

Beke András

GÉPI BESZÉLŐDETEKTÁLÁS
MAGYAR NYELVŰ SPONTÁN TÁRSALGÁSOKBAN

Beszéd • Kutatás • Alkalmazás

Beke András

**GÉPI BESZÉLŐDETEKTÁLÁS
MAGYAR NYELVŰ
SPONTÁN TÁRSALGÁSOKBAN**

Budapest, 2015



A kötet megjelenését az OTKA PUB-K 114596 ny. pályázat támogatta.

Lektorálták:

Adamikné Jászó Anna

Gósy Mária

Olaszy Gábor

© Beke András, 2015

ISBN 978-963-312-234-1

ISSN 2064-4442



Felelős kiadó: Hunyady András ügyvezető igazgató

Felelős szerkesztő: Pál Dániel Levente

Nyomdai munkák: Multiszolg Bt.

Tördelés: Windor Bt.

Borítóterv: Csele Kmotrik Ildikó

Tartalom

Sorozatszerkesztői előszó.....	9
Előszó.....	11
1. Bevezetés.....	13
2. A beszélődetektáló általános felépítése.....	23
2.1. Akusztikai jellemzők a beszélődetektáláshoz.....	25
2.2. Beszélőszegmentálás.....	27
2.2.1. Metrikus alapú szegmentáló algoritmusok.....	28
2.2.1.1. Bayes-féle információs kritérium (BIC: Bayesian Information Criterion).....	28
2.2.1.2. Általánosított valószínűségarány (GLR: Generalized Likelihood Ratio).....	31
2.2.1.3. Gish-távolság (Gish-distance)	32
2.2.1.4. Kullback–Leibler-távolság (KL vagy KL2)	32
2.2.1.5. Más távolságmérési eljárások	33
2.2.2. Nem metrikán alapuló szegmentálók	33
2.2.2.1. Szünetalapú beszélőszegmentáló.....	33
2.2.2.2. Modellalapú szegmentáló	34
2.2.3. A beszélőszegmentáló algoritmusok összegzése.....	35
2.3. Beszélőklaszterezés.....	35
2.3.1. Hierarchikus klaszterezési technikák.....	36
2.3.1.1. Alulról felfelé (egyesítő, bottom-up) klaszterező eljárások	37
2.3.1.2. Fentről lefelé (lebontó, top-down) klaszterező technikák	40
2.4. Beszéddetektálás	40
2.4.1. A beszéddetektáló általános leírása.....	41
2.4.2. Jellemzőkinyerés a beszéddetektáló megvalósításához	41
2.4.3. A beszéddetektáló döntési modulja.....	42
2.4.4. A beszéddetektáló utófeldolgozása (simítás)	42
2.5. Beszélőspecifikus jellemzők a gépi beszélőfelismerésben	43
2.5.1. Kevert Gauss-beszélőmodell.....	48
2.5.1.1. Kevert Gauss-modell	48
2.5.1.2. Univerzális háttérmodell.....	50
2.5.1.3. A beszélőegyezés mérése	51
2.6. Az egyszerre beszélés detektálása	52

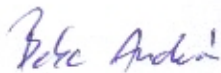
3. A kutatás célja, kutatási kérdések és hipotézisek.....	57
3.1. Kutatási kérdések.....	57
3.2. A kutatás célja	57
3.3. A kutatás hipotézisei	58
4. Kísérleti személyek, általános anyag és módszer	59
4.1. Anyag és kísérleti személyek.....	59
4.1.1. A beszélődetektáló kiértékeléséhez használt korpusz	60
4.1.2. A beszélőspecifikus jellemzők kialakításához használt korpusz	60
4.1.3. A beszédetektáléhoz használt korpusz.....	61
4.1.4. Az egyszerrebeszélés-detektáléhoz használt korpusz	61
4.2. Kiértékelési módszer	61
4.2.1. Beszélődetektálási hibaarány (DER: Diarization Error Rate).....	62
4.2.2. További kiértékelési technikák (DET: Detection Error Tradeoff)	64
5. Beszélődetektálás társalgásokban	67
5.1. A korpusz általános statisztikai jellemzői	67
5.2. A beszélődetektáló felépítése.....	69
5.2.1. Beszélőszegmentálás	69
5.2.1.1. Jellemzőkinyerés a beszélőszegmentáláshoz	69
5.2.1.2. Bayes-féle információs kritérium (BIC: Bayesian Information Criterion).....	69
5.2.1.2.1. Növekedő ablakhosszmetódus a ΔBIC számításához....	71
5.2.1.2.2. A BIC paraméterei.....	71
5.2.1.3. Téves riasztások csökkentése (False Alarm Compensation).....	72
5.2.1.3.1. A KL2-alapú utófeldolgozás beállításai.....	72
5.2.2. Beszélőklaszterezés	73
5.2.2.1. Jellemzőkinyerés a beszélőklaszterezéshez	73
5.2.2.2. GMM-szupervektor.....	73
5.2.2.3. BIC-alapú klaszterezés	74
5.3. A beszédetektálás felépítése.....	75
5.3.1. Jellemzőkinyerés	76
5.3.2. A beszédetektáló döntési módszere.....	76
5.3.3. A beszédetektáló utófeldolgozása.....	77
5.3.4. Az általunk javasolt eljárás a küszöb meghatározására	78
5.3.5. A beszédetektáló kiértékelése.....	79
5.4. Az egyszerre beszélések automatikus osztályozása	
spontán magyar társalgásokban.....	79
5.4.1. Jellemzőkinyerés	80
5.4.2. Lényegkiemelés	82
5.4.2.1. Korlátozott Boltzmann-gép	82

5.4.2.2. Az RBM előtanítási paraméterei	84
5.4.3. Osztályozás	85
5.4.3.1. Szupport vektor gép (SVM: Support Vector Machine)	85
5.4.3.1.1. Az SVM tanítási paraméterei	88
6. Eredmények	89
6.1. A beszélőszegmentálás eredménye az alapbeállítások mellett	89
6.2. A BIC beszélődetektáló beszélőspecifikus akusztikai jellemzővel	90
6.2.1. Beszélőspecifikus jellemzők	90
6.2.2. A beszélőspecifikus jellemzők implementálása a beszélődetektálóba	92
6.3. A BIC λ paraméterének optimális megválasztása	93
6.4. A beszéddetektálás implementálása	93
6.4.1. A beszéddetektáló eredményei spontán társalgásban	93
6.4.2. A beszéddetektáló implementációja a beszélődetektálóba	95
6.5. Az egyszerrebeszélés-detektáló eredménye	96
6.5.1. Az egyszerrebeszélés-detektáló eredménye spontán társalgásban	96
6.5.2. Az egyszerrebeszélés-detektáló implementációja a beszélődetektálóba ...	98
7. Következtetések	101
7.1. Beszéddetektáló	101
7.2. Beszélőspecifikus jellemzők a gépi beszélőfelismerésen keresztül	102
7.3. Az egyszerre beszélések automatikus osztályozása spontán magyar társalgásokban	103
7.4. Beszélődetektálás	104
7.4.1. A beszélődetektáló alaprendszere	105
7.4.2. Beszélőspecifikus akusztikai jellemzők implementálása	105
7.4.3. A BIC λ paraméterének beállítása	106
7.4.4. A beszéddetektálás implementálása	106
7.4.5. Az egyszerrebeszélés-detektáló implementálása	106
7.4.6. A kifejlesztett rendszer végső eredménye	106
8. Összegzés	107
9. Kitekintés	109
10. Irodalom	111
Automatic speaker diarization in Hungarian spontaneous conversations	131

NYILATKOZAT

Alulírott Beke András ezúton nyilatkozom, hogy az OTKA/NKFI 114596. sz. támogatási projektben részt alvállalkozóként részt vevő szerződéses partnerekkel részemről összeférhetetlenség nem áll fenn.

Budapest, 2015. december 16.



Beke András