanie się (dla eksploracji otoczenia w odległości większej niż zasięg rąk lub białej laski) oraz zmysł słuchu, m.in. do rozpoznawania rodzaju otoczenia (np. środowisko miejskie) lub oceny wielkości i kształtu pomieszczenia, np. na podstawie pogłosu lub echa.



Rys. 1.9. Definicja różnej skali przestrzeni wg Halla oraz modalności zmysłowe wykorzystywane przez osobę niewidomą do jej spostrzegania

Źródło: opracowanie własne

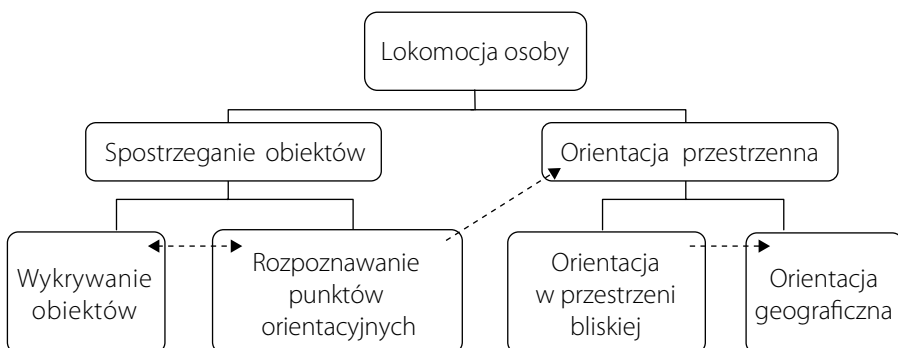
W definicji przestrzeni wg Halla został przyjęty tzw. model egocentryczny, w którym relacje przestrzenne obiektów i zdarzeń jest ustalane względem układu odniesienia związanego z odbiorcą (*body-centred frame of reference*) (Millar, 2002).

Innym stosowanym modelem jest tzw. model allocentryczny (*allocentric frame of reference*), w którym punktami orientacyjnymi są stałe punkty umiejscowione w otoczeniu oraz kierunki wyznaczone wg stron świata (Klatzky, 1999).

W wielu pracach dotyczących orientacji przestrzennej osób niewidomych stwierdzono, że niewidomi stosują zarówno model egocentryczny jak i allocentryczny w spostrzeganiu przestrzeni. Wykazano, że ci którzy stosują model egocentryczny lepiej rozpoznają bliską przestrzeń, zaś osoby niewidome stosujące model allocentryczny potrafią lepiej spostrzegać złożone zależności przestrzenne w otoczeniu odległym (Ungar i inni, 2004). W wielu innych pracach podkreśla się różnice w spostrzeganiu otoczenia przez osoby niewidome od urodzenia i osoby ociemniałe. Te drugie, tj. osoby które utraciły wzrok w późniejszym okresie życia, lepiej rozumieją pojęcie perspektywy i skuteczniej budują trójwymiarowe mentalne mapy otoczenia. Z kolei niewidomi od urodzenia mają lepiej wypracowane techniki zarówno niewizualnego spostrzegania przestrzeni jak i sprawnego poruszania się w otoczeniu.

W odniesieniu do poruszania się osób niewidomych definiuje się pojęcie mobilności, której związłą ale trafną definicję zaproponował Foulke (1971): *mobilność* to zdolność do samodzielnego, bezpiecznego, sprawnego i nieuciążliwego (wygodnego) poruszania się. Z kolei Pissaloux (2007) definiuje mobilność w szerszym kontekście obejmującym zdefiniowane wcześniej: spostrzeganie przestrzeni i orientację przestrzenną ale również: samą czynność fizyczną związaną z marszem i utrzymaniem równowagi ciała oraz umiejętnością nawigacji w otoczeniu (której definicja zostanie przytoczona później).

Z kolei Brambring (1985) zaproponował model poruszania się osoby niewidomej, który nazwał lokomocją i wyróżnił w nim aspekt związany ze spostrzeganiem obiektów przestrzeni oraz orientację przestrzenną w różnych skalach. Na rys. 1.10 w postaci schematu blokowego pokazano model lokomocji osoby niewidomej wg Brambringa.



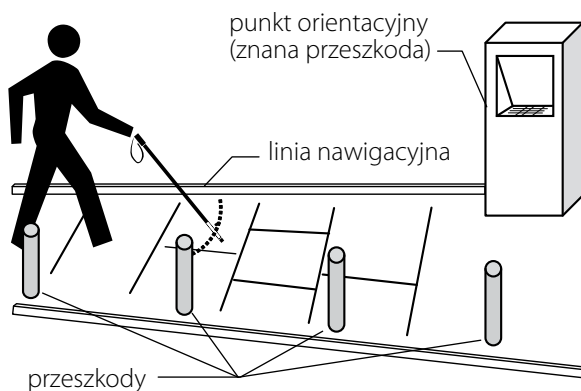
Rys. 1.10. Model Brambringa definiujący lokomocję osoby niewidomej

Źródło: opracowanie własne na podstawie (Hersh i Johnson, 2008)

W modelu lokomocji osoby niewidomej w części obejmującej spostrzeganie obiektów można wyróżnić:

- przeszkody (*obstacles*) – tj. elementy otoczenia (stałe lub ruchome) znajdujące się na ścieżce marszu, mogące stanowić zagrożenie dla niewidomego pieszego;
- punkty orientacyjne (*landmarks*) – tj. elementy i właściwości otoczenia (również rozpoznane wcześniej przeszkody), które mogą służyć jako wskazówki nawigacyjne (m.in. również charakterystyczne dźwięki i zapachy).

Wykryte obiekty otoczenia (jak zaznaczono strzałką na rys. 1.10) służą też jako wskazówki lokalizacyjne w orientacji przestrzennej. Szczególnymi elementami otoczenia są również obiekty rozciągłe (ściany, krawężniki), które mogą stanowić linie nawigacyjne (*shorelines*), pomocne w utrzymaniu kierunku marszu. Na rys. 1.11 zilustrowano przykład różnej interpretacji obiektów otoczenia przez osobę niewidomą.



Rys. 1.11. Ilustracja różnej interpretacji obiektów otoczenia przez niewidomego pieszego

Źródło: opracowanie własne

W modelu lokomocji osoby niewidomej wg Brambringa (rys. 1.10) są również wyróżnione dwie kategorie orientacji przestrzennej.

1. *Orientacja w przestrzeni bliskiej* – związana z umiejętnością utrzymywania kierunku marszu bez zbaczenia z celu, np. wzdłuż chodnika.
2. *Orientacja geograficzna* – tj. orientacja w przestrzeni odległej związana ze zdolnością określania własnego położenia na trasie marszu względem punktu startowego i celu.

W obu sposobach orientacji przestrzennej osoba niewidoma korzysta zarówno ze spostrzeganych na bieżąco obiektów i jak i wiedzy *a priori* o pokonywa-

nej trasie. Jeżeli jest to trasa pokonywana pierwszy raz to zadanie lokomocji osoby niewidomej w takim terenie jest znacznie bardziej utrudnione (McNamara i Shelton, 2002). Jednym z rozwiązań jest uprzednie przejście zadanej trasy z przewodnikiem lub, co jest bardziej preferowane przez osoby niewidome, z inną osobą niewidomą, która potrafi zwrócić uwagę na przeszkody i charakterystyczne punkty orientacyjne dla lepszego nawigowania wzdłuż trasy. Dużą pomoc w tym zakresie stanowią urządzenia nawigacji satelitarnej z interfejsem dźwiękowym, które opisano w rozdziale 5.

Zdolność do budowania i zapamiętywania tzw. mapy wyobraźniowej otoczenia zwanej też mapą mentalną lub poznawczą (*cognitive map*) jest kluczową umiejętnością warunkującą dobrą orientację geograficzną u ludzi i zwierząt (Tolman, 1948). Umiejętność tę często określa się mianem „wewnętrznej pomocy nawigacyjnej” (Golledge, 1999). Osoba niewidoma może budować taką mapę na kilka sposobów:

- osobiste pokonanie nowej trasy (najlepiej z inną osobą niewidomą),
- na podstawie opisu słownego uzyskanego od innej osoby (lub programu komputerowego wyposażonego w syntezy mowy),
- haptyczne badanie tzw. map dotykowych (*tactile maps*), tj. druku wypukłego obrazującego wycinek terenu w odpowiedniej skali (Ungar i inni, 2004).

Wykazano, że ten ostatni sposób budowania mapy wyobraźniowej u osób niewidomych jest szczególnie skuteczny (Espinoza i inni, 1998).

Problemowi budowania map wyobraźniowych poświęcono wiele prac badawczych i wykonano wiele praktycznych testów z udziałem osób niewidomych (Bradley i Dunlop, 2008). Opisano dwa zasadnicze sposoby, wg których osoby niewidome organizują informacje składające się na budowane mapy:

- budowanie mapy na podstawie sekwencyjnie uporządkowanych punktów orientacyjnych, tzw. punktów-kotwic (*anchor-point theory*) (Golledge, 1999),
- budowanie mapy na podstawie sieci ścieżek i ich wzajemnej geometrycznej konfiguracji (Gärling i inni, 1981).

Stwierdzono, że pierwszy sekwencyjny sposób budowania mapy jest stosowany przede wszystkim przy zapamiętywaniu nowych ścieżek, a drugi (o przestrzennej równoległej strukturze) przy uszczegóławianiu wiedzy o poznanych wcześniej obszarach terenu. Keevin Lynch (1960) amerykański urbanista uważa, że poznane na różne sposoby mapy wyobraźniowe są zapamiętywane jako zbiór następujących elementów składowych: drogi, linie brzegowe (krawężniki, ściany budynków), obszary (np. plac, tereny zielone), węzły (np. punkty przecięcia dróg) oraz punkty orientacyjne.

Obecnie w anglojęzycznej literaturze poświęconej orientacji przestrzennej i mobilności osób niewidomych rozróżnia się dwa bliskoznaczne ale różne pojęcia:

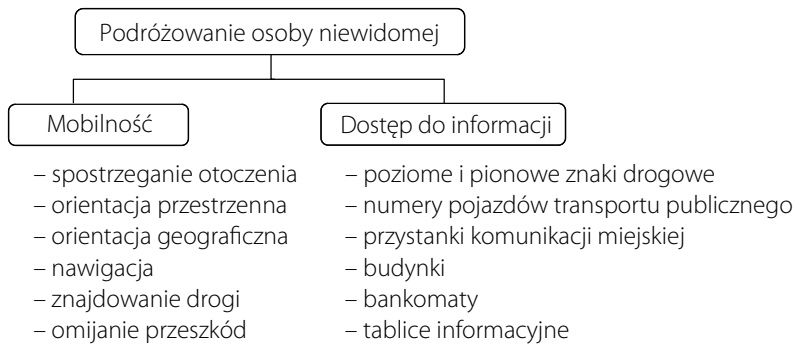
- *znajdowanie drogi (wayfinding)* – rozumiane jako zdolność do odnajdywania drogi do celu wśród sieci innych dróg (Golledge, 1999); jest to umiejętność ściśle zależna od wcześniej zapamiętanej mapy wyobraźniowej terenu i wymaga określonej aktywności kognitywnej osoby niewidomej (autorstwo tego pojęcia przypisuje się Lynchowi),
- *nawigacja (navigation)* – definiowana (w odniesieniu do poruszania się osób niewidomych) jako zaplanowane przemieszczanie się wg wybranej trasy od punktu startowego do celu; zadaniem nawigacji jest odpowiedź na trzy podstawowe zagadnienia nawigacji: Gdzie jestem? Gdzie znajduje się cel mojego marszu? Jak się tam dostać?

Dobrym sposobem rozróżnienia pojęć nawigacji i znajdowania drogi jest interpretowanie znajdowania drogi (*wayfinding*) jako czynności przede wszystkim kognitywnej, w odróżnieniu od nawigacji, którą należy rozumieć jako praktyczną (fizyczną) realizację planu przemieszczania się do celu. Niektórzy badacze dodatkowo dzielą nawigację na tzw. mikro-nawigację, jako czynność polegającą na omijaniu przeszkód oraz makro-nawigację jako przemieszczanie się z uświadomionym własnym położeniem na ścieżce względem punktu startowego i docelowego (Loomis i inni, 1999).

Zdefiniowane i opisane pojęcia związane z orientacją przestrzenną i mobilnością osób niewidomych nabierają dodatkowego znaczenia w kontekście pieszego przemieszczania się i podróżowania w mieście. Duży ruch kołowy i pieszy, hałas, duża liczba nietypowych przeszkód (nierówne chodniki, różnej wysokości krawężniki, słupki i znaki drogowe umieszczone w obszarze chodnika, ogrodzenia terenu bodowy), schody, przejścia podziemne powodują, że poruszanie się w środowisku miejskim jest szczególnie utrudnione i może być niebezpieczne dla osób niewidomych (Zadrozny, 2009).

W terenie miasta jest zatem uzasadnione rozszerzenie modelu lokomocji osoby niewidomej o zakres czynności związanych z pozyskiwaniem informacji, która powinna być dostępna dla osoby niewidomej w równym stopniu jak dla osób sprawnych wzrokowo. Na schemacie z rys. 1.12 uwzględniono zarówno zakres zadań związanych z mobilnością jak i dostępem do potrzebnej informacji w terenie miasta.

Ważną informacją dla projektantów urządzeń wspomagających mobilność osób niewidomych jest określenie głównych problemów i potrzeb osób z niepełnosprawnością wzroku odnośnie orientacji przestrzennej, pieszego przemieszczania się i podróżowania.



Rys. 1.12. Schemat podróży osoby niewidomej, w którym wyróżniono zakres aktywności związanych z mobilnością i wykaz informacji pomocnej w podróżowaniu (najczęściej niedostępnej dla osoby niewidomej)

Źródło: opracowanie własne

W ramach grantu badawczego dotyczącego systemów wspomaganie osób niewidomych w samodzielnym poruszaniu się realizowanego w Instytucie Elektroniki Politechniki Łódzkiej przeprowadzono ankietę z osobami niewidomymi i słabowidzącymi zrzeszonymi w Polskim Związku Niewidomych – Okręg Łódzki (IE PŁ, 2007). W ankiecie wzięło udział 17 osób (9 kobiet i 8 mężczyzn) w wieku od 24 do 75 lat. Osobom biorącym udział w ankiecie zadano pytania dotyczące podstawowych barier i uciążliwości związanych z samodzielnym poruszaniem się. Wyniki ankiet zebrano w tabelach 1.2÷1.4.

Tabela. 1.2. Odpowiedzi na pytanie ankietowe: *Jakie sytuacje życiowe są dla Pani/Pana najbardziej uciążliwe w związku z niepełnosprawnością wzrokową?*

Rodzaj uciążliwej sytuacji	Liczba osób sygnalizujących daną sytuację	Liczba osób wymieniających sytuację jako pierwszą
Poruszanie się w nieznanym otoczeniu	11	5
Przeszkody trwałe (lub tymczasowe) na chodnikach	11	
Brak dźwiękowej identyfikacji autobusów i tramwajów	8	2
Ograniczony dostęp do informacji	6	3
Brak oznakowania przedmiotów (np. produktów) i urządzeń	5	
Nieprawidłowo zaparkowane samochody	5	
Nierówności podłoża i schody w dół	5	

Tabela. 1.2. cd. Odpowiedzi na pytanie ankietowe: *Jakie sytuacje życiowe są dla Pani/Pana najbardziej uciążliwe w związku z niepełnosprawnością wzrokową?*

Rodzaj uciążliwej sytuacji	Liczba osób sygnalizujących daną sytuację	Liczba osób wymieniających sytuację jako pierwszą
Niekorzystne ukształtowanie podłoża (np. płynne zejście na ulicę)	5	
Brak pomocy ze strony osób widzących, brak wyrozumiałości kierowców i kultury jazdy	4	2
Brak sygnalizacji dźwiękowej na skrzyżowaniach (lub słabo czytelna)	4	
Poruszanie się po rozległych skrzyżowaniach (np. z torami tramwajowymi)	4	1
Nadmierny hałas, nadmierna akustyka pomieszczeń	4	
Konieczność nawigacji przy ubogiej informacji o punktach odniesienia	4	
Poruszanie się w nieznanych budynkach	2	
Nadmierne zagęszczenie przechodniów	2	
Zmiana natężenia oświetlenia (w przypadku poczucia światła)	1	1
Nieznane rozmieszczenie meblowania w nieznanych pomieszczeniach	1	
Złe warunki atmosferyczne, szczególnie śnieg	1	

Źródło: opracowanie własne

Tabela. 1.3. Odpowiedzi na pytanie ankietowe: *Jakie elementy nieznanego otoczenia należy zdaniem Pani/Pana sygnalizować aby zwiększyć bezpieczeństwo i komfort poruszania się?*

Rodzaj obiektu	Liczba osób sygnalizujących dany obiekt	Liczba osób wymieniających obiekt jako pierwszy
Niska przeszkoda: kosz, słupek, krzew	12	4
Słup, latarnia, znak drogowy	8	2
Krawężnik, nierówność chodnika	6	2
Zagłębienie w ulicy, uskok	7	

Tabela. 1.3. cd. Odpowiedzi na pytanie ankietowe: *Jakie elementy nieznanego otoczenia należy zdaniem Pani/Pana sygnalizować aby zwiększyć bezpieczeństwo i komfort poruszania się?*

Rodzaj obiektu	Liczba osób sygnalizujących dany obiekt	Liczba osób wymieniających obiekt jako pierwszy
Szyld, reklama wisząca na wysokości głowy	6	
Początek schodów w dół, schody	7	1
Zaparkowany samochód	4	1
Stoisko handlowe, ławka	4	
Miejsce użyteczne, przystanek (np. na bazie GPS)	4	
Kierunek wzdłuż ulicy lub korytarza	4	1
Przejście dla pieszych	3	
Drzwi	2	
Rozkład umeblowania	1	
Schody wychodzące na chodnik	2	
Nadjeżdżający samochód	2	1
Natężenie ruchu ulicznego	1	1
Ściany, mury i ich zakończenia	1	

Źródło: opracowanie własne

Tabela. 1.4. Odpowiedzi na pytanie ankietowe: *Czy jest Pani/Pan skłonny korzystać z urządzenia prezentującego dźwiękowo otoczenie w postaci okularów wyposażonych w miniaturowe kamery i słuchawki stereofoniczne?*

Zdecydowanie tak	6
Pod warunkiem estetycznego wyglądu i łatwości używania	5
Pod warunkiem znaczącego wzbogacenia informacji o otoczeniu i czytelności kodu	3
Zdecydowanie nie	3
Brak odpowiedzi	2

Źródło: opracowanie własne

Morston (2009) badał podstawowe problemy jakie niewidomi napotykają w znajdowaniu drogi i podróżowaniu w nieznanym terenie (m.in. w mieście). W badaniach brało udział 30 osób niewidomych z dobrą orientacją przestrzenną

i sprawnie poruszających się z białą laską lub z psem przewodnikiem. Uczestnicy badania oceniali stopień trudności przy wykonywaniu różnych zadań związanych ze znajdowaniem drogi i podróżowaniem środkami komunikacji miejskiej. Wyniki badań przeprowadzonych w USA zebrano w tabeli 1.5. Zwróćmy uwagę, że zadanie polegające na przejściu ruchliwej ulicy w nieznanej okolicy zostało uznane za jedno z łatwiejszych zadań do wykonania dla osób niewidomych biorących udział w testach.

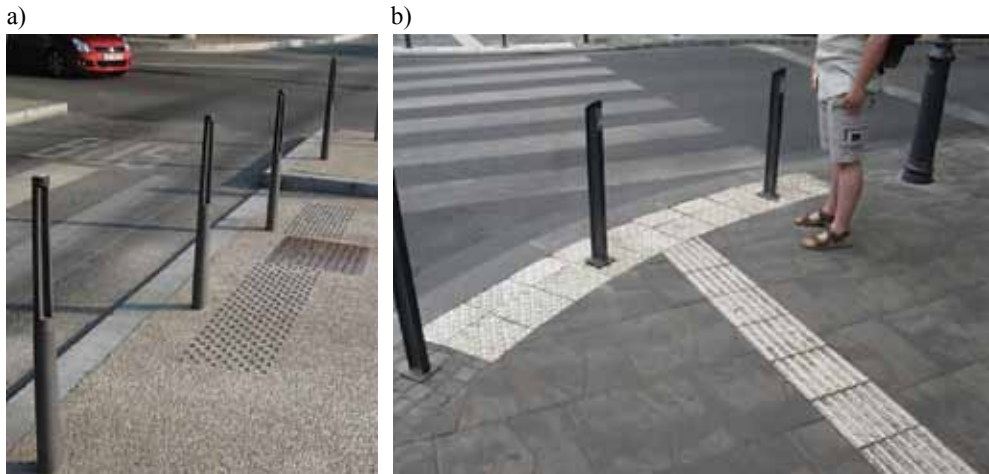
Tabela. 1.5. Różne zadania podróżowania i ocena stopnia trudności ich wykonania w nieznanym terenie miasta przez osobę niewidomą w skali 1÷5 (1 – najtrudniejsze, 5 – najłatwiejsze)

Zadanie podróżne	Ocena trudności
Odszukanie właściwego peronu i bramek wejściowych na perony na stacji kolejowej*	2,0
Uzyskanie tej samej dostępności do budynków i łatwości przemieszczania się pomiędzy nimi jaką mają osoby bez niepełnosprawności	2,3
Przesiadka z autobusu na autobus na ruchliwym terminalu autobusowym	2,3
Odszukanie okienka informacji, kas biletowych, toalet, budek telefonicznych na dworcu kolejowym lub autobusowym	2,3
Odszukanie przystanku autobusowego w mieście	2,3
Informacja o numerach autobusów zatrzymujących się na konkretnym przystanku	2,3
Poruszanie się w obszarze nieznanego dworca kolejowego lub autobusowego	2,4
Zdobycie informacji o pociągach lokalnych zatrzymujących się na peronie	2,5
Transfer z dworca i zmiana środka podróży	2,5
Wyjście z dworca i odszukanie przystanku taksówek	2,5
Odszukanie konkretnego stanowiska na dworcu autobusowym	2,6
Identyfikacja rogu ulic na skrzyżowaniu	2,7
Przesiadka z autobusu na autobus na przystanku	2,8
Informacja po której stronie peronu nadjedzie pociąg	2,9
Odszukanie drzwi do pociągu stojącego na peronie	3,2
Przejście przez ruchliwą ulicę w nieznanej okolicy	3,2

* Badania wykonano w Stanach Zjednoczonych Ameryki.

Źródło: opracowanie własne na podstawie (Morston, 2009)

W wielu miastach za granicą jak i w niektórych miastach polskich są wprowadzane jednostkowe udogodnienia w podróżowaniu osób niewidomych. Standardem staje się dźwiękowa sygnalizacja świateł drogowych. W pojazdach komunikacji miejskiej podróżni są informowani głosowo o kolejnych przystankach. Niewidomi są wyposażeni w nadajniki radiowe uaktywniające komunikat głosowy informujący o numerze nadjeżdżającego pojazdu komunikacji miejskiej. Przykład dobrych praktyk stosowanych w przestrzeni miejskiej zwiększających bezpieczeństwo poruszania się niewidomych w krytycznych punktach miasta, np. skrzyżowaniach ulic i przejść dla pieszych pokazano na rys. 1.13. W Polsce, tylko w niektórych dużych miastach takie udogodnienia są wprowadzane (Wysocki, 2009).



Rys. 1.13. Przykład umieszczania w powierzchni chodnika elementów (tzw. płyt kierunkowych i fakturowanych) ułatwiających osobie niewidomej rozpoznanie charakterystycznych miejsc w terenie miasta takich jak skrzyżowania i przejścia dla pieszych

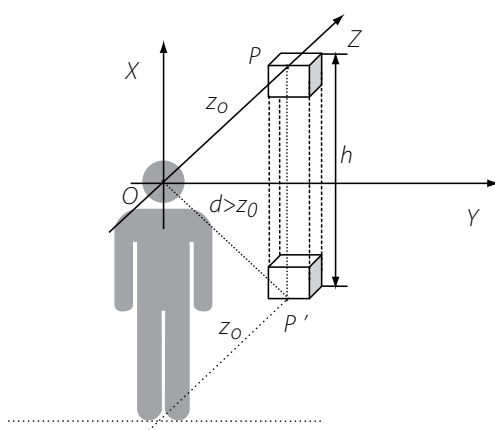
a) Lyon (2010 r.), b) Budapeszt (2011 r.)

Źródło: fotografie własne

Istnieje wiele prototypowych rozwiązań polegających na wbudowywaniu w infrastrukturę miejską systemów o architekturze rozproszonej, m.in. nadajników oraz czujników małego zasięgu (radiowych i komunikujących się światłem podczerwonym). Rozwijane są systemy informacji pasażerskiej udostępnionych dla użytkowników telefonów komórkowych. W coraz większym stopniu niewidomi korzystają ze specjalizowanych systemów nawigacji satelitarnej i map cyfrowych dla użytkowników pieszych. Podstawy techniczne osobistych systemów nawigacji projektowanych specjalnie dla osób z niepełnosprawnością wzroku omówiono w dalszych rozdziałach monografii.

Techniki estymacji głębi i modelowanie sceny trójwymiarowej

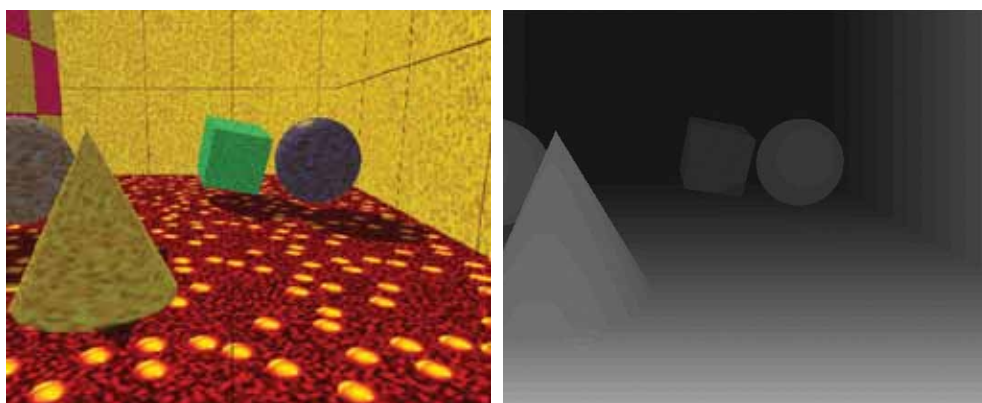
Zmysł wzroku człowieka posiada trudną do osiągnięcia w systemach technicznych zdolność do spostrzegania trójwymiarowej struktury otoczenia, tj. wyodrębniania obiektów, rozpoznawania ich kształtu, barwy (i tekstury), oceny ich wzajemnych relacji przestrzennych oraz oceny własnego położenia względem otoczenia. Te zdolności spostrzegania wzrokowego pozwalają człowiekowi sprawnie poruszać się w otoczeniu, m.in. przemieszczać się do celu wzdłuż wybranej ścieżki i skutecznie omijać przeszkody nieruchome i znajdujące się w ruchu. Umiejętności te są możliwe m.in. dzięki zdolności do spostrzegania głębi, tj. oceny własnej odległości (liczonej wzdłuż linii prostej, równoległej do podłoża) do najbliższych obiektów otoczenia (rys. 2.1).



Rys. 2.1. Widoczne na rysunku prostopadłości mają tę samą głębokość Z_0 w układzie współrzędnych $OXYZ$, a znajdują się w różnej odległości od jego początku

Źródło: opracowanie własne

Informacja o głębi jest podstawowym parametrem wykorzystywanym w systemach nawigacji robotów (m.in. robotów medycznych) (Tadeusiewicz, 1992), (Diamantas i inni, 2010) oraz systemach ostrzegania osób niewidomych o przeszkodach (Hersh i Johnson, 2008). Istnieją techniki rejestracji obrazów umożliwiające wyznaczanie tzw. map głębi (*depth map*). Przykład takiej mapy wyznaczonej dla komputerowego modelu sceny trójwymiarowej pokazano na rys. 2.2. Mapy takie określa się mianem tzw. obrazów dwu i półwymiarowych ($2\frac{1}{2}D$). Pomimo, że są one dane w postaci funkcji lub tablic dwuwymiarowych zawierają informację o trójwymiarowej strukturze obrazowanej sceny. Obrazy takie są jednak tylko uproszczonymi modelami sceny trójwymiarowej, gdyż na ich podstawie nie jest możliwe odtworzenie głębi obiektów (lub ich fragmentów) przesłoniętych przez inne obiekty sceny (Materka i Strumiłło, 2010). Ograniczenie to wynika z zastosowanej techniki obserwacji sceny trójwymiarowej za pomocą jednej kamery.



Rys. 2.2. Obraz wirtualnej sceny trójwymiarowej (z lewej) oraz tzw. gęsta mapa głębi (tzw. obraz $2\frac{1}{2}D$) wyznaczona dla tej sceny (z prawej); jasność punktu mapy oddaje głębię odwzorowanego punktu sceny (im bardziej odległy punkt od obserwatora tym ma mniejszą jasność)

Źródło: opracowanie własne

Najogólniej definiując, metody estymacji głębi można podzielić na metody kontaktowe i bezkontaktowe (Remondino i El-Hakim, 2006). Te pierwsze wymagają zastosowania urządzeń mechanicznych dających zwykle pomiar punktowy głębi, np. z zastosowaniem robota wyposażonego w ramię sterowane silnikami krokowymi. Kontaktową technikę pomiaru głębi jest zmuszona stosować osoba niewidoma wykorzystując białą laskę. W systemach technicznych jednoczesną, wielopunktową estymację głębi sceny trójwymiarowej uzyskuje się za pomocą

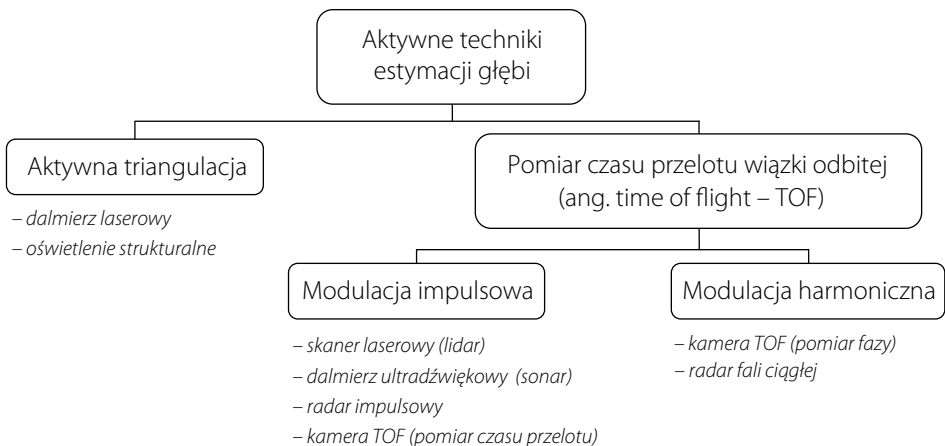
wielu technik bezkontaktowych, które można dalej podzielić na techniki aktywne, m.in. z zastosowaniem skanowania laserowego lub metod pasywnych wykorzystujących m.in. obrazowanie wizyjne.

2.1. Aktywne i pasywne techniki estymacji głębi

Aktywne techniki estymacji głębi polegają na zastosowaniu źródeł promieniowania wysyłających wiązki energii i pomiarze parametrów sygnałów odbitych od obiektów otoczenia. W technikach tych wykorzystuje się fale:

- dźwiękowe o szerokim zakresie długości, stosowane np. w sonarach (*sound navigation and ranging*),
- ultradźwiękowe (stosowane m.in. w ultrasonografach i dalmierzach ultradźwiękowych),
- radiowe, stosowane m.in. w radarach (*radio detection and ranging*),
- świetlne, których spójne i zbieżne wiązki są wytwarzane przez lasery i stosowane w lidarach (*light detection and ranging*),
- podczerwone, rzadziej obecnie wykorzystywane w systemach automatycznego ogniskowania aparatów fotograficznych.

Aktywne techniki estymacji głębi pokazane na rys. 2.3 pogrupowano wg rodzaju parametrów sygnałów odbitych od obiektów otoczenia. Zamieszczono również przykłady urządzeń do pomiaru głębi działających wg wymienionych metod.

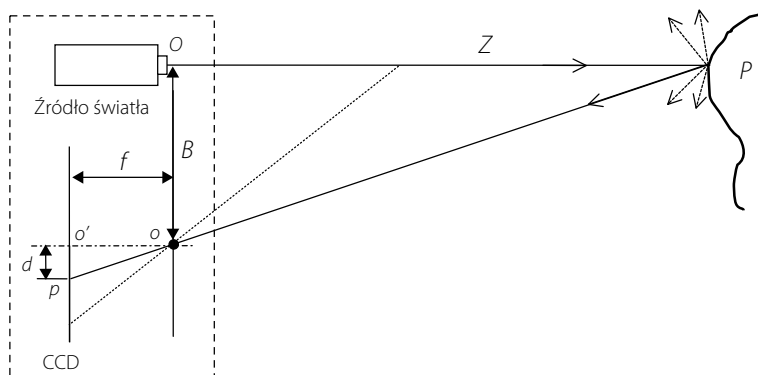


Rys. 2.3. Podstawowe techniki aktywnej estymacji głębi

Źródło: opracowanie własne

W metodach aktywnej triangulacji do wyznaczania głębi stosuje się prawa optyki geometrycznej. W metodach określanych wspólnie mianem pomiaru czasu przelotu (*time of flight – ToF*) wyróżnia się metody, w których emitowana wiązka jest impulsem (modulacja impulsowa) lub jest falą ciągłą (natężenie wiązki modulowane harmonicznie). W tej pierwszej pomiar głębi odbywa się na podstawie pomiaru czasu powrotu wiązki odbitej od elementów sceny, a w tej drugiej pomiarowi podlega opóźnienie fazy sygnału harmonicznego modulującego natężenie emitowanej wiązki.

Metodę aktywnej triangulacji wykorzystuje się w dalmierzach laserowych oraz w metodzie oświetlenia strukturalnego. W istocie, w obu tych technikach stosuje się zasadę pomiaru triangulacyjnego objaśnionego na rys. 2.4. W metodzie oświetlenia strukturalnego wysyłanych jest wiele wiązek światła, a w dalmierzach laserowych stosuje się pojedynczą wiązkę światła lasera. Na rys. 2.4 rozpatrzono przypadek dla pojedynczej wiązki światła. Emitowana wiązka oświetla punkt P obiektu otoczenia i odbija się w różnych kierunkach. Rejestracji podlega promień odbity w kierunku układu optycznego urządzenia. Promień ten pada na punkt p matrycy lub linijki CCD. Linia wyznaczona przez emitowaną wiązkę i oś optyczną urządzenia rejestrującego leżą w jednej płaszczyźnie ale niekoniecznie są równoległe (jak w przypadku pokazanym na rys. 2.4).



Rys. 2.4. Zasada pomiaru wg metody aktywnej triangulacji stosowanej w dalmierzach laserowych oraz w metodach oświetlenia strukturalnego; położenie punktu padania wiązki na przetwornik optoelektroniczny zależy od odległości oświetlanego obiektu

Źródło: opracowanie własne

Odległość Z punktu P od urządzenia pomiarowego wyznacza się na podstawie wzajemnego podobieństwa trójkątów OPo i $o'op$ (stąd nazwa metody). Z proporcji boków tych trójkątów otrzymujemy:

$$Z = \frac{fB}{d} \quad (2.1)$$

gdzie: f – jest ogniskową układu optycznego, B – jest odległością pomiędzy emitowaną wiązką światła a środkiem układu optycznego urządzenia rejestrującego oraz d – jest współrzędną obrazu punktu P na matrycy CCD.

Zakres pomiaru odległości w ręcznych dalmierzach laserowych stosowanych w systemach ostrzegania o przeszkodach dla osób niewidomych wynosi 10 cm÷10 m przy względnym błędzie pomiaru ok. 1% (Farcy i inni, 2006). Punktowy pomiar odległości (1-D) jest wykonywany z częstością ok. 40Hz. Wyznaczanie głębi dla szerszego zakresu kąтового wymaga ręcznego skanowania otoczenia. Zakłócenia pomiaru mogą wystąpić w warunkach istnienia innych źródeł światła (np. słonecznego) o dużym natężeniu (>130klx).

Techniki oświetlenia strukturalnego (*structured light*) są metodami wizyjnymi polegającymi na oświetleniu sceny zadaniem a priori regularnym wzorcem (linie lub siatki). Kamera znajduje się w dokładnie określonym położeniu względem źródła oświetlenia i rejestruje obrazy odkształconych wzorców utworzonych na powierzchni oświetlanych obiektów sceny. Dla uzyskania mapy głębi należy tu dodatkowo zastosować odpowiednie algorytmy analizy obrazów wzorców (Le Moigne i Waxman, 1988). Przykład obrazu sceny oświetlonej wzorcem w postaci siatki prostokątnej zilustrowano na rys. 2.5. Technika oświetlenia strukturalnego w praktyce nie znajduje zastosowania w systemach nawigacji lokalnej i urządzeniach ostrzegania o przeszkodach dla niewidomych.



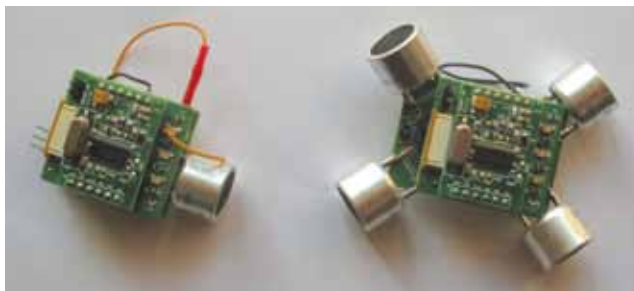
Rys. 2.5. Obraz sceny widziany w oświetleniu strukturalnym

Źródło: opracowanie własne

W drugiej grupie wymienionych metod aktywnej estymacji głębi (rys. 2.3) parametrem mierzonym jest czas przelotu wiązki promieniowania (laserowego, podczerwonego, ultradźwiękowego lub radiowego). Szeroko znanym urządzeniem, działającym wg tej zasady pomiaru jest radar impulsowy. Dla tej klasy urządzeń odległość jest wyznaczana na podstawie zależności: $Z = vt/2$, gdzie: v – jest prędkością propagacji wiązki promieniowania w powietrzu oraz t – jest czasem liczącym od chwili emisji impulsu wiązki do jej powrotu do urządzenia po odbiciu od obiektu otoczenia ($v \cong 340$ m/s dla fali dźwiękowej i $v \cong 3 \cdot 10^8$ m/s dla fali elektromagnetycznej). Dla przykładu, dla odległości $Z = 15$ m mierzonej za pomocą wiązki lasera czas powrotu wiązki wynosi $t \cong 10$ ns, a dla fali dźwiękowej $t \cong 88$ ms.

Urządzenia typu ToF wykonują pomiar punktowy głębi lub są wyposażone w odpowiedni system skanowania umożliwiający generowanie obrazów głębi, tzw. obrazów $2\frac{1}{2}D$ (Nowakowski i Podsędkowski, 2002).

W dalmierzach ultradźwiękowych, laserowych oraz dalmierzach światła podczerwonego uzyskuje się pomiar punktowy głębi z wykorzystaniem metody TOF z modulacją impulsową wiązki. Dalmierze tego typu są wykorzystywane w urządzeniach ostrzegania o przeszkodach montowanych w białej łasce niewidomego (zob. rozdz. 5.). Względny błąd pomiaru głębi w tych urządzeniach nie przekracza 1%. Wadą dalmierzy ultradźwiękowych jest rozbieżność wiązki fali dźwiękowej dla tzw. pola dalekiego (tj. zakresu głębi: $Z > d^2/\lambda$, gdzie: d – apertura źródła wiązki, λ – długość fali ultradźwiękowej o częstotliwości 40 kHz, tj. dla ultradźwięków $\lambda \cong 8,5$ mm (Dobrucki, 2007)). W polu dalekim kąt rozbieżności wiązki ultradźwięków wynosi $30^\circ \div 40^\circ$ (kąt przekroju osiowego stożka promieniowania). Na rys. 2.6 pokazano moduły elektroniczne wykonane w Instytucie Elektroniki Politechniki Łódzkiej, w których zastosowano metodę ToF do pomiaru odległości.



Rys. 2.6. Moduły układów elektronicznego nadajnika (po lewej) i odbiornika (po prawej) ultradźwiękowego do pomiaru wzajemnej odległości metodą ToF (Barański i Morański, 2007)

Źródło: fotografia własna

Pomiar głębi w szerszym zakresie kątowym wymaga zastosowania odpowiednich technik skanowania otoczenia za pomocą emitowanej wiązki promieniowania. W lidarach jest stosowane skanowanie mechaniczne (Borenstein i inni, 1996). Wiązka promieniowania laserowego odbija się od lustra obracającego się z zadaną częstotliwością (typowo 25÷35 Hz) i omiata otaczającą scenę. Lidary mają duże rozmiary i są stosowane głównie w geodezji, fotogrametrii i robotyce (Nowakowski i Podsędkowski, 2002). Charakteryzują się małym względnym błędem pomiaru odległości, tj. poniżej 0,1%.

Bardzo duży postęp odnotowuje się ostatnio w konstrukcji kamer ToF, w których jest stosowana zarówno modulacja impulsowa jak i harmoniczna natężenia emitowanej wiązki światła laserowego lub podczerwonego (www.mesa-imaging.ch). W kamerach tych scena jest oświetlana matrycą składającą się z kilkudziesięciu diod elektroluminescencyjnych, a odbite od elementów sceny światło jest rejestrowane przez kamerę wyposażoną w matrycę CCD lub CMOS (Lange i Seitz, 2001).

Zasada działania kamer ToF z modulacją impulsową jest prosta ale wymaga bardzo szybkich elektronicznych układów przełączających, a dla uzyskania dużej energii krótkiego impulsu są stosowane specjalnej konstrukcji diody elektroluminescencyjne lub źródła promieniowania laserowego (Piekara, 1976). W kamerach ToF z falą ciągłą jako źródeł oświetlenia stosuje się tanie diody elektroluminescencyjne. W kamerach tych natomiast bardziej złożona jest procedura pomiaru odległości nazywana techniką demodulacyjną, w której jest wyznaczane opóźnienie fazy modulowanej harmonicznie wiązki światła. Faza funkcji harmonicznej jest wyznaczana wg następującej procedury obliczeniowej (Lange, 2000).

Niech sygnał reprezentujący natężenie rejestrowanej wiązki światła jest dany zależnością:

$$x(t) = 1 + A \cos(\omega t - \varphi) \quad (2.2)$$

gdzie: $A < 1$ i ω – są odpowiednio amplitudą i pulsacją modulacji oraz φ jest opóźnieniem fazowym. Rejestrowany sygnał $x(t)$ jest demodulowany (korelowany) z wytworzonym w kamerze sygnałem referencyjnym $y(t) = 2 \cos(\omega t)$. Funkcja korelacji wzajemnej tych sygnałów jest dana w postaci:

$$c(\tau) = 2 \int_{-\infty}^{\infty} (1 + A \cos(\omega t - \varphi)) \cos(\omega t + \omega \tau) dt = A \cos(\varphi + \omega \tau) \quad (2.3)$$

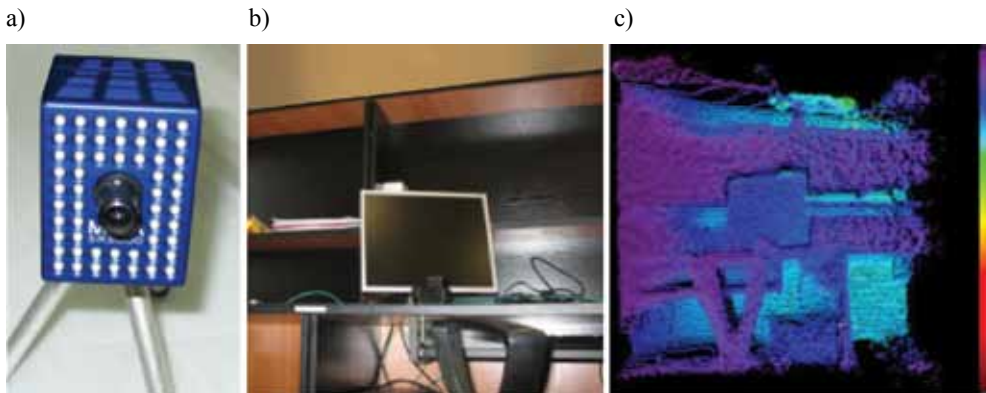
Funkcja korelacji (2.3) jest zatem funkcją okresową. Wyznamy wartości tej funkcji dla następujących faz $\omega \tau_0 = 0$, $\omega \tau_1 = \pi/2$, $\omega \tau_2 = \pi$, $\omega \tau_3 = 3\pi/2$ wtedy:

$$\begin{aligned}
c(\tau_0) &= A\cos(\varphi) \\
c(\tau_1) &= -A\sin(\varphi) \\
c(\tau_2) &= -A\cos(\varphi) \\
c(\tau_3) &= A\sin(\varphi)
\end{aligned}
\tag{2.4}$$

Odejmując odpowiednio parametry w równaniach i biorąc odpowiedni iloraz tych różnic, szukaną wartość opóźnienia fazowego dostajemy z zależności:

$$\varphi = \operatorname{atan}\left(\frac{c(\tau_3) - c(\tau_1)}{c(\tau_0) - c(\tau_2)}\right)
\tag{2.5}$$

Na rys. 2.7. pokazano zdjęcie kamery ToF z falą ciągłą, w której odległość jest wyznaczana wg opisanej procedury obliczeniowej. Na rys. 2.7b pokazano obraz sceny w świetle widzialnym, a na rys. 2.7c obraz tej sceny, tzw. mapę odległości uzyskaną z kamery ToF. Na zdjęciu jest widoczna matryca oświetlająca scenę, która składa się z 70 diod (7×10) emitujących wiązki światła podczerwonego o dł. fali 850 nm. Kamera rejestruje obrazy sceny z szybkością 50 obrazów na sekundę i mierzy odległość (a nie głębść) w zakresie 0,8÷10 m z błędem nie przekraczającym 15 mm. Podstawową zaletą kamer ToF jest możliwość rekonstrukcji sceny trójwymiarowej przy braku oświetlenia, a podstawową wadą jest zakłócanie pomiaru w warunkach dużego natężenia oświetlenia zewnętrznego oraz niemały pobór mocy (ok. 10 W).

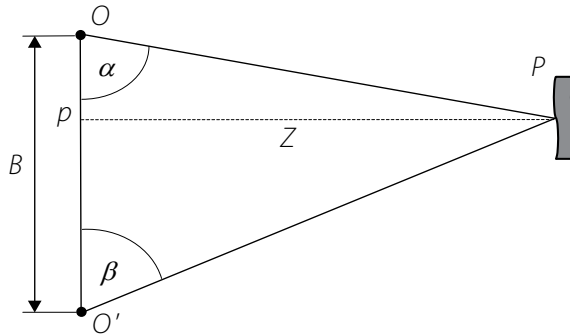


Rys. 2.7. Pomiar odległości za pomocą kamery ToF:

a) zdjęcie kamery ToF Swiss Ranger SR3000 firmy Mesa Imaging (rozmiar: $65 \times 65 \times 68$ mm, waga: ok. 0,5 kg), b) obraz sceny w świetle widzialnym, c) $2\frac{1}{2}$ wymiarowy obraz odległości, w którym skali odległości przyporządkowano paletę barw umownych

Źródło: opracowanie własne

W pasywnych technikach estymacji głębi wykorzystuje się naturalne oświetlenie sceny. Znaną od dawna i szeroko stosowaną metodą pomiaru głębi jest technika triangulacji pasywnej, w której otoczenie podlega obserwacji z co najmniej dwóch punktów o znanej wzajemnej odległości. Na rys. 2.8. punktami obserwacji są punkty O i O' oddalone od siebie o odległość B zwaną bazą układu.



Rys. 2.8. Ilustracja zasady pomiaru triangulacyjnego polegającego na wyznaczeniu kątów widzenia punktu P sceny z dwóch miejsc obserwacji wzajemnie oddalonych o znaną odległość B zwaną bazą

Źródło: opracowanie własne

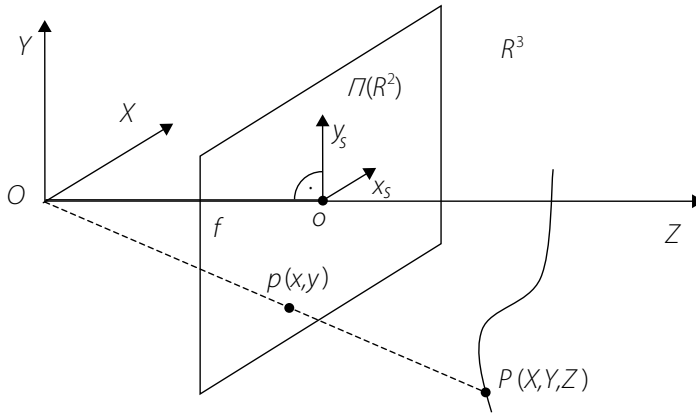
Zapisując funkcje trygonometryczne dla trójkątów OPp i $O'Pp$ (stąd nazwa metody zwana triangulacją) można wyznaczyć głębię obserwowanego punktu P sceny na podstawie znajomości bazy układu $B = |Op| + |pO'|$ i kątów α, β pod którymi jest widoczny punkt P sceny odpowiednio z punktów obserwacji O i O' :

$$Z = \frac{B}{\operatorname{ctg}\alpha + \operatorname{ctg}\beta} \quad (2.6)$$

Na podobnej zasadzie działają dalmierze optyczne.

We współczesnych, pasywnych systemach estymacji głębi (w szczególności urządzeniach mobilnych) wykorzystuje się przede wszystkim techniki obrazowania otoczenia za pomocą optycznych urządzeń rejestrujących nazywanych kamerami. Wyróżnia się metody obrazowania jednokamerowego lub obrazowania dwukamerowego (w niektórych zastosowaniach wielokamerowego).

Rozważmy model tzw. kamery perspektywicznej (*pin-hole camera*) pokazany na rys. 2.9, w którym środek rzutowania O kamery (tzw. ognisko) jest umieszczony w początku układu współrzędnych $OXYZ$, a płaszczyzna rzutowania π jest prostopadła do osi OZ (tzw. głównej osi kamery) i przecina ją wyznaczając punkt o nazywany punktem głównym płaszczyzny rzutowania.

**Rys. 2.9.** Model kamery perspektywicznej

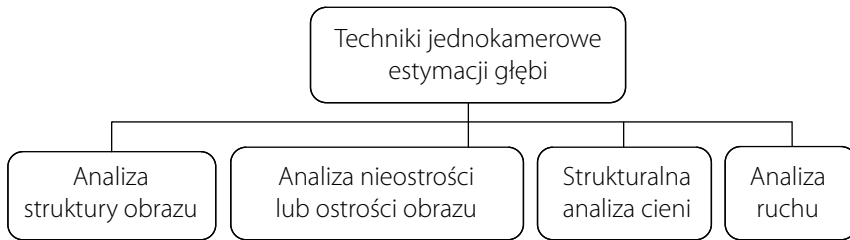
Źródło: opracowanie własne

Płaszczyzna rzutowania jest umieszczona w odległości f (zwanej ogniskową kamery) od środka rzutowania. Zgodnie z zasadą transformacji perspektywicznej, każdy widoczny ze środka rzutowania punkt sceny trójwymiarowej $P(X, Y, Z)$ jest odwzorowywany w punkt $p(x, y)$ w ten sposób, że punkt p jest przecięciem płaszczyzny rzutowania i prostej łączącej środek rzutowania z punktem sceny P . Współrzędne punktów P i p wiąże zależność nieliniowa:

$$x = f \frac{X}{Z}, \quad y = f \frac{Y}{Z} \quad (2.7)$$

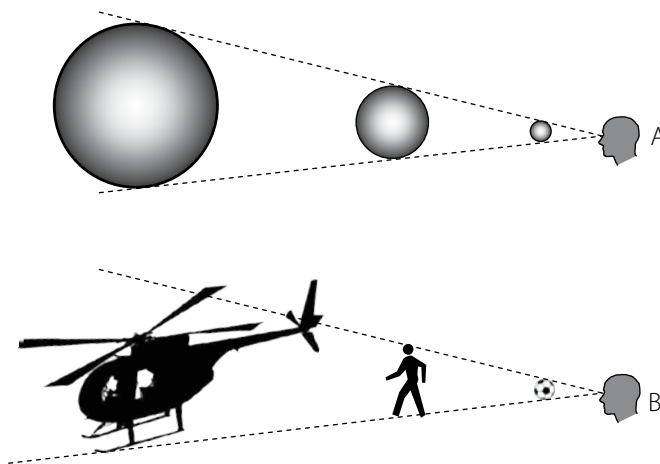
Istnieje nieskończenie duża liczba par współrzędnych (X, Z) i (Y, Z) spełniających ten układ dwóch równań o trzech niewiadomych. Zatem na podstawie współrzędnych punktu $p(x, y)$ nie można jednoznacznie wyznaczyć położenia obrazowanych punktu sceny $P(X, Y, Z)$. Wnioskujemy, że za pomocą pojedynczej kamery estymacja głębi na podstawie geometrycznych zależności wynikających z transformacji perspektywicznej nie jest możliwa. Opracowano jednak szereg metod (rys. 2.10), w których na podstawie analizy pojedynczego obrazu sceny jest możliwa ocena względnych i bezwzględnych wartości głębi.

W metodach estymacji głębi na podstawie struktury sceny analizie podlegają cechy globalne obrazów takie jak perspektywa (w tym obszary tekstur), wzajemne przesłanianie obiektów oraz jest wykorzystywana wiedza a priori o obrazowanych obiektach sceny (Saxena i inni, 2007). Metody te wymagają jednak zastosowania złożonych algorytmów przetwarzania obrazów, wydzielania obiektów sceny oraz ich analizy semantycznej (Torralba i Oliva, 2002) (rys. 2.11).



Rys. 2.10 Podstawowe techniki estymacji głębi za pomocą technik obrazowania jednokamerowego

Źródło: opracowanie własne



Rys. 2.11. Obserwator A nie może ocenić głębi dla poszczególnych obiektów bez wiedzy *a priori* o ich rozmiarach; obserwator B dzięki możliwej kategoryzacji semantycznej obiektów może ocenić głębie dla obiektów sceny

Źródło: opracowanie własne na podstawie (Torralba i Oliva, 2002)

W metodach estymacji głębi na podstawie ostrości lub nieostrości obrazu (*depth from focus or defocus*), wykorzystuje się właściwość układu optycznego kamery nazywanego głębią ostrości, tj. wycinkiem głębi sceny, w której obiekty sceny są odwzorowywane w postaci ostrych (nierozmytych) obrazów (rys. 2.12).

W metodach analizy ostrości obrazu wymagany jest taki dobór parametrów optycznych obiektywu kamery aby uzyskać najlepszą ostrość obrazowania wybranego obiektu sceny (Ens i Lawrence, 1993). Jest to metoda czasochłonna. Jest jednak wykorzystywana w systemach automatycznego ustawiania ostrości w aparatach fotograficznych. Znacznie skuteczniejszymi technikami estymacji głębi, ale wymagającymi bardziej złożonych obliczeń, są metody analizy nieostrości obrazu. W metodach tych parametry układu optycznego kamery są tak

dobierane aby uzyskać skończoną głębię ostrości. Nieostrość obrazu jest modelowana za pomocą tzw. funkcji rozmycia punktu (*point spread function*) danej funkcją cylindryczną (Materka i Strumiłło, 2009):

$$h(x, y) = \begin{cases} 1, & x^2 + y^2 \leq R^2 \\ 0, & x^2 + y^2 > R^2 \end{cases} \quad (2.8)$$

gdzie: x, y – są współrzędnymi punktów obrazu, a R – jest promieniem podstawy cylindra.

W prekursorskiej pracy Pentland'a (1982) wykazano, że przy dostępnej wiedzy *a priori* o kształcie obrazowanego obiektu można skutecznie estymować głębię na podstawie jednokrotnej rejestracji obrazu i analizy jego nieostrości. Tematyka ta jest nadal rozwijana m.in. w pracach Stasiaka i Materki (2003) oraz Leroya i innych (2008). W tej pierwszej zastosowano sztuczną sieć neuronową do estymacji głębi w szerokim zakresie, a w drugiej estymację głębi uzyskano na podstawie porównania rozmycia dwóch obrazów rejestrowanych z tego samego miejsca ale przy różnych nastawach obiektywu kamery. Uzyskiwane względne błędy estymacji głębi wynoszą w tych metodach kilka procent.



Rys. 2.12. Przykład obrazu zarejestrowanego dla skończonego zakresu głębi ostrości, stopień rozmycia obrazów obiektów sceny (kłosy skrzypu polnego) zależy od głębi

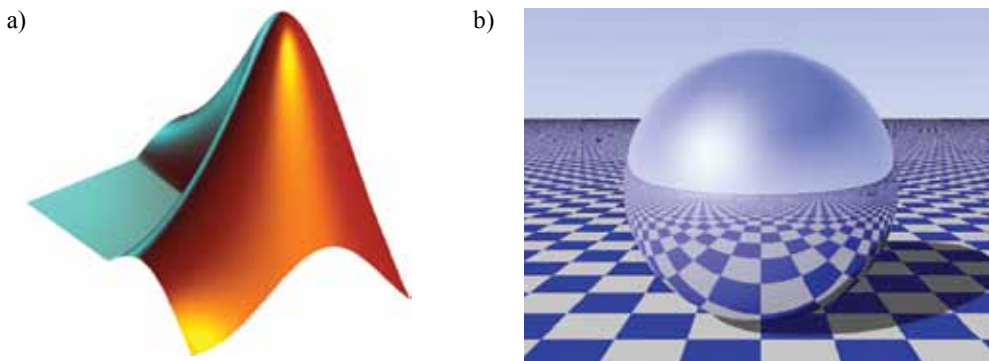
Ź r ó d ł o: (Materka i Strumiłło, 2009)

Metody polegające na analizie cieni (*shape from shading – SFS*) w obrazowanych scenach są skuteczne przy odpowiednim oświetleniu sceny i dla zawężonej klasy obrazowanych scen (rys. 2.13). Bardzo dobry przegląd tych technik analizy

struktury sceny i estymacji głębi jest podany w pracy (Zhanga i innych, 1999). Ważnym ograniczeniem tych metod jest konieczność przyjęcia założenia o lambertowskim charakterze modelowanych kształtów powierzchni¹. Na podstawie tego założenia i lokalnego rozkładu jasności obrazu:

$$C(x, y) = \mathbf{L} \mathbf{N} = |\mathbf{L}| |\mathbf{N}| \cos \alpha \quad (2.9)$$

można śledzić kształt powierzchni obiektu, tj. zmianę kierunku wektora normalnego $\mathbf{N}$ dla wycinka analizowanej powierzchni obiektu, gdzie $\mathbf{L}$ – jest wektorem wskazującym promieni oświetlających scenę. W niektórych sytuacjach znak zmiany kierunku wektora normalnego jest trudny do jednoznacznego wyznaczenia. W pracy Roblesa-Kelly i Hancocka (2007) skutecznie rozwiązano ten problem przez zastosowanie analogii fizycznej, w której zmiana kierunku wektora normalnego $\mathbf{N}$ jest skutecznie śledzona przez zastosowanie równania przewodnictwa cieplnego definiowanego dla obszarów modelowanych powierzchni obiektów sceny.



Rys 2.13. Przykład obrazu sceny trójwymiarowej: a) – której struktura (i głębia) może być skutecznie modelowana na podstawie strukturalnej analizy cieni oraz obrazu obiektu b) – którego głębia nie może być wiarygodnie estymowana metodą SFS (powierzchnie nielambertowskie)

Źródło: a) opracowanie własne z wykorzystaniem pakietu Matlab, b) Wikimedia Commons

Modelowanie struktury sceny na podstawie wzajemnego przemieszczenia kamery i obiektów sceny jest trudnym problemem obliczeniowym wymagającym rejestracji sekwencji obrazów i ich analizy (Ma i inni, 2003). Rozwiązanie tego problemu wiąże się z przeprowadzeniem trzech etapów analizy obrazów.

¹ Powierzchnie matowe tzw. lambertowskie rozpraszają światło równomiernie we wszystkich kierunkach w odróżnieniu od powierzchni nielambertowskich wykazujących kierunkowość rozpraszanego lub odbijanego światła.

1. Wyznaczenie przemieszczenia obrazów obiektów w płaszczyźnie rzutowania kamery na skutek wzajemnego ruchu kamery i obiektów sceny.
2. Powiązania ruchu obiektów w scenie trójwymiarowej z przemieszczeniem ich obrazów w płaszczyźnie rzutowania kamery.
3. Wyznaczenie struktury sceny i parametrów głębi.

Pierwsze zagadnienie wiąże się z pojęciem przepływu optycznego (*optical flow*), tj. wyznaczaniem chwilowych prędkości przemieszczania wskazanych punktów obrazu (tzw. punktów kluczowych, np. brzegów obiektów, przecięć linii). Szybkości te można powiązać z ruchem obiektów sceny wykorzystując równania rzutu perspektywicznego (2.7):

$$u = \frac{dx}{dt} = f \left(\frac{dX}{dt} \frac{1}{Z} - \frac{X}{Z^2} \frac{dZ}{dt} \right) \quad (2.10)$$

$$v = \frac{dy}{dt} = f \left(\frac{dY}{dt} \frac{1}{Z} - \frac{Y}{Z^2} \frac{dZ}{dt} \right) \quad (2.11)$$

Zależności (2.10-2.11) pokazują, że związek pomiędzy przepływem optycznym danym polem wektorowym $\vec{P}(x,y) = (u,v)$ przyporządkowującym każdemu punktowi obrazu sceny wektor prędkości jest dany niedookreślonym układem równań nieliniowych (dane: u, v , szukane: $dX/dt, dY/dt, dZ/dt$). Jednak uwzględniając zasady rzutu perspektywicznego (oraz często dostępną wiedzę *a priori* o scenie lub ruchu kamery) można na podstawie obrazu pola wektorowego estymować strukturę przestrzenną sceny i głębię obiektów. Na rys. 2.14 pokazano przykład dwóch klatek obrazu i wyznaczony na ich podstawie przepływ optyczny. Wielkość tego przepływu jest parametrem wykorzystywanym do wyznaczenia głębi obiektów w scenie.



Rys. 2.14. Obrazy tzw. sekwencji garden (www.m4if.org) oraz obraz przepływu optycznego wyznaczony dla tych obrazów

Źródło: opracowanie własne z wykorzystaniem programu VcDemo (<http://ict.tudelft.nl/vcdemo>)

Podstawowymi metodami stosowanym do estymacji pola przepływu optycznego są techniki gradientowe oraz klasyczne metody zaproponowane przez Horna i Schuncka (1981) oraz Lucasa i Kanadego (1981). Metoda Horna-Schuncka jest podejściem regularyzacyjnym nakładającym warunek „gładkości” ruchu w obszarze obrazu, a metoda Lucasa-Kanadego jest metodą lokalną, w której przyjmuje się stałość przepływu optycznego w pewnym zadanym otoczeniu analizowanego punktu obrazu. Więcej szczegółowych informacji nt estymacji przepływu optycznego zawiera praca (Fleet i Weiss, 2005).

Warto zaznaczyć, że równania przepływu optycznego są wykorzystywane w standardach MPEG (*Moving Picture Experts Group*) kompresji obrazów wideo (Sayood, 2002). Zamiast przysyłać kolejne ramki obrazu, można dla dużych obszarów obrazów przysyłać tylko informacje o przepływie optycznym wyznaczonym dla tych obszarów i znacznie skuteczniej kodować sekwencje obrazów.

2.2. Wybrane zagadnienia stereowizji

Stereowizja jest techniką wizyjną umożliwiającą wyznaczanie współrzędnych punktów sceny trójwymiarowej na podstawie ich obrazów uzyskiwanych w co najmniej dwóch urządzeniach rejestrujących, umieszczonych w różnych punktach sceny (Hartley i Zisserman, 2006), (Cyganek, 2002). Kluczową własnością obrazowania stereowizyjnego jest możliwość wyznaczenia współrzędnych przestrzennych obrazowanych punktów sceny trójwymiarowej na podstawie następujących danych:

- parametrów optycznych kamer, tzw. parametrów wewnętrznych kamer, na które składają się m.in.: ogniskowa, liczba punktów matrycy obrazu, jej rozmiar oraz parametry zniekształceń geometrycznych torów optycznych kamer,
- informacji o wzajemnym położeniu przestrzennym kamer, tzw. parametrów zewnętrznych układu stereowizyjnego definiowanego wektorami translacji i obrotu układów współrzędnych związanych z każdą z kamer,
- współrzędnych punktów w płaszczyznach obrazowych kamer będących odwzorowaniami tego samego punktu $P(X, Y, Z)$ sceny trójwymiarowej.

Dalsze rozważania dotyczą układu stereowizyjnego składającego się z dwóch kamer. Do rozwiązania zagadnienia rekonstrukcji sceny trójwymiarowej, tj. wyznaczenia współrzędnych przestrzennych wszystkich punktów sceny, które znajdują się w polu widzenia obu kamer należy wykonać następujące kroki obliczeniowe:

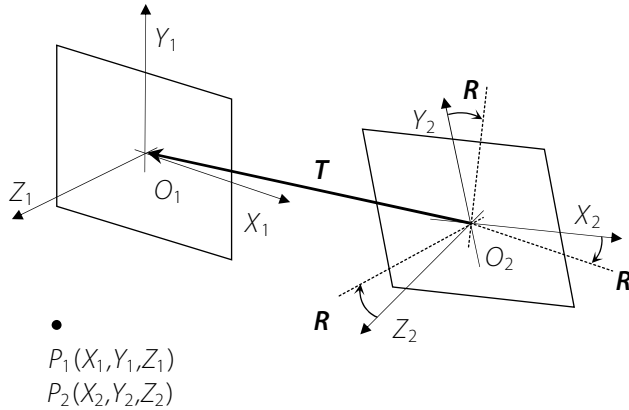
- 1) kalibrację torów optycznych kamer (korekcja zniekształceń geometrycznych i tzw. rektyfikacja kamer),
- 2) rozwiązanie zagadnienia odpowiedniości (*correspondence problem*), tj. odszukanie par współrzędnych w obrazie lewej i prawej kamery, które są rzutem perspektywicznym tych samych punktów sceny trójwymiarowej (Brown i inni, 2003),
- 3) rekonstrukcję głębi sceny trójwymiarowej, polegającą na wyznaczeniu tzw. gęstej mapy głębi (*dense depth map*), tj. obrazu, zwanego też obrazem dwu i pół wymiarowym ($2\frac{1}{2}D$), w którym poszczególnym punktom przyporządkowano wartość głębi obrazowanego punktu sceny.

Przez kalibrację torów optycznych kamer rozumie się znalezienie tzw. parametrów wewnętrznych i zewnętrznych kamer (Cyganek, 2002).

Znajomość parametrów wewnętrznych kamer układu stereowizyjnego jest konieczna do korekcji zniekształceń geometrycznych wprowadzanych przez układy optyczne kamer. Opracowano skuteczne metody korekcji zniekształceń geometrycznych, które polegają na określeniu klasy przekształceń wiążących współrzędne idealnej kamery perspektywicznej i kamery ze zniekształceniami geometrycznymi (Fusiello i inni, 2000). W większości zaproponowanych metod stosuje się przekształcenie modelujące te zniekształcenia za pomocą wielomianów drugiego lub trzeciego stopnia. W procedurze korekcji zniekształceń geometrycznych są identyfikowane współczynniki wielomianów określające niedokładności w wykonaniu soczewki obiektywu kamer (tzw. zniekształcenia radialne) oraz współczynniki określające niedokładność montażu matrycy kamery względem osi optycznej obiektywu kamery (tzw. zniekształcenia tangencjalne).

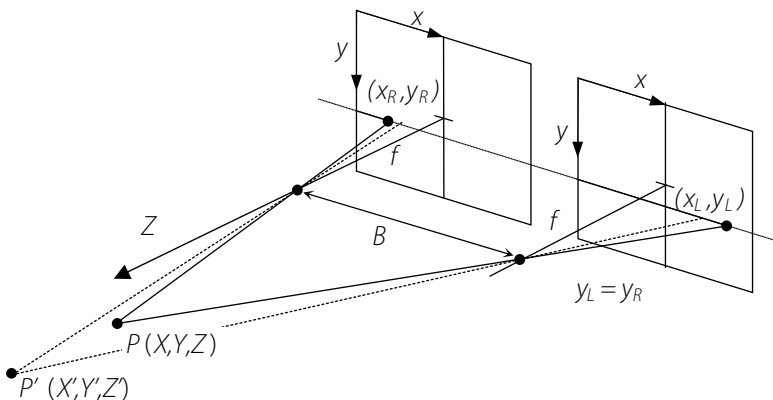
Parametry zewnętrzne kamer układu stereowizyjnego określają wzajemne położenie kamer względem siebie. Są one definiowane przez trójelementowy wektor translacji $\mathbf{T}$ i macierz obrotu $\mathbf{R}$ o rozmiarze 3×3 (rys. 2.15). Wektor $\mathbf{T}$ określa przesunięcie między środkami O_1 i O_2 układu współrzędnych związanych z kamerami, a macierz $\mathbf{R}$ określa obrót między osiami tych układów współrzędnych (Hartley i Zisserman, 2006). Jeżeli ten sam punkt sceny jest dany jednocześnie w układach współrzędnych $O_1X_1Y_1Z_1$ i $O_2X_2Y_2Z_2$, tj. $\mathbf{P}_1 = [X_1, Y_1, Z_1]^T$ i $\mathbf{P}_2 = [X_2, Y_2, Z_2]^T$ to parametry zewnętrzne kamery 1 względem kamery 2 określają związek pomiędzy współrzędnymi $\mathbf{P}_1$ i $\mathbf{P}_2$ wg zależności:

$$\mathbf{P}_1 = \mathbf{R} \times \mathbf{P}_2 + \mathbf{T} \quad (2.12)$$

**Rys. 2.15.** Układ nieskalibrowanych kamer stereowizyjnych

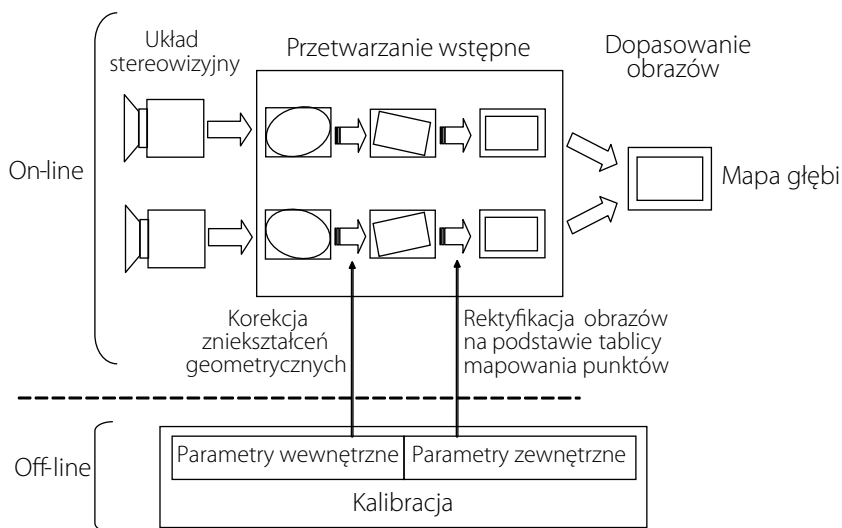
Źródło: opracowanie własne

Na podstawie obliczonych parametrów torów optycznych układu stereowizyjnego oraz parametrów zewnętrznych układu dla obrazów z każdej z kamer są wyznaczane tzw. tablice mapowania punktów obrazu. Rozmiar tych tablic jest równy rozmiarowi obrazu, a elementami tablic są pary liczb określające poprawkę współrzędnych $(\delta x, \delta y)$ dla każdego punktu odpowiednio lewego i prawego obrazu. Zastosowanie tych poprawek przekształca kolejno rejestrowane pary obrazów do tzw. układu kanonicznego, tj. obrazów po korekcji zniekształceń geometrycznych i dla których hipotetyczne płaszczyzny przetworników kamer znajdują się w tej samej płaszczyźnie, a wiersze obrazów są współliniowe i zgodne co do indeksów wierszy (rys. 2.16).

**Rys. 2.16.** Obrazy stereowizyjne rejestrowane w układzie kanonicznym kamer

Źródło: opracowanie własne

W pracy (Rzeszotarski i inni, 2005) opisano szczegółowo procedurę identyfikacji parametrów wewnętrznych i zewnętrznych układu stereowizyjnego, której schemat blokowy pokazano na rys. 2.17. W procedurze tej macierze przekształceń perspektywicznych niezbędnych do przeprowadzenia kalibracji zostały wyznaczone zgodnie z metodą opisaną w pracy (Fusiello i inni, 2000), dostępnej również na stronie internetowej: <http://profs.sci.univr.it/~fusiello/demo/rect/>.



Rys. 2.17. Schemat blokowy procedury kalibracji układu stereowizyjnego składającej się z korekcji zniekształceń geometrycznych i rektyfikacji kamer
Źródło: opracowanie własne na podstawie (Rzeszotarski i inni, 2005)

Na rys. 2.17 wyróżniono dwa moduły schematu blokowego procedury rektyfikacji układu stereowizyjnego. Procedury modułu *off-line* są realizowane przed docelowym zastosowaniem układu stereowizyjnego. Wyznaczona *off-line* tablica mapowania punktów jest wykorzystywana *on-line* do kalibracji układu kamer, tj. przekształcania rejestrowanych par obrazów do układu kanonicznego.

Na rys. 2.18 pokazano jeden z układów stereowizyjnych wykorzystywany do weryfikacji procedur kalibracji.

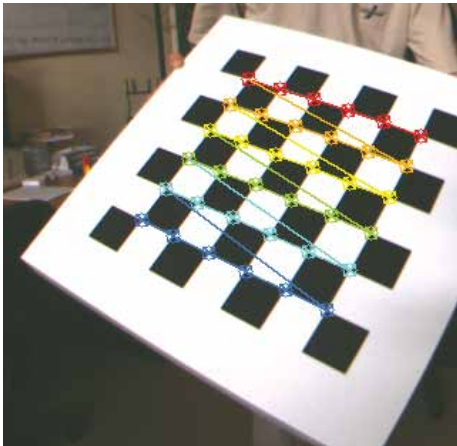
Na rys. 2.19 zilustrowano wynik kalibracji systemu stereowizyjnego z rys. 2.18 po korekcji zniekształceń geometrycznych układu optycznego kamer. Do kalibracji kamer zastosowano obiekt kalibracyjny w postaci tablicy z szachownicą o rozmiarze pojedynczego pola 60×60 mm.



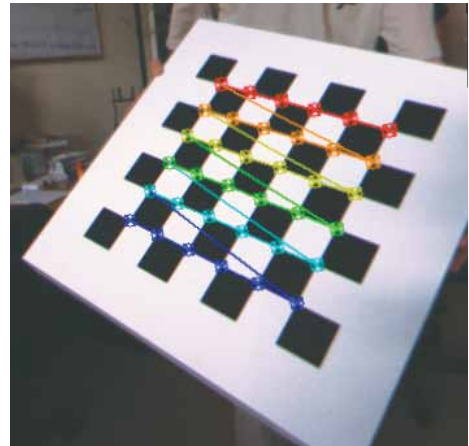
Rys. 2.18. Widok testowego układu stereowizyjnego o bazie $B = 8$ cm; zastosowano kamery cyfrowe Flea firmy Pointgrey (www.pointgrey.com); parametry kamery: matryca CCD o rozmiarze 1/3 cala i rozmiarze pojedynczej komórki (piksele): $\Delta_x \cong 8,8 \mu\text{m}$, maksymalna rozdzielczość: 1024×768 , maksymalna szybkość klatkowa: 30 fps (*frames per second*), ogniskowa obiektywu: $f = 3,5$ mm

Źródło: fotografia własna

a)



b)

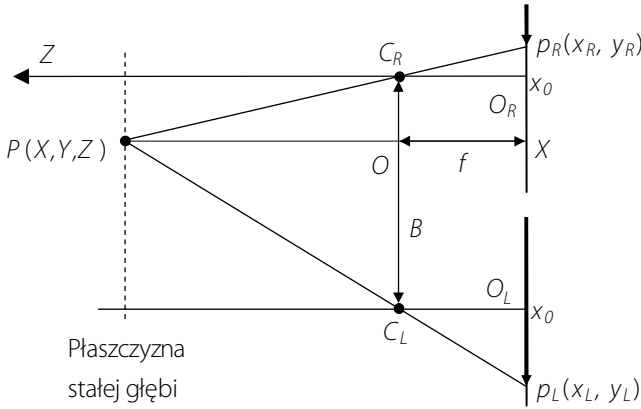


Rys. 2.19. Obraz tablicy kalibracyjnej przed korekcją zniekształceń geometrycznych układu optycznego kamery (a) oraz obraz tej samej tablicy po korekcji zniekształceń geometrycznych (b); zauważmy prawidłowe odwzorowanie prostych krawędzi tablicy kalibracyjnej

Źródło: opracowanie własne

Uzyskany w procedurze kalibracyjnej układ kanoniczny kamer ma tę właściwość, że obrazowane punkty sceny trójwymiarowej (rys. 2.16) są rzutowane na te same wiersze w obrazie kamery lewej i kamery prawej, tj. $y_L = y_R$. Można zauważyć, że $x_L \neq x_R$ i różnica tych współrzędnych maleje wraz z oddalaniem się obrazowanych punktów sceny. Wielkość $d = x_L - x_R$ nazywamy przesunięciem lub dysparycją (*disparity*), dla głębi $Z \rightarrow \infty$ dysparycja $d \rightarrow 0$.

Głębina obrazowanego punktu sceny może być zatem estymowana na podstawie dysparycji, tj. wielkości przesunięcia pomiędzy rzutami perspektywicznymi tego punktu w obrazach kamery lewej i prawej. Dla kamer w układzie kanonicznym związek pomiędzy dysparycją a głębią można wyznaczyć z prostych zależności geometrycznych zilustrowanych na rys. 2.20, w którym punkt główny prawej kamery jest umieszczony w początku układu współrzędnych O_XZ .



Rys. 2.20. Para obrazów stereowizyjnych w układzie kanonicznym kamer
Źródło: opracowanie własne

Związane z obrazem prawym trójkąty POC_R i $C_R O_R p_R$ oraz odpowiednio trójkąty POC_L i $C_L O_L p_L$ związane z obrazem lewym są podobne. Z proporcji długości boków tych trójkątów otrzymujemy:

$$\frac{X}{Z-f} = \frac{x_0 - x_R}{f} \quad (2.13)$$

$$\frac{B-X}{Z-f} = \frac{x_L - x_0}{f} \quad (2.14)$$

gdzie: B – odległość między osiami optycznymi kamer, tzw. *baza układu stereowizyjnego*, f – ogniskowa kamer, x_0 – lokalne współrzędne punktów głównych środka obrazów.

Po podzieleniu równań (2.13) i (2.14) stronami oraz dodaniu tych równań stronami można wyznaczyć współrzędne X i Z punktu P sceny:

$$X = \frac{B(x_O - x_R)}{d} \quad (2.15)$$

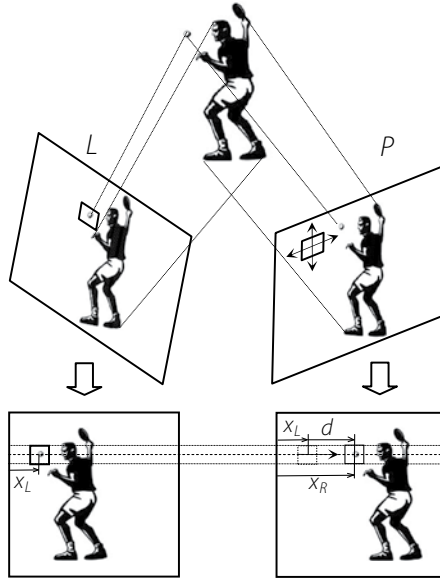
$$Z = \frac{Bf}{d} + f \cong \frac{Bf}{d} \quad (2.16)$$

gdzie: d – dysparycja: $d = x_L - x_R$ wyrażona w jednostkach fizycznych odległości. W równaniu (2.16) pominęto ogniskową kamery f , z uwagi na jej pomijalnie małą wartość (pojedyncze milimetry) w porównaniu z estymowanymi wartościami głębi. Stosując podobne własności geometryczne jak dla równań (2.13–2.14) można uzyskać zależność dla współrzędnej Y pamiętając, że dla kanonicznego układu kamer $y_L = y_R = y$.

$$Y = \frac{B(y - y_0)}{d} \quad (2.17)$$

Wyznaczenie przesunięcia (dysparycji) i obliczenie współrzędnych przestrzennych X , Y , Z przy obrazowaniu pojedynczego punktu sceny trójwymiarowej, jak w modelowym przykładzie z rys. 2.20, jest zadaniem łatwym. Przy obrazowaniu rzeczywistych scen trójwymiarowych należy rozwiązać tzw. problem odpowiedniości (korespondencji), tj. odszukania dla każdego punktu obrazu w wybranej kamerze (np. lewej) odpowiadającego mu punktu w obrazie prawym, takiego, że obydwa te punkty są rzutami perspektywicznymi tego samego punktu sceny trójwymiarowej. Rozwiązanie tego problemu znacznie upraszcza się dla kanonicznego układu kamer (rys. 2.21).

Opracowano skuteczne metody rozwiązywania problemu odpowiedniości dla obrazów stereowizyjnych (Brown i inni, 2003). Szeroki przegląd tych metod, które dzielą się na metody tzw. lokalne i globalne, zawarto w pracy (Scharstein i Szeliski, 2002). W systemach przetwarzania sekwencji obrazów, jako szybsze, znajdują zastosowania metody lokalne. Te z kolei dzielą się na metody dopasowania bloków i metody dopasowania cech obrazów. Mapę głębi dla wszystkich punktów obrazu, tzw. gęstą mapę głębi, można wyznaczyć za pomocą metod blokowych. Aby wyznaczyć taką mapę należy rozwiązać problem odpowiedniości dla wszystkich punktów obrazu sceny. Rozwiązanie tego problemu przebiega wg schematu pokazanego na rys. 2.21. Dla niekanonicznego układu kamer należy przeszukiwać całe pole obrazu prawego dla wyznaczenia położenia



Rys. 2.21. Rozwiązanie problemu odpowiedności dla niekanonicznego układu kamer (górna część rysunku) oraz dla układu kanonicznego kamer (dolna część rysunku)

Źródło: opracowanie własne

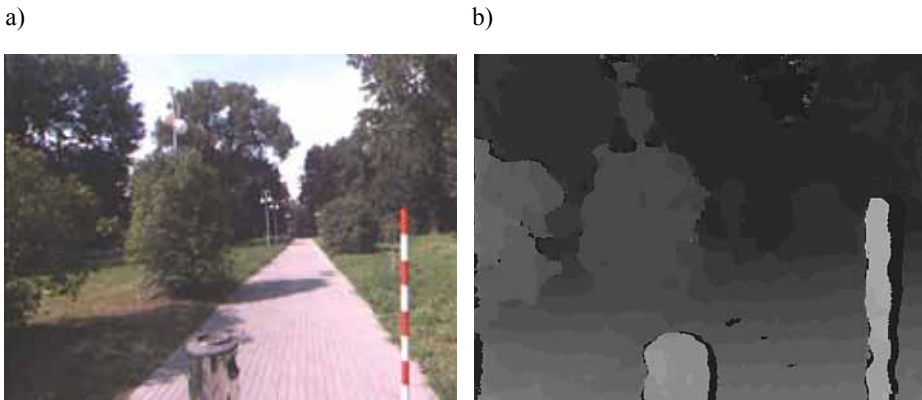
obiektu wskazanego w obrazie lewym. Rozwiązanie problemu odpowiedności upraszcza się dla układu kanonicznego kamer, w którym poszukiwanie obiektu sceny wystarczy przeprowadzić w obrazie prawym tylko wzdłuż wierszy obrazu, tj. dla każdego położenia bloku x_L obrazu lewego jest poszukiwane położenie bloku x_R w obrazie prawym, zawierającym najbardziej „podobny” blok obrazu do bloku w obrazie lewym. Najczęściej stosowaną i skuteczną miarą podobieństwa bloków jest suma bezwzględnych różnic (*SAD – sum of absolute differences*) definiowana następująco:

$$\begin{aligned}
 SAD(x_L, y_L, d) = & \\
 = & \sum_{i=-\frac{1}{2}(winx-1)}^{\frac{1}{2}(winx-1)} \sum_{j=-\frac{1}{2}(winy-1)}^{\frac{1}{2}(winy-1)} \left[|R_L(x_L + i, y_L + j) - R_R(x_R + i + d, y_R + j)| + \right. \\
 & |G_L(x_L + i, y_L + j) - G_R(x_R + i + d, y_R + j)| + \\
 & \left. |B_L(x_L + i, y_L + j) - B_R(x_R + i + d, y_R + j)| \right]
 \end{aligned} \quad (2.18)$$

gdzie: $winx \times winy$ – rozmiar bloku, i, j – indeksy punktów obrazu w obszarze bloku, R, G, B – składowe kolorów odpowiednio koloru czerwonego, zielonego i nie-

bieskiego dla obrazu kamery lewej i prawej. Miara SAD jest obliczana dla każdego punktu obrazu odniesienia (obraz kamery lewej) w zadanym zakresie dysparycji (przesunięcia) d , tj. $d_{\max} - d_{\min}$. Dla obrazu o rozmiarze $w \times h$ należy wyznaczyć $(w + winx + 1) \times (h + 1) \times (d_{\max} - d_{\min} + 1)$ miar SAD. Dla przykładu dla obrazu o rozdzielczości 1024×768 , rozmiarze bloku 7×7 i zakresie poszukiwania bloku $d_{\max} - d_{\min} = 20$ punktów, liczba wyznaczanych miar SAD wynosi 16 665 768. Jest to duży koszt obliczeniowy. W pracy (Pełczyński i Strumiłło, 2005) zaproponowano metodę szybkiego wyznaczania odpowiedniości bloków polegającą na połączeniu metod dopasowania cech i dopasowania bloków.

W następnym kroku obliczeń, spośród $n = d_{\max} - d_{\min} + 1$ miar SAD wyznaczonych dla każdego punktu obrazu referencyjnego należy odszukać miarę o najmniejszej wartości i zanotować odpowiadającą jej wartość dysparycji (przesunięcia) d . W wyniku przeprowadzenia takiej procedury budowana jest gęsta mapa dysparycji. Na rys. 2.22 pokazano przykładowy obraz sceny i wyznaczoną dla niej mapę dysparycji.



Rys. 2.22. Obraz sceny trójwymiarowej (a) i wyznaczony dla tej sceny obraz tzw. gęstej mapy dysparycji (b); jasność punktów w mapie dysparycji maleje wraz z powiększającą się głębią; zakres zmian dysparycji $0 \div 20$ został przeskalowany do zakresu $0 \div 255$ i wyświetlony jako jasności punktów obrazu

Źródło: opracowanie własne

Szybkość obliczeń mapy dysparycji dla różnych rozdzielczości obrazów i zakresu przeszukiwania przesunięcia zestawiono w tabeli 2.1 (kolumny z nagłówkiem CPU). Program do tych obliczeń został napisany w języku kompilatora C++ i uruchomiony na komputerze PC z procesorem AMD Athlon 64X2 Dual Core 2,61 GHz. W algorytmie zastosowano tzw. metodę „przesuwanych okien” (*sliding window*) (Mulhmann i inni, 2002).

W ostatnich latach obserwuje się dynamiczny wzrost mocy obliczeniowej procesorów graficznych (*graphical processing units – GPU*). Są to procesory o równoległej architekturze, dla których udostępniono narzędzia do programowania równoległego, m.in. CUDA (*compute unified device architecture*), umożliwiające opracowywanie wydajnych obliczeniowo algorytmów przetwarzania danych (nie tylko graficznych) w językach wysokiego poziomu (Meuth, 2007). W pracy (Strumiłło i inni, 2009a) zaproponowano implementację algorytmu wyznaczania gęstej mapy dysparycji z wykorzystaniem procesorów rodziny NVIDIA. Uzyskano ok. 20-krotne przyspieszenie czasu obliczeń map dysparycji (tabela 2.1). W pracy (Brown i inni, 2003) zawarto szeroki przegląd różnych implementacji sprzętowych algorytmów wyznaczania dysparycji, m.in. z wykorzystaniem układów reprogramowanych FPGA (*field programmable gate array*) oraz procesorów DSP (*digital signal processor*).

Tabela 2.1. Szybkość obliczania mapy dysparycji dla bloków SAD o rozmiarze 7×7 punktów i rozdzielczości obrazów dysparycji 512×384 punktów uzyskane dla procesorów ogólnego przeznaczenia (CPU) i procesorów graficznych (GPU)

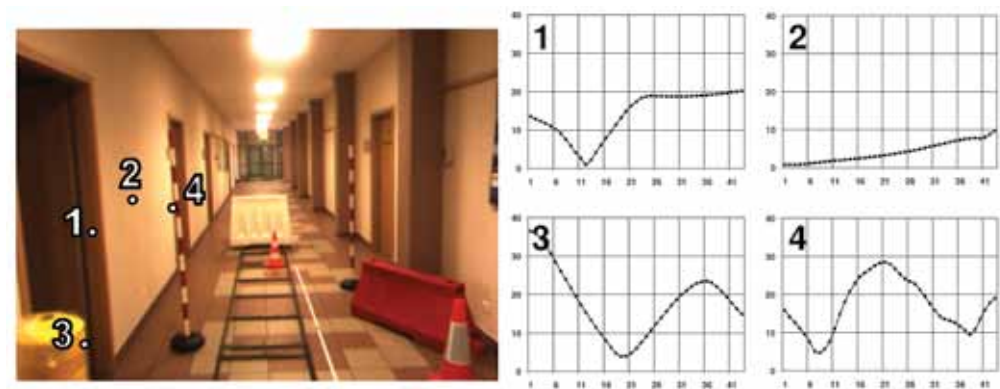
Zakres poszukiwania dysparycji	CPU	GPU 8600M	GPU 9800GT**
n = 20	9 fps	67 fps	200* fps
n = 40	5 fps	33 fps	100* fps

* teoretyczna szybkość obliczona na podstawie czasu wyznaczania pojedynczego obrazu dysparycji;

** dla procesora 9800GT szybkość jest niezależna od rozmiaru bloku w zakresie $3 \times 3 \div 17 \times 17$.

Źródło: opracowanie własne

Wyznaczenie minimum miary SAD nie zawsze jest problemem dobrze uwarunkowanym. Na rys. 2.23 pokazano przykład obrazu sceny trójwymiarowej oraz miary SAD wyznaczone dla zaznaczonych cyframi 1÷4 punktów obrazu (Skulimowski, 2009a). Spośród czterech wskazanych punktów sceny, dysparycja (zatem również głębia) dla punktu 2 zostanie wyznaczona błędnie z uwagi na jednolitą jasność obszaru obrazu w otoczeniu tego punktu (bloki dla tego obszaru mają podobną miarę SAD). Problem błędnego wyznaczania dysparycji dla obszarów o jednolitej jasności (lub kolorze) jest podstawową wadą estymacji głębi na podstawie technik stereowizyjnych. W pracy (Rzeszutarski i inni, 2007) zaproponowano metodę oceny wiarygodności wyznaczanych map dysparycji dla takich obszarów w obrazach scen trójwymiarowych.



Rys. 2.23. Obraz sceny trójwymiarowej wraz z zaznaczonymi punktami 1, 2, 3, 4, dla których wyznaczono miary SAD; jednoznaczne i poprawne minima można wyznaczyć dla punktów 1, 3, 4, natomiast dla punktu 2 minimum miary SAD nie odpowiada faktycznej głębici obrazowanego punktu sceny
Źródło: opracowanie własne

Dyskusji wymaga problem rozdzielczości wyznaczania głębici za pomocą układów stereowizyjnych. W kamerach cyfrowych wyznaczana wartość dysparycji jest wielokrotnością odstepu próbkowania obrazu Δx . Zatem równanie (2.16) można zapisać w postaci:

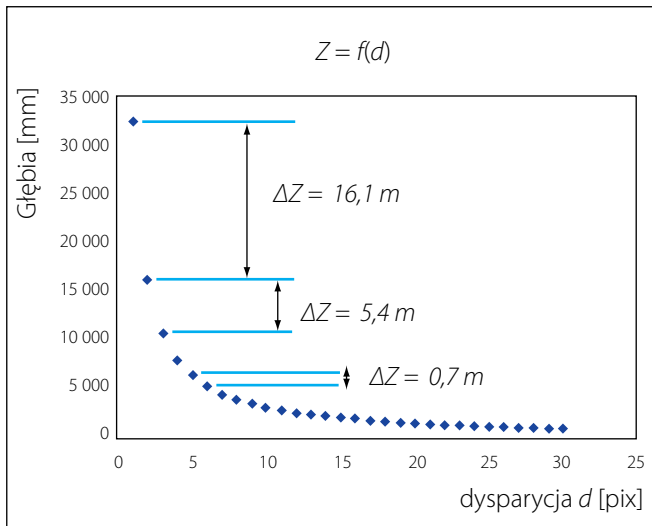
$$Z = \frac{Bf}{d} = \frac{Bf}{n\Delta x} \tag{2.19}$$

gdzie n – jest przesunięciem wyrażonym w liczbie punktów obrazu. Dla testowego systemu stereowizyjnego pokazanego na rys. 2.18 kolejnym skwantowanym wartościom dysparycji odpowiadają kolejne, również skwantowane wartości głębici zestawione w tabeli 2.2 i wykreślone na rys. 2.24.

Tabela. 2.2. Wartości dysparycji (n) i odpowiadające im wartości głębici (Z) dla układu stereowizyjnego z rys. 2.18 o parametrach ($B = 8\text{ cm}$, $f = 3,5\text{ m}$, $\Delta x = 8,8\text{ }\mu\text{m}$)

n	0	1	2	3	4	5	6	7	8	9	10
$Z[\text{m}]$	∞	31,8	15,9	10,6	7,95	6,36	5,30	4,55	3,98	3,54	3,18
n	11	12	13	14	15	16	17	18	19	20	21
$Z[\text{m}]$	2,89	2,65	2,45	2,27	2,12	1,99	1,87	1,77	1,67	1,59	1,51

Źródło: opracowanie własne



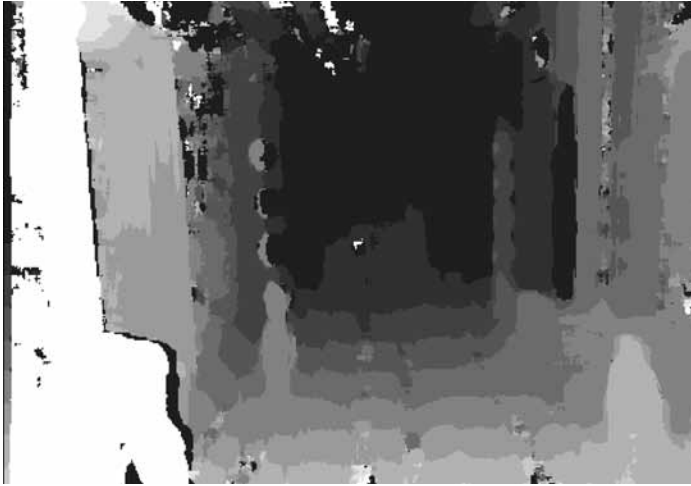
Rys. 2.24. Wykres zależności pomiędzy dysparcją a głębokością dla układu stereowizyjnego o parametrach ($B = 8\text{ cm}$, $f = 3,5\text{ m}$, $\Delta x = 8,8\text{ }\mu\text{m}$)

Źródło: opracowanie własne

Ogólny związek pomiędzy rozdzielczością obliczania dysparycji a rozdzielczością wyznaczania głębi można wyznaczyć z zależności:

$$\Delta Z \cong -\frac{Bf}{d^2} \Delta d = -\frac{Z^2}{Bf} \Delta n \Delta x \quad (2.20)$$

Zatem bezwzględna rozdzielczość wyznaczania głębi rośnie z kwadratem estymowanej głębi, jest proporcjonalna do odstępów próbkowania obrazu oraz odwrotnie proporcjonalna do bazy układu stereowizyjnego i ogniskowej kamer. Powiększanie bazy poprawia rozdzielczość głębi ale skutkuje zwiększeniem rozmiarów układu stereowizyjnego, trudniejszym dopasowaniem bloków (pary bloków są mniej „podobne”) i większymi obszarami sceny widzianych tylko przez jedną kamerę. Podobnie też zwiększenie ogniskowej poprawia rozdzielczość wyznaczania głębi ale duża ogniskowa zawęży kąt widzenia kamer. Dobór parametrów układu stereowizyjnego powinien ściśle zależeć od zastosowania (Cyganek, 2002). Dla rozważanego systemu substytucji sensorycznej wzroku zastosowano bazę układu stereowizyjnego zbliżoną do rozstawu oczu człowieka, tak aby system mógł być umieszczony na głowie osoby niewidomej (www.naviton.pl). Na rys. 2.25 uwidoczniono rozdzielczość głębi dla sceny zilustrowanej na rys. 2.23.



Rys. 2.25. Gęsta mapa dysparycji wyznaczona dla tzw. rozdzielczości pikselowej (wyraźnie widoczne kwantowanie głębi obrazowanej sceny)

Źródło: opracowanie własne

Opracowano metody wyznaczania map głębi z tzw. rozdzielczością subpikselową, w których zastosowano metody interpolacji dla ciągu miar SAD. Metody te wymagają dodatkowych i znacznych nakładów obliczeń (Szeliski i Scharstein, 2002). W rozdziale 3. omówiono technikę poprawy dokładności estymacji głębi na podstawie sekwencji obrazów stereoskopowych rejestrowanych przez ruchomy układ stereowizyjny (Skulimowski i Strumiłło, 2008).

W tabeli 2.3 zestawiono podstawowe zalety i wady stereowizji jako pasywnej techniki estymacji głębi.

Tabela. 2.3. Podstawowe zalety i wady stereowizji

Zalety:

- skuteczność działania w naturalnych warunkach oświetlenia (metoda pasywna)
 - rekonstrukcja sceny trójwymiarowej w postaci gęstej mapy głębi
 - małe rozmiary i waga (możliwość zastosowania miniaturowych kamer cyfrowych)
 - istnienie skutecznych algorytmów rekonstrukcji
 - możliwość rekonstrukcji scen trójwymiarowych z częstotliwością wyświetlania obrazów jak w standardzie TV (przy zastosowaniu odpowiednio szybkich procesorów, np. GPU)
-

Tabela. 2.3. cd. Podstawowe zalety i wady stereowizji

Zalety:

- duża rozdzielczość estymacji głębi obiektów sceny umieszczonych w małej odległości od kamer
 - mały koszt
-

Wady:

- konieczność kalibracji torów optycznych kamer
 - wymagana wzajemna synchronizacja kamer (w szczególności przy obrazowaniu scen oświetlanych światłem lamp fluorescencyjnych)
 - duży nakład obliczeń związany z rekonstrukcją obrazów
 - mała rozdzielczość estymacji głębi obiektów sceny znajdujących się w dużej odległości od kamer (istnieją metody estymacji głębi z rozdzielczością „subpikselową”)
 - estymacja głębi ograniczona do fragmentów sceny widzianych jednocześnie przez obie kamery (m.in. brak możliwości estymacji głębi fragmentów sceny zasłoniętych przez inne obiekty)
 - błędna estymacja głębi dla obszarów o jednolitej jasności, struktur obrazowanych w postaci poziomych linii o jednolitej barwie i obszarów o regularnie powtarzających się wzorach
-

Źródło: opracowanie własne

2.3. Modelowanie sceny trójwymiarowej

Modelowanie sceny trójwymiarowej definiuje się jako proces składający się z rejestracji (akwizycji) obrazów sceny (np. za pomocą technik aktywnych lub pasywnych) oraz przetwarzania i analizy tych danych dla uzyskania modelu komputerowego sceny budowanego za pomocą figur geometrycznych (brył, powierzchni i płaszczyzn). Modelowanie scen i obiektów trójwymiarowych jest stosowane w wielu dziedzinach nauki i techniki, m.in. w biologii i medycynie, fotogrametrii, grafice i animacji komputerowej oraz robotyce i nawigacji (Remondino i El-Hakim, 2006). Zależnie od zastosowania wykorzystuje się różne techniki modelowania scen trójwymiarowych. W grafice komputerowej złożone figury geometryczne buduje się głównie z siatek wielokątów, głównie trójkątów

(*triangular meshes*), gdyż takie modelowanie figur upraszcza tzw. renderowanie, tj. tworzenie dwuwymiarowych zobrazowań scen trójwymiarowych (Blythe, 2008). Jakość dzisiejszych technik modelowania scen trójwymiarowych umożliwia uzyskanie modeli o fotorealistycznej jakości (Evers-Senne i Koch, 2005), w których obiektom można nadawać barwę z wielu milionów kolorów, praktycznie dowolny kształt, teksturę oraz zastosować realistyczny model oświetlenia sceny uwzględniający własności optyczne powierzchni (Singh i Narayanan, 2010). Możliwości takie udostępniają m.in. programy 3ds Max, Lightwave i Maya oraz szybko rozwijające się technologie procesorów graficznych umożliwiających szybkie renderowanie i animowanie złożonych modeli trójwymiarowych obiektów i scen (Akenine-Moller i Strom, 2008).

W wielu zastosowaniach, m.in. robotyce i nawigacji dąży się z kolei do budowania uproszczonych modeli scen trójwymiarowych, na podstawie których jest możliwa automatyczna, szybka i niezawodna nawigacja (Santosh i inni, 2008). Opracowano skuteczne algorytmy budowania modeli scen trójwymiarowych dla ruchomych układów akwizycji obrazów (*structure from motion*). Do klasy tych algorytmów można zaliczyć metodę analizy modelu sceny i jej syntezy z zastosowaniem iteracyjnej procedury porównującej bieżący model sceny z rejestrowanymi obrazami sceny (Koch, 1993)). Znany w literaturze jest również algorytm SLAM (*simultaneous localization and mapping*), umożliwiający jednoczesną lokalizację układu rejestracji obrazów i budowanie modelu sceny (Davidson i inni, 2007).

W niniejszej części pracy omówiono metodę modelowania sceny trójwymiarowej z zastosowaniem stereowizji, koncepcyjnie podobną do algorytmu SLAM, opracowaną specjalnie do prezentacji dźwiękowej sceny w osobistym systemie wspomagania nawigacji dla osoby niewidomej.

Przyjęto uproszczony model sceny, który jest budowany z wycinków płaszczyzn oraz prostopadłościanów (Skulimowski i Strumiłło, 2006). Uzasadnienie dla wyboru takiego modelu sceny jest następujące:

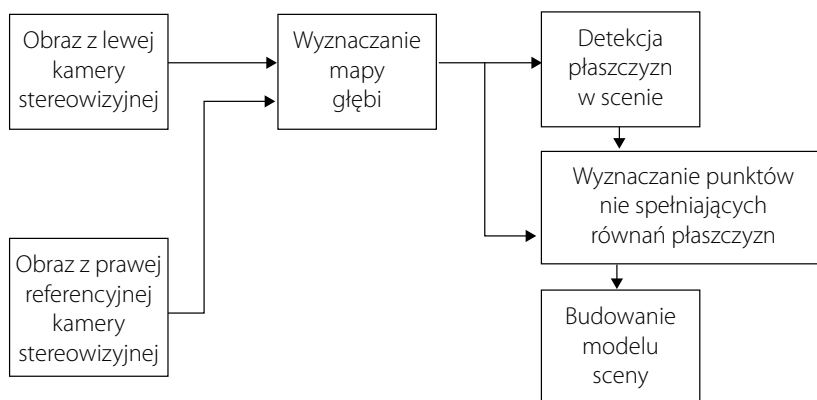
- przeważająca większość scen wewnątrz budynków jak i w obszarze zurbanizowanym spełnia przyjęte założenie (fragmenty płaszczyzny stanowią: chodnik, ulica, ściany wewnątrz budynków i elewacje budynków), rys. 2.26,
- osoby niewidome zwracają uwagę, że nie kształt przeszkody, a sposób jej ominięcia jest ważny; prostopadłościenny model przeszkody dobrze przystaje do tych wymagań,
- wycinki płaszczyzn można skutecznie wykrywać w obrazach głębi i na tej podstawie dokonywać segmentacji obrazów na fragmenty płaszczyzn i obszary niespełniające równań płaszczyzn,

- podział sceny na wycinki płaszczyzn i inne obiekty jest ważne dla zastosowanej techniki sonifikacji sceny trójwymiarowej (opisanej w rozdz. 4.), w której fragmentom płaszczyzn przypisuje się inne dźwięki niż obiektom klasyfikowanym jako przeszkody.



Rys. 2.26. Przykłady scen trójwymiarowych, których model można budować z płaszczyzn i brył
Źródło: opracowanie własne

Uproszczony schemat blokowy algorytmu budowania modelu sceny trójwymiarowej składającej się z fragmentów płaszczyzn i brył pokazano na rys. 2.27.



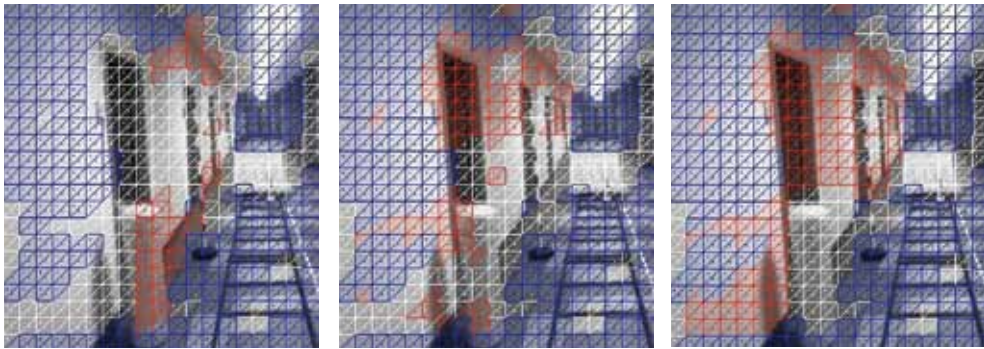
Rys. 2.27. Schemat blokowy algorytmu detekcji fragmentów płaszczyzn i brył w scenie trójwymiarowej
Źródło: opracowanie własne

Opracowany algorytm składa się z dwóch etapów obliczeniowych, w których jest przeprowadzana detekcja wycinków płaszczyzn, a następnie detekcja obiektów nie spełniających równań płaszczyzn.

Algorytm detekcji płaszczyzn przebiega wg następujących kroków (bardziej szczegółowy opis algorytmu można znaleźć w pracy (Skulimowski, 2009)):

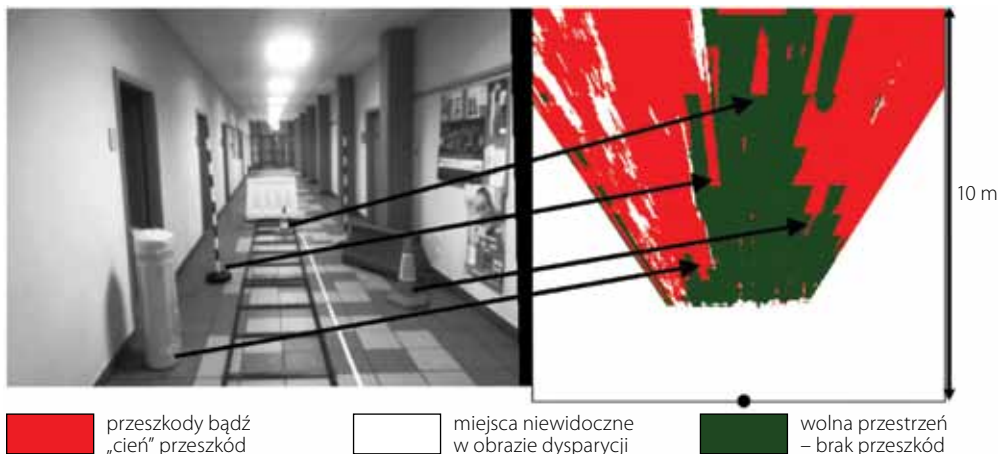
1. Podziel obraz sceny trójwymiarowej na trójkąty tworzące regularną siatkę (rys. 2.28).
2. Dla wierzchołków trójkątów wyznacz głębię, zastosuj odpowiedni test weryfikujący poprawność wyznaczonej głębii (Rzeszutarski i inni, 2007) i wybierz trójkąty dla których głębia jest poprawnie wyznaczona.
3. Dla każdego i -tego trójkąta (na podstawie współrzędnych jego wierzchołków) oblicz równanie płaszczyzny: $A_ix + B_iy + C_iz + 1 = 0$ oraz składowe wektora normalnego $\vec{n}_i$ do tej płaszczyzny.
4. Znajdź i^* -ty trójkąt podobny do największej liczby trójkątów (niezakwalifikowanych do płaszczyzny) stosując jako miarę podobieństwa iloczyn skalarny wektorów: $S_{ij} = \vec{n}_i \cdot \vec{n}_j$ i określ i^* -ty trójkąt jako trójkąt wzorcowy stanowiący zaczątek płaszczyzny.
5. Dla każdego i -tego ($i \neq i^*$) trójkąta oblicz odległości $l(P_i)$ jego wierzchołków od płaszczyzny wyznaczonej przez trójkąt wzorcowy.
6. Jeżeli odległość każdego wierzchołka i -tego trójkąta spełnia warunek $l(P_i) < \varepsilon$, gdzie ε jest doświadczalnie ustaloną wartością, to włącz ten trójkąt do płaszczyzny i uaktualnij równanie płaszczyzny dla powiększonej o 3 liczby punktów stosując kryterium minimum błędu średniokwadratowego.
7. Powtarzaj kroki 5 i 6 (zastępując trójkąt wzorcowy zbiorem trójkątów zaliczonych do płaszczyzny).

Na rys. 2.28 zilustrowano kolejne iteracje algorytmu detekcji płaszczyzn. Na podstawie badań eksperymentalnych stwierdzono, że dla 10 iteracji uzyskuje się prawidłową detekcję płaszczyzn sceny.



Rys. 2.28. Ilustracja iteracyjnego algorytmu detekcji płaszczyzn (dla iteracji 1, 3 i 5). Kolorem niebieskim zaznaczono trójkąty leżące w obszarach obrazu zakwalifikowanych do płaszczyzn; kolorem czerwonym zaznaczono trójkąty zakwalifikowane jako należące do aktualnie przetwarzanej płaszczyzny; kolorem białym zaznaczono trójkąty niezakwalifikowane do płaszczyzn (trójkąty są „rozpięte” na powierzchni obiektów sceny, a ich wierzchołki znajdują się w miejscu wyznaczonej dla nich głębi)
Źródło: opracowanie własne

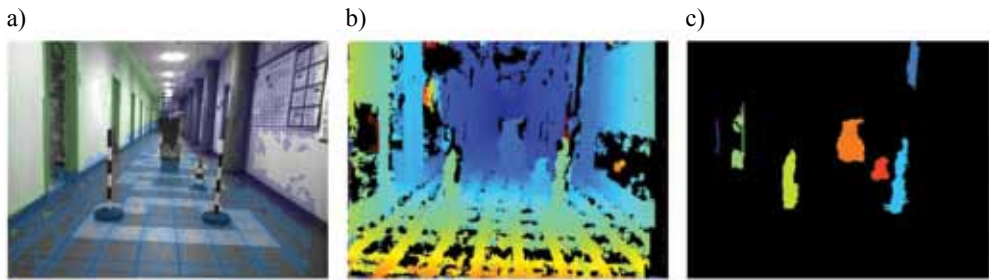
Algorytm detekcji płaszczyzn umożliwia m.in. wyznaczenie płaszczyzny podłoża i zastosowanie jej do określenia wolnego od przeszkód pola podłoża. Wykorzystanie tej informacji rozważano w jednym ze schematów sonifikacji otoczenia w systemie nawigacji dla osoby niewidomej (Bujacz i inni, 2011). Na rys. 2.29 pokazano obraz sceny oraz widok „z góry” na tę scenę z zaznaczonym kolorem zielonym polem podłoża wolnym od przeszkód.



Rys. 2.29. Obraz sceny oraz obraz wydzielonych przeszkód wyznaczony dla hipotetycznej płaszczyzny równoległej do płaszczyzny podłoża rozpiętej na wysokości 30 cm

Źródło: opracowanie własne

Jak pokazano na schemacie blokowym z rys. 2.27 po wyznaczeniu fragmentów płaszczyzn, detekcji podlegają obiekty sceny nie spełniające równań płaszczyzn. Opracowano algorytmy detekcji obiektów wykorzystujące progowanie wartości punktów mapy dysparycji i metody rozrostu obszaru. Zastosowano prostą metodę detekcji progowej wg warunku: $|d(x, y) - d_{pl}(x, y)| > T$ gdzie: $d(x, y)$ jest dysparycją rozpatrywanego punktu mapy, $d_{pl}(x, y)$ jest dysparycją obliczoną dla modelu płaszczyzny znalezionej w poprzednim kroku algorytmu, a T jest dobraną doświadczalnie wartością progową. Odpowiedni algorytm rozrostu obszaru służy do łączenia takich punktów w spójne obszary i ich etykietowania. Wynik segmentacji obrazu sceny na płaszczyzny i obiekty interpretowane jako przeszkody pokazano na rys. 2.30.

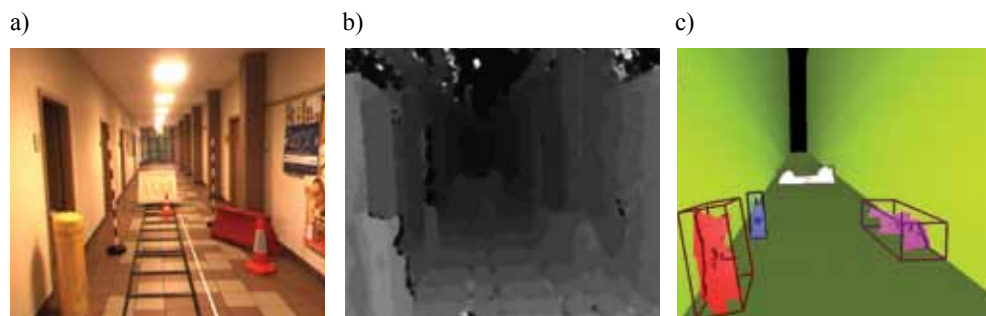


Rys. 2.30. Detekcja płaszczyzn i obiektów w scenie trójwymiarowej:

- a) obraz sceny trójwymiarowej z zaznaczonymi obszarami wyznaczonych fragmentów płaszczyzn, b) mapa głębi wyświetlona z zastosowaniem pseudokolorów;
- c) wynik detekcji przeszkód w scenie (każda oznaczona innym kolorem)

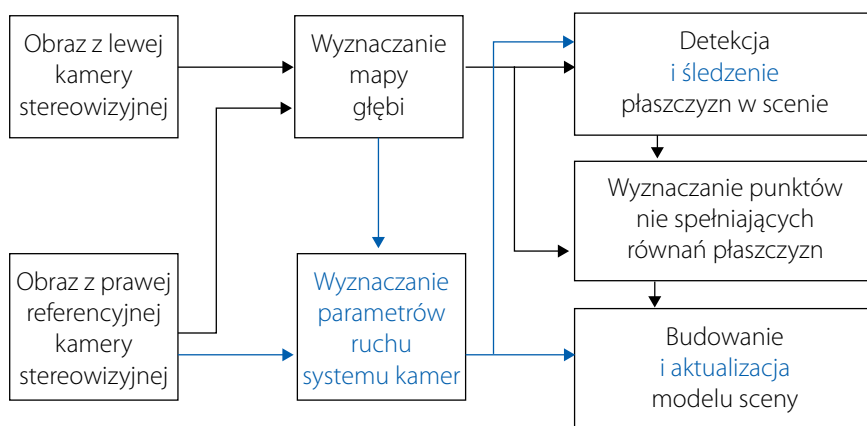
Źródło: opracowanie własne

Końcowym etapem budowanego modelu sceny trójwymiarowej jest przyporządkowanie bryły prostopadłościennej każdemu wydzielonemu obiektowi sceny nie należącemu do płaszczyzny. Dla współrzędnych przestrzennych zbioru punktów należących do obiektów (rys. 2.30) wyznaczono tzw. trzy główne kierunki wg metody analizy składowych głównych (*PCA – principal component analysis*). Wzajemnie prostopadłe kierunki określają orientację przestrzenną, a długości wektorów je wyznaczających określają długości boków prostopadłościanu. Początki wektorów są zaczepione w punkcie odpowiadającym „środkowi ciężkości” współrzędnych punktów należących do obiektów. Końcowy wynik segmentacji pokazano na rys. 2.31.



Rys. 2.31. Obraz sceny trójwymiarowej (a), mapa głębi wyznaczona za pomocą układu stereowizyjnego (b) i model komputerowy sceny składający się z fragmentów płaszczyzn i prostopadłościanów utworzony z zastosowaniem biblioteki OpenGL (c)
Źródło: opracowanie własne

Opisany algorytm budowania modelu sceny poddano pomyślnej weryfikacji dla różnych scen otoczenia oraz wirtualnych scen testowych. Opracowany model był podstawą do opracowania algorytmu nawigacji wizyjnej, w którym wydzielone obiekty sceny w postaci zbudowanego modelu są śledzone w kolejno rejestrowanych obrazach (rys. 2.32).



Rys. 2.32. Schemat blokowy algorytmu detekcji i śledzenia fragmentów płaszczyzn i przeszkód w scenie trójwymiarowej
Źródło: opracowanie własne

Obiekty sceny wydzielone i śledzone w kolejnych ramach obrazów są opisane zbiorem parametrów. Dla fragmentów płaszczyzn są to współczynniki równań płaszczyzn $A(t)$, $B(t)$, $C(t)$, a dla przeszkód współrzędne środka geometrycznego

prostopadłościanu ($X_0(t)$, $Y_0(t)$, $Z_0(t)$) i jego długości boków ($dX_0(t)$, $dY_0(t)$, $dZ_0(t)$), gdzie t oznacza chwilę rejestracji kolejnej pary obrazów stereowizyjnych. Zbiór tych zmieniających się w czasie parametrów (na skutek ruchu układu akwizycji) steruje algorytmem sonifikacji obiektów tj. parametrami dźwięków przypisanych wydzielonym obiektom sceny. Uzyskana szybkość obliczeń algorytmu segmentacji sekwencji obrazów sceny trójwymiarowej, wyznaczania parametrów ruchu układu kamer i śledzenia wydzielonych obiektów wynosi ok. 4 obrazów na sekundę o rozdzielczości 512×384 punktów i obliczeń wykonanych z zastosowaniem komputera klasy PC z procesorem Intel Quad 6600 2,4 GHz i kompilatora Visual C++ 2008 .NET.

Zastosowany schemat sonifikacji i testy systemu do dźwiękowego obrazowania otoczenia z udziałem osób niewidomych opisano w rozdziale 5. książki.

W niniejszym rozdziale omówiono podstawowe techniki estymacji głębi, które znajdują zastosowanie w systemach ostrzegania o przeszkodach i wspomagania orientacji przestrzennej dla osób niewidomych². Pozyskana informacja o strukturze geometrycznej otoczenia podlega odpowiedniej analizie dla wytworzenia niewizualnego zobrazowania otoczenia i jego prezentacji dla zmysłów słuchu lub dotyku. Zagadnienia te omówiono w rozdziale 4. monografii.

Następny, 3. rozdział poświęcono systemom wspomagającym osoby niewidome w tzw. makronawigacji (nawigacji globalnej), tj. systemom określania położenia geograficznego oraz systemom integracji danych nawigacyjnych.

² Systemy tej klasy określa się też mianem systemów mikronawigacji lub nawigacji lokalnej.

Droga lądowa z punktu startowego do celu tylko w rzadkich przypadkach może być planowana wzdłuż linii prostej, nawet w sytuacji, w której cel podróży jest widoczny z punktu startowego. Lokalne ukształtowanie terenu i naturalne przeszkody, sieć dróg, a w terenie miejskim sieć ulic i chodników oraz gęsta i wysoka zabudowa powodują, że wyznaczona trasa jest linią o złożonym kształcie. Poruszanie się wzdłuż takiej trasy (szczególnie w nieznanym terenie) nie jest zadaniem prostym nawet dla człowieka w pełni sprawnego. Dlatego też tradycyjnie wspomaga się on mapami, kompasem oraz korzysta z charakterystycznych punktów orientacyjnych i znaków drogowych. Wymienione pomoce nawigacyjne dla osoby niewidomej są dostępne w bardzo ograniczonym stopniu, a odpowiedź na trzy podstawowe zagadnienia nawigacji (Leonard i Durrant-Whyte, 1991), tj.: „Gdzie jestem?”, „Gdzie znajduje się cel podróży?”, „Jak się tam dostać?” jest znacznie trudniejsza.

Odpowiedź na pierwsze pytanie osoba niewidoma może formułować na podstawie wykrytych punktów orientacyjnych i ich porównywaniu z informacją *a priori* o lokalnym otoczeniu (np. z zapamiętaną mapą wyobrażeniową). Odpowiedź na drugie pytanie wymaga wiedzy *a priori* o szerszym obszarze terenu, w którym osoba niewidoma się porusza (pozyskanej np. na podstawie mapy dotykowej lub opisu słownego). Odpowiedź na trzecie pytanie jest zwykle najtrudniejsza i wiąże się zarówno z wiedzą *a priori* o terenie jak i lokomocją osoby niewidomej, m.in. jej orientacją przestrzenną oraz umiejętnością znajdowania drogi i omijania przeszkód.

Warto nadmienić, że problemy określania bieżącego położenia, planowania ścieżki (kinematyki) i dynamiki ruchu są przedmiotem badań związanych

z robotyką (Tadeusiewicz, 1992), (Jezierski, 2005) w szczególności z robotami mobilnymi (Granosik, 2011). Należy jednak podkreślić, że systemy nawigacji dla osoby pieszej (w szczególności osoby niewidomej) wymagają specjalnych technik obliczeniowych, odmiennych niż w systemach nawigacji samochodowej lub nawigacji robotów mobilnych.

3.1. Rodzaje technik nawigacji

Jedna z definicji nawigacji (Gallistel, 1990) mówi:

Nawigacja jest procesem polegającym na określaniu własnego położenia na linii kursu i utrzymywaniu kursu prowadzącego do celu podróży.

Skuteczna realizacja procesu nawigacji wymaga zatem ciągłego (lub regularnego) ustalania następujących parametrów dla nawigowanego obiektu.

1. Położenia wg zadanego układu współrzędnych.
2. Kierunku i prędkości przemieszczania się.

Pierwsze z zadań nawigacji określa się mianem wyznaczania położenia bezwzględnego (*absolute positioning*), a drugie wyznaczaniem położenia względnego, nazywane też metodą zliczeniową (*dead-reckoning*).

Wśród podstawowych metod wyznaczania położenia bezwzględnego można wyróżnić:

- *nawigację według modelu*, tj. map dwu- lub trójwymiarowych lub zapamiętanych map wyobrażeniowych (Filliat i Meyer, 2003),
- *nawigację według punktów orientacyjnych*, tj. naturalnych punktów otoczenia (*natural landmarks*), np. układu gwiazd lub specjalnie rozmieszczonych punktów (*artificial landmarks*), np. latarni i radiolatarni o znanej lokalizacji (Kurek i Czerwiński, 2003), (Kulyukin i inni, 2004),
- *nawigację z zastosowaniem multilateracji (m.in. trilateracji)*, w tej metodzie są stosowane co najmniej trzy nadajniki o znanej lokalizacji, a położenie wyznacza się na podstawie pomiarów pośrednich według trzech podstawowych technik:
 - *pomiar czasu propagacji sygnału (time of arrival – TOA)* lub *różnicy czasu propagacji sygnałów (time difference of arrival – TDOA)* od nadajników, które wysyłają synchronizowane sygnały; pomiar czasu wzajemnych przesunięć odebranych sygnałów stanowi podstawę do wyliczania odle-

głości do tych nadajników i położenia wg metody trilateracji¹; te techniki wyznaczania położenia są wykorzystywane w globalnych systemach pozycjonowania, m.in. GPS oraz w lokalnych sieciach radiowych WLAN (*wireless local area network*); w tej grupie metod wyróżnia się również metody wyznaczania położenia na podstawie pomiaru czasu powrotu wysłanego sygnału (*roundtrip time of flight* – RTOF), które są stosowane m.in. w technice radarowej (Buczowski i Radecki, 2010);

- *pomiar poziomu sygnału odbieranego (received signal strength – RSS)*: pomiary natężenia pola wraz z danymi z modeli propagacji (Hauzman, 2007), służą do wyznaczenia odległości do nadajników, pozycja jest następnie obliczana wg metody trilateracji (Li i inni, 2006); technika ta jest stosowana m.in. do określania lokalizacji wewnątrz budynków, na podstawie pomiaru pola elektromagnetycznego sieci WLAN oraz w terenie otwartym na podstawie pomiaru pola stacji bazowych telefonii komórkowej GSM (Wang i inni, 2008);
- *pomiar kierunku odebranego sygnału (angle of arrival – AOA)*, technika określania położenia, w której stosuje się triangulację, metodę pomiaru stosowaną od setek lat w geodezji i nawigacji morskiej.

Należy zaznaczyć, że położenie bezwzględne może być wyznaczane wg globalnego układu odniesienia związanego z powierzchnią Ziemi (*earth frame*) lub współrzędnych lokalnych, stosowanych np. w niektórych mapach cyfrowych (Titterton i Weston, 2004). Położenie globalne można powiązać ze współrzędnymi w lokalnym układzie odniesienia przez zastosowanie odpowiednich transformacji przeliczających współrzędne (Kadaj, 2006).

Do podstawowych metod zliczeniowych należą:

- *odometria*, tj. technika stosowana głównie w robotyce do pomiaru długości pokonanej drogi, m.in. z wykorzystaniem tzw. enkoderów impulsowych mierzących szybkość obrotów mechanizmu napędowego (Skulimowski i Świątczak, 2009b),

¹ *Trilaterację* (na przykładzie globalnego systemu pozycjonowania) można w uproszczeniu objaśnić następująco: jeżeli wiemy, że jesteśmy w odległości d_A od satelity A to możemy znajdować się w dowolnym miejscu sfery o tym promieniu (której środek jest wyznaczony położeniem tego satelity). Jeżeli również wiemy, że jesteśmy w odległości d_B od satelity B to nasze położenie ogranicza się do okręgu wyznaczonego linią przecięcia sfer odpowiadających satelitom A i B . Mierząc następnie odległość d_C od satelity C , nasze położenie ogranicza się już do dwóch punktów stanowiących przecięcie sfery o promieniu d_C z wcześniej wyznaczonym okręgiem. Powierzchnia Ziemi, na której się znajdujemy jest czwartą sferą (w przybliżeniu) jednoznacznie rozstrzygającą, w którym z dwóch punktów się znajdujemy.

- *odometria wizyjna (visual odometry)*, polegająca na analizie ruchu (przepływu optycznego) tzw. punktów kluczowych sceny w rejestrowanych sekwencjach obrazów i na tej podstawie estymacji ruchu własnego układu nawigacji (*ego-motion*) układu (Nister i inni, 2004)),
- *nawigacja inercyjna (inertial navigation)*, w której współcześnie wykorzystuje się elektroniczne czujniki inercyjne (żyroskopy i czujniki przyspieszeń) do pomiaru przyspieszeń liniowych i prędkości obrotowych (Titterton i Weston, 2004)).

Wadą metod zliczeniowych jest akumulacja błędów wynikająca z przyrostowego sposobu wyznaczania kolejnych położeń.

W tabeli 3.1 zebrano i porównano wybrane techniki wyznaczania położenia bezwzględnego i względnego (zliczeniowego). Zaznaczono podstawowe zalety i wady każdej z technik (Grejner-Brzezińska i inni, 2008).

W rozdz. 5. omówiono podstawowe systemy nawigacji osobistej dla osób niewidomych, w których zastosowano wybrane techniki nawigacji osobistej.

Tabela. 3.1. Podstawowe metody wyznaczania położenia bezwzględnego i względnego, wykorzystywane w systemach nawigacji osobistej, gdzie: X, Y, Z i X_G, Y_G, Z_G – oznaczają odpowiednio współrzędne w układzie lokalnych i globalnym, $\dot{X}, \dot{Y}, \dot{Z}$ i $\dot{X}_G, \dot{Y}_G, \dot{Z}_G$ – oznaczają odpowiednio prędkości i przyspieszenia liniowe; α, β, γ – oznaczają kąty obrotu odpowiednio wokół osi OX, OY i OZ .

Technika pozycjonowania	Parametr nawigacji	Typowa dokładność	Uwagi: wady/zalety, zastosowania
Wyznaczanie położenia bezwzględnego			
Globalne systemy pozycjonowania	X_G, Y_G, Z_G $\dot{X}_G, \dot{Y}_G, \dot{Z}_G$	10÷100 m 0,05÷0,2 m/s	• położenie wyznaczone w globalnym układzie współrzędnych • skuteczna w terenie otwartym • duże błędy pozycjonowania w obszarze gęstej zabudowy miejskiej
WLAN	X, Y, Z	1÷6 m	• stosowana wewnątrz budynków • pomiar natężenia pola • dokładniejsze metody empiryczne (<i>fingerprinting</i>) (Li i inni, 2006)
Siec stacji bazowych GSM	X_G, Y_G	50÷300 m	• multilateracja • mała dokładność dla obszarów o małej gęstości stacji bazowych • wymagane odpowiednie oprogramowanie w urządzeniu mobilnym • zastosowanie: usługi lokalizacyjne (<i>location based services</i>)

Tabela. 3.1. (cd).

Technika pozycjonowania	Parametr nawigacji	Typowa dokładność	Uwagi: wady/zalety, zastosowania
Elektroniczny kompas	kąt azymutu	$1^\circ \div 180^\circ$	• duże zakłócenia w terenie miejskim (zakrzywienie linii sił pola magnetycznego od obiektów metalowych o dużej masie: słupy, pojazdy, szyny) • wymaga stabilizacji w położeniu poziomym
Wyznaczanie położenia względnego (metody zliczeniowe)			
Odometria*	X, Y $\dot{X}, \dot{Y}, \dot{Z}$	$\sim 3\%$	• błąd pomiaru spowodowany poślizgiem kół względem podłoża
Odometria wizyjna	X, Y, Z $\dot{X}, \dot{Y}, \dot{Z}$ α, β, γ	$\sim 1 \text{ m}$	• systemy jedno- i dwukamerowe wymagane zastosowanie złożonych algorytmów analizy obrazu • zawodzi dla szybkich przemieszczeń układu akwizycji (np. obrotów)
Inercyjna: czujniki przyspieszeń	$\ddot{X}, \ddot{Y}, \ddot{Z}$	$< 0,05 \text{ m/s}^2$	• skuteczna dla szybkich przemieszczeń • problem eliminacji przyspieszenia ziemskiego dla czujników 3DoF (<i>degrees of freedom</i>) • występuje dryft
Inercyjna: żyroskopy	α, β, γ	$1^\circ \div 3^\circ$	• skuteczna dla szybkich obrotów • występuje dryft

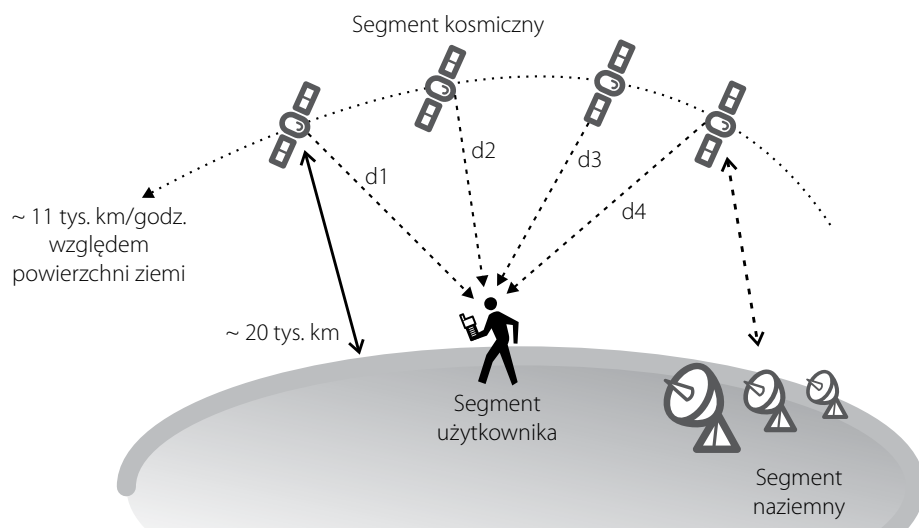
* Odometrię (typowo stosowaną w robotyce) zastosowano w urządzeniu o nazwie Mygo, które jest specjalnej budowy białą laską; u jej dolnego końca wbudowano enkoder z małym kołem pomiarowym (Ritzler, 2011).

Źródło: opracowanie własne na podstawie (Grejner-Brzezińska i inni, 2008)

3.2. Globalne systemy pozycjonowania

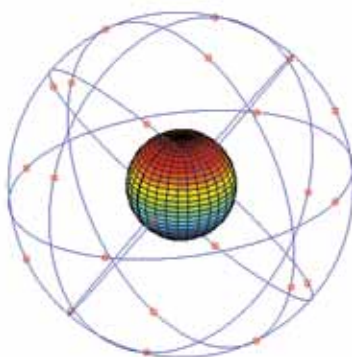
Pierwszym powszechnie udostępnionym globalnym systemem pozycjonowania jest GPS–NAVSTAR (*Global Positioning System – Navigation Signal Timing and Ranging*) opracowany i nadzorowany przez Departament Obrony USA (Narkiewicz, 1999). Umożliwia on wyznaczenie położenia geograficznego i czasu w dowolnym miejscu na powierzchni Ziemi (lub nad powierzchnią ziemi na wysokości małej w porównaniu z wysokością orbitowania satelitów). System składa się z tzw. segmentów (rys. 3.1):

- *kosmicznego*, który jest konstelacją nominalnie 24 satelitów okrążających Ziemię ok. dwukrotnie na dobę, poruszających się po 6 orbitach na wysokości ok. 20 tys. km, (układ satelitów umożliwia bezpośrednią widoczność co najmniej 4 satelitów z powierzchni Ziemi),
- *segmentu kontroli naziemnej*, na który składają się ośrodki rozmieszczone w kilku miejscach na Ziemi nadzorujące pracę systemu, m.in. korygujące położenie satelitów,
- *segmentu użytkownika*, stanowiącego odbiornik GPS umożliwiający jednoczesny odbiór sygnałów nadawanych z kilku satelitów.



Rys. 3.1. Segmenty globalnego systemu pozycjonowania GPS (w uproszczeniu)

Źródło: opracowanie własne



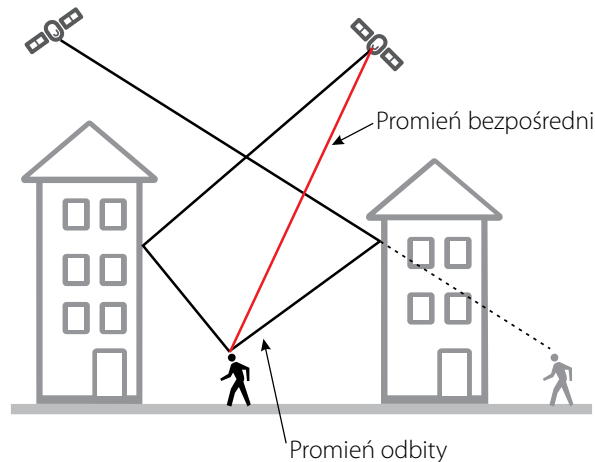
Rys. 3.2. Konstelacja i orbity satelitów GPS

Źródło: rysunek udostępniony przez P. Barańskiego

Każdy z satelitów systemu GPS jest wyposażony w zegar atomowy. Synchronicznie nadaje on sygnały radiowe o częstotliwościach nośnych $L_1=1227$ MHz i $L_2=1575$ MHz. Sygnały te, nazywane depeszami nawigacyjnymi zawierają informacje o położeniu satelity i czasie wysłania depeszy. Ustalenie położenia geograficznego odbiornika odbywa się na podstawie metody trilateracji (multilateracji), wg której jednoczesny odbiór sygnałów z co najmniej trzech satelitów jest wymagany do wyznaczenia współrzędnych geograficznych odbiornika znajdującego się na powierzchni Ziemi. Dostępność sygnałów z czterech satelitów umożliwia wyznaczenie wysokości nad poziomem morza.

Pomiar czasu propagacji sygnałów nawigacyjnych od satelitów jest obarczony błędami, na które składają się:

- refrakcja transmitowanych sygnałów w niejednorodnych warstwach jonosfery i troposfery,
- wielodrogowa propagacja sygnałów (do odbiornika docierają sygnały biegnące bezpośrednio od satelitów ale też sygnały odbite od struktur terenu i budynków, rys. 3.3),
- niedokładność zegarów satelitów i rozbiegania się trajektorii satelitów,
- jakość działania odbiornika (stabilność zegara, układy i algorytmy przetwarzania sygnałów).



Rys. 3.3. Ilustracja wielodrogowej propagacji sygnałów od satelitów do odbiornika; od jednego z satelitów tylko sygnał odbity dobiega do odbiornika

Źródło: opracowanie własne

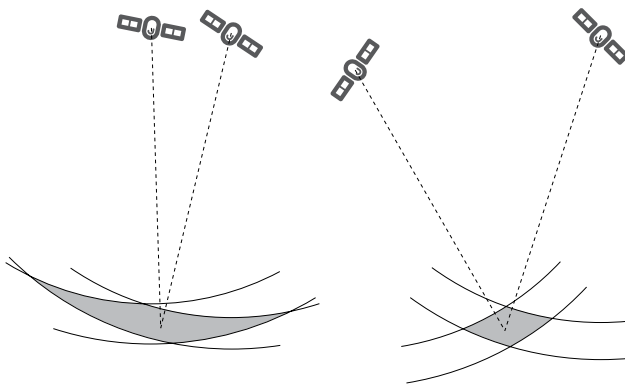
Według ostatniej, 4. edycji raportu Departamentu Obrony USA (Department of Defense, 2008) średni błąd pozycjonowania uzyskiwany w systemie GPS (przez 95% czasu przy pełnej widoczności nieba) wynosi ≤ 9 m w kierunku

poziomym oraz ≤ 15 m dla wysokości). Odbiornik GPS oprócz obliczonego położenia geograficznego i tzw. uniwersalnego czasu koordynowanego UTC (*universal time clock*) udostępnia wskaźniki niepewności odczytu położenia (Langley, 1999), m.in.:

- tzw. rozmycie dokładności pozycjonowania, PDOP (*positioning dilution of precision*), $PDOP \cong \sqrt{\sigma_X^2 + \sigma_Y^2 + \sigma_Z^2}$,
- tzw. rozmycie dokładności pozycjonowania w kierunku poziomym, HDOP (*horizontal dilution of precision*), $HDOP \cong \sqrt{\sigma_X^2 + \sigma_Y^2}$.

gdzie: σ_X , σ_Y , σ_Z oznaczają odpowiednio standardowe odchylenia serii pomiarów długości (X) i szerokości (Y) geograficznej oraz wysokości n.p.m. (Z).

Niepewność odczytu położenia może być powodowana niekorzystną konfiguracją geometryczną satelitów w stosunku do odbiornika (rys. 3.4).



Rys. 3.4. Układ geometryczny satelitów względem odbiornika wpływa na dokładność pozycjonowania (z lewej duże „rozmycie położenia”, z prawej małe rozmycie)

Źródło: opracowanie własne

Odbiorniki globalnych systemów pozycjonowania wraz z danymi pomocniczymi (m.in. liczba widzianych satelitów, wskaźniki rozmycia DOP, HDOP, VDOP) dają odczyt współrzędnych geograficznych w formacie tekstowym, w tzw. protokole NMEA (*National Marine Electronics Association*). Dla przykładu sekwencję znaków w tym formacie: \$GPGLL,4913.55,N,1959.00,E,125123,A*31 należy interpretować jako komunikat:

\$GPGLL	Geographic position, latitude/longitude
4913.55,N	Latitude 49 deg. 13.55 min. North
1959.00,E	Longitude 19 deg. 59.00 min. East
125123	Fix taken at 12:51:23 UTC
A	Data valid
*31	checksum

który wskazuje współrzędne geograficzne Kasprowego Wierchu.

Śledzenie wartości wskaźników DOP jest szczególnie istotne w terenie miasta. Błędy wyznaczania położenia w obszarze wysokiej zabudowy miejskiej (spowodowane przesłanianiem satelitów przez budynki i wielodrogowością sygnałów) mogą sięgać 100 m (Misra i inni, 1999), (Modsching i inni, 2006). Na podstawie badań dokładności wyznaczania położenia wykonanych w Łodzi zanotowane błędy dochodziły do 50 m. Bieżące wartości HDOP wykorzystano w hybrydowym systemie nawigacji łączącym dane z globalnego systemu pozycyjnego i dane z nawigacji zliczeniowej. Dla wartości HDOP przekraczających założoną wartość progową system przełączano na nawigację zliczeniową (Barański i Strumiłło, 2009a).

Innymi systemami globalnego pozycjonowania są: rosyjski GLONASS (www.glonass-ianc.rsa.ru) oraz znajdujące się w budowie europejski system Galileo (www.esa.int) oraz chiński Compass (www.navchina.com).

3.3. Zastosowanie map elektronicznych w nawigacji pieszej

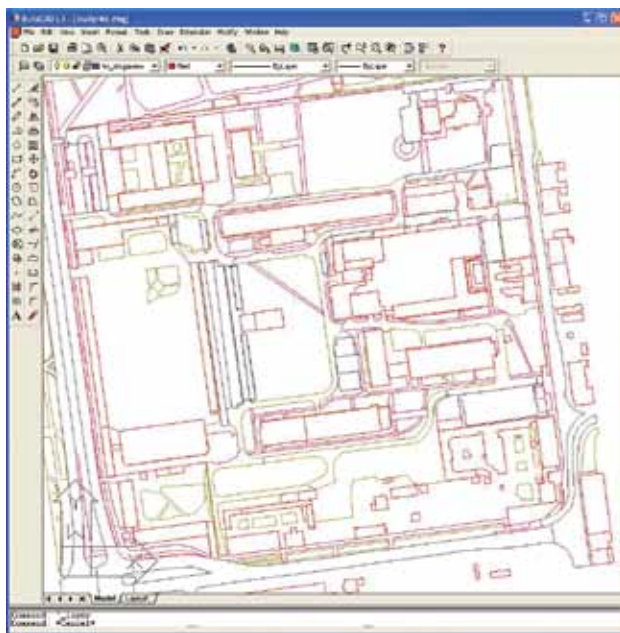
Praktyczną korzyść z globalnych systemów pozycjonowania uzyskuje się po połączeniu danych o położeniu geograficznym z informacjami pozyskanymi z map terenu. Rozwiązania tę rozpowszechniły się szczególnie w systemach nawigacji samochodowej, w której wykorzystuje się cyfrowe mapy terenu, zawierające m.in. dane o lokalizacji budynków, ulic i sieci dróg (Enge i Misra, 1999).

W ostatnich kilku latach są wdrażane systemy nawigacji osobistej pieszej i rowerowej, w tym systemy nawigacji specjalnie przeznaczone dla osób niewidomych. Te ostatnie wymagają uwzględnienia następujących uwarunkowań wynikających z niepełnosprawności wzrokowej i pieszego przemieszczania się:

- mapy terenu powinny zawierać informacje o obszarach udostępnionych do ruchu pieszego (chodniki, place, przejścia podziemne); niewiele map terenu udostępnia te informacje (zob. m.in. OpenRouteService.org, www.loadstone-gps.pl, www.nokia.pl/uslugi-i-aplikacje/nawigacja)
- mapy przeznaczone dla nawigacji pieszej dla niewidomych powinny być regularnie aktualizowane (remonty dróg, chodników, zmiana organizacji ruchu pieszego),
- mapy dla osób niewidomych powinny zawierać dokładne położenia przystanków komunikacji miejskiej (autobusów, tramwajów).

Na potrzeby testów systemu nawigacji dla osób niewidomych budowanego w Instytucie Elektroniki PŁ opracowano odpowiednie mapy terenu (IE PŁ, 2010).

Na rys. 3.5. pokazano mapę geodezyjną fragmentu kampusu Politechniki Łódzkiej wykonaną przez Zakład Geodezji PŁ. Mapa taka jest łatwa do interpretacji przez człowieka, zawiera m.in. obrysy budynków i ciągów komunikacyjnych. Jej format (plik *dwg* pakietu AutoCAD) nie jest odpowiedni do komputerowej analizy, m.in. do celów nawigacji. Opracowano program do konwersji formatu pliku, automatycznej interpretacji obiektów mapy i transformacji współrzędnych geograficznych (z tzw. układu współrzędnych państwowych do współrzędnych geograficznych) (Kadaj, 2006).



Rys. 3.5. Przykład mapy geodezyjnej terenu (kampus Politechniki Łódzkiej, kwartał pomiędzy ulicami Al. Politechniki, ul. Wólczańską, ul. Wróblewskiego i parkiem im. ks. Klepacza)

Źródło: opracowanie własne

Na rys. 3.6. pokazano mapę uzyskaną na podstawie mapy geodezyjnej z rys. 3.5. Jest to mapa przygotowana w formacie wektorowym Shape (*ESRI Shapefile*), w którym obiekty są podzielone na typy (punkty, linie, wielokąty) i zapamiętane w tzw. warstwach. Mapy w formacie Shape wykorzystano do testowania algorytmów korekcji estymacji położenia geograficznego, m.in. testowania położenia punktu (np. rozstrzygania czy znajduje się wewnątrz czy na zewnątrz obiektu) oraz obliczania odległości od najbliższej ściany budynku w zadanym kierunku (Barański i Strumiłło, 2009a).



Rys. 3.6. Widok okna programu wyświetlającego mapę cyfrową (format Shape) z barwami przyporządkowanymi do różnego typu obiektów: czerwony – budynki, szary – jezdnie ulic, żółty – chodniki i ścieżki, zielony – trawniki, ciemno-niebieski – zbiorniki wodne, jasno-niebieski – latarnie, słupki

Źródło: opracowanie własne

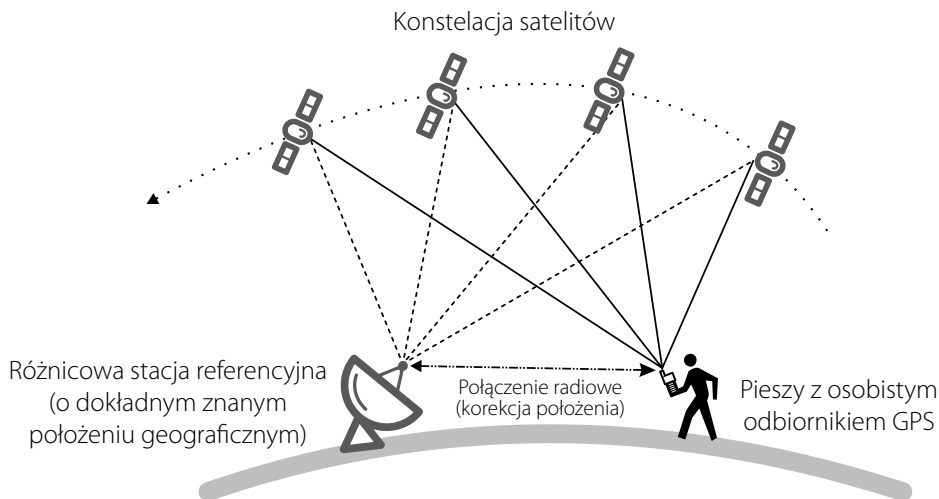
3.4. Wyznaczanie położenia z zastosowaniem nadajników naziemnych

W ostatnich latach dynamicznie rozwija się duża liczba różnych technik wyznaczania położenia z wykorzystaniem sieci radiowych, m.in. sieci telefonii komórkowej. Systemy te można podzielić na dwa podstawowe rodzaje:

- systemy pozycjonowania w terenie (w tym systemy wspierające globalne systemy pozycjonowania (van Diggelen, 2009), oraz systemy telefonii komórkowej (Adusei i inni, 2002))
- systemy pozycjonowania wewnątrz budynków (Gu i inni, 2009), (Bensky, 2008).

Coraz szerzej stosowaną techniką korekcji błędów GPS jest DGPS (*differential GPS*). System ten korzysta z własności propagacji sygnałów od satelitów polegającej na występowaniu podobnych błędów wyznaczania położenia dla odbiorników znajdujących się w niewielkiej odległości od siebie (do kilkunastu kilometrów). Ideę działania tego systemu zilustrowano na rys. 3.7. Wykorzystuje

on dodatkową sieć radiową stacji referencyjnych o znanych dokładnie położeniach geograficznych. W stacjach tych są na bieżąco obliczane błędy pozycjonowania i drogą radiową ($f \approx 300$ kHz) rozsyłane do odbiorników tzw. poprawki różnicowe. DGPS umożliwia uzyskanie $1\div 2$ m dokładności. Należy podkreślić, że DGPS jest nieskuteczny w terenie gęstej zabudowy miejskiej, w której występują dodatkowe błędy pozycjonowania powodowane wielodrogową propagacją sygnałów.



Rys. 3.7. Zasada działania techniki DGPS do poprawy dokładności wyznaczania położenia w systemie GPS

Źródło: opracowanie własne

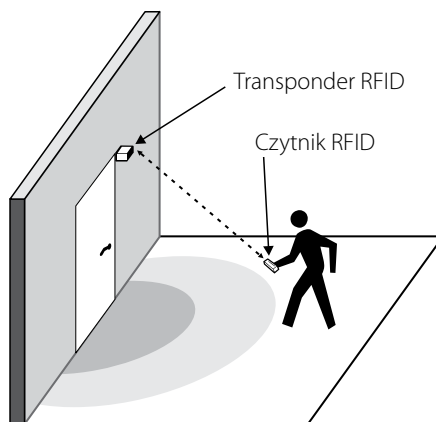
Wdrażanymi systemami poprawek różnicowych są amerykański WAAS (*Wide Area Augmentation System*) oraz europejski EGNOS (*European Geostationary Navigation Overlay Service*). Są to systemy o dużym zasięgu, działające na podobnej zasadzie jak DGPS, przy czym funkcje stacji referencyjnych pełnią w tych systemach satelity geostacjonarne.

Ważnym parametrem w globalnych systemach pozycjonowania jest czas potrzebny do wyznaczenia położenia geograficznego po załączeniu odbiornika (*time to first fix – TTFF*). Zależy czy jest to włączenie odbiornika po kilkuminutowej przerwie (*warm or hot start*) czy też pierwsze włączenie odbiornika (*cold start*) TTFF może wynosić od kilku sekund do kilkunastu minut (van Diggelen, 2009). Satelity GPS rozsyłają co 12,5 minuty informacje o położeniu wszystkich satelitów (tzw. almanach). Utrata informacji o układzie satelitów wymaga przestrajania częstotliwości odbiornika w poszukiwaniu sygnałów od satelitów w całym zakresie pasma transmisji używanych w GPS z uwagi

na przesunięcie częstotliwości widzianej przez odbiornik powodowane efektem Dopplera (satelity poruszają się w różnych kierunkach względem powierzchni Ziemi z prędkością ok. 11 tys. km/godz.).

Operatorzy sieci komórkowych udostępniają usługę lokalizacyjną o nazwie AGPS (*assisted GPS*) umożliwiającą skrócenie TTFF do ok. 1 s. Sieć rozproszonych odbiorników GPS operatora nieustannie monitoruje położenie satelitów i przesyła te informacje do odbiornika użytkownika (np. z wykorzystaniem pakietowej transmisji danych GPRS systemu GSM).

Zagadnienie wyznaczania położenia wewnątrz budynków ma mniejsze znaczenie w systemach nawigacji dla osób niewidomych. Niewidomi poruszają się dość sprawnie w budynkach o znanych rozkładzie przestrzennym (np. w swoim mieszkaniu). W nieznanach budynkach bardziej istotna jest dla nich informacja o infrastrukturze budynku i jego obiektach (np. lokalizacja wind, schodów, informacja o numerze piętra, rozkładzie pomieszczeń) niż informacja o dokładnym własnym położeniu (Kwapiszowie, 1990). Istnieje duża liczba technologii wyznaczania położenia wewnątrz budynków m.in. z użyciem ultradźwięków, lokalnych sieci radiowych WLAN i systemów wizyjnych (zob. obszerną pracę przeglądową Gu i innych (2009)). Jednak, z uwagi na wspomniane uwarunkowania, w systemach nawigacji wewnątrz budynków jest wykorzystywana głównie technologia układów identyfikacji radiowej RFID (*radio-frequency identification*). Idea działania tego systemu opiera się na koncepcji latarni nawigacyjnych rozmieszczonych w charakterystycznych punktach otoczenia (rys. 3.8), które komunikują się z użytkownikiem wyposażonym w odpowiednie urządzenie zwane czytnikiem (*reader*).



Rys. 3.8. Zastosowanie układów identyfikacji radiowej RFID do nawigacji osób niewidomych wewnątrz budynków

Źródło: opracowanie własne

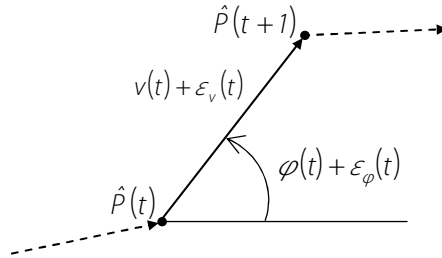
Funkcje radiolatarni pełnią tzw. transpondery RFID (*RFID tag*). Wyróżnia się dwa rodzaje transponderów: transpondery pasywne, które nie wymagają zasilania i transpondery aktywne (Want, 2006). Transpondery pasywne są załączane energią pola elektromagnetycznego wysyłanego z czytnika. Zaindukowana energia w antenie transpondera wystarcza na zwrotne przesłanie do czytnika krótkiego kodu, np. numeru identyfikacyjnego transpondera. Transpondery aktywne wymagają własnego źródła zasilania ale mogą wysłać sygnał zwrotny do czytnika na odległość do ok. 6m. Zasięg komunikacji czytnika z transponderem zależy od zastosowanych mocy nadajników (typowo w zakresie $10\text{ }\mu\text{W}$ ÷1 mW) oraz częstotliwości transmisji sygnałów (125÷134 kHz, 10÷15 MHz, 868÷956 MHz i 2,4÷5,8 GHz) (Ward i van Kranenburg, 2006).

W ostatnich kilku latach opracowano prototypowe systemy nawigacji dla osób niewidomych z wykorzystaniem układów RFID. Niektóre z nich są na etapie wdrażania. Krótki przegląd takich systemów instalowanych wewnątrz budynków jak i w terenie otwartym omówiono w rozdziale 5. Obecnie podstawowym zastosowaniem układów RFID jest elektroniczne znakowanie pojazdów i produktów w sieciach zaopatrzenia.

3.5. Nawigacja zliczeniowa

Ideę nawigacji zliczeniowej (*dead reckoning*) pokazano na rys. 3.9. Polega ona na przyrostowym wyznaczaniu kolejnych położenia obiektu, tj. jego bieżącego położenia $\hat{P}(t)$ na podstawie poprzedniego położenia $\hat{P}(t-1)$ oraz predykcji następnego położenia $\hat{P}(t+1)$ na podstawie bieżącego położenia itd. Kolejno wyznaczone położenia są obliczane na podstawie parametrów ruchu obiektu, tj. jego chwilowej prędkości $v(t)$ oraz kierunku ruchu $\varphi(t)$. W nawigacji zliczeniowej kolejno estymowane parametry kinematyczne są uzyskiwane z pewną niepewnością. Skutkuje to kumulacją błędów dla kolejno wyliczanych położenia obiektu.

W zależności od zastosowanej techniki nawigacji zliczeniowej ruch obiektu jest wygodniej wyznaczać względem nawigacyjnego układu współrzędnych (*navigation frame*, *n-frame*) lub względem układu współrzędnych związanych z poruszającym się obiektem (*body frame*, *b-frame*).



Rys. 3.9. Ilustracja idei nawigacji zliczeniowej polegającej na kolejnym wyznaczaniu położenia obiektu; wielkości $\varepsilon_v(t)$ oraz $\varepsilon_\varphi(t)$ oznaczają odpowiednio niepewności pomiaru prędkości oraz kierunku ruchu obiektu

Źródło: opracowanie własne

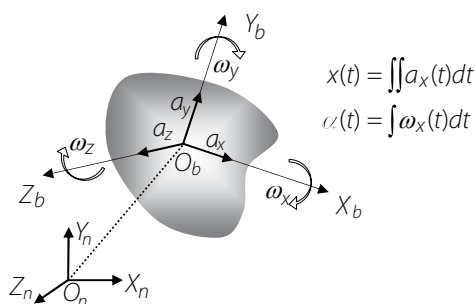
3.5.1. Nawigacja inercyjna

W nawigacji inercyjnej (bezwładnościowej), parametry ruchu obiektu wyznacza się przyjmując, że obiekt jest materialny (posiada niezerową masę) oraz porusza się ruchem niejednostajnym względem pewnego nawigacyjnego układu współrzędnych (Narkiewicz, 1999). Zgodnie z II zasadą dynamiki Newtona na taki obiekt działa siła bezwładności $\vec{F} = -m\vec{a}$ gdzie: m – jest masą obiektu oraz $\vec{a}$ – jest wektorem przyspieszenia. Wektor siły ma zwrot przeciwny do zwrotu wektora przyspieszenia.

Jeżeli obiekt przemieszcza się ruchem swobodnym (bez więzów) w przestrzeni trójwymiarowej to do jednoznacznego opisu ruchu tego obiektu należy rejestrować 6 parametrów (trzy składowe przesunięcia i trzy kąty obrotów), które są definiowane względem nawigacyjnego układu współrzędnych $O_n X_n Y_n Z_n$ (n -frame) (rys. 3.10). Obiekt taki określa się mianem układu o 6 stopniach swobody (*six degrees of freedom* – 6DoF).

W nawigacji inercyjnej pomiar przemieszczenia jest wyznaczany na podstawie podwójnego całkowania składowych przyspieszeń a_x , a_y , a_z (3 DoF) wynikających z ruchu obiektu (należy pamiętać, że pomiary przyspieszeń są wykonywane dla układu, na który oddziałuje przyspieszenie ziemskie) (Wu i inni, 2002).

Do pomiaru przyspieszenia, w szczególności w małych urządzeniach mobilnych stosuje się obecnie mikrosystemy elektromechaniczne (*micro electro-mechanical systems* – MEMS), (Wiak, 2009). W urządzeniach MEMS wykorzystuje się pojemnościowe czujniki przyspieszenia, w których jedna z okładek kondensatora jest ruchoma i zmienia swoje położenie wraz ze zmianą przyspieszenia, co powoduje zmianę pojemności kondensatora pomiarowego. Pomiar kątów obrotu jest realizowany w układach MEMS, m.in. przez pomiar prędkości obrotu



Rys. 3.10. Sposób wyznaczania przemieszczenia oraz obrotu bryły o sześciu stopniach swobody na podstawie odpowiednich operacji całkowania składowych przyspieszenia i prędkości obrotowych; układ współrzędnych $O_b X_b Y_b Z_b$ (*b-frame*) jest tzw. nieinercyjnym układem odniesienia względem układu nawigacyjnego $O_n X_n Y_n Z_n$ (*n-frame*)

Źródło: opracowanie własne

na podstawie pomiaru siły Coriolisa oddziałującej na mikroelement rezonujący wzdłuż promienia wodzącego obrotu. Moduł elektroniczny wyposażony w czujnik inercyjny MEMS, scalony odbiornik GPS oraz mikrokontroler został wykonany w Instytucie Elektroniki PŁ (rys. 3.11).

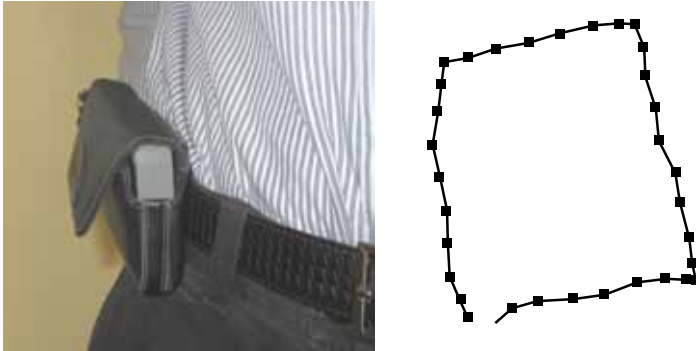


Rys. 3.11. Moduł elektroniczny zawierający trójosiowy czujnik przyspieszenia, trójosiowy żyroskop oraz odbiornik GPS (wykonano w Instytucie Elektroniki PŁ)

Źródło: fotografia własna

Nawigację inercyjną stosuje się m.in. do zliczania i wyznaczania długości kroków osoby pieszej (Jirawimut i inni, 2003). Pomiar długości kroków jest szczególnie ważnym parametrem w systemach nawigacji dla osób niewidomych. Estymacja przebytej drogi na podstawie liczby kroków w przypadku osób niewidomych zawodzi (m.in. stwierdzono, że długość stawianych kroków u osób niewidomych jest mniejsza przy pokonywaniu nowych ścieżek). Na rysunku 3.12, pokazano prototypowy układ z czujnikiem inercyjnym przy-

twierdzonym do pasa. Pomiar długości kroku jest estymowany na podstawie empirycznie ustalonej metody pomiaru wielkości składowej przyspieszenia w kierunku przyspieszenia ziemskiego (poddanej filtracji górnoprzepustowej dla usunięcia składowej przyspieszenia grawitacyjnego). System umożliwił znaczną poprawę estymacji długości przebytej drogi w porównaniu z typowym krokomierzem (Barański i Strumiłło, 2009a).



Rys. 3.12. Prototypowy układ do detekcji kroków i pomiaru ich długości (z lewej) oraz linia ścieżki marszu wzdłuż boków kwadratu o dł. 5 m wyznaczona na podstawie nawigacji inercyjnej (z prawej); wykryte kolejne kroki zaznaczono małymi kwadratami

Źródło: opracowanie własne

Moduł elektroniczny pokazany na rys. 3.12. wykorzystano również w hybrydowym systemie nawigacji, w którym estymacja położenia geograficznego jest wyznaczana na podstawie algorytmu łączącego dane z nawigacji bezwładnościowej i globalnego systemu pozycyjnego. Opis tego systemu zamieszczono w dalszej części tego rozdziału.

3.5.2. Nawigacja wizyjna

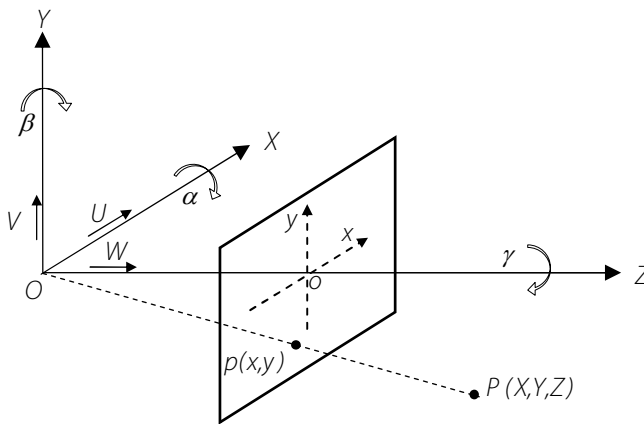
W nawigacji wizyjnej, zwanej też odometrią wizyjną (*visual odometry*) (Bruss i Horn, 1983) wyznaczanie położenia ruchomego obiektu odbywa się wyłącznie na podstawie rejestracji sekwencji obrazów sceny trójwymiarowej. Ruch systemu wizyjnego, tzw. ruch własny (*egomotion*), jest estymowany na podstawie analizy przesunięcia punktów sceny (przepływu optycznego, *optical flow*) wykrywanego w sąsiednich obrazach sekwencji wideo.

W systemach odometrii wizyjnej stosuje się zarówno układy z pojedynczymi kamerami (Garcia i Tziritas, 2001), (Rumeliotis i inni, 2002) jak i układy stereo-

skopowe (Nister i inni, 2004), (Milella i Siegart, 2006). Wykorzystanie informacji o głębi znacznie upraszcza algorytmy estymacji ruchu systemu wizyjnego. W systemach jednokamerowych dla wyznaczenia ruchu w jednostkach fizycznych jest wymagana dodatkowa wiedza *a priori* o strukturze geometrycznej sceny lub ograniczenie stopni swobody estymowanego ruchu (Saxena i inni, 2009).

Zostanie rozważony przypadek ruchu swobodnego kamery stereoskopowej (np. przytwierdzonej do ciała osoby pieszej) oraz nałożone następujące warunki na ten ruch i parametry układu akwizycji obrazu:

- kamery są wzajemnie skalibrowane,
- przesunięcia $\Delta \mathbf{T}$ i kąty obrotu α, β, γ układu kamery w interwałach pomiędzy rejestracją kolejnych obrazów są niewielkie (ruch kamery można przybliżyć odcinkiem linii),
- w czasie ekspozycji pojedynczego obrazu przemieszczenie obrazów elementów sceny jest dwukrotnie mniejsze niż odstęp próbkowania obrazu.



Rys. 3.13. Układ współrzędnych ze zdefiniowanymi parametrami przesunięć

$\Delta \mathbf{T} = [U, V, W]^T$ i obrotów α, β, γ tego układu

Źródło: opracowanie własne

Niech początek układu współrzędnych *b-frame* $OXYZ$ znajduje się w ognisku prawej kamery układu stereowizyjnego (rys. 3.13). Na skutek ruchu translacyjnego $\mathbf{T}$ i obrotowego² $\mathbf{R}$ tego układu współrzędne punktu P sceny przyjmą nowe wartości, które można wyznaczyć z zależności (Horn i Schunck, 1981):

² Obrót w przestrzeni trójwymiarowej można definiować jako złożenie kolejnych obrotów $\mathbf{R}_\alpha, \mathbf{R}_\beta, \mathbf{R}_\gamma$ wokół odpowiednich osi, tj. $\mathbf{R} = \mathbf{R}_\alpha \mathbf{R}_\beta \mathbf{R}_\gamma$ gdzie:

$$\mathbf{R}_\alpha = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}, \mathbf{R}_\beta = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix}, \mathbf{R}_\gamma = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$P' = -RP - T \quad (3.1)$$

Dla przyjętych niewielkich przemieszczeń kamery w kolejnych interwałach pomiędzy rejestrowanymi obrazami równanie (3.1) można zapisać w postaci (Ma i inni, 2003):

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = \begin{bmatrix} 0 & \gamma & -\beta \\ -\gamma & 0 & \alpha \\ \beta & -\alpha & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \begin{bmatrix} U \\ V \\ W \end{bmatrix} \quad (3.2)$$

gdzie: $\dot{X}$, $\dot{Y}$, $\dot{Z}$ – są składowymi prędkości przemieszczania punktów sceny względem kamery, U , V , W – są składowymi prędkości przemieszczenia kamery w scenie oraz α , β , γ – są prędkościami kątowymi obrotu kamery względem osi układu współrzędnych *b-frame*. Przyjmując model perspektywiczny kamery, w którym współrzędne punktu obrazu $p(x,y)$ będącego rzutem punktu sceny $P(X,Y,Z)$ na płaszczyznę obrazu wiążą zależności: $x = fX/Z$, $y = fY/Z$, składowe (u, v) prędkości przemieszczania się punktu p w obrazie na skutek ruchu kamery określają równania przepływu optycznego:

$$u = \frac{dx}{dt} = f \left(\dot{X} \frac{1}{Z} - \frac{X}{Z^2} \dot{Z} \right) \quad (3.3)$$

$$v = \frac{dy}{dt} = f \left(\dot{Y} \frac{1}{Z} - \frac{Y}{Z^2} \dot{Z} \right) \quad (3.4)$$

Po podstawieniu wzorów na prędkości $\dot{X}$, $\dot{Y}$, $\dot{Z}$ z równania (3.2) do równań (3.3) i (3.4) oraz ponownym zastosowaniu równań rzutu perspektywicznego równania te przyjmują postać (Horn i Schunck, 1981):

$$u = \frac{-Uf + xW}{Z} + \alpha \frac{xy}{f} - \beta \left(\frac{x^2}{f} + f \right) + \gamma y \quad (3.5)$$

$$v = \frac{-Vf + yW}{Z} - \beta \frac{xy}{f} + \alpha \left(\frac{y^2}{f} + f \right) - \gamma x \quad (3.6)$$

Należy zaznaczyć, że kolejność obrotów nie jest przemienne, zmiana ich kolejności prowadzi do różnych położeń końcowych obiektu. Bardziej szczegółowe omówienie zagadnień związanych z tym aspektem nawigacji zliczeniowej można znaleźć w pracy Farrell (2008). Wprowadzono w niej pojęcie kwaternionów (szczególnej klasy liczb zespolonych), dla których obroty można definiować jako macierze o elementach będących funkcjami liniowymi. Uzyskuje się wtedy znaczne uproszczenie całkowania numerycznego stosowanego w inercjalnych układach nawigacji.

Są to równania wiążące prędkość przemieszczania punktu w obrazie (u, v) z prędkością przemieszczania układu wizyjnego wyznaczoną wektorem parametrów $[U, V, W, \alpha, \beta, \gamma]$. Pamiętajmy z rozdz. 2, że głębie punktu sceny opisuje zależność $Z = fB/d$, gdzie: B – jest bazą układu stereowizyjnego, a d jest dysparycją. Po podstawieniu tej zależności do równań (3.5) i (3.6) oraz przedstawieniu tych zależności w zapisie macierzowym równania te można zapisać w postaci:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -\frac{d}{B} & 0 & \frac{xd}{fB} \\ 0 & -\frac{d}{B} & \frac{yd}{fB} \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix} + \begin{bmatrix} \frac{xy}{f} & -\left(\frac{x^2}{f} + f\right) & y \\ \left(\frac{x^2}{f} + f\right) & -\frac{xy}{f} & -x \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \quad (3.7)$$

w którym, wyodrębniono składowe translacyjne ruchu kamery i składowe rotacyjne.

Wyznaczenie sześciu szukanych parametrów ruchu kamery $[U, V, W, \alpha, \beta, \gamma]$ występujących w równaniu (3.7) można wykonać wg następujących kroków obliczeniowych:

- 1) wykryj N ($N \geq 3$) punktów obrazu dla uzyskania minimum trzech układów równań (3.7), których składowe prędkości (u'_n, v'_n), $n = 1, \dots, N$ będą wyznaczone (są to tzw. punkty kluczowe (*key points*)), dla których wyznaczono dokładną estymatę głębi (Rzeszotarski i inni, 2007);
- 2) wyznacz współrzędne punktów kluczowych w kolejnych dwóch obrazach, tj. dla chwil t i $t+1$;
- 3) oblicz przesunięcia dla każdego z punktów $n = 1, \dots, N$ w kolejnych obrazach:

$$\begin{aligned} u'_n(t) &= x_n(t+1) - x_n(t) \\ v'_n(t) &= y_n(t+1) - y_n(t) \end{aligned} \quad (3.8)$$

- 4) wyznacz błąd średniokwadratowy postaci:

$$\begin{aligned} \varepsilon &= \sum_{n=1}^N \left[(u_n - u'_n)^2 + (v_n - v'_n)^2 \right] = \\ &= \sum_{n=1}^N \left[(u_n + x_n(t) - x_n(t+1))^2 + (v_n + y_n(t) - y_n(t+1))^2 \right] \end{aligned} \quad (3.9)$$

gdzie: u_n, v_n są dane równaniami (3.8);

- 5) wyznacz 6 poszukiwanych wartości parametrów ruchu kamery $U, V, W, \alpha, \beta, \gamma$ na podstawie rozwiązania tzw. równań normalnych otrzymanych przez ob-

liczenie pochodnych cząstkowych równania (3.9) względem poszukiwanych parametrów i przyrównanie ich do zera.

6) powtarzaj obliczenia wg kroków 2÷5 dla każdego kolejnego obrazu.

Należy podkreślić, że parametry ruchu kamery są wyznaczane na drodze obliczeń analitycznych bez potrzeby stosowania czasochłonnych obliczeń optymalizacyjnych.

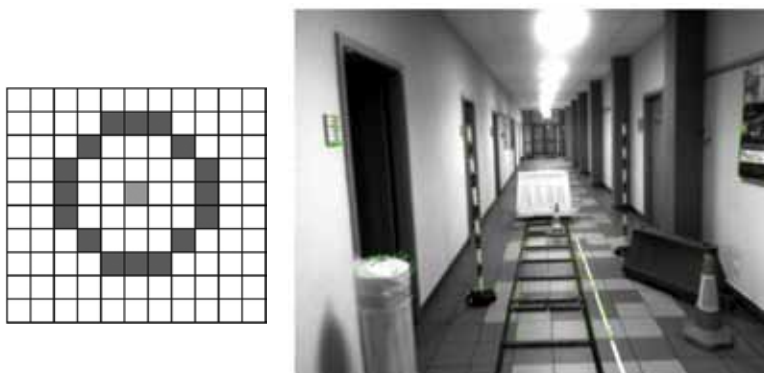
Odrębnym problemem obliczeniowym jest zadanie wykrywania i śledzenia punktów kluczowych (zauważmy, że dla kolejnych obrazów należy również obliczać współrzędne głębi tych punktów (zob. rozdz. 2)). Za dobre punkty kluczowe uważa się obszary obrazu o dużej lokalnej zmianie jasności obrazu, np. brzegi obiektów, przecięcia linii. Dla takich obszarów sceny głębię można estymować z dużą wiarygodnością (zob. rys. 2.23 w rozdz. 2 oraz (Rzeszotarski i inni, 2007)).

Podstawowym detektorem punktów kluczowych, jest tzw. detektor Harrisa³, w którym są wyznaczane gradienty kierunkowe obrazu. Za bardziej skuteczne uznaje się obecnie detektory: FAST (*features from accelerated segment test*) (Rosten i Drummond, 2005) i detektor zaproponowany przez Shi i Tomasiego (1994) będący modyfikacją detektora Harrisa (dostępny w bibliotece przetwarzania i analizy obrazów Open Source Computer Vision (OpenCV)). Prostym koncepcyjnie ale skutecznym i szybkim sposobem wyznania punktów kluczowych jest zastosowanie detektora FAST. W detektorze tym badane są jasności punktów obrazu leżące na tzw. okręgu Bresenham, którego środek pokrywa się z analizowanym punktem obrazu. Punkt obrazu jest uznawany za kluczowy jeżeli jasność zadanej liczby sąsiednich punktów okręgu jest większa (lub mniejsza) od jasności punktu leżącego w środku tego okręgu. Na rys. 3.14 pokazano okrąg Bresenham o promieniu $r = 3$.

Do śledzenia punktów kluczowych w kolejnych obrazach można zastosować metodę dopasowania bloków wg miary sum wartości bezwzględnych (SAD) zdefiniowanej w rozdz. 2 i zastosowanej do rozwiązywania problemu korespondencji w stereowizji (Skulimowski, 2009). Na rys. 3.15 pokazano pojedynczy obraz sekwencji z wykrytymi punktami kluczowymi i ich przemieszczeniami wyznaczonymi wg minimów miary SAD. Ostatnio zaproponowano metody szybkiego przeszukiwania wektorów ruchu na potrzeby kompresji sekwencji stereoskopowych (Yang i inni, 2009), (Konieczny i Domański, 2010).

³ Detektor Harrisa jest definiowany macierzą postaci: $H(x, y) = \begin{bmatrix} I_x^2(x, y) & I_x I_y(x, y) \\ I_x I_y(x, y) & I_y^2(x, y) \end{bmatrix}$,

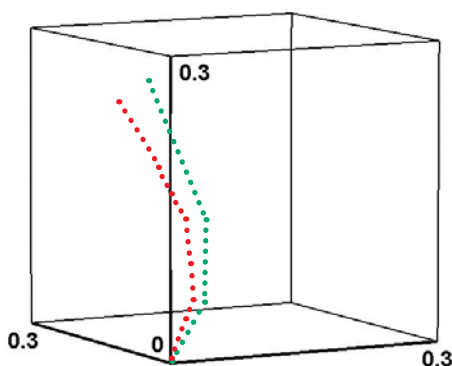
gdzie: I_x, I_y oznaczają odpowiednio pochodne cząstkowe obrazu wzdłuż wierszy i kolumn. Dla obszaru obrazu zawierającego punkt kluczowy (np. przecięcie linii) wartości własne macierzy H są dużymi liczbami dodatnimi.



Rys. 3.14. Okrąg Breshenhama ($r = 3$) detektora FAST (z lewej) i pojedynczy obraz sekwencji wideo z wykrytymi punktami kluczowymi i ich przesunięciami (zielone odcinki linii) w kolejnych dwóch obrazach (z prawej)

Źródło: opracowanie własne

Opisana metoda wyznaczania parametrów ruchu kamery jest metodą zliczeniową, w której błędy estymacji położenia kamery kumulują się dla kolejnych przetwarzanych ramek obrazu. W pracy Skulimowskiego (2009) przeprowadzono ilościową weryfikację tej techniki dla sekwencji obrazów wytworzonych z zastosowaniem biblioteki graficznej OpenGL, w których są dokładnie znane parametry ruchu kamery w scenie. Na rys. 3.15. zilustrowano efekt rozbiegania się trajektorii ruchu kamery i estymowanego ruchu kamery wyznaczonego wg opisanej metody odometrii wizyjnej.



Rys. 3.15. Weryfikacja metody wyznaczania ruchu kamery stereowizyjnej (skala na osiach niemianowana); w liniach kropkowanych każdy punkt odpowiada kolejno rejestrowanym obrazom; kolor zielony – zadany ruch kamery, kolor czerwony – ruch wyznaczony na podstawie odometrii wizyjnej

Źródło: Skulimowski (2009)

Nawigacja zliczeniowa na podstawie odometrii wizyjnej charakteryzuje się następującymi wadami:

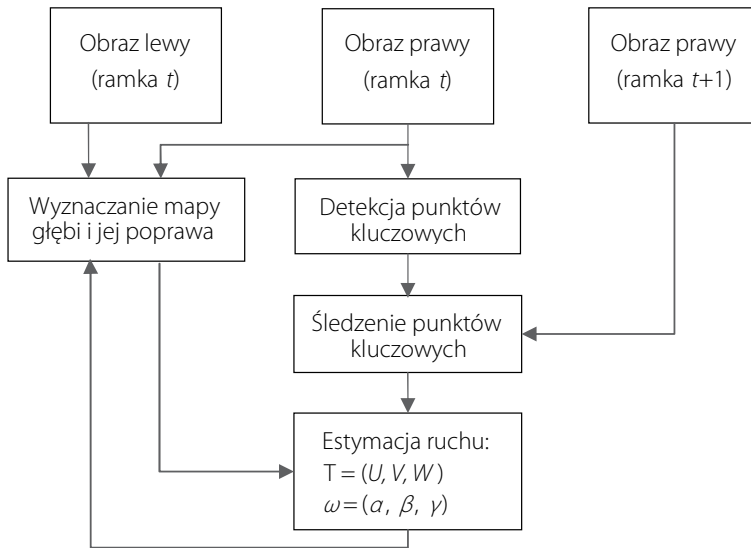
- jest techniką nawigacji zliczeniowej, w której kumulują się kolejne błędy estymacji ruchu,
- może być stosowana do nawigacji w scenach tylko przy odpowiednim oświetleniu i występowaniu w scenie odpowiedniej liczby punktów kluczowych,
- procedura dopasowania blokowego w dziedzinie przestrzennej obrazu do obliczania przesunięć punktów kluczowych wymaga dużego nakładu obliczeń, w szczególności dla dużych przesunięć punktów pomiędzy kolejno rejestrowanymi obrazami,
- dla szybkich przemieszczeń kamery np. szybkich rotacji (w wyniku ograniczeń związanych z szybkością akwizycji obrazów i ich przetwarzania (<50 obrazów/s)), następuje „rozmycie” obrazów i skuteczna detekcja punktów kluczowych nie jest możliwa,
- odometria wizyjna jest utrudniona lub niemożliwa dla scen, w których występują obiekty ruchome.

Pomimo kumulacji błędów estymacji ruchu kamery, wyznaczone parametry ruchu mogą posłużyć do następujących celów:

- poprawy dokładności wyznaczanych sekwencji map głębi (Skulimowski i Strumiłło, 2008); problemu szczególnie istotnego dla małej bazy mobilnego układu stereoskopowego,
- poprawy dokładności nawigacji zliczeniowej w systemie hybrydowym składającym się z nawigacji inercyjnej i odometrii wizyjnej (Pęczniński i inni, 2011).

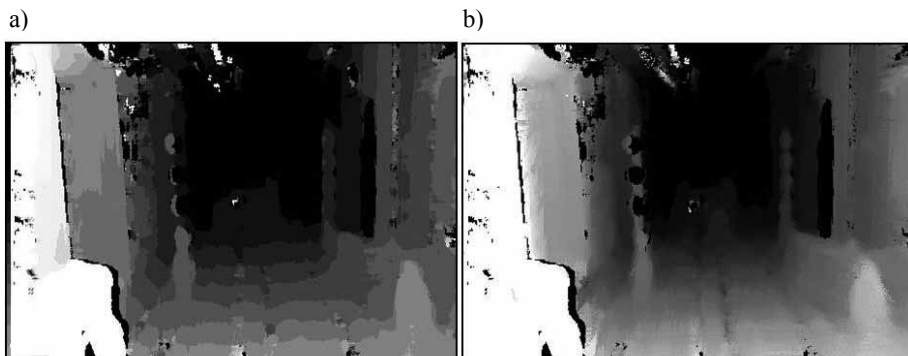
Poprawę dokładności wyznaczanych map głębi uzyskano korzystając z mapy głębi wyznaczonej w chwili t , estymowanego ruchu kamery $[U, V, W, \alpha, \beta, \gamma]$ oraz wyznaczonych z równania (3.7) przesunięć (u, v) do predykcji obrazów w chwili $t+1$ i na tej podstawie predykcji mapy głębi. Dla uniknięcia kumulacji błędów predykcji tak wyznaczone mapy głębi są na bieżąco porównywane z mapami wyznaczonymi metodami blokowymi i odpowiednio korygowane. Na rys. 3.16 pokazano schemat blokowy opisanej procedury, a na rys. 3.17 zilustrowano wynik korekcji mapy głębi. Widoczna jest poprawa gładkości mapy głębi, która dokładniej odwzorowuje strukturę geometryczną sceny. Procedurę tę wykorzystano w algorytmie detekcji i śledzenia obiektów sceny w systemie obrazowania dźwiękowego otoczenia (Skulimowski i Strumiłło, 2006, 2007), zob. też rozdział 5.

Zastosowanie systemu integrującego techniki nawigacji bezwładnościowej i odometrii wizyjnej umożliwia poprawę skuteczności śledzenia ruchu własnego obiektu (Klein i Drummond, 2004), (Pęczniński i inni, 2011).



Rys. 3.16. Schemat blokowy procedury do jednoczesnej estymacji ruchu układu stereoskopowego i poprawy dokładności obliczanych map głębi

Źródło: opracowanie własne



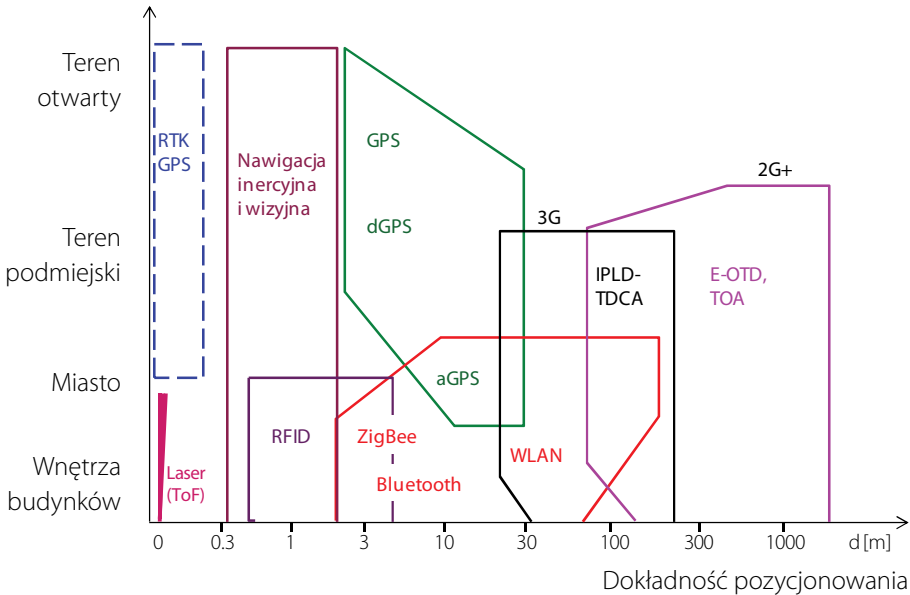
Rys. 3.17. Mapa głębi wyznaczona z dokładnością „pikselową” na podstawie metody dopasowania blokowego (a), mapa głębi wyznaczona z dokładnością „subpikselową” z wykorzystaniem parametrów ruchu kamery (b)

Źródło: opracowanie własne

3.5.3. Techniki integracji danych nawigacyjnych

Na rys. 3.18 porównano metody wyznaczania położenia z zastosowaniem różnych technik pozycjonowania bezwzględnego (głównie z zastosowaniem nadajników radiowych) oraz technik pozycjonowania względnego (nawigacja

inercyjna i wizyjna). Jak pokazano na rysunku, zarówno zakres zastosowania (skala przestrzenna i rodzaj terenu) oraz dokładność pozycjonowania z wykorzystaniem wskazanych technik są różne (Adusei i inni, 2002). Skłania to do łączenia tych technik dla osiągnięcia lepszej dokładności pozycjonowania oraz rozszerzenia zakresu ich zastosowań.



Rys. 3.18. Porównanie szacunkowych dokładności wyznaczania położenia z zastosowaniem różnych technik pozycjonowania bezwzględnego i względnego; objaśnienia skrótów:

RTK GPS (*real time kinematic GPS*)⁴, TOA (*time of arrival*), IPLD-TDCA (*idle period downlink-observed time difference of arrival*), E-OTD (*enhanced observed time difference*), 2G, 3G – generacje sieci telefonii komórkowej

Źródło: opracowanie własne na podstawie (Adusei i inni, 2002)

W wielu pracach badawczych jest podejmowany problem poprawy dokładności wyznaczania położenia z zastosowaniem rozwiązań hybrydowych, w których łączy się globalne systemy pozycjonowania z innymi technikami, m.in. z zastosowaniem lokalnych sieci radiowych (Fang i inni, 2005), (Bensky, 2008) sieci stacji bazowych telefonii komórkowej (Adusei i inni, 2002) oraz metod nawigacji zliczeniowej (inercyjnej i wizyjnej) (Farell, 2008). Podejmowany jest również problem poprawy położenia dla zastosowań w nawigacji pieszej (Ceranka i Niedźwiecki, 2003), (Mezentsev i inni, 2005). Stosowane są w tych systemach różne

⁴ RTK GPS – bardzo dokładna i szybka (ale bardzo droga) technika wyznaczania położenia (van Diggelen, 2009)

techniki integracji danych nawigacyjnych od prostych reguł heurystycznych do bardziej zaawansowanych metod łączenia (*data fusion*) i filtracji danych z różnych źródeł, m.in. z zastosowaniem filtracji Kalmana i filtracji cząsteczkowej (*particle filtering*) (Gustaffson i inni, 2002). Potrzeba budowania rozwiązań umożliwiających uzyskanie dokładniejszego określania położenia geograficznego jest szczególnie istotna w nawigacji terenowej dla osób niewidomych. Wynika ona ze specyficznej charakterystyki pieszego przemieszczania się osoby niewidomej w porównaniu do innych zadań nawigacji, np. samochodowej, tj.:

- osoba niewidoma może zboczyć z zaplanowanej ścieżki marszu, np. z chodnika (w nawigacji samochodowej algorytmy estymacji położenia geograficznego pojazdów ograniczają zbiór możliwych lokalizacji do sieci dróg),
- marsz osoby niewidomej przebiega wg nieregularnego rytmu (częsty „marsz w miejscu” wokół własnej osi) i o znacznym zróżnicowaniu długości stawianych kroków,
- osoba niewidoma porusza się wzdłuż toru ruchu o złożonym kształcie (częste zmiany kierunku marszu, zatrzymania i obroty), tor ruchu w nawigacji samochodowej wynika z gładkiej linii pasa drogi,
- pokonywana pieszo trasa może przebiegać w obszarach braku sygnałów GPS (przejścia podziemne, dworce, centra handlowe).

Dodatkowym utrudnieniem jest ciągle niedostateczna dostępność map terenu uwzględniających ciągi piesze.

Wymienione uwarunkowania czynią zagadnienie budowania bezpiecznych systemów nawigacji terenowej dla osób niewidomych szczególnie trudnym (Hersh i Johnson, 2008). W niniejszej części książki skupiono się na problemie osobistej nawigacji terenowej, w której zastosowano integrację układów nawigacji globalnej i zliczeniowej dla poprawy dokładności wyznaczanego położenia geograficznego.

3.5.4. Zastosowanie filtracji Kalmana i filtracji cząsteczkowej w nawigacji pieszej

Piesze przemieszczanie się można interpretować jako pewien proces dynamiczny, którego zmienne stanu podlegają określonym regułom zmian. Na potrzeby rozważanego w dalszej części opisu problemu estymacji położenia, zmienne stanu zostaną ograniczone do bieżącego położenia pieszego $s(t) = [x(t) \ y(t)]^T$ po wykonaniu kroku o długości $d(t)$ w wybranym kierunku $\varphi(t)$. Wektor $u(t) = [d(t) \ \varphi(t)]^T$ jest tzw. wektorem wymuszenia (sterowania) takiego układu. Zarówno zmienne stanu jak i sterowanie nie są bezpośrednio dostępne

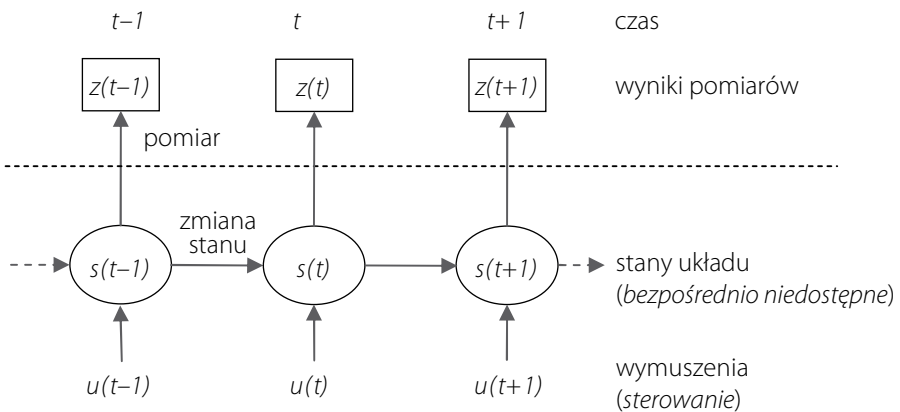
i podlegają pomiarom, które są obarczone błędami, np. błędami odczytów GPS w przypadku pomiaru położenia, błędami pomiaru długości kroku oraz błędami wyznaczania chwilowego kierunku marszu z czujników przyspieszeń i żyroskopów. Tak określony układ można opisać równaniami:

$$\begin{cases} s(t) = f[s(t-1), u(t), n(t-1)] \\ z(t) = h[s(t), v(t)] \end{cases} \quad (3.10)$$

gdzie:

$f[.]$ jest różniczkowalną funkcją wektorową (w ogólności nieliniową) opisującą związek pomiędzy bieżącym stanem układu $s(t)$, a bieżącym sterowaniem $u(t)$ i stanem układu w kroku poprzednim $s(t-1)$ w obecności zakłóceń sterowania $n(t-1)$, $h[.]$ jest różniczkowalną funkcją wektorową (również w ogólności nieliniową) opisującą związek pomiędzy aktualnym stanem układu $s(t)$, a pomiarem $z(t)$ tego stanu w obecności zakłóceń pomiaru $v(t)$.

Układ równań (3.10) określa się mianem rozszerzonego filtra Kalmana (*extended Kalman filter* – EKF) w odróżnieniu od filtra Kalmana (KF), w którym na funkcje $f[.]$, $h[.]$ nakłada się warunek liniowości (Kalman, 1960). Równania (3.10) opisują układ autonomiczny (stacjonarny) o nieliniowej dynamice i nieliniowej zależności pomiędzy wynikiem pomiaru a stanem układu. Pierwsze z równań (3.10) jest nazywane równaniem przejść stanu (ewolucji stanu) a drugie równaniem pomiaru (obserwacji). Dla EKF i KF przyjmuje się, że rozkłady statystyczne zakłóceń są szumami gaussowskimi o zerowej wartości średniej. Na rys. 3.19



Rys. 3.19. Ciąg stanów układu dynamicznego (bezpośrednio niedostępnego obserwatorowi), których estymaty są wyznaczane na podstawie kolejnych pomiarów i wymuszeń zewnętrznych układu z zastosowaniem filtracji Kalmana

Źródło: opracowanie własne

pokazano ciąg stanów takiego systemu. Bieżący stan układu zależy bezpośrednio od poprzedniego stanu i poprzedniego wymuszania. Taki proces nazywamy procesem Markowa pierwszego rzędu.

Celem filtracji jest wyznaczenie niedostępnego stanu układu $\mathbf{s}(t)$ (tj. kolejnego położenia) będącego wynikiem ciągu wymuszeń $\dots, \mathbf{u}(t-2), \mathbf{u}(t-1), \mathbf{u}(t)$ (tj. kolejno stawianych kroków) na podstawie dostępnego ciągu pomiarów $\dots, \mathbf{z}(t-2), \mathbf{z}(t-1), \mathbf{z}(t)$ (tj. odczytów położenia pieszego uzyskiwanych z odbiornika GPS oraz przyjętego modelu układu dynamicznego). Rekursywne estymowanie ciągu stanów dla równań nieliniowych danych w postaci (3.10) w obecności zakłóceń (które podlegają nieliniowym transformacjom) jest w praktyce niemożliwe (Korbicz i inni, 1994). Jednym z możliwych rozwiązań problemu jest linearyzacja funkcji (za pomocą pierwszych wyrazów szeregu Taylora) lub takie zdefiniowanie problemu filtracji, dla której można zastosować liniowy filtr Kalmana (Anderson i Moore, 1984).

Filtr liniowy Kalmana (KF) jest zdefiniowany układem równań:

$$\begin{cases} \mathbf{s}(t) = \mathbf{F}\mathbf{s}(t-1) + \mathbf{G}\mathbf{u}(t) + \mathbf{n}(t-1) \\ \mathbf{z}(t) = \mathbf{H}\mathbf{s}(t) + \mathbf{v}(t) \end{cases} \quad (3.11)$$

gdzie dla rozważanego problemu estymacji położenia określono dwie zmienne stanu dane wektorem $\mathbf{s}(t) = [x(t) \ y(t)]^T$, dwa sygnały sterujące $\mathbf{u}(t) = [d(t) \ \varphi(t)]^T$ oraz dwa sygnały mierzone $\mathbf{z}(t) = [\text{GPS}_x(t) \ \text{GPS}_y(t)]^T$.

$\mathbf{F}$ – jest macierzą (2×2) wiążącą stan układu w chwili $t-1$ i w chwili t ,

$\mathbf{G}$ – jest macierzą (2×2) określającą związek pomiędzy sterowaniem i stanem układu w chwili t ,

$\mathbf{H}$ – jest macierzą (2×2) wiążącą stan układu z jego pomiarem w chwili t ,

$\mathbf{n}(t-1)$ i $\mathbf{v}(t)$ – są odpowiednio wektorami (2×1) określającymi statystycznie niezależne zakłócenia stanu układu w chwili $t-1$ i zakłócenia pomiaru w chwili t ; przyjmuje się również, że wektory zakłóceń są dane wielowymiarowymi rozkładami normalnymi: $\mathbf{n}(t) \sim N(0, \mathbf{Q}(t))$ i $\mathbf{v}(t) \sim N(0, \mathbf{R}(t))$, gdzie: $\mathbf{Q}(t)$ i $\mathbf{R}(t)$ są odpowiednimi macierzami kowariancji.

Rozwiązanie zadania filtracji Kalmana umożliwia wyznaczenie najlepszych estymat wg kryterium minimum błędu średniokwadratowego, tzw. estymat *a posteriori* wektora zmiennych stanu $\hat{\mathbf{s}}(t/t)$ dla kolejnych chwil t na podstawie serii zakłóconych pomiarów (np. z różnych źródeł) wykonanych do chwili t włącznie. Dodatkowo uzyskuje się również niepewność tych estymatorów danych, opisanych macierzą kowariancji błędu $\mathbf{P}(t/t)$. Wyznaczenie tych wielkości dla

kolejnych chwil czasu jest możliwe na podstawie układu równań rekursywnych, których wyprowadzenie można znaleźć w m.in. w pracach Andersona i Moore'a (1984) oraz Farella (2008).

W pracach badawczo-rozwojowych dotyczących opracowania systemu wspomaganie osób niewidomych w nawigacji terenowej stosowano metody integracji różnych technik nawigacji, tj. nawigacji wykorzystującej globalny system pozycjonowania i nawigację zliczeniową (Barański i Strumiłło, 2009a). Badano również skuteczność algorytmów filtracji danych nawigacyjnych, m.in. z wykorzystaniem filtracji Kalmana i tzw. filtracji cząsteczkowej:

- zastosowano filtr Kalmana do korekcji odczytów położenia z odbiornika GPS,
- zastosowano filtrację cząsteczkową do łączenia danych z odczytów GPS i czujników inercyjnych z uwzględnieniem oraz bez uwzględnienia danych *a priori* uzyskanych z map cyfrowych terenu.

W pierwszym etapie badań zastosowano filtr Kalmana w wersji bez sterowania, dla którego równanie przejść stanu zdefiniowano następująco:

$$\begin{bmatrix} x(t) \\ y(t) \\ V_x(t) \\ V_y(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(t-1) \\ y(t-1) \\ V_x(t-1) \\ V_y(t-1) \end{bmatrix} + \mathbf{n}(t) \quad (3.12)$$

gdzie zmienne stanu układu: $x(t)$, $y(t)$ oraz $V_x(t)$, $V_y(t)$ są wielkościami reprezentującymi odpowiednio długość i szerokość geograficzną oraz składowe prędkości pieszego przemieszczania się. Równanie pomiaru zdefiniowano w postaci:

$$\begin{bmatrix} GPS_x(t) \\ GPS_y(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \\ V_x(t) \\ V_y(t) \end{bmatrix} + \mathbf{w}(t) \quad (3.13)$$

gdzie: $GPS_x(t)$, $GPS_y(t)$ są odpowiednio odczytami GPS długości i szerokości geograficznej. W równaniach (3.12), (3.13) wektory $\mathbf{n}(t-1)$ i $\mathbf{w}(t)$ reprezentują statystycznie niezależne zakłócenia stanu układu i zakłócenia pomiaru o rozkładach normalnych (macierze kowariancji są diagonalne).



Rys. 3.20. Zastosowanie filtracji Kalmana do estymacji położenia geograficznego na podstawie odczytów GPS; kolory linii na obrazie mapy oznaczają: jasno zielony – ścieżka marszu, czerwony – odczyty współrzędnych GPS, niebieski – estymacja położenia z zastosowaniem filtra Kalmana

Źródło: opracowanie własne z wykorzystaniem mapy Łodzi dostępnej w systemie InterSit Miejskiego Ośrodka Dokumentacji Geodezyjnej i Kartograficznej w Łodzi, www.mapa.lodz.pl

Na rys. 3.20 pokazano fragment tzw. ortofotomapy Łódzkiego Internetowego Systemu Informacji o Terenie (InterSit) dla terenu kampusu Politechniki Łódzkiej. Cyframi (1÷4) oznaczono charakterystyczne obszary trasy nawigacyjnej, które skomentowano poniżej.

1. Filtracja Kalmana odczytów GPS (z uwagi na pamięć filtru Kalmana) spowodowała pogorszenie estymaty w obszarze nagłej zmiany kierunku marszu.
2. W tym obszarze odczyty GPS, pomimo stałej prędkości marszu, grupują się w kilku dyskretnych obszarach (odległych o ok. 10 m); ten niekorzystny efekt został skutecznie usunięty za pomocą filtracji Kalmana.
3. W obszarze wysokiej zabudowy odczyty GPS są obarczone dużym błędem systematycznym; filtr Kalmana nie jest skuteczny w eliminacji takich błędów z uwagi na średniokwadratowe kryterium estymacji.
4. Jednostkowe, duże odchyłki odczytów GPS są skutecznie korygowane przez filtr Kalmana.

Wnioskujemy, że filtr Kalmana zastosowany do korekcji położenia geograficznych, wyłącznie na podstawie odczytów GPS nie eliminuje szeregu błędów wyznaczania położenia geograficznego. Znacznej poprawy dokładności estymowanych położenia można się spodziewać po zastosowaniu techniki nawigacji zliczeniowej (np. inercyjnej).

W ostatnich latach wykazano bardzo dużą skuteczność metod symulacji Monte Carlo (SMC) w modelowaniu układów i procesów dla których niemożliwe jest zastosowania modelowania analitycznego (m.in. w ekonometrii i medycynie).

W szczególności pokazano, że sekwencyjna wersja SMC, nazywana filtracją cząsteczkową, jest skuteczną techniką obliczeniową rozwiązywania zadań estymacji stanu układów nieliniowych o złożonej dynamice i o niegaussowskim charakterze zakłóceń (Arulampalam i inni, 2002). Do takiej klasy zagadnień należą zadania estymacji położenia obiektów, w których wykorzystuje się hybrydowe techniki nawigacji m.in. łączące nawigację globalną z lokalną (np. systemu GPS z nawigacją zliczeniową) (Gustafsson i inni, 2002). W polskojęzycznej literaturze szerokie wprowadzenie do metody filtracji cząsteczkowej na gruncie podstaw teorii wnioskowania statystycznego można znaleźć w pracy Brzozowskiej-Rup i Dawidowicza (2009). Jedne z pierwszych badań nt zastosowania filtracji cząsteczkowej w nawigacji pieszej prowadzili Ceranka i Niedźwiecki (2003).

Podstawą koncepcji filtracji cząsteczkowej jest zastąpienie rekursywnych równań estymujących stan układu dynamicznego przez losowo generowany zbiór estymat (cząsteczek), który w iteracyjnym procesie wielokrotnych, selektywnych losowań jest ukierunkowywany na poprawę przybliżenia szukanego rozwiązania. Warto zaznaczyć, że dla liniowych modeli estymacji stanu i dla zakłóceń gaussowskich filtr cząsteczkowy jest zbieżny (dla odpowiednio dużej liczby cząsteczek) do rozwiązań definiowanych analitycznie przez filtr Kalmana (Cappe i inni, 2007).

W rozważanym zastosowaniu i -tą cząsteczkę spośród K dużej liczby cząsteczek losowanych w chwili t definiuje się jako dwójkę:

$$c_i(t) = \{X_i(t), w_i(t)\}, i = 1, 2, \dots, K \quad (3.14)$$

gdzie: $X_i(t)$ – jest wektorem estymowanych parametrów układu, np. zmiennych stanu układu przyporządkowanych i -tej cząsteczce, a $w_i(t)$ – jest nieujemną wagą przyporządkowaną i -tej cząsteczce w chwili t . Estymowany wektor parametrów jest obliczany w każdym kroku iteracji jako średnia ważona parametrów przyporządkowanych cząsteczkom, tj.:

$$\hat{X}(t) = \sum_{i=1}^K w_i(t) X_i(t) \quad (3.15)$$

Filtrację cząsteczkową zastosowano w hybrydowym układzie nawigacji osobistej, w którym użyto typowego odbiornika GPS (układ GPMMOPA6B) i moduł elektroniczny zawierający czujniki inercyjne (czujniki przyspieszeń i żyroskopy) (Barański i inni, 2011).

W rozważanym zagadnieniu cząsteczki zdefiniowano następująco:

$$c_i(t) = \{x_i(t), y_i(t), \varphi_i(t), w_i(t)\}, i = 1, 2, \dots, K \quad (3.16)$$

gdzie: $x_i(t)$, $y_i(t)$ – są współrzędnymi położenia geograficznego cząsteczki i w chwili t , $\varphi_i(t)$ – jest względnym położeniem kątowym związanym z cząsteczką (względem zadanego kierunku odniesienia) oraz $w_i(t)$ – jest wagą przyporządkowaną i -tej cząsteczce w chwili t .

Algorytm filtracji cząsteczkowej w dyskutowanym zastosowaniu obejmuje cztery kroki obliczeń:

Krok 1

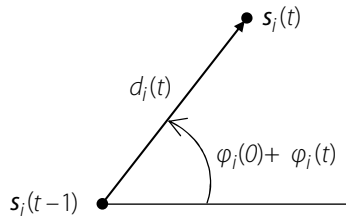
Nadaj losowe wartości początkowe $x_i(0)$, $y_i(0)$ oraz $\varphi_i(0)$ a wagom przyporządkuj jednakowe wartości: $\{w(0)_{i=1}^K\} = 1/K$.

Krok 2

Dokonaj predykcji stanu układu dla każdej cząsteczki, tj. przesun cząsteczki stosując równanie:

$$s_i(t) = \begin{bmatrix} x_i(t) \\ y_i(t) \end{bmatrix} = \begin{bmatrix} x_i(t-1) \\ y_i(t-1) \end{bmatrix} + (d_i(t) + n_{di}(t)) \begin{bmatrix} \cos(\varphi_i(0) + \varphi_i(t) + n_{\varphi_i}(t)) \\ \sin(\varphi_i(0) + \varphi_i(t) + n_{\varphi_i}(t)) \end{bmatrix} \quad (3.17)$$

Znaczenie wybranych składników równania wyjaśniono na rys. 3.21, a $n_{di}(t)$ i $n_{\varphi_i}(t)$ są wzajemnie niezależnymi zakłóceniami sterowania związanymi z wyznaczaniem długości kroku i kierunku. W równaniu (3.17), w odróżnieniu od równań filtracji Kalmana (3.11) związek pomiędzy wektorem sterowania $u_i(t) = [d_i(t), \varphi_i(t)]^T$ i zakłóceniami losowymi, a wektorem stanu jest nieliniowy.



Rys. 3.21. Związek między kolejnymi stanami układu $s(t-1)$ i $s(t)$ a sterowaniem $u_i(t) = [d_i(t), \varphi_i(t)]^T$ długość kroku $d(t)$ jest odczytywana na podstawie analizy sygnałów z czujnika przyspieszeń 3DoF, a kierunek wykonanego kroku $\varphi(t)$ jest kątem wyznaczonym za pomocą żyroskopu, φ_0 jest początkowym kątem odniesienia (z odczytów żyroskopu jest wyznaczana względna zmiana kąta)

Źródło: opracowanie własne

Krok 3:

Aktualizuj wagi cząsteczek na podstawie wykonanego pomiaru GPS, wg następującej procedury.

3a. Zastosuj równanie pomiaru postaci:

$$\begin{bmatrix} GPS_x(t) \\ GPS_y(t) \end{bmatrix} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} n_x(t) \\ n_y(t) \end{bmatrix} \quad (3.18)$$

gdzie: $GPS_x(t)$ i $GPS_y(t)$ są odpowiednio odczytami długości i szerokości geograficznej uzyskanymi z odbiornika GPS w chwili t oraz $n_x(t)$, $n_y(t)$ są zakłóceniami pomiaru współrzędnych, $n_x(t)$ oraz $n_y(t)$ są odpowiednio losowane z rozkładów normalnych $N(0, \sigma_{GPSx}(t))$ i $N(0, \sigma_{GPSy}(t))$, empirycznie przyjęto, że $\sigma_{GPSx} = \sigma_{GPSy} = \sigma_{GPS} = HDOP^2$ oraz $HDOP$ są parametrami rozmycia w kierunku poziomym podawanym wraz z kolejnymi odczytami współrzędnych GPS.

3b. Zastosuj następujący wzór do aktualizacji wag cząsteczek:

$$\tilde{w}_i(t) = w_i(t-1) \cdot p_{GPS}(\delta_i(t)) \quad (3.19)$$

gdzie: $\delta_i(t)$ – jest odległością euklidesową między współrzędnymi i -tej cząsteczki $[x_i(t), y_i(t)]$, a współrzędnymi $[GPS_x(t), GPS_y(t)]$ odczytanymi z systemu GPS oraz $p_{GPS}(\delta_i(t))$ – jest funkcją Gaussa $N(0, \sigma_{GPSx}(t))$.

3c. Normalizuj wagi cząstek stosując podstawienie:

$$w_i(t) = \frac{\tilde{w}_i(t)}{\sum_{i=1}^K \tilde{w}_i(t)} \quad (3.20)$$

w wyniku którego $\sum_{i=1}^K w_i(t) = 1$.

Krok 4:

Wyznacz współczynnik określający efektywną liczbę cząstek biorących udział w symulacji⁵:

$$K_{eff}(t) = \frac{K}{\sum_{i=1}^K w_i(t)^2} \quad (3.21)$$

⁵ W pracy Arulampalam i innych (2002) wykazano, że wielokrotnie powtarzane kroki filtracji cząsteczkowej prowadzą do asymptotycznego zerowania wartości wszystkich wag cząsteczek z wyjątkiem jednej. Zatem efektywna liczba cząsteczek decydujących o estymowanym stanie układu zmniejsza się z każdą iteracją. Dla przeciwdziałania temu zjawisku wykonuje się tzw. *resampling*, tj. ponowne selektywne losowanie cząsteczek biorących udział w dalszych symulacjach.

Jeżeli wartość współczynnika $K_{eff}(t)$ zmniejszy się poniżej pewnej arbitralnej wartości progowej (zalecaną wartością progu jest $K_{eff} = \frac{2}{3} K$) jest wykonywane tzw. przepróbkowanie (*resampling*). Nowy zbiór o liczności K cząsteczek jest generowany ze starego zbioru cząsteczek stosując losowanie z powtarzaniem. Prawdopodobieństwo wylosowania cząsteczki ze starego zbioru jest proporcjonalne do $w_i(t)$. Nowym cząsteczkom ponownie nadaje się jednakowe wagi, tj. $w_i(t) = 1/K$.

Krok 5:

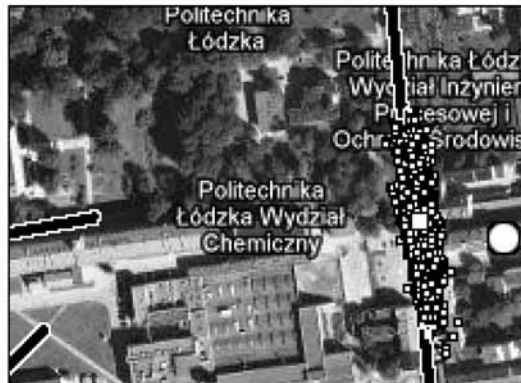
Wyznacz estymatę stanu układu $s(t) = [x(t) \ y(t)]^T$ jako średnią ważoną stanów przyporządkowanych wszystkim cząsteczkom:

$$x(t) = \sum_{i=1}^K w_i(t) x_i(t), \quad y(t) = \sum_{i=1}^K w_i(t) y_i(t) \quad (3.22)$$

Krok 6:

Wykonaj podstawienie $t = t + 1$ i idź do kroku 2 lub zakończ symulację.

Na rys. 3.22 zilustrowano przykład zastosowania filtracji cząsteczkowej do estymacji położenia osoby przemieszczającej się pieszo w terenie miasta. W symulacji zastosowano $K = 10000$ cząsteczek.



Rys. 3.22. Graficzna ilustracja filtracji cząsteczkowej do estymacji położenia w terenie miasta; małe kwadraty to cząsteczki reprezentujące hipotetyczny stan układu $(x_i(t), y_i(t))$; dużym kwadratem oznaczono prawdziwe położenie osoby przemieszczającej się pieszo; jasne koło wskazuje na lokalizację odczytaną z odbiornika GPS

Źródło: opracowanie własne

Na podstawie testów opracowanego algorytmu dla kilku ścieżek w obrębie kampusu Politechniki Łódzkiej uzyskano ilościową poprawę dokładności estymacji położenia osoby przemieszczającej się pieszo. Wyniki zebrano w tabeli 3.2.

Tabela 3.2. Średni błąd oraz odchylenie standardowe oszacowania położenia geograficznego: oszacowanie tylko na podstawie odczytów GPS (m_{GPS} , σ_{GPS}) oraz oszacowanie na podstawie łączenia danych z odczytów GPS i z odczytów z czujników inercyjnych z zastosowaniem filtracji cząsteczkowej (m_{fcz} , σ_{fcz}).

Błąd średni [m]		Odchylenie standardowe [m]	
m_{GPS}	m_{fcz}	σ_{GPS}	σ_{fcz}
9,8	6,1	15,5	8,8

Źródło: opracowanie własne

Podstawową zaletą zastosowania nawigacji zliczeniowej była eliminacja dużych błędów pozycjonowania. Pomimo statystycznej poprawy dokładności estymowanych położen geograficznych uzyskane błędy są nadal znaczące i mało przydatne dla bezpiecznej nawigacji osoby niewidomej. Koszt obliczeń związany z realizacją filtracji cząsteczkowej jest ściśle związany z liczbą symulowanych cząsteczek i z zasady jest wiele większy od metody filtracji Kalmana. W rozważnym zastosowaniu nawigacji dla osoby przemieszczającej się pieszo czas trwania obliczeń jednej iteracji wykonywanych na komputerze PC (zegar 2,5 GHz) dla $K = 10000$ cząsteczek wyniósł ok. 20 ms.

Zaletą algorytmów filtracji cząsteczkowej, z uwagi na empiryczny charakter obliczeń i wielokrotnie generowaną pulę estymat jest możliwość łatwego wprowadzenia ograniczeń na przestrzeń rozwiązań (Cappe i inni, 2007). Mogą one wynikać z ograniczeń fizycznych modelowanych obiektów lub procesów (np. maksymalna prędkość marszu), a także z ograniczeń wynikających z wiedzy *a priori* o modelowanym układzie np. obszary terenu niedostępne dla pieszego.

W testach algorytmów filtracji cząsteczkowej dla osobistej nawigacji pieszej zastosowano wiedzę *a priori* pozyskaną z dostępnych map terenu miasta (Barański i inni, 2011). Mapę terenu podzielono na trzy klasy obszarów:

- zabronione (mury, ściany, ogrodzenia, zbiorniki wodne, ulice wielopasmowe), którym nadano mnożnik $p = 0,1$,
- dostępne (trawniki, parkingi), którym nadano mnożnik $p = 0,5$,
- zalecane (chodniki, aleje, przejścia dla pieszych, przejścia podziemne), którym nadano mnożnik $p = 1$.

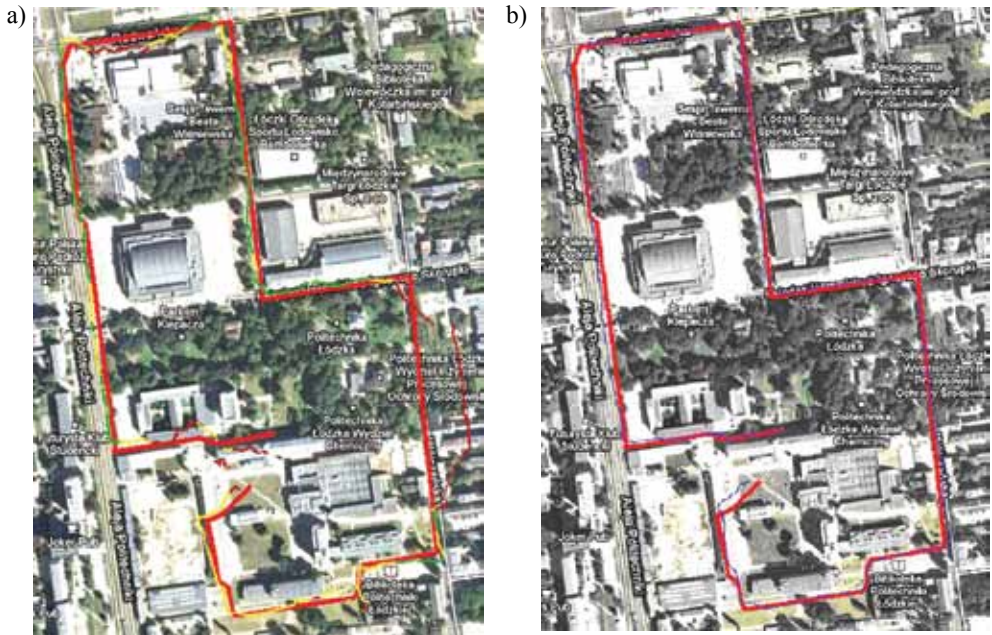
Na rys. 3.23 pokazano ortofotomapę terenu miasta i ten sam fragment terenu, któremu przydzielono 3 ww obszary. Wagi te zastosowano w algorytmie filtracji cząsteczkowej, tj. jeżeli położenie cząsteczki znajduje się obrębie danego obszaru to jej waga jest mnożona przez odpowiedni czynnik związany z tym obszarem wg objaśnionej reguły. Obszarom zabronionym nadano wagę różną od zera dla uniknięcia sytuacji, w której wszystkim cząsteczkom byłyby przydzielone zerowe wagi.



Rys. 3.23. Ortofotomapa fragmentu miasta Łodzi (a) oraz mapa rastrowa tego kwartału miasta, w której wydzielono 3 klasy obszarów: obszary czarne – zabronione, obszary szare – dostępne, obszary białe – zalecane

Źródło: a) Miejski Ośrodek Dokumentacji Geodezyjnej i Kartograficznej w Łodzi, b) opracowanie własne

Zastosowanie algorytmu cząsteczkowego do filtracji odczytów GPS, sygnałów z czujników inercyjnych i danych *a priori* o mapie terenu umożliwiło uzyskanie dalszej poprawy dokładności śledzenia ścieżki marszu w obrębie terenu miasta. Na rys. 3.24b wykreślono wytyczoną ścieżkę marszu (kolor czerwony) i linię uzyskaną z kolejno estymowanych położen wg algorytmu filtracji cząsteczkowej (linia niebieska). Dla porównania, dla tej samej ścieżki marszu wykreślono na rys. 3.24a położenia geograficzne uzyskane z odczytów GPS. Położenia te oznaczono kolorami odzwierciedlającymi wartość parametru HDOP: zielony – $HDOP < 1$ – dokładne odczyty (błędy pozycjonowania do 5 m); żółty – $HDOP < 2$ – dokładność średnia (błędy do 20 m); bordowy – $HDOP > 2$ – mała dokładność (błędy pozycjonowania większe niż 20 m).

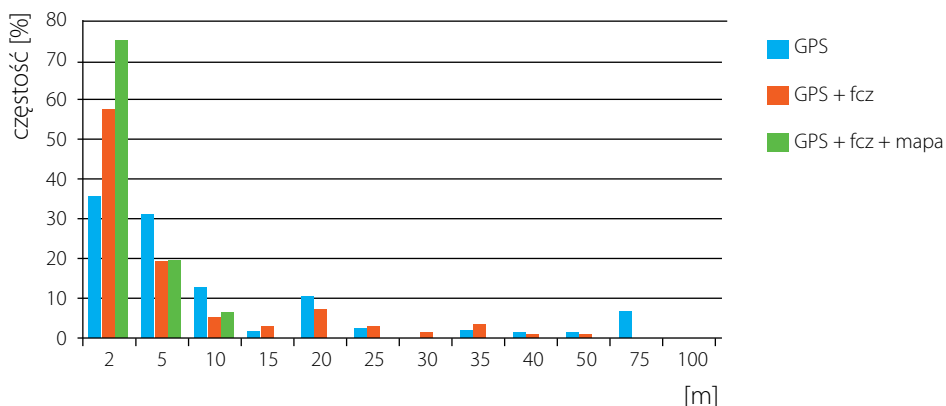


Rys. 3.24. Wytyczona ścieżka marszu (po lewej) wykreślona na tle mapy terenu (kolor czerwony) oraz linia marszu (po prawej) wyznaczona z wykorzystaniem filtracji cząsteczkowej uwzględniającej dane z odczytów GPS, czujników inercyjnych i dane z mapy terenu (linia niebieska)

Źródło: opracowanie własne z wykorzystaniem mapy Łodzi dostępnej w systemie InterSit Miejskiego Ośrodka Dokumentacji Geodezyjnej i Kartograficznej w Łodzi, www.mapa.lodz.pl

Ilościowe porównanie trzech testowanych metod estymacji położenia geograficznego pokazano na rys. 3.25 w postaci histogramów błędów. Na zamieszczonym wykresie zastosowano następujące oznaczenia dla poszczególnych histogramów:

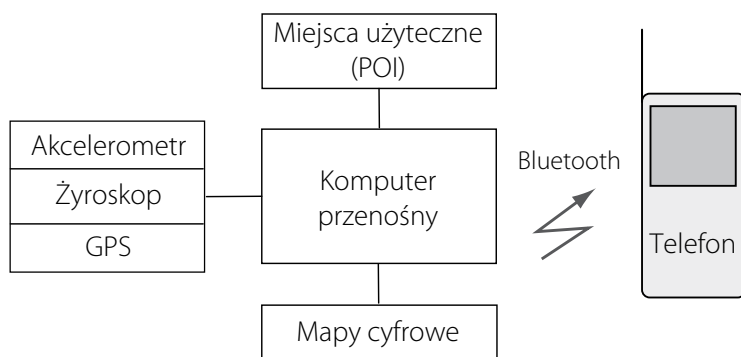
- GPS – położenie geograficzne wyznaczone tylko na podstawie odczytów GPS (nawigacja globalna), rys. 3.20,
- GPS+fcz – położenie geograficzne wyznaczone w hybrydowym systemie nawigacji na podstawie odczytów GPS i sygnałów z czujników inercyjnych z zastosowaniem filtracji cząsteczkowej (filtrowana nawigacja globalna i zliczeniowa),
- GPS+fcz+mapa – położenie geograficzne wyznaczone w hybrydowym systemie nawigacji na podstawie odczytów GPS, sygnałów z czujników inercyjnych, danych z map cyfrowych terenu (filtrowana nawigacja z ograniczeniami, globalna i zliczeniowa).



Rys. 3.25. Histogramy błędów dla estymowanych położenia geograficznych w terenie miasta (objaśnienie w tekście)

Źródło: opracowanie własne

Na podstawie wykreślonych histogramów zauważamy najlepszą dokładność wyznaczanych położenia geograficznych dla metody nawigacyjnej GPS+fcz+mapa. W metodzie tej 95% wyznaczanych położenia jest estymowanych z błędem nie większym niż 5 m. Nie jest to jednak dokładność, na podstawie której można zbudować automatyczny system do bezpiecznego prowadzenia osoby niewidomej w terenie miasta. Zasadniczą wartością uzyskanych wyników jest możliwość rozwijania prac nad systemami nawigacji osobistej tej klasy. W Instytucie Elektroniki Politechniki Łódzkiej zbudowano w tym celu testowy system nawigacji składający się z modułu czujników inercyjnych i odbiornika GPS



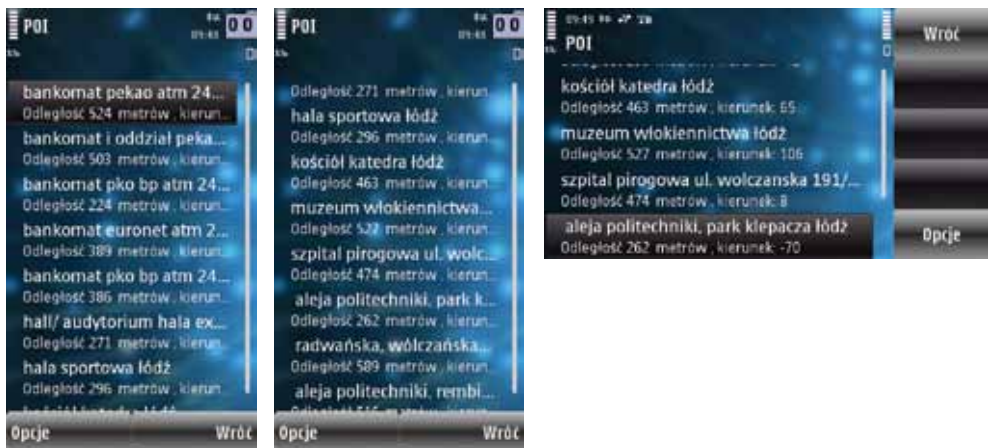
Rys. 3.26. Uproszczony schemat blokowy układu wyznaczania położenia geograficznego na podstawie połączonych informacji z odczytów odbiornika GPS, danych o przemieszczeniu oraz mapy numerycznej terenu

Źródło: opracowanie własne

(zob. rys. 3.11), komputera przenośnego i inteligentnego telefonu komórkowego. Schemat blokowy systemu pokazano na rys. 3.26.

Ważnym elementem systemu są informacje pozyskiwane z map cyfrowych w postaci tzw. miejsc użytecznych (*points of interest – POI*). Są to dane zapisane w mapie cyfrowej zawierające informacje o wybranych elementach struktury miasta i ich lokalizacji geograficznej, np. przystanki komunikacji miejskiej, przejścia dla pieszych, adresy ulic, bankomaty, budynki użyteczności publicznej. Na potrzeby rozwijanego systemu uruchomiono serwer www udostępniający informacje o miejscach użytecznych oraz utworzono bazę danych takich punktów. Baza zawiera zarówno własne dane o utworzonych miejscach POI oraz punkty pobrane ze strony projektu loadstone: <http://umpdlals.loadstone-gps.pl/tekstowe/>. Budowane są aplikacje umożliwiające aktualizację informacji przechowywanych w tej bazie danych.

Uzyskana dokładność nawigacji osobistej wynosząca w opisanym systemie ok. 5 m umożliwia skuteczne informowanie osoby niewidomej o punktach zainteresowania znajdujących się w jej otoczeniu oraz wskazywania ich kierunku. Dla uzyskania tej funkcji systemu opracowano aplikacje dla inteligentnego telefonu komórkowego, który pełni funkcje mobilnego terminala osoby niewidomej (Skulimowski i Sulmowski, 2011). Klawiszom numerycznym telefonu przyporządkowano różne kategorie miejsc użytecznych. Po wskazaniu określonego miejsca POI osoba niewidoma słyszy komunikat głosowy informujący o odległości i kierunku szukanego miejsca. Na rys. 3.27 pokazano przykładowe komu-



Rys. 3.27. Ilustracja działania modułu programowego telefonu komórkowego określania miejsc POI

Źródło: (Skulimowski i Sulmowski, 2011)

nikaty, które są odczytywane osobie niewidomej z wykorzystaniem syntezy mowy. Przeprowadzone testy w terenie wykazały 90% skuteczność wykrywania miejsc POI i ich wskazywania.

W poprzednim i niniejszym rozdziale omówiono odpowiednio techniki mikro- oraz makro-nawigacji, wykorzystywane we wspomaganiu osób niewidomych w samodzielnym poruszaniu się i orientacji przestrzennej. Następny rozdział jest poświęcony metodom prezentacji pozyskanych informacji nawigacyjnych z zastosowaniem technik niewizualnych.

Graficzny interfejs użytkownika (*graphical user interface – GUI*) jest podstawowym elementem systemów komputerowych i maszyn wykorzystywanych do interakcji z użytkownikiem (*human computer interaction – HCIs*) (Dix i inni, 2004). Przy tendencji budowania coraz bardziej złożonych systemów HCI (Strumiłło i inni, 2011b) i potrzeby przekazywania ich operatorom coraz większej ilości informacji, interfejsy graficzne są uzupełniane o odpowiednie interfejsy dźwiękowe. Istnieje też wiele zastosowań systemów HCI, w których użyteczność interfejsów graficznych jest niewielka (np. w sytuacjach alarmowych, które są sygnalizowane przede wszystkim dźwiękami) lub gdy ich użycie nie jest możliwe ze względu na niepełnosprawność wzrokową operatora systemu (Nesbitt, 2005). W takich sytuacjach są stosowane interfejsy prezentujące informacje dla innych, niewizualnych modalności zmysłowych. Wyróżnić można dwie podstawowe grupy takich systemów:

- interfejsy dźwiękowe (*sonic user interfaces, auditory user interfaces*),
- interfejsy haptyczne i dotykowe (*haptic and tactile user interfaces*).¹

Szczególnym zastosowaniem tej klasy interfejsów jest prezentacja otoczenia osobie niewidomej odpowiednio za pomocą dźwięków i technik haptycznych w systemach substytucji sensorycznej wzroku. Kluczowym problemem w skutecznym zastosowaniu tych technik jest wytworzenie odpowiednich bodźców

¹ Rozróżnienie tych dwóch pojęć nie jest ściśle przestrzegane w literaturze tematu, spostrzeganie dotykowe dotyczy czucia powierzchniowego, a spostrzeganie haptyczne wiąże się również z czuciem głębokim i obejmuje receptory umiejscowione w mięśniach, stawach ścięgnach i więzadłach (Lederman i Klatzky, 2009).

dla zmysłu słuchu lub dotyku, które umożliwią poprawną orientację niewidomego w otaczającej przestrzeni (Bach-y-Rita i Kercel, 2003a). Należy tu podkreślić odmienną naturę fizyczną bodźców i różne parametry „metrologiczne” tych zmysłów w porównaniu ze zmysłem wzroku.

4.1. Porównanie własności sensorycznych zmysłów wzroku, słuchu i dotyku

Słuch i dotyk są podstawowymi niewizualnymi zmysłami wykorzystywanymi w systemach substytucji sensorycznej zmysłu wzroku. W tym kontekście ważne jest porównanie podstawowych własności zmysłów wzroku, słuchu i dotyku w odniesieniu do funkcji związanych z niewizualną prezentacją otoczenia osobie niewidomej. W porównaniu tym uwzględniono zarówno podstawy biofizyczne działania zmysłów (Kalat, 2011) oraz prawa psychologii percepcji, wg których człowiek buduje obrazy wielozmysłowe otoczenia (Maruszewski, 2002), (Nęcka i inni, 2006).

W tabeli 4.1 zestawiono podstawowe własności sensoryczne zmysłów wzroku, słuchu i dotyku. W porównaniu tym użyto następujących pojęć :

- *pole receptorowe*, tj. wycinek przestrzeni otaczającej odbiorcę, w którym receptory poszczególnych zmysłów rejestrują bodźce; zmysł wzroku i słuchu jest wyposażony w tzw. telereceptory umożliwiające odbiór bodźców z dużych odległości i w szerokim zakresie kątowym, zaś receptory zmysłu dotyku odbierają bodźce wytworzone w bezpośrednim kontakcie z powierzchnią ciała,
- *przepływność informacyjna* definiowana jako ilość informacji przekazywanej w jednostce czasu (jednostka 1 bod = 1 bit/s); przepływność informacyjna jest podobna dla zmysłów wzroku i dotyku oraz blisko 1000 razy mniejsza dla zmysłu słuchu; należy zaznaczyć, że tylko część pojemności informacyjnej zmysłów niewizualnych może być wykorzystana w systemach substytucji sensorycznej,
- *pamięć sensoryczna* służąca do przechowania wrażenia zmysłowego po ustaniu bodźca w charakterystycznym dla danego zmysłu okresie czasu; odpowiednio dla wzroku, słuchu i dotyku rozróżnia się: *pamięć ikoniczną* – odgrywającą ważną rolę w krótkotrwałym przechowywaniu bodźców świetlnych; wynika z bezwładności wzroku wynoszącej ok. 150 ms (Coltheart, 1980); *pamięć echoiczną* – służącą do przechowywania dźwięków, które są „odtwa-

Tabela. 4.1. Porównanie właściwości sensorycznych zmysłów wzroku, słuchu i dotyku człowieka

Lp.	Parametry	Wzrok	Słuch	Dotyk
1	Natura fizyczna bodźca	Światło (część widma fali elektromagnetycznej)	Fala akustyczna	Siła nacisku
2	Pole receptorowe	Widzenie przestrzenne obuoczne w zakresie od kilku cm do nieskończoności; zakres kątowy dla ustalonego kierunku widzenia: ok. 150° (ruch ciała, głowy oraz gałek ocznych);	Słyszenie obouszne (ruch ciała, głowy): zakres kątowy: dookólny; rejestracja dźwięków z małych i bardzo dużych odległości	Receptory o różnej gęstości na powierzchni ciała (zob. rys. 4.18.)
3	Rodzaj receptorów	Fotoreceptory siatkówki oka (czopki i pręciki)	Komórki włosowate ślimaka	Różne rodzaje mechanoreceptorów
4	Pasmo	~10 ¹⁵ Hz (długość fali: 300-770nm)	16 Hz–20 kHz	~0 Hz (nacisk) – kilkaset Hz (wibracje)
5	Rozdzielczość	1 minuta kątowa (norma)	Kątowa: 1°–10° Częstotliwościowa: liczba rozróżnianych tonów: ~3000	5 mm÷50 mm; wynika z różnej gęstości receptorów dotyku na powierzchni ciała (rys. 4.18.)
6	Rozróżnialność wielkości bodźca	wg prawa Webera*	wg prawa Webera	wg prawa Webera

Tabela. 4.1. cd. Porównanie właściwości sensorycznych zmysłów wzroku, słuchu i dotyku człowieka

Lp.	Parametry	Wzrok	Słuch	Dotyk
7	Próg detekcji bodźca	Widzenie nocne (skotopowe), luminancja: 10^{-6} – 10^{-2} cd/m ² ; widzenie dzienne (fotopowe), luminancja: >1 cd/m ²	Poziom ciśnienia dźwięku (<i>sound pressure level</i> – SPL):	20 mN/mm ²
8	Zakres dynamiczny	10^{10} , adaptacja do zmiennych warunków oświetlenia	SPL = 120 dB (próg bólu)	20 mN/mm ² –409 mN/mm ²
9	Przepustowość informacji	10^7 bodów	$5 \cdot 10^4$ bodów	10^7 bodów
10	Pamięć sensoryczna	Pamięć ikoniczna	Pamięć echoiczna	Pamięć haptyczna

* Prawo psychometryczne sformułowane przez niemieckiego fizjologa Heinricha Webera (1795–1878) jako wynik obserwacji subiektywnej miary wielkości spostrzeganego bodźca odniesionej do fizycznej (obiektywnej) miary wielkości bodźca. Prawo to mówi, że: *Wielkość bodźca postrzeganego przez człowieka (tzw. wrażenie zmysłowe) jest proporcjonalne do logarytmu miary fizycznej bodźca:*

$$w = k \ln \frac{S}{S_0}$$

gdzie: k – stała proporcjonalności, S – wartość bodźca, S_0 – wartość bodźca odniesienia (tzw. ledwie wyczuwalnego bodźca). Zgodnie z tym prawem zwiększanie bodźca wg postępu geometrycznego powoduje zwiększanie wrażenia zmysłowego wg postępu arytmetycznego.

Źródło: opracowanie własne na podstawie (Kalat, 2011), (Longstaff, 2006), (Moore, 2004)

rzane” w korze słuchowej w okresie 3–4 s po ustaniu bodźca dźwiękowego; pojemność pamięci echoicznej jest znacznie mniejsza niż pamięci ikonicznej (Darwin i inni, 1972); oraz *pamięć haptyczną* – utrzymująca informację o wrażeniu dotykowym przez czas ok. 2 s po jego zaniku; (Shih i inni, 2009).

Rejestrowane w pamięci sensorycznej bodźce uaktywniają złożone procesy związane z budowaniem całościowych wrażeń zmysłowych, ich interpretacją semantyczną i zapamiętywaniem. Całość tych procesów określa się mianem *sposptrzegania* (Gerrig i Zimbardo, 2009). Na gruncie takich nauk jak fizjologia, neurologia i psychologia zaproponowano i eksperymentalnie potwierdzono szereg praw wyjaśniających wybrane mechanizmy *sposptrzegania* u człowieka (Nęcka i inni, 2008). Są to tzw. paradygmaty percepcji. Najważniejsze z nich zebrano i krótko scharakteryzowano w tabeli 4.2.

Duży wkład w wyjaśnienie procesów percepcji u człowieka wnieśli psychologowie postaci (Koffka, 1935). Sformułowali szereg zasad tzw. grupowania percepcyjnego, wg których człowiek wyodrębnia tzw. figurę (jej postać) spośród innych elementów składających się na tło. Pokazali przekonujące przykłady słuszności swej teorii dla percepcji wzrokowej (tabela 4.3). W wielu późniejszych pracach wykazano, że zasady grupowania percepcyjnego obowiązują dla innych modalności zmysłowych, m.in. dla percepcji słuchowej (Moore, 2004), (Bregman, 1990) oraz dotykowej (Chang i inni, 2007). Podstawowe zasady psychologii postaci w odniesieniu do zmysłów wzroku, słuchu i dotyku zebrano w tabeli 4.3. Wybrano te zmysły dla pokazania jak właściwości *sposptrzegania* wzrokowego przekładają się (wg postaciowej teorii percepcji) na właściwości zmysłów wykorzystywanych w systemach substytucji sensorycznej wzroku (Munkong i Juang, 2008), (Klawiter i Preis, 2006).

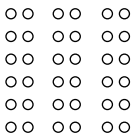
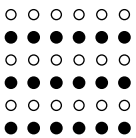
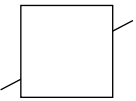
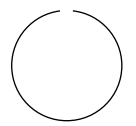
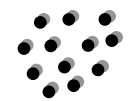
Tabela. 4.2. Zestawienie najważniejszych teorii spostrzegania (percepcji)

Teoria percepcji	Opis	Literatura
Asocjacyjna (atomistyczna)	Spostrzeżenie to suma odrębnych, wielomodalnych wrażeń zmysłowych łączonych w spójny obraz percepcyjny na podstawie praw kojarzenia częściowych wrażeń.	(Maruszewski, 2002)
Cech*	Obiekty są spostrzegane na podstawie składania elementów percepcyjnych jednej modalności zmysłowej w całościowy obraz zmysłowy otoczenia, np. łączenia tonów w melodie, składania odcinków linii w kontury kształtu.	(Selfridge i Neisser, 1960)
Schematów (prototypów)	Rozpoznawanie kategorii obiektów polega na porównaniu obrazu percepcyjnego z zapamiętanymi wcześniej wzorcami kategorii, tzw. schematami, a nie ich instancjami.	(Reed, 1972)
Percepcji bezpośredniej	Otoczenie jest spostrzegane bezpośrednio, tj. bez odwoływania się do schematów pamięciowych (jak w teorii wzorców); Gibson wskazuje na duże znaczenie tzw. aktywności poznawczej polegającej m.in. na świadomej aktywności ruchowej w spostrzeganiu, np. obrót głowy w kierunku źródła dźwięku i skupienie uwagi wzrokowej na obiekcie, tzw. mechanizm „gdzie”–„co” („where”–„what”), (Duchowski, 2003).	(Gibson, 1979)
Obliczeniowa*	Spostrzeganie wzrokowe wg Marra (1982) przebiega w 3 etapach: 1) budowania pierwotnych reprezentacji (<i>primal sketch</i>), np. obraz dwuwymiarowy (2D) na siatkówce oka, 2) wytworzeniu reprezentacji 2½ wymiarowej (2½D) uwzględniającej głębię sceny, 3) odwzorowania obiektów w postaci wyobrażeniowych modeli trójwymiarowych (3D) uwzględniających niewidoczne części obiektów.	(Marr, 1982)
Postaci (psychologia Gestalt)	Postuluje prymat całości spostrzeżenia (tzw. postaci) nad jej częściowymi składowymi. Całość jest czymś innym niż suma części. Z teorii tej wynikają ważne zasady tzw. grupowania percepcyjnego.	(Koffka, 1935)

*) teoria cech i teoria obliczeniowa są najbliższe komputerowym modelom wykorzystywanym w systemach rozpoznawania obrazów, m.in. neocognitron (wielowarstwowa sztuczna sieć neuronowa) zaproponowana przez Fukushima (2000) oraz algorytm SIFT (*Scale Invariant Feature Transform*) opracowany przez Davida Lowe (2001) i wykorzystywany m.in. w rozpoznawaniu biometrycznym (Ślot, 2010).

Źródło: opracowanie własne na podstawie (Maruszewski, 2002), (Nęcka i inni, 2006)

Tabela. 4.3. Prawa grupowania percepcyjnego dla zmysłów wzroku, słuchu i dotyku

Lp.	Zasady Gestalt	Wzrok	Słuch	Dotyk
1	Bliskości (łączenie w całość blisko położonych obiektów)		Styczność tonów w czasie lub częstotliwości	Styczność w przestrzeni (np. punktów w znakach alfabetu Braille'a)
2	Podobieństwa (łączenie w całość obiektów wzajemnie podobnych)		Podobieństwo barwy lub głośności	Podobieństwo siły nacisku, tekstury powierzchni, wzoru wibracji
3	Ciągłości lub dobrej kontynuacji (łączenia w całość, „uciągłania” obiektów)		Zlewanie się dźwięków w tło	Łączenie ciągu wzorców dotykowych w całość np. znaków alfabetu Braille'a w wyrazy i wiersze
4	Domykania (tendencja do uzupełniania i domykania kształtów)		Śledzenie zmieniających się parametrów dźwięku pomimo okresowego ich maskowania (np. szumem)	Dotykowe rozpoznawanie fragmentów kształtów w całość
5	Wspólnej drogi (grupowanie obiektów podlegających wspólnej zmianie, np. położenia, tonu)		Postrzeganie wspólnych zmian kilku dźwięków i ich wydzielanie od dźwięków tła	Łączenie w jedno wrażenie dotykowe wspólnie przemieszczanych bodźców dotykowych

Źródło: opracowanie własne

Teoria strumieni dźwiękowych wg Bregmana

Albert Bregman (1990) rozwinął prawa grupowania percepcyjnego dla spostrzegania słuchowego. Na podstawie wieloletnich badań psychoakustycznych dowiódł, że mechanizm interpretacji zdarzeń dźwiękowych przebiega u człowieka wg następujących zasad:

- spostrzegane dźwięki podlegają organizacji percepcyjnej, w wyniku której jest budowana tzw. scena dźwiękowa (*auditory scene*) i dopiero ona podlega rozpoznaniu,
- na scenę dźwiękową składa się pewna liczba tzw. strumieni dźwiękowych (*auditory streams*),
- zasady integracji dźwięków w strumień dźwiękowy (*stream fusion*) lub ich rozdzielania na oddzielne strumienie (*stream segregation*) przebiega wg zasad grupowania percepcyjnego,
- w grupowaniu percepcyjnym biorą udział dwa mechanizmy tzw. analiza prymitywna, w której grupowanie jest warunkowane cechami psychoakustycznymi dźwięku (m.in. wysokość tonu, barwy, harmoniczności) lub analiza schematyczna przebiegająca na poziomie kognitywnym i polega na dopasowaniu porównawczym dźwięków do wcześniej zapamiętanego schematu dźwięków,
- liczba jednocześnie rozróżnianych strumieni dźwiękowych jest ograniczona (typowo 4–5 strumieni).

Ważnym spostrzeżeniem badań Bregmana jest stwierdzenie, że im bardziej jest utrwalony dany schemat dźwiękowy (strumień) tym jego rozpoznanie przebiega szybciej i przy małych zasobach uwagi (Wickens, 2002). Wniosek ten jest ważny dla diskutowanego zastosowania systemu prezentacji dźwiękowej otoczenia dla osób niewidomych.

Dla dźwiękowego obrazowania otoczenia jest również ważne określenie parametrów dźwięków, przy których zachodzi rozdzielanie strumieni, są to parametry definiowane:

- w czasie: stwierdzono, że 20 ms to minimalny odstęp czasu pomiędzy dwoma dźwiękami potrzebny do określenia, który z dźwięków wystąpił pierwszy,
- w przestrzeni: im większa różnica kierunków, z których dochodzą dźwięki tym lepiej są one rozróżnialne,
- w dziedzinie widma: częstotliwość dwóch jednocześnie występujących dźwięków powinna się różnić przynajmniej o 20%,
- w dziedzinie cech dźwięku: człowiek bardzo dobrze rozróżnia dźwięki na podstawie ich barwy.

Wcześniej opisane właściwości sensoryczne i psychoakustyczne zmysłu słuchu uzupełniono o właściwości, które powinny być uwzględnione przez projektanta systemu dźwiękowej prezentacji otoczenia dla osoby niewidomej (pełniejszy opis użytych pojęć można znaleźć w pracach Moora (2004) i Bujacza (2010)):

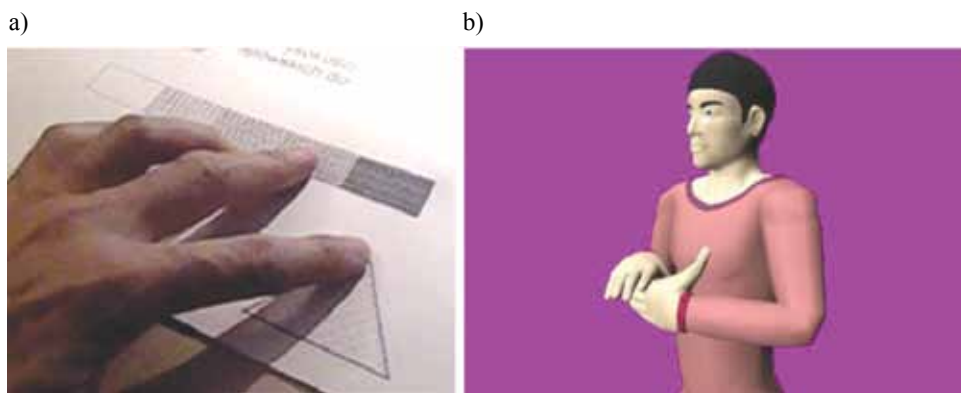
- dźwięki z pasma 2÷5 kHz są najlepiej słyszalne i spostrzegane jako głośniejsze w porównaniu do innych dźwięków o tym samym natężeniu (zob. krzywe jednakowej głośności),
- dźwięki z pasma 5÷7 kHz są niezbędne do dokładnej lokalizacji źródła dźwięku,
- tony muzyczne i dźwięki harmoniczne ze środkowych trzech oktaw (131–988 Hz) są subiektywnie najprzyjemniejsze dla ludzkiego ucha,
- człowiek lepiej rozpoznaje względne zmiany wysokości i głośności dźwięku niż ich bezwzględne wartości,
- spostrzegana wysokość tonu zależy od głośności dźwięku,
- dźwięki występujące blisko siebie w czasie lub mające podobne częstotliwości mogą podlegać maskowaniu.

W podanej skrótowo charakterystyce właściwości zmysłów należy uwzględnić ważny mechanizm świadomej selekcji bodźców oddziałujących na zmysły, który określa się mianem selektywności uwagi, dzięki której tylko część bodźców podlega analizie percepcyjnej (Gerrig i Zimbardo, 2009).

Selektywność uwagi można interpretować jako skupienie aktywności poznawczej tylko na części bodźców docierających z otoczenia, np. możemy skupiać uwagę na wypowiedzi wybranego mówcy (strumieniu dźwięków jego mowy) spośród wielu głośno rozmawiających osób (*cocktail-party problem*) lub czytać konkretny wiersz tekstu na zadrukowanej stronie. Uwzględnienie procesu uwagi w systemach substytucji sensorycznej ma bardzo istotne znaczenie. Z wielu prac nad mechanizmami uwagi, które badano dla poszczególnych zmysłów wynika, że człowiek może jednocześnie i świadomie skupiać uwagę na niewielkiej liczbie zdarzeń lub obiektów otoczenia (Bregman, 1990), (Nęcka i inni, 2006). Zdolności te można trenować tak, aby czynności pierwotnie wymagające skupienia uwagi mogły być wykonywane w sposób automatyczny (Maruszewski, 2002), np. osoby sprawnie piszące na klawiaturze komputera nie muszą koncentrować uwagi wzrokowej na palcach, a wolne zasoby uwagi mogą przeznaczyć na śledzenie wyświetlanego tekstu.

4.2. Systemy substytucji sensorycznej

Substytucję sensoryczną (*sensory substitution*) definiuje się jako zastąpienie określonej reprezentacji (modalności) zmysłowej reprezentacją dostosowaną do odbioru przez receptory innych zmysłów (Bach-y-R, 1972), (Bach-y-Rita i Ker-cel, 2003a). Systemy tej klasy są przede wszystkim stosowane jako pomoce dla osób z niepełnosprawnością sensoryczną. Systemy sensoryczne zmysłu wzroku to m.in.: druk wypukły, np. alfabet Braille’a, programy do syntezy mowy, tzw. technologie *text-to-speech* (zob. www.ivona.com), techniki narracyjne nazywane audiodeskrypcją (Szymańska i Strzyński, 2011), a dla głuchoniemych aplikacje do translacji tekstu na sekwencje obrazów języka migowego (*speech to sign language*) (rys. 4.1).



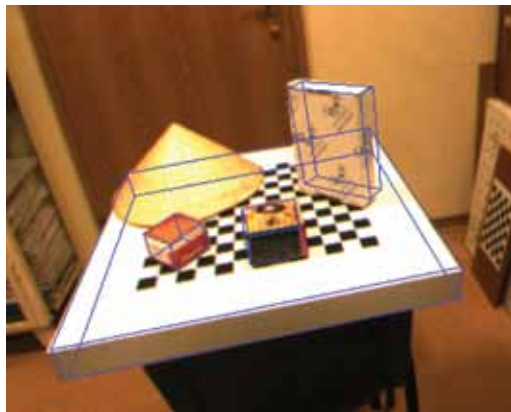
Rys. 4.1. Przykłady systemów substytucji sensorycznej:

(a) druk wypukły (technika drukarska zwana tyflografią) umożliwia niewidomym rozpoznawać znaki alfanumeryczne, obrazy, kształty i mapy (b) widok okna programu Thetos z wirtualnym animowanym tłumaczem (awatarem) tekstu na język migowy

Źródło: a) opracowanie własne, b) (Szmaj i Suszczyńska, 2002)

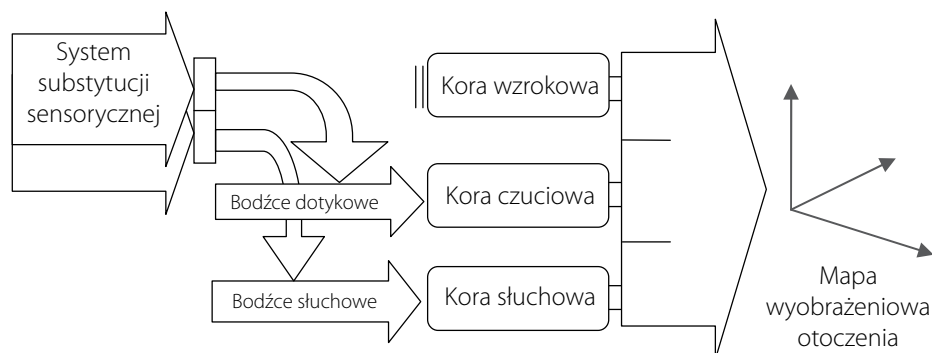
Pełne zastąpienie funkcji utraconego zmysłu nie jest możliwe. Dlatego też, określenie „substytucja sensoryczna” wg podanej definicji jest przez wielu badaczy krytykowane (Lenay i inni, 2003). Postulują oni aby stosować określenia takie jak „kompensacja sensoryczna” lub „suplementacja sensoryczna”, jako technik kompensujących lub uzupełniających funkcje utraconego zmysłu, a nie w pełni je zastępujących. Suplementacja sensoryczna może również dotyczyć, uzupełnienia (wzbogacania) postrzeganych wrażeń sensorycznych dla tego samego zmysłu, np. aparat słuchowy i okulary korekcyjne, a także noktowizor i kamery na podczerwień można zaliczyć do tej klasy urządzeń.

Idea substytucji sensorycznej jest w coraz większym stopniu wykorzystywana w tzw. systemach *rzeczywistości rozszerzonej* (*augmented reality – AR*), w których postrzegane otoczenie jest wzbogacane o nowe, generowane komputerowo obiekty (Bimber i inni, 2005). Obiekty te mogą rozszerzać (wzbogacać) daną modalność, np. rejestrowane obrazy otoczenia mogą być łączone z obiektami generowanymi komputerowo (rys. 4.2) lub przekształcać obiekty przeznaczone dla danej modalności zmysłowej do innej modalności zmysłowej. Do klasy systemów *AR* można zatem zaliczyć tzw. „wyświetlacze dźwiękowe” (*auditory displays*) i „wyświetlacze dotykowe” (*tactile displays*), w których obrazy otoczenia są przekształcane odpowiednio do wzorców dźwiękowych lub dotykowych. Systemy tej klasy znajdują zastosowanie w urządzeniach suplementacji sensorycznej dla niewidomych oraz w systemach zwiększających możliwości percepcyjne operatorów złożonych systemów sterowania. Tematyka ta jest przedmiotem regularnie organizowanych konferencji naukowych (m.in. *International Conference on Auditory Display* (ICAD, 2011)), publikacji naukowych (Kaczmarek, 1999) i opracowań monograficznych (Zadeh, 2010).



Rys. 4.2. Łączenie obrazu sceny i jej trójwymiarowego modelu
Źródło: (Pełczyński i inni, 2009)

Zadaniem systemu substytucji sensorycznej zmysłu wzroku jest wytworzenie specjalnych bodźców dla niewizualnych modalności zmysłowych, które wzbogacą zdolność spostrzegania otoczenia przez osobę niewidomą. Jest to zatem system klasy *rzeczywistości wzbogaconej*, w którym (jak zaznaczono na rys. 4.3) do spostrzeganych naturalnie bodźców są dodawane nowe bodźce, np. dźwięki lub bodźce dotykowe, w celu uzyskania poprawy orientacji przestrzennej osoby niewidomej.



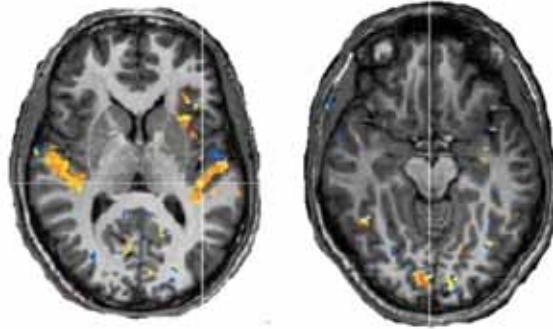
Rys. 4.3. Idea systemu substytucji sensorycznej wzbogacającego spostrzeganie otoczenia przez osobę niewidomą poprzez wytwarzanie dodatkowych bodźców dla zmysłów niewizualnych

Źródło: opracowanie własne

Zastąpienie spójnych wrażeń wzrokowych (tzw. perceptów) bodźcami dla innego zmysłu jest zadaniem niezmiernie trudnym z uwagi na odmienną naturę fizyczną (lub chemiczną) bodźców postrzeganych przez zmysły człowieka. Trudność ta wynika także z różnej pojemności informacyjnej receptorów poszczególnych zmysłów, ich odmiennego anatomicznego rozmieszczenia oraz różnego sposobu (i szybkości) przetwarzania, scalania i wytwarzania spójnej reprezentacji zmysłowej.

Stosowanie systemów substytucji zmysłu wzroku, w których jest przetwarzana duża ilość informacji o przestrzennej strukturze otoczenia, jest szczególnie użyteczna jeżeli osoba niewidoma ma zachowane funkcje ośrodka mózgu (kora wzrokowa) odpowiedzialnego za widzenie (Longstaff, 2006). Uwarunkowanie to jest w większości przypadków spełnione, gdyż utrata wzroku w przeważającej liczbie przypadków jest skutkiem uszkodzenia peryferyjnych odcinków układu wzrokowego (np. rogówki, siatkówki lub nerwu wzrokowego). W wielu badaniach psychofizycznych potwierdzono, że osoby niewidome (w szczególności ociemniałe) korzystające z systemów substytucji sensorycznej informowały o wrażeniach zmysłowych, które określały jako odwzorowanie otoczenia „widziane” dźwiękiem lub dotykiem (Arno i inni, 2001). Na podstawie badań eksperymentalnych przeprowadzonych z udziałem osób niewidomych potwierdzono, że bodźce kierowane dla zmysłu dotyku lub słuchu związane z orientacją przestrzenną aktywują korę wzrokową (Weeks i inni, 2000), (Amedi i inni, 2010). Świadczą o tym wyniki badań, w których analizowano aktywność obszarów kory wzrokowej na podstawie obrazów cyfrowych mózgu rejestrowanych technikami funkcjonalnego rezonansu magnetycznego (fMRI – *functional magnetic resonance imaging*) oraz tomografii z emisją pozytronów (PET – *positron emission*

tomography) (Steven i Blakemore, 2004), (Voss i inni, 2006). Na rys. 4.4 pokazano wyniki prac opisanych w (Merabet i inni, 2008) pokazujących aktywowanie kory słuchowej ale również kory wzrokowej u osób niewidomych, które uczestniczyły w sesjach treningowych polegających na korzystaniu z dźwiękowego systemu substytucji wzroku.

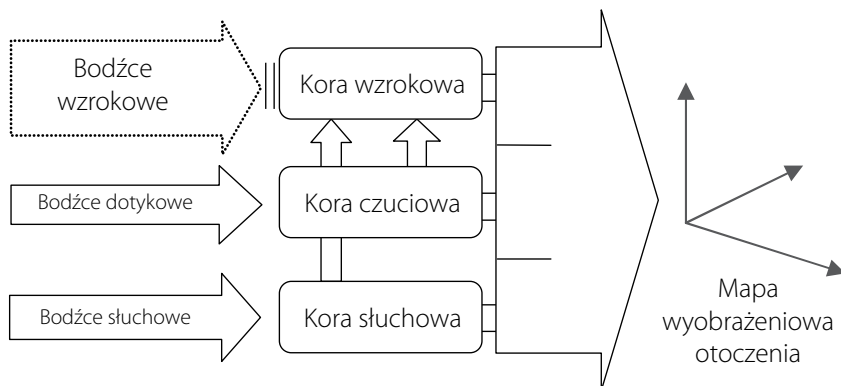


Rys. 4.4. Obrazy fMRI mózgu osoby niewidomej uzyskane po serii sesji treningowych z systemem dźwiękowej substytucji wzroku; z lewej – obraz aktywności ośrodków kory słuchowej (obszary oznaczone kolorami) oraz obraz pokazujący aktywność ośrodków kory wzrokowej u niewidomego uczestnika testów

Źródło: (Marabet i inni, 2008) za zgodą autorów pracy

Uzyskane wyniki świadczą o wytworzeniu nowych połączeń synaptycznych pomiędzy ośrodkami wzrokowymi mózgu (które na skutek uszkodzenia drogi wzrokowej nie są stymulowane) a ośrodkami zmysłów słuchu i dotyku (rys. 4.5). Proces ten jest w neurobiologii określany mianem plastyczności (tzw. neuroplastyczności) mózgu. Pojęcie to zostało zdefiniowane i opisane po raz pierwszy przez polskiego neurofizjologa Jerzego Konorskiego (Konorski, 1948). Plastyczność mózgu definiuje się jako zdolność komórek neuronowych do reorganizacji synaptycznej, tj. modyfikacji liczby, struktury oraz siły połączeń między komórkami neuronowymi (Żernicki, 1983). W szczególności mianem kompensacyjnej plastyczności mózgu (Kinałski, 2008) określa się mechanizm wytwarzania nowych połączeń, dzięki którym aktywowane są komórki neuronowe ośrodka mózgu pozbawione „swoich” naturalnych bodźców.

W badaniach z udziałem osób ze sprawnym wzrokiem, którym na czas badań (ok. 5 dni) zablokowano spostrzeganie bodźców wzrokowych zauważono, że lepiej rozróżniały one znaki pisma Braille’a, niż osoby, u których nie blokowano bodźców wzrokowych. Stwierdzono również, że po ok. 20 godz. po usunięciu blokady bodźców wzrokowych zwiększona ostrość czucia dotykiem zanikała (Hamilton i Pascual-Leone, 1998).

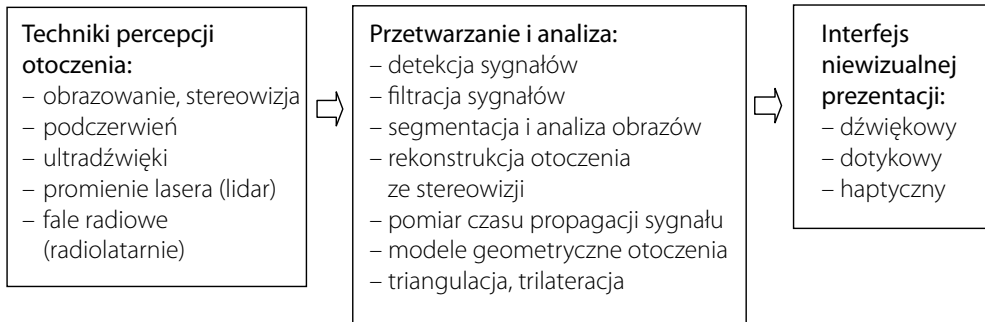


Rys. 4.5. Ilustracja mechanizmu plastyczności mózgu: w wyniku uszkodzenia drogi wzrokowej bodźce wizualne nie docierają do kory wzrokowej ale dzięki mechanizmowi plastyczności mózgu kora wzrokowa jest stymulowana z ośrodków kory czuciowej i słuchowej i może uczestniczyć w procesie budowania „obrazów” otoczenia

Źródło: opracowanie własne

U osób niewidomych stwierdzono występowanie mechanizmu plastyczności kompensacyjnej pomiędzy korą wzrokową a korą czuciową i słuchową (Liu i inni, 2007). Co więcej, zauważono że regularne aktywowanie kory somatosensorycznej (np. przez naukę alfabetu Braille’a) i kory słuchowej (słuchanie dźwięków przestrzennych) wzmacnia mechanizm plastyczności kompensacyjnej (Ptito i inni, 2009), (Collington i inni, 2009). Można zatem wnioskować, że osoby niewidome poddane takiemu treningowi będą mogły z większą skutecznością posługiwać się niewizualnymi zmysłami w percepcji otoczenia. Spostrzeżenie to jest szczególnie ważne w odniesieniu do dyskutowanych technik substytucji sensorycznej.

Biała laska niewidomego jest podstawową techniką substytucji sensorycznej wzroku, w której utracone funkcje zmysłu wzroku są zastępowane bodźcami haptycznymi pochodzącymi od drgań przenoszonych przez laskę. Współczesne technologie czujników elektronicznych (kamery cyfrowe, czujniki ultradźwiękowe, lasery) oraz techniki obliczeniowe (np. techniki analizy i syntezy obrazów) oferują o wiele doskonalsze sposoby rozpoznawania otoczenia. Dostępne są również bardziej złożone techniki niewizualnej prezentacji informacji o otoczeniu, m.in. interfejsy haptyczne (Zadeh, 2010) oraz techniki dźwięku przestrzennego (Moore, 2004). Na rys 4.6 pokazano ogólny model zawierający podstawowe bloki systemu substytucji sensorycznej zmysłu wzroku. Zadaniem takiego systemu jest pozyskiwanie odpowiednich danych o strukturze otoczenia, odpowiednie ich przetworzenie i wytwarzanie bodźców dla zmysłów niewizualnych.



Rys. 4.6. Uproszczony schemat blokowy systemu substytucji sensorycznej zmysłu wzroku
Źródło: opracowanie własne

Na pokazanym schemacie można wyróżnić trzy moduły.

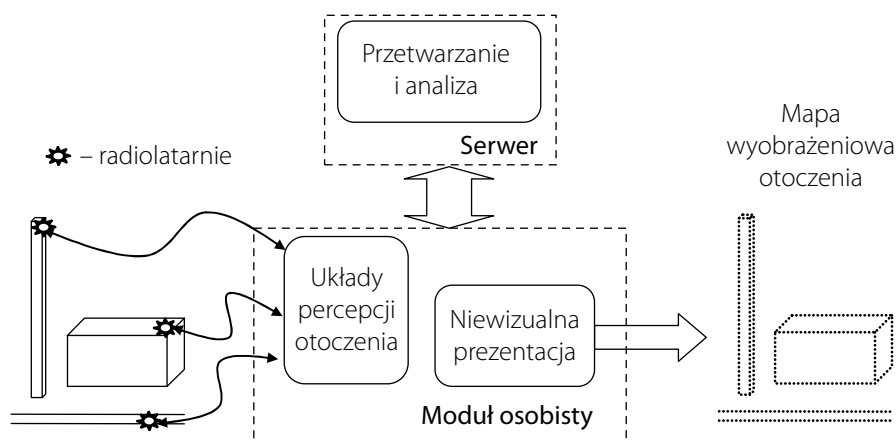
1. Moduł percepcji otoczenia, którego zadaniem jest zbieranie danych o strukturze otoczenia; są tu wykorzystywane zarówno techniki pasywne (wizyjne) jak i aktywne (działające na zasadzie emisji wiązek energii do otoczenia i rejestracji energii odbitej od struktur otoczenia), np. laser, podczerwień, ultradźwięki.
2. Moduł przetwarzania i analizy, którego zadaniem jest przetwarzanie i analiza sygnałów z układów percepcji otoczenia.
3. Moduł niewizualnej prezentacji otoczenia, którego zadaniem jest wytwarzanie bodźców niewizualnych.

Podstawowe konstrukcje systemów substytucji sensorycznej to urządzenia osobiste, w których wszystkie moduły systemu, tj. układy percepcji, przetwarzania, analizy i prezentacji są zintegrowane w jednym urządzeniu przenośnym.

W ostatnich latach podlegają szybkiemu rozwojowi systemy o rozproszonej strukturze (*cross reality environments*), w tym systemy tzw. inteligentnego otoczenia (*ambient intelligence*). W systemach tych wbudowane w otoczenie układy czujników i bezprzewodowo sterowanych manipulatorów otwierają nowe możliwości interakcji człowieka z otoczeniem (Paradiso i Landlay, 2009), (Mikołajewska i Mikołajewski, 2011).

Systemy substytucji sensorycznej można również skutecznie realizować wg tzw. *architektury rozproszonej*. W systemach takich wybrane bloki systemu są umiejscowione poza modułem osobistym osoby niewidomej. Takie zadania obliczeniowe systemu jak analiza obrazu, synteza mowy, analiza mapy cyfrowej terenu mogą być realizowane w odległym serwerze połączonym siecią bezprzewodową z modułem osobistym (rys. 4.7). Układy percepcji otoczenia o charakterze rozproszonym to sieci mikroukładów elektronicznych wykonanych m.in. w technologii RFID (Na, 2006) lub nadajników światła podczer-

wonego (SKERI, 2011). Układy te mogą być wbudowywane w infrastrukturę otoczenia np. wejścia do budynków, przystanki komunikacji miejskiej i stanowić system tzw. radiolatarni aktywowanych modulem osobistym i wspomagających niewidomego w orientacji przestrzennej (Kurek i Czerwiński, 2003). Na rys. 4.7 pokazano schemat przykładowego systemu substytucji sensorycznej dla osoby niewidomej (o architekturze rozproszonej).



Rys. 4.7. System substytucji sensorycznej o architekturze rozproszonej; węzły sieci rozproszonych czujników łączą się bezprzewodowo z modulem osobistym, który komunikuje się z odległym serwerem i wytwarza niewizualne komunikaty dla osoby niewidomej

Źródło: opracowanie własne

4.2.1. Dźwiękowe techniki obrazowania

Zmysł słuchu jest jedynym zmysłem człowieka, który rejestruje bodźce pochodzące z dowolnego kierunku otoczenia i z dowolnej odległości. Człowiek reaguje na dźwięk kierując wzrok w stronę źródła tego dźwięku. Jest to znany w psychologii poznawczej mechanizm „gdzie-co” (*where-what*), w którym uwaga (gdzie?) przechwycona przez bodziec dźwiękowy jest następnie ukierunkowana na obiekt w celu jego dokładniejszego rozpoznania wzrokowego (co?) (Duchowski, 2003). Osoba niewidoma po wykryciu zdarzenia dźwiękowego jest pozbawiona możliwości takiego dokładniejszego rozpoznania, np. wzrokowego potwierdzenia lokalizacji przestrzennej źródła dźwięku. Zakres tego rozpoznania jest ograniczony jedynie do zasięgu dotykowej eksploracji otoczenia, która jest przydatna w rozpoznawaniu kształtów i faktury obiektów ale nie jest przydatna w skutecznym poruszaniu się.

Funkcje sensoryczne zmysłu słuchu u osób niewidomych nie różnią się zasadniczo od osób ze sprawnym wzrokiem. W licznych badaniach psychoakustycznych stwierdzono, że osoby które utraciły wzrok we wczesnym okresie życia lepiej potrafią rozróżniać strumienie dźwiękowe zakłócone innymi dźwiękami, np. mowę zakłócaną szumem (Hollins, 1989), (Kwapiszowie, 1990). Intrigującą umiejętnością stwierdzoną u wielu osób niewidomych jest zdolność wykrywania przeszkód, którą nazwano naturalną echolokacją. Pierwotnie, błędnie próbowano ją tłumaczyć zdolnością osób niewidomych do spostrzegania subtelного przepływu powietrza wokół twarzy. Dopiero w systematycznych badaniach Davey (1968) wykazano, że niektóre osoby niewidome potrafią wykrywać niewielkie zmiany dźwięku własnych kroków przy zbliżaniu się do przeszkody (przy zatkanych obu kanałach słuchowych zdolność tę traciły). Wyostrenie zdolności sensorycznych słuchu u osób niewidomych nie dotyczy lokalizacji źródeł dźwięku. Stwierdzono, że osoby niewidome z nieznacznie gorszą dokładnością niż osoby ze sprawnym wzrokiem i słuchem potrafią wskazywać kierunek położenia źródła dźwięku (Bujacz i inni, 2011). Przyczyn tego faktu upatruje się w braku sprzężenia między zmysłami wzroku i słuchu, tj. zdolności, która wykształca się w trakcie rozwoju osobniczego i polega na wzrokowym potwierdzeniu i uszczegółowieniu lokalizacji źródła dźwięku (Zwiers i inni, 2001) (Zahorik, 2002).

W wielu badaniach pokazano jednak, że interfejsy dźwiękowe mogą być skutecznie wykorzystywane przez osoby niewidome, m.in. w dźwiękowej prezentacji danych graficznych (Talbot i Cowan, 2006) (Wersenyi, 2010), korzystaniu z dźwiękowych środowisk w nauczaniu (Cooper i Petrie, 2004) oraz w systemach nawigacji dla osób niewidomych (Ohuchi i inni, 2006).

4.2.1.1. Techniki sonifikacji

Stosowane są dwie podstawowe techniki dźwiękowego przekazania informacji osobie niewidomej o otoczeniu:

- opis werbalny (zwany też audiodeskrypcją) – czyli technika opisu słownego informacji wizualnej (również z wykorzystaniem syntezatorów mowy),
- sonifikacja (*sonification*) – w której do prezentacji informacji o otoczeniu stosuje się niewerbalne, syntezywane sygnały dźwiękowe.

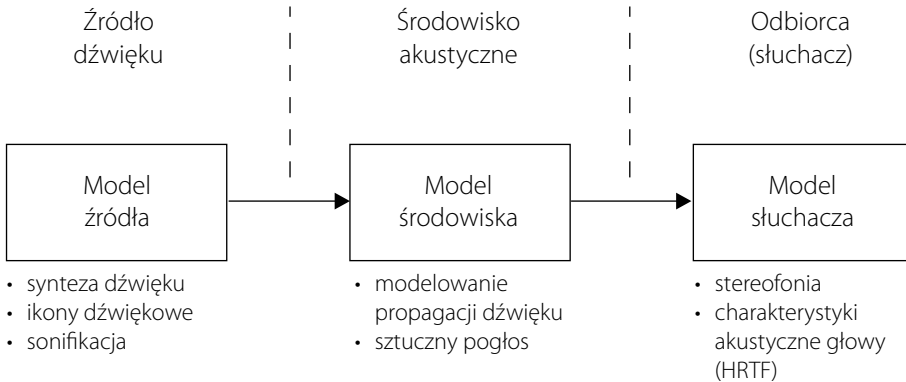
Audiodeskrypcja jest wykorzystywana przez przewodników osób niewidomych oraz jest stosowana do opisu słownego treści wizualnych, m.in. w edukacji, opisie dzieł sztuki (filmów, obrazów, rzeźb), komentowaniu słownym stron internetowych. Opracowano odpowiednie zalecenia odnośnie zasad tworzenia opisów werbalnych dla osób niewidomych (ITC, 1996).

Sonifikacja jest techniką szeroko stosowaną w systemach interakcji człowiek-komputer (zob. obszerny raport nt. tych technik (Kramer i inni, 1994)) oraz w tzw. wirtualnych wyświetlaczach dźwiękowych (*virtual auditory displays* – VAD) jako elementu systemów rzeczywistości wirtualnej (Lokki i inni, 2003). Stosowanymi źródłami dźwięku w takich systemach są ciągłe lub rozdzielone w czasie sztucznie generowane dźwięki o różnym stopniu złożoności, które służą do budowania niewerbalnych komunikatów o ustalonym znaczeniu (Kramer, 1994). Dźwięki te wg terminologii zastosowanej w pracy (McGookin i Brewster, 2004) określa się mianem ikon dźwiękowych (*auditory icons*) lub innego rodzaju dźwięków tzw. *earcons*. Ikony dźwiękowe są akustycznymi odpowiednikami prawdziwych zdarzeń wytwarzających dźwięki (np. dźwięk gniecionej kartki papieru informującej o umieszczeniu pliku w koszu), a *earcons* są dźwiękami wirtualnymi nie mającymi odpowiednika w dźwiękach naturalnych. W pracach (McGookin i Brewster, 2004) proponuje się aby takie komunikaty dźwiękowe interpretować jako elementy składowe pewnego „języka dźwiękowego”, z regułami semantycznego i syntaktycznego ich łączenia podobnymi do stosowanych w językach naturalnych.

Jednym z podstawowych parametrów takich systemów jest wytwarzanie u słuchacza realistycznego wrażenia akustycznego polegającego na słyszeniu przestrzennej, trójwymiarowej sceny dźwiękowej. Przykładem systemu, w którym uzyskano efekt dźwięku dookólnego z zastosowaniem wielokanałowych torów dźwiękowych i odpowiedniego rozmieszczenia zestawów głośnikowych był system Ambisonics (Fellget, 1975).

Ogólny schemat blokowy współczesnych systemów wytwarzania wirtualnego środowiska dźwiękowego pokazano na rys. 4.8. W schemacie tym wyróżniono bloki reprezentujące źródło dźwięku, środowisko propagacji fal dźwiękowych oraz odbiorcę (słuchacza).

Systemy substytucji sensorycznej, w których stosuje się techniki sonifikacji do dźwiękowego obrazowania otoczenia dla osób niewidomych można zaliczyć do szczególnej klasy systemów wirtualnych wyświetlaczy dźwiękowych (Sánchez, Lumbreira, 1999). Ze względu na wymóg mobilności systemu prezentowane dźwięki są wytwarzane w słuchawkach stereofonicznych lub miniaturowych głośnikach (Cooper i Petrie, 2004). Za pomocą technik sonifikacji można w bardziej skuteczny i szybszy sposób niż za pomocą komunikatów słownych prezentować otoczenie i przekazywać wskazówki nawigacyjne osobom niewidomym (Ardito i inni, 2007). Praktycznie wszystkie systemy substytucji sensorycznej wzroku wykorzystujące prezentację dźwiękową stosują techniki sonifikacji do akustycznego obrazowania otoczenia (Hersh i Johnson, 2008). Szczególnie interesujące są



Rys. 4.8. Ogólny schemat blokowy systemu do wytwarzania wirtualnego środowiska dźwiękowego
Źródło: opracowanie własne

metody wyświetlania dźwiękowego z zastosowaniem technologii uwzględniającej charakterystyki akustyczne głowy słuchacza umożliwiające uzyskanie efektu przestrzenności wytwarzanych dźwięków (Hartmann i Wittenberg, 1996).

4.2.1.2. Modele słyszenia przestrzennego i ich identyfikacja

Słyszenie przestrzenne, tj. zdolność człowieka do lokalizacji źródeł dźwięku jest tradycyjnie wyjaśniane na drodze porównania sygnałów dźwiękowych odbieranych przez ucho lewe i ucho prawe. Znany fakt jest jednak umiejętność określania kierunku źródła dźwięku na podstawie słyszenia jednousznego.

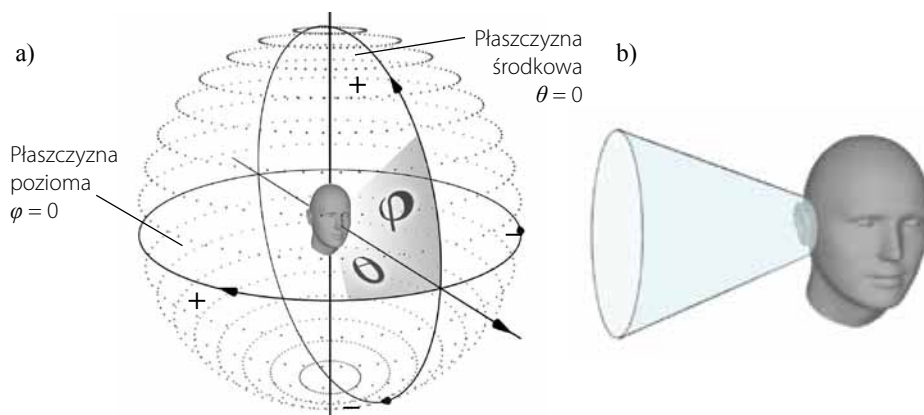
Mechanizm słyszenia przestrzennego u człowieka wyjaśnia się zatem na gruncie dwóch modeli (Moore, 2004):

- słyszenia dwuusznego (tzw. teoria duplex),
- słyszenia jednousznego.

W układzie współrzędnych biegunowych związanych z umownym środkiem głowy słuchacza kierunek źródła dźwięku jest definiowany przez kąt azymutu – θ i kąt elewacji (wzniesienia) – φ (rys. 4.9a). Dokładność spostrzegania kierunku źródła dźwięku jest zależna od jego położenia kąтового i charakterystyki promieniowania źródła. Dokładność ta jest definiowana dla źródeł statycznych oraz ruchomych. Pomiary dla źródeł ruchomych służą do określenia najmniejszej zauważalnej zmiany położenia kąтового źródła (*minimum audible angle* – MAA).

Do lokalizacji przestrzennej źródła dźwięku jest również konieczna informacja o odległości źródła od słuchacza (Strumiłło, 2011). Jest ona wypracowana na podstawie własności propagacyjnych fali dźwiękowej w środowisku (Funkhouser i inni, 2002), (Loomis i inni, 1999) na które składają się:

- tłumienie dźwięku (dwukrotnie zmniejszanie natężenia dźwięku z każdym podwojeniem odległości źródła),
- zmiana barwy dźwięku wraz z oddaleniem źródła (składowe o większej częstotliwości podlegają większym tłumieniom),
- wielokrotne odbicia fali dźwiękowej w środowisku pogłosowym (efekt pogłosu i echa).



Rys. 4.9. Geometria słyszenia przestrzennego: a) biegunowy układ współrzędnych związany z umownym środkiem głowy słuchacza; „+” i „-” oznaczają odpowiednio dodatni i ujemny zakres kątów azymutu i elewacji; okręgi wykreślone linią ciągłą wyznaczają płaszczyznę poziomą i środkową, b) kształt i położenie stożka niejednoznaczności, dla którego wartości ILD i ITD są stałe

Źródło: opracowanie własne

W modelu słyszenia dwuusznego (*binaural hearing*) zaproponowanym przez Rayleigha na początku XX w. przyjmuje się, że ucho człowieka nie ma charakterystyki kierunkowej, a lokalizacja źródła dźwięku jest określana na podstawie dwóch przesłanek: różnicy czasu (*interaural time difference – ITD*) oraz różnicy natężenia (*interaural level difference – ILD*) fali dźwiękowej dochodzącej do lewego i prawego ucha. Różnica natężenia dźwięku *ILD* wynika głównie z tłumienia fali dźwiękowej przez głowę, która dla dużego zakresu kątów azymutu przesłania źródło dźwięku dla ucha bardziej oddalonego od źródła dźwięku. Przyjmuje się, że w paśmie częstotliwości do 1500 Hz główną rolę w postrzeganiu kierunku źródła dźwięku pełni *ITD* (*ILD* ma małą wartość z uwagi na zjawisko dyfrakcji tej długości fali dźwiękowej (ok. 40 cm) na powierzchni głowy). Dla częstotliwości powyżej umownej granicy 1500 Hz główną rolę w lokalizacji dźwięku pełni *ILD* (gdyż określenie międzyusznnej różnicy fazy fal dźwiękowych o dużej częstotliwości jest niejednoznaczne). Model słyszenia dwuusznego nie wyjaśnia jednak następujących własności percepcyjnych słuchu:

- zdolności do wskazywania kierunku dźwięku w odsłuchu jednousznym,
- zdolności do rozróżniania kierunku dźwięku dla źródeł leżących w płaszczyźnie środkowej, dla której $ILD = ITD = 0$ oraz źródeł położonych na powierzchni bocznej tzw. stożka niejednoznaczności (Moore, 2004), dla której $ILD = const_{ILD}$ i $ITD = const_{ITD}$ (rys. 4.8b).

Batteau (1967) wykazał, że decydujące znaczenie w przestrzennym spostrzeganiu dźwięków pełni małżowina uszna, której kształt decyduje o dużej kierunkowości słyszenia. Z kolei Gardner (1973) na podstawie wielu badań odsłuchowych stwierdził, że jeżeli źródło dźwięku zawiera składowe o częstotliwości z zakresu 7÷10 kHz to jego kierunek może być najdokładniej spostrzegany. Zauważmy, że zakres tych częstotliwości dźwięku odpowiada fali dźwiękowej o długości ok. 30÷43 mm, tj. porównywalnej z rozmiarem małżowiny usznej.

Na podstawie wyżej opisanych i wielu innych badań akustycznych wykazano, że małżowiny uszne wraz z głową i torsem tworzą kierunkowy filtr akustyczny, którego parametry są określane przez zjawiska odbicia, ugięcia, rozproszenia i rezonansu fali dźwiękowej w interakcji z małżowiną uszną (Freedman i Fisher, 1968), (Duda, 1996). Zatem sygnały akustyczne dobiegające do błony bębenkowej ucha lewego i prawego są poddawane różnym filtracjom i na tej podstawie odpowiednio interpretowane w korze słuchowej dla wypracowania informacji o kierunku źródła dźwięku (Shamma, 2001).

Zauważmy, że jeżeli dźwięki są wytwarzane w słuchawkach stereofonicznych to nie podlegają takiej naturalnej filtracji kierunkowej, gdyż ciśnienie akustyczne działa bezpośrednio na błonę bębenkową. Efekt przestrzenności dźwięku próbuje się zatem uzyskać poprzez oddzielne sterowanie parametrami definiowanymi w modelu dwuusznym słyszenia ILD i ITD (głównie ILD). Wykazano wcześniej, że model ten nie umożliwia jednoznacznego odtworzenia wszystkich kierunków źródeł dźwięku.

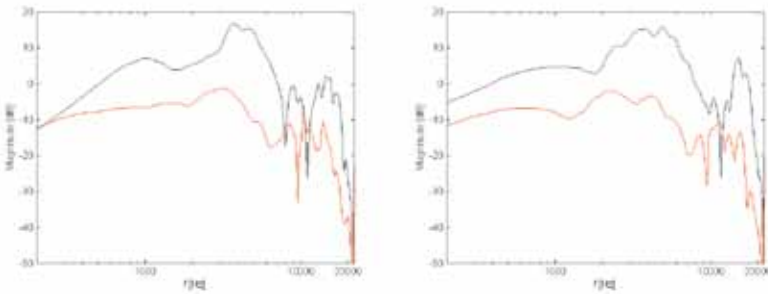
Wnioskujemy zatem, że uzyskanie tzw. efektu eksternalizacji dźwięku (*sound externalization*) tj. wrażenia percepcyjnego lokalizacji źródeł dźwięku jako pochodzących „z zewnątrz głowy” jest możliwe na drodze identyfikacji parametrów filtru kierunkowego wynikającego z anatomii ciała słuchacza (Hartmann i Wittenberg, 1966). Charakterystyki widmowe takich filtrów uwzględniają zjawiska akustyczne związane z interakcją dźwięku z ciałem słuchacza, w szczególności głową i małżowinami usznymi. Przyjmując liniowość tych zjawisk, transmitancję filtrów można modelować porównując charakterystykę widmową źródła dźwięku oraz widmo sygnału rejestrowanego u wejścia przewodów słuchowych ucha lewego i prawego (Cheng i Wakefield, 2001). Stosunek tych dwóch widm, wyznaczany dla szerokiego zakresu kierunków

kątowych (θ, φ) położenia źródła dźwięku wokół głowy słuchacza określany jest mianem funkcji transmitancji akustycznej odniesionej do głowy (*head related transfer functions – HRTFs*). Funkcje te, które są funkcjami trzech zmiennych $H_L(\theta, \varphi, \omega)$, $H_R(\theta, \varphi, \omega)$ wyznacza się oddzielnie dla ucha lewego i prawego. Na rys. 4.10 wykreślono przykładowe funkcje HRTF wyznaczone dla dwóch położeń źródła dźwięku umieszczonego po lewej stronie słuchacza ($\theta_1 = \theta_2 = -45^\circ$) dla różnych kątów wzniesienia, tj. $\varphi_1 = -11^\circ$ i $\varphi_2 = 11^\circ$.

Reprezentacją funkcji HRTF w dziedzinie czasu są akustyczne odpowiedzi impulsowe odniesione do głowy (*head related impulse response – HRIR*):

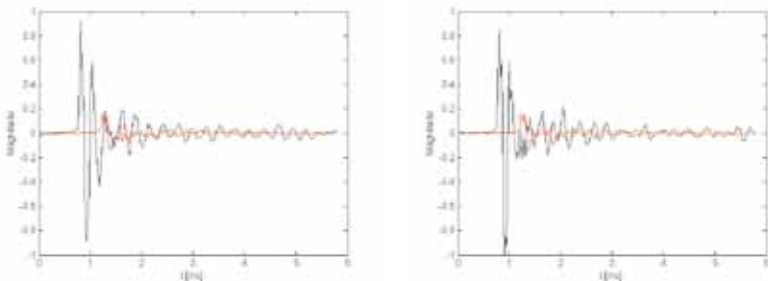
$$HRIR(\theta, \varphi, t) = \mathfrak{F}^{-1}(HRTF(\theta, \varphi, \omega)) \quad (4.1)$$

gdzie: $\mathfrak{F}^{-1}$ – jest operatorem odwrotnego przekształcenia Fouriera. Na rys. 4.11 wykreślono odpowiedzi impulsowe odpowiadające funkcjom HRTF pokazanym na rys. 4.10 (Pec i inni, 2007).



Rys. 4.10. Przykładowa para charakterystyk HRTF (widma amplitudowe) wyznaczona dla położeń kątowych $\theta_1 = -50^\circ$ i $\varphi_1 = -9^\circ$ (po lewej) i $\theta_2 = -50^\circ$ i $\varphi_2 = 9^\circ$ (po prawej); linią czarną wykreślono charakterystykę dla ucha lewego a czerwoną dla ucha prawego (zauważmy większe amplitudy widma dla ucha lewego w porównaniu z uchem prawym)

Źródło: opracowanie własne



Rys. 4.11. Pary odpowiedzi impulsowych HRIR odpowiadające funkcjom HRTF pokazanym na rys. 4.10

Źródło: opracowanie własne

Funkcje HRTF można uzyskać z pomiaru lub odpowiedniego modelowania. Dostępne są bazy danych z pomiarami charakterystyk HRTF, m.in. baza CIPIC zawierająca pomiary HRTF 43 osób (Algazi i inni, 2011) oraz baza opracowana w Media Lab MIT z wykorzystaniem manekina KEMAR (Gardner i inni, 1994). Opracowano też techniki modelowania numerycznego HRTF na podstawie geometrycznego modelu głowy i małżowiny usznej (Plaskota, 2006). Zastosowane techniki modelowania są jednak drogie i w praktyce rzadko stosowane.

Ponieważ funkcje HRTF są unikatową cechą osobniczą, najlepsze, potwierdzone testami odsłuchowymi wyniki eksternalizacji uzyskuje się na podstawie indywidualnych pomiarów charakterystyk akustycznych głowy przeprowadzanych w warunkach bezpogłosowych (Cheng i Wakefield, 2001).

Pomiar funkcji HRTF polega na rejestracji sygnałów emitowanych przez szerokopasmowe źródło dźwięku rozmieszczane w polu dalekim w różnych położeniach kątowych wokół głowy słuchacza. Pomiar jest wykonywany za pomocą miniaturowych mikrofonów umieszczonych w kanałach słuchowych osoby badanej. W celu eliminacji wpływu transmitancji środowiska pomiarowego $A(\omega)$ pomiar jest wykonywany dwukrotnie (rys. 4.12) (Dobrucki i inni, 2010).

1. Pomiar z udziałem osoby badanej, w którym jest wyznaczane widmo sygnału mierzonego w mikrofonach umieszczonych w kanałach słuchowych. Widmo sygnału mierzonego dla znanej charakterystyki widmowej źródła dźwięku $S(\omega)$ można zapisać w postaci:

$$M(\theta, \varphi, \omega) = H(\theta, \varphi, \omega)S(\omega) \quad (4.2)$$

gdzie: $H(\theta, \varphi, \omega)$ – jest charakterystyką akustyczną głowy z uwzględnieniem charakterystyki akustycznej środowiska pomiarowego.

2. Pomiar bez udziału osoby badanej (tzw. pomiar odniesienia), w którym mikrofony pozostają w niezmiennym położeniu geometrycznym, tak jak w pkt. 1. Widmo sygnału mierzonego można zapisać jako iloczyn transmitancji widmowych środowiska pomiarowego $A(\omega)$ i charakterystyki widmowej źródła dźwięku $S(\omega)$:

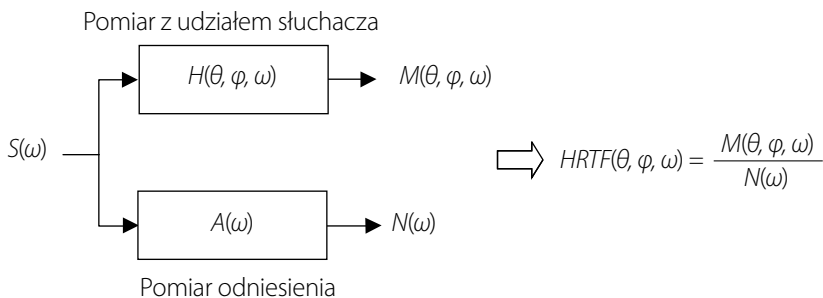
$$N(\omega) = A(\omega)S(\omega) \quad (4.3)$$

Poszukiwaną funkcję HRTF w postaci widma amplitudowego i fazowego uzyskujemy odpowiednio z zależności:

$$|HRTF(\theta, \varphi, \omega)| = \frac{|M(\omega)|}{|N(\omega)|} = \frac{|M(\theta, \varphi, \omega)|}{|A(\omega)||S(\omega)|} \quad (4.4)$$

$$\angle HRTF(\theta, \varphi, \omega) = \angle M(\theta, \varphi, \omega) - \angle N(\omega) = \angle M(\theta, \varphi, \omega) - \angle A(\omega) - \angle S(\omega) \quad (4.5)$$

gdzie: $| |$ – jest modulem liczby zespolonej a $\angle$ – jest jej argumentem.



Rys. 4.12. Charakterystyki widmowe wyznaczone w procedurze pomiaru funkcji HRTF:

$M(\theta, \varphi, \omega)$ – widmo sygnału zmierzonego z udziałem osoby badanej,

$N(\omega)$ – widmo sygnału zmierzonego bez udziału słuchacza

Źródło: opracowanie własne

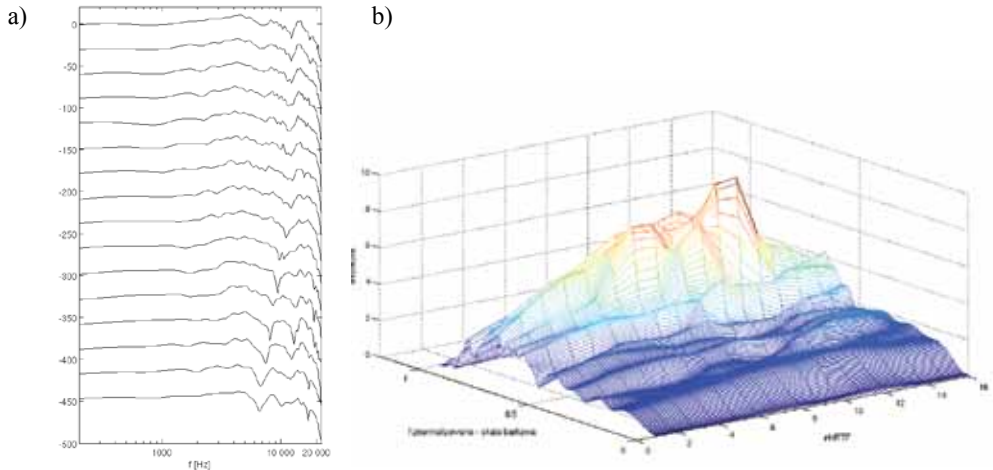
Na rys. 4.13 pokazano zdjęcie użytkowanego w Instytucie Elektroniki systemu do pomiaru charakterystyk HRTF wykonanego we współpracy z Zakładem Akustyki Politechniki Wrocławskiej (Dobrucki i inni, 2010). System składa się z 16-głośników rozmieszczonych na łuku o promieniu ok. 1 metra, obrotowego fotela sterowanego elektrycznie oraz dwukanałowego toru pomiarowego. System umożliwia automatyczny pomiar charakterystyk HRTF w szerokim zakresie kątów azymutu: $\theta = \langle -180^\circ, \dots, 180^\circ \rangle$ z regulowanym skokiem i kątów wzniesienia $\varphi = \langle -45^\circ, \dots, 90^\circ \rangle$ ze stałym skokiem $\Delta\varphi = 9^\circ$. Maksymalna liczba punktów pomiarowych funkcji HRTF wynosi 5760 (dla skoku kąta azymutu $\Delta\theta = 1^\circ$). Procedura pomiaru charakterystyk HRTF trwa około pół godziny, przy czym sam proces rejestracji charakterystyk trwa ok. 10 min, a dodatkowe 10 minut zajmuje pomiar charakterystyk odniesienia. Rys. 4.13 ilustruje widok systemu w czasie przygotowania procedury pomiarowej.

Dla ilustracji zmiany kształtu funkcji HRTF zachodzących przy zmianie położenia źródła dźwięku na rys. 4.14a wykreślono widma amplitudowe funkcji HRTF dla kolejno zmienianych kątów wzniesienia i stałego azymutu kierunku źródła dźwięku. Największą zmienność w kolejno wykreślonych charakterystykach można zauważyć dla zakresu częstotliwości 6÷12 kHz. Potwierdza to hipotezę o kluczowym znaczeniu tego pasma częstotliwości dla lokalizacji źródeł dźwięku. Na rys. 4.14b wykreślono funkcje HRTF w układzie współrzędnych,



Rys. 4.13. Zdjęcie urządzenia do pomiaru charakterystyk akustycznych głowy (HRTF) wykonane w chwili przygotowania pomiaru

Źródło: fotografia własna



Rys. 4.14. Wykres a): widma amplitudowe HRTF ucha lewego wykreślone dla stałego azymutu $\theta = -80^\circ$ i wzniesień $\varphi = \langle 90^\circ \div -11^\circ \rangle$, (kolejne charakterystyki przesunięte względem siebie o 30dB); wykres b): obraz zbioru funkcji HRTF wykreślonych w rzucie aksonometrycznym dla jednego kąta azymutu ($\theta = 50^\circ$) i szesnastu kątów wzniesienia $\varphi = \langle 90^\circ \div -45^\circ \rangle$

Źródło: opracowanie własne

w którym oś częstotliwości jest wyrażona w tzw. skali barkowej². Pec (2010) stosował tę skalę psychoakustyczną do projektowania filtrów cyfrowych, dających bardzo dobre odwzorowanie funkcji HRTF przy małym rzędzie filtru.

4.2.1.3. Wytwarzanie dźwięków przestrzennych

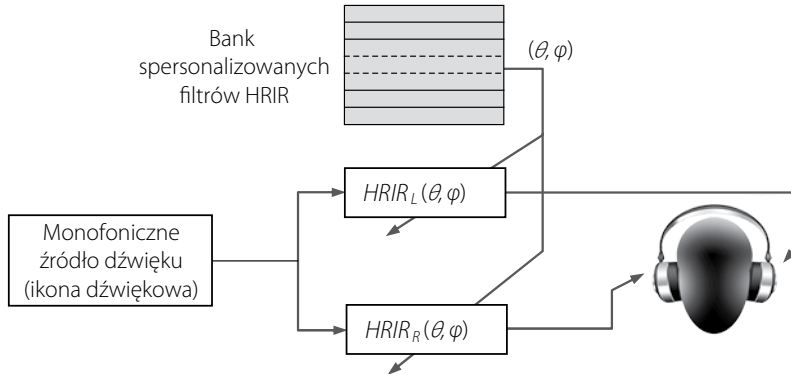
W najprostszych systemach ostrzegania o przeszkodach są wytwarzane monofoniczne dźwięki informujące o odległości czujnika systemu do wskazywanego ręcznie obiektu (przeszkody). Takie cechy strumieni dźwiękowych jak: częstotliwość powtarzania krótkich tonów (analogicznie do stosowanych w czujnikach parkowania), wysokość tonu lub natężenie dźwięku są stosowane do odwzorowania odległości (lub głębi) do wskazywanej przeszkody.

W bardziej zaawansowanych systemach, w których wykorzystuje się obrazowanie scen trójwymiarowych (np. z zastosowaniem stereowizji) są stosowane bardziej złożone metody sonifikacji i wytwarzania efektu przestrzennej sceny dźwiękowej. Najprostszym, ale niepraktycznym sposobem zbudowania takiego systemu jest przytwierdzenie źródeł dźwięku do obiektów sceny oraz zastosowanie odpowiednich ikon dźwiękowych (lub *earcons*) jako akustycznych reprezentacji tych obiektów. Na podstawie takiej sceny dźwiękowej osoba niewidoma mogłaby lokalizować przeszkody na swojej drodze, uzyskiwać informacje o ich cechach i skutecznie je omijać.

Dzięki opracowanej technice rejestracji funkcji HRTF jest możliwe wytworzenie efektu przestrzenności dźwięków w osobistych systemach obrazowania dźwiękowego otoczenia przez zastąpienie (pominiętego w odsłuchu słuchawkowym) układu akustycznego głowa-małowiny uszne jego odpowiednim modelem w postaci osobistych funkcji HRTF. Jeżeli użyć funkcji $HRIR = \mathfrak{I}^{-1}(HRTF)$ jako filtrów cyfrowych (odpowiednio dla ucha lewego i prawego) w stereofonicznym torze dźwiękowym to można wytworzyć u słuchacza subiektywne wrażenie spostrzegania strumieni dźwiękowych jako umiejscowionych w przestrzeni otaczającej słuchacza. Na rys. 4.15 pokazano uproszczony schemat blokowy systemu do wytwarzania dźwięków przestrzennych. W praktycznej realizacji systemu przydatne jest oddzielne modelowanie widm amplitudowych i fazowych

² Skala barkowa jest jedną ze stosowanych w psychoakustyce tzw. perceptualnych skal częstotliwości przybliżających spostrzeganie dźwięków przez człowieka. Utworzono ją na podstawie badań szerokości tzw. pasm krytycznych, tj. właściwości słuchu człowieka polegającej na maskowaniu cichszych dźwięków przez głośniejsze dźwięki o podobnej częstotliwości (Fastl i Zwicker, 1999). Skala barowa jest przybliżana odpowiednimi skalami logarytmicznymi (pasma krytyczne zawężają się ze wzrostem częstotliwości) i jest wykorzystywana w algorytmach kompresji sygnałów dźwiękowych.

HRTF (zob. równania 4.4 i 4.5), dzięki któremu w torze dźwiękowym systemu można zastosować kaskadowe połączenie bloku opóźniającego i minimalno-fazowy filtr HRIR (tj. o minimalnym opóźnieniu grupowym i w konsekwencji mniejszym rzędzie) (Pec, 2010).



Rys. 4.15. Uproszczony schemat wytwarzania dźwięku przestrzennego z zastosowaniem akustycznych odpowiedzi impulsowych odniesionych do głowy (HRIR); monofoniczne źródło dźwięku jest spostrzegane przez słuchacza jako dobiegające z kierunku (θ, φ)

Źródło: opracowanie własne

W pracach (Pec i inni, 2007) do weryfikacji zmierzonych u 15 osób (w tym 8 osób niewidomych) funkcji HRTF zastosowano program SLAB opracowany w Spatial Auditory Displays Lab w NASA (SLAB, 2011). Jest to program typu *open source* umożliwiający wytwarzanie scen dźwiękowych (m.in. ze statycznymi i ruchomymi źródłami dźwięków) z zastosowaniem zadanych charakterystyk HRTF. Zbadano występowanie efektu eksternalizacji dźwięku. Na rys. 4.16 zilustrowano zdjęcia z przeprowadzonych testów odsłuchowych, a w tabeli 4.4 zamieszczono wyniki pokazujące, że tylko u jednej osoby niewidomej taki efekt nie wystąpił.



Rys. 4.16. Test lokalizacji; osoby niewidome wskazują położenie wirtualnego źródła dźwięku

Źródło: opracowanie własne

Tabela. 4.4. Wyniki testów eksternalizacji dźwięków

Efekt eksternalizacji dźwięku	Liczba osób
Eksternalizacja i dokładna lokalizacja źródeł dźwięku	8 (4 osoby niewidome)
Eksternalizacja i mało dokładna lokalizacja źródeł dźwięku	4 (2 osoby niewidome)
Eksternalizacja po krótkim treningu	2 (2 osoby niewidome)
Nie uzyskano efektu eksternalizacji źródła dźwięku	1 (niewidomy)

Źródło: opracowanie własne

Wiele prac badawczych poświęcono weryfikacji systemów wytwarzania dźwięków przestrzennych z zastosowaniem funkcji HRTF. Badano przede wszystkim dokładność spostrzegania kierunku wirtualnych źródeł dźwięków w funkcji kąta elewacji i wzniesienia oraz porównywano te dokładności do dokładności spostrzegania rzeczywistych źródeł dźwięków (Bujacz, 2010). W tabeli 4.5 zebrano wyniki testów odsłuchowych opisanych w literaturze. Warto zauważyć, że dla badań wykonanych z zastosowaniem uśrednionych funkcji HRTF błędy lokalizacji źródeł dochodzą do 60°.

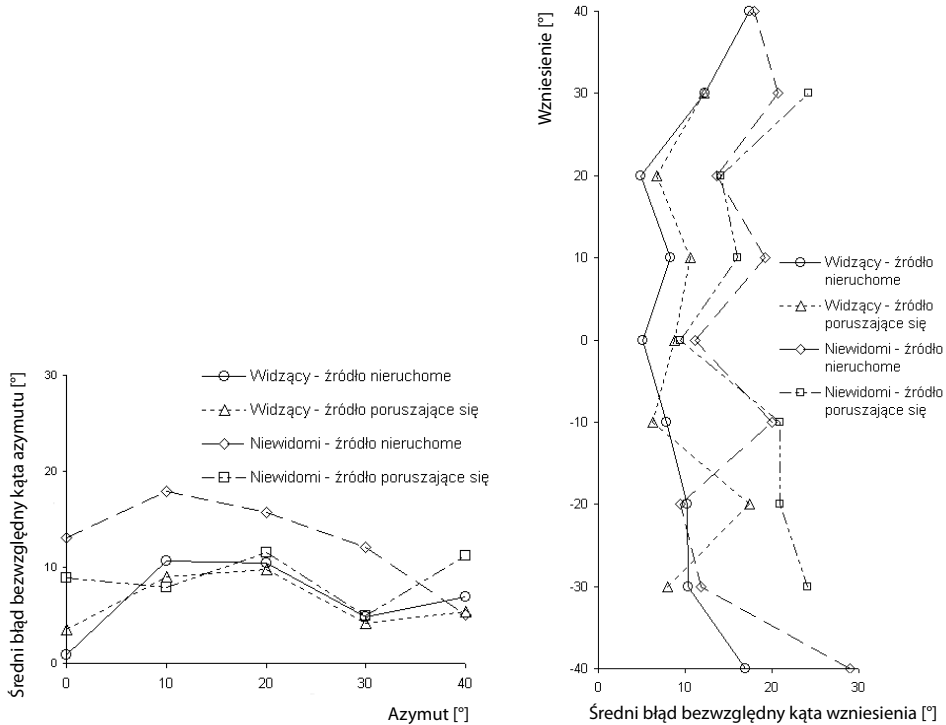
Tabela. 4.5. Dokładność lokalizacji źródeł dźwięku (wyrażona jako średni błąd wskazania lokalizacji katowej źródła)

System wytwarzania dźwięku	Rozdzielczość dla kąta azymutu*	Rozdzielczość dla kąta wzniesienia*	Literatura
Źródło rzeczywiste	1°–10°	10°–20°	(Calamia, 1998), (Pec i inni, 2007), (Dobrucki i inni, 2010), (Pedersen i Jorgensen, 2005)
Konwencjonalna tereofonia (ILD)	5°–15°	brak danych	(Calamia, 1998), (Pedersen i Jorgensen, 2005)
Z zastosowaniem funkcji HRTF z bazy danych	4°–15°	12°–60°	(Duda, 1996), (Pec i inni, 2007), (Dobrucki i inni, 2010), (Pedersen i Jorgensen, 2005)
Z zastosowaniem osobistych funkcji HRTF	5°–22°	10°–38°	(Angel i inni, 2002), (Duda, 1996) (Pec i inni, 2007), (Wersenyi, 2007) (Dobrucki i inni, 2010)

* Zamieszczone porównanie jest szacunkowe z uwagi na stosowane różne źródła dźwięku (szum szerokopasmowy, impulsy szumów, inne), różne metody oceny błędów oraz różne osoby biorące udział w badaniach.

Źródło: opracowanie własne na podstawie podanego w tabeli wykazu literatury

Interesujące jest też porównanie zdolności lokalizacji źródeł dźwięku przez osoby sprawne wzrokowo i osoby niewidome przeprowadzone w pracach (Pec i inni, 2007) i (Dobrucki i inni, 2010). Na rys. 4.17 oraz w tabeli 4.6 zestawiono wyniki takiego testu porównawczego wykonanego dla 15 osób (w tym 9 osób



Rys. 4.17. Porównanie dokładności z jaką lokalizują źródła dźwięku osoby niewidome i osoby sprawne wzrokowo
Źródło: opracowanie własne

Tabela. 4.6. Dokładność lokalizacji źródeł dźwięku (wyrażona jako średni błąd wskazania lokalizacji kątowej źródła)

Rodzaj źródła dźwięku	Średni błąd lokalizacji dźwięku [°]			
	Sprawni wzrokowo		Osoby niewidome	
	Azymut	Wzniesienie	Azymut	Wzniesienie
Źródło nieruchome – samogłoska	8,1	15,4	14,4	22,9
Źródło nieruchome – sygnał chirp*	6,7	10,2	12,8	16,8
Źródło ruchome	6,4	9,5	8,9	16,0

* tzw. sygnał świergotowy – uzyskiwany z modulacji częstotliwościowej sygnału harmonicznego

Źródło: opracowanie własne

niewidomych) dla różnych rodzajów źródeł dźwięku oraz oddzielnie dla źródeł nieruchomych i poruszających się. Wyniki te pokazują, że osoby niewidome wskazują źródła dźwięku z mniejszą dokładnością kątową.

Zamieszczony w niniejszej części pracy przegląd literatury i wyników badań dotyczących metod obrazowania dźwiękowego można podsumować następująco:

- potwierdzono skuteczność techniki generacji dźwięków przestrzennych z wykorzystaniem funkcji HRTF w słuchawkowym torze stereofonicznym,
- efekt eksternalizacji występuje u zdecydowanej większości badanych osób dla spersonalizowanych funkcji HRTF (zarówno dla osób sprawnych wzrokowo jak i osób niewidomych),
- mniejsze błędy lokalizacji dźwięków uzyskuje się dla źródeł o szerokim zakresie częstotliwości (dla osób niewidomych średnio 13° dla azymutu i 17° dla wzniesienia),
- błąd lokalizacji źródeł poruszających się jest mniejszy niż dla źródeł nieruchomych,
- osoby niewidome lokalizują źródła dźwięku z nieznacznie mniejszą dokładnością niż osoby sprawne wzrokowo (nie we wszystkich pracach wynik ten jest potwierdzony (Wersenyi, 2007)).

Warto również zwrócić uwagę na wyniki pracy (Kato i inni, 2003), z której wynika, że dokładność postrzegania lokalizacji źródeł statycznych i ruchomych poprawia się przy tzw. aktywnym poznaniu, tj. przy niewielkich ruchach głowy wykonywanych mimowolnie przez słuchacza dla uzyskania lepszej lokalizacji przestrzennej źródła.

W niniejszym rozdziale nie dyskutowano również pewnych ważnych praw percepcji słuchowej, które należy uwzględnić w systemach dźwiękowego obrazowania otoczenia. Należą do nich m.in. (Moore, 2004):

- możliwość skupiania uwagi słuchowej na niewielkiej liczbie strumieni dźwiękowych (Bregman, 1990),
- brak możliwości natychmiastowego przeniesienia uwagi słuchowej z jednego strumienia dźwiękowego na drugi strumień, jeżeli jest on położony w dużej odległości kątowej od strumienia pierwszego,
- efekt maskowania dźwięku o mniejszej głośności przez dźwięki o większej głośności (efekt ten silnie zależy od różnicy częstotliwości tych dźwięków i od bezwzględnych wartości tych częstotliwości),
- efekt pierwszeństwa (*precedence effect*), jest to zjawisko stereofoniczne występujące w rezultacie odtwarzania w kanałach dźwięków nieznacznie przesuniętych w czasie ($5\text{ ms} \div 40\text{ ms}$); efekt ten polega na tym, że słuchacz przypisuje dźwięk do tego źródła, od którego dźwięk dociera pierwszy do ucha (pomimo jego mniejszej głośności); efekt ten jest silnie zależny od widma i głośności dźwięków.

Podstawowym wnioskiem jaki można sformułować na podstawie przeprowadzonych badań i referowanych prac jest stwierdzenie, że techniki sonifikacji uwzględniające prawa psychologii słyszenia i charakterystyki HRTF umożliwiają wytworzenie przestrzennej sceny dźwiękowej, która może służyć do dźwiękowego obrazowania otoczenia osobom niewidomym.

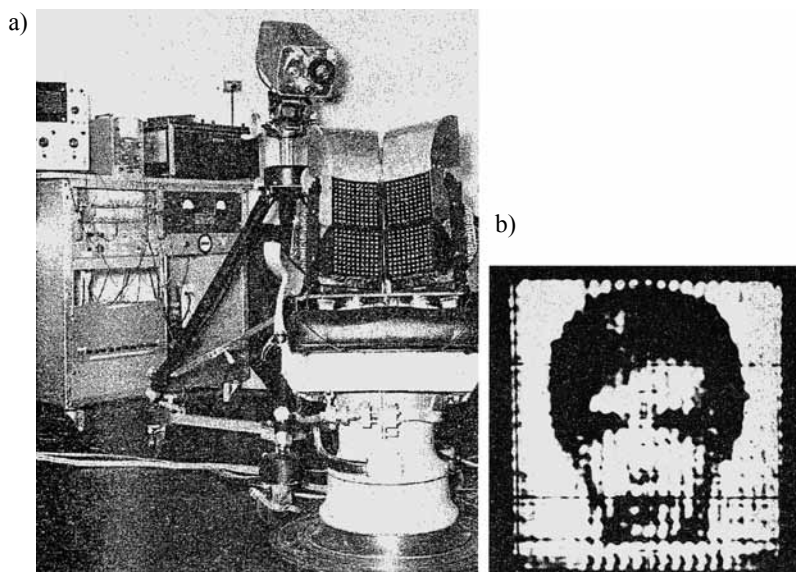
4.2.2. Dotykowe i haptyczne techniki obrazowania

Zmysł dotyku jest umiejscowiony w receptorach skórnych i pozwala eksplorować w sposób bezpośredni tylko tzw. przestrzeń osobistą wg Halla, tj. otoczenie w zasięgu kończyn. Zasięg tego zmysłu można w łatwy sposób ale w niewielkim stopniu powiększać. Biała laska jest podstawową pomocą (*primary travel aid*) w samodzielnym poruszaniu się niewidomego. Znamiennym spostrzeżeniem jest fakt, że niewidomi odbierają drgania białej laski, których źródło umiejscawiają u jej dolnego końca, a nie w rękojeści. Potrafią zatem wytworzyć mentalny obraz otoczenia odwzorowujący wycinek przestrzeni wykraczającej poza pole przestrzeni osobistej. Loomis (1992) określił ten efekt mianem atrybucji bodźca dystalnego (odległego). Jednym z pierwszych, elektronicznych urządzeń do dotykowego obrazowania otoczenia był prototypowy system zbudowany przez Bach-y-Ritę i innych (1972), (rys. 4.18).

Autorzy systemu stwierdzili, że po odpowiednim treningu osoby niewidome potrafiły wytworzyć wrażenia zmysłowe, w których spostrzegały obrazy otoczenia umiejscowione w przestrzeni odległej, a nie bezpośrednio w miejscu ich wytworzenia na powierzchni skóry. Wyniki prac Bach-y-Rity potwierdzone w innych badaniach (Klatzky i Lederman, 2007), (Visell, 2009), (Jansson, 2008) pokazały, że dotyk może być zmysłem, który można skutecznie wykorzystać do obrazowania otoczenia. Zaletą takiego systemu substytucji sensorycznej wzroku (w porównaniu do systemów obrazowania dźwiękowego) jest nieangażowanie zasobów percepcyjnych słuchu, który jest wykorzystywany zarówno do werbalnego komunikowania się jak i spostrzegania dźwiękowych wskazówek nawigacyjnych (Rouben i Terveen, 2007), (Kaczmarek, 1999).

Występują dwa mechanizmy związane z układem czuciowym człowieka (Lederman i Klatzky, 2009):

- czucie powierzchniowe (dotykowe), którego receptory znajdują się w skórze,
- czucie głębokie, którego receptory są umiejscowione wewnątrz ciała, tj. w mięśniach, stawach, ścięgnach i więzadłach.



Rys. 4.18. Prototypowy system dotykowego obrazowania otoczenia TVSS (*tactile vision substitution system*) składający się z kamery TV i matrycy aktuatorów mechanicznych zamocowanych w oparciu fotela (a), obraz dotykowy utworzony przez matrycę aktuatorów o rozdzielczości 20×20 (każdy element wibruje z amplitudą proporcjonalną do jasności odwzorowywanego fragmentu obrazu)

Źródło: (Bach-y-Rita i inni, 1969)

Czucie głębokie jest źródłem informacji zarówno o aktywności motorycznej ciała i kończyn oraz o świadomości ich wzajemnego położeniu w bezruchu. Czuć to określa się też mianem zmysłu kinestetycznego.

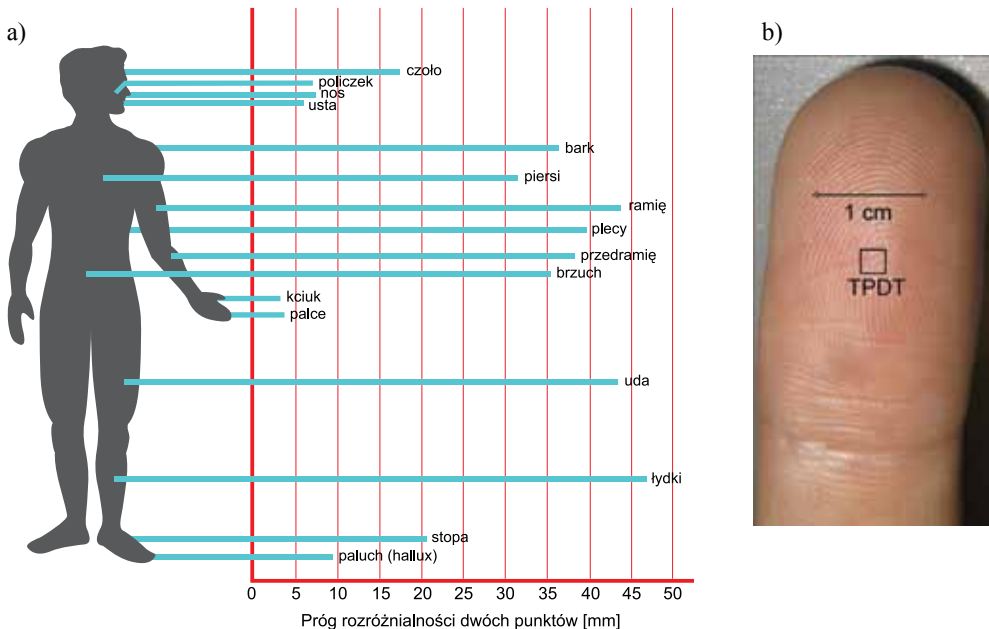
W związku z ww rodzajami układów czuciowych rozróżnia się tzw. spostrzeganie dotykowe (*tactile*) i spostrzeganie haptyczne (*haptic*). Pierwsze odnosi się do czucia powierzchniowego i nie musi wiązać się ruchem ciała. Drugie natomiast obejmuje zarówno czucie dotykowe jak i czucie głębokie. Jest ono też nazywane aktywnym dotykiem. Odpowiednio też system Bachy-y-Rity pokazany na rys. 4.18 można określić mianem interfejsu dotykowego, a klawiaturę komputera interfejsem haptycznym wymagającym ruchu dłoni i palców przy jego obsłudze.

Przy budowie interfejsów dotykowych należy uwzględnić gęstość, rodzaj i właściwości sensoryczne receptorów czuciowych występujących w skórze. Można je podzielić na trzy zasadnicze rodzaje: mechanoreceptory, termoreceptory i receptory bólu (Longstaff, 2006). Dla budowy interfejsów haptycznych najistotniejsze znaczenie mają mechanoreceptory. Wyróżnia się dwa podstawowe rodzaje mechanoreceptorów:

- receptory powierzchniowe występujące na granicy naskórka i skóry (ciałka Meissnera i łąkolki dotykowe Merkela – tzw. receptory I typu),
- receptory umiejscowione w głębszych warstwach skóry (ciałka Ruffiniego i Paciniego) – tzw. receptory II typu.

Receptory I typu mają małe pole recepcyjne ale dużą gęstość (największą na opuszkach palców ($\sim 140/\text{cm}^2$), ustach i języku) i wrażliwość (od 20 mN siły nacisku). Pełnią one podstawową rolę w percepcji kształtu i faktury dotykanej powierzchni. Największa wrażliwość receptorów powierzchniowych występuje dla częstotliwości bodźców dotykowych ok. 50 Hz. Receptory II typu mają większe pole recepcyjne, występują rzadziej i z największą skutecznością wykrywają silniejsze bodźce mechaniczne o częstotliwości wibracji w zakresie 200÷300 Hz. Potwierdzono, m.in. że osoby w czasie marszu najlepiej zauważają właśnie ten zakres częstotliwości wibracji telefonu komórkowego (Yim i inni, 2007).

Na rys. 4.19a pokazano tzw. próg rozróżnialności dwóch punktów (*two-point discrimination threshold* – TPDT) dla ciała człowieka, który należy interpretować jako minimalną odległość przy której są rozróżniane dwa punktowe bodźce jako oddzielne. Obok na rys. 4.19b pokazano zdjęcie opuszka palca z naniesioną skalą i progiem TPDT.



Rys. 4.19. Progi rozróżnialności dwóch punktów (TPDT) na ciele człowieka (a) oraz parametr TPDT~2,5 mm dla opuszka palca zaznaczony na zdjęciu (b); taką odległość stosuje się między punktami w znakach alfabetu Braille’a

Źródło: a) opracowanie własne na podstawie (Lederman i Klatzky, 2009), b) opracowanie własne

Do dotykowej eksploracji otoczenia człowiek wykorzystuje przede wszystkim kończyny. Wyczuwa stopami rodzaj podłoża, a dłońmi potrafi ocenić kształt, wagę, temperaturę, twardość i teksturę obiektów. Lederman i Klatzky (1987) usystematyzowali sposób w jaki osoba niewidoma bada dłońmi właściwości przedmiotów o różnej wielkości i kształcie. Osoby niewidome stosują aktywny sposób poznania wymagający wykonywania odpowiednich ruchów dłoni i palców. Podstawowe schematy takich ruchów to: przemieszczanie dłoni po powierzchni przedmiotu (dla rozpoznania tekstury stosowany m.in. przy czytaniu alfabetu Braille'a), wywieranie nacisku (dla oceny twardości), stosowanie statycznego dotyku (dla oceny temperatury), trzymanie przedmiotu w dłoniach (dla oceny wagi), objęcie przedmiotu dłońmi (lub jego fragmentów) dla rozpoznania kształtu oraz koncentracja dotyku lub śledzenie charakterystycznych cech geometrycznych przedmiotu (krawędzi, narożników, wypukłości).

Na podstawie omówionych własności sensorycznych zmysłów czuciowych i sposobu eksploracji otoczenia przez ruch i dotyk wnioskujemy, że możliwa jest prezentacja informacji w postaci „obrazów” dotykowych odwzorowujących otoczenie za pomocą odpowiednich urządzeń zwanych interfejsami (Velázquez i Pissaloux, 2008). Urządzenia takie można podzielić na interfejsy haptyczne i interfejsy dotykowe.

Wśród interfejsów dotykowych wyróżnia się trzy podstawowe metody wytwarzania bodźców dotykowych (Visell, 2009):

- nacisk statyczny; wytwarzany za pomocą układów elektromechanicznych, materiałów z pamięcią kształtu oraz mikrodysze wytwarzające strumień powietrza,
- stymulacja za pomocą układów drgających; w postaci pojedynczych aktuatorów lub odpowiednich matryc aktuatorów (zob. system Bach-y-Rity); stosowany zakres częstotliwości drgań to 10÷300 Hz,
- mikroprądy wytwarzane za pomocą elektrod o niewielkiej powierzchni ($1\text{ mm}^2\div 100\text{ mm}^2$) przytwierdzanych do skóry i pobudzających zakończenia nerwowe receptorów czuciowych.

Wymienione interfejsy można określić mianem aktywnych w odróżnieniu od interfejsów pasywnych, np. ekranów dotykowych (*touch screens*), które działają na zasadzie pomiaru zmiany rezystancji lub pojemności elektrycznej powierzchni czynnej przy jej dotknięciu. Osoby niewidome w coraz większym stopniu korzystają z telefonów komórkowych wyposażonych w takie ekrany. Dzięki technologii rozpoznawania dużej liczby tzw. gestów dotykowych (*touch gestures*) osoby niewidome skutecznie obsługują w ten sposób urządzenia mobilne.

Do interfejsów haptycznych zaliczyć można podstawowe osobiste pomoce nawigacyjne wykorzystywane przez osoby niewidome, tj.:

- białą laskę, której użycie wymaga wykonania regularnych ruchów skanujących otoczenie; wyczuwane przez niewidomego drgania i położenie kątowne laski służą do wypracowania informacji o przeszkodach,
- psa przewodnika – osoba niewidoma trzyma psa za sztywną uprząż (zob. rys. 5.1) i otrzymuje odpowiedni sygnał haptyczny.

oraz interfejsy haptyczne, które mogą być stosowane do zapoznania się z nieznanym otoczeniem:

- tyflografika, technika grafiki wypukłej (tzw. reliefu) wykorzystywana do obrazowania kształtów obiektów i drukowania map dotykowych,
- urządzenia z tzw. zwrotnym sprzężeniem siły (*force feedback*), które mierzą wywieraną przez użytkownika siłę lub wprowadzane przesunięcie i wytwarzają proporcjonalny do nich opór.

Jak wspomniano wcześniej, za pomocą takich interfejsów niewidomi potrafią budować tzw. mapy wyobrażeniowe otoczenia, bardzo pomocne w znajdowaniu drogi i skutecznym poruszaniu się w nieznanym terenie (Ungar i inni, 2004).

Mapy dotykowe są typowo wykonywane techniką druku wypukłego na specjalnym papierze lub polimerze (Talukder i Jakubowski, 2003). Na rys. 4.20 zilustrowano przykładowe mapy dotykowe obrazujące zarówno mapy terenu jak i plany wnętrz budynków. Zawierają one opisy tekstowe w postaci wypukłych znaków alfanumerycznych lub alfabetu Braille’a. Mapy dotykowe są obecnie produkowane zarówno w postaci dużych plansz jak i w wersji kieszonkowej (zob. www.niewidomi.com.pl).



Rys. 4.20. Mapy dotykowe obrazujące fragmenty terenu i plany budynków

Źródło: IntouchGraphics (www.intouchgraphics.com)

Projekt osobistej mapy dotykowej pokazano na rys. 4.21. Jest to koncepcja urządzenia mocowanego opaską do ramienia osoby niewidomej. Specjalny materiał z pamięcią kształtu ma umożliwiać „dotykowe” obrazowanie wycinka terenu (w skali 1:1000). Obrazowana mapa przesuwana się w miarę pieszego

przemieszczania się użytkownika, którego położenie (oznaczone wypukłym punktem) jest utrzymywane w środku pola mapy dotykowej.



Rys. 4.21. Osobista mapa dotykowa dla osoby niewidomej – wizja artysty

Źródło: Tuvie design, www.tuvie.com, autor: Natalia Panomariewa

Wadą map dotykowych jest ograniczenie wrażeń czuciowych do przestrzeni dwuwymiarowej (Paneels i Roberts, 2010). Ograniczenia takiego nie posiadają urządzenia ze sprzężeniem zwrotnym siły, których elementy można przemieszczać z określoną liczbą stopni swobody, np. z sześcioma stopniami swobody (6DoF). Urządzenia te można odpowiednio programować dla symulacji wirtualnej przestrzeni z umieszczonymi w niej obiektami (np. fragment cyfrowej mapy trójwymiarowej miasta). Co więcej w urządzeniach tych można zadawać odpowiednie sprzężenia siły symulujące efekt określonej tekstury i twardości symulowanych obiektów. Urządzenia tej klasy mają szerokie zastosowania, m.in. w robotyce (np. roboty chirurgiczne), szkoleniu operatorów maszyn, przemyśle kosmicznym i rozrywkowym (systemy rzeczywistości wirtualnej). Urządzenia te od wielu lat znajdują zastosowania w tyflopedagogice (Yu i inni, 2001), a ostatnio dzięki ich doskonalszym konstrukcjom są również wykorzystywane jako pomoce w nauce orientacji przestrzennej i nawigacji dla niewidomych (McGookin i inni, 2010).

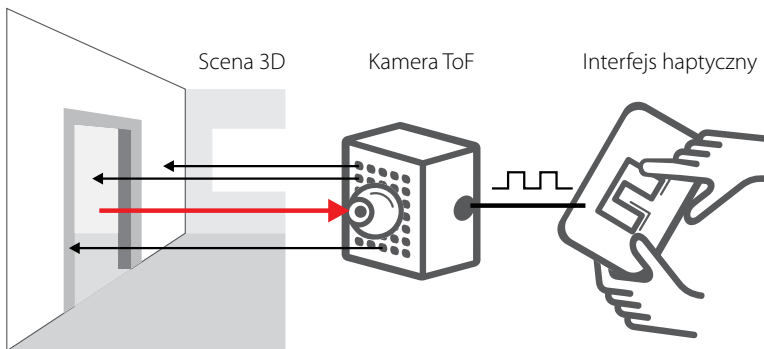
Na rys. 4.22 pokazano zdjęcie interfejsu haptycznego firmy Falcon Novit o trzech stopniach swobody (ruchy translacyjne 3DoF). Urządzenie to jest wykorzystywane w Instytucie Elektroniki Politechniki Łódzkiej w badaniach rozwojowych nad systemami orientacji przestrzennej dla osób niewidomych. Napisało oprogramowanie umożliwiające generację zadanych scen trójwymiarowych, które są następnie odtwarzane w interfejsie haptycznym. Osoby niewidome biorące udział w testach uczą się obsługi urządzenia i porównują modele scen wykonanych z kartonu ze scenami wirtualnymi wytworzonymi w urządzeniu haptycznym. W pracy (Morański i Materka, 2011) opisano badania, w których osoby niewidome uczą się nawigacji w wirtualnie wytwarzanych scenach.



Rys. 4.22. Zdjęcie stanowiska do nauki obsługi interfejsu haptycznego; z lewej model kartonowy sceny, w środku model komputerowy sceny trójwymiarowej i z prawej interfejs haptyczny symulujący zaprogramowaną wirtualną scenę

Źródło: fotografia własna

Prace zmierzają do opracowania systemu, w którym trójwymiarowe modele scen będą budowane na podstawie analizy obrazów rejestrowanych przez kamerę rekonstruującą głębokość sceny oraz odtwarzanych w osobistym interfejsie haptycznym. Koncepty takiego systemu pokazano na rys. 4.23. Przewiduje się, że interfejsy haptyczne o małych rozmiarach będą mogły być budowane z wykorzystaniem tzw. elektroaktywnych polimerów – sztucznych materiałów, które znacząco zmieniają swój kształt pod wpływem przyłożonego napięcia (Zyłka, 2004).



Rys. 4.23. Konceptja systemu obrazowania haptycznego otoczenia z wykorzystaniem kamery ToF (*time of flight*) oraz osobistego interfejsu haptycznego

Źródło: opracowanie własne na podstawie (Morański i Materka, 2011)

Skuteczność interfejsów haptycznych pokazano w spektakularnym pokazie prowadzenia samochodu przez niewidomego kierowcę. W styczniu 2011 roku niewidomy przejechał bez kolizji tor wyścigowy Daytona International Speedway, na którym rozmieszczono w przypadkowych miejscach przeszkody. Zastosowano interfejs haptyczny w rękawicach do ostrzegania o przeszkodach oraz interfejs dotykowy w pasie przymocowanym do tułowia sygnalizującym prędkość jazdy (Matacz, 2011).

Podsumowując, problematykę niewizualnych technik obrazowania otoczenia z zastosowaniem interfejsów dotykowych i haptycznych należy stwierdzić, że jest to dziedzina znajdująca się ciągle u progu pełniejszego rozwoju. Bardzo dobrze są zbadane właściwości sensoryczne zmysłów czuciowych człowieka. Ograniczeniem są jednak nadal mało przyjazne dla użytkownika techniki stosowane w interfejsach haptycznych (duże urządzenia z dużą liczbą elementów elektromechanicznych). Warunkiem rozwoju tej klasy interfejsów człowiek-maszyna jest postęp w inżynierii materiałowej, m.in. opanowanie produkcji materiałów, w których w łatwy sposób można sterować zmianą ich kształtu, np. z wykorzystaniem polimerów elektroaktywnych.

W niniejszym rozdziale omówiono podstawowe sposoby niewizualnej prezentacji otoczenia osobom niewidomym za pomocą urządzeń wytwarzających bodźce dla zmysłów słuchu i dotyku. Układy realizujące te funkcje określa się mianem odpowiednio wyświetlaczy dźwiękowych (*auditory displays*) lub wyświetlaczy dotykowych (*tactile displays*). Są one podstawowym elementem składowym systemów substytucji sensorycznej wspomagających osoby niewidome w samodzielnym poruszaniu się i orientacji przestrzennej. W następnym rozdziale monografii zawarto szeroki przegląd takich systemów, zarówno prototypów jak i systemów wdrożonych do produkcji.

Wybrane elektroniczne systemy nawigacji osobistej dla niewidomych

Naturalną pomocą nawigacyjną dla niewidomego jest osoba stałego przewodnika (opiekuna lub członka rodziny) lub pomoc tymczasowego przewodnika. Podstawową wadą takiej metody asystowanego poruszania się jest uzależnienie osoby z taką niepełnosprawnością od pomocy innych. Pomoc nie zawsze jest dostępna i fachowa (Kwapiszowie, 1990). Osoby niewidome korzystające z takiej pomocy informują o poczuciu bycia ciężarem dla innych.

Stosowaną od wieków przez niewidomych osobistą pomocą w poruszaniu się był kij lub laska. Dzisiaj jest to biała laska stosowana od lat 30-tych XX wieku (rys. 5.1). Długość białej laski jest dopasowana do wzrostu osoby i powinna sięgać splotu słonecznego. Współcześnie produkowane białe laski są wykonywane z aluminium, grafitu lub innych lekkich materiałów. Istnieją wersje składane. Czas treningu potrzebny do prawidłowego posługiwania się białą laską wynosi ok. 100 godzin. Biała laska jest pomocna w wykrywaniu przeszkód w bliskim otoczeniu na poziomie stóp (w tym o nieciągłości podłoża: uskoki, krawężniki), jest tania, lekka i informuje inne osoby o niepełnosprawności wzrokowej jej posiadacza. Podstawowymi wadami białej laski to wymóg ręcznego przeszukiwania otoczenia i kontaktu mechanicznego z przeszkodą, ograniczony zakres ostrzegania o przeszkodach (tylko na poziomie stóp) i ograniczony promień zasięgu (ok. 1,5 m).

Pies przewodnik jest uważany za bardzo skuteczną pomoc w poruszaniu się niewidomych (rys. 5.1). Niestety z uwagi duży koszt szkolenia psa (ok. 60 tys. zł) i konieczność ponoszenia regularnych kosztów jego utrzymania tylko 1% niewidomych korzysta z pomocy tych zwierząt. Należy podkreślić, że zadaniem psa przewodnika nie jest nawigowanie niewidomego, a wytresowane zachowanie w kon-

kretnych sytuacjach, np. zatrzymanie przed schodami lub przejściem dla pieszych. Niektórzy niewidomi poruszający się z psem przewodnikiem nadal zgłaszają trudności z dostępem do miejsc publicznych, m.in. sklepów i centrów handlowych.



Rys. 5.1. Osoby niewidome korzystające z podstawowych pomocy w poruszaniu się (*primary travel aids*), z lewej z psem przewodnikiem, z prawej z białą laską

Źródło: fotografie własne

Jak wspomniano we wstępie książki, pierwsze elektroniczne urządzenie wspomagające niewidomych w poruszaniu się zbudował w 1898 r. polski okulista Kazimierz Noiszewski. Wykorzystał on fotokomórki selenowe do przetwarzania światła na dźwięki. Jego prototypowe urządzenie było niestety niepraktyczne w użyciu z uwagi na dużą wagę i pobór prądu.

Od czasu pionierskich prac Noiszewskiego powstała duża liczba osobistych urządzeń i systemów, w których wykorzystano różne metody wspomagania osób niewidomych w samodzielnym poruszaniu się i nawigacji. Można je podzielić na trzy zasadnicze rodzaje (Strumiłło, 2010).

1. Elektroniczne detektory przeszkód.
2. Elektroniczne systemy niewizualnej prezentacji obrazów otoczenia.
3. Systemy nawigacji osobistej o architekturze rozproszonej.

Podstawowe rozwiązania techniczne stosowane w tych urządzeniach i systemach omówiono w poprzednich rozdziałach pracy. W niniejszym rozdziale skupiono się na zwięzłym, ale możliwie szerokim przeglądzie urządzeń tej klasy, ich podstawowych parametrach i własnościach użytkowych. Bardziej szczegółowo

omówiono system do sonifikacji otoczenia, w którym zastosowano stereowizję, techniki analizy obrazów oraz charakterystyki akustyczne HRTF do wytworzenia przestrzennych „obrazów” dźwiękowych sceny. Prototyp tego systemu został zbudowany w Instytucie Elektroniki Politechniki Łódzkiej (IE PŁ, 2010).

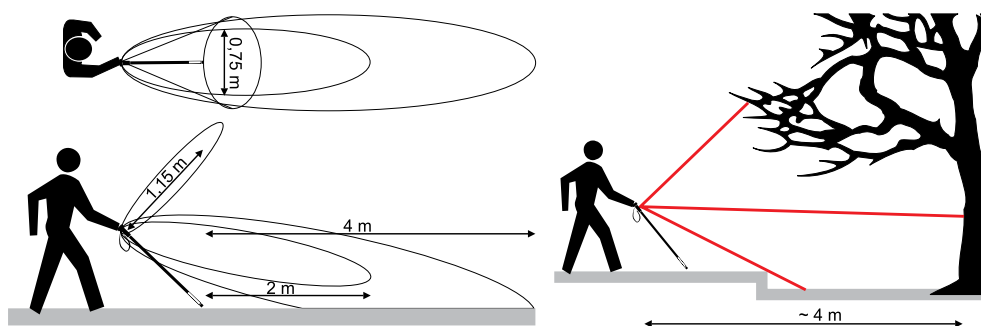
Więcej informacji dotyczących technologii wspomagających osoby niewidome można znaleźć w obszernej pracy zbiorowej (Hersh i Johnson, 2008) oraz w pracach przeglądowych (Roentgen, 2008) i (Ball, 2008). Cennym źródłem informacji w języku polskim na temat ww technologii jest czasopismo Tyfloświat (www.new.tyfloswiat.pl), strona internetowa Uniwersytetu Gdańskiego (<http://bariery.ug.edu.pl/orientacja.html>), artykuł Chałcampowicza (2009) oraz praca Buczkowskiego i Radeckiego (2010).

5.1. Elektroniczne detektory przeszkód

W elektronicznych detektorach przeszkód (*obstacle detectors*) są stosowane aktywne metody estymacji odległości z wykorzystaniem dalmierzy ultradźwiękowych lub laserowych. Użytkownik jest informowany o odległości do przeszkody za pomocą sygnałów dźwiękowych lub dotykowych. Są to urządzenia o małych rozmiarach, które mogą być trzymane w dłoni lub montowane na białej lasce (rys. 5.2).

Pierwszym szerzej rozpowszechnionym detektorem przeszkód dla niewidomych był Kay Sonic Torch (w polskim tłumaczeniu „Dźwiękowa latarka Kaya”) zbudowany przez Amerykanina Leslie Kaya w 1964 roku. Odległość do przeszkody jest w tym urządzeniu przetwarzana na odpowiednie sygnały dźwiękowe. Obecnie jest dostępna miniaturowa wersja (rys. 5.3) tego urządzenia o nazwie ‘K’ Sonar-Cane (Kay, 2011). Detektorem wykorzystującym ultradźwięki jest UltraCane (Hoyle i Waters, 2008) opracowany w Uniwersytecie Leeds w Wielkiej Brytanii. Nadajniki i odbiorniki ultradźwiękowe są wbudowane w rękojeść białej laski o specjalnej konstrukcji (rys. 5.3). Zakres czułości detektorów obejmuje wycinek przestrzeni o szerokości ok. 75 cm i wysokości chroniącej głowę użytkownika (rys. 5.2). Intensywność wibracji informuje o odległości do przeszkody.

Ultradźwięki wykorzystywano także w australijskim detektorze przeszkód o nazwie MiniGuide (GDP Research, 2011), pierwotnie produkowanym w dwóch wersjach Audio MiniGuide (ostrzeganie dźwiękowe) i Tactile MiniGuide (ostrzeganie wibracjami). Obecnie jest oferowana wersja, w której użytkownik może przełączać się pomiędzy sygnalizacją dźwiękową lub dotykową (rys. 5.3). Wadą



Rys. 5.2. Sposób skanowania otoczenia w elektronicznych detektorach przeszkód dla osób niewidomych; z lewej UltraCane i z prawej LaserCane

Źródło: opracowanie własne

ultradźwiękowych detektorów przeszkód jest niezbyt duża dokładność mierzonej odległości oraz duża rozbieżność wiązki ultradźwiękowej (ok. 40°).

Bardziej dokładny pomiar odległości za pomocą wiązki o małej rozbieżności kątowej jest możliwy w dalmierzach laserowych. Stosuje się w nich wiązki laserowe o bardzo małym natężeniu uważanym za bezpieczne dla oczu. Pierwszym oferowanym dla niewidomych laserowym detektorem przeszkód był amerykański LaserCane (Malvern i Nazir, 1973). W ostatnim modelu detektora przeszkód produkowanym od 1973 roku o symbolu N-2000 zastosowano trzy wiązki laserowe (rys. 5.2). Są one skierowane pod różnymi kątami dla ochrony twarzy (wiązka górna), detekcji przeszkód znajdujących się na wprost w odległości $2\div 4$ m (wiązka środkowa) i wykrywania nieciągłości podłoża (wiązka dolna).

Europejskim laserowym detektorem przeszkód opracowanym przez Rene Farcy (2006) z Université Paris Sud jest urządzenie TeleTact. Najnowszą wersję tego urządzenia TeleTact II pokazano na rys.5.3. Jest to detektor z pojedynczą wiązką lasera, którą użytkownik powinien omiatać otoczenie w sposób podobny jak białą laską. Zastosowano 32. nutowy sposób sygnalizacji odległości, w którym wysokość tonu wzrasta wraz ze zmniejszającą się odległością wykrywanej przeszkody. Jest dostępna wersja urządzenia, w której różne odległości są sygnalizowane 8 różnymi rodzajami wibracji. TeleTact ma wbudowany dodatkowy nadajnik i odbiornik podczerwieni ostrzegający przed przeszkodami przezroczystymi dla światła lasera lub przeszkodami pochłaniającymi światło lasera. Producent sprzedaje urządzenie pod warunkiem zaliczenia 50-godzinowego szkolenia przez przyszłego użytkownika urządzenia.

W tabeli 5.1 zebrano podstawowe właściwości omówionych detektorów przeszkód a na rys. 5.3 pokazano ich zdjęcia.

Tabela. 5.1. Podstawowe parametry ultradźwiękowych i laserowych detektorów przeszkód dla niewidomych

Nazwa (strona internetowa producenta)	Zasada działania	Zasięg	Zalety	Wady
Bat 'K' Sonar-Cane (www.batforblind.co.nz)	Dalmierz ultradźwiękowy. Sygnalizacja odległości za pomocą dźwięków o różnej wysokości tonów.	Przełączany: 2 m, 5 m.	Mały rozmiar (12x2x9 cm). Barwa dźwięku odwzorowuje właściwości akustyczne elementów otoczenia.	Konieczne używanie słuchawek (wg producenta nie tłumiących dźwięków z otoczenia), mała rozdzielczość kątowa detekcji przeszkód.
UltraCane* (www.ultracane.com)	Dwa dalmierze ultradźwiękowe. Sygnalizacja odległości za pomocą kilku elementów wibrujących zamontowanych w rękojeści laski.	Dalmierz dla kierunku na wprost: 4 m. Dalmierz na wysokości twarzy: 1 m.	Ostrzega o przeszkodach na wprost w pasie o szerokości ok. 75 cm oraz na wysokości twarzy. Nie jest wymagany ruch skanujący laski.	Mała rozróżnialność informacji o odległości i kierunku. Masywna budowa.
Miniguide (www.gdp-research.com.au)	Dalmierz ultradźwiękowy. Sygnalizacja odległości za pomocą dźwięków lub wibracji.	Przełączane zakresy: 0,5, 1, 2, 4, 8 m.	Bardzo mały rozmiar: 80x38x23 mm. Długi czas pracy na baterii.	Mała rozdzielczość kątowa detekcji przeszkód, wymaga ręcznego skanowania otoczenia.
LaserCane**	Trzy dalmierze laserowe. Jednoczesna sygnalizacja dźwiękiem i wibracjami.	1–4 m.	Duża dokładność pomiaru odległości, szeroki zakres detekcji przeszkód w kierunku pionowym.	Wymaga ręcznego skanowania otoczenia; praca może być zakłócana w warunkach silnego nasłonecznienia.
Teletact (www.bel108.it)	Dalmierz laserowy i dalmierz w podczerwieni. Wersje urządzenia z sygnalizacją odległości za pomocą dźwięków lub wibracji.	10 cm–15 m.	Bardzo duża dokładność pomiaru odległości (ok. 1%); ostrzega przed przeszkodami szklanymi lub pochłaniającymi światło lasera	Wymaga ręcznego skanowania otoczenia; praca może być zakłócana w warunkach silnego nasłonecznienia.

*) Wznowiono produkcję w 2011 r. po 3-letniej przerwie; **) Zaprzestano produkcji w 2007 r.

Źródło: opracowanie własne na podstawie zawartych w tabeli adresów stron internetowych



'K' Sonar-Cane



UltraCane



MiniGuide



TeleTact

Rys. 5.3. Zdjęcia elektronicznych detektorów przeszkód dla osób niewidomych
Źródło: fotografie ze stron internetowych wg wykazu w tabeli 5.1 (za zgodą producentów)

5.2. Elektroniczne systemy niewizualnej prezentacji obrazów otoczenia

W elektronicznych detektorach przeszkód decyzje o kierunku detekcji przeszkód dla kątów elewacji i azymutu pozostawia się użytkownikowi. Pewnym odstępstwem od tej reguły są tu urządzenia UltraCane i LaserCane, w których zakres kątowy wykrywania przeszkód został powiększony przez zastosowanie więcej niż jednego dalmierza. W detektorach tych wymaga się od użytkownika wykonywania regularnych ruchów skanujących otoczenie, angażujących jedną rękę.

Podstawowym założeniem systemów niewizualnej prezentacji obrazów otoczenia jest automatyczne pozyskiwanie informacji (za pomocą metod aktywnych lub pasywnych) o strukturze geometrycznej sceny trójwymiarowej, tak aby użytkownik był uwolniony od wymogu cyklicznego wykonywania ruchów ska-

nujących otoczenie. Systemy takie określamy mianem systemów obrazowania otoczenia, gdyż pozyskiwana informacja jest dana w postaci sekwencji jednej lub kilku jednocześnie wyznaczanych odwzorowań dwuwymiarowych, np. obrazów rejestrowanych przez kamery stereowizyjne i wyznaczanych map głębi. Na podstawie takich obrazów, stosując odpowiednio zaawansowane algorytmy analizy i rozpoznawania obrazów (Ma i inni, 2003), można wyznaczać następujące parametry obiektów sceny:

- wielkość obiektu,
- kształt,
- cechy optyczne: barwa, tekstura, cienie (wynikające m.in. ze sposobu oświetlenia sceny),

oraz parametry związane z położeniem obiektu względem układu rejestracji obrazów, tj.:

- odległość,
- kierunek,
- orientacja przestrzenna,
- kierunek przemieszczania się obiektu.

W rozważanym zastosowaniu systemów obrazowania tylko niektóre parametry obiektów są istotne dla bezpiecznego poruszania się osoby niesprawnej wzrokowo. Osoby niewidome zwracają m.in. uwagę, że nie sam kształt przeszkody jest istotny, a jej szerokość (dla jej bezpiecznego ominięcia). Przywiązują z kolei wagę do obiektów nie znajdujących na ich ścieżce marszu, które mogą służyć za punkty lub linie nawigacyjne. Zatem sposób niewizualnej prezentacji otoczenia dla osób niewidomych nie jest typowym zadaniem modelowania scen trójwymiarowych stosowanym m.in. w grafice komputerowej i fotogrametrii¹ (Kurczyński i Preuss, 2002). W systemach takich zachodzi konieczność odpowiedniej selekcji danych o scenie z uwagi na mniejszą przepływność informacyjną zmysłów niewizualnych.

Większość systemów niewizualnej prezentacji obrazów otoczenia to opracowania prototypowe, tylko niektóre z nich zostały wdrożone do produkcji. Krótką charakterystykę tej klasy systemów podano w tabeli 5.2, a na rys. 5.4 pokazano zdjęcia wybranych z nich.

¹ Dziedzina obrazowania fotograficznego służąca do wyznaczania parametrów geometrycznych, optycznych i wzajemnego rozmieszczenia obiektów w przestrzeni trójwymiarowej; stosowana m.in. do fotorealistycznej rekonstrukcji scen oraz w kartografii, geodezji i planowaniu przestrzennym.

Tabela 5.2. Podstawowe parametry i zasada działania urządzeń do niewizualnej prezentacji obrazów otoczenia dla niewidomych

Nazwa (rok opracowania, źródło)	Zasada działania	Zalety	Wady
PREZENTACJA DOTYKOWA OBRAZÓW			
Forehead Retina System (FSR) – 1999 (FRS, 2011)	Detektor: Kamera cyfrowa wbudowana w oprawkę okularów. Interfejs użytkownika: Bodźce dotykowe. Matryca ~500 punktów elektrostymulacyjnych rozmieszczonych na czole. Obraz z kamery podlega przetwarzaniu dla podkreślenia krawędzi obiektów i ich ruchu w scenie.	Nie angażuje zmysłu słuchu, mały rozmiar.	Niepotwierdzona użyteczność urządzenia. Noszenie szerokiej opaski na czole może być kłopotliwe. Nie odwzorowuje głębi.
Tyflos – 2004 (Bourbakis, 2008)	Detektor: Dwie kamery cyfrowe (bez wyznaczania głębi) i skaner laserowy podłączone do komputera przenośnego. Interfejs użytkownika: Bodźce dotykowe, matryca 32x32 stymulatorów dotykowych; zastosowano miniaturowe silniki elektryczne, których częstotliwość wibracji wzrasta ze zmniejszającą się odległością do przeszkody.	Oryginalne połączenie obrazowania wizyjnego i aktywnej metody estymacji głębi. Prowadzone testy metod odwzorowania głębi za pomocą matrycy stymulatorów dotykowych.	Projekt w fazie prototypu.
BrainPort – 2010 (pierwszy prototyp o nazwie Tongue Display Unit – 2005) (BrainPort, 2011)	Detektor: Kamera cyfrowa z obiektywem o regulowanej ogniskowej wbudowana w oprawkę okularów z zoomem i regulacją kontrastu. Interfejs użytkownika: Umieszczana na języku matryca 100÷600 elektrod (zależnie od wersji) wytwarzających małe napięcie elektryczne.	Innowacyjna koncepcja wykorzystania receptorów dotykowych języka do obrazowania otoczenia. Pierwsze testy wykazały, że 2÷10 godz. trening	Urządzenie może być niewygodne w użyciu, utrudnia lub uniemożliwia mówienie. Brak udokumentowanych testów działania systemu.

Tabela 5.2. cd.

Nazwa (rok opracowania, źródło)	Zasada działania	Zalety	Wady
	Matryca o powierzchni $\sim 9 \text{ cm}^2$ wytwarza elektryczny obraz, w którym wartość lokalnego napięcia jest proporcjonalna do lokalnej jasności obrazu.	wystarcza do korzystania z urządzenia. Wzmocnienie napięcia elektrycznego wytwarzanego przez elektrody może być regulowane przez użytkownika.	
PREZENTACJA DŹWIĘKOWA OBRAZÓW			
Sonicguide – 1974 KASPA – 1993 ¹	Detektor: Szerokokątny nadajnik ultradźwiękowy i dwa odbiorniki; montowane w oprawce okularów (Sonicguide) i w opasce na czoło (KASPA). Prezentacja: Stereofonia, wysokość tonów odwzorowuje odległość przeszkody, a efekt stereofoniczny kierunek położenia przeszkody.	Pierwsze zastosowanie trzech cech dźwięku (wysokość tonu, barwa, kierunek) do obrazowania otoczenia. Barwa dźwięku odwzorowuje właściwości akustyczne elementów otoczenia. Zastosowano miniaturowe głośniczki zamiast słuchawek aby zmniejszyć tłumienie dźwięków otoczenia.	Odwzorowuje przekrój sceny na poziomie głowy (schylanie głowy konieczne do wykrywania przeszkód pod nogami). Bezpośrednie przetwarzanie ultradźwięków na sygnały akustyczne.
Sonic Pathfinder – 1984 (Heyes, 1984)	Detektor: Pięć przetworników ultradźwiękowych (dwa nadajniki, trzy odbiorniki) przytwierdzonych do opaski umieszczanej na czołe.	Pierwsze urządzenie, w którym zastosowano algorytmy adaptacyjnej	Odwzorowuje przekrój sceny na poziomie głowy (schylanie głowy konieczne

¹ Obecnie brak dostępnych oryginalnych źródeł na temat systemów Sonicguide i KASPA.

Tabela 5.2. cd. Podstawowe parametry i zasada działania urządzeń do niewizualnej prezentacji obrazów otoczenia dla niewidomych

Nazwa (rok opracowania, źródło)	Zasada działania	Zalety	Wady
PREZENTACJA DŹWIĘKOWA OBRAZÓW cd.			
(La Grow, 1999)	Interfejs użytkownika: Stereofonia, odległość kodowana wysokością i szybkością powtarzania tonów muzycznych odtwarzanych przez miniaturowe głośniczki nie stykające się bezpośrednio z uchem.	metody selekcji obiektów do sonifikacji uwzględniającej prędkość poruszania się niewidomego i lokalizację przeszkód. Posiada tryb detekcji linii orientacyjnej (np. ściany) i prezentacji bieżącej odległości do niej.	do wykrywania przeszkód pod nogami).
Navbelt – 1990 (Shoval i inni, 1998)	Detektor: 16 dalmierzy ultradźwiękowych (8 skierowanych w dół i 8 w górę) przymocowanych do pasa. Szerokość przemiatania otoczenia 170°. Interfejs użytkownika: Stereofonia, dwa tryby prezentacji dźwiękowej: tryb obrazowy – sekwencja tonów słyszanych od lewej do prawej strony odwzorowuje cykliczne przemiatanie otoczenia, wysokość i głośność tonu oddaje odległość do przeszkody, tryb nawigacyjny – cykliczny ton wskazujący kierunek wolnej przestrzeni.	Duża liczba dalmierzy, szeroki kąt przemiatania otoczenia w kierunku poziomym i pionowym. Dwa tryby pracy. Zastosowano algorytmy automatycznej regulacji szybkości przemiatania otoczenia zależnie od liczby przeszkód i szybkości marszu.	Duże i ciężkie urządzenie. Wymagany trening użytkownika. W trybie obrazowym duża liczba regularnie powtarzanych tonów.

Tabela 5.2. cd.

Nazwa (rok opracowania, źródło)	Zasada działania	Zalety	Wady
The vOlCe – 1992 (Meijer, 1992)	Detektor: Kamera cyfrowa wbudowana w oprawkę okularów. Interfejs użytkownika: Rejestrowane obrazy podlegają cyklicznemu przemiataniu kolumn obrazu (z okresem 1 s). Głośność wytwarzanego tonu odwzorowuje średnią jasność kolumny obrazu, a częstotliwość tonu jest tym wyższa im wyżej w obrazie znajduje się przeszkoda.	Bardzo bogaty sygnał dźwiękowy – możliwość odtworzenia wielu szczegółów sceny. Ergonomiczna budowa.	Wymagane podłączenie kamery do komputera (łączem USB). Z uwagi na mało intuicyjny sposób prezentacji dźwiękowej konieczny wielomiesięczny trening. Nie odwzorowuje głębi.
Espacio Acustico Virtual (EAV) – 1998 (Gonzales-Mora i inni, 1999)	Detektor: Stereoskopowy układ kamer w oprawie okularów wyznacza gęstą mapę głębi. Interfejs użytkownika: Stereofonia z zastosowaniem indywidualnych funkcji HRTF dla uzyskania efektu dźwięku przestrzennego.	Przestrzenny sygnał dźwiękowy dostarcza informacji o obrazie sceny jak i odległości elementów sceny.	Brak strukturalnej interpretacji elementów sceny. Duża liczba jednocześnie generowanych dźwięków
Naviton – 2010 (Bujacz i inni, 2011)	Detektor: Stereoskopowy układ kamer w oprawie okularów wyznacza gęstą mapę głębi. Interfejs użytkownika: Stereofonia z zastosowaniem indywidualnych funkcji HRTF dla uzyskania efektu dźwięku przestrzennego. Cykliczne skanowanie głębi sceny i kolejna sonifikacja wykrytych obiektów. Moduł wytwarzania dźwięku zintegrowany z układem stereofonicznym w jednej obudowie.	Zaawansowany algorytm budowania trójwymiarowego modelu sceny zawierającego obiekty klasyfikowane jako przeszkody i płaszczyzny. Każdy wydzielony ze sceny obiekt jest oddzielnie sonifikowany.	Wymagana duża moc obliczeniowa wynikająca z wyznaczania gęstych map głębi, budowania modelu sceny i wytwarzania dźwięków.

Źródło: opracowanie własne na podstawie zawartych w tabeli adresów stron internetowych



Sonic Pathfinder



The vOICe



Espacio Acustico Virtual



BrainPort

Rys. 5.4. Fotografie systemów do niewizualnej prezentacji obrazów otoczenia dla osób niewidomych

Źródło: fotografie ze stron internetowych wg wykazu w tabeli 5.2 (za zgodą autorów)

W wymienionych systemach zastosowano selekcję informacji do prezentacji, wynika ona z ograniczonej przepływności danych niewizualnych. W systemach tych zastosowano następujące metody selekcji danych:

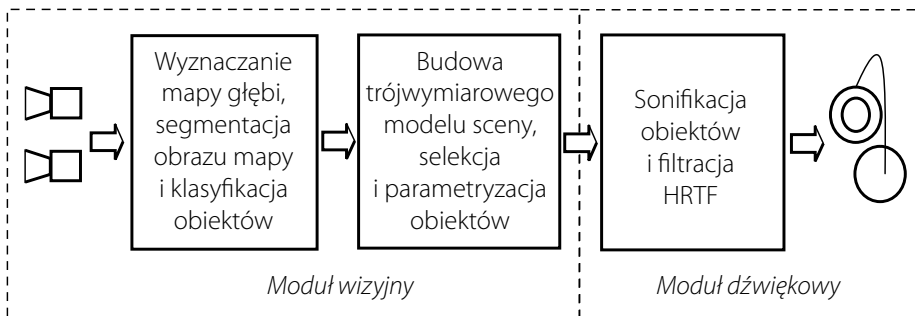
- ograniczenie rozdzielczości prezentowanych obrazów (w przypadku interfejsów dotykowych i systemu EAV),
- skanowanie każdego obrazu w cyklach znacznie dłuższych niż w standardzie TV (system vOICe i NavBelt),
- hierarchiczna selekcja obiektów do sonifikacji w zależności od ich położenia względem ścieżki marszu (Sonic Pathfinder),
- strategia dźwiękowej prezentacji wolnej przestrzeni, a nie sonifikacji przeszkód (NavBelt); jest to podejście krytykowane przez niektórych badaczy,
- segmentacja obrazów głębi, klasyfikacja i selekcja obiektów, składających się na uproszczony model sceny (Naviton).

W omówionych systemach (z wyjątkiem the vOICe i BrainPort, w których informacja o głębi nie jest bezpośrednio odwzorowywana), zastosowano co najmniej jedną cechę bodźca niewizualnego do odwzorowania głębi. Dla prezentacji dotykowej stosowano wzrastającą amplitudę bodźca m.in. w Tyflos, a dla prezentacji audio cechy dźwięku jakie wykorzystano do odwzorowania głębi to: wysokość tonu, głośność i szybkość powtarzania tonów (stosowano jedną lub dwie cechy jednocześnie). Autorzy systemów the vOICe i BrainPort, w których informacja o głębi nie jest bezpośrednio odtwarzana dokumentują, że osoby niewidome w czasie marszu potrafiły wypracować informacje o głębi na podstawie szybkości powiększającego się pola obszarów odwzorowujących obiekty (Meijer, 1992). Podkreślali jednak konieczność odpowiednio długich treningów (np. kilka miesięcy w przypadku the vOICe).

Projekt Naviton

W ramach projektu rozwojowego MNiSW (IE PŁ, 2010) w Instytucie Elektroniki Politechniki Łódzkiej zaprojektowano i wykonano modele oraz prototypy systemów do niewizualnego obrazowania otoczenia dla osób niewidomych. Badania te poprzedzono ankietami z grupą 20 osób niewidomych, które wskazały na główne problemy związane z samodzielnym poruszaniem się w nieznanym otoczeniu (wyniki ankiet zamieszczono w rozdz. 1).

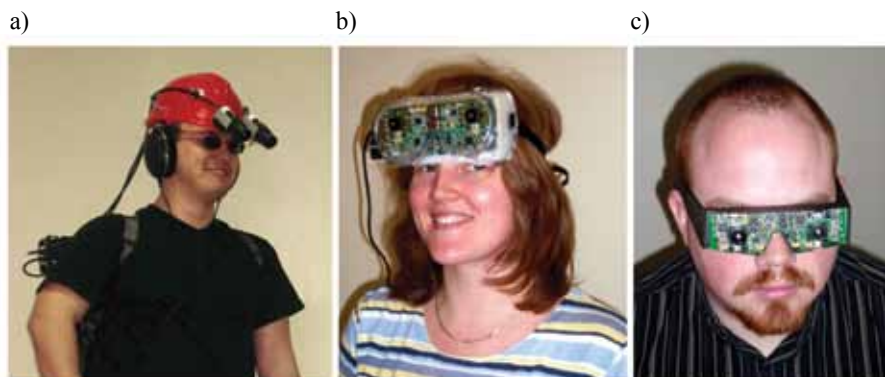
Zastosowano stereowizję do estymacji głębi, oraz algorytmy analizy i segmentacji obrazów do budowania trójwymiarowych modeli sceny oraz przestrzenną sonifikację wydzielonych obiektów sceny. Ogólny schemat blokowy opracowanych algorytmów pokazano na rys. 5.5.



Rys. 5.5. Uproszczony schemat blokowy algorytmów do prezentacji dźwiękowej otoczenia

Źródło: opracowanie własne

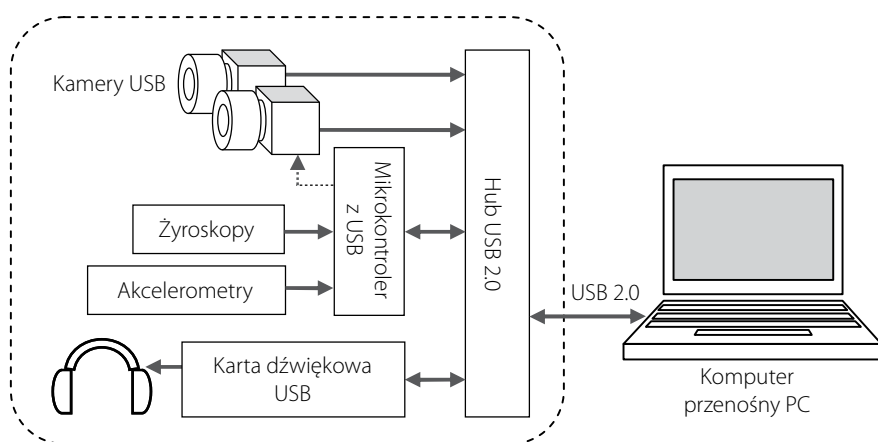
Zaprojektowano i wykonano kilka wersji modułów elektronicznych do akwizycji obrazów i wytwarzania dźwięków przestrzennych. Kolejne wersje układów sprzętowych systemu pokazano na rys. 5.6.



Rys. 5.6. Wersje modelowe i prototypowe układów do dźwiękowego obrazowania otoczenia wykonanych w Instytucie Elektroniki PŁ; (a) kamery TV zamontowane na kasku i podłączone do komputera przenośnego (2008 r.), (b) pierwsza wersja modułu elektronicznego (2010 r.) oraz (c) druga wersja modułu elektronicznego z 2011 r. (ze zdjętą pokrywą „okularów”)

Źródło: fotografie własne

Schemat blokowy układów elektronicznych montowanych na płycie drukowanej pokazanej na rys. 5.8 uwidoczniono w postaci schematu blokowego na rys. 5.7. Zastosowane czujniki przyspieszeń służą obecnie do celów badawczych (Pełczyński i inni, 2009).



Rys. 5.7. Schemat blokowy modułu rejestracji obrazów stereowizyjnych i wytwarzania dźwięków komunikujący się z przenośnym komputerem PC za pomocą interfejsu USB 2.0

Źródło: opracowanie własne

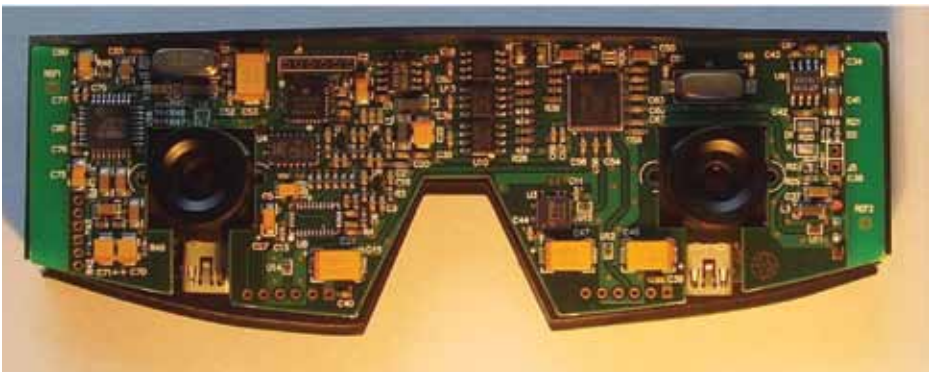
W systemie pokazanym na zdjęciu z rys. 5.8 zastosowano następujące układy elektroniczne:

- dwie synchronizowane programowo kamery cyfrowe o rozdzielczości 752×480 punktów wyposażone w interfejs USB,
- trójosiowy żyroskop elektroniczny o zakresie pomiarowym: $\pm 1600 \pm 6$ st./s,
- akcelerometr liniowy trójosiowy o zakresie pomiarowym $\pm 6g$,
- mikrokontroler,
- scalony moduł dźwiękowy (16-bitowe próbkowanie sygnału z częstotliwością 48 kHz).

Opracowany układ wykonuje następujące funkcje:

- synchroniczna rejestracja pary obrazów stereowizyjnych o rozdzielczości 752×480 z programowaną częstością rejestrowanych obrazów $1 \div 30$ na sekundę,
- rejestracja sygnałów z czujników inercyjnych synchronizowana z odczytami obrazów,
- transmisja on-line rejestrowanych danych do komputera przenośnego,
- wytwarzanie dźwięku stereofonicznego przesyłanego z komputera przenośnego.

Do realizacji ww funkcji układu opracowano odpowiednie biblioteki programów do synchronicznej rejestracji obrazów stereoskopowych i odczytu sygnałów z czujników inercyjnych. Wykonany układ elektroniczny został umieszczony w obudowie z tworzywa sztucznego ABS z przednią przezroczystą ścianką wykonaną z pleksiglasu, zaciemnioną w obszarach poza obiektywami kamer. Urządzenie przypomina masywne okulary i waży ok. 200 g. Zdjęcie płytki drukowanej z zamontowanymi układami elektronicznymi i kamerami pokazano na rys. 5.8.



Rys. 5.8. Widok płytki drukowanej pełniącej funkcje personalnego układu obrazowania otoczenia, wyposażonego w kamery stereoskopowe, czujniki przyspieszeń i kartę dźwiękową

Źródło: fotografia własna

Zgodnie ze schematem blokowym algorytmów pokazanym na rys. 5.5, w module wizyjnym systemu jest wyznaczana mapa głębi, a następnie jest przeprowadzana segmentacja, klasyfikacja i selekcja obiektów do sonifikacji. Zastosowane metody przetwarzania i analizy obrazów omówiono w rozdz. 2. Na rys. 5.9 i 5.10 pokazano przykładowe wyniki etapu przetwarzania obrazów w module wizyjnym systemu.



Rys. 5.9. Obraz sceny trójwymiarowej (z lewej) i model komputerowy utworzony do celów prezentacji dźwiękowej (z prawej)

Źródło: opracowanie własne

Budowany model sceny trójwymiarowej składa się z dwóch klas obiektów:

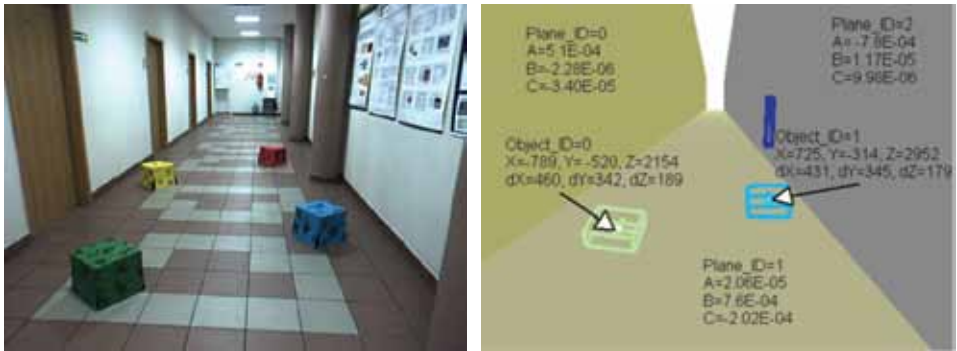
- płaszczyzn, tj. obiektów, które stanowią linie nawigacyjne i nie są klasyfikowane jako przeszkody²,
- brył z zamkniętą objętością (prostopadłościanów), tj. obiektów, które nie spełniają równań płaszczyzn i są klasyfikowane jako przeszkody.

Przyjęty model sceny składający się z płaszczyzn i prostopadłościanów ma szereg ograniczeń, ale jest adekwatny dla większości scen w terenie miasta i wewnątrz budynków (Skulimowski i Strumiłło, 2007).

W zastosowanym schemacie prezentacji dźwiękowej obie klasy obiektów podlegają sonifikacji. Schemat sonifikacji jest zależny od parametrów geometrycznych obiektów i ich lokalizacji w scenie. Na rys. 5.10. pokazano obraz sceny i jej model komputerowy, w którym zaznaczono parametry wyznaczone dla każdego z obiektów.

Każdemu wydzielonemu obiektowi sceny przydzielono numer identyfikacyjny w określonej klasie oraz zestaw parametrów. Dla płaszczyzn są to współczynniki A , B , i C równań płaszczyzny danych w postaci: $AX + BY + CZ + 1 = 0$, gdzie

² Płaszczyzny są interpretowane jako przeszkody jeżeli znajdują się w założonym *a priori* polu analizy sceny i mają określoną orientację przestrzenną.



Rys. 5.10. Obraz sceny trójwymiarowej (z lewej) oraz jej model komputerowy z zaznaczonymi parametrami wydzielonych obiektów sceny (wartości parametrów podano w milimetrach)
Źródło: opracowanie własne

X, Y, Z są współrzędnymi układu kamer. Parametry reprezentujące prostopadłościany (klasyfikowane jako przeszkody) stanowią: współrzędne środka geometrycznego i -tego prostopadłościanu (X_i, Y_i, Z_i) i jego rozmiary (dX_i, dY_i, dZ_i).

Rozważano różne schematy sonifikacji wydzielonych obiektów scen. W tabeli 5.3 scharakteryzowano zastosowany schemat sonifikacji.

Tabela. 5.3. Zastosowane schematy sonifikacji cech obiektu na cechy dźwięku

Parametry obiektów (geometryczne, położenie i ruch)	Parametry dźwięków zastosowane do sonifikacji obiektów
Lokalizacja obiektu:	
• Odległość	• Opóźnienie dźwięku, echo
• Kierunek	• Głośność
	• HRTF
• Ruch	• HRTF (przełączanie filtrów odpowiadających różnym kierunkom)
Cechy obiektu:	
• Wielkość	• Częstotliwość podstawowa
• Typ obiektu - - - - - (np. przeszkoda, linia nawigacyjna)	• Barwa dźwięku (rodzaj instrumentu), szerokość pasma
—————>	Naturalne skojarzenie cechy obiektu z cechą dźwięku
—————>	Instynktowne skojarzenie psychoakustyczne
- - - - ->	Wypracowanie skojarzenia wymagające treningu

Źródło: opracowanie własne

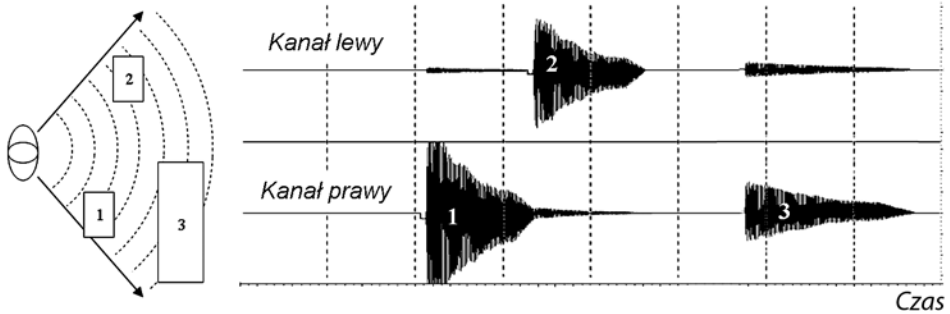
Opracowano program komputerowy, który dla symulowanych obiektów w scenie trójwymiarowej umożliwiał zastosowanie różnych schematów sonifikacji obiektów o zadawanych i zmienianych przez użytkownika parametrach i położeniu w scenie (Bujacz, 2010). Udostępniono możliwość jednoczesnej sonifikacji do 4 obiektów z zastosowaniem tonów krtaniowych (Pełczyński i inni, 2006), wykorzystaniem dźwięków z biblioteki MIDI (*Musical Instrument Digital Interface*) oraz syntezowanych tonów harmoniczných o różnych parametrach obwiedni amplitudy ADSR (*attack decay sustain release*). Program poddano testom przez grupę 10 osób niewidomych. Z uczestnikami testu przeprowadzono ankietę, której wyniki można streścić następująco:

- spośród wielu syntezowanych i naturalnych tonów zastosowanych do sonifikacji obiektów wybrano pełne tony muzyczne,
- wskazano głośność dźwięków jako parametr najlepiej oddający odległość obiektu, połowa osób wskazywała również na wysokość tonu jako cechę dźwięku dobrze oddającą odległość obiektu,
- do odwzorowania wielkości obiektu większość osób wybrała czas trwania dźwięku a połowa wskazała na wysokość tonu jako naturalną cechę dźwięku dobrze oddającą wielkość obiektu,
- jednoczesna sonifikacja kilku obiektów była oceniona jako zbyt bogata w dźwięki i wymagająca dużego skupienia uwagi,
- większość osób preferowała schemat obrazowania dźwiękowego, w którym obiekty były sonifikowane w sekwencji czasowej uporządkowanej wg rosnącej odległości.

Wskazany w ostatnim punkcie schemat sonifikacji zastosowano w prototypowym systemie do dźwiękowego obrazowania otoczenia. Nazwano go skanowaniem głębi. Jest to metoda sonifikacji, w której hipotetyczna płaszczyzna (prostopadła do osi optycznej kamery) oddala się od obserwatora i po napotkaniu obiektu włącza odpowiednie strumienie dźwiękowe (Bujacz, 2010).

Do sonifikacji obiektów zastosowano dźwięki z biblioteki MIDI zakłócone białym szumem. Szum zapewnia odpowiednio szerokie pasmo dźwięków konieczne dla odpowiednio dokładnej ich lokalizacji przestrzennej wynikającej z filtracji HRTF (zob. rozdz. 4). Dźwięki pełnych tonów w tzw. diatonicznej skali (oktawy od 2. do 4.) zapisano jako 5 sekundowe pliki wave. Sonifikowanym obiektom przydzielono tony muzyczne od G2 (98 Hz) do B4 (493 Hz) dla obiektów interpretowanych jako przeszkody oraz tony z zakresu G3 (196 Hz) do G4 (392 Hz) dla płaszczyzn. Dla przeszkód zastosowano dźwięk pianina a dla płaszczyzn syntezowany dźwięk kontrabas. Taki wybór dźwięków dla dwóch klas obiektów, tj. różniących się zarówno wysokością tonu jak i barwą był podyktowany regułami segregacji i integracji strumieni dźwiękowych wg Bregmana (1990).

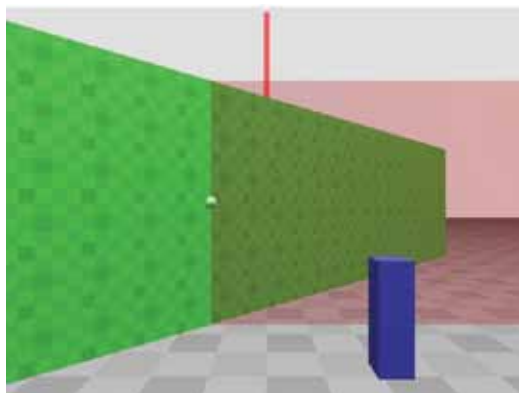
Na rys. 5.11 wykreślono sygnały dźwiękowe wytwarzane w kanale dźwiękowym lewym i prawym wg metody sonifikacji polegającej na skanowaniu głębi. Sekwencja czasowa generowanych dźwięków odpowiada układowi przestrzennemu obiektów pokazanemu po lewej stronie rysunku.



Rys. 5.11. Rzut pionowy sceny (z lewej) i wykres sygnałów dźwiękowych (z prawej) wytwarzanych w lewym i prawym kanale stereofonicznym (oznaczenia numeryczne sygnałów odpowiadają numerom obiektów)

Źródło: (Bujacz, 2010)

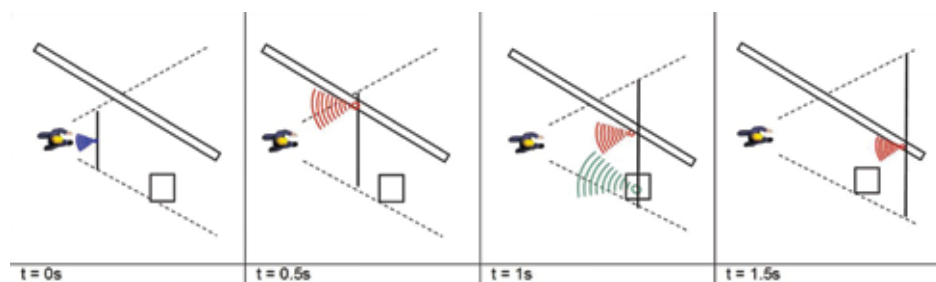
Na rys. 5.12 pokazano obraz uzyskany z symulacji komputerowej omawianej metody sonifikacji, a na rys. 5.13. objaśnienie sposobu sonifikacji takiej sceny. W chwili rozpoczęcia skanowania, $t = 0$, jest generowany krótki, referencyjny sygnał dźwiękowy informujący o rozpoczęciu cyklu przemiatania głębi. Dla



Rys. 5.12. Obraz sceny trójwymiarowej z widoczną częścią płaszczyzny (kolor jasnoróżowy) znajdującej się za niebieskim obiektem; płaszczyzna przecina ścianę wykreśloną kolorem zielonym (mała kula oznacza miejsce, w którym znajduje się źródło dźwięku sonifikujące ścianę)

Źródło: opracowanie własne

chwili $t=0,5$ s wycinek płaszczyzny (widziany w polu widzenia układu akwizycji obrazu) przecina ścianę i łączy ją z sonifikacją. W chwili $t=1$ s płaszczyzna przecina zarówno prostopadłościan jak i ścianę, generowane są jednocześnie dwa dźwięki: krótki informujący o detekcji prostopadłościanu i trwający od chwili $t=0,5$ s dźwięk sonifikujący ścianę. W chwili $t=1,5$ s ściana nadal podlega sonifikacji za pomocą dźwięku o zmniejszającej się głośności. Cykl skanowania głębi powtarza się z okresem regulowanym przez użytkownika systemu. W zastosowanym schemacie obrazowania dźwiękowego otoczenia płaszczyzna podłoża nie podlega sonifikacji.



Rys. 5.13. Metoda sonifikacji polegająca na skanowaniu głębi za pomocą płaszczyzny przemieszczającej się w scenie (rzut wycinka płaszczyzny określonego polem widzenia układu wizyjnego jest wykreślony pogrubioną linią ciągłą)

Źródło: (Bujacz i inni, 2011)

Omówiony system zastosowano do testów nawigowania i omijania przeszkód w środowisku modelowym, w którym przeszkodami były kartonowe pudełka. W teście wzięło udział 5 osób sprawnych wzrokowo i 5 osób niewidomych i słabowidzących. W czasie testów osobom sprawnym wzrokowo i słabowidzącym zasłonięto oczy opaską. Bezpośrednio przed rozpoczęciem testów uczestnicy otrzymali krótkie wyjaśnienie schematu sonifikacji otoczenia. Przeprowadzono dwa rodzaje testów. Testy wstępne polegały na wskazywaniu kierunku przeszkód (bez marszu) a następnie dojściu do przeszkody spostrzeganej jako bardziej odległej. Druga grupa testów to testy nawigacyjne. Testy te składały się kilku serii, w których zadaniem uczestników było przejście trasy z rozmieszczonymi dwoma, a następnie czterema przeszkodami oraz wskazanie małego obiektu (telefonu komórkowego) odtwarzającego muzykę i znajdującego się na końcu trasy. Na zdjęciu z rys. 5.14 pokazano przebieg testu wstępnego, w którym uczestnik wskazuje białą laską kierunki, z których spostrzega sonifikowane przeszkody, a na rys. 5.15 pokazano zdjęcie z przebiegu testów nawigacyjnych z udziałem osoby niewidomej.



Rys. 5.14. Przebieg testu spostrzegania kierunku przeszkód, uczestnik testu białą laską wskazuje kolejno kierunki spostrzeganych przeszkód

Źródło: opracowanie własne



Rys. 5.15. Test przejścia trasy z rozmieszczonymi przeszkodami i odszukaniem obiektu odtwarzającego melodię (wskazanego strzałką)

Źródło: opracowanie własne

Szczegółowe wyniki testów można znaleźć w pracy (Bujacz, 2010). W tabeli 5.4 zamieszczono ich syntetyczne podsumowanie.

Tabela. 5.4. Wyniki testów nawigacyjnych za pomocą prototypu systemu do dźwiękowego obrazowania otoczenia

Liczba błędów nawigacyjnych dla 100 prób	Test 1	Test 2
Niegroźne kolizje z przeszkodą w wyniku błędnie ocenionego kierunku	3	0
Niegroźne kolizje z przeszkodą w wyniku błędnie ocenionej odległości	3	4
Nieznaleziony cel w wyniku błędnie ocenionego kierunku	3	6
Nieznaleziony cel w wyniku błędnie ocenionej odległości	5	2
Liczba błędów nawigacyjnych dla 100 prób	Test 3	Test 4
Błędnie wybrana ścieżka	2	11
Niegroźne kolizje spowodowane błędnie wybraną ścieżką	0	6
Niegroźne kolizje z przeszkodą w wyniku błędnie ocenionego kierunku	3	0
Niegroźne kolizje z przeszkodą w wyniku błędnie ocenionej odległości	4	4
Utrata orientacji przestrzennej	1	0

Test 1: Wskazanie jednej przeszkody i marsz w jej kierunku,

Test 2: Wskazanie dwóch przeszkód i marsz w kierunku bardziej odległej przeszkody,

Test 3: Przejsie trasy o dł. 5 m do celu z rozmieszczonymi 2 przeszkodami,

Test 4: Przejsie trasy o dł. 5 m do celu z rozmieszczonymi 4 przeszkodami.

Źródło: opracowanie własne

Wyniki podano łącznie dla osób niewidomych i osób ze sprawnym wzrokiem z uwagi na statystycznie podobne liczby błędów. Osoby niewidome szybciej pokonywały trasy w testach nawigacyjnych. Dla testów 3 i 4 uzyskały one odpowiednio czasy 21 s i 25 s podczas gdy czasy pokonania trasy przez osoby sprawne wzrokowo wynosiły odpowiednio 27 s i 42 s. Jako błędnie wybraną ścieżkę (11 na 100 prób testowych) uznawano trasę wybraną przez uczestników (spośród kilku alternatywnych tras), w której należało przejść między dwoma przeszkodami odległymi od siebie o 60 cm. Część osób zwracała uwagę, że świadomie wybrało tę trasę kierując się prawidłowo ocenioną odległością pomiędzy przeszkodami. W opisanych testach zastosowano dwusekundowy cykl przemiatania głębi. Po zakończeniu testu połowa uczestników wybrała krótsze cykle przemiatania głębi, tj. 1 s i 1,5 s i z dobrym skutkiem powtórzyła zadania nawigacyjne.

Osoby niewidome oceniły system jako bardzo pomocny w nawigowaniu w nieznanym terenie (Charłampowicz, 2011). Przeprowadzono 100 serii testowych, jednak z uwagi na małą liczbę niewidomych uczestników testu (5 osób) oraz warunki modelowe testu, referowane wyniki badań prototypu systemu należy nadal uznać jako wstępne ale warte kontynuacji.

Podsumowując przegląd systemów niewizualnej prezentacji obrazów dla osób niewidomych należy stwierdzić, że systemy te są nadal przedmiotem badań rozwojowych lub testów użytkowych. Pomimo opracowanych wielu prototypów wykorzystujących zarówno interfejs dźwiękowy jak i dotykowy urządzenia te nie są jeszcze w praktyce stosowane przez niewidomych.

5.3. Systemy nawigacji osobistej o architekturze rozproszonej

W systemach nawigacji osobistej o architekturze rozproszonej wykorzystuje się układy i sieci transmisji bezprzewodowej sygnałów lokalizacyjnych i nawigacyjnych (Radecki i Buczkowski, 2010). Osoba niewidoma jest wyposażona w osobiste urządzenie nawigacyjne odbierające sygnały lub dwustronnie komunikujące się z węzłami takich sieci. Systemy o architekturze rozproszonej wykorzystywane do wspierania osoby niewidomej w nawigacji można podzielić na:

- systemy dalekiego zasięgu, do których zaliczają się m.in. globalny system pozycjonowania GPS i sieć telefonii komórkowej,
- systemy krótkiego zasięgu z wykorzystaniem latarni nawigacyjnych (m.in. radiolatarni), umieszczanych w stałych punktach otoczenia na zewnątrz lub wewnątrz budynków.

Oryginalnym rozwiązaniem systemu nawigacji osobistej o architekturze rozproszonej są systemy wspomagania nawigacji osób niewidomych, w których zdalny operator komunikuje się z osobą niewidomą i pełni funkcję jego zdalnego przewodnika. Prototypowy system tej klasy opracowano w Instytucie Elektroniki Politechniki Łódzkiej (Barański i inni, 2010a).

Pierwsze zastosowania systemu GPS do nawigacji pieszej dla niewidomych powstały już w latach osiemdziesiątych ubiegłego wieku (Collins, 1985). W roku 1996 w ramach projektu MoBIC (*Mobility of Blind and Elderly People Interacting with Computer*) opracowano system wyznaczania lokalizacji specjalnie przystosowany do obsługi przez osoby niewidome i starsze (Petrie i inni, 1996). Jednak dopiero od roku 2000, po udostępnieniu przez Departament Obrony USA systemu GPS o nominalnej dokładności użytkownikom cywilnym, datuje się szybki rozwój systemów nawigacji samochodowej i pieszej, w tym systemów nawigacji dla osób niewidomych (Loomis, 2001).

Dostępne systemy nawigacji satelitarnej projektowane dla osób niewidomych (tabela 5.5) charakteryzują się następującymi cechami ułatwiającymi ich obsługę przez osobę niewidomą:

Tabela 5.5. Wybrane systemy nawigacji osobistej dla niewidomych wykorzystujące system GPS

Nazwa (producent, strona www)	Charakterystyka
BrailleNote GPS (Sendero Group: www.senderogroup.com)	Odbiornik GPS współpracujący z notatnikiem Braille'a (BrailleNote) lub syntezatorem (VoiceNote). Funkcje: planuje ścieżkę dla wyznaczonego adresu, POI lub zadanej współrzędnej geograficznej, informuje głosem o kierunku marszu i bieżącym miejscu, zapisuje komentarze dla bieżącego miejsca trasy, rejestruje ścieżkę marszu (ma możliwość jej odtworzenia), oblicza długość drogi.
Trekker GPS (HumanWare: www.humanware.com)	Odbiornik GPS współpracujący z urządzeniem klasy PDA. Funkcje: wykorzystuje mapy Navteq, ma dwa tryby wyznaczania tras wg map, tj. samochodowy i pieszy oraz tryb bez mapy z nawigacją wg punktów kontrolnych, zapisuje komentarze dla bieżącego miejsca trasy, ma możliwość uprzedniego planowania i zapamiętania ścieżki.
Trekker Breeze (HumanWare: www.humanware.com)	Odbiornik GPS zintegrowany z interfejsem użytkownika w jednej obudowie (12,7x5x2,5 cm). Funkcje: urządzenie z założenia prostsze w obsłudze niż Trekker i o mniejszej liczbie funkcji (m.in. nie tworzy automatycznie ścieżek), odtwarza ścieżkę marszu i automatycznie tworzy ścieżkę dla drogi powrotnej, łatwy w obsłudze.
Nawigator (Migraf: www.migraf.pl)	Odbiornik GPS zintegrowany z interfejsem użytkownika w jednej obudowie (11x6x2 cm). Funkcje: nie wykorzystuje map, ma bogaty zestaw POI. Informuje o położeniu względem miejsca POI, przykładowy komunikat: „Miejsce dwa, sklep, osiemdziesiąt metrów, trzydzieści stopni w lewo”, wyznacza ścieżki jako uporządkowaną sekwencję POI, umożliwia definiowanie własnych POI, uruchomiona lista dyskusyjna użytkowników, łatwy w obsłudze.

Źródło: opracowanie własne na podstawie stron internetowych wyszczególnionych w tabeli

- interfejs użytkownika wyposażony w odpowiednią klawiaturę dotykową (lub brajlowską) służącą m.in. do wpisywania danych o miejscach użytecznych (*Points of Interest – POI*), syntezy dźwiękowej i/lub syntezy mowy oraz dyktafon do wprowadzania komentarzy słownych dla miejsc POI (niektóre wersje interfejsów są również wyposażane w tzw. linijki brajlowskie lub interfejsy dotykowe z elementami wibrującymi),
- funkcje planowania tras pieszych na podstawie sekwencji zapisanych punktów kontrolnych lub orientacyjnych (np. POI) w postaci współrzędnych geograficznych i nawigowania wzdłuż takiej trasy (m.in. podawania głosowych komunikatów o bieżącym położeniu, nazwie ulicy i dalszym kierunku marszu),
- funkcje planowania tras pieszych (*routing*) z wykorzystaniem map cyfrowych i nadzorowanego nawigowania wzdłuż wyznaczonej trasy.

Urządzenia osobistej nawigacji GPS dla osób niewidomych mogą zawierać tylko niektóre z ww funkcji. Trzeba pamiętać, że każda osoba niewidoma ma inne umiejętności obsługi takich urządzeń oraz odmienne potrzeby w zakresie wsparcia w nawigacji terenowej, np. inne przy marszu wzdłuż znanej ścieżki, a inne przy pokonywaniu nowej trasy. W tabeli 5.5 podano wykaz najpopularniejszych systemów GPS. W większości systemów wykorzystuje się odbiorniki GPS współpracujące z urządzeniami klasy PDA (*Personal Digital Assistant*) lub telefonami komórkowymi. Niektórzy producenci systemów GPS dla niewidomych (m.in. amerykański Trekker Breeze i polski Nawigator) dla poprawy ergonomii urządzeń oferują dedykowane rozwiązania sprzętowe, w których odbiorniki GPS są zamknięte w jednej obudowie z interfejsem użytkownika systemu. Zdjęcia tych urządzeń pokazano na rys. 5.16.

Jedyne znane autorowi opracowanie prototypowe, w którym system GPS jest jednym z modułów zintegrowanego systemu nawigacji wykorzystującego, m.in.



Rys. 5.16. Urządzenia nawigacji satelitarnej zaprojektowane specjalnie dla osób niewidomych: Nawigator produkowany przez polską firmę Migraf (z lewej) oraz Trekker Breeze produkowany przez firmę HumanWare

Źródło: fotografie ze stron internetowych www.migraf.pl i www.humanware.com (za zgodą producentów)

technologię RFID i informacje z bazy danych systemu informacji pasażerskiej jest wdrażane w Finlandii (NOPPA, 2011). W Politechnice Poznańskiej została opracowana aplikacja o nazwie Itiner dla telefonów komórkowych udostępniająca informacje o rozkładach jazdy pojazdów komunikacji miejskiej. Zawiera ona tzw. planer podróży i dostarcza informacji o czasie oczekiwania na pojazdy (na podstawie uzyskanej informacji lokalizacyjnej z systemu GPS). Aplikacja jest udostępniona bez opłaty (www.itiner.pl). Trwają prace nad jej przystosowaniem do obsługi przez osoby niewidome.

Warto również wspomnieć o aplikacjach nawigacji pieszej dla urządzeń przenośnych, specjalnie przystosowanych do obsługi przez osoby niewidome. Oprócz aplikacji sprzedawanych wraz z produktami firmowymi takimi jak Trekker czy BrailleNote są również dostępne aplikacje dla wielu modeli telefonów komórkowych wyposażonych w system operacyjny. Są to m.in.:

- Loadstone (www.loadstone-gps.pl), darmowy program do nawigacji satelitarnej dla telefonów z systemem operacyjnym Symbian serii 60 (istnieje wersja polska, która nie wykorzystuje map ale zawiera stale powiększany zbiór miejsc POI regularnie aktualizowany przez społeczność użytkowników); przy braku dostępu do sygnału GPS określa przybliżone położenie użytkownika na podstawie identyfikacji najbliższej stacji bazowej telefonii komórkowej GSM,
- Wayfinder Access (www.wayfinder.com), płatna aplikacja dla telefonów komórkowych z systemem operacyjnym Symbian serii 60, korzysta z map cyfrowych i udostępnia dwa tryby nawigacji: pieszy i samochodowy, zawiera ok. 60 mln POI dla map Europy.

Ocenę wybranych systemów i aplikacji do nawigacji satelitarnej opracowaną przez niewidomych użytkowników można znaleźć na stronie internetowej amerykańskiej federacji osób niewidomych (*National Federation of the Blind* – NFB, 2011).

Znaną i podstawową wadą systemów nawigacji wykorzystujących globalne systemy pozycjonowania (dyskutowaną w rozdz. 3.) jest mała dokładność wyznaczanej lokalizacji w obszarach wysokiej zabudowy oraz brak sygnałów nawigacyjnych w przestrzeniach zamkniętych (wewnątrz budynków, w halach, dworcach, przejściach podziemnych). Jest to szczególnie istotna wada w nawigacji pieszej dla osób niewidomych. Dlatego też prowadzonych jest wiele badań, w których globalne systemy pozycjonowania są wspierane nawigacją zliczeniową (zob. m.in. (Barański i Strumiłło, 2009a)). Alternatywnym rozwiązaniem problemu zapewnienia ciągłej informacji lokalizacyjnej jest zastosowanie systemów bezprzewodowych krótkiego zasięgu wykorzystujących koncepcję latarni nawigacyjnych. Systemy takie pracują w ogólnodostępnym paśmie radiowym

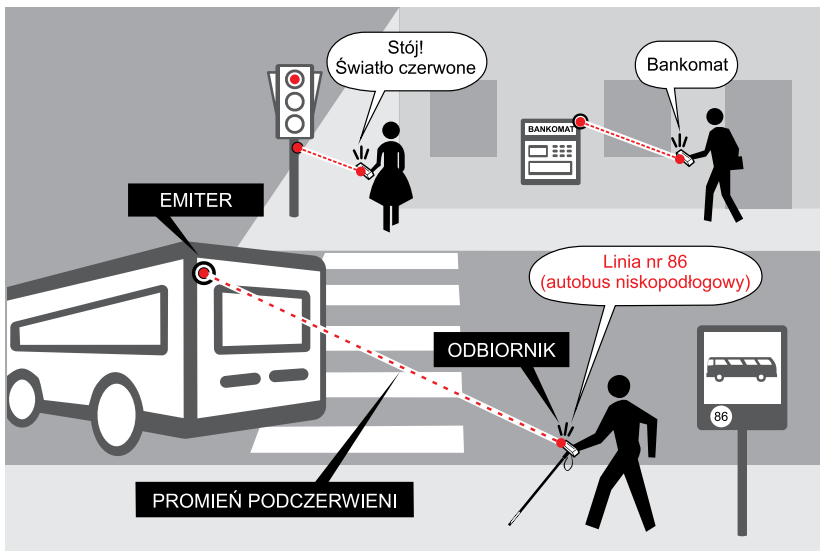
ISM (*Industrial Scientific Medical*) i wykorzystują różne technologie radiowe m.in.: RFID, WLAN (IEEE 802.11), WPAN/ZigBee (IEEE 802.15.4), Bluetooth (IEEE 802.15.1), oraz technologie optyczne wykorzystujące promieniowanie w bliskiej podczerwieni (850÷900nm), m.in. IrDA (*Infrared Data Association*). Wybrane systemy tej klasy zestawiono w tabeli 5.6.

Podstawowe wady systemów nawigacji osobistej wg latarni nawigacyjnych są następujące:

- duży koszt instalacji i utrzymania systemu,
- konieczne utworzenie gęstej sieci nadajników w obszarze działania systemu,
- brak standardów dla systemów tej klasy (istnieje wiele rozwiązań prototypowych wykorzystujących różne protokoły komunikacji i rozwiązania sprzętowe (Willis i Helal, 2005), (Bohonos i inni, 2007), (Ni i Zhang, 2011)).

Ze względu na minimalizację kosztów systemów nawigacji wewnątrz budynków stosowane są rozwiązania wykorzystujące coraz powszechniejsze sieci WLAN z zastosowaniem metod pomiaru poziomu sygnału (RSS) i czasu propagacji sygnału (RTOF) (Gu i inni, 2009).

Nowych technik lokalizacji w terenie otwartym można oczekiwać po szerszym rozpowszechnieniu sieci szerokopasmowej transmisji danych WiMAX (*Worldwide Interoperability for Microwave Access*) (Bshara i inni, 2010).



Rys. 5.17. Idea działania systemu TalkingSigns; sieć emiterów światła podczerwonego wbudowanych w infrastrukturę miasta i pojazdy komunikacji informują niewidomego o miejscach użytecznych (POI) i numerach pojazdów komunikacji miejskiej

Źródło: opracowanie własne na podstawie www.talkingsigns.com

Tabela 5.6. Systemy nawigacji dla osób niewidomych wykorzystujące latarnie nawigacyjne

Nazwa	Charakterystyka systemu (adres www bądź źródło)
Talking Signs	System emiterów światła podczerwonego jest rozmieszczany w stałych miejscach (wewnątrz lub na zewnątrz budynków) oraz pojazdach komunikacji miejskiej. Emitery są uaktywniane urządzeniem noszonym przez osobę niewidomą i wysyłają unikatowy kod, który jest przetwarzany przez odbiornik użytkownika na komunikaty głosowe. System wskazuje kierunek położenia emitera przez wykorzystanie efektu stereofonicznego. Komunikaty są okresowo powtarzane (rys. 5.17). (www.talkingsigns.com)
Blind-enT	System prototypowy opracowany w Akademii Podlaskiej w Siedlcach. Zastosowano ideę podobną do TalkingSigns z nadaniem struktury obiektowej elementom otoczenia; osoba niewidoma jest wyposażona w inteligentny telefon komórkowy, jest informowana o klasie obiektów kanałem IrDA. Komunikuje się z serwerami systemu przez GPRS lub Bluetooth. (http://ent.ipipan.waw.pl/blind-ent/pages/about.jsp)
SESAMONET	System 1260 transponderów RFID rozmieszczonych w mieście Ispra w północnych Włoszech (w roku 2008). Czytnik RFID zamontowany w białej lasce niewidomego komunikuje się z najbliższymi transponderami i po odbiorze kodu przesyła go łączem Bluetooth do telefonu komórkowego, który odtwarza odpowiedni komunikat głosowy o miejscach POI otoczenia. System na etapie testów. (http://spectrum.ieee.org/biomedical/devices/rfid-tags-guide-the-blind)
PAVIP – Personal Assistant for Visually Impaired People	System wspierający podróżowanie komunikacją miejską dla niewidomych. W systemie PAVIP wykorzystano technologie RFID. Niewidomy otrzymuje czytnik RFID wyposażony w syntezytor mowy. Sieć pasywnych transponderów znajduje się w rozkładach jazdy na przystankach oraz transponderów aktywnych w pojazdach komunikacji miejskiej. Niewidomy uzyskuje informacje o rozkładzie jazdy pojazdów

Tabela 5.6. cd.

Nazwa	Charakterystyka systemu (adres www bądź źródło)
	na przystanku i wybiera w urzędzeniu numer pojazdu, jest automatycznie informowany o nadjeżdżającym pojeździe wybranej linii. W czasie jazdy (z osobistego urządzenia) otrzymuje komunikaty głosowe o kolejnych przystankach. (www.bones.ch)
ASBUS	Prototypowy system wspomagania orientacji przestrzennej dla niewidomych działający na terenie kampusu Uniwersytetu w Stuttgarcie wykorzystuje bardzo dokładne mapy trójwymiarowe budynków i terenu kampusu zebrane z wykorzystaniem skanerów laserowych. Niewidomy jest wyposażony w tablet PC z podłączonymi do niego kompasem, czujnikami inercyjnymi, kamerą stereowizyjną, odbiornikiem GPS i czytnikiem RFID. Urządzenie automatycznie rozpoznaje otoczenie na podstawie dopasowania obrazów zapamiętanych w mapie oraz informacji o bieżącej lokalizacji niewidomego uzyskanych z sieci transponderów RFID, odczytów GPS i sygnałów z czujników inercyjnych. Autorzy systemu określają dokładność wyznaczania położenia na terenie kampusu na ok. 1 m. Położenie to jest automatycznie przesuwane na środek ekranu dotykowego. Wskazanie tego położenia automatycznie odczytuje informacje o miejscach POI zdefiniowanych w bazie systemu. (http://cumbia.informatik.uni-stuttgart.de/eng/research/proj/sfb627/asbus/)
RNIB React	System wykorzystujący sieć radiolatarni w.cz. oraz emiterów podczerwieni (InfraVoice). Osoba niewidoma jest wyposażona w urządzenie nadawczo-odbiorcze wielkości telefonu komórkowego. Użytkownik po aktywowaniu osobistego urządzenia otrzymuje komunikaty słowne o swojej lokalizacji i miejscach POI otoczenia. System opracowany w 2004 roku w Wielkiej Brytanii i testowany w kilku miastach Anglii i Szkocji. (www.rnib.org.uk/professionals/accessibleenvironments/signagewayfinding/rnibreact/Pages/rnib_react.aspx)

Tabela 5.6. cd. Systemy nawigacji dla osób niewidomych wykorzystujące latarnie nawigacyjne

Nazwa	Charakterystyka systemu (adres www bądź źródło)
ESOT	System wykorzystujący nadajniki o krótkim zasięgu nadające sygnał w.cz, z którymi komunikuje się osobisty odbiornik niewidomego. Odbiornik odtwarza słowny komentarz o wykrytej radiolatarni. W najnowszej wersji systemu zastosowano pomiar poziomu sygnału (<i>Received Signal Strength – RSS</i>) dla uzyskania informacji o odległości i kierunku położenia radiolatarni. Pierwsze wersje systemu powstały w latach 1999-2001 we współpracy Instytutu Radioelektroniki Politechniki Warszawskiej i Instytutu Biocybernetyki i Inżynierii Biomedycznej PAN. Wersja prototypowa systemu jest testowana przez niewidomych użytkowników. <i>(Radecki i inni, 2004), (Marski i inni, 2010)</i>
Step-Hear	Zastosowano sieć nadajników w.cz. (433 MHz) pełniących funkcje radiolatarni i w sposób ciągły nadający sygnał radiowy. Gdy użytkownik znajdzie się w zasięgu radiolatarni odbiornik sygnalizuje wykryty nadajnik i informuje użytkownika krótkim dźwiękiem i jednoczesną wibracją. Po naciśnięciu odpowiedniego przycisku użytkownik nadaje sygnał zwrotny do nadajnika, który odtwarza z głośnika zapisany wcześniej komunikat. Opracowano wersje radiolatarni montowanych w pojazdach komunikacji miejskiej. Kierowca pojazdu jest automatycznie informowany o osobie niewidomej na przystanku. <i>(www.step-hear.com)</i>

Źródło: opracowanie własne na podstawie stron internetowych wyszczególnionych w tabeli

5.4. Telenawigacja przy pomocy zdalnego przewodnika

Najbezpieczniejszą pomocą nawigacyjną osoby niewidomej jest inna sprawna i odpowiednio przeszkolona osoba pełniąca rolę przewodnika ostrzegającego zarówno przed przeszkodami w przestrzeni bliskiej (mikronawigacja) jak i wytaczająca drogę marszu do celu (makronawigacja). Taki sposób wsparcia mobilności niewidomego ma określone wady. Czyni go w dużym stopniu zależnym od dyspozycyjności swojego przewodnika. Ma on poczucie obciążania innych swoją niepełnosprawnością. Uważa się również, że nadmierna pomoc przewodnika prowadzi do pogorszenia orientacji przestrzennej i sprawności psychofizycznej niewidomego (Hollins, 1989).

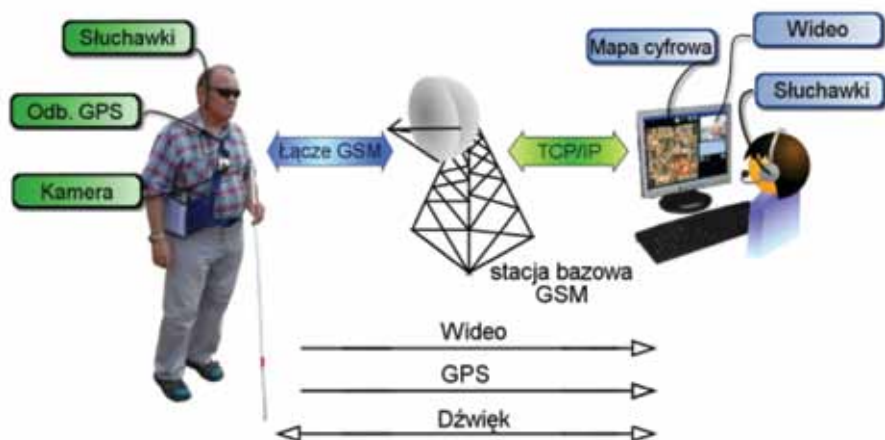
Z zamieszczonego w poprzednich rozdziałach pracy przeglądu elektronicznych systemów osobistej nawigacji osób niewidomych można wnioskować, że największy postęp dokonał się w zakresie wsparcia makronawigacji. Dzięki globalnym systemom pozycjonowania (m.in. GPS), technologii map cyfrowych i aplikacjom nawigacji pieszej wyposażonym w odpowiednie interfejsy użytkownika osoba niewidoma uzyskuje skuteczną pomoc w orientacji geograficznej. Trudniejszym zagadnieniem jest zapewnienie odpowiedniej orientacji w przestrzeni bliskiej i mikronawigacji.

Dzięki dynamicznemu postępowi technologii teleinformatycznych jaki dokonał się w ostatniej dekadzie, stała się technicznie możliwa realizacja oryginalnej metody wsparcia nawigacyjnego osoby niewidomej. Polega ona na utworzeniu łącza komunikacyjnego pomiędzy przewodnikiem a osobą niewidomą. Przewodnik może znajdować się w dowolnym miejscu i na odległość prowadzić osobę niewidomą na podstawie przesyłanego mu obrazu otoczenia rejestrowanego przez osobiste urządzenie niewidomego. Pierwszy prototypowy system tej klasy został opracowany w Wielkiej Brytanii w Brunel University (Garaj i inni, 2003). Niewidomy użytkownik był wyposażony w dwa telefony komórkowe GSM, jeden do łączności głosowej ze zdalnym przewodnikiem, a drugi do przesyłania danych lokalizacyjnych GPS. Zastosowano oddzielne, dedykowane łącze radiowe w.cz. do przesyłania obrazów rejestrowanych przez kamerę wideo. Pomimo utrudnionej obsługi terminala mobilnego składającego się z czterech urządzeń (dwa telefony GSM, odbiornik GPS i kamera z nadajnikiem radiowym), autorzy systemu przeprowadzili pierwsze jego testy (*proof of concept*) pokazując skuteczność takiej metody zdalnego prowadzenia osoby niewidomej (Hunaiti i inni, 2005). Niestety, prac nad prototypem prawdopodobnie nie kontynuowano (od opublikowania cytowanego referatu nie ukazały się późniejsze doniesienia literaturowe dokumentujące wyniki dalszych prac). System nawigacji

z udziałem zdalnego przewodnika został też zbudowany przez polską fundację o nazwie Strefa Designu Innowacji Integracji. Jego działanie było ograniczone do wnętrza budynków w obszarze zasięgu sieci lokalnych WiFi.

Prace nad systemem nawigacji dla osób niewidomych, w którym wykorzystuje się idee prowadzenia osoby niewidomej przez zdalnego przewodnika zostały również podjęte w roku 2008 w Instytucie Elektroniki Politechniki Łódzkiej w ramach grantu rozwojowego finansowanego przez Ministerstwo Nauki i Szkolnictwa Wyższego (IE PŁ, 2010). Celem projektu było opracowanie prototypu systemu zdalnej nawigacji dla osób niewidomych poruszających się na zewnątrz budynków z wykorzystaniem systemu GPS, map cyfrowych oraz transmisji obrazów wideo do terminala zdalnego przewodnika. Pierwszą wersję systemu zbudowano z wykorzystaniem komunikatora internetowego Skype i poddano testom z udziałem osób niewidomych. Wyniki tych testów, które zreferowano w pracy (Bujacz i inni, 2008a) pokazały dużą skuteczność tej techniki nawigacji. Osoby niewidome pewnie i bezpieczniej pokonywały ścieżki testowe.

W drugim etapie prac napisano dwie dedykowane aplikacje na komputery klasy PC uruchamiane w terminalu mobilnym osoby niewidomej i terminalu zdalnego przewodnika. Główne elementy takiego systemu pokazano na rys. 5.18.



Rys. 5.18. Idea systemu prowadzenia osoby niewidomej przez zdalnego przewodnika

Źródło: opracowanie własne

Terminal mobilny składa się z komputera przenośnego oraz podłączonych do niego: kamery cyfrowej, odbiornika GPS i zestawu słuchawkowego (rys. 5.19). W kamerze zastosowano własnej konstrukcji obiektyw szerokokątny typu FishEye, który umożliwił zwiększenie kąta widzenia kamery do ok. 160°. Terminal

zdalnego przewodnika stanowił komputer stacjonarny PC oraz zestaw słuchawkowy. Widok interfejsu graficznego aplikacji tego terminala pokazano na rys. 5.19.



Rys. 5.19. Widok okna aplikacji terminala zdalnego przewodnika (z lewej) wyświetlającego obraz z kamery terminala mobilnego (A) oraz mapę terenu (B) z zaznaczoną ścieżką marszu osoby niewidomej oraz zdjęcie osoby niewidomej z terminalem mobilnym systemu (z prawej):

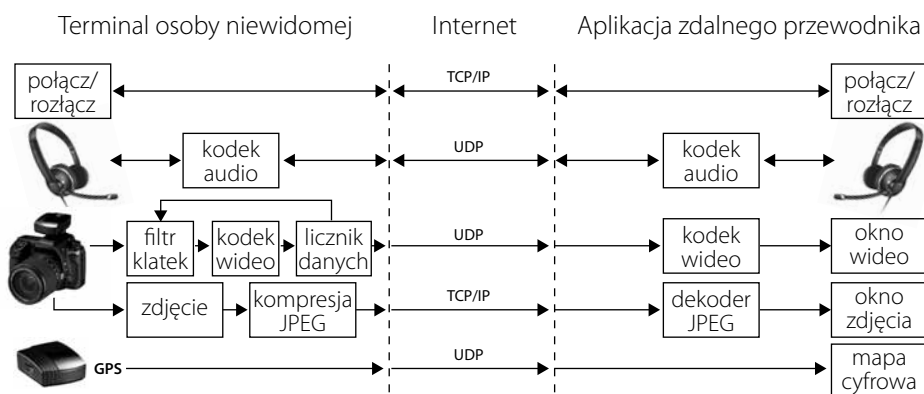
A – słuchawka, B – mikrofon, C – odbiornik GPS, D – kamera, E – komputer przenośny

Źródło: opracowanie własne

Schemat blokowy systemu z zaznaczonymi protokołami transmisji pomiędzy terminalem mobilnym a terminalem zdalnego przewodnika pokazano na rys. 5.20. Do transmisji danych zastosowano dwa protokoły: protokół internetowy TCP/IP (*Transmission Control Protocol/Internet Protocol*) ze śledzeniem pakietów i ewentualną ich retransmisją oraz protokół UDP (*User Datagram Protocol*) bez śledzenia pakietów ale oferujący większą szybkość transmisji.

Połączenie terminali systemu może być inicjowane zarówno przez zdalnego przewodnika jak i osobę niewidomą. Proces nawiązywania połączenia odbywa się z wykorzystaniem protokołu TCP/IP.

Głos jest poddawany kompresji za pomocą kodeka dźwięku z rozdzielczością 8 bitów na próbkę i częstotliwością próbkowania 8kHz. Próbkę dźwięku są składane w pakiety i przesyłane protokołem UDP. Sekwencja obrazów rejestrowana przez kamerę terminala mobilnego (w wersjach rozdzielczości: 640 x 480, 320 x 240 i 160 x 120) jest poddawana kompresji za pomocą kodeka MPEG-2 i przesyłana protokołem UDP. Ważną zastosowaną procedurą systemu jest sterowana szybkość transmisji obrazów strumienia wideo zależnie od regularnie testowanej chwilowej przepływności łącza UDP (maksymalna częstość jest równa



Rys. 5.20. Uproszczony schemat blokowy ilustrujący transmisję danych pomiędzy terminalami systemu

Źródło: opracowanie własne

15 obrazów na sekundę). System umożliwia również przesyłanie pojedynczych zdjęć o maksymalnej rozdzielczości kamery i jest kompresowany bez utraty jakości (funkcja przydatna do przesyłania zdjęć o rozdzielczości umożliwiającej odczytanie osobie niewidomej tekstu przez zdalnego przewodnika). Odczyty odbiornika GPS są przesyłane do terminala zdalnego przewodnika co 1 sek. Do transmisji danych pomiędzy terminalami wykorzystano sieć GSM i technologie HSDPA/UMTS z komutacją pakietów.

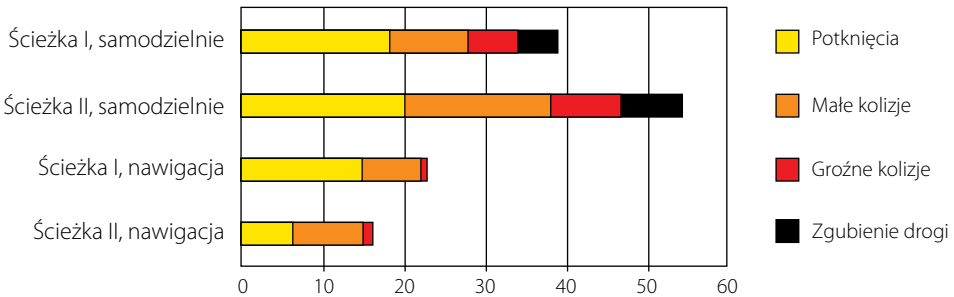
Opisany system poddano testom z udziałem 3 osób niewidomych o różnej sprawności posługiwania się białą laską. Testy przeprowadzono na terenie kampusu Politechniki Łódzkiej w roku 2008. Każdy z uczestników miał za zadanie przejść dwie ścieżki o długości ok. 150 m i 190 m. Zaplanowane ścieżki były pokonywane przez osoby niewidome w następującej kolejności:

1. Przejście z osobistym przewodnikiem, który objaśniał drogę.
2. Przejście ze zdalnym przewodnikiem – osoba niewidoma była „prowadzona” komunikatami głosowymi zdalnego asystenta.
3. Przejście samodzielne – osoba niewidoma pokonywała ścieżkę samodzielnie z użyciem białej laski.

Dla każdego testowego przejścia ścieżki notowane były następujące parametry:

- czas przejścia ścieżki,
- liczba potknięć,
- liczba małych kolizji (np. niegroźny kontakt z przeszkodą),
- liczba potencjalnie groźnych kolizji (konieczna interwencja instruktora towarzyszącego osobie niewidomej),
- zgubienie ścieżki.

Bardziej szczegółowy opis przebiegu testów zawarto w pracy (Bujacz i in., 2008b). Na rys. 5.21 zebrano uśrednione ich wyniki.



Rys. 5.21. Porównanie uśrednionej liczby zdarzeń dla trzech uczestników testu i dwóch przejść ścieżek: samodzielnie z użyciem białej laski i przy pomocy zdalnego przewodnika

Źródło: opracowanie własne

Wyniki testów pokazały, że prowadzenie osoby niewidomej przez zdalnego przewodnika zwiększyło bezpieczeństwo poruszania się osób niewidomych. Zmniejszyła się liczba groźnych kolizji, nie stwierdzono też zagubienia drogi, a czas przejścia ścieżek został skrócony o 15%–20%.

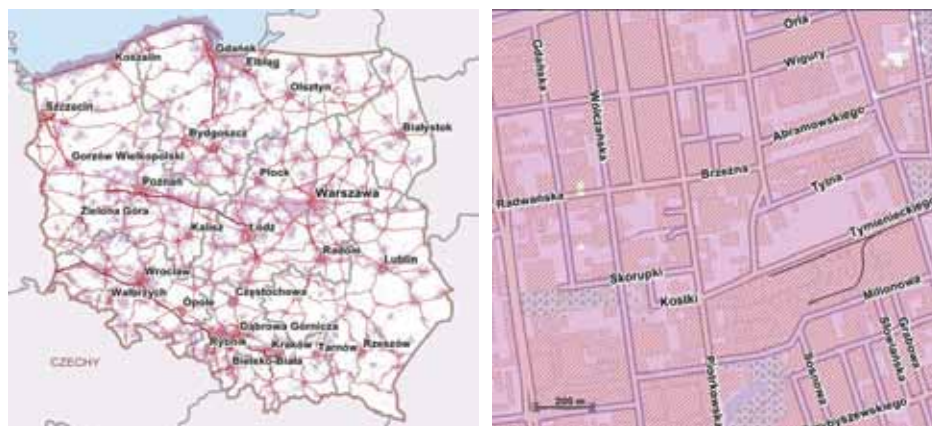
W czasie testów zwrócono uwagę na istotne elementy związane z jakością transmisji i sposobem komunikacji słownej pomiędzy zdalnym przewodnikiem i osobą przez niego prowadzoną. Osoby niewidome przedkładały krótkie instrukcje typu: „ławka po lewej”, „kosz na wprost” nad dokładne i zbyt częste instrukcje nawigacyjne. Stwierdzono też, że każda z osób niewidomych miała już wypracowany indywidualny sposób korzystania z punktów orientacyjnych otoczenia i prosiła o informacje na ich temat, np. „trzymaj się krawężnika z prawej”. Ustalono też, że umieszczenie kamery na głowie osoby niewidomej jest niekorzystne. Osoby niewidome w czasie marszu obracają głowę dla lepszego spostrzegania otaczających dźwięków i często patrzą nie na wprost kierunku marszu.

W czasie testów rejestrowano opóźnienia w transmisji danych (w tym sekwencji obrazów) oraz zdarzenia polegające na utracie transmitowanych pakietów. Znotowano opóźnienia transmisji dochodzące do ok. 2,5 s. Na rys. 5.22 na tle mapy miasta wykreślono linie marszu i zaznaczono monitorowane parametry transmisji. Uzyskana średnia szybkość transmisji obrazów wyniosła ok. 11 kB/s a transmisji audio ok. 7,5 KB/s (zauważmy duże odchyłki odczytów GPS od linii pokonywanej ścieżki; problem dyskutowany w rozdz. 3). Na rys. 5.23 pokazano aktualny zasięg sieci T-Mobile dla technologii transmisji danych UMTS/HSDPA wykorzystywanej w systemie. W Polsce, zasięg ten obejmuje głównie tereny miejskie.



Rys. 5.22. Ocena parametrów transmisji sekwencji obrazów wykonana dla testów systemu nawigacji dla osoby niewidomej przy pomocy zdalnego przewodnika; średnie opóźnienia oznaczają opóźnienie transmisji obrazu o ok. 200 ms, a duże o ok. 1000 ms; wyniki uzyskane w roku 2008 (zob. legenda mapy)

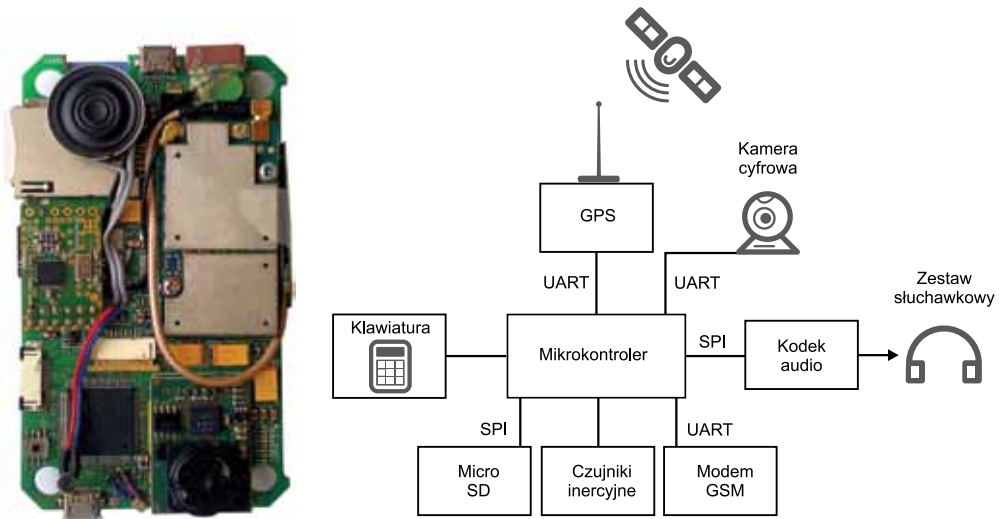
Źródło: opracowanie własne



Rys. 5.23. Zasięg usług sieciowych w technologii UMTS/HSDPA (operator T-Mobile): dla terenu Polski (z lewej) oraz dla obszaru miasta Łodzi, w którym przeprowadzano testy systemu (ciemniejsze obszary odpowiadają większej mocy sygnału radiowego w sieci),

Źródło: www.t-mobile.pl/pl/biznes/stali_klienci/uslugi_dla_firm/zasieg, stan z 7/08/2011 r.

Napotkane problemy z utratą pakietów, niezawodnością aplikacji terminala mobilnego oraz niedostateczną ergonomią modelu terminala mobilnego (w którym wykorzystano typowy komputer przenośny klasy PC) skłoniły autorów do podjęcia prac nad wykonaniem własnego modułu elektronicznego integrującego opisane funkcje terminala osoby niewidomej (Barański i inni, 2009b, 2010a). Schemat blokowy urządzenia oraz widok płytki elektronicznej wykonanego modułu pokazano na rys. 5.24. W pracy (Strumiłło i inni, 2011) zawarto bardziej szczegółowy opis tego układu elektronicznego.



Rys. 5.24. Schemat blokowy oraz zdjęcie układu elektronicznego terminala mobilnego systemu zdalnej nawigacji dla osoby niewidomej (rozmiar płytki drukowanej 56 mm×98 mm, waga układu z obudową i baterią: ok. 250 g)

Źródło: opracowanie własne

Układem steruje mikrokontroler z rdzeniem ARM z 32-bitowym zestawem instrukcji RISC. Urządzenia peryferyjne układu są obsługiwane w trybie bezpośredniego dostępu do pamięci DMA (*direct memory access*) dla zmniejszenia obciążenia mikroprocesora. Rozdzielczość transmitowanych obrazów wideo może być definiowana na bieżąco przez zdalnego przewodnika w zakresie od 160 x 120 do 640 x 480 punktów RGB. Zastosowano kąt widzenia kamery o wartości 100°. Urządzenie ma wbudowany modem GSM pełniący funkcję telefonu komórkowego oraz modemu transmisji danych. Strumienie wideo, audio i GPS są transmitowane przez port TCP/IP. Widok okna terminala zdalnego przewodnika pokazano na rys. 5.25.



Rys. 5.25. Widok okna programu terminala zdalnego przewodnika (z lewej obraz przesłany z terminala mobilnego, z prawej mapa cyfrowa miasta z zaznaczonym zieloną strzałką położeniem terminala mobilnego)

Źródło: opracowanie własne

Szczególną uwagę zwrócono na ergonomię terminala mobilnego. Zastosowano klawiaturę membranową własnego projektu (rys. 5.26). Połączenie z terminalem zdalnego przewodnika odbywa się wg prostej procedury, w trakcie której użytkownik otrzymuje komunikaty słowne o statusie połączenia. Komunikaty



Rys. 5.26. Zdjęcie osoby niewidomej korzystającej z terminala mobilnego systemu oraz aplikacja terminala zdalnego przewodnika w wersji dla urządzenia mobilnego

Źródło: fotografie własne

są odtwarzane z zastosowaniem difonowego syntezyatora mowy polskiej zaprojektowanego i wykonanego na potrzeby urządzenia (Barański i inni, 2010b). Od kilku miesięcy trwają testy długoterminowe systemu prowadzone przez niewidomego użytkownika (Charłampowicz, 2010). W przygotowaniu jest aplikacja serwera nadzorującego transmisje danych, która docelowo umożliwi skalowanie systemu dla odpowiednio większej liczby jednocześnie obsługiwanych użytkowników. Pierwsze testy rozpoczęto z aplikacją terminala zdalnego przewodnika zainstalowaną w urządzeniu mobilnym (rys. 5.26).

Można oczekiwać, że system nawigacji przy pomocy zdalnego przewodnika jest sposobem wsparcia mobilności niewidomych, który będzie szybko zyskiwał na znaczeniu. Regularnie zwiększająca się przepływność sieci telefonii komórkowej (np. sieć LTE (*Long Term Evolution*)) i sieci radiowych (np. WiMAX) umożliwi przesyłanie obrazów wideo o lepszej jakości i z większą prędkością.

W zakresie użyteczności tej klasy systemów należy zwrócić uwagę na nowe problemy związane z wypracowania odpowiednich komunikatów nawigacyjnych używanych przez przewodnika. Ważne jest też zrozumienie przez użytkowników systemu jego ograniczeń technicznych, m.in. czasu opóźnienia transmisji sygnałów i ograniczonej jakości przesyłanych obrazów.

Zebrane doświadczenia projektantów i użytkowników tego systemu nawigacji dla niewidomych można podsumować następująco:

- osoby niewidome oceniają system jako szczególnie przydatny w pokonywaniu i uczeniu się nowych ścieżek oraz jako nieodzowną pomoc w przypadku zagubienia ścieżki;
- kontakt głosowy ze zdalnym przewodnikiem zwiększa poczucie bezpieczeństwa i poprawia orientację przestrzenną,
- w wielu sytuacjach wystarcza krótka sesja „nawigacyjna” ze zdalnym przewodnikiem, np. pomoc w odszukaniu zgubionej drogi lub odszukaniu peronu na dworcu,
- użytkowników systemu należy traktować jako dwuosobowy zespół, który powinien być odpowiednio przeszkolony we wspólnym posługiwaniu się systemem; stwierdzono m.in., że każda osoba niewidoma lub słabowidząca wymaga odmiennego sposobu nawigacji słownej.

Ważnym niedyskutowanym zagadnieniem jest problem odpowiedzialności jaką bierze zdalny operator za bezpieczeństwo osoby niewidomej.

W przygotowaniu są testy z większą liczbą niewidomych użytkowników systemu.

W niniejszym rozdziale zawarto przegląd elektronicznych systemów nawigacji osobistej dla osób niewidomych. W systemach tych są stosowane zarówno różne techniki pozyskiwania informacji o otoczeniu jak i różne sposoby niewizualnej ich prezentacji. Różne są też konstrukcje tych elektronicznych systemów poczynając od prostych detektorów przeszkód, przez systemy obrazowania otoczenia aż po systemy o architekturze rozproszonej (wykorzystujące globalne systemy pozycjonowania, radiolatarnie nawigacyjne oraz systemy prowadzenia niewidomych przy pomocy zdalnych przewodników). Niektóre z tych systemów, w szczególności systemy nawigacji GPS są w coraz większym stopniu wykorzystywane przez osoby niewidome.

Podsumowanie

W ostatnich dziesięcioleciach dokonał się znaczący postęp w technice i medycynie. Układy elektroniczne wielkiej skali integracji, nanotechnologia, lasery, obliczenia i sieci komputerowe, telekomunikacja naziemna i satelitarna, inżynieria materiałowa są dziedzinami, które zmieniły życie każdego z nas. Postęp w ww. technologiach dał również podstawę do przełomowych osiągnięć w medycynie – diagnozie (elektrografie, tomografie komputerowe, telemedycyna) i terapii (roboty chirurgiczne, sterowane elektronicznie protezy, implanty, transplantologia). Na świecie przeprowadza się setki tysięcy transplantacji organów i miliony wszczepień protez rocznie.

Duży postęp dokonał się również w technologiach wspomagających osoby z niezbyt poważnymi niepełnosprawnościami sensorycznymi. Elektroniczne aparaty słuchowe są powszechnie stosowanymi urządzeniami korekcji słuchu. Najnowszej generacji aparaty cyfrowe umożliwiają programowanie widma wzmacnianych dźwięków np. selektywnie dla dźwięków mowy aby ułatwiać jej zrozumienie. Szkła korekcyjne, w tym szkła kontaktowe oraz wszczepiane sztuczne soczewki skutecznie poprawiają sprawność widzenia.

Znacznie trudniej jest pomóc osobom, które całkowicie utraciły sprawność określonego zmysłu. Należy pamiętać, że bodźce odbierane przez receptory naszych zmysłów podlegają złożonym procesom przetwarzania, analizy i ich kognitywnej interpretacji. Przywrócona sprawność części sensorycznej systemu zmysłowego nie jest jednoznaczna z przywróceniem pełnej jego sprawności. Osoby długotrwale pozbawione słuchu lub wzroku z trudem nabywają (lub w ogóle nie nabywają) umiejętności rozpoznawania świata dźwięków lub obrazów.

Coraz bardziej rozpowszechnia się technika wszczepiania implantu ślimakowego. Znaczące osiągnięcia w tym zakresie ma Międzynarodowe Centrum Słuchu i Mowy w Kajetanach pod Warszawą. Osoby po zabiegu przywrócenia słuchu przechodzą wieloletnią rehabilitację polegającą na „uczeniu” się słyszenia z wykorzystaniem implantu słuchowego (który należy wielokrotnie przestawiać w miarę postępów uczenia).

W poprzednich rozdziałach książki nawiązano do głośnego przypadku Mikea Maya, któremu po 40 latach życia przywrócono wzrok utracony w wieku 3 lat. Mike nie rozpoznaje osób i przedmiotów, ma problemy w percepcji głębi. Po przebytej operacji wszczepienia rogówki regularnie bierze silne leki hamujące odrzucenie przeszczepu. Pomimo tego rodzaju trudności, wielkie nadzieje wiąże się z zabiegami medycznymi umożliwiającymi przywrócenie widzenia, do których należy zaliczyć (Wieland i Humayun, 2008):

- terapię genową, procedurę medyczną polegającą na wszczepianiu zdrowych komórek w obszar zmienionej chorobowo siatkówki,
- implanty elektroniczne zastępujące funkcje siatkówki; pierwsze udane wszczepienia przeprowadziły firmy Optobionics (www.optobionics.com) i SecondSight; prowadzone są prace w innych ośrodkach (www.retina-implant.de),
- bezpośrednią stymulację elektryczną nerwu wzrokowego,
- bezprzewodową stymulację elektryczną kory wzrokowej; w tym zakresie zaawansowane prace są prowadzone przez zespół prof. Sawana z Montrealu (Hashemi i inni, 2009); sygnałami dostarczanych do elektrod wszczepianych do kory mózgowej są odpowiednio przetworzone obrazy rejestrowane przez kamery cyfrowe.

Tylko niewielki procent osób niewidomych kwalifikuje się do ww inwazyjnych metod przywracania wzroku. Są to obecnie procedury bardzo drogie o trudnym do określenia terminie wdrożenia do szerszej praktyki medycznej.

W niniejszej monografii skupiono się na zagadnieniach dotyczących wspomagania orientacji przestrzennej i mobilności osób niewidomych z zastosowaniem elektronicznych systemów nawigacji osobistej. Nabycie umiejętności samodzielnego poruszania się jest warunkiem pokonania wielu barier ograniczających udział osób niewidomych w życiu społecznym i zawodowym, m.in. ich uzależnienia od pomocy innych.

Historia badań nad elektronicznymi systemami osobistej nawigacji dla osób niewidomych liczy ponad sto lat (Elektroftalm Noiszewskiego opracowany w 1898 r.). Duży postęp dokonał się w zakresie makronawigacji (nawigacji globalnej), w której skutecznie wykorzystuje się systemy globalnego pozycjonowania, mapy cyfrowe i programowane urządzenia mobilne (rozdz. 4 pracy). Natomiast zagadnienie mikro-

nawigacji (nawigacji lokalnej), tj. orientacji w przestrzeni bliskiej, omijanie przeszkód statycznych i ruchomych oraz bezpieczne poruszanie się w mieście okazało się trudnym do rozwiązania zagadnieniem badawczym. Obecnie dysponujemy wieloma skutecznymi technologiami pozyskiwania informacji o lokalnym otoczeniu, do których należą pasywne i aktywne systemy estymacji głębi i trójwymiarowej rekonstrukcji otoczenia, m.in. stereowizja (omówione dokładniej w rozdziale 2.). Podstawową trudnością jest natomiast odpowiednie przetworzenie, analiza i interpretacja danych pomiarowych. Jedno z rozwiązań tego problemu zaproponowano w Instytucie Elektroniki Politechniki Łódzkiej (IE PŁ). Opracowany prototyp systemu nawigacji wykorzystuje stereowizję, modelowanie scen trójwymiarowych i dźwięki przestrzenne do obrazowania otoczenia. Pierwsze testy prototypu przeprowadzone w warunkach kontrolowanych pokazały skuteczność tej techniki nawigacji.

Za szczególnie obiecujący należy uznać kierunek badań, w którym są stosowane algorytmy integracji danych makronawigacyjnych (dane GPS) i danych mikronawigacyjnych pozyskiwanych z czujników inercyjnych i obrazowych. Istnieją skuteczne algorytmy łączenia takich danych z zastosowaniem metod filtracji Kalmana i filtracji cząsteczkowej omówione w rozdziale 3.5.3. W tym zakresie omówiono również wyniki prac prowadzonych w IE PŁ, pokazujące znaczną poprawę wyznaczania położenia w terenie wysokiej zabudowy miejskiej. Użytkiwana obecnie dokładność pozycjonowania wynosząca kilka metrów nie jest wystarczająca do bezpiecznej nawigacji dla osób niewidomych, może być jednak z pożytkiem wykorzystana do dokładniejszego planowania ścieżek i łatwiejszego znajdowania miejsc użytecznych POI.

Niedocenianym aspektem projektowanych osobistych urządzeń wspomagania orientacji przestrzennej i nawigacji dla osób niewidomych jest interfejs użytkownika i ergonomia tych urządzeń. Metodom niewizualnej prezentacji informacji o otoczeniu poświęcono rozdział 4. Interfejsy dźwiękowe umożliwiają zastosowanie technik generacji dźwięków przestrzennych i budowanie „akustycznych obrazów” otoczenia. Wadą tych systemów jest zagłuszanie naturalnych dźwięków otoczenia, wśród których osoby niewidome potrafią skutecznie selekcjonować wskazówki nawigacyjne. W rozdziale 5. omówiono technikę sonifikacji, w której dla zmniejszenia liczby strumieni dźwiękowych dokonano selekcji obiektów otoczenia oraz zastosowano sekwencyjny sposób ich sonifikacji (z wykorzystaniem technologii HRTF) z kontrolowaną liczbą wytwarzanych strumieni dźwiękowych. Za pomocą tego systemu osoby niewidome skutecznie omijają przeszkody rozmieszczone w przypadkowych miejscach ścieżek testowych (Charłampowicz, 2011). W najbliższych latach można oczekiwać rozwoju technik dotykowej prezentacji otoczenia z zastosowaniem nowych materiałów zmieniających kształty, m.in. polimerów elektroaktywnych (rozd. 4.2.2).

Rozwijanym obecnie kierunkiem badań nad systemami nawigacji dla osób niewidomych są systemy o architekturze rozproszonej. Obserwuje się tu dwa podejścia: zastosowanie nawigacji lokalnej wg punktów orientacyjnych zwanych radiolatarniami oraz nawigację z zastosowaniem zdalnych przewodników. Pierwsza z wymienionych technik nawigacji ma szanse na szybkie rozpowszechnienie się dzięki dynamicznie rozwijającej się technologii układów identyfikacji radiowej RFID (rozdz. 5). Ich zastosowanie wewnątrz budynków w większym stopniu może pełnić funkcje informacyjne niż nawigacyjne. Drugi kierunek rozwoju systemów o architekturze rozproszonej to rozwiązania, w których osobiste urządzenie nawigacyjne osoby niewidomej łączy się bezprzewodowo z terminalem stacjonarnym lub osobistym zdalnego przewodnika. Przesyłane na odległość obrazy i dźwięki są modalnościami wykorzystywanymi przez zdalnego przewodnika do przekazywania słownych wskazówek nawigacyjnych osobie niewidomej. System tej klasy, opracowany w IE PŁ i opisany w rozdz. 5., jest od kilku miesięcy testowany przez niewidomych użytkowników. Badania te pokazały, że nawet krótki czas połączenia ze zdalnym przewodnikiem może być skuteczną pomocą nawigacyjną. Przewidywany wzrost szybkości transmisji danych w sieciach bezprzewodowych, powinien wkrótce poprawić użyteczność takich systemów.

Różnorodność stosowanych pomocy nawigacyjnych dla osób niewidomych, m.in. urządzeń ostrzegania o przeszkodach, systemów do nawigacji lokalnej oraz systemów nawigacji globalnej powoduje, że ich jednoczesne wykorzystanie przez osobę niewidomą nie jest możliwe ze względów ergonomicznych. Oczekiwanym kierunkiem rozwoju takich systemów jest odpowiednia integracja ich funkcji w postaci jednego, łatwego w obsłudze osobistego urządzenia nawigacyjnego. Prace nad takim systemem podjęto w roku 2010 w Instytucie Elektroniki Politechniki Łódzkiej w ramach grantu rozwojowego pt. *System ułatwiający samodzielne poruszanie się i dostęp do infrastruktury miejskiej dla osób niewidomych i słabowidzących, integrujący technologie bezprzewodowej transmisji danych oraz systemy nawigacji globalnej i lokalnej*. Projekt ten jest realizowany przy współpracy z Okręgiem Łódzkim Polskiego Związku Niewidomych oraz Miejskim Przedsiębiorstwem Komunikacyjnym w Łodzi.

Autor wraz z kilkunastoosobowym zespołem badawczym wyraża nadzieję, że zarówno monografia jak i wyniki opisanych prac, będą pożytecznym wkładem w wielowątkowy nurt badań prowadzonych w kraju i na świecie, które w niedługim czasie przyniosą wdrożenia konkretnych rozwiązań technicznych dla poprawienia jakości życia osób niewidomych i ich udziału w życiu społecznym i zawodowym.

- Adusei I.K., Kyamakya K., Jobmann K. (2002) Mobile positioning technologies in cellular networks: an evaluation of their performance metrics. *IEEE Proceedings of the Military Communications Conference*, vol. 2, pp. 1239–1244.
- Akenine-Moller T., Strom J. (2008) Graphics processing units for handhelds. *Proceedings of the IEEE*, vol. 96, no. 5, pp. 779–789.
- Algazi V.R., Duda R.O., Thompson D.M., Avendano C. (2001) The CIPIC HRTF database. *IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pp. 99–102.
- Amedi A., Raz N., Azulay H., Malach R., Zohary E. (2010) Cortical activity during tactile exploration of objects in blind and sighted humans. *Restorative Neurology and Neuroscience*, vol. 28, pp. 143–156.
- Anderson D.O., Moore J.B. (1984) *Filtracja optymalna*. WNT Warszawa.
- Angel E., Algazi R., Duda R. (2002) On the design of canonical sound localization environments. In: *Audio Engineering Society 113th Convention*.
- Ardito C., Costabile M.F., De Angeli A., Pittarello F. (2007) Navigation help in 3D worlds: some empirical evidences on use of sound. *Multimedia Tools and Applications*, vol. 33, no. 2, pp. 201–216, Springer.
- Arno P., De Volder A.G., Vanlierde A., Wanet-Defalque M.C., Streel E., Robert A., Sanabria-Bohórquez S., Veraart C. (2001) Occipital activation by pattern recognition in the early blind using auditory substitution of vision. *Neuroimage*, vol. 13, pp. 632–645.
- Arulampalam M.S., Maskell S., Gordon N., Clapp T. (2002) A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188.
- Bach-y-Rita P., Collins C.C., Sauters F., White B., and L. Scadden L. (1969) Vision substitution by tactile image projection. *Nature*, vol. 221, pp. 963–964.

- Bach-y-Rita P. (1972) *Brain Mechanisms in Sensory Substitution*. Academic Press.
- Bach-y-Rita P., Kercel S.W. (2003a) Sensory substitution and the human-machine interface. *Trends in Cognitive Neuroscience*, vol. 7, no. 12, pp. 541–546.
- Bach-y-Rita P., Tyler M.E., Kaczmarek K.A., (2003b) Seeing with the brain. *International Journal of Human–Computer Interaction*, vol. 15, no. 2, pp. 285–295.
- Balcerzak-Paradowska B. (red.) (2002) *Sytuacja osób niepełnosprawnych w Polsce*. Raport Instytutu Pracy i Spraw Socjalnych, Warszawa.
- Ball E. M. (2008) Electronic travel aids: an assessment. In: Hersh M.A. and Johnson M.A. (eds.) *Assistive Technology for the Visually Impaired and Blind People*, Springer-Verlag, London Limited, pp. 289–321.
- Barański P., Morański M. (2007) *System bezprzewodowej lokalizacji obiektów do wspomagania pracy architektów*. Praca magisterska, Instytut Elektroniki Politechniki Łódzkiej.
- Barański P., Strumiłło P. (2009a) Układ nawigacji zliczeniowej do korekcji odczytów GPS w terenie miejskim. *Przegląd Elektrotechniczny*, nr 11, str. 160–163.
- Barański P., Strumiłło P., Bujacz M., Materka A. (2009b) A remote guidance system aiding the blind in urban travel. *Proceedings of the Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Past Successes and Future Challenges CVHI 2009*, M.A. Hersh (ed.), Wrocław, Poland (conference–CD).
- Barański P., Polańczyk M., Strumiłło P. (2010a) A remote guidance system for the blind. *12th IEEE International Conference on e-Health Networking, Applications and Services* (HealthCom 2010), Lyon, pp. 386–390.
- Barański P., Bronakowski Ł., Strumiłło P. (2010b) System syntezy mowy polskiej do zastosowań w urządzeniach mobilnych. *IX Krajowa Konferencja Elektroniki*, Dąbrowa Tarnobrzeg, str. 376–381.
- Barański P., Polańczyk M., Strumiłło P. (2011) Fusion of data from inertial sensors, raster maps and GPS for estimation of pedestrian geographic location in urban terrain. *Metrology and Measurement Systems*, vol. 18, no.1, pp. 145–158.
- Batteau D.W. (1967). The role of the pinna in human localization. *Proceedings of the Royal Society B*, vol. 168, pp. 158–180.
- Bensky A. (2008) *Wireless Positioning Technologies and Applications*. Artech House, Inc.
- Bimber O., Raskar W., Raskar R. (2005) *Spatial Augmented Reality – Merging Real and Virtual World*. A.K. Peters, Ltd.
- Blythe D (2008) Rise of the graphics processor. *Proceedings of the IEEE*, vol. 96, no. 5, pp.761–778.
- Bohonos S., Lee A., Malik A., Thai C., Manduchi R. (2007) Universal real-time navigational assistance (URNA): an urban bluetooth beacon for the blind. *HealthNet'07*, San Juan, Puerto Rico, USA.
- Borenstein J., Everett H. R., Feng L. (1996) *Where am I? Sensors and Methods for Mobile Robot Positioning*. University of Michigan, Technical Report.
- Bourbakis N. (2008) Sensing surrounding 3-D space for navigation of the blind. *IEEE Engineering in Medicine and Biology Magazine*, pp. 49–55.

- Bradley N., Dunlop M. (2008) Navigation AT: context aware computing. In: Hersh M.A., Johnson M.A. (eds.) *Assistive Technology for Visually Impaired and Blind People*, Springer-Verlag, London Limited, pp. 231–260.
- Brambring M. (1985) Mobility and orientation processes of the blind. In: Warren D.H., Strelow E.R. (eds.), *Electronic Spatial Sensing for the Blind*, Dordrecht, The Netherlands, Nijhoff, pp. 493–508.
- Bregman A.S. (1990) *Auditory Scene Analysis*. Cambridge Press.
- BrainPort Vision Technology (2011) webpage: <http://vision.wicab.com>
- Bronzino J.D. (ed.) (1999) *The Biomedical Engineering Handbook*. Second Edition, CRC Press.
- Brown M.Z., Burschka D., Hager G.D. (2003) Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 993–1008.
- Bruss A.R., Horn K.P. (1983) Passive navigation, *Computer Vision, Graphics, and Image Processing*, vol. 21, pp. 3–20.
- Brzozowska-Rup K., Dawidowicz A.L. (2009) Metoda filtru cząsteczkowego. *Matematyka Stosowana: matematyka dla społeczeństwa*, tom. 10/51, str. 69–107.
- Bshara M., Orguner U., Gustafsson F., Van Biesen L. (2009) Fingerprinting localization in wireless networks based on received-signal-strength measurements: a case study on WiMAX networks. *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 283–294.
- Buczowski T., Radecki K. (2010) Urządzenia i systemy do wspomagania orientacji i nawigacji terenowej osób niewidomych. *Przegląd Telekomunikacyjny*, nr 12, str. 1732–1737.
- Bujacz M., Barański P., Morański M., Strumiłło P., A. Materka (2008a) Remote guidance for the blind – a proposed teleassistance system and navigation trials. *International Conference on Human System Interaction*, Cracow, Poland, pp. 888–892.
- Bujacz M., Barański P., Morański M., Strumiłło P., Materka A. (2008b) Remote mobility and navigation aid for the visually disabled. *Proceedings of the 7th International Conference on Disability, Virtual Reality and Associated Technologies*, in P.M. Sharkey, P. Lopes-dos-Santos, P.L. Weiss & A. L. Brooks (eds.), pp. 263–270, Maia, Portugal.
- Bujacz M. (2010) *Representing 3D Scenes through Spatial Audio in an Electronic Travel Aid for the Blind*. Ph.D. Thesis, Technical University of Lodz.
- Bujacz M., Pec M., Skulimowski P., Strumiłło P., Materka A. (2011) Sonification of 3D scenes in an electronic travel aid for the blind. In: Strumiłło P. (ed.), *Advances in sound localization*, InTech, Zagreb, Croatia, pp. 251–268.
- Calamia P. (1998) *Three-Dimensional Localization of a Close-Range Acoustic Source using Binaural Cues*. Master's thesis, University of Texas.
- Cappe O., Godsill S.J., Moulines E. (2007) An overview of existing methods and recent advances in sequential Monte Carlo. *Proceedings of the IEEE*, vol. 95, no. 5, pp. 899–924.

- Ceranka Sz., Niedźwiecki M. (2003) Application of particle filtering in navigation system for the blind. *Proceedings of the IEEE Seventh International Symposium on Signal Processing and its Applications*, pp. 495–498.
- Chang D., Nesbitt K.V., Wilkins, K. (2007) The Gestalt principles of similarity and proximity apply to both the haptic and visual grouping of elements. In *Proceedings of the 8th Australasian User Interface Conference*, Ballarat, Australia, ACS. pp. 79–86.
- Charłampowicz R., (2009) Elektroniczne pomoce w orientacji przestrzennej. *Tyfloświat* nr 1(3), str. 27–35.
- Charłampowicz R., (2010) Orientacja w mieście. *Tyfloświat*, nr 2(8), str. 32–37.
- Charłampowicz R., (2011) Naviton. *Tyfloświat*, nr 1(10), str. 29–32.
- Cheng C. I., Wakefield G. H. (2001) Introduction to head-related transfer functions (HRTFs): Representations of HRTFs in time, frequency, and space. *Journal of Acoustic Society of America*, vol. 49, pp. 231–249.
- Collingnon O., Davare M., Olivier E., De Volder A.G. (2009) Reorganization of the right occipito-parietal stream for auditory spatial processing in early blind humans: a transcranial magnetic stimulation study. *Brain Topography*, vol. 21, pp. 232–240.
- Collins C.C. (1985) On mobility aids for the blind. In: Warren D.H., Strelow E. R. (eds.): *Electronic Spatial Sensing for the Blind*, Dordrecht: Martinus Nijhoff, pp. 35–64.
- Coltheart M. (1980) Iconic memory and visible persistence. *Perception & Psychophysics*, vol. 27, no. 3, pp. 183–228.
- Cooper M., Petrie H. (2004) Three dimensional auditory display: issues in applications for visually impaired students. *Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display*, Sydney, Australia.
- Cyganek B. (2002) *Komputerowe przetwarzanie obrazów trójwymiarowych*. Akademicka Oficyna Wydawnicza EXIT, Warszawa.
- Darwin C.J., Turvey M.T., Crowder R.G. (1972) An auditory analogue of the sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology*, vol. 3, pp. 255–267.
- Davey B.L. (1968) *A Survey of History Philosophy and Research in Orientation and Mobility*. Old Dominion University, Norfolk, Virginia.
- Davidson A.J., Reid I.D., Molton N.D., Stasse O. (2007) MonoSLAM: real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 1052–1067.
- Davison A. (2003) Real-time simultaneous localization and mapping with a single camera. *IEEE International Conference on Computer Vision*, pp. 1403–1410.
- DeAngelis G.C. (2000) Seeing in three dimensions: the neurophysiology of stereopsis. *Trends in Cognitive Sciences*, vol. 4, no. 3, pp. 80–90.
- Department of Defense (2008) *Global Positioning System Standard Positioning Service Performance Standard*, 4th Edition.
- Diamantas S., Oikonomidis A., Crowder R. (2010) Depth estimation for autonomous robot navigation: a comparative approach. *IEEE International Conference on Imaging Systems and Techniques*, Thessaloniki, Greece.

- Dix A., Finlay J., Abowd G., Beale R. (2004) *Human Computer Interaction*. 3rd Edition Prentice Hall.
- Dobrucki A. (2007) *Przetworniki elektroakustyczne*. WNT, Warszawa.
- Dobrucki A., Plaskota P., Pruchnicki P., Pec M., Bujacz M., Strumillo P. (2010) Measurement system for personalized head-related transfer functions and its verification by virtual source localization trials with visually impaired and sighted individuals. *Journal Audio Engineering Society*, vol. 58, no. 9, pp. 724–738
- Duchowski A.T. (2003) *Eye Tracking Methodology: Theory and Practice*. Springer Verlag, London.
- Duda R.O. (1996) Auditory localization demonstrations. *Acustica/Acta Acustica*, vol. 82, pp. 346–355.
- EAV (2011) Espacio Acustico Virtual, web page: <http://www.iac.es/proyecto/eavi/english/>
- Enge P., Misra P. (1999) Special issue on Global Positioning System. *Proceedings of the IEEE*, vol. 187, no. 1.
- Ens J., Lawrence P. (1993) An investigation of methods for determining depth from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 2, pp. 97–108.
- Espinoza M.A., Ungar S., Ochaita E., Blades M., Spencer C. (1998) Comparing methods for introducing blind and visually impaired people to unfamiliar urban environments. *Journal of Environmental Psychology*, vol. 18, pp. 277–287.
- Evers-Senne J.-F., Koch R. (2005) Image-based rendering of complex scenes from a multi-camera rig. *IEE Proceedings Vision, Image and Signal Processing*, vol. 152, no. 4, pp. 470–480.
- Fang L., Antsaklis P.J., Montestruque L., McMickell M.B., Lemmon M., Sun Y., Fang H., Koutroulis I., Haenggi M., Xie M., Xie X. (2005) Design of a wireless assisted pedestrian dead reckoning system – The NavMote experience. *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 6, pp. 2342–2358.
- Farcy R., Leroux R., Jucha A., Damaschini R., Grégoire C., Zogaghi A. (2006) Electronic travel aids and electronic orientation aids for blind people: technical, rehabilitation and everyday life points of view. *Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments Technology for Inclusion*, CVHI, M.A. Hersh (ed.).
- Farrell J.A. (2008) *Aided Navigation*. McGraw-Hill Book Co., New York.
- Fastl H., Zwicker E. (1999) *Psychoacoustics: Facts and Models*. Springer-Verlag, Berlin.
- Fellgett P. (1975) Ambisonics: general system description. *Studio Sound*, vol. 40, no. 1, pp. 20–22.
- Filliat D., Meyer J.-A. (2003) Map-based navigation in mobile robots. I.: a review of localization strategies. *Cognitive Systems Research*, vol. 4, no. 4, pp. 243–282.
- Fine I., Wade A.R., Brewer A.A., May M.G., Goodman D.F., Boynton G.M., Wandell B.A., MacLeod D.I. (2003) Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, vol. 6, no. 9, pp. 915–916.

- Fleet D.J., Weiss Y. (2005) Optical flow estimation. In: Paragios N. (ed.) *Handbook of Mathematical Models in Computer Vision*, Springer, pp. 238–258.
- Foulke E. (1971) The perceptual basis for mobility. *American Foundation for the Blind Research Bulletin*, vol. 23, pp. 1–8
- Freedman S.J., Fisher H.G. (1968) The role of the pinna in auditory localization. In: S.J. Freedman (ed.) *Neuropsychology of Spatially Oriented Behavior*, Dorsey Press, Illinois.
- FRS (2011) Forehead Retina System, Technology.com webpage: www.technovelgy.com/ct/Science-Fiction-News.asp?NewsNum=708
- Fukushima K. (1980) Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, vol. 36, no. 4, pp. 93–202.
- Funkhouser T., Jot J-M., Tsingos N. (2002) Sounds good to me! *Computational Sound for Graphics, Virtual Reality, and Interactive Systems. SIGGRAPH Course Notes*.
- Fusiello A., Trucco E., Verri A. (2000) A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, vol. 12, pp. 16–22.
- Gallistel C.R. (1990) *The organization of learning*, MIT Press. Cambridge, MA.
- Garaj V., Jirawimut R., Ptasiński P., Cecelja F., Balachandran W. (2003) A system for remote sighted guidance of visually impaired pedestrians. *British Journal of Visual Impairment*, vol. 21, pp. 55–63.
- Garcia C., Tziritas G. (2001) 3D Translational motion estimation from 2D displacements. *Proceedings of International Conference on Image Processing*, vol. 2, pp. 945–948.
- Gardner M.B., Gardner R.S. (1973) Problem of localization in the median plane: effect of pinnae cavity occlusion. *The Journal of the Acoustical Society of America*, vol. 53, no. 2, pp. 400–408.
- Gardner M.B., Martin K. (1994) HRTF measurements of a KEMAR dummy-head microphone. *MIT Media Lab Perceptual Computing - Technical Report*, no. 280.
- Gärling T., Book A., Lindberg E., Nilsson T. (1981) Memory for the spatial layout of the everyday physical environment: Factors affecting rate of acquisition. *Journal of Environmental Psychology*, vol. 1, pp. 263–277.
- GDP Research (2011) web page: www.gdp-research.com.au
- Gerrig R.J., Zimbardo P.G. (2009) *Psychologia i życie*. WNT, Warszawa.
- Gibson, J.J. (1979) *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin
- Gonzalez-Mora J.L., Rodriguez-Hernandez A., Rodriguez-Ramos L.F., Diaz-Saco L., Sosa N. (1999) *Development of a New Space Perception System for Blind People, Based on the Creation of a Virtual Acoustic Space: Engineering Applications of Bio-Inspired Artificial Neural Networks*. Springer Berlin/Heidelberg, pp. 321–330.
- Golledge R.G. (1999) (ed.) *Wayfinding Behaviour: Cognitive Mapping and other Spatial Processes*. The John Hopkins University Press.
- Granlund G.H. (1999) The complexity of vision. *Signal Processing*, vol. 74, pp. 101–126.

- Granosik G. (2011) *Hypermobile Robots: Concept, Construction, and Control*. Rozprawa habilitacyjna, Zeszyty Naukowe Politechniki Łódzkiej, nr 1088.
- Gregory R.L. (2005) Knowledge for vision: vision for knowledge. *Philosophical Transactions of the Royal Society B*, vol. 360, pp. 1231–1251.
- Grejner-Brzezińska D.A., Toth C.K., Moafipoor S. (2008) Performance assesment of a multi-sensor personal navigator supported by an adaptive knowledge based system. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, part B5, Beijing, pp. 857–866.
- Gu Y., Lo A., Niemegeers I. (2009) A survey of indoor positioning systems for wireless personal networks. *IEEE Communications Surveys & Tutorials*, vol. 11, no. 1, pp. 13–31.
- Gustafsson F., Gunnarsson F., Bergman N., Forssell U., Jansson J., Karlsson R., Nordlund P.-J. (2002) Particle filters for positioning, navigation and tracking. *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 425–437.
- Hamilton R.H., Pascual-Leone A. (1998) Cortical plasticity associated with Braille learning. *Trends in Cognitive Neuroscience*, vol. 2, no. 5, pp. 168–174.
- Hartley R., Zisserman A. (2006) *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge.
- Hartmann W.H., Wittenberg A. (1966) On the externalization of sound images. *Journal of Acoustic Society of America*, vol. 99, pp. 3678–3688.
- Hashemi A.B. Saeid S., Sawan M., Savaria Y. (2009) A novel low-drop voltage CMOS active rectifier for RF powered devices: experimental results. *Elsevier Microelectronics Journal*, vol. 40, no. 11, pp. 1547–1554
- Hausman S. (2007) *Modelowanie propagacji fal radiowych do projektowania naziemnych systemów łączności bezprzewodowej*. Rozprawa habilitacyjna, Zeszyty Naukowe Politechniki Łódzkiej, nr 1009.
- Helal A.A., Mokhtari M., Abdulrazak B. (eds.) (2008) *The Engineering Handbook of Smart Technology for Aging, Disability and Independence*. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Hersh M.A., Johnson M.A (eds.) (2008) *Assistive Technology for Visually Impaired and Blind People*. Springer-Verlag, London Limited.
- Heyes D.A. (1984) The Sonic Pathfinder: a new electronic travel aid. *Journal of Visual Impairment and Blindness*, vol. 77, no. 2, pp. 200–202.
- Hollins M. (1989) *Understanding Blindness: a Multidisciplinary Approach*. Lawrence Erlbaum Associates Publishers, Hillsdale, New Jersey.
- Horn B.K.P., Schunck B.G. (1981) Determining optical flow. *Artificial Intelligence*, vol. 17, pp. 185–203.
- Hoyle B., Waters D. (2008) Mobility AT: The Batcane (UltraCane). In: Hersh M.A., Johnson M.A. (eds.) *Assistive technology for the visually impaired and blind people*, Springer-Verlag, London Limited, pp. 209–230.

- Hunaiti Z., Garaj V., Balachandran W., Cecelja F. (2005) Use of remote vision in navigation of visually impaired pedestrians. *International Congress Series*, vol. 1282, pp. 1026–1030.
- ICAD (2011) *17th Annual Conference on Auditory Display*, Budapest, Hungary, web page: <http://www.icad2011.com>
- IE PŁ (2007) *Zintegrowany system wspomagający osobę niewidomą, z dźwiękowym interfejsem użytkownika*. Raport grantu badawczego MNiSW nr 3T11B 038 27, Instytut Elektroniki Politechniki Łódzkiej.
- IE PŁ (2010) *Projekt i badania prototypu mobilnego systemu wspomagania osoby niewidomej w samodzielnym poruszaniu się w terenie miejskim z wykorzystaniem obrazowania dźwiękowego otoczenia, systemu nawigacji satelitarnej i mapy elektronicznej miasta*. Raport grantu rozwojowego MNiSW nr R02 013 03, Instytut Elektroniki Politechniki Łódzkiej.
- ITC (1996) *Independent Television Commission*. Guidance on Standards for Audio Description.
- Jansson G. (2008) Haptics as a substitute for vision. In: Hersh M.A., Johnson M.A. (eds.) *Assistive Technology for the Visually Impaired and Blind People*, Springer-Verlag, London Limited, pp. 135–166.
- Jezierski E. (2005) *Dynamika robotów*, WNT, Warszawa.
- Jirawimut R., Ptasiński P., Garaj V. (2003) A method for dead reckoning parameter correction in pedestrian navigation system. *IEEE Transactions on Instrumentation and Measurement*, vol. 52, no. 1, pp. 209–215.
- Kabsh A. (red.) (2009) *Konwencja praw osób z niepełnosprawnościami: godność i sprawiedliwość dla nas wszystkich*. Zeszyty Promocji Rehabilitacji, zeszyt 3, Ośrodek Wydawnictw Naukowych, Poznań.
- Kaczmarek K.A. (1999) Sensory augmentation and substitution. In: Bronzino J.D. (eds.) *The Biomedical Engineering Handbook: Second Edition*, CRC Press, chapt. 143.
- Kadaj R. (2006) *Zasady zastosowania metody transformacyjnej do przeliczeń punktów z układu „1965” lub lokalnego do układu „2000”*. Strona internetowa Głównego Urzędu Geodezji i Kartografii: www.gugik.gov.pl
- Kalat J.W. (2011) *Biologiczne podstawy psychologii*. Wydawnictwo Naukowe PWN, Warszawa.
- Kalman R.E. (1960) A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45.
- Kato M., Uematsu H., Kashino M., Hirahara T. (2003) The effect of head motion on the accuracy of sound localization. *Acoustical Science and Technology*, vol. 24, no. 5, pp. 315–317.
- Kay L. (2011) Bay Advanced Technologies ‘K’ Sonar, web page: www.batforblind.co.nz
- Kinałski R. (2008) *Neurofizjologia kliniczna dla neurorehabilitacji*. Podręcznik dla studentów i absolwentów wydziałów fizjoterapii, Wydawnictwo MedPharm Polska, Wydanie I, Wrocław, str. 193–289.

- Klatzky R.L., Beall A.C., Loomis J.M., Golledge R.G., Philbeck J.W. (1999) Human navigation ability: tests of the encoding-error model of path integration. *Spatial Cognition and Computation*, vol. 1, pp. 31–65.
- Klatzky R.L., Lederman S.J. (2007) Object recognition by touch. In: Rieser J., Ashmeed D., Ebner F., Com A. (eds.) *Blindness and Brain Plasticity in Navigation and Object Perception*, Mahwah, Erlbaum.
- Klawiter A., Preis A. (2006) Percepcja słuchowa przedmiotów, szkic teorii i jej testowanie. *Kolokwia Psychologiczne*, nr 14, Instytut Psychologii PAN, Warszawa.
- Klein G.S.W., Drummond T.W. (2004) Tightly integrated sensor fusion for robust visual tracking. *Image and Vision Computing*, vol. 22, pp. 769–776.
- Koch R. (1993) Dynamic 3D scene analysis through synthesis feedback control. *IEEE Transaction on Pattern Analysis Machine Intelligence*, vol. 15, no. 6, pp. 556–568.
- Koffka K. (1935) *Principles of Gestalt psychology*. New York: Harcourt Brace.
- Konieczny J., Domański M. (2010) Depth-based inter-view prediction of motion vectors for improved multiview video coding. *3DTV-Conference: The True Vision – Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp.1–4.
- Konorski J. (1948) *Conditioned Reflexes and Neuron Organization*. Cambridge Univ. Press, Cambridge.
- Korbicz J., Obuchowicz A., Uciński D. (1994) *Sztuczne sieci neuronowe: podstawy i zastosowania*. Akademicka Oficyna Wydawnicza PLJ, Warszawa.
- Kramer G. (ed.) (1994) *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Santa Fe Institute Studies in the Sciences of Complexity, vol. XVIII. Reading, MA: Addison-Wesley.
- Krawczyk-Pasławska E. (2010) *Niepelnosprawność: bariery i szanse*. Fundacja Bariera, Kraków.
- Kulyukin V., Gharpure C., John Nicholson J. (2004) RFID in Robot-Assisted Indoor Navigation for the Visually Impaired. *Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2, pp. 1979–1984.
- Kurczyński Z., Preuss R. (2002) *Podstawy fotogrametrii*. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa.
- Kurek A., Czerwieński K. (2003) System orientacji terenowej dla osób niewidomych. *Krajowa Konferencja Radiokomunikacji, Radiofonii i Telewizji*, Wrocław.
- Kurson R. (2008) *Odzyskać wzrok*. Świat Książki, Warszawa.
- Kwapiszowie J., J. (1990) *Orientacja przestrzenna i poruszanie się niewidomych oraz słabowidzących: poradnik metodyczny*. Wydawnictwo Szkolne i Pedagogiczne, Warszawa.
- La Grow S. (1999) The use of the Sonic Pathfinder as a secondary mobility aid for travel in business environments: a single-subject design. *Journal of Rehabilitation Research and Development*, vol. 36, no. 4, pp. 1–13.
- Lange R. (2000) *3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-technology*. PhD dissertation, University of Siegen.
- Lange R., Seitz P. (2001) Solid-state time-of-flight range camera. *IEEE Journal of Quantum Electronics*, vol. 37, no. 3, pp. 390–397.

- Langley R.B. (1999) Dilution of precision. *GPS World*, pp. 52–59.
- Lecewicz-Bartoszewska J., Polak-Sopińska A. (red.) (2008) *Ergonomia niepełnosprawnym w organizacji pracy i zarządzaniu – projektowanie*. Media Press, Łódź.
- Le Moigne J.J., Waxman A.M. (1988) Structured light patterns for robot mobility. *IEEE Transactions on Robotics and Automation*, vol. 4, no. 5, pp. 541–548.
- Lederman S.J., Klatzky R.L. (1987) Hand movements: a window into haptic object recognition. *Cognitive Psychology*, vol. 19, pp. 342–368.
- Lederman S.J., Klatzky R.L. (2009) Haptic perception: a tutorial. *Attention, Perception & Psychophysics*, vol. 71, no. 7, pp. 1439–1459.
- Lenay C., Gapenne O., Hanneton S., Marque C., Geouelle C. (2003) Sensory substitution: limits and perspectives, touching for knowing, *Cognitive Psychology of Haptic Manual Perception*, pp. 275–292.
- Leonard J., Durrant-Whyte H.F. (1991) Mobile robot localization by tracking geometric beacons. *IEEE Transactions on Robotics and Automation*, vol. 7, no. 3, pp. 376–382.
- Leroy J.V., Simon T., Deschenes F. (2008) An efficient method for monocular depth from defocus. *Proceedings of the 50th International Symposium ELMAR*, vol. 1, pp. 133–136.
- Lewandowski J., Boczkowska K. (red.) (2008) *Ergonomia niepełnosprawnym w organizacji pracy i zarządzaniu – badania, oceny, wytyczne*. Media Press, Łódź.
- Li B., Salter J., Dempster A.G., Rizos C. (2006) Indoor positioning techniques based on wireless LAN. *Proceedings of the 1st IEEE International Conference on Wireless Broadband and Ultra Wideband Communications*, Sydney, Australia, CD ROM, paper 113.
- Liu Y., Yu C., Liang M., Tian L., Zhou Y., Qin W., Li K., Jiang T. (2007) Whole brain functional connectivity in the early blind. *Brain*, vol. 130, pp. 2085–2096.
- Lokki T., Savioja L., Vaananen R., Huopaniemi J., Takala T. (2003) Creating interactive virtual auditory environments. *IEEE Computer Graphics and Applications*, vol. 22, no. 4, pp. 49–57.
- Longstaff A. (2006) *Krótkie wykłady, neurobiologia*. Wydawnictwa Naukowe PWN, Warszawa.
- Loomis J.M. (1992) Distal attribution and presence. *Forum: Spotlight on: the Concept of Telepresence*, vol. 1, no. 1, pp. 113–119.
- Loomis J.M., Klatzky R.L., Golledge R.G., Philbeck J.W. (1999) Human navigation by path integration. In: Golledge R.G. (ed.) *Wayfinding: Cognitive Mapping and other Spatial Processes*, Baltimore: Johns Hopkins, pp. 125–151.
- Loomis J.M., Golledge R.G., Klatzky R.L. (2001) GPS-based navigation systems for the visually impaired. In: Barfield W., Caudell T. (eds.) *Fundamentals of Wearable Computers and Augmented Reality*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 429–446.
- Lowe D.G. (2001) Local feature view clustering for 3D object recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, Kauai, Hawaii, pp. 682–688.
- Lucas B.D., Kanade T. (1981) An iterative image registration technique with an application to stereo vision. *Proceedings of Imaging Understanding Workshop*, pp. 121–130.

- Lynch K. (1960) *The Image of the City*, MIT Press, Cambridge MA.
- Ma Y., Soatto S., Kosecka J., Sastry S. (2003) *An Invitation to 3D Vision*, Springer.
- Malvern B., Nazir A. (1973) An improved laser cane for the blind developments in laser technology II. *SPIE Proceedings*, vol. 41.
- Marr D. (1982) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. MIT Press.
- Marski J., Bajurko P., Radecki K., Buczkowski T. (2010) Miniaturowe radiolatarnie i terminale z sygnalizacją RSSI do wspomagania orientacji osób niewidomych. *Przegląd Telekomunikacyjny-Wiadomości Telekomunikacyjne*, vol. 79, nr 6, str. 320–324.
- Maruszewski T. (2002) *Psychologia poznania*. Gdańskie Wydawnictwo Psychologiczne, Gdańsk.
- Massaro D.W. (1999) Speechreading: illusion or window into pattern recognition. *Trends in Cognitive Sciences*, no. 3, pp. 310–317.
- Matacz M. (2011) Jazda bez patrzenia. *Focus*, Sierpień, str. 88–90.
- Materka A., Strumiłło P. (2009) *Wstęp do komputerowej analizy obrazów*. Politechnika Łódzka.
- McGookin D., Cole R., Brewster S. (2010) Virtual navigator: developing a simulator for independent route learning. *Workshop on the Haptic Audio Interaction Design*, Copenhagen, Denmark, pp. 42–44.
- McGookin D.K., Brewster S.A. (2004) Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition. *ACM Transactions on Applied Perception*, vol. 1, no. 2, pp. 130–155.
- McNamara T.P., Shelton A.L. (2002) Cognitive maps and the hippocampus. *Trends in Cognitive Sciences*, vol. 7, no. 8, pp. 333–335.
- Meijer P. (1992) An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 2, pp. 112–121.
- Merabet L., Poggel D., Stern W., Bhatt E., Hemond C., Maguire S., Meijer P., Pascual-Leone A. (2008) Activation of visual cortex using crossmodal retinotopic mapping. *The 14th Annual Meeting of the Organization for Human Brain Mapping (OHBM)*, Melbourne, Australia.
- Meuth R.J. (2007) GPUs surpass computers at repetitive calculations. *IEEE Potentials*, pp. 12–15.
- Mezentsev O., Lachapelle G., Collin J. (2005) Pedestrian dead reckoning – a solution to navigation in GPS signal degraded areas. *Geomatica*, vol. 59, no. 2, pp. 175–182.
- Mikołajewska E., Mikołajewski D. (2011) Zastosowanie medyczne systemów Ambient Intelligence, *Inżynieria Biomedyczna*, vol. 17, nr 3, str. 207–210.
- Milella A., Siegwart R. (2006) Stereo-based ego-motion estimation using pixel tracking and iterative closest point. *Proceedings of the IEEE International Conference on Computer Vision Systems ICVS '06*, pp. 21–21.
- Millar S. (2002) *Understanding and Representing Space: Theory and Evidence from Studies with Blind and Sighted Children*. Oxford University Press.

- Milner A.D., Goodale M.A. (2008) *Mózg wzrokowy w działaniu*. Wydawnictwo Naukowe PWN.
- Misra P., Burke B.P., Pratt M.M. (1999) GPS performance in navigation. *Proceedings of the IEEE*, vol. 187, no. 1, pp. 65–85.
- Modsching M., Kramer R., Hagen K. (2006) Field trial on GPS accuracy in a medium size city: the influence of built-up. *Proceedings of the 3rd Workshop on Positioning, Navigation and Communication (WPNC'06)*, pp. 209–218.
- Moore B.C. (2004) *Wprowadzenie do psychologii słyszenia*. Państwowe Wydawnictwa Naukowe, Warszawa.
- Morański M., Materka A. (2011) *A Haptic Interface for Presenting 3D Objects in Virtual Reality for the Blind*. Internal report, Institute of Electronics, Technical University of Lodz, Poland.
- Morston J. (2009) Researching end-user requirements with application to wayfinding technologies. *Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments*, (CD-edition).
- Muhlmann K., Maier D., Hesser J., Manner R. (2002) Calculating dense disparity maps from color stereo images, an efficient implementation. *International Journal on Computer Vision*, vol. 47, no. 1–3, pp. 79–88.
- Munkong R., Juang B.H. (2008) Auditory perception and cognition. *IEEE Signal Processing Magazine*, vol. 25, no. 3, pp. 98–117.
- Na J. (2006) The blind interactive guide system using RFID-based indoor positioning system. In *Computers Helping People With Special Needs, Lecture Notes in Computer Science*, vol. 4061, pp. 1298–1305.
- Narkiewicz J. (1999) *Podstawy układów nawigacyjnych*, WKiŁ, Warszawa.
- Nęcka E., Orzechowski J., Szymura B. (2006) *Psychologia poznawcza*. Wydawnictwo Naukowe PWN, Warszawa.
- Nesbitt K. (2005) A framework to support the designers of haptic, visual and auditory displays. *GOTHI'05 Guidance on Tactile and Haptic Interactions*, Saskatoon, Saskatchewan, Canada.
- NFB (2011) National Federation of the Blind, GPS Technology for the blind, A Product Evaluation. Braille Monitor, web page: www.nfb.org/Images/nfb/Publications/bm/bm06/bm0602/bm060206.htm
- Ni L.M., Zhang D. (2011) RFID-based localization and tracking technologies. *IEEE Wireless Communications*, pp. 45–51.
- Nister D., Naroditsky O., Bergen J. (2004) Visual Odometry. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol.1, pp. I-652–I-659.
- NOPPA (2011) *Personal Navigation and Information system for Users of Public Transport*, web page: <http://virtual.vtt.fi/virtual/noppa/noppaeng.htm>
- Nowakowski J., Podsędkowski L. (2002) Analiza obrazu ze skanera laserowego do identyfikacji przeszkód i planowania trajektorii robotów mobilnych. *X Konferencja: Sieci i Systemy Informatyczne*, Łódź, str. 583–588.

- Ohuchi M., Iwaya Y., Suzuki Y., Muneakata T. (2006) Cognitive-map formation of blind persons. In: *Proceedings of the 12th International Conference on Auditory Display*, London, UK.
- Paneels S., Roberts J.C. (2010) Review of designs for haptic data visualization. *IEEE Transactions on Haptics*, vol. 3, no. 2, pp. 119–137.
- Paplińska M. (red.) (2008) *Edukacja równych szans*. Uniwersytet Warszawski.
- Paradiso J.A., Landlay J.A. (2009) Guest editors introduction to special theme issue: Cross-reality environments, *Pervasive Computing*, July–September.
- Pec M., Bujacz M., Strumiłło P. (2007) Personalized head related transfer function measurement and verification through sound localization resolution. *Proceedings of the 15th European Signal Processing Conference (EUSIPCO 2007)*, str. 2326–2330.
- Pec M. (2010) *Modelowanie transmitancji akustycznej głowy przy użyciu filtrów cyfrowych małego rzędu*. Rozprawa doktorska, Politechnika Łódzka, 2010.
- Pelczyński P., Strumiłło P. (2005) Szybkie wyznaczanie głębi w scenie trójwymiarowej ze stereoskopii. *Materiały XIII Konferencji Sieci i Systemy Informatyczne*, Łódź, str. 501–507.
- Pelczyński P., Strumiłło P., Bujacz M., (2006) Formant-based speech synthesis in auditory presentation of 3d scene elements to the blind. *ACOUSTICS High Tatras, 33rd International Acoustical Conference – EAA Symposium*, Štrbské Pleso, Slovakia, October 4th–6th, 2006, pp. 346–349.
- Pelczyński P., Ostrowski B., Rzeszotarski D. (2009) Mobilny system pasywnej nawigacji w scenie trójwymiarowej. *VIII Krajowa Konferencja Elektroniki KKE 2009*, Darłówko Wschodnie, str. 682–687.
- Pelczyński P., Ostrowski B., Rzeszotarski D. (2011) Wyznaczanie wektora ruchu kamery stereowizyjnej z czujnikami inercyjnymi. *X Krajowa Konferencja Elektroniki KKE 2011*, Darłówko Wschodnie.
- Pentland A.P. (1982) Depth of scene from depth of field. *Proceedings of the Image Understanding Workshop (Palo Alto, CA)*, pp. 253–259.
- Petrie H., Johnson V., Strothotte T., Raab A., Fritz S., Michel R. (1996) MoBIC: designing a travel aid for blind and elderly people, *Journal of Navigation*, vol. 49, no. 1, pp. 45–52.
- Piekara A. (1976) *Nowe oblicze optyki*. Państwowe Wydawnictwo Naukowe, Warszawa.
- Pissaloux E., Maingreud F., Velazquez R., Fontaine E. (2007) Concept of the walking cognitive assistance: experimental validation. *ASME Int. Journal on Advances in Modeling*, C, vol. 67, pp. 75–86.
- Pizlo Z. (2008) *3D shape: its unique place in visual perception*. The MIT Press.
- Plaskota P. (2006) *Symulacja HRTF uwzględniająca akustyczne parametry głowy*. Praca doktorska, Politechnika Wrocławska.
- Ptito M., Matteau I., Gjedde A., Kupers R. (2009) Recruitment of the middle temporal area by tactile motion in congenital blindness. *NeuroReport*, vol. 20, pp. 543–547.
- Radecki K., Łukaszewicz K., Kulikowski J. (2004) Application of miniature FSK transmitters for improving the orientation of the blind person in urban environment. *Pro-*

- ceedings of the TRANSED 2004: The 10th International Conference on Mobility and Transport for Elderly and Disabled People*, Hamamatsu, Japan, pp. 833–840.
- Reed S.K. (1972) Pattern recognition and categorization. *Cognitive Psychology*, vol. 3, no. 3, pp. 382–407.
- Remondino F., El-Hakim S. (2006) Image-based 3D modeling: a review. *The Photogrammetric Record*, vol. 21, no. 115, pp. 269–291.
- Ritzler S. (2011) Mygo Cane for the blind to replace guide dog web page: www.invention-reaction.com/new-inventions/Mygo-Cane-For-The-Blind-To-Replace-Guide-Dog
- Robles-Kelly A., Hancock E.R. (2007) Shape-from-shading using the heat equation. *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 7–21.
- Roentgen U.R., Gelderblom G.J., Soede M., de Witte L. P. (2008) Inventory of electronic mobility aids for persons with visual impairments: a literature review. *Journal of Visual Impairment & Blindness*, vol. 102 (on-line access).
- Rosten E., Drummond T. (2005) Fusing points and lines for high performance tracking. *10-th IEEE International Conference on Computer Vision*, vol. 2, pp. 1508–1515.
- Rouben A., Terveen L. (2007) Speech and non-speech audio: navigational information and cognitive load. *Proceedings of the 13th International Conference on Auditory Display*, Montreal, Canada, pp. 468–475.
- Roumeliotis S.I., Johnson A.E., Montgomery J.F. (2002) Augmenting Inertial Navigation with Image-Based Motion Estimation. *Proceedings of IEEE International Conference on Robotics and Automation*, vol. 4, pp. 4326–4333.
- Rzeszotarski D., Strumiłło P., Pelczyński P., Więcek B., Lorenc A. (2005) System obrazowania stereoskopowego sekwencji scen trójwymiarowych. *Zeszyty Naukowe Elektrotechnika*, nr 10, str. 165–184.
- Rzeszotarski D., Skulimowski P., Strumiłło P. (2007) A method for verification of dense disparity maps computed from the matching algorithm implemented in the stereo-vision system. *Proceedings of SPIE*, 69372R.
- Sánchez J., Lumbrera M. (1999) Virtual environment interaction through 3D audio by blind children. *Journal of Cyberpsychology and Behavior*, vol. 2, pp. 101–111.
- Santosh D., Achar S., Jawahar C.V. (2008) Autonomous image-based exploration for mobile robot navigation. *IEEE International Conference on Robotics and Automation, ICRA 2008*, pp. 2717–2722.
- Sawicki D. (2007) Nielambertowskie właściwości emisyjne źródła światła. *Przegląd Elektrotechniczny*, nr 1, str. 51–54.
- Saxena A., Chung S.H., Ng A.Y. (2007) 3-D depth reconstruction from a single still image. *International Journal of Computer Vision*, vol. 76, no. 1, pp. 53–69.
- Saxena A., Sun M., Ng A.Y. (2009) Make3D: Learning 3D scene structure from a single still image. *IEEE Transactions of Pattern Analysis and Machine Intelligence (PAMI)*, vol. 30, no. 5, pp. 824–840.
- Sayood K. (2002) *Kompresja danych – wprowadzenie*. Wydawnictwo RM, Warszawa.

- Scharstein D., Szeliski R. (2002) A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal on Computer Vision*, vol. 47, no. 1–3, pp. 7–42.
- Selfridge O.G., Neisser U. (1960) Pattern recognition by machine. *Scientific American*, vol. 203, pp. 60–68.
- Shamma S. (2001) On the role of space and time in auditory processing. *Trends in Cognitive Sciences*, vol. 5, no. 8, pp. 340–342.
- Shi J., Tomasi C. (1994) Good features to track. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 593–600.
- Shih R., Dubrowski A., Carnahan H. (2009) Evidence for haptic memory. In: *Proceedings of the World Haptic Conference*, pp. 145–149.
- Shoval S., Borenstein J., Koren, Y. (1998) Auditory guidance with the Navbelt – a computerized travel aid for the blind. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 28, no. 3, pp. 459–467.
- Singh J.M., Narayanan P.J. (2010) Real-time ray tracing of implicit surfaces on the GPU. *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 2, pp. 261–272.
- SKERI (2011) Smith-Kettlewell Eye Research Institute web page: www.ski.org
- Skulimowski P., Strumiłło P. (2006) Detekcja płaszczyzn w sekwencji obrazów stereowizyjnych. *V Sympozjum Naukowe Techniki Przetwarzania Obrazu*, str. 282–287.
- Skulimowski P., Strumiłło P. (2007) Obstacle localization in 3D scenes from stereoscopic sequences. *15th European Signal Processing Conference EUSIPCO 2007*, pp. 2095–2099.
- Skulimowski P., Strumiłło P. (2008) Refinement of depth from stereo camera ego-motion parameters. *Electronics Letters*, vol. 12, pp. 729–730.
- Skulimowski P. (2009a) *Segmentacja i opis parametryczny obiektów w sekwencjach obrazów z ruchomego układu stereowizyjnego do celów dźwiękowej prezentacji sceny trójwymiarowej*. Rozprawa doktorska, Politechnika Łódzka.
- Skulimowski P., Świątczak T. (2009b) Urządzenie mobilne do synchronizowanego zbierania danych z kompasu elektronicznego, odometru i odbiornika GPS. *Elektronika*, nr 10, str. 80–83.
- Skulimowski P., Sulmowski D. (2011) System przechowywania danych i zarządzania punktami charakterystycznymi terenu. *Przegląd Telekomunikacyjny*, nr 8–9, str. 752–754.
- SLAB (2011) Spatial Auditory Displays Lab, web page: <http://human-factors.arc.nasa.gov/SLAB/>
- Stasiak A., Materka A. (2003) Zastosowanie sztucznej sieci neuronowej do oceny położenia kamery w systemach wizyjnych o skończonej głębi ostrości. *Elektronika*, str. 119–146.
- Steven M.S., Blakemore C. (2004) Cortical plasticity in the adult human brain. In: Gazzaniga M.S. (ed.) *The Cognitive Neurosciences*, Cambridge, MA: MIT Press, pp. 1243–1254.
- Strumiłło P., Pelczyński P., Bujacz M., Pec M. (2006) Space perception by means of acoustic images: an electronic travel aid for the blind, ACOUSTICS High Tatras. *33rd International Acoustical Conference – EAA Symposium*, Slovakia, pp. 296–299.

- Strumiłło P., Szajerman D., Pełczyński P., Materka A. (2009a) Implementation of stereo matching algorithms on graphics processing units. In: Choraś R.S., Zabłudowski A. (eds.): *Image Processing & Communications Challenges*, Academy Publishing House EXIT, Warsaw, pp. 286–293.
- Strumiłło P., Skulimowski P., Polańczyk M. (2009b) Programming Symbian smartphones for the blind and visually impaired. In: Kącki E., Rudnicki M., Stempeżyńska J. (eds.): *Computers in Medical Activity, Advances in Intelligent and Soft Computing*, Springer 2009, pp. 129–136.
- Strumiłło P. (2010) Electronic interfaces aiding the visually impaired in environmental access, mobility and navigation. *IEEE Conference on Human System Interaction*, 13–15 May, Rzeszów, pp. 17–24.
- Strumiłło P. (ed.) (2011) *Adances in Sound Localization*. InTech Publishing House, Zagreb, Croatia.
- Strumiłło P., Barański P., Polańczyk M. (2011a) Systemy telenawigacji dla osoby niewidomej. *Przegląd Telekomunikacyjny*, nr 8–9, str. 730–732.
- Strumiłło P., Materka A., Królak A. (2011b) Systemy interakcji człowiek–komputer dla osób niepełnosprawnych. *Biuletyn Techniczno-Informacyjny Stowarzyszenia Elektryków Polskich*, nr 1, str. 2–9.
- Strumiłło P., Materka A. (2011c) Systemy teleinformatyczne i urządzenia elektroniczne dla niewidomych. *Przegląd Telekomunikacyjny*, nr 8–9, str. 710–712.
- Szeliski R., Scharstein D. (2002) Symmetric sub-pixel stereo matching. *7th European Conference on Computer Vision*, Copenhagen, Denmark, vol. 2, pp. 525–540.
- Szmal P., Suszczańska N. (2002) Translating texts into sign language in the TGT–1 system: principles and realization. In: Demenko G., Karpiński M., Jassem, K. (eds.) *Speech and Language Technology*, vol. 6, pp.113–124. (thetos.polsl.pl)
- Szymańska B., Strzymiński T. (2011) *Standardy tworzenia audiodeskrypcji do produkcji audiowizualnych*: www.audiodeskrypcja.pl
- Ślot K. (2010) *Rozpoznawanie biometryczne: nowe metody ilościowej reprezentacji obiektów*. Wydawnictwa Komunikacji i Łączności, Warszawa.
- Tadeusiewicz R. (1992) Systemy wizyjne robotów przemysłowych, WNT, Warszawa.
- Talbot M., Cowan B. (2006) Trajectory capture in frontal plane geometry for visually impaired. In: *Proceedings of the 12th International Conference on Auditory Display*, London, UK.
- Talukder A., Jakubowski M. (2003) Technologia tworzenia map i planów wypukłych dla niewidomych i słabowidzących. W: *Nowoczesne techniki kształcenia dzieci niewidomych i słabowidzących*. Europejska Konferencja, Owińska, 25–26 kwietnia.
- Titterton D.H., Weston J.L. (2004) Strapdown inertial navigation technology. *Progress in Astronautics and Aeronautics*, MIT Lincoln Laboratory, 2nd Edition.
- Tolman E.C. (1948) Cognitive maps in rats and men. *The Psychological Review*, vol. 55, no. 4, pp. 189–208.
- Torralba A., Oliva A. (2002) Depth estimation from image structure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1226–1238.

- Ungar S., Simpson A., Blades M. (2004) Strategies for organizing information while learning a map by blind and sighted people. In: Ballesteros S. J., Heller M.A. (eds.), *Touch, Blindness and Neuroscience*, pp. 271–280, UNED Press, Madrid, Spain.
- Ustawa RP (1997) *Ustawa Rzeczypospolitej Polskiej z dnia 27 sierpnia 1997 r. o rehabilitacji zawodowej i społecznej oraz zatrudnianiu osób niepełnosprawnych*. Dz. U. Nr 123, poz. 776 z późn. zm.
- van Diggelen F. (2009) *A-GPS: Assisted GPS, GNSS, and SBAS*. Artech House, Boston, London.
- Velázquez R., Pissaloux E.E. (2008) Tactile displays in human-machine interaction: four case studies. *International Journal of Virtual Reality*, vol. 7, no. 2, pp. 51–58.
- Visell Y. (2009) Tactile sensory substitution: models for enactment in HCI. *Interacting with Computers*, vol. 21, no. 1–2, pp. 38–53.
- Voss P., Gougoux F., Lassonde M., Zatorre R.J., Lepore F. (2006) A positron emission tomography study during auditory localization by late-onset blind individuals. *Auditory and Vestibular Systems*, vol. 17, no. 4, pp. 383–288.
- Wallis G., Bühlhoff H. (1999) Learning to recognize objects. *Trends in Cognitive Sciences*, vol. 3, no. 1, pp. 22–25.
- Wang S., Min J., Yi B.K. (2008) *Location Based Services for Mobiles: Technologies and Standards IEEE International Conference on Communication (ICC)*, Beijing, China.
- Want R. (2006) An Introduction to RFID technology. *Pervasive Computing*, pp. 25–33.
- Wapiennik E., Piotrowicz R. (2002) *Niepełnosprawny pełnoprawny obywatel Europy*. Urząd Komitetu Integracji Europejskiej, Warszawa.
- Ward M., van Kranenburg R. (2006) RFID: frequency, standards, adoption and innovation. *JISC Technology and Standards Watch*.
- Weeks R., Horwitz B., Aziz-Sultan A., Tian B., Wessinger C.M., Cohen L.G., Hallett M., Rauschecker J.P. (2000) A Positron Emission Tomographic study of auditory localization in the congenitally blind. *Journal of Neuroscience*, vol. 20, no. 7, pp. 2664–2672.
- Weiland J.D., Humayun M.S. (2008) Visual prosthesis. *Proceedings of the IEEE*, vol. 96, no. 7, pp. 1076–1084.
- Wersényi G.Y. (2010) Auditory representations of a graphical user interface for a better humancomputer interaction. In: Ystad S. (ed.) *Auditory Display, Lecture Notes in Computer Science* 5954, Springer Verlag, Berlin, pp. 80–102.
- Wersényi G. Y. (2007) Simulation of small head-movements on a virtual audio display using headphone playback and HRTF synthesis. In: *Proceedings of the 13th International Conference on Auditory Display*, Montreal, Canada.
- Wexler M. (2002) More to 3-D vision than meets the eye. *Trends in Cognitive Sciences*, vol. 6, no. 12, pp. 497–498.
- Wexler M., van Boxtel J.J.A. (2005) Depth perception by the active observer. *Trends in Cognitive Sciences*, vol. 9, no. 9, pp. 431–438.
- WHO (2011) International Classification of Diseases, web page: www.who.int/classifications/icd/en/

- Wiak S. (red.) (2009) *Mechatronika. Tom I – Podstawy, Metody, Modele Komputerowe, Technologie*, EXIT.
- Wickens Ch. (2002) Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, vol. 3, no. 2, , pp. 150–177.
- Willis S., Helal S. (2005) RFID information grid and wearable computing solution to the problem of wayfinding for the blind user in a campus environment. *9-th IEEE International Symposium on Wearable Computers (ISWC'05)*.
- Woźniak Z. (2009) Niepełnosprawni – status społeczny największej mniejszości. W: Kabsh A. *Konwencja praw osób z niepełnosprawnościami: godność i sprawiedliwość dla nas wszystkich*. Zeszyty Promocji Rehabilitacji, zeszyt 3, Ośrodek Wydawnictw Naukowych, Poznań, str. 11–30.
- Wu Z.C., Wang Z.F., Ge Y. (2002) Gravity based online calibration for monolithic triaxial accelerometers' gain and offset drift. *Proceedings of the 4th World Congress on Intelligent Control and Automation*, vol. 3, pp. 2171–2175.
- Wysocki M. (2009) *Dostępna przestrzeń publiczna*. Fundacja Instytut Rozwoju Regionalnego, Kraków.
- Yang H., Chang Y., Huo J. (2009) Fine-granular motion matching for inter-view motion skip mode in multiview video coding. *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 6, pp. 887–892.
- Yeadon A., (2000) Najważniejsze zrozumieć. W: Adamowicz–Hummel A., Guzowska H. (red.) *Poradnik pracodawcy osób niewidomych i słabo widzących*, Fundacja AWARE Europe, Warszawa.
- Yim J., Myung R., Lee B. (2007) The mobile phone's optimal vibration frequency in mobile environments, usability and internationalization. *HCI and Culture, Lecture Notes in Computer Science*, vol. 4559, pp. 646–652.
- Yu W., Ramloll R., Brewster S. (2001) Haptic graphs for blind computer users. *Lecture Notes in Computer Science*, vol. 2058, pp. 41–51.
- Zadeh M.H. (ed.) (2010) *Advances in Haptics*, InTech, Zagreb, Croatia.
- Zadrozny P. (2009) *Dostępna komunikacja miejska*. Fundacja Instytut Rozwoju Regionalnego, Kraków.
- Zahorik P. (2002) Auditory display of sound source distance. In: *Proceedings of the 2002 International Conference on Auditory Display (ICAD 2002)*.
- Zhang R., Tsai P.S., Cryer J.E., Shah M. (1999) Shape-from-shading: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 8, pp. 690–706.
- Zwiers M.P., Van Opstal A.J., Cruysberg, J.R.M. (2001) A spatial hearing deficit in early-blind humans. *The Journal of Neuroscience*, vol. 21, pp. 1–5.
- Zyłka P. (2004) Polimery elektroaktywne. *Przegląd Elektrotechniczny*, nr 1K, str. 277–280.
- Zernicki B. (1983) *Mózg*. Ossolineum, Warszawa.



Politechnika Łódzka

Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki

Z recenzji profesora Jerzego Lewandowskiego:

Tematyka monografii dotycząca systemów nawigacji osobistej dla niewidomych i słabowidzących jest jak najbardziej aktualna z naukowego i społecznego punktu widzenia. (...) Szczegółowym rozważaniom poddał Autor zagadnienia dotyczące rodzajów technik nawigacyjnych, (...) zastosowania map elektronicznych w nawigacji pieszej, nawigacji inercyjnej i wizyjnej oraz technik integracji danych nawigacyjnych. (...) Omówił sposoby niewizualnej prezentacji otoczenia osobom niewidomym za pomocą urządzeń wytwarzających bodźce dla zmysłów słuchu i dotyku, w systemach substytucji sensorycznej wspomagających osoby niewidome w samodzielnym poruszaniu się i orientacji przestrzennej...

Z recenzji profesora Andrzeja Materki:

(...) Chociaż główne wątki monografii są skoncentrowane na układach elektronicznych i systemach teleinformatycznych oraz metodach i algorytmach przetwarzania i analizy obrazów oraz sygnałów, to oceniane dzieło jako całość ma charakter interdyscyplinarny. Obejmuje ono zagadnienia anatomii, fizjologii i psychologii spostrzegania, (...) zagadnienia ergonomii i projektowania urządzeń do komunikacji człowieka pozbawionego zmysłu wzroku z otoczeniem, ekscytującą pod względem naukowym i ważną utylitarnie tematykę „substytucji zmysłów”. (...) Nie znam innego podobnego opracowania w piśmiennictwie krajowym; wydano tylko kilka książek o zbliżonej tematyce na świecie.

O Autorze:



Paweł Strumiłło jest profesorem Politechniki Łódzkiej. Jego specjalnością naukową jest elektronika medyczna. Dziedzinie tej poświęcił pracę doktorską obronioną w 1993 roku w Uniwersytecie Strathclyde w Szkocji oraz rozprawę habilitacyjną, którą obronił w Politechnice Łódzkiej w 2002 roku. Opublikował ponad 100 prac naukowych obejmujących tematykę przetwarzania sygnałów biomedycznych, analizy obrazów, inteligencji obliczeniowej oraz systemów komunikacji człowiek–komputer. Jest kierownikiem Zakładu Elektroniki Medycznej. Prowadzi projekty badawcze i rozwojowe, których celem jest wdrażanie systemów wspomagających osoby z niepełnosprawnościami wzroku i ruchu (m.in. we współpracy z Orange Labs Polska). Jest członkiem IEEE w stopniu senior member oraz członkiem Sekcji Sygnałów, Układów i Systemów Elektronicznych PAN.

ISBN 978-83-933359-0-9



Publikacja dystrybuowana bezpłatnie