

Spis treści

| | | |
|----------|--|-----------|
| 1 | Wstęp | 11 |
| 1.1 | Do kogo adresowana jest ta książka | 12 |
| 1.2 | Historia badań nad mową i językiem | 12 |
| 1.3 | Obecne główne trendy badań | 16 |
| 1.4 | Opis zawartości rozdziałów | 18 |
| 2 | Wyzwania i możliwe korzyści | 21 |
| 2.1 | Problemy związane z technologiami mowy | 25 |
| 2.2 | Komunikacja, a komputer i matematyka | 26 |
| 2.3 | Zastosowania | 27 |
| 3 | Komunikacja człowieka z maszyną | 31 |
| 3.1 | Zasady udanej interakcji człowieka z maszyną | 34 |
| 3.2 | Kilka słów o fonetyce | 36 |
| 3.3 | Czego ludzie wymagają od technologii mowy? | 36 |
| 4 | Zasoby językowe | 39 |
| 4.1 | Korpusy obcojęzyczne | 43 |
| 4.2 | Korpusy języka polskiego | 45 |
| 4.3 | Słowniki | 49 |
| 4.4 | Sposoby gromadzenia danych | 51 |
| 5 | Klasyfikacja sygnałów mowy i ich modele | 53 |
| 5.1 | Definicja sygnału mowy | 53 |
| 5.2 | Modele matematyczne sygnałów mowy | 54 |
| 5.2.1 | Sygnały analogowe | 55 |
| 5.2.2 | Sygnały cyfrowe | 56 |

| | | |
|----------|--|-----------|
| 6 | Analiza częstotliwościowa sygnału | 59 |
| 6.1 | Widmo sygnałów okresowych | 60 |
| 6.2 | Transformacja Fouriera | 63 |
| 6.3 | Widma sygnałów mowy | 64 |
| 6.4 | Własności transformacji Fouriera | 68 |
| 6.4.1 | Parzystość widma amplitudowego i nieparzystość widma fazowego | 68 |
| 6.4.2 | Liniowość transformacji Fouriera | 69 |
| 6.4.3 | Zachowanie iloczynu skalarnego | 70 |
| 6.4.4 | Zachowanie energii | 70 |
| 6.4.5 | Dualność transformacji Fouriera | 71 |
| 6.4.6 | Zmiana skali czasu sygnału | 71 |
| 6.4.7 | Przesunięcie w dziedzinie czasu | 72 |
| 6.4.8 | Przesunięcie w dziedzinie częstotliwości | 72 |
| 6.4.9 | Splot w dziedzinie czasu | 73 |
| 6.4.10 | Różniczkowanie w dziedzinie czasu | 74 |
| 6.4.11 | Różniczkowanie w dziedzinie częstotliwości | 75 |
| 6.4.12 | Całkowanie w dziedzinie czasu | 75 |
| 6.5 | Przykłady transformat Fouriera | 76 |
| 6.6 | Uogólnienie transformacji Fouriera | 79 |
| 6.6.1 | Widmo impulsu Diraca | 80 |
| 6.6.2 | Transformata Fouriera sygnału stałego | 81 |
| 6.6.3 | Transformaty Fouriera sygnałów okresowych | 82 |
| 6.6.4 | Transformacja sygnału z niezerową wartością średnią | 82 |
| 6.7 | Transformacja kosinusowa | 83 |
| 7 | Lokalna analiza widmowa mowy | 85 |
| 7.1 | Okna transformacji lokalnych | 89 |
| 7.1.1 | Okno prostokątne | 90 |
| 7.1.2 | Okno Bartletta | 91 |
| 7.1.3 | Okno Hanna | 91 |
| 7.1.4 | Okno Hamminga | 92 |
| 7.1.5 | Okno Parzena | 93 |
| 7.1.6 | Okno paraboliczne | 95 |
| 7.1.7 | Okno Gaussa | 95 |
| 7.2 | Transformacja Gabora | 102 |

| | | |
|-----------|---|------------|
| 7.3 | Analiza falkowa | 103 |
| 7.3.1 | Zastosowanie falek do reprezentacji sygnałów | 110 |
| 7.3.2 | Transformacja falkowa | 112 |
| 8 | Filtry analogowe | 117 |
| 8.1 | Definicje i modele matematyczne filtrów | 117 |
| 8.2 | Idealny filtr dolnoprzepustowy | 122 |
| 8.3 | Rzeczywiste filtry dolnoprzepustowe | 124 |
| 8.3.1 | Filtr Butterwortha | 124 |
| 8.3.2 | Filtry Czebyszewa | 126 |
| 8.3.3 | Filtr eliptyczny | 130 |
| 8.4 | Porównanie własności filtrów rzeczywistych | 134 |
| 8.5 | Przekształcenie filtrów dolnoprzepustowych w filtry innego typu | 135 |
| 9 | Przetwarzanie A/C | 139 |
| 9.1 | Próbkowanie sygnałów mowy | 139 |
| 9.2 | Aliasing | 142 |
| 10 | Analiza częstotliwościowa sygnałów cyfrowych | 145 |
| 10.1 | Dyskretna transformacja Fouriera sygnałów akustycznych | 145 |
| 10.2 | Szybka transformacja Fouriera | 152 |
| 10.3 | Dyskretna transformacja kosinusowa | 158 |
| 11 | Filtry cyfrowe | 161 |
| 11.1 | Z-transformacja | 161 |
| 11.1.1 | Przykłady z -transformat | 163 |
| 11.1.2 | Własności z -transformacji | 164 |
| 11.1.3 | Związek z -transformacji z transformacją Fouriera | 167 |
| 11.2 | Filtry FIR | 168 |
| 11.2.1 | Charakterystyki częstotliwościowe filtrów FIR | 173 |
| 11.2.2 | Filtry FIR z liniową charakterystyką fazową | 174 |
| 11.2.3 | Optymalizacyjna metoda projektowania filtrów FIR | 179 |
| 11.3 | Filtry IIR | 183 |
| 11.3.1 | Stabilność filtrów IIR | 184 |
| 11.3.2 | Projektowanie filtrów IIR w oparciu o filtry analogowe | 194 |

| | |
|---|------------|
| 11.4 Banki filtrów | 196 |
| 12 Teoria falek w lokalnej analizie sygnałów | 197 |
| 13 Kodowanie i kompresja mowy | 203 |
| 13.1 Kompresja bezstratna | 205 |
| 13.1.1 Kodowanie różnicowe | 205 |
| 13.1.2 Kodowanie entropowe | 207 |
| 13.1.3 Kodowanie Huffmana | 210 |
| 13.1.4 Kodowanie ciągów | 213 |
| 13.1.5 PCM | 213 |
| 13.1.6 LPC | 214 |
| 13.2 Kompresja stratna | 214 |
| 13.2.1 Kwantyzacja skalarna | 215 |
| 13.2.2 Kodowanie transformatowe sygnałów | 216 |
| 13.2.3 Kompresja wykorzystująca modelowanie traktu gło- sowego | 217 |
| 13.2.4 Systemy kompresji stratnej dźwięku | 218 |
| 14 Parametryzacja i segmentacja | 221 |
| 14.1 Odszumianie | 222 |
| 14.1.1 Systemy wielomikrofonowe | 222 |
| 14.2 Segmentacja | 224 |
| 14.3 Parametryzacja | 227 |
| 14.3.1 Metody oparte o liniowe współczynniki predykcji | 227 |
| 14.3.2 Metody oparte o banki filtrów i faleki | 231 |
| 15 Modele akustyczne | 235 |
| 15.1 Zapis fonetyczny | 235 |
| 15.2 Reguła Bayesa | 236 |
| 15.3 HMM – niejawne modele Markowa | 237 |
| 15.4 Algorytm Viterbiego | 243 |
| 15.5 SVM | 247 |
| 15.6 Sieci Bayesa | 249 |
| 15.7 Sieci neuronowe | 251 |

| | |
|--|------------|
| 16 Modele językowe | 255 |
| 16.1 <i>N</i> -gramy | 255 |
| 16.2 Wygładzanie <i>n</i> -gramów | 266 |
| 16.3 Filtr Blooma | 270 |
| 16.4 Modele bez zachowania kolejności słów | 271 |
| 16.5 Łączenie modeli – grafy i listy | 275 |
| 17 Przetwarzanie języka naturalnego | 277 |
| 17.1 Parsery zdań | 277 |
| 17.2 Analizatory morfologiczne | 279 |
| 17.3 Metody semantyczne | 279 |
| 17.4 Słowosieć | 280 |
| 18 Synteza mowy | 285 |
| 18.1 Techniki syntezy mowy | 285 |
| 18.2 Syntezatory | 287 |
| 19 Rozpoznawanie mówcy | 289 |
| 19.1 Weryfikacja a identyfikacja | 289 |
| 19.2 Dalsza klasyfikacja | 290 |
| 19.3 Algorytmy | 291 |
| 19.4 Istniejące rozwiązania | 293 |
| 20 Systemy dla innych języków | 297 |
| 20.1 HTK | 298 |
| 20.2 Dragon firmy Nuance | 298 |
| 20.3 Sphinx | 299 |
| 20.4 Julius | 300 |
| 20.5 Microsoft | 300 |
| 20.6 Fiński | 301 |
| 21 Przetwarzanie mowy polskiej | 303 |
| 21.1 Wiek XX | 303 |
| 21.2 XXI wiek | 304 |

| | |
|---|------------|
| 22 Inne technologie mowy | 315 |
| 22.1 Rozpoznawanie emocji | 316 |
| 22.2 Systemy dialogowe | 317 |
| 22.3 Automatyczne tłumaczenia | 325 |
| 22.4 Nauka języków obcych | 326 |
| 23 Podsumowanie | 327 |

