



Универзитет у Новом Саду
Факултет техничких наука

Департман за енергетику, електронику и
телекомуникације

Аутор:

Бранислав Поповић

Ментор:

Проф. др Владо Делић

**Хијерархијско кластерованье модела Гаусових смеша у
апликацијама за континуално препознавање говора**

ДОКТОРСКА ДИСЕРТАЦИЈА

Нови Сад, 2012.

Универзитет у Новом Саду

Факултет техничких наука

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Редни број: РБР	
Идентификациони број: ИБР	
Тип документације: ТД	Монографска документација
Тип записа: ТЗ	Текстуални штампани материјал
Врста рада (дипл., маг., докт.): ВР	Докторска дисертација
Име и презиме аутора: АУ	Бранислав Поповић, М.Sc.
Ментор (титула, име, презиме, звање): МН	Проф. др Владо Делић, ванредни професор
Наслов рада: НР	Хијерархијско кластеровање модела Гаусових смеша у апликацијама за континуално препознавање говора српски
Језик публикације: ЈП	
Језик извода: ЈИ	српски / енглески
Земља публикавања: ЗП	Република Србија
Уже географско подручје: УГП	Аутономна Покрајина Војводина
Година: ГО	2012.
Издавач: ИЗ	Ауторски репринт
Место и адреса: МА	Трг Доситеја Обрадовића 6 21000 Нови Сад

Физички опис рада: ФО	(број поглавља / страница / слика / графика / референци / прилога) 11/102/10/22/40/3
Научна област: НО	
Научна дисциплина: НД	
Предметна одредница, кључне речи: ПО	Модел Гausових смеша - Операција поделе и спајања - Хијерархијско кластерованье - Континуално препознавање говора - Селекција Гаусијана
УДК	
Чува се: ЧУ	У библиотеци Факултета техничких наука у Новом Саду, Трг Доситеја Обрадовића 6, 21000 Нови Сад
Важна напомена: ВН	-
Извод: ИЗ	<p>У оквиру докторске дисертације представљен је нови алгоритам хијерархијског кластерованья модела Гаусових смеша, заснован на операцији поделе и спајања. Алгоритам тежи побољшању локално оптималног решења одређеног иницијалном констелацијом. Иницијализује се локално оптималним параметрима, добијеним коришћењем референтног приступа сличног k-means-у и тежи ка приближавању глобалном оптимуму функције циља, итеративном применом операција поделе и спајања над кластерима Гаусових компоненти, одређеним применом референтног алгоритма.</p> <p>Додатно побољшање алгоритма осварено је увођењем принципа селекције модела у сврху одређивања најповољнијег односа тачности и рачунске сложености, у задатку</p>

селекције гаусијана унутар реалног система за препознавање. Предложени метод тестиран је како над вештачки генерисаним подацима, тако и у оквиру алгоритма селекције гаусијана, на примеру система за континуално препознавање говора. У оба случаја забележено је побољшање резултата у односу на резултате остварене применом референтног алгоритма.

Побољшања алгоритма селекције гаусијана избором оптималног скупа системских параметара такође су размотрена.

Датум прихватања теме од стране

НН већа:

ДП

Датум одбране:

ДО

Чланови комисије:

КО

председник: др Жељен Трповски,
ванредни професор, Факултет
техничких наука, Нови Сад

члан: др Слободан Јовичић, редовни
професор, Електротехнички факултет,
Београд

члан: др Татјана Грбић, доцент,
Факултет техничких наука, Нови Сад

члан: др Милан Сечујски, доцент,
Факултет техничких наука, Нови Сад

члан: др Владо Делић, ванредни
професор, Факултет техничких наука,
Нови Сад (ментор)

University of Novi Sad

Faculty of Technical Sciences

KEY WORD DOCUMENTATION

Accession number:
ANO
Identification number:
INO
Document type: Monograph documentation
DT
Type of record: Textual printed material
TR
Contents code: Ph.D. Thesis
CC
Author: Branislav Popović, M.Sc.
AU
Mentor: Prof. Vlado Delić, Ph.D., Associate
MN Professor
Title: Hierarchical Clustering of Gaussian
TI Mixture Models in Applications for
Continuous Speech Recognition
Language of text: Serbian
LT
Language of abstract: English / Serbian
LA
Country of publication: Republic of Serbia
CP
Locality of publication: Autonomous Province of Vojvodina
LP
Publication year: 2012.
PY
Publisher: Author's reprint
PU
Publication place: Trg Dositeja Obradovića 6
PP 21000 Novi Sad

Physical description: PD	(chapters / pages / images / charts / references / appendixes) 11/102/10/22/40/3
Scientific field SF	
Scientific discipline SD	
Subject, Key words SKW	Gaussian mixture models - Split-and- merge operation - Hierarchical clustering - Continuous speech recognition - Gaussian selection
UC	
Holding data: HD	Library of the Faculty of Technical Sciences, Trg Dositeja Obradovića 6, 21000 Novi Sad
Note: N	-
Abstract: AB	<p>The dissertation presents a novel split-and-merge algorithm for hierarchical clustering of Gaussian mixture models. The algorithm tends to improve on the local optimal solution determined by the initial constellation. It is initialized by local optimal parameters obtained by using a baseline approach similar to k-means, and it tends to approach more closely to the global optimum of the target clustering function, by iteratively splitting and merging the clusters of Gaussian components obtained as the output of the baseline algorithm.</p> <p>The algorithm is further improved by introducing model selection in order to obtain the best possible trade-off between recognition accuracy and computational load in a Gaussian selection task applied within an actual recognition system. The proposed method is tested both on artificial data and in the framework of Gaussian selection performed within a real</p>

continuous speech recognition system. In both cases an improvement over the baseline method has been observed.

Additional improvements of Gaussian selection algorithm by using the optimal set of system parameters are also discussed.

Accepted on Scientific Board on:

AS

Defended:

DE

Thesis Defend Board:

DB

president: Dr Željko Trpovski, Associate Professor, Faculty of Technical Sciences, Novi Sad

member: Dr Slobodan Jovičić, Full Professor, School of Electrical Engineering, Belgrade

member: Dr Tatjana Grbić, Assistant Professor, Faculty of Technical Sciences, Novi Sad

member: Dr Milan Sečujski, Assistant Professor, Faculty of Technical Sciences, Novi Sad

member: Dr Vlado Delić, Associate Professor, Faculty of Technical Sciences, Novi Sad (mentor)

Садржај

Поглавље 1	Увод.....	1
1.1.	Мотивација и предмет истраживања.....	3
1.2.	Организација рада	5
Поглавље 2	Теоријске основе докторске дисертације	8
2.1.	Модел Гausових смеша	8
2.2.	Естимација параметара модела Гаусових смеша	12
2.2.1	EM алгоритам	14
2.2.2	K-means алгоритам.....	17
2.3.	Алгоритми поделе и спајања	18
Поглавље 3	Метод селекције гаусијана	22
Поглавље 4	Хијерархијско кластеровање модела Гаусових смеша	27
4.1.	Формулација проблема	27
4.2.	Опис референтног алгоритма	28
Поглавље 5	Нови алгоритам поделе и спајања за хијерархијско кластеровање модела Гаусових смеша	32
5.1.	Операција спајања	36
5.2.	Операција поделе	37
Поглавље 6	Метод селекције модела у предложеном алгоритму поделе и спајања	43
Поглавље 7	Додатно побољшање перформанси алгоритма селекције гаусијана	46

Поглавље 8	Карактеристике система за континуално препознавање говора.....	49
8.1.	Опис система за континуално препознавање говора.....	49
8.2.	Говорна база	52
Поглавље 9	Експериментални резултати	53
9.1.	Експерименти над синтетичким подацима, конструкција узорка	53
9.2.	Примена хијерархијског кластеровања у оквиру реалног CSR система	64
9.3.	Перформансе алгорита селекције гаусијана	70
Поглавље 10	Закључак	76
Поглавље 11	Додаци	78
11.1.	Додатак А: Приказ HGMMC алгорита	78
11.2.	Додатак Б: Приказ S&M HGMMC алгорита.....	79
11.3.	Додатак В: Приказ S&M MS HGMMC алгорита	81
Литература	83

Списак слика

СЛИКА 2.1	
ТРОДИМЕНЗИОНАЛНИ ПРИКАЗ GMM МОДЕЛА	8
СЛИКА 2.2	
ЈЕДНОДИМЕНЗИОНАЛНА НОРМАЛНА (ГАУСОВА) РАСПОДЕЛА.....	10
СЛИКА 2.3	
ГЕОМЕТРИЈСКА ИНТЕРПРЕТАЦИЈА КАРАКТЕРИСТИЧНИХ ВЕКТОРА И КАРАКТЕРИСТИЧНИХ ВРЕДНОСТИ	12
СЛИКА 2.4	
КОРАЦИ ЕМ АЛГОРИТМА	16
СЛИКА 2.5	
ИТЕРАЦИЈЕ K-MEANS АЛГОРИТМА	18
СЛИКА 2.6	
КОРАЦИ АЛГОРИТМА ПОДЕЛЕ И СПАЈАЊА (Blekas, Lagaris)	20
СЛИКА 5.1	
ПРОЦЕДУРА ПОДЕЛЕ И СПАЈАЊА GMM КЛАСТЕРА	34
СЛИКА 5.2	
ОДСТУПАЊЕ У ОДНОСУ НА MCSA АПРОКСИМАЦИЈУ СА МИЛИОН ОДБИРАКА	41
СЛИКА 8.1	
СТАБЛО ФОНЕТСКЕ СЛИЧНОСТИ	50
СЛИКА 8.2	
ДИЈАГРАМ ПРОЦЕДУРЕ ПОВЕЗИВАЊА НММ СТАЊА	51

Списак табела

ТАБЕЛА 9.1

СМАЊЕЊЕ ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНО КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M MS HGMMС И S&M HGMMС АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 55

ТАБЕЛА 9.2

ПРОСЕЧНО СМАЊЕЊЕ СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНО КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M MS HGMMС И S&M HGMMС АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 58

ТАБЕЛА 9.3

ПРОЦЕНТУАЛНИ БРОЈ СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ S&M MS HGMMС И S&M HGMMС АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 61

ТАБЕЛА 9.4

РЕЗУЛТАТИ ПРИМЕНЕ ПРЕДЛОЖЕНИХ S&M HGMMС И S&M MS HGMMС АЛГОРИТАМА, РЕФЕРЕНТНОГ HGMMС АЛГОРИТАМА И ПОЛАЗНОГ СИСТЕМА У ТЕРМИНИМА WER И CF ФАКТОРА, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА. 66

ТАБЕЛА 9.5

УПОРЕДНИ ПРИКАЗ ПЕРФОРМАНСИ АЛГОРИТАМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И CF ФАКТОРА ПРИ ПРОМЕНИ ТИПА РАСТОЈАЊА КОРИШЋЕНОГ ПРИ ДЕКОДОВАЊУ. 71

ТАБЕЛА 9.6

ПЕРФОРМАНСЕ АЛГОРИТАМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И CF ФАКТОРА ПРИ ПРОМЕНИ ПРОСЕЧНОГ БРОЈА ГАУСИЈАНА ПО ЈЕДНОМ КЛАСТЕРУ. СИСТЕМ КОРИСТИ МАХАЛАНОБИСОВО РАСТОЈАЊЕ У ПРОЦЕСУ ЕВАЛУАЦИЈЕ. 73

ТАБЕЛА 9.7

УПОРЕДНИ ПРИКАЗ ПЕРФОРМАНСИ АЛГОРИТАМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И CF ФАКТОРА. ВРШИ СЕ ЕВАЛУАЦИЈА САМО ЈЕДНОГ КЛАСТЕРА ПО ФРЕЈМУ, ПРИ ЧЕМУ КЛАСТЕРИ ЗАУЗИМАЈУ ЗНАЧАЈАН ДЕО ПРОСТОРА ОБЕЛЕЖЈА. СИСТЕМ КОРИСТИ МАХАЛАНОБИСОВО РАСТОЈАЊЕ У ПРОЦЕСУ ЕВАЛУАЦИЈЕ. 75

Списак графикана

ГРАФИКОН 9.1

УПОРЕДНИ ПРИКАЗ СМАЊЕЊА ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ LBVA АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА..... 55

ГРАФИКОН 9.2

УПОРЕДНИ ПРИКАЗ СМАЊЕЊА ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ MCSA АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА..... 56

ГРАФИКОН 9.3

УПОРЕДНИ ПРИКАЗ СМАЊЕЊА ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА..... 56

ГРАФИКОН 9.4

УПОРЕДНИ ПРИКАЗ СМАЊЕЊА ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M MS АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА..... 57

ГРАФИКОН 9.5

УПОРЕДНИ ПРИКАЗ ПРОСЕЧНОГ СМАЊЕЊА СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ LBVA АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 59

ГРАФИКОН 9.6

УПОРЕДНИ ПРИКАЗ ПРОСЕЧНОГ СМАЊЕЊА СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ MCSA АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 59

ГРАФИКОН 9.7

УПОРЕДНИ ПРИКАЗ ПРОСЕЧНОГ СМАЊЕЊА СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 60

ГРАФИКОН 9.8

УПОРЕДНИ ПРИКАЗ ПРОСЕЧНОГ СМАЊЕЊА СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M MS АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 60

ГРАФИКОН 9.9

УПОРЕДНИ ПРИКАЗ БРОЈА СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ LBVA АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 62

ГРАФИКОН 9.10

УПОРЕДНИ ПРИКАЗ БРОЈА СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ MCSA АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 62

ГРАФИКОН 9.11

УПОРЕДНИ ПРИКАЗ БРОЈА СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ S&M АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 63

ГРАФИКОН 9.12

УПОРЕДНИ ПРИКАЗ БРОЈА СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ S&M MS АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА. 63

ГРАФИКОН 9.13

WER ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 40\%$]. 67

ГРАФИКОН 9.14

CF ФАКТОР ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 40\%$]. 67

ГРАФИКОН 9.15

WER ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 30\%$]. 68

ГРАФИКОН 9.16

CF ФАКТОР ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 30\%$]. 68

ГРАФИКОН 9.17

WER ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 20\%$]..... 69

ГРАФИКОН 9.18

SF ФАКТОР ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 20\%$]. 69

ГРАФИКОН 9.19

WER У ЗАДАТКУ СЕЛЕКЦИЈЕ ГАУСИЈАНА ПРИ ПРОМЕНИ ТИПА РАСТОЈАЊА КОРИШЋЕНОГ ПРИ ДЕКОДОВАЊУ ($n_{avr} = 50$). 72

ГРАФИКОН 9.20

SF ФАКТОР У ЗАДАТКУ СЕЛЕКЦИЈЕ ГАУСИЈАНА ПРИ ПРОМЕНИ ТИПА РАСТОЈАЊА КОРИШЋЕНОГ ПРИ ДЕКОДОВАЊУ ($n_{avr} = 50$). 72

ГРАФИКОН 9.21

УПОРЕДНИ ПРИКАЗ ПЕРФОРМАНСИ АЛГОРИТМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И SF ФАКТОРА ПРИ ПРОМЕНИ ПРОСЕЧНОГ БРОЈА ГАУСИЈАНА ПО ЈЕДНОМ КЛАСТЕРУ. СИСТЕМ КОРИСТИ МАХАЛАНОВИСОВО РАСТОЈАЊЕ У ПРОЦЕСУ ЕВАЛУАЦИЈЕ. 74

ГРАФИКОН 9.22

УПОРЕДНИ ПРИКАЗ ПЕРФОРМАНСИ АЛГОРИТМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И SF ФАКТОРА. ВРШИ СЕ ЕВАЛУАЦИЈА САМО ЈЕДНОГ КЛАСТЕРА ПО ФРЕЈМУ, ПРИ ЧЕМУ КЛАСТЕРИ ЗАУЗИМАЈУ ЗНАЧАЈАН ДЕО ПРОСТОРА ОБЕЛЕЖЈА. СИСТЕМ КОРИСТИ МАХАЛАНОВИСОВО РАСТОЈАЊЕ У ПРОЦЕСУ ЕВАЛУАЦИЈЕ..... 75

Листа скраћеница

CSR	Continuous Speech Recognition	Континуално препознавање говора
EDGS	Eigenvalues Driven Gaussian Selection	Селекција Гаусијана применом карактеристичних вредности
EM	Expectation Maximization	EM алгоритам
GMM	Gaussian Mixture Model	Модел Гаусове смеше
GS	Gaussian Selection	Селекција гаусијана (метод)
HGMMC	Hierarchical Clustering of Gaussian Mixture Models	Хијерархијско кластеровање модела Гаусових смеша (алгоритам)
KL	Kullback-Leibler (divergence)	KL дивергенца
LBVA	Lower Bound Variational Approximation	-
MCSA	Monte Carlo Sampling Approximation	-
MFCC	Mel Frequency Cepstral Coefficients	Мел фреквенцијски кепстрални коефицијенти
ML	Maximum Likelihood	Критеријум максималне веродостојности
OWA	Ordered Weighted Averaging	Оператор усредњавања
S&M	Split & Merge	(Операција) поделе и спајања
TBC	Tree Based Clustering	-
VQ	Vector Quantization	(Алгоритам) векторске квантизације

Захвалница

Захваљујем се ментору, проф. др Владу Делићу, како на корисним и конструктивним саветима које ми је пружио у току израде докторске дисертације и целокупних докторских студија, тако и на подршци, поверењу, љубазности, отворености и разумевању.

Захваљујем се колегама и сарадницима, пре свега Марку Јаневу, Милану Сечујском и Никши Јаковљевићу, који су својим стручним залагањем значајно допринели квалитету рада и саме докторске дисертације, АлфаНум тиму, као и члановима Катедре за телекомуникације и обраду сигнала.

Посебну захвалност дугујем породици и пријатељима.

Истраживање је подржано средствима Министарства просвете и науке (пројекат *Говорна комуникација човек-машина* ТР 11001 и пројекат *Развој дијалогских система за српски и друге јужнословенске језике* ТР 32035).

Бранислав Поповић

Поглавље 1

Увод

Модели Гаусових смеша (*Gaussian Mixture Model*, GMM) широко су распрострањени у апликацијама за обраду слике и говора [1] [2], препознавање емоција, препознавање говорника и говора [3]. Користе се и у процесу кластеровања података, као и за естимацију густине вероватноћа. Процедура кластеровања представља једну од најзначајнијих компоненти било ког задатка препознавања [4]-[6]. Хијерархијско кластеровање модела Гаусових смеша [7] [8] (*Hierarchical Clustering of Gaussian Mixture Models* – HGMMC) кључна је компонента метода селекције гаусијана [9]-[12], који се примењује за повећање брзине система за препознавање говора (говорника или емоција) базираном на моделима Гаусових смеша, и то посебно у апликацијама за континуално препознавање говора (*Continuous Speech Recognition* – CSR), верификацију гласом [3] и адаптацију на говорника [13].

Алгоритам селекције гаусијана спроводи се у две одвојене фазе. У оквиру прве фазе, врши се груписање гаусијана, компоненти оригиналне Гаусове смеше, а на основу карактеристичних вредности њихових коваријансних матрица. Након груписања врши се кластеровање свих Гаусових компоненти које се користе у систему препознавања, као и додела јединствених хипер-гаусијана сваком од ових кластера. Параметри хипер-гаусијана одређују се директно. Алгоритам кластеровања заснован је на примени алгоритма векторске квантизације и извршава се појединачно за сваку од ових група.

Друга фаза (фаза декодовања) подразумева директан прорачун вероватноћа, само и искључиво оних Гаусових компоненти које одговарају предефинисаном проценту кластера, чији хипер-гаусијани имају највеће вероватноће за дати улазни вектор обележја. Преосталим Гаусовим компонентама вероватноће се додељују апроксимативно, коришћењем унапред договорених вредности. Најчешће се додељује вероватноћа хипер-гаусијана коме је одговарајућа Гаусова компонента придружена, одређена за дати улазни вектор обележја.

Сврха HGMMC алгоритма јесте кластеровање компоненти полазне Гаусове смеше у циљу добијања компактне, тј. поједностављене репрезентације оригиналне Гаусове смеше, уз ограничење да ниједна компонента оригиналне Гаусове смеше не може бити подељена у оквиру редуковане репрезентације. Технике селекције гаусијана осмишљене су зарад повећања брзине система, тј. смањења рачунске сложености, или тачније, смањења потребног броја рачунских операција на уштрб незнатног смањења тачности. То је омогућено директним прорачуном вероватноћа, искључиво за селековани подскуп гаусијана, за дати улазни вектор обележја.

Метод селекције гаусијана први пут је описан у раду [8]. У овом раду изложена је основна идеја алгоритма селекције гаусијана, која подразумева оптимизацију система смањењем потребног броја рачунских операција, директним прорачуном вероватноћа за мањи број полазних Гаусијана. Алгоритам је редефинисан и након тога ефикасно примењен на проблем континуалног препознавања говора у систему заснованом на примени скривених Марковљевих модела [9]. Међутим, метод описан у раду [9] није укључио пуне коваријансне матрице у процес кластеровања и селекције. Стога је додатно уопштен у радовима [10] [11] и [12]. Додатна побољшања алгоритма селекције гаусијана представљена су у раду [14]. Математички аспекти сличног, али донекле формалнијег приступа, представљени у раду [7], искоришћени су као основ развоја ефикаснијег алгоритма кластеровања Гаусових компоненти који се може применити у било ком горе поменутом задатку препознавања, што је задатак докторске дисертације [15].

HGMMC алгоритам предложен у раду [7] примењен је као полазни, односно референтни алгоритам, у циљу поређења експерименталних резултата. Имајући у виду да је овај алгоритам заснован на алгоритму векторске квантизације, тј. *Lindo-Buzo-Gray* приступу [16] [17], (другим речима, алгоритам је сличан *k-means* алгоритму), он конвергира само ка локално оптималном решењу након коначног броја извршених итерација. Слична особина важи и за друге методе селекције гаусијана тренутно присутне у литератури ([5] [6] [8]).

Поред метода селекције гаусијана постоје разни други приступи убрзању система, засновани на примени модела Гаусових смеша, као што су технике просторног повезивања [18] [19] и [20], технике које користе *Kd* стабла

(структуре које обезбеђују партиционисање простора обележја, тј. организацију тачака k -димензионалног простора, [21] [22] [23]) и друге.

У оквиру докторске дисертације, предложен је нови алгоритам хијерархијског кластеровања модела Гаусових смеша (S&M HGMMC) [15], заснован на сукцесивном дељењу и међусобном повезивању кластера (*Split & Merge*, S&M). Сврха овог алгоритма јесте да се у процесу кластеровања, итеративним дељењем и међусобним повезивањем компонентних кластера Гаусове смеше, добијених након конверзије референтног алгоритма, оствари решење ближе глобалном оптимуму посматране функције циља. Полазне, локално оптималне вредности, одређене су иницијалном констелацијом, тј. коришћењем референтног HGMMC алгоритма, представљеног у раду [7]. Предложени метод успешно је примењен у задатку селекције гаусијана, при чему се остварује знатно повољнији однос брзине и тачности у задатку препознавања (посебно у случају апликације за континуално препознавање говора, у односу на референтни алгоритам [7], као и њему сличне алгоритме.

1.1. Мотивација и предмет истраживања

Мотивација алгоритма предложеног у оквиру докторске дисертације произилази из истраживања спроведеног у раду [24], у коме аутори предлажу решење у правцу превазилажења проблема локалне конвергенције стандардног EM (*Expectation Maximization*) алгоритма за естимацију параметара модела Гаусових смеша, засновано на сукцесивном извођењу операција поделе и спајања. Неколико других истраживача такође се бавило проналажењем критеријума за ефикасну селекцију кандидата за поделу и спајање [25] приликом одређивања параметара GMM модела на основу посматраног скупа података, применом EM алгоритма.

Поменуте идеје нашироко се користе у областима као што су компјутерска визија, препознавање облика и обрада сигнала. У свим овим радовима, операција спајања је добро постављена операција, с обзиром да за дату операцију постоји јединствено решење. С друге стране, предложена операција поделе уобичајено је локална Kullback-Leibler (KL) дивергенца. У том смислу, врши се подела гаусијана са највећом локалном KL

дивергенцом. То је лоше постављен (*ill-posed*) проблем, с обзиром да за тако постављен проблем не постоји решење у затвореној форми.

У оквиру докторске дисертације, концепт сукцесивне примене операција поделе и спајања, иницијално предложен у раду [24], примењен је на случај HGMMC алгоритма предложеног у раду [7], који одговара процесу селекције гаусијана. Као критеријум спајања, одабрана је минимална KL дивергенца између два произвољна гаусијана, одређена у процесу кластеровања. За предложену операцију спајања решење постоји у затвореној форми, стога је она добро постављена. Предложен је оптималан начин за проналажење параметара хипер-гаусијана који одговара резултујућем кластеру.

С друге стране, уопштење критеријума поделе веома је проблематично. Кластер са највећом KL дивергенцом између придруженог хипер-гаусијана и припадајућег GMM модела намеће се као природан избор. Међутим, решење тог проблема у виду израза за KL дивергенцу између две Гаусове смеше не постоји у затвореној форми, изузев у тривијалном случају када се обе Гаусове смеше састоје из само једне Гаусове компоненте. Стога су, у циљу добијања једноставне и изводљиве имплементације метода сукцесивне поделе и спајања, искоришћене две различите апроксимације, обе предложене у раду [26]. Прва апроксимација (*Monte Carlo Sampling Approximation*, MCSA) прецизнија је од друге апроксимације (*Lower Bound Variational Approximation*, LBVA), која је рачунски далеко мање захтевна, али и даље довољно прецизна. Полазећи од чињенице да је операција поделе лоше постављен проблем, у овом раду предложено је једно могуће решење тог проблема. Штавише, описан је начин укључивања механизма селекције модела у оквиру предложеног S&M HGMMC алгоритма, чиме се остварује бољи однос рачунске сложености и тачности, у случају када је алгоритам кластеровања примењен у оквиру задатка селекције гаусијана (GS), у реалном систему препознавања.

Експерименти су планирани и након тога извршени у две одвојене фазе. У оквиру прве фазе, спроведени су експерименти над синтетичким подацима. У оквиру друге фазе, експерименти се изводе у оквиру реалног GS система, коришћеног у задатку континуалног препознавања говора (CSR).

У експериментима спроведеним над вештачки генерисаним подацима, за предложени S&M HGMMC алгоритам, остварено је побољшање у виду смањења вредности функције цене у великом проценту случајева, за сваку од тестираних димензионалности простора, као и за обе апроксимације KL

дивергенце које су коришћене у оквиру експеримената. Када су у питању експерименти спроведени у оквиру реалног GS система, постиже се бољи однос тачности препознавања с једне, и рачунске сложености с друге стране, у случају примене предложеног S&M HGMMC алгоритма, у односу на референтни HGMMC алгоритам. Још бољи резултати постижу се применом механизма селекције модела унутар предложеног алгоритма, као и одговарајућим одабиром полазних параметара алгоритма селекције гаусијана [14].

1.2. Организација рада

Рад је организован на следећи начин:

- Поглавље 2

Представљене су теоријске основе докторске дисертације. Поред прегледа литературе од значаја за тему докторске дисертације, објашњени су математички појмови и концепти неопходни за потпуно разумевање наредних поглавља. Представљени су модели Гаусових смеша и њихове особине, алгоритми кластеровања, уз посебан осврт на EM и k-means алгоритам, као и алгоритми засновани на примени операција поделе и спајања. Додатна објашњења везана за одабир критеријума поделе и спајања представљена су у оквиру петог поглавља.

- Поглавље 3

Детаљно је описан алгоритам селекције гаусијана, примењен у задатку континуалног препознавања говора. Описане су обе фазе алгоритма, метод груписања, начин декодовања, као и алгоритам кластеровања заснован на алгоритму векторске квантизације, чија су побољшања представљена у оквиру наредних поглавља.

- Поглавље 4

Изложена је дефиниција проблема и детаљно описан референтни алгоритам хијерархијског кластеровања модела Гаусових смеша, коришћен у процесу иницијализације предложеног алгоритма поделе и спајања у виду стартне тачке итерација поделе и спајања, као и за поређење експерименталних резултата у оквиру деветог поглавља.

- Поглавље 5

Предложен је нови алгоритам за хијерархијско кластеровање модела Гаусових смеша, применом операција поделе и спајања. Представљени су одговарајући критеријуми и детаљно описане примењене апроксимације. Овај алгоритам може се посматрати и као уопштење алгоритма за естимацију параметара модела Гаусових смеша, описаних у оквиру првог поглавља, примењено на проблем хијерархијског кластеровања модела Гаусових смеша. Описане су операције поделе и спајања, као и методе прорачуна при свакој итерацији алгоритма. Детаљан приказ алгоритма по фазама приложен је у додатку.

- Поглавље 6

Предложен је начин за укључивање механизма селекције модела у алгоритам описан у оквиру петог поглавља. Одабир оптималног броја кластера врши се у само једном пролазу алгоритма, чиме се постижу знатно бољи резултати у смислу односа тачности и рачунске сложености у задатку континуалног препознавања говора.

- Поглавље 7

Дискусија о утицају полазних параметара система на перформансе алгоритма селекције гаусијана. Овај утицај посматра се у контексту одабира типа растојања коришћеног у процесу декодовања, броја хипер-гаусијана чији се придружени гаусијани евалуирају директно, оптималног броја кластера, као и одговарајућег типа кластеровања у смислу степена преклапања између појединачних кластера.

- Поглавље 8

Укратко су описане најзначајније карактеристике система за континуално препознавање говора, као и говорна база коришћена у задатку континуалног препознавања говора.

- Поглавље 9

Приказани су експериментални резултати настали применом предложених алгоритама поделе и спајања, као и применом референтног алгорита, како над вештачки генерисаним подацима, тако и у задатку селекције гаусијана у оквиру реалног система препознавања. Представљене су перформансе алгорита селекције гаусијана у зависности од одабраног скупа параметара.

- Поглавље 10

Дискусија везана за резултате описане у оквиру претходних поглавља, уз кратак осврт на даље правце примене предложених метода.

- Поглавље 11

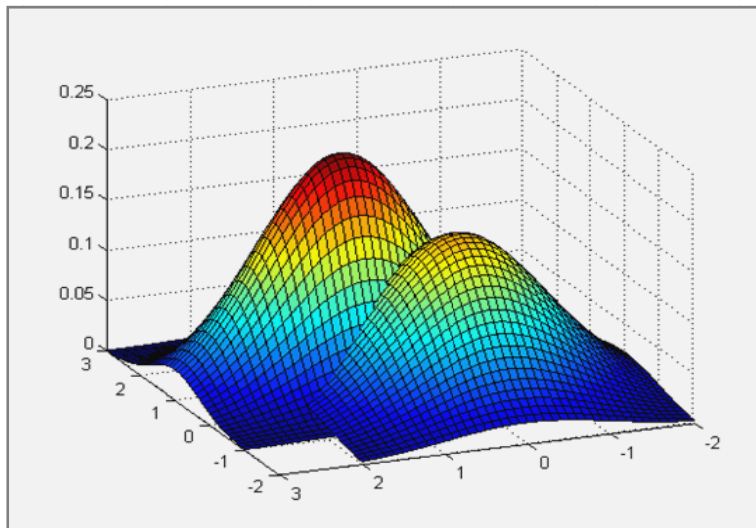
Приказ предложених алгоритама, као и референтног алгорита, по фазама.

Поглавље 2

Теоријске основе докторске дисертације

2.1. Модели Гаусових смеша

Модели Гаусових смеша спадају у групу параметарских расподела вероватноћа. То у практичном смислу подразумева претпоставку да се ради о случајним одбирцима генерисаним у складу са одговарајућом (у овом случају нормалном) расподелом, чији се параметри накнадно одређују у зависности од доступних опсервација.



СЛИКА 2.1 ТРОДИМЕНЗИОНАЛНИ ПРИКАЗ GMM МОДЕЛА

GMM модели уобичајено се користе при естимацији густина вероватноћа, као и у оквиру разноврсних биометријских система, углавном за представу појединачних, континуалних величина. Користе се и као незаобилазан алат у оквиру апликација заснованих на принципима неуронских мрежа, примени

скривених Марковљевих модела, у оквиру алгоритама кластеровања, па тако и у оквиру система за континуално препознавање говора (говорника или емоција). Естимација параметара GMM модела уобичајено се изводи применом EM алгорита, односно ML критеријума, описаног у одељку 2.3.

GMM модели математички се описују као пондерисане суме компонентних гаусијана. Представљени су следећим изразом:

$$f(y) = \sum_{i=1}^k \alpha_i N(y; \mu_i, \Sigma_i) = \sum_{i=1}^k \alpha_i f_i(y) \quad (2.1)$$

при чему је y вектор обележја, α_i , $i = 1, \dots, k$ су тежине Гаусових компоненти, док су $f_i(y) = N(y; \mu_i, \Sigma_i)$, $i = 1, \dots, k$ компонентни гаусијани.

Свака Гаусова компонента представљена је изразом

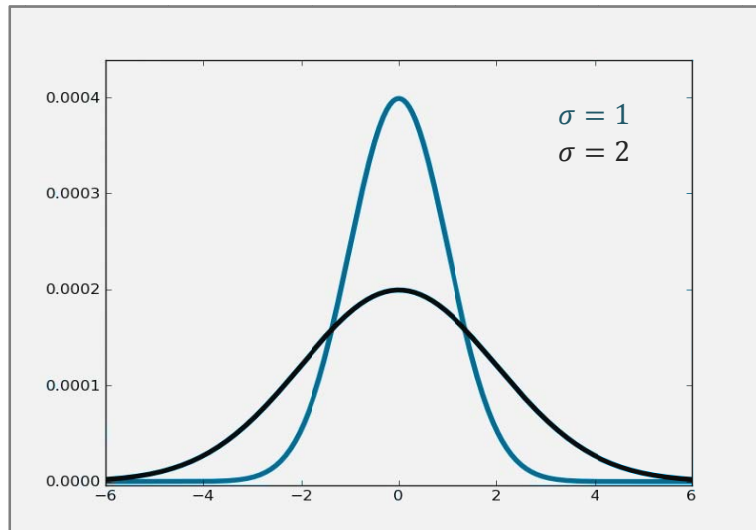
$$f_i(y) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (y - \mu_i)^T \Sigma_i^{-1} (y - \mu_i) \right\} \quad (2.2)$$

при чему је μ_i вектор средњих вредности, Σ_i коваријансна матрица, d је димензионалност простора, T је оператор транспоновања.

Параметри μ и Σ довољни су за јединствени опис нормалне (Гаусове) расподеле, што је један од главних разлога њене велике популарности и широке применљивости. Маргиналне и условне расподеле такође су нормалне, линеарна трансформација N случајних променљивих које имају здружену Гаусову расподелу остаје здружено Гаусова, а важи и централна гранична теорема описана у наставку (теорема у наставку односи се на случај једнодимензионалне Гаусове расподеле). Једнодимензионална нормална расподела приказана је на Слици 2.2.

ТЕОРЕМА 2.1

Нека је дата произвољна расподела чија је средња вредност μ и варијанса σ^2 . Расподела узорачке средње вредности тежи нормалној расподели са средњом вредношћу μ и варијансом σ^2/N када величина узорка N расте.



СЛИКА 2.2 ЈЕДНОДИМЕНЗИОНАЛНА НОРМАЛНА (ГАУСОВА) РАСПОДЕЛА

Другим речима, централна гранична теорема тврди да сума N независних случајних променљивих са истом расподелом тежи ка нормалној расподели када је N велико. Постоје и знатно општије формулације централне граничне теореме, као на пример формулација која дозвољава корелисане случајне променљиве, као и формулација теореме која се може применити и у вишедимензионалном случају. Докази ових теорема доступни су у литератури [27].

КАРАКТЕРИСТИЧНЕ ВРЕДНОСТИ И КАРАКТЕРИСТИЧНИ ВЕКТОРИ

ДЕФИНИЦИЈА 2.1

Посматрајмо квадратну матрицу A . v је карактеристични вектор матрице A ако постоји скалар λ такав да важи:

$$Av = \lambda v \quad (2.3)$$

λ је карактеристична вредност матрице.

Карактеристичне вредности матрице A одређују се решавањем карактеристичне једначине [28]

$$|A - \lambda I| = 0 \tag{2.4}$$

$$\lambda^N + a_1 \lambda^{N-1} + \dots + a_{N-1} \lambda + a_0 = 0$$

где је $|\cdot|$ ознака за детерминанту.

ДЕФИНИЦИЈА 2.2

Модална матрица матрице A представља матрицу формирану од карактеристичних вектора постављених један до другог.

$$A_{mod} = \begin{bmatrix} \uparrow & \uparrow & \uparrow & \uparrow \\ v_1 & v_2 & \dots & v_N \\ \downarrow & \downarrow & \downarrow & \downarrow \end{bmatrix} \tag{2.5}$$

ДЕФИНИЦИЈА 2.3

Канонска форма матрице A која има једноструке карактеристичне корене представља дијагоналну матрицу формирану постављањем карактеристичних вредности матрице A по главној дијагонали.

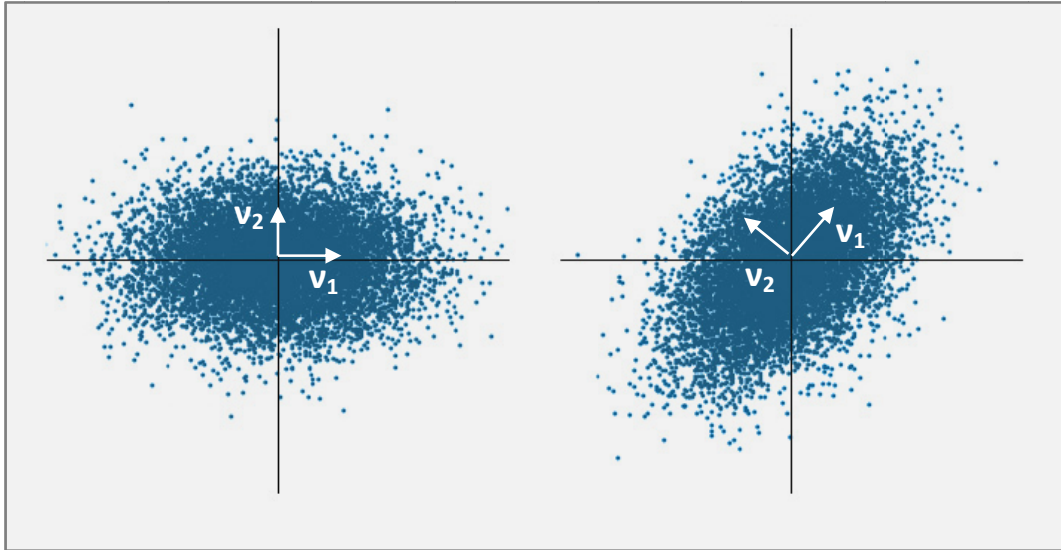
$$A_{can} = \begin{bmatrix} \lambda_1 & & \\ & \lambda_1 & \\ & & \lambda_N \end{bmatrix} \tag{2.6}$$

ДЕФИНИЦИЈА 2.4

Нека је A квадратна и симетрична матрица. Кажемо да је матрица A позитивно дефинитна ако је $\varphi^T A \varphi > 0$ за свако $\varphi \neq 0$.

Карактеристичне вредности реалних и симетричних матрица су реални бројеви, док су карактеристични вектори придружени различитим

карактеристичним вредностима линеарно независни. Уколико је матрица A позитивно дефинитна и симетрична, њене карактеристичне вредности су позитивне.



СЛИКА 2.3 ГЕОМЕТРИЈСКА ИНТЕРПРЕТАЦИЈА КАРАКТЕРИСТИЧНИХ ВЕКТОРА И КАРАКТЕРИСТИЧНИХ ВРЕДНОСТИ

Уколико је реч о Гаусовој расподели, карактеристични вектори коваријансне матрице Σ представљају главне правце расподеле, док карактеристичне вредности представљају варијансе дуж одговарајућих главних праваца, као што је приказано на слици 2.3.

2.2. Естимација параметара модела Гаусових смеша

Expectation Maximization (EM) алгоритам представља једноставну, итеративну процедуру за одређивање *Maximum Likelihood* (ML) естимације параметара, у ситуацији када један део података недостаје [29]. Под ML естимацијом подразумева се одређивање параметара претпостављене расподеле применом критеријума максималне веродостојности. Овај критеријум описан је у наставку.

ML естимација полази од претпоставке да су параметри расподеле фиксни, али непознати. Усвајају се оне вредности параметара које се најбоље слажу са доступним узорцима из расподеле.

Претпоставимо да је дат скуп узорака расподеле чије параметре одређујемо $X = \{x_1, \dots, x_K\}$. За дати скуп узорака потребно је одредити параметре модела који максимизују логаритам веродостојности података (*Log Likelihood*).

$$\begin{aligned}\hat{\theta} &= \operatorname{argmax}[p(X|\theta)] = \operatorname{argmax} \left[\sum_{k=1}^K \log p(x_k|\theta) \right] \\ \hat{\theta} &= \operatorname{argmax}[p(X|\theta)] = \operatorname{argmax} \left[\sum_{k=1}^K \log \sum_{m=1}^M \alpha_m p(x_k|\theta_m) \right]\end{aligned}\tag{2.7}$$

Параметри $\theta_m = (\mu_m, \Sigma_m)$ и α_m представљају вектор средњих вредности, коваријансну матрицу и тежину компонентог гаусијана, која се у овом случају може интерпретирати и као апостериорна вероватноћа. $p(X|\theta)$ представља расподелу чији се непознати параметри одређују на основу познатих узорака.

Изједначавањем првих извода са нулом, може се показати да је ML естимација параметара вишедимензионалне нормалне расподеле једнака узорачкој средњој вредности и узорачкој коваријансној матрици

$$\hat{\mu} = \frac{1}{K} \sum_{k=1}^K x_k\tag{2.8}$$

$$\hat{\Sigma} = \frac{1}{N} \sum_{k=1}^K (x_k - \hat{\mu})(x_k - \hat{\mu})^T$$

где је T оператор транспоновања [29].

У случају да се ради о GMM моделу, изједначавањем првих извода са нулом добија се нелинеаран систем једначина. Максимизација логаритма

веродостојности изједначавањем првог извода са нулом у овом случају није могућа.

Процене параметара одређују се применом ЕМ алгоритма, тј. одабиром почетне вредности параметара модела и итеративним прерачунавањем израза са десне стране једначина, при чему се добијају нове вредности параметара. При свакој новој итерацији, користе се вредности параметара израчунате у оквиру претходне итерације. На тај начин, вредност логаритма веродостојности повећава се итеративно.

2.2.1 ЕМ алгоритам

ЕМ алгоритам примењује се у ситуацији када подаци заиста недостају, или су оштећени као последица неадекватног процеса прикупљања података, али исто тако и у случају када претпоставка о постојању параметара који недостају може у великој мери да поједностави функцију веродостојности, која би у противном довела до нерешивог оптимизационог проблема [25].

Посматрајмо GMM модел у d -димензионалном простору, $d \geq 1$, са $M > 1$ компонентних гаусијана.

$$p(x|\theta) = \sum_{m=1}^M \alpha_m p(x|\theta_m) \quad (2.9)$$

$$\alpha_m \geq 0 \quad \sum_{m=1}^M \alpha_m = 1$$

$p(x|\theta_m)$ представља d -димензионалну нормалну расподелу, која одговара m -тој Гаусовој компоненти, која се записује на следећи начин

$$p(x|\theta_m) = \frac{1}{(2\pi)^{d/2} |\Sigma_m|^{1/2}} \exp \left\{ -\frac{1}{2} (y - \mu_m)^T \Sigma_m^{-1} (y - \mu_m) \right\} \quad (2.10)$$

Уколико би Гаусова компонента која је генерисала узорак x била позната, одређивање непознатих параметара $\theta = \{\alpha_1, \dots, \alpha_m, \theta_1, \dots, \theta_M\}$, при чему је $\theta_m = \{\mu_m, \Sigma_m\}$, не би представљало проблем. Такође, уколико би нам били познати параметри θ , лако бисмо одредили расподелу која је генерисала узорак x . Нажалост, није нам познато ни једно ни друго.

Нека је дат скуп обележја $X = \{x_1, \dots, x_K\}$. Комплетан скуп података $Z = (X; Y)$ састоји се из скупа обележја X чија је вредност позната (непотпуни подаци) и скупа обележја Y чија нам вредност није позната (недостајући подаци). Естимација параметара ЕМ алгоритмом врши се sukcesивном применом корака очекивања (Е корак) и корака максимизације (М корак).

У оквиру Е корака, израчунава се очекивана вредност логаритма веродостојности комплетног скупа података, при чему је позната тренутна естимација непознатих параметара $\theta^{(t)}$ након извршеног i -тог корака.

$$Q(\theta|\theta^{(i)}) = \sum_{k=1}^K \sum_{m=1}^M P(m|x_k; \theta^{(i)}) \log \alpha_m p(x_k|\theta_m) \quad (2.11)$$

$P(m|x_k; \theta^{(i)})$ представља вероватноћу да је узорак x_k генерисала m -та Гаусова компонента, при тренутној естимацији параметара $\theta^{(i)}$.

$$P(m|x_k; \theta^{(i)}) = \frac{\alpha_m^{(i)} p(x_k|\theta_m^{(i)})}{\sum_{l=1}^M \alpha_l^{(i)} p(x_k|\theta_l^{(i)})} \quad (2.12)$$

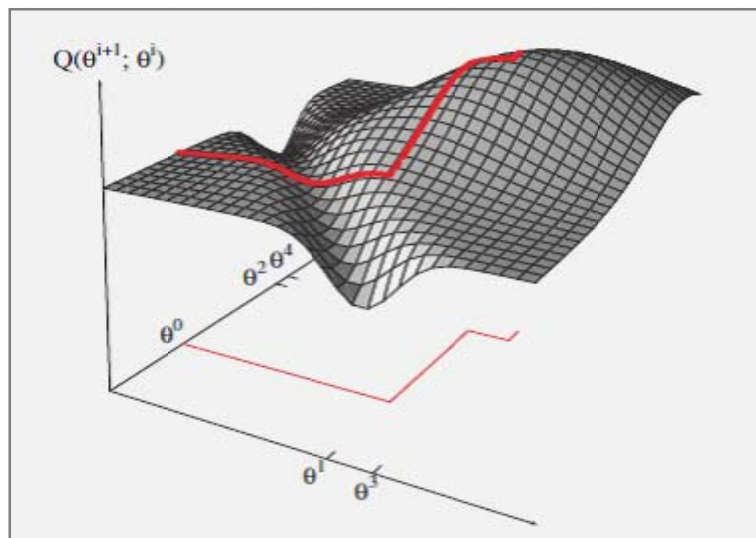
У оквиру М корака, врши се максимизација вредности логаритма веродостојности комплетног скупа података, при чему се добија нова процена вредности параметара $\theta^{(i+1)}$.

$$\alpha_m^{(i+1)} = \frac{1}{K} \sum_{k=1}^K P(m|x_k; \theta^{(i)}) \quad (2.13)$$

$$\mu_m^{(i+1)} = \frac{\sum_{k=1}^K x_k P(m|x_k; \theta^{(i)})}{\sum_{k=1}^K P(m|x_k; \theta^{(i)})} \quad (2.14)$$

$$\Sigma_m^{(i+1)} = \frac{\sum_{k=1}^K (x_k - \mu_m^{(i+1)})(x_k - \mu_m^{(i+1)})^T P(m|x_k; \theta^{(i)})}{\sum_{k=1}^K P(m|x_k; \theta^{(i)})} \quad (2.15)$$

На Слици 2.4 илустровани су кораци ЕМ алгоритма [29].



СЛИКА 2.4 КОРАЦИ ЕМ АЛГОРИТМА

Претрага започиње одабиром иницијалих вредности параметара модела θ^0 . Применом M корака, проналазе се оптималне вредности параметара θ^1 . Затим се одређују вредности параметара θ^2 које максимизују очекивану вредност логаритма веродостојности, при чему се претходно израчунате вредности параметара θ^1 посматрају као константе.

Приказ ЕМ алгоритма

– Иницијализација:

- За предефинисано $\varepsilon > 0$ одабрати иницијално решење $\theta = \theta^0$. $i = 0$.

– Итерације:

- Е корак:

$$\text{Одредити } Q(\theta|\theta^{(i)})$$

- М корак:

$$\theta^{(i+1)} \leftarrow \operatorname{argmax} Q(\theta|\theta^{(i)})$$

- Понављати Е и М корак све док се не добије

$$Q(\theta^{(i+1)}|\theta^{(i)}) - Q(\theta^{(i)}|\theta^{(i-1)}) \leq \varepsilon$$

- Усвојити коначно решење $\hat{\theta} = \theta^{(i+1)}$

Показује се да свака итерација ЕМ алгоритма (Е и М корак) повећава логаритам веродостојности, тј. да ЕМ алгоритам гарантовано конвергира ка локалном максимуму функције веродостојности комплетног скупа података [29].

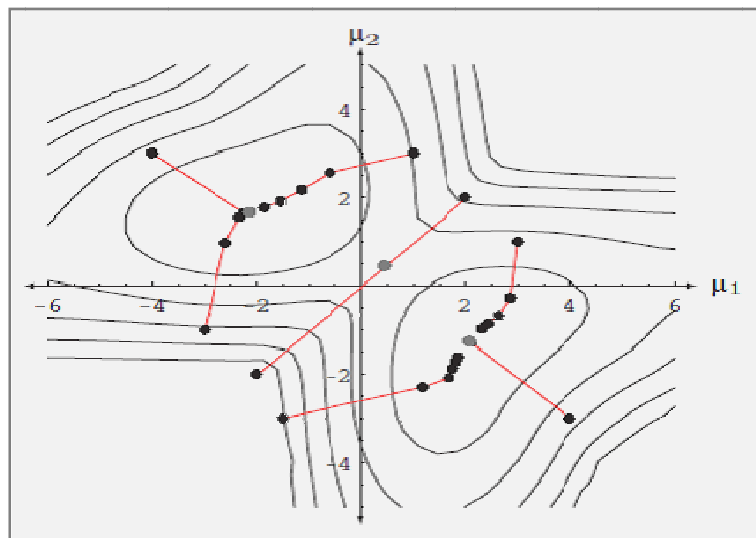
Проблем ЕМ алгоритма управо је чињеница да овај алгоритам конвергира само ка локално оптималном решењу, док је сам алгоритам веома осетљив на иницијални избор параметара [25]. Стога се у процесу иницијализације често примењује нека од метода векторске квантизације, попут k-means алгоритма.

2.2.2 K-means алгоритам

K-means алгоритам спада у групу алгоритама нехијерархијског кластеровања [29]. Може се посматрати и као специјалан случај ЕМ алгоритма за моделе Гаусових смеша. Представља итеративну процедуру кластеровања којом се вредност критеријумске функције постепено смањује.

K-means алгоритам започиње иницијализацијом броја кластера. Познати узорци расподеле произвољно се додељују сваком од кластера, или се произвољно бирају центри кластера, случајним избором међу постојећим

узорцима расподеле. Након тога, врши се поновно прерачунавање центара кластера. Сваки од узорака додељује се њему најближем кластеру. Процедура се итеративно понавља, све док има промене стања између две узастопне итерације. Имајући у виду чињеницу да и сам k-means алгоритам конвергира само ка локално оптималном решењу, примена овог алгоритма убрзава конвергенцију EM алгоритма, али свакако не гарантује конвергенцију ка глобалном оптимуму [25].



СЛИКА 2.5 ИТЕРАЦИЈЕ K-MEANS АЛГОРИТМА

Један од начина да се реши овај проблем јесте да се одабере неколико различитих скупова иницијалних вредности параметара који се одређују применом EM алгоритма. EM алгоритам спроводи се над сваким од ових скупова појединачно и бира се најповољније решење. На тај начин, повећава се вероватноћа да ће EM алгоритам конвергирати ка глобалном оптимуму, али се тиме истовремено знатно повећава рачунска сложеност алгоритма. Итерације k-means алгоритма приказане су на Слици 2.5 [29].

2.3. Алгоритми поделе и спајања

Други начин за превазилажење проблема локалне конвергенције стандардног EM алгоритма јесте примена операција поделе и спајања. У

раду [24], предложена је примена ових операција након конвергенције стандардног EM алгоритма. При свакој итерацији поделе и спајања врши се прво локална, затим и глобална оптимизација, поновном применом EM алгоритма. На тај начин добија се решење ближе глобалном оптимуму посматране критеријумске функције, и истовремено се остварује мања рачунска сложеност у односу на решење предложено у оквиру претходног одељка. Један од проблема при дефинисању оваквог алгоритма свакако је избор одговарајућих критеријума поделе и спајања. Алгоритам естимације применом критеријума поделе и спајања описан је у наставку.

Приказ EM S&M алгоритма

– Иницијализација:

- Применити стандардни EM алгоритам до конвергенције.

– Итерације:

- Сортирати кандидате за операције поделе и спајања на основу одабраних критеријума.
- Применити операције поделе и спајања и извршити реестимацију параметара.
- Применити стандардни EM алгоритам, иницијализован новим вредностима параметара. (Парцијално за новоформиране гаусијане, а потом и за цео скуп).
- Понављати поступак докле год вредност критеријумске функције опада.
- Усвојити коначно решење.

У већини радова на тему естимације параметара модела Гаусових смеша применом операција поделе и спајања, као критеријум поделе, предлаже се локална KL дивергенца, описана у оквиру петог поглавља (на пример [24][25]). У раду [24], као критеријум спајања, предложен је

критеријум заснован на прорачуну апостериорних вероватноћа појединачних гаусијана.

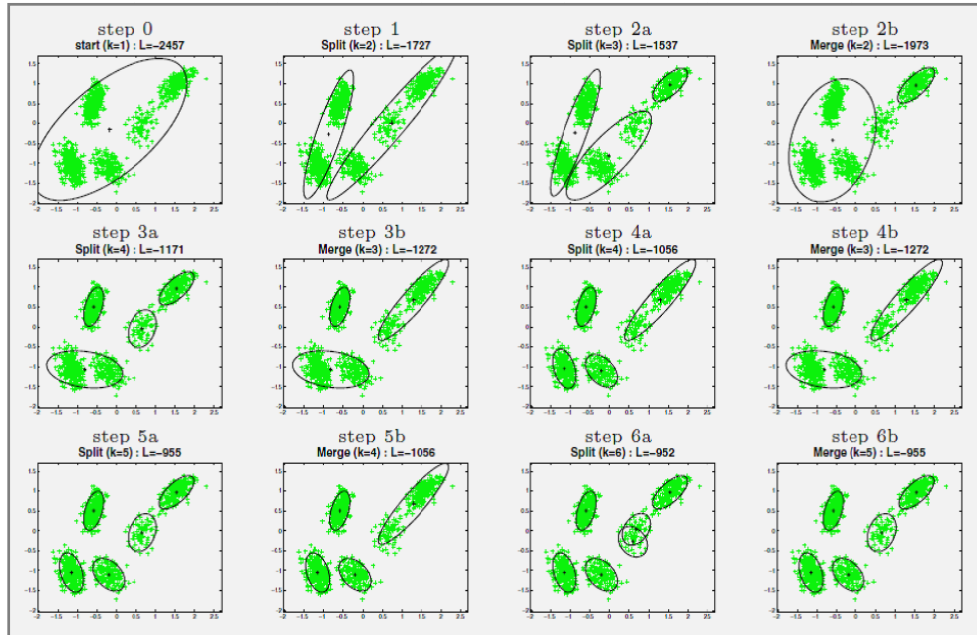
$$J_{merge}(i, j; \theta^*) = P_i(\theta^*)^T P_j(\theta^*) \quad (2.16)$$

при чему је $P_i(\theta^*)$

$$P_i(\theta^*) = (P(i|x_1\theta^*), \dots, P(i|x_K\theta^*))^T \quad (2.17)$$

K -димензионални вектор апостериорних вероватноћа за i -ти гаусијан. Са становишта процедуре кластеровања, овај критеријум није проблематичан [25].

Број Гаусових компоненти често је иницијално задат, и не мења се у току једног извођења алгоритма, изузев у оквиру појединачне S&M итерације. С друге стране, у раду [30], предложена је имплементација алгоритма поделе и спајања, у оквиру које се број Гаусових компоненти итеративно повећава. Кораци овог алгоритма приказани су на Слици 2.6 [30].



СЛИКА 2.6 КОРАЦИ АЛГОРИТМА ПОДЕЛЕ И СПАЈАЊА (Blekas, Lagaris)

Горе поменути алгоритми за естимацију параметара модела Гаусових смеша представљају како мотивацију, тако и основ за дефинисање новог алгоритма поделе и спајања, примењеног у задатку кластеровања GMM модела.

Референтни алгоритам, описан у Поглављу 4, примењен је у оквиру предложеног алгоритма хијерархијског кластеровања аналогно EM алгоритму. Истовремено, потребно је дефинисати аналогне критеријуме поделе и спајања, што је у случају критеријума поделе води ка нерешивом оптимизационом проблему. С обзиром да решење тог проблема у виду израза за KL дивергенцу између две Гаусове смеше не постоји у затвореној форми, примењују се различите апроксимације, описане у оквиру наредних поглавља.

Поглавље 3

Метод селекције гаусијана

У системима за континуално препознавање говора заснованим на примени скривених Марковљевих модела уобичајено се користе расподеле континуалног типа, као што је нормална (Гаусова) расподела. Свако НММ стање моделује се великим бројем Гаусових компоненти, које се при одређивању вероватноћа стања засебно евалуирају, што ове системе чини неупотребљивим за апликације у реалном времену. Прорачун вероватноћа стања представља значајан део укупног броја рачунских операција (процењује се између 30 и 70%).

У данашње време постоји велики број техника које се примењују у циљу смањења потребног броја рачунских операција и самим тим, повећања брзине система за континуално препознавање говора (говорника, емоција и слично). То се постиже смањењем димензионалности (линеарна дискриминантна анализа), повезивањем акустичких стања применом полуконтинуалних скривених Марковљевих модела, применом техника пројекционе претраге и бројних других.

Метод селекције гаусијана [8] представља алтернативни приступ убрзању система, при чему се време извршавања алгорита препознавања смањује директним одређивањем вероватноћа само за одређени подскуп гаусијана, за дати улазни вектор обележја, док се преостале вероватноће попуњавају апроксимативним вредностима. Сврха алгорита селекције гаусијана јесте повећање брзине система за континуално препознавање говора (говорника, емоција и слично), при чему перформансе система остају непромењене, или је деградација перформанси у односу на полазни систем занемарљиво мала.

У раду [12], аутори предлажу измењени приступ алгоритму селекције гаусијана, који решава проблем који настаје применом овог алгорита у случајевима када је преклапање између гаусијана у оквиру полазне констелације велико. У системима за континуално препознавање говора

такво преклапање неизбежно је и готово увек присутно. У том смислу, поред алгоритма кластеровања, предлаже се груписање полазних Гаусових компоненти у оквиру прве фазе алгоритма.

Основна идеја је да се сваки од гаусијана из полазне констелације, пре него што се изврши процедура кластеровања, придружи једној од неколико унапред дефинисаних група. Придруживање се врши у зависности од вредности израчунате агрегацијом карактеристичних вредности коваријансне матрице посматране Гаусове компоненте, користећи OWA (*Ordered Weighted Averaging*) операторе. Овај метод предложен је у литератури под називом селекција гаусијана на основу сопствених вредности (*Eigenvalues Driven Gaussian Selection*, EDGS [12]).

Претпоставимо супротно. У случају да се кластеровање примени без претходно извршеног груписања, по завршетку процедуре, сваки од кластера могао би да садржи како гаусијане чије су карактеристичне вредности коваријансне матрице доминантно велике, тако и гаусијане чије су карактеристичне вредности коваријансне матрице доминантно мале. Ово је нарочито проблематично у случајевима када је преклапање између гаусијана унутар полазне констелације велико. У таквим случајевима долази до маскирања гаусијана чија је варијанса мала гаусијанима чија је варијанса велика. Уколико је вероватноћа посматраног хипер-гаусијана изнад предефинисаног прага, евалуација сваког од придружених гаусијана врши се директно за дати улазни вектор обележја. То практично значи да ће се вероватноће директно прорачунавати и за оне гаусијане чија је вероватноћа за дати улазни вектор обележја мала, што доводи не само до непотребног повећања рачунске сложености алгоритма препознавања, већ и до значајног повећања грешке препознавања.

Применом EDGS метода, остварује се значајно побољшање у смислу односа између тачности препознавања с једне, и рачунске сложености с друге стране. Овај метод комбинује најзначајније карактеристичне вредности коваријансних матрица, у циљу њиховог груписања у предефинисани број група, да би се тек након тога процедура кластеровања извршила за сваку од ових група појединачно. Највеће сопствене вредности уједно су и најзначајније са становишта груписања полазних гаусових компоненти.

Идеја је да се у процесу агрегације веће тежине придруже већим, тј. значајнијим сопственим вредностима коваријансне матрице сваког од

гаусијана. За дати вектор сопствених вредности $\lambda = (\lambda_1, \dots, \lambda_p)$, OWA оператор промењује се на следећи начин:

$$OWA_{\omega}(\lambda_1, \dots, \lambda_p) = \sum_{j=1}^p \omega_j \lambda_{\sigma(j)} \quad (3.1)$$

при чему је $0 \leq \lambda_{\sigma(1)} \leq \lambda_{\sigma(p)}$. Коефицијенти $\omega \in R^p$ задовољавају услов $0 \leq \omega_j \leq 1$. Њихова сума једнака је јединици.

OWA оператори припадају фамилији агрегационих оператора, у коју спадају и веома познати оператори, као што су минимум, максимум, медијана, аритметичка средина и други. Могу се посматрати као параметарске интерполације између минималне и максималне вредности у процесу агрегације. У случају алгоритма селекције гаусијана, потребно је да ова вредност буде нешто ближа максималној вредности, јер се на тај начин фаворизују најзначајније карактеристичне вредности у процесу груписања.

Сваки од гаусијана чије су карактеристичне вредности $\lambda = (\lambda_1, \dots, \lambda_p)$ придружује се групи g из скупа $g \in \{1, \dots, G\}$, ако и само ако важи да се $A_{\omega}(\lambda)$ налази унутар интервала $[\tau_{min}, \tau_{max}]$ предвиђеног за дату групу. $A_{\omega}(\cdot)$ је OWA оператор, тежине ω добијају се на начин описан у раду [12]. Циљ је да свака група обухвати само и искључиво оне гаусијане за које важи да је већина њихових најзначајнијих сопствених вредности унутар неког предефинисаног опсега.

По завршетку груписања полазних гаусијана врши се кластеровање засновано на алгоритму векторске квантизације (VQ), засебно за сваку од ових група. Један од таквих алгоритама јесте и VQGS алгоритам, примењен у раду [12]. VQGS алгоритам, као и предложени EDGS метод, приказани су у наставку, док су алгоритми кластеровања модела Гаусових смеша предложени у оквиру докторске дисертације, укључујући и референтни алгоритам, приказани у оквиру наредних поглавља.

Имајући у виду да су карактеристичне вредности коваријансних матрица унутар одговарајућег опсега, може се претпоставити да ће већина гаусијана унутар било ког кластера, за дати улазни вектор обележја, имати приближне вредности вероватноћа. Тачније, мало је вероватно да ће вероватноћа хипер-гаусијана за дати улазни вектор обележја бити велика, а вероватноће

великог броја појединачних гаусијана унутар датог кластера мале, или обратно. Таква ситуација значајно би угрозила перформансе система у погледу односа тачности и рачунске комплексности.

Друга фаза алгоритма селекције гаусијана јесте фаза декодовања. Након што је извршено груписање, а потом и кластероване полазних гаусијана, одређује се предефинисани проценат најбољих хипер-гаусијана у зависности од одабране мере блискости. Гаусијани чији су хипер-гаусијани унутар предефинисаног процента евалуирају се директно. Ова фаза детаљно је описана у оквиру седмог поглавља.

Приказ VQGS алгоритма

– Иницијализација:

- За предефинисани просечан број гаусијана по једном кластеру n_{avr} и укупан број гаусијана K , израчунати број кластера као $m = \left\lfloor \frac{K}{n_{avr}} \right\rfloor$
- На случај одабрати m различитих центроида и доделити им јединичне коваријансне матрице

– Итерације:

За предефинисани праг и критеријумску функцију

- Сваком од гаусијана доделити њему најближи хипер-гаусијан.
- Применом МЛ естимације одредити параметре новодобијених хипер-гаусијана
- У случају да се неки од хипер-гаусијана испразни у процесу кластерованја, смањити број кластера за један.

Понављати итерације све док вредност критеријумске функције не падне испод предефинисаног прага.

Приказ EDGS алгоритма

– Иницијализација:

- Специфицирати број група
- Одредити оптималне вредности тежина OWA оператора
- Одредити интервале $[\tau_{min}, \tau_{max}]$ предвиђене за сваку од група

– Груписање:

За сваки од K полазних гаусијана

- Одредити карактеристичне вредности коваријансних матрица полазних гаусијана
- Груписати гаусијане применом OWA оператора
- Применити алгоритам кластеровања над сваком од група појединачно

– Декодовање:

За сваку од опсервација и сваки кластер

- Одредити логаритме веродостојности између опсервација и хипер-гаусијана који одговарају кластерима
- Уколико је логаритам веродостојности посматраног хипер-гаусијана изнад предефинисаног прага, егзактно одредити вероватноће сваког од придружених гаусијана. У супротном, доделити вредност одређену за дати хипер-гаусијан.

Поглавље 4

Хијерархијско кластеровање модела Гаусових смеша

4.1. Формулација проблема

Посматрајмо оригинални (полазни) модел Гаусове смеше f са $k > 1$ d -димензионалних Гаусових компоненти:

$$f(y) = \sum_{i=1}^k \alpha_i N(y; \mu_i, \Sigma_i) = \sum_{i=1}^k \alpha_i f_i(y) \quad (4.1)$$

при чему α_i представља тежину i -те Гаусове компоненте смеше, μ_i је вектор средњих вредности, Σ_i коваријансна матрица, док $f_i = N(y; \mu_i, \Sigma_i)$ представља i -ту Гаусову компоненту.

Задатак кластеровања модела Гаусових смеша јесте кластеровање компонентних гаусијана оригиналне Гаусове смеше f , у циљу формирања нове, компактне, тј. поједностављене репрезентације g оригиналне Гаусове смеше f са $m \ll k$ компоненти, поштујући ограничење да ниједна оригинална компонента не може бити подељена у поједностављеној репрезентацији. Расподела g представља хипер-смешу, тј. хипер-GMM, чије су компоненте хипер-гаусијани. Практично, задатак кластеровања модела Гаусових смеша своди се на проналажење компактне репрезентације оригиналне Гаусове смеше g , најближе посматраном моделу f , имајући у виду одабрану меру блискости, при чему структура полазног модела остаје у потпуности очувана.

У оквиру наредног одељка описан је референтни алгоритам кластеровања модела Гаусових смеша, сличан k-means алгоритму. Сврха овако дефинисаног алгоритма је вишеструка. Поред примене у задатку хијерархијског кластеровања, алгоритам је могуће применити и у задатку

обуке система, груписања дискретних категорија, као и раније поменутом задатку препознавања говора (говорника, емоција). Посебно је важна примена овог алгоритма у сврху повећања ефикасности система, тј. смањења рачунске комплексности, у оквиру апликација које се извршавају у реалном времену. Алгоритам представљен у наредном одељку, тестиран је у задатку селекције гаусијана у оквиру реалног CSR система, као и над вештачки генерисаним подацима.

4.2. Опис референтног алгоритма

У раду [7] (Goldberger, Roweis [2005]) аутори предлажу проширени приступ хијерархијском кластеровању модела Гаусових смеша, при чему структура оригиналног модела остаје у потпуности очувана, у складу с постављеним задатком. KL дивергенца намеће се као природна мера блискости између два GMM модела. С обзиром да такав израз не постоји у затвореној форми, те да тако одабран критеријум води ка нерешивом оптимизационом проблему, аутори уводе нову, аналитички решиву меру растојања између модела f и g :

$$d(f, g) = \sum_{i=1}^k \alpha_i \min_{j=1}^m KL(f_i || g_j) \quad (4.2)$$

где су f и g представљени у следећој форми:

$$f = \sum_{i=1}^k \alpha_i f_i \quad (4.3)$$

$$g = \sum_{j=1}^m \beta_j g_j$$

и важи да је $m \ll k$.

Означимо скуп свих d -димензионалних модела Гаусових смеша са највише m компонената са $\text{MoG}(m)$ и скуп свих партиција са m елемената скупа $\Omega_k = \{1, \dots, k\}$ са \mathcal{P}_m , при чему Ω_k представља скуп који индексира све Гаусове компоненте оригиналне Гаусове смеше f . То значи да се свака партиција $\pi \in \mathcal{P}_m$ може представити као $\pi = (\pi_1, \dots, \pi_m)$, при чему важи: $\Omega_k = \bigcup_{j=1}^m \pi_j$, $\pi_j \neq \emptyset$, $i \neq j \Rightarrow \pi_i \cap \pi_j = \emptyset$ за свако $i, j \in \{1, \dots, m\}$. Свака партиција $\pi \in \mathcal{P}_m$ одговара једном кластеровању на m кластера скупа Гаусових компоненти $\{f_i\}_{i=1}^k$, где су $\pi_j, j \in \{1, \dots, m\}$ кластери Гаусових компоненти, а π је одговарајућа партиција. Нека је $\pi \in \mathcal{P}_m$ фиксно. Тада за свако $i \in \Omega_k$ кажемо да је $\pi(i) = \pi_j$ ако и само ако важи да је $i \in \pi_j$. π_i представља класу еквиваленције којој $i \in \Omega_k$ припада, за јединствену релацију еквиваленције коју генерише партиција $\pi \in \mathcal{P}_m$. Уколико дефинишемо помоћну функцију $d(f, g, \pi)$ на следећи начин

$$d(f, g, \pi) = \sum_{i=1}^k \alpha_i \text{KL}(f_i \parallel g_{\pi(i)}) \quad (4.4)$$

за ма које $\pi \in \mathcal{P}_m$, за дато $g \in \text{MoG}(m)$, партиција $\pi^g \in \mathcal{P}_m$ дата је као

$$\pi^g(i) = \arg \min_{j=1}^m \text{KL}(f_i \parallel g_j), \quad i = 1, \dots, k \quad (4.5)$$

Може се показати да важи

$$d(f, g) = d(f, g, \pi^g) = \min_{\pi \in \mathcal{P}_m} d(f, g, \pi) \quad (4.6)$$

тј. за дато g, π^g је оптимална партиција скупа Ω_k на скуп од m елемената, компонентних гаусијана посматраног модела g . На тај начин, долазимо до проблема двоструке минимизације.

$$(\tilde{g}, \tilde{\pi}) = \arg \min_{g \in \text{MoG}(m)} \min_{\pi \in \mathcal{P}_m} d(f, g, \pi) \quad (4.7)$$

За дату партицију $\pi \in \mathcal{P}_m$, функција $g^\pi \in \text{MoG}(m)$ дефинише се као Гаусова смеша у којој је Гаусова компонента g_j^π добијена обједињавањем расподеле

$$f_j^\pi = \frac{\sum_{i \in \pi_j} \alpha_i f_i}{\sum_{i \in \pi_j} \alpha_i} \quad (4.8)$$

у јединствени гаусијан. Средња вредност и варијанса од f_j^π израчунавају се као:

$$\begin{aligned} \tilde{\mu}_j &= \frac{1}{\beta_j} \sum_{i \in \pi_j} \alpha_i \mu_i \\ \tilde{\Sigma}_j &= \frac{1}{\beta_j} \sum_{i \in \pi_j} \alpha_i \left(\Sigma_i + (\mu_i - \tilde{\mu}_j)(\mu_i - \tilde{\mu}_j)^T \right) \\ \beta_j &= \sum_{i \in \pi_j} \alpha_i \end{aligned} \quad (4.9)$$

где су $f_i = N(\mu_i, \Sigma_i)$ полазне Гаусове компоненте. Може се показати да за гаусијан $g_j^\pi = N(\tilde{\mu}_j, \tilde{\Sigma}_j)$ тежине β_j важи следећи израз:

$$g_j^\pi = \arg \min_{g \in \text{MoG}(1)} \text{KL}(f_j^\pi \| g) = \arg \min_{g \in \text{MoG}(1)} d(f_j^\pi \| g) \quad (4.10)$$

при чему се минимизација изводи над коначним скупом d -димензионалних Гаусових расподела, тј. g_j^π се добија обједињавањем f_j^π у јединствени гаусијан. Означимо једну такву представу f на основу $\pi \in \mathcal{P}_m$ са g^π , тј.

$$g^\pi = \sum_{j=1}^m \beta_j g_j^\pi \quad (4.11)$$

Може се показати (погледати [7], Лема 1) да за дату смешу гаусијана (MoG) f и за дату партицију π , g^π представља јединствену тачку минимума функције $d(f, g, \pi)$. С обзиром да проблем двоструке минимизације $(\tilde{g}, \tilde{\pi}) = \arg \min_{g \in \text{MoG}(m)} \min_{\pi \in \mathcal{P}_m} d(f, g, \pi)$ није аналитички решив за произвољно $(g, \pi) \in \text{MoG}(m) \times \mathcal{P}_m$, предложена је процедура наизменичне минимизације ([7]) у циљу достизања локалног минимума функције $d(f, g)$. Дефинише се као наизменична и итеративна примена операција прерасподеле и реестимације (*Regroup* i *Refit*).

$$\pi^g = \operatorname{argmin}_{\pi} d(f, g, \pi) \quad \text{Regroup} \quad (4.12)$$

$$g^\pi = \operatorname{argmin}_g d(f, g, \pi) \quad \text{Refit} \quad (4.13)$$

Извођење ових операција монотono снижава вредност функције $d(f, g)$ дефинисане формулом (2) (видети [7]). С обзиром да је \mathcal{P}_m коначно, ова функција конвергира само ка локалном минимуму \hat{g} функције $d(f, g)$ над $\text{MoG}(m)$.

Поглавље 5

Нови алгоритам поделе и спајања за хијерархијско кластеровање модела Гаусових смеша

Студија описана у раду [24] представља мотивацију за примену принципа сукцесивне поделе и спајања у циљу превазилажења постојећег проблема локалног оптимума EM алгоритма, примењеног у функцији одређивања параметара GMM модела на основу расположивих опсервација. Као критеријум поделе, одабрана је локална KL дивергенца, представљена изразом

$$J_{split}(i; \psi) = \int f_i(x; \psi) \log \frac{f_i(x; \psi)}{p(x; \psi_i)} dx \quad (5.1)$$

Она представља растојање између две расподеле, локалне густине података обухваћених i -том Гаусовом компонентом и i -те Гаусове расподеле одређене естимацијом параметара у датом моменту [31]. Гаусова компонента са највећим растојањем $J_{split}(i; \psi)$ подразумева најмање прецизну естимацију, стога је треба поделити на две нове Гаусове компоненте, чији се параметри уобичајено одређују применом EM алгоритма и ML критеријума, описаног у Поглављу 2. Када је у питању критеријум спајања, уобичајено се користе критеријуми засновани на прорачуну апостериорних вероватноћа појединачних гаусијана, максимални коефицијент корелације између две Гаусове компоненте [25] и други. Овај концепт успешно је примењен у раду [25].

У оквиру овог поглавља, предложен је нови алгоритам поделе и спајања (S&M) за хијерархијско кластеровање модела Гаусових смеша (S&M HGMMC). Основна идеја овог алгоритма иста је као и за претходно описане S&M алгоритме, коришћене за одређивање параметара GMM модела [24][25][30], у смислу тежње да се локално оптимално решење

функције цене (4.2) превазиђе итеративном применом S&M операција. Ова идеја искоришћена је на следећи начин:

За дату полазну расподелу $f \in \text{MoG}(k)$, произвољно $g \in \text{MoG}(m)$ и $\pi \in \mathcal{P}_m$, за фиксно $m \in \mathbb{N}$, формула (4.3) приказује се у следећој еквивалентној форми:

$$d(f, g, \pi) = \sum_{j=1}^m A_j \tag{5.2}$$

$$A_j = \sum_{i \in \pi_j} \alpha_i KL(f_i || g_j)$$

Претпоставимо да је $(\hat{g}, \hat{\pi}) \in \text{MoG}(m) \times \mathcal{P}_m$ локални минимум функције $d(f, g, \pi)$ дате формулама (4.4) и (5.2) и нека је $m > 3$. Функција цене (5.2) за дато $(\hat{g}, \hat{\pi})$ може се записати на следећи начин

$$d(f, \hat{g}, \hat{\pi}) = \hat{A}_{\hat{k}_1} + \hat{A}_{\hat{k}_2} + \hat{A}_j + \hat{A} \tag{5.3}$$

$$\hat{A} = \sum_{\substack{j \in \{1, \dots, m\} \\ j \notin \{\hat{k}_1, \hat{k}_2, j\}}} \hat{A}_j$$

за фиксно $\hat{k}_1, \hat{k}_2, j \in \{1, \dots, m\}$, где је

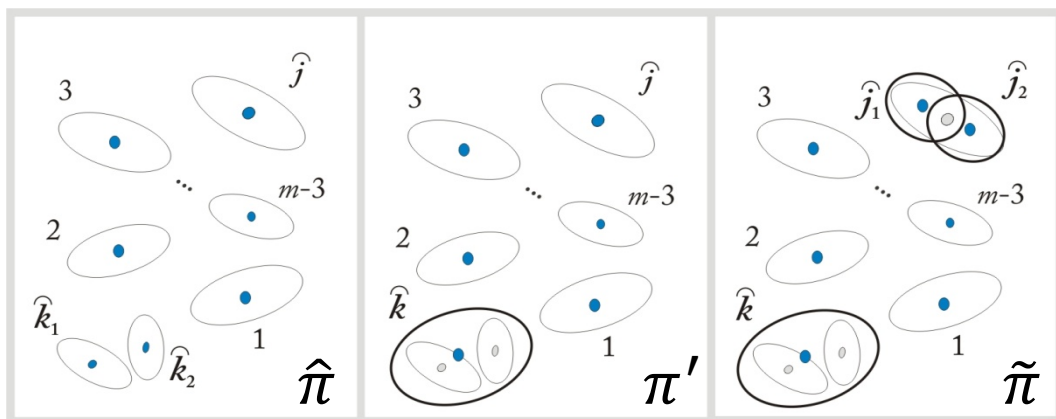
$$\hat{A}_j = \sum_{i \in \hat{\pi}_j} \alpha_i KL(f_i || \hat{g}_j) \tag{5.4}$$

Претпоставимо, без губитка општости, да су кластери $\hat{\pi}_{\hat{k}_1}, \hat{\pi}_{\hat{k}_2} \in \hat{\pi}$ одабрани за операцију спајања, а на основу претходно усвојеног критеријума спајања. Притом се формира нови кластер $\hat{\pi}_{\hat{k}} = \hat{\pi}_{\hat{k}_1} \cup \hat{\pi}_{\hat{k}_2}$ и њему придружена партиција $\pi' = \{\hat{\pi}_{\hat{k}}\} \cup \{\hat{\pi}_j | j \in \{1, \dots, m\} \setminus \{\hat{k}_1, \hat{k}_2\}\}$. Гаусова компонента $\hat{g}_{\hat{k}}$ и тежина $\hat{\beta}_{\hat{k}}$ које одговарају новоформираном

кластеру $\hat{\pi}_{\hat{k}}$ процењују се применом формула (4.9). Приметимо да је партиција $\pi' \in P_{m-1}$, као и да се ова партиција добија на основу партиције $\hat{\pi}$, једноставним повезивањем кластера $\hat{\pi}_{\hat{k}_1}$ и $\hat{\pi}_{\hat{k}_2}$ у јединствени кластер. Приметимо и то да се модел $g' \in MoG(m-1)$, који одговара партицији π' , може приказати у виду компонентних гаусијана: $\hat{g}_j, j \in \{1, \dots, m\} \setminus \{\hat{k}_1, \hat{k}_2\}$ и $\hat{g}_{\hat{k}}$ и њима одговарајућих тежина. Додатно, како за π' , тако и за компонентне гаусијане модела g' , врши се реиндексирање коришћењем индексног скупа $\{1, \dots, m-1\}$.

Претпоставимо даље, без губитка општости, и да је кластер $\hat{\pi}_j \in \hat{\pi}$ одабран за операцију поделе, а на основу претходно усвојеног критеријума поделе. Применом поменутог критеријума, формирају се два нова кластера $\hat{\pi}_{j_1}$ и $\hat{\pi}_{j_2}$, при чему је $\hat{\pi}_j = \hat{\pi}_{j_1} \cup \hat{\pi}_{j_2}$, док се процене параметара одговарајућих Гаусових компоненти $\hat{g}_{j_1}, \hat{g}_{j_2}$ и њихових тежина $\hat{\beta}_{j_1}, \hat{\beta}_{j_2}$ добијају применом формула (4.9). На тај начин, добија се нова партиција $\tilde{\pi} \in P_m$, дефинисана као $\tilde{\pi} = (\pi' \setminus \{\hat{\pi}_j\}) \cup \{\hat{\pi}_{j_1}, \hat{\pi}_{j_2}\}$, уз одговарајући хипер-GMM $\tilde{g} \in MoG(m)$, који садржи $m-3$ компоненте \hat{g}_j за $j \notin \{\hat{k}_1, \hat{k}_2, j\}$, као и три додатне Гаусове компоненте $\hat{g}_{\hat{k}}, \hat{g}_{j_1}$ и \hat{g}_{j_2} , одређене обједињавањем Гаусових смеша које одговарају кластерима $\hat{\pi}_{\hat{k}}, \hat{\pi}_{j_1}$ и $\hat{\pi}_{j_2}$ (погледати формуле (4.2) и (4.3)), заједно за одговарајућим тежинама. Након тога, врши се поновно реиндексирање компоненти за $\tilde{\pi}$ и \tilde{g} коришћењем индексног скупа $\{1, \dots, m\}$.

Процедура поделе и спајања илустрована је на слици 5.1, док су предлози за избор критеријума поделе и спајања детаљно описани у оквиру одељака 5.1 и 5.2.



СЛИКА 5.1 ПРОЦЕДУРА ПОДЕЛЕ И СПАЈАЊА GMM КЛАСТЕРА

Дефинишимо сада

$$\tilde{A}_j = \sum_{i \in \tilde{\pi}_j} \alpha_i KL(f_i || \tilde{g}_j) \quad (5.5)$$

$$\tilde{A} = \sum_{\substack{j \in \{1, \dots, m\} \\ j \notin \{\hat{k}, j_1, j_2\}}} \tilde{A}_j$$

и посматрајмо услов

$$\hat{A}_{\hat{k}_1} + \hat{A}_{\hat{k}_2} + \hat{A}_j \geq \tilde{A}_{\hat{k}} + \tilde{A}_{j_1} + \tilde{A}_{j_2} \quad (5.6)$$

Уколико је поменути услов задовољен, имајући у виду релацију $\tilde{A} = \hat{A}$, важи да је $d(f, \tilde{g}, \tilde{\pi}) \leq d(f, \hat{g}, \hat{\pi})$. С обзиром на (4.6) као и на чињеницу да је $\hat{\pi} = \pi^{\hat{g}}$ и $\tilde{\pi} = \pi^{\tilde{g}}$, то коначно имплицира да је $d(f, \tilde{g}) \leq d(f, \hat{g})$.

Уколико критеријуме поделе и спајања, као и адекватан метод поделе кластера $\hat{\pi}_j$ на кластере $\hat{\pi}_{j_1}$ и $\hat{\pi}_{j_2}$ одаберемо тако да релација (5.6) буде задовољена у задовољавајуће великом проценту случајева, увођење S&M операција у алгоритам поновног груписања и придруживања (погледати изразе (4.12) и (4.13)) имало би за последицу додатно смањење вредности локалног минимума функције $d(f, g)$ над $\text{MoG}(m)$. Имајући то у виду, може се рећи да је правилан одабир критеријума поделе и спајања, као и метода за увођење операције поделе, од суштинског значаја за ефикасно укључивање S&M операција у алгоритам итеративног груписања и придруживања, а све у циљу смањења вредности функције цене (4.2). Одабир критеријума поделе и спајања врши се на природан начин, коришћењем KL дивергенце као добро познатог информационо псеудо-растојања. Јединствено решење понуђено је и за прорачун параметара $\hat{g}_{\hat{k}}$, што операцију спајања чини изводљивом.

Полазећи од чињенице да не постоји израз за KL дивергенцу између два произвољна GMM модела, уводе се различите апроксимације, како би усвојени критеријум поделе постао функционалан. Штавише, с обзиром да је поменута операција поделе лоше постављена операција, у овом раду предложено је једно једноставно и ефикасно решење тог проблема.

Потребно је нагласити да се разматрања која се тичу образаца (5.3) и (5.4) односе се на случај када се операција спајања изводи пре операције поделе. Аналогни закључци могу се извести и у случају када је редослед поменутих операција супротан. Ова чињеница је искоришћена у оквиру Поглавља 6, унутар ког је метод селекције модела имплементиран у оквиру предложеног S&M алгорита кластеровања модела Гаусових смеша, у циљу остваривања бољег односа између рачунске сложености с једне, и тачности алгорита препознавања с друге стране, у задатку селекције гаусијана унутар реалног CSR система. Резултати експеримената у оквиру реалног CSR система изложени су у оквиру одељка 9.2.

5.1. Операција спајања

За дату партицију $\hat{\pi} \in P_m$ и њој одговарајући хипер-гаусијан $\hat{g} \in MoG(m)$, бира се информациони критеријум заснован на KL растојању. Овај критеријум предложен је као потенцијално најефикаснији, с обзиром да његова примена доводи до задовољења услова израженог формулом (5.6) у највећем проценту случајева. Два произвољна кластера $\hat{\pi}_{\hat{k}_1}, \hat{\pi}_{\hat{k}_2} \in \hat{\pi}$, заједно са њима придруженим Гаусовим компонентама, биће одабрани за процедуру спајања уколико је следећи услов задовољен

$$(\hat{k}_1, \hat{k}_2) \in \arg \min_{(k_1, k_2) \in (\{1, \dots, m\})^2} KL(\hat{g}_{k_1} \parallel \hat{g}_{k_2}) \quad (5.7)$$

Израз за $KL(h_1 \parallel h_2)$ за ма која два d -димензионална гаусијана h_1 и h_2 може се дефинисати у затвореној форми

$$KL(h_1 \parallel h_2) = \frac{1}{2} \left(\log \frac{|\Sigma_2|}{|\Sigma_1|} + Tr(\Sigma_2^{-1} \Sigma_1) + (\mu_1 - \mu_2)^T \Sigma_2^{-1} (\mu_1 - \mu_2) - d \right) \quad (5.8)$$

при чему су (μ_1, Σ_1) и (μ_2, Σ_2) параметри гаусијана h_1 и h_2 респективно. Уколико решење није јединствено, произвољно се бира једно од њих. Након тога, изводи се операција спајања која је добро постављена (*well-posed*) операција: одређују се кластер $\hat{\pi}_{\hat{k}} = \hat{\pi}_{\hat{k}_1} \cup \hat{\pi}_{\hat{k}_2}$ и њему одговарајући хипер-

гаусијан $\hat{g}_{\hat{k}}$ (погледати Поглавље 4), користећи процене изражене формулом (4.9) за новодобијену смешу:

$$f_{\hat{k}}^{\pi} = \frac{\sum_{i \in \hat{\pi}_{\hat{k}_1}} \alpha_i f_i + \sum_{i \in \hat{\pi}_{\hat{k}_2}} \alpha_i f_i}{\sum_{i \in \hat{\pi}_{\hat{k}_1} \cup \hat{\pi}_{\hat{k}_2}} \alpha_i} \quad (5.9)$$

Потребно је још напоменути да S&M GMM обука ([24][25]) експлицитно садржи EM корак оптимизације за GMM модел који се састоји од само једне Гаусове компоненте која одговара новодобијеном кластеру (погледати [24]). Када се операција спајања изводи у оквиру предложеног концепта S&M операција у задатку кластеровања модела Гаусових смеша, стварна оптимизација врши се имплицитно, с обзиром да хипер-гаусијан $\hat{g}_{\hat{k}}$ који одговара кластеру $\hat{\pi}_{\hat{k}}$ производи минималну KL дивергенцу за дати GMM модел (5.8) пошто је одређен применом процена представљених формулом (4.9).

5.2. Операција поделе

Из истих разлога као и у случају операције спајања, KL дивергенцу користимо при дефинисању критеријума поделе. Нека је $\hat{\pi} \in P_m$ партиција добијена у оквиру актуелног корака итерације поделе и спајања, и нека је $\hat{g} \in \text{MoG}(m)$ одговарајућа хипер-смеша. За операцију поделе бира се кластер $\hat{j} \in \hat{\pi}$ који задовољава следећи услов

$$\hat{j} \in \arg \max_{j \in \{1, \dots, m\} \setminus \{\hat{k}_1, \hat{k}_2\}} \text{KL}(f_j^{\hat{\pi}} \parallel \hat{g}_j) \quad (5.10)$$

при чему је $f_j^{\hat{\pi}}$ Гаусова меша која одговара кластеру $\hat{\pi}_j$, \hat{g}_j је компонента од \hat{g} која одговара j -ом кластеру, а кластери \hat{k}_1 и \hat{k}_2 су изабрани за операцију спајања која се изводи непосредно пре операције поделе. Уколико решење није јединствено, као у и случају операције спајања, једно од понуђених решења бира се произвољно.

Познато је (погледати нпр. [26]), да за разлику од KL дивергенце између два гаусијана за коју се дефинише израз у затвореној форми на начин

представљен изразом (5.8), не постоји рачунски изводљив израз за KL дивергенцу између две произвољне Гаусове смеше. У том смислу, како би критеријум одређен формулом (5.10) постао рачунски изводљив, користе се две различите апроксимације за $KL(f_j^{\hat{\pi}} \parallel \hat{g}_j)$, обе предложене у раду [26]. Поменуће апроксимације укратко су описане у наставку. Прва је MCSA апроксимација, представљена у форми:

$$D_{MCSA}(f_j^{\hat{\pi}} \parallel \hat{g}_j) = \frac{1}{n} \sum_{i=1}^n \ln \frac{f_j^{\hat{\pi}}(x_i)}{\hat{g}_j(x_i)} \quad (5.11)$$

где су $\{x_i\}_{i=1}^n$ независни, подједнако расподељени, случајни одбирци расподеле $f_j^{\hat{\pi}}$. Ови одбирци независно се генеришу за сваку од појединачних Гаусових компоненти f_i , $i \in \hat{\pi}_j$. Број опсервација n_i који се генерише за неко одређено f_i одређује се у складу с изразом

$$n_i = \left\lfloor \frac{\alpha_i}{\sum_{i \in \pi_j} \alpha_i} n \right\rfloor \quad (5.12)$$

при чему је

$$n = \sum_{i \in \pi_j} n_i \quad (5.13)$$

Како варијанса естимације тежи нули када $n \rightarrow \infty$ (погледати [26]), за велики број генерисаних одбирака n може се добити произвољно добра процена стварне вредности функције $KL(f_j^{\hat{\pi}} \parallel \hat{g}_j)$. С друге стране, без обзира на чињеницу да се процес кластеровања одвија *off-line*, укупан број рачунских операција потребан за добијање описаних естимација могао би бити превелик за данашње апликације.

Друга апроксимација јесте LBVA апроксимација. Примењена је у оквиру докторске дисертације у циљу смањења рачунске сложености предложеног S&M HGMMC алгоритма. Представљена је следећим изразом:

$$D_{LBVA}(f_j^{\hat{\pi}} \parallel \hat{g}_j) = \sum_{i \in \hat{\pi}_j} \alpha_i \ln \frac{\sum_{k \in \hat{\pi}_j} \alpha_k e^{-\text{KL}(f_i \parallel f_k)}}{e^{-\text{KL}(f_i \parallel \hat{g}_j)}} \quad (5.14)$$

Ово је израз у затвореној форми, с обзиром да су $\text{KL}(f_i \parallel f_k)$ и $\text{KL}(f_i \parallel \hat{g}_j)$ KL дивергенце између одговарајућих Гаусових компоненти такође дате као изрази у затвореној форми (5.8). Ова апроксимација подразумева знатно мањи број рачунских операција у односу на MCSA апроксимацију, али је довољно прецизна, што је потврђено одговарајућим експериментима приказаним у оквиру Поглавља 9.

LBVA апроксимација добија се на следећи начин [26]: Полази се од израза за KL дивергенцу између функција густине вероватноћа f и g

$$D(f \parallel g) = \int f(x) \log \frac{f(x)}{g(x)} dx \quad (5.15)$$

при чему су f и g два GMM модела представљена изразима (4.3). Уколико веродостојност $L_f(g)$ дефинишемо као

$$L_f(g) = E_{f(x)} \log g(x) \quad (5.16)$$

где је $E_{f(x)}$ очекивање по $f(x)$, израз за KL дивергенцу можемо преписати у форми

$$D(f \parallel g) = L_f(f) - L_f(g) \quad (5.17)$$

Дефинишемо варијационе параметре $\phi_{j|i} > 0$ чија је сума по j једнака јединици $\sum_{j=1}^m \phi_{j|i} = 1$. Применом Јенсенове неједнакости добијамо

$$\begin{aligned} L_f(g) &= E_f \log g \\ &= E_f \log \sum_j \beta_j g_j \end{aligned} \quad (5.18)$$

$$\begin{aligned}
 &= E_f \log \sum_j \phi_{j|i} \frac{\beta_j g_j}{\phi_{j|i}} \\
 &\geq E_f \sum_j \phi_{j|i} \log \frac{\beta_j g_j}{\phi_{j|i}} \\
 &\stackrel{\text{def}}{=} \mathcal{L}_f(g, \phi)
 \end{aligned}$$

Максимизацијом горњег израза по ϕ добијамо

$$\hat{\phi}_{j|i} = \frac{\beta_j e^{-\text{KL}(f_i || g_j)}}{\sum_l \beta_l e^{-\text{KL}(f_i || g_l)}} \quad (5.19)$$

Аналогно томе можемо дефинисати $\mathcal{L}_f(f, \psi)$

$$\mathcal{L}_f(f, \psi) \stackrel{\text{def}}{=} E_f \sum_k \psi_{k|i} \log \frac{\alpha_k f_k}{\psi_{k|i}} \quad (5.20)$$

и одредити оптималну вредност коефицијената ψ

$$\hat{\psi}_{k|i} = \frac{\alpha_k e^{-\text{KL}(f_i || f_k)}}{\sum_l \alpha_l e^{-\text{KL}(f_i || f_l)}} \quad (5.21)$$

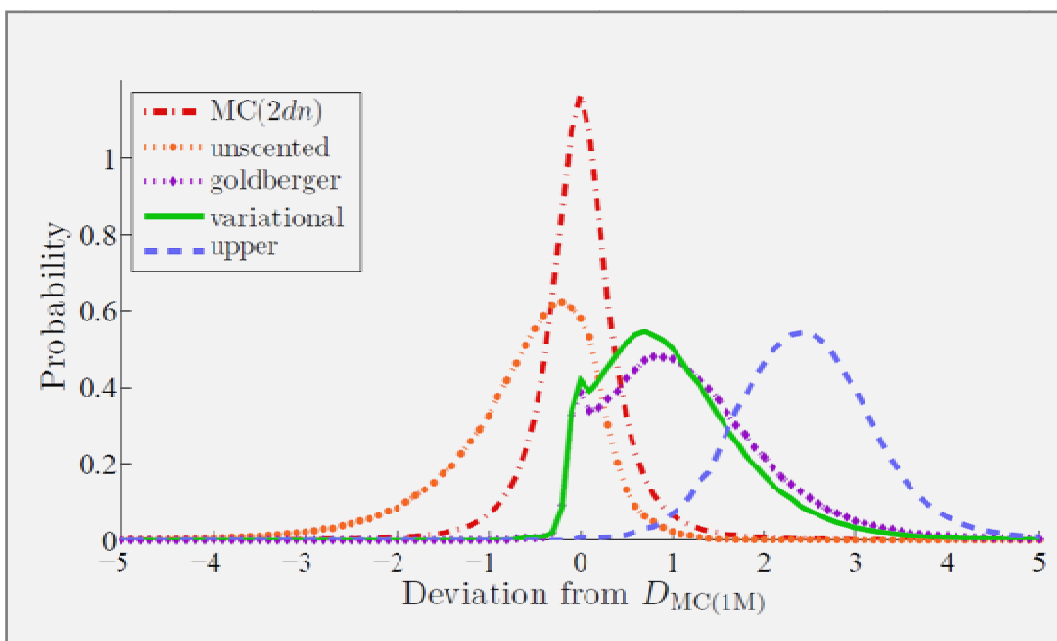
Уколико $D(f || g)$ апроксимирамо изразом

$$D_{LBVA} = \mathcal{L}_f(f, \hat{\psi}) - \mathcal{L}_f(g, \hat{\phi}) \quad (5.22)$$

заменом оптималних вредности коначно добијамо

$$D_{LBVA}(f || g) = \sum_i \alpha_i \log \frac{\sum_k \alpha_k e^{-\text{KL}(f_i || f_k)}}{\sum_j \beta_j e^{-\text{KL}(f_i || g_j)}} \quad (5.23)$$

На Слици 5.2 [26] приказано је одступање резултата остварених применом LBVA апроксимације („*variational*“) од резултата остварених применом MCSA апроксимације за милион одбирака. Међу рачунски једноставним апроксимацијама чији су изрази дати у затвореној форми, LBVA апроксимација показује се као најефикаснија.



СЛИКА 5.2 ОДСТУПАЊЕ У ОДНОСУ НА MCSA АПРОКСИМАЦИЈУ СА МИЛИОН ОДБИРАКА

Претпоставимо да је кластер $\hat{\pi}_j$ одабран за операцију поделе, чиме се формирају два нова кластера $\hat{\pi}_{j_1}$ и $\hat{\pi}_{j_2}$. Параметри одговарајућих хипергаусијана $\hat{g}_{j_1}^{(0)}$, $\hat{g}_{j_2}^{(0)}$ и њихове тежине иницијализују се на следећи начин:

$$\hat{\beta}_{j_1}^{(0)} = \hat{\beta}_{j_2}^{(0)} = \frac{\hat{\beta}_j}{2} \quad (5.24)$$

$$\hat{\Sigma}_{j_1}^{(0)} = \hat{\Sigma}_{j_2}^{(0)} = \frac{\hat{\Sigma}_j}{2} \quad (5.25)$$

$$\hat{\mu}_{j_1}^{(0)} = \hat{\mu}_j + \frac{\sqrt{\lambda_{\max}}}{2} v_{\max} \quad (5.26)$$

$$\hat{\mu}_{j_2}^{(0)} = \hat{\mu}_j - \frac{\sqrt{\lambda_{\max}}}{2} v_{\max} \quad (5.27)$$

где су $(\hat{\mu}_j, \hat{\Sigma}_j)$ параметри гаусијана \hat{g}_j који одговара кластеру $\hat{\pi}_j$, док λ_{\max} и ν_{\max} представљају максималну карактеристичну вредност и њој одговарајући карактеристични вектор, израчунат за $\hat{\Sigma}_j$. За тако добијене $\hat{g}_{j_1}^{(0)}, \hat{g}_{j_2}^{(0)}$ изводи се операција прерасподеле дата у (4.12) како би се добили кластери $\hat{\pi}_{j_1}^{(0)}, \hat{\pi}_{j_2}^{(0)}$. Даље се над паром кластера $(\hat{\pi}_{j_1}^{(0)}, \hat{\pi}_{j_2}^{(0)})$ извршава алгоритам прерасподеле и реестимације описан формулама (4.12) и (4.13), у оквиру Поглавља 4, све до конвергенције у неких S корака, на тај начин добијајући $(\hat{\pi}_{j_1}, \hat{\pi}_{j_2}) = (\hat{\pi}_{j_1}^{(S)}, \hat{\pi}_{j_2}^{(S)})$, уз одговарајуће хипер-гаусијане $(\hat{g}_{j_1}, \hat{g}_{j_2}) = (\hat{g}_{j_1}^{(S)}, \hat{g}_{j_2}^{(S)})$ и тежине $(\hat{\beta}_{j_1}, \hat{\beta}_{j_2}) = (\hat{\beta}_{j_1}^{(S)}, \hat{\beta}_{j_2}^{(S)})$. Ово је заправо корак оптимизације. С обзиром да је операција поделе лоше постављена операција, предложено је решење у оквиру кога се параметри иницијализују на једноставан начин. Имајућу и виду да иницијализација може утицати чак и на случај са два кластера, а посебно велики утицај има у случајевима када се ради о високо димензионалним просторима (као што је то случај у CSR задатку), постоји простор за побољшање перформанси предложеног алгоритма, коришћењем повољнијег начина иницијализације, што може бити тема будућег рада.

У оквиру последњег корака сваке итерације S&M алгоритма, тј. након корака поделе и спајања, изводи се референтни алгоритам, иницијализован излазним резултатом претходних корака $(\tilde{\pi}, \tilde{g})$ све до конвергенције и добијања (π^*, g^*) . Уколико важи да је $d(f, g^*, \pi^*) < d(f, \hat{g}, \hat{\pi})$, при чему је $(\hat{\pi}, \hat{g})$ резултат претходне итерације S&M алгоритма, узима се да је $(\hat{\pi}, \hat{g}) \leftarrow (\pi^*, g^*)$ и алгоритам се наставља следећом итерацијом. Уколико то није испуњено, процес је завршен, а $(\hat{\pi}, \hat{g})$ се усваја као коначно решење.

Поглавље 6

Метод селекције модела у предложеном алгоритму поделе и спајања

Селекција модела један је од најзначајнијих процеса у оквиру сваког задатка кластеровања. Суштински, своди се на одређивање оптималног броја кластера m^* за дати GMM модел. Приликом одређивања параметара GMM модела, примењују се разне технике селекције модела, као што су нпр. Акаике информациони критеријум (AIC), Бајесов информациони критеријум (БИЦ), или још ефикасније крос-валидационе технике [32]. Док AIC и BIC за већину апликација нису довољно поуздане, крос-валидација је рачунски изузетно захтевна. Па ипак, сва три информациона критеријума теже да се што више приближе оптималном или субоптималном решењу за величину GMM модела који описује дату полазну расподелу, на начин да одређено информационо растојање између полазне расподеле и одговарајућег GMM модела буде што је могуће мање (тачније, расподела није егзактно дата, доступан је само скуп опсервација за дату расподелу). У случају HGMMС алгоритма, уколико претпоставимо да је полазна расподела једнака f , и уколико тражимо минималну KL дивергенцу и занемаримо чињеницу да поменути алгоритам поновног груписања и придруживања даје само субоптимално решење, јасно је да је $m^* = k$ и $g^* = f$, што само по себи намеће закључак да би коришћење таквог информационог критеријума било бесмислено. Циљ је да се селекција модела, тј. процес одређивања таквог броја кластера m којим се остварује најбољи однос између тачности и рачунске комплексности, укључи у предложени алгоритам у задатку селекције гаусијана у оквиру реалног система за препознавање.

Гаусове компоненте $\{f_i\}_{i=1}^k$ сматрају се објектима у алгоритму сличном k-means-у, где се KL дивергенца користи као псеудо-информационо растојање, као што је то дефинисано у оквиру Поглавља 4. У даљем тексту

описана је једна од најједноставнијих техника за селекцију модела која успешно може да се примени на било које кластеровање засновано на k-means алгоритму (погледати [17] и [32]). Ова техника ефикасна је у великој већини случајева.

Полази се од мањег броја кластера (на пример $m = 2$), затим се изводи алгоритам заснован на k-means алгоритму до конвергенције, и наставља са извођењем истог тог алгоритма са повећаним бројем кластера. Вредност функције цене опада релативно брзо, све док се не приближимо „оптималном“ броју кластера, након чега се брзина којом се смањује вредност функције цене значајно смањује, тако да се „оптималан“ број кластера може одабрати као број који одговара колону претходно описане карактеристике (функције цене). Проблем је у чињеници да се процес кластеровања увек изводи изнова, за сваки предефинисани број кластера, што резултује комбинаторном експлозијом и чини описани метод неподобним у практичном смислу.

У овом раду предложена је верзија S&M алгоритма за кластеровање модела Гаусових смеша, која користи операције поделе и спајања описане у оквиру Поглавља 5, при чему се S&M итерације изводе над растућим бројем кластера. На тај начин, избор модела врши се у само једном пролазу алгоритма (“on the fly”).

Предложени метод отпочиње референтним алгоритмом описаним у Поглављу 4, при чему је иницијални број кластера постављен на $m^{(0)} = 2$. Затим се извршава алгоритам прерасподеле и реестимације описан формулама (4.12) и (4.13), до конвергенције, при чему се добија констелација $(\hat{\pi}_0^{(0)}, \hat{g}_0^{(0)}) \in \mathbb{P}_2 \times MoG(2)$. Након тога, врше се операције поделе и спајања, описане у одељцима 5.1 и 5.2, почевши од $(\hat{\pi}_0^{(0)}, \hat{g}_0^{(0)})$. У k-том S&M кораку добијамо $(\hat{\pi}_{s,k}^{(0)}, \hat{g}_{s,k}^{(0)})$ и $(\hat{\pi}_{m,k}^{(0)}, \hat{g}_{m,k}^{(0)})$ респективно, при чему су $(\hat{\pi}_{k-1}^{(0)}, \hat{g}_{k-1}^{(0)})$ партиција и хипер-гаусијан добијени по извршењу претходног, K-1-вог покушаја. Уколико важи да је $d(f, \hat{g}_{m,k}^{(0)}, \hat{\pi}_{m,k}^{(0)}) < d(f, \hat{g}_{k-1}^{(0)}, \hat{\pi}_{k-1}^{(0)})$, усваја се $(\hat{\pi}_{m,k}^{(0)}, \hat{g}_{m,k}^{(0)})$ као тренутно стање, и наставља са извршавањем операција поделе и спајања. С друге стране, уколико је $d(f, \hat{g}_{m,k}^{(0)}, \hat{\pi}_{m,k}^{(0)}) \geq d(f, \hat{g}_{k-1}^{(0)}, \hat{\pi}_{k-1}^{(0)})$, узима се да је $m^{(1)} = 3$, а такође и $(\hat{\pi}_0^{(1)}, \hat{g}_0^{(1)}) \leftarrow (\hat{\pi}_{s,k}^{(0)}, \hat{g}_{s,k}^{(0)})$, $(\hat{\pi}_0^{(1)}, \hat{g}_0^{(1)}) \in \mathbb{P}_3 \times MoG(3)$. Број кластера се надаље повећава као $m^{(p)} \leftarrow m^{(p-1)} + 1$ коришћењем горе описане процедуре, све док у некој K-тој итерацији не

добијемо $d(f, \hat{g}_0^{(K)}, \hat{\pi}_0^{(K)}) \geq d(f, \hat{g}_0^{(K-1)}, \hat{\pi}_0^{(K-1)}) - T$, где је $T > 0$ предефинисани праг. „Оптимални“ број кластера усваја се као $m^* = m^{(K)}$ и добија се коначно стање $(\pi^*, g^*) = (\hat{\pi}_0^{(K)}, \hat{g}_0^{(K)}) \in P_{m^*} \times MoG(m^*)$. То значи да ће се број кластера повећавати све док смањење функције цене не постане довољно безначајно, тј. испод предефинисаног прага T . Праг T контролише однос тачности и рачунске комплексности унутар стварног GS система који користи предложени алгоритам, у смислу да ниже вредности за T производе већи број кластера и самим тим већу тачност, али истовремено подразумевају и већи број рачунских операција.

Избором одговарајућег прага T остварује се повољнији однос тачности и рачунске комплексности у оквиру GS задатка на реалном CSR систему, за различито $\theta[\%]$ (удео хипер-гаусијана у процентима за које се придружене Гаусове компоненте директно евалуирају у GS процесу [9]), што се може закључити на основу експерименталних резултата приказаних у Табели 9.4, у оквиру Поглавља 9. Штавише, извођење S&M операција над растућим бројем кластера, као што је предложено у претходном разматрању, показује боље резултате у смислу ниже просечне вредности функције цене (4.2), ниже просечне варијансе функције цене, као и у смислу повећања броја случајева у којима долази до побољшања у односу на референтни алгоритам у случају вештачки генерисаних података, као што је приказано у Табелама 9.1, 9.2 и 9.3 респективно.

Поглавље 7

Додатно побољшање перформанси алгоритма селекције гаусијана

Алгоритам хијерархијског кластеровања модела Гаусових смеша заснован на операцији поделе и спајања, представљен у оквиру претходна два поглавља, постиже знатно боље резултате у поређењу с референтним алгоритмом у смислу односа тачности и рачунске комплексности, у задатку селекције гаусијана у оквиру реалног CSR система. Унутар овог поглавља представљено је неколико различитих приступа којима се остварују додатна побољшања перформанси алгоритма селекције гаусијана. Додатно побољшање у поређењу с полазним системом остварено је оптималним избором скупа системских параметара, односно, одговарајућом комбинацијом предефинисаног броја кластера, типа растојања коришћеног у процесу декодовања, броја хипер-гаусијана чији се придружени гаусијани евалуирају директно и коначно, избором одговарајућег типа кластеровања у смислу степена преклапања између појединачних кластера.

Полазни концепт селекције гаусијана подразумева кластеровање у коме нема преклапања новодобијених хипер-гаусијана. Овај принцип има за последицу значајно смањење рачунске сложености, али истовремено захтева директан прорачун за велики број полазних гаусијана да би се одржао висок ниво тачности. Стога је извршено поређење случајева с различитим степеном преклапања између новодобијених хипер-гаусијана. У оквиру прве етапе, тестирана је процедура кластеровања с великим степеном преклапања између новодобијених кластера. То је учињено полазећи од основног система, при чему се сваки гаусијан придружује сваком хипер-гаусијану, уколико је растојање између гаусијана и хипер-гаусијана испод одговарајућег прага, који се одређује засебно за сваки од хипер-гаусијана. Вредност прага поставља се се емпиријски, јер се на тај начин контролише степен преклапања између новодобијених кластера. Сличан приступ описан је у радовима [9] и [33], користећи дијагоналне

матрице. Остварени резултати говоре у прилог процедуре кластеровања у којој има преклапања између новодобијених хипер-гаусијана, али је преклапање мало. Повећање броја директно евалуираних кластера подразумева повећану тачност система препознавања, али истовремено, и знатно повећан број рачунских операција.

Уколико говоримо о броју кластера, може се позати да већи предефинисани број кластера за дати број гаусијана и хипер-гаусијана који се директно евалуирају, показује боље резултате у односу на случајеве с мањим предефинисаним бројем кластера, при чему не долази до значајније деградације перформанси система за континуално препознавање говора.

Параметри хипер-гаусијана одређују се као ML процене, на начин описан у оквиру претходних поглавља. На тај начин, остварује се минимална KL дивергенца између сваког хипер-гаусијана и њему придружених Гаусових компоненти полазне смеше. У циљу евалуације хипер-гаусијана у фази декодовања, користи се логаритамска веродостојност (*log-likelihood*), као што је описано у [12]. Процедура селекције гаусијана посматра се у термину односа између тачности у задатку препознавања и смањења потребног броја рачунских операција.

Упоредени су резултати остварени применом неколико најчешће коришћених мера блискости, у комбинацији са другим системским параметрима. У процесу декодовања тестирани су логаритамска веродостојност, еуклидско и Махаланобисово растојање. Показује се да примена Махаланобисовог растојања у фази декодовања даје боље резултате не само у односу на еуклидско растојање, већ и у односу на логаритамску веродостојност која се у пракси чешће користи. То је могуће објаснити чињеницом да величина гаусијана има велики утицај преко параметра $\ln(|\Sigma|)$, што није погодно за GS задатак.

ДЕФИНИЦИЈА 7.1

За дати улазни вектор обележја x и за дату вишедимензионалну нормалну (Гаусову) расподелу, логаритамска веродостојност дефинише се на следећи начин

$$D_{Log} = -\frac{d}{2} \log(2\pi) - \frac{1}{2} \log|\Sigma| - \frac{1}{2} (x_i - \mu)^T \Sigma^{-1} (x_i - \mu) \quad (7.1)$$

ДЕФИНИЦИЈА 7.2

За дати улазни вектор обележја x и за дату вишедимензионалну нормалну (Гаусову) расподелу чији је вектор средњих вредности μ , еуклидско растојање израчунава се применом обрасца

$$D_{EUC} = \sqrt{(x_i - \mu)^T (x_i - \mu)} \quad (7.2)$$

ДЕФИНИЦИЈА 2.4

За дати улазни вектор обележја x и за дату вишедимензионалну нормалну (Гаусову) расподелу чији је вектор средњих вредности μ и коваријансна матрица Σ , Махаланобисово растојање дефинише се као

$$D_{MAH} = \sqrt{(x_i - \mu)^T \Sigma^{-1} (x_i - \mu)} \quad (7.3)$$

Уколико је коваријансна матрица једнака јединичној матрици, Махаланобисово растојање своди се на претходно дефинисано еуклидско растојање.

Процена вредности вероватноћа појединачних гаусијана чији се хипер-гаусијани не одређују директно веома је важна и најчешће се спроводи тако што се вероватноће појединачних Гаусових компоненти апроксимирају вредношћу вероватноће за одговарајући хипер-гаусијан и дати улазни вектор обележја. И у случају процеса у коме постоји преклапање између новодобијених хипер-гаусијана, вредности вероватноћа компонентних гаусијана одређују се на потпуно исти начин, тј. ове вредности замењују се вредностима које одговарају хипер-гаусијану коме је дата Гаусова компонента придружена у почетној фази процеса у којој нема преклапања. Резултати експеримената детаљно су приказани у оквиру одељка 9.3.

Поглавље 8

Карактеристике система за континуално препознавање говора

8.1. Опис система за континуално препознавање говора

Систем за континуално препознавање говора коришћен у оквиру експеримената [34] заснован је на примени скривених Марковљевих модела (CSR HMM), при чему су емитујуће вероватноће стања представљене моделима Гаусових смеша. Свака од смеша дефинисана је вектором средњих вредности и пуном коваријансном матрицом. Сврха декодовања јесте проналажење највероватније секвенце речи која одговара улазном говору, при чему се користи Витербијев алгоритам. У циљу убрзања процедуре претраге, примењује се претходно описан принцип селекције гаусијана.

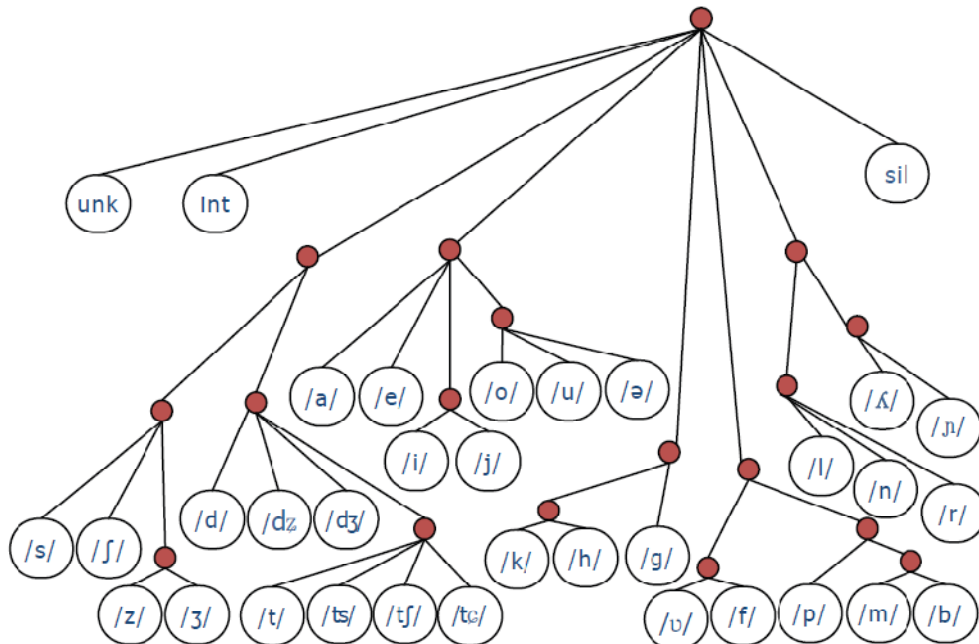
Систем користи трифоне као основне јединице моделовања, с обзиром да су акустичке особине фона под утицајем њему суседних фона (овај ефекат познат је под називом коартикулација). Фони су третирани као подјединице, при чему се тишина, као и разни типови импулсног шума, моделују као контексно независне јединице.

Укупан број HMM стања по једном моделу пропорционалан је просечном трајању свих инстанци одговарајућег фона у бази података. На тај начин, омогућено је прецизније моделовање путање у простору обележја, по цени смањења броја опсервација по једном стању.

Број гаусијана по једном HMM стању одређује се полуаутоматски. Број се постепено повећава, све док просечна вредност логаритма веродостојности на датом валидационом скупу не почне да опада, или док се не добије максималан број гаусијана за дато стање. Максималан број гаусијана по једном стању зависи од тога ком моделу дато стање припада. Модели

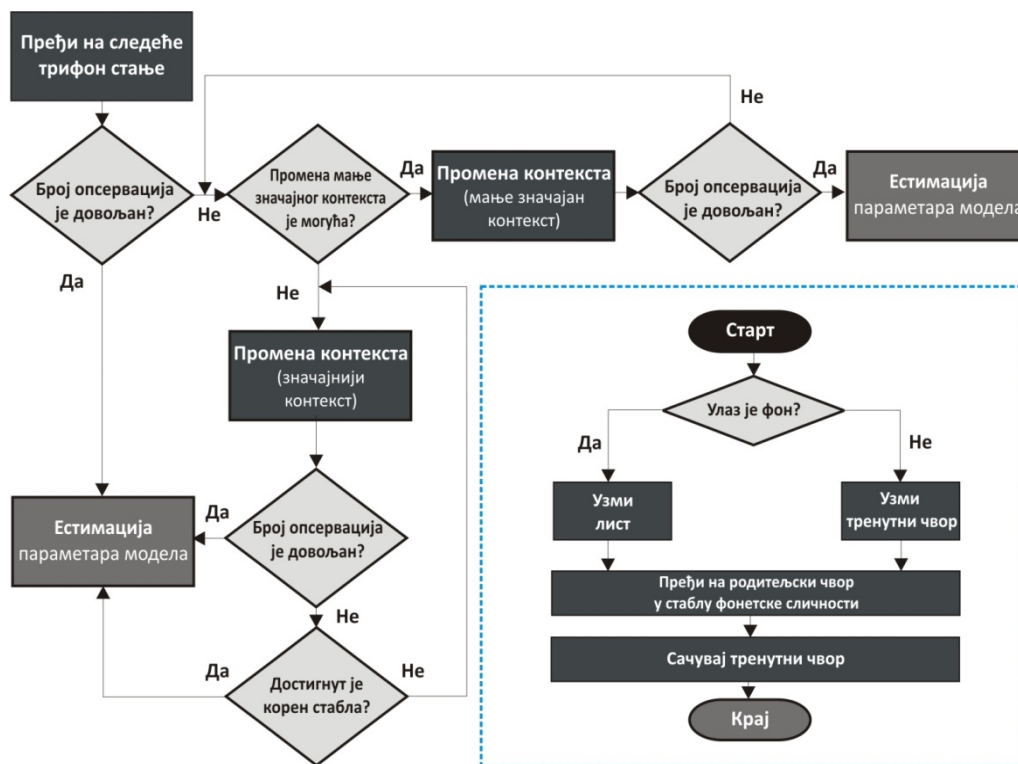
којима се моделују они фони код којих коартикулациони ефекти не долазе у великој мери до изражаја представљају се мањим бројем стања и обрнуто. Посебне методе примењују се у циљу робусније естимације сваког НММ стања, као што је на пример процедура повезивања, која подразумева додавање опсервација којима се моделују акустички блиска стања [35].

Обука система извршена је применом Tree Based Clustering (TBC) алгоритма [36][37] који омогућава дељење параметара између различитих фонетских модела. На тај начин, разрешен је проблем недостатка довољне количине података у бази за обуку, као и њихове неравномерне распоређености. TBC алгоритам користи метод повезивања НММ стања применом фонетских стабала одлуке, чиме се одређују контекстуално слична НММ стања. Под контекстуално сличним НММ стањима подразумевају се стања једног истог фона, која се налазе на истој позицији у оквиру НММ-а, који у свом контексту садржи фоне сличне према месту и начину изговора. Претходни фон у оквиру трифона остварује највећи утицај на почетна НММ стања, док наредни фон у оквиру трифона највећи утицај остварује на финална НММ стања. Позиција стања у оквиру НММ-а одређује значај контекста. Стабло фонетске сличности за српски језик приказано је на Слици 8.1 (преузето из [35]).



СЛИКА 8.1 СТАБЛО ФОНЕТСКЕ СЛИЧНОСТИ

Процедура повезивања („*tying procedure*“) примењује се само на она стања која имају недовољан број опсервација [38] **Error! Reference source not found.** Повезивање стања врши се на основу сличности левог и десног контекста. Уколико је број опсервација недовољан, позајмљују се опсервације придружене стањима чији се мање значајан контекст разликује. Уколико је потребно, процедура се наставља променом значајнијег контекста, све док се не достигне корен стабла, или постигне довољан број опсервација за дато стање. Дијаграм процедуре повезивања НММ стања илустрован је на Слици 9.2 [35].



СЛИКА 8.2 ДИЈАГРАМ ПРОЦЕДУРЕ ПОВЕЗИВАЊА НММ СТАЊА

Систем препознавања користи 26 обележја, при чему прва 24 обележја описују обвојницу спектра говорног сигнала (12 статичких и 12 динамичких мел фреквенцијских кепстралних коефицијената, тј. њихових првих извода) [15]. Преостала два обележја представљају нормализовану енергију и њен први извод. Укупно се користи 5826 акустичких стања и 29588 гаусијана [38][39]. У случају експерименталних резултата представљених у

одељку 9.4, број стања, као и укупан број гаусијана био је нешто већи, с обзиром да је на тај начин остварена нешто виша тачност препознавања (10020 акустичких стања, 37525 гаусијана [14]).

8.2. Говорна база

За потребе обуке и тестирања, коришћена је говорна база снимљена на Факултету техничких наука у Новом Саду, формирана у складу са SpeechDat(E) стандардом [40][40]. База садржи исказе од око 800 говорника (мушкарци и жене подједнако су заступљени) снимљене на излазу из телефонског канала, у укупном трајању од око 20 часова, односно 12 часова изузимајући оштећене сегменте и тишину. Транскрипције су уписане на нивоу фона, док су границе између суседних гласова ручно кориговане. Тест скуп садржи исказе од око 180 говорника (100 мушкараца и 80 жена), укупног трајања 50 минута. Фајлови су забележени у A-law формату, уз фреквенцију одабирања од 8 kHz.

Поглавље 9

Експериментални резултати

У оквиру овог поглавља представљени су резултати експеримената остварени применом предложеног алгоритма за хијерархијско кластеровање модела Гаусових смеша заснованог на принципу поделе и спајања у односу на референтни алгоритам. Експерименти су планирани, а потом и извршени у две одвојене фазе. У оквиру прве фазе извођени су експерименти над вештачки генерисаним подацима. Ова фаза имала је за циљ да предвиди понашање алгоритама у оквиру реалног система препознавања. Описана је у оквиру одељка 9.1. Друга фаза обухватала је експерименте над реалним GS системом у оквиру CSR задатка. Овај систем, као и пратећа говорна база, описани су у оквиру претходног поглавља. Тиме су потврђени резултати остварени у оквиру прве фазе. Друга фаза детаљно је описана у одељку 9.2. Приказ побољшања перформанси алгоритма селекције гаусијана остварених применом метода размотрених у оквиру Поглавља 7 представљен је у одељку 9.5.

9.1. Експерименти над синтетичким подацима, конструкција узорка

У експериментима извршеним над вештачки генерисаним подацима, применом предложених S&M MS HGMMC и S&M HGMMC алгоритама, остварена су побољшања у виду смањења просечне вредности функције цене (4.2), као и њене варијансе, у великом проценту случајева у односу на референтни HGMMC алгоритам. Побољшања су остварена за различите димензионалности простора, као и за обе апроксимације KL дивергенце коришћене у овом раду.

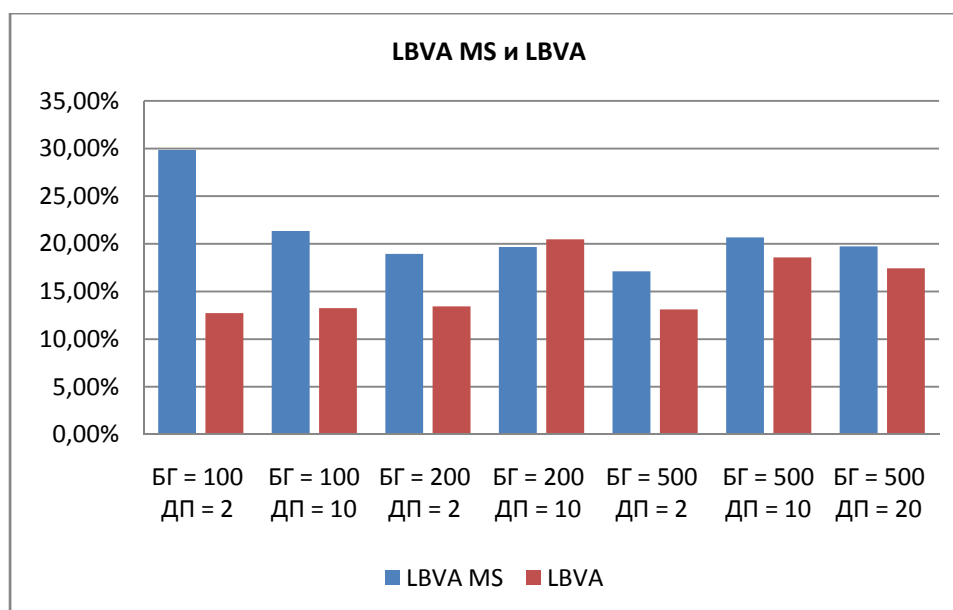
Симулације су извршене коришћењем MCSA и LBVA апроксимације за $KL(f_j^{\hat{\pi}} \parallel \hat{g}_j)$. Укупно је извршено 2800 симулација над вештачки генерисаним подацима, односно, по 100 симулација за дати алгоритам и коришћену KL апроксимацију, за сваку тестирану комбинацију броја димензија и броја Гаусових компоненти. Поље „димензионалност простора“ у Табелама 9.1 до 9.3 односи се на димензионалност простора обележја, односно на димензионалност Гаусових компоненти, док поље „број гаусијана“ представља укупан број гаусијана коришћених у оквиру датог експеримента. Резултати се добијају на следећи начин: Прво се извршава S&M MS HGMMC алгоритам, који проналази „оптималан“ број кластера $m^* = m^{(K)}$ у некој K -тој итерацији. K -та итерација представља ону итерацију у оквиру које се вредност функције цене смањила за мање од једног процента у односу на вредност из претходне, $K-1$ -ве итерације. Након конвергенције S&M MS HGMMC алгоритма, стартује C&M ХГММЦ алгоритам са задатим „оптималним“ бројем кластера m^* .

У Табели 9.1 и на Графиконима 9.1 до 9.4 приказано је побољшање у виду просечног смањења функције цене (4.2), тј. просечног растојања (KL дивергенце) између кластера и њима придружених гаусијана, остварено коришћењем предложених S&M MS HGMMC и S&M HGMMC алгоритама, у поређењу са референтним HGMMC алгоритмом, иницијализованим са истим бројем кластера. Експериментални резултати остварени применом предложених алгоритама бољи су у односу на резултате остварене применом референтног алгоритма.

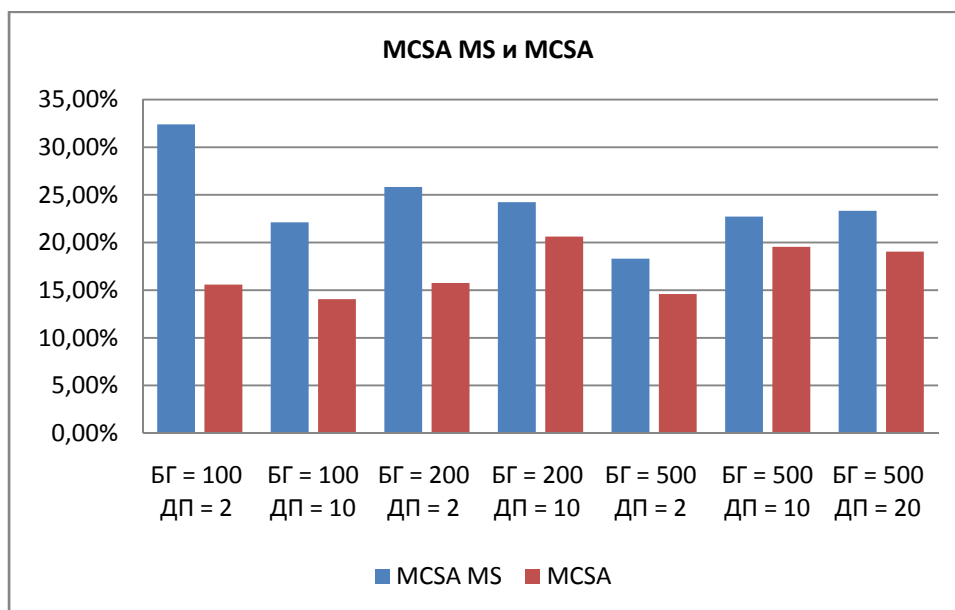
Додатно смањење просечне вредности функције цене остварено је применом предложеног S&M MS HGMMC алгоритма са растућим бројем кластера, чиме се у једном пролазу одређује „оптимални“ број кластера m^* , у односу на S&M HGMMC алгоритам иницијализован истим тим бројем кластера m^* . Може се приметити да је смањење просечне вредности функције цене у директној спрези с тачношћу препознавања у GS задатку, у оквиру реалног система препознавања.

ТАБЕЛА 9.1 СМАЊЕЊЕ ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНО КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M MS HGMMC И S&M HGMMC АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.

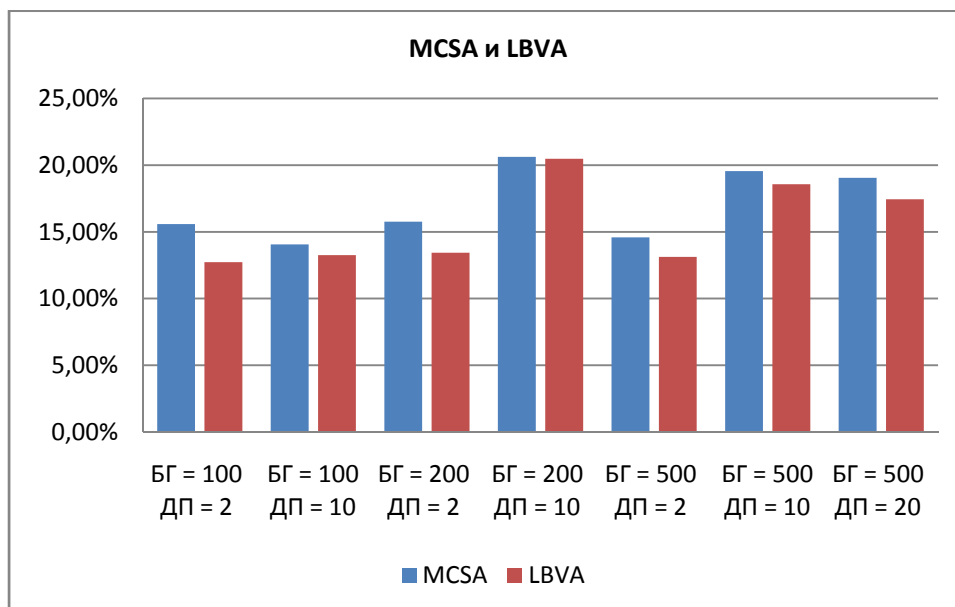
Број гаусијана	Дименз. простора	Просечно побољшање [%]			
		LBVA MS	LBVA	MCSA MS	MCSA
100	2	29.84	12.73	32.39	15.58
100	10	21.35	13.25	22.12	14.06
200	2	18.94	13.44	25.83	15.76
200	10	19.67	20.48	24.23	20.62
500	2	17.11	13.12	18.31	14.59
500	10	20.68	18.57	22.73	19.55
500	20	19.73	17.44	23.33	19.05
Средња вредност:		21.05	15.58	24.13	17.03



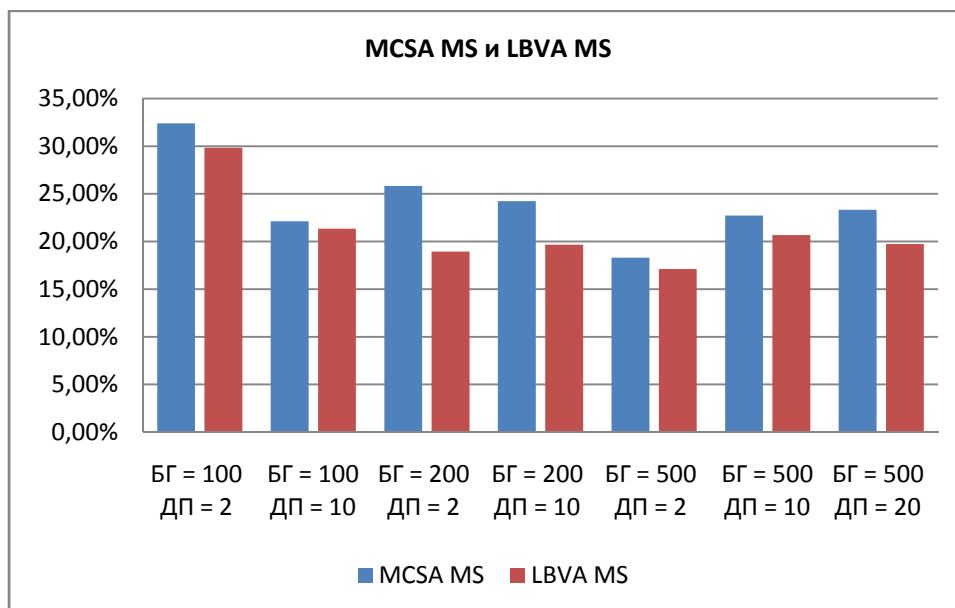
ГРАФИКОН 9.1 УПОРЕДНИ ПРИКАЗ СМАЊЕЊА ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ LBVA АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.



ГРАФИКОН 9.2 УПОРЕДНИ ПРИКАЗ СМАЊЕЊА ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ MCSA АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.



ГРАФИКОН 9.3 УПОРЕДНИ ПРИКАЗ СМАЊЕЊА ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.



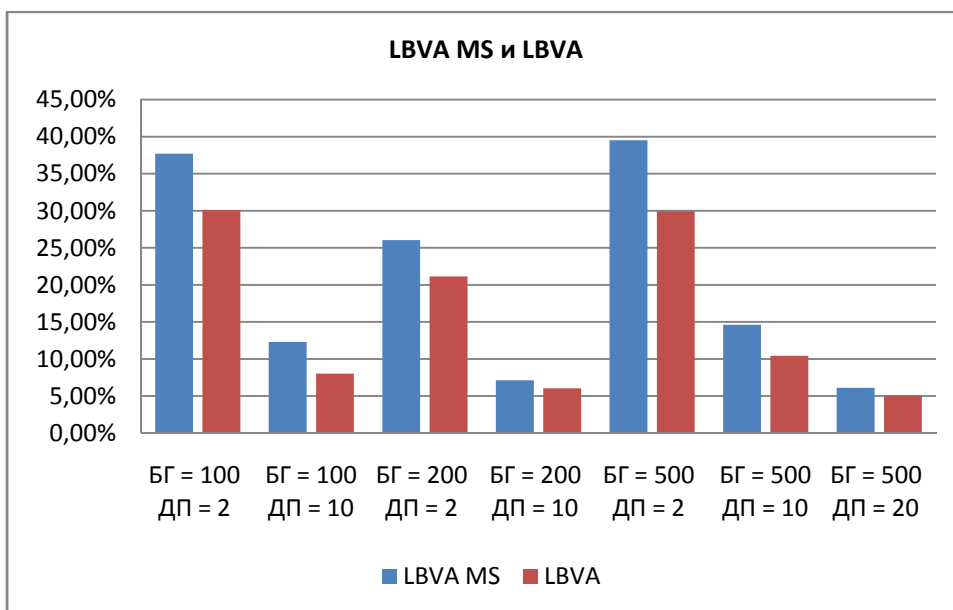
ГРАФИКОН 9.4 УПОРЕДНИ ПРИКАЗ СМАЊЕЊА ПРОСЕЧНЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M MS АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.

У Табели 9.2 и на Графиконима 9.5 до 9.8 приказано је просечно смањење стандардне девијације функције цене (4.2), за предложене S&M MS HGMMC и S&M HGMMC алгоритме, у поређењу са референтним HGMMC алгоритмом, иницијализованим са истим бројем кластера. Приметно је да су експериментални резултати за оба предложена алгоритма бољи у смислу смањења просечне вредности стандардне девијације функције цене, у поређењу с референтним HGMMC алгоритмом. Селекција модела у оквиру S&M MS HGMMC алгоритма додатно доприноси смањењу просечне вредности стандардне девијације функције цене у односу на S&M HGMMC алгоритам иницијализован са “оптималним” бројем кластера m^* . Приметно је и да нижа вредност стандардне девијације функције цене утиче на перформансе алгоритма кластеровања, у смислу смањења вероватноће да

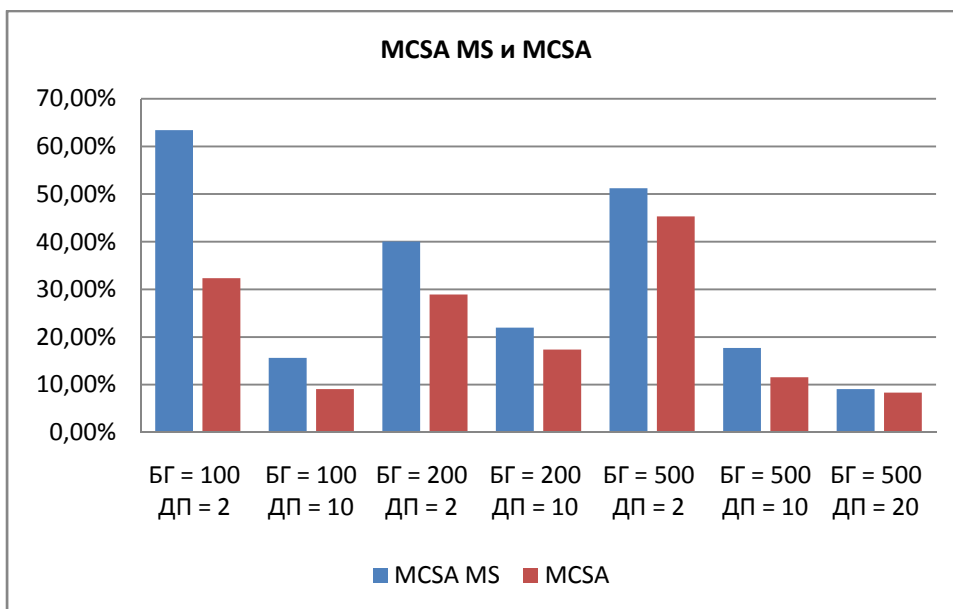
ће одређена инстанца кластеровања довести до повећања грешке препознавања у задатку селекције гаусијана у оквиру реалног система препознавања, а у односу на просечну грешку препознавања за дати метод кластеровања.

ТАБЕЛА 9.2 ПРОСЕЧНО СМАЊЕЊЕ СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНО КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M MS HGMMC И S&M HGMMC АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.

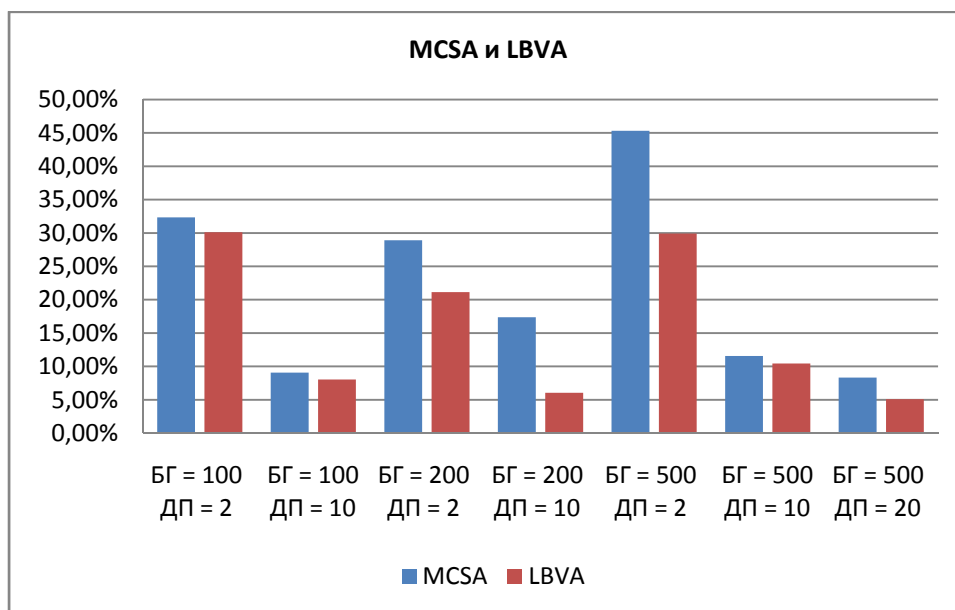
Broj gausijana	Dimenz. prostora	Prosečno poboljšanje [%]			
		LBVA MS	LBVA	MCSA MS	MCSA
100	2	37.71	30.11	63.42	32.35
100	10	12.31	8.03	15.63	9.08
200	2	26.05	21.14	40.04	28.91
200	10	7.13	6.05	21.98	17.36
500	2	39.51	29.92	51.23	45.31
500	10	14.62	10.44	17.71	11.56
500	20	6.12	5.08	9.07	8.32
Srednja vrednost:		20.49	15.82	31.30	21.84



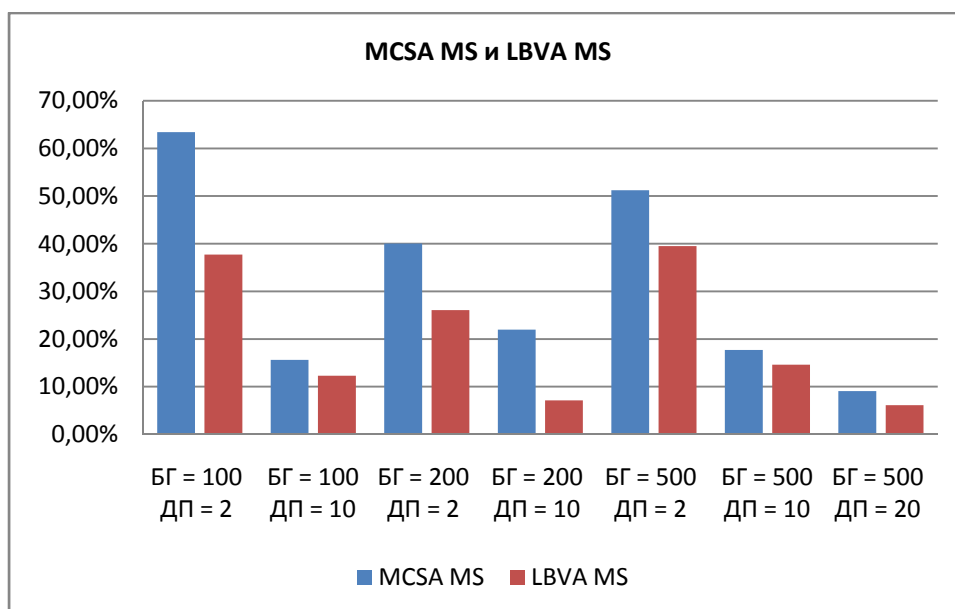
ГРАФИКОН 9.5 УПОРЕДНИ ПРИКАЗ ПРОСЕЧНОГ СМАЊЕЊА СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ LBVA АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.



ГРАФИКОН 9.6 УПОРЕДНИ ПРИКАЗ ПРОСЕЧНОГ СМАЊЕЊА СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ MCSA АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.



ГРАФИКОН 9.7 УПОРЕДНИ ПРИКАЗ ПРОСЕЧНОГ СМАЊЕЊА СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.

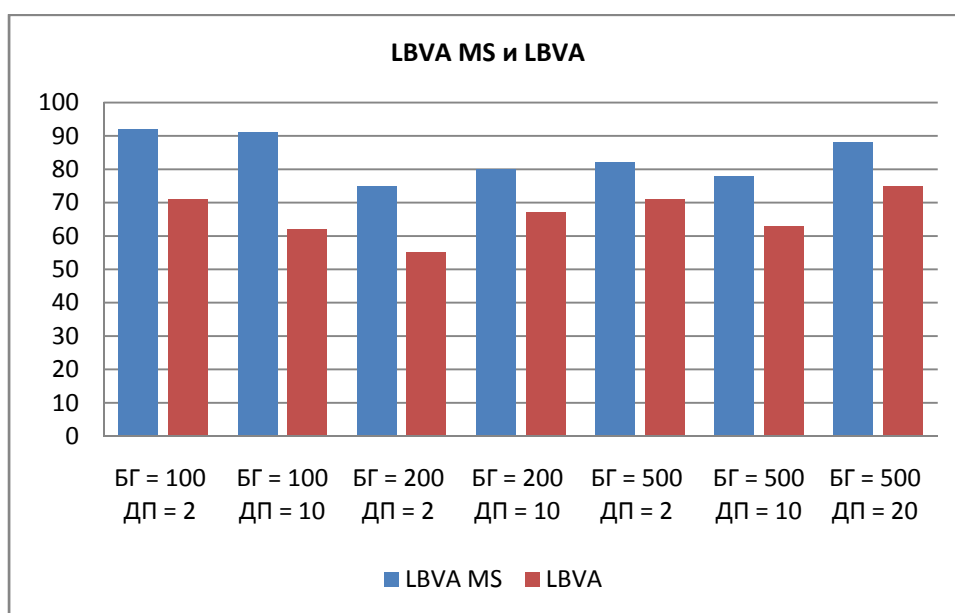


ГРАФИКОН 9.8 УПОРЕДНИ ПРИКАЗ ПРОСЕЧНОГ СМАЊЕЊА СТАНДАРДНЕ ДЕВИЈАЦИЈЕ ФУНКЦИЈЕ ЦЕНЕ, ОСТВАРЕНОГ КОРИШЋЕЊЕМ ПРЕДЛОЖЕНИХ S&M MS АЛГОРИТАМА, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.

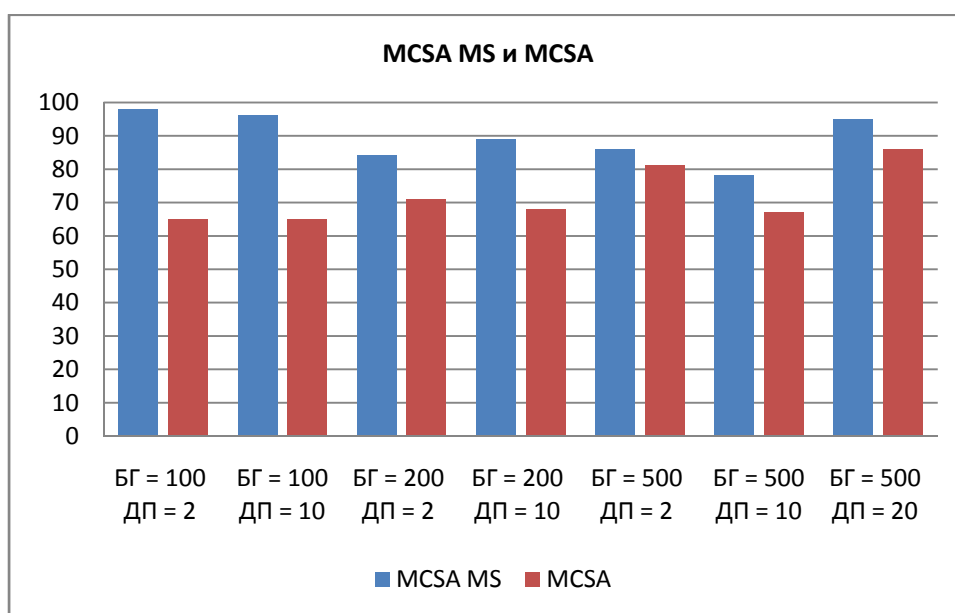
У Табели 9.3 и на Графиконима 9.9 до 9.12 приказан је процентуални удео симулација, односно број симулација у којима је дошло до смањења вредности функције цене, на 100 извршених симулација за предложене S&M MS HGMMC и S&M HGMMC алгоритме. Поставке експеримента исте су за експерименте чији су резултати приказани у све три табеле. Премда предложени приступ сукцесивног дељења и спајања кластера експлицитно не гарантује смањење вредности функције цене, статистички је могуће добити информацију о томе колико често се такво смањење дешава. Приметно је да у великом проценту случајева примена S&M алгоритама доводи до смањења вредности функције цене у односу на референтни алгоритам, за сваки од тестираних случајева. У оквиру наредног одељка, приказани су резултати остварени применом предложеног концепта поделе и спајања на реалном CSR систему.

ТАБЕЛА 9.3 ПРОЦЕНТУАЛНИ БРОЈ СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ S&M MS HGMMC И S&M HGMMC АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.

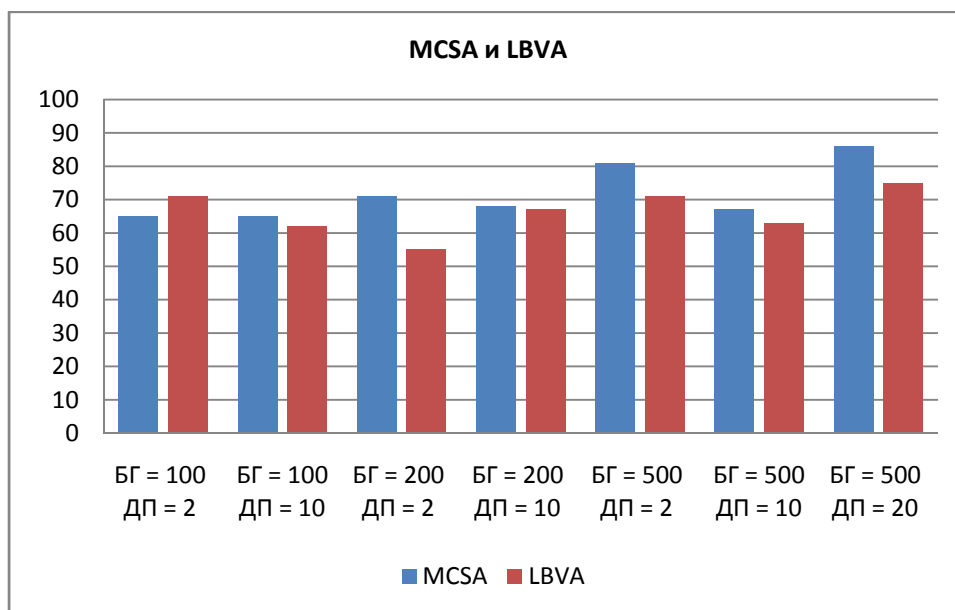
Број гаусијана	Дименз. простора	Број симулација у којима је дошло до побољшања			
		LBVA MS	LBVA	MCSA MS	MCSA
100	2	92	71	98	65
100	10	91	62	96	65
200	2	75	55	84	71
200	10	80	67	89	68
500	2	82	71	86	81
500	10	78	63	78	67
500	20	88	75	95	86
Средња вредност:		83.71	66.29	89.43	71.86



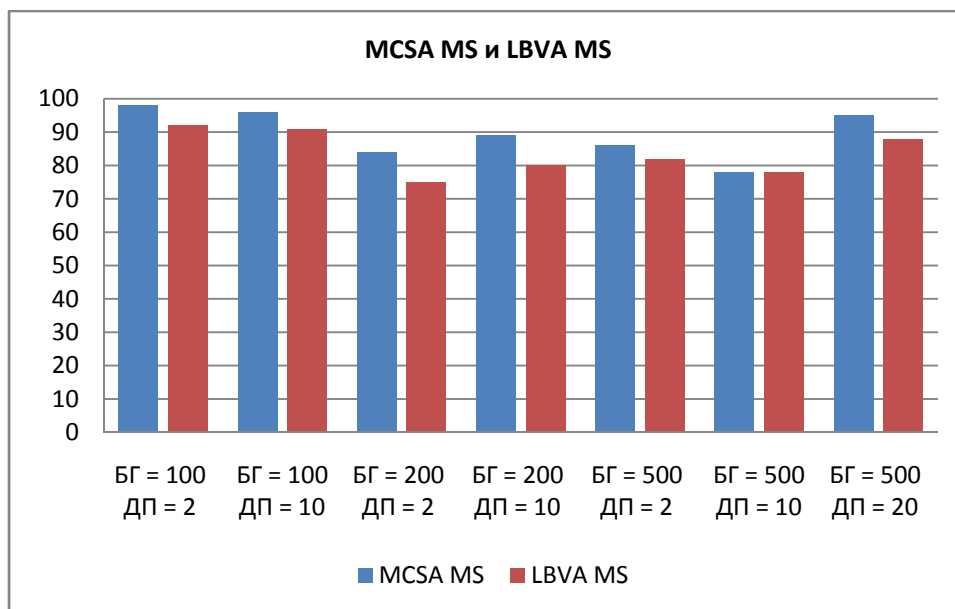
ГРАФИКОН 9.9 УПОРЕДНИ ПРИКАЗ БРОЈА СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ LBVA АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.



ГРАФИКОН 9.10 УПОРЕДНИ ПРИКАЗ БРОЈА СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ MCSA АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ НГММС АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.



ГРАФИКОН 9.11 УПОРЕДНИ ПРИКАЗ БРОЈА СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ S&M АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.



ГРАФИКОН 9.12 УПОРЕДНИ ПРИКАЗ БРОЈА СИМУЛАЦИЈА У КОЈИМА ЈЕ ЗАБЕЛЕЖЕНО СМАЊЕЊЕ ВРЕДНОСТИ ФУНКЦИЈЕ ЦЕНЕ ЗА ПРЕДЛОЖЕНЕ S&M MS АЛГОРИТМЕ, У ОДНОСУ НА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ. ЕКСПЕРИМЕНТИ СУ ИЗВРШЕНИ НАД СИНТЕТИЧКИМ ПОДАЦИМА.

9.2. Примена хијерархијског кластеровања у оквиру реалног CSR система

Алгоритми предложени у овом раду примењени су у задатку селекције гаусијана, у оквиру реалног CSR система описаног у оквиру претходног поглавља, и упоређени с резултатима референтног HGMMC алгоритма, описаног у Поглављу 4, као и с резултатима система који не користи селекцију гаусијана (полазни систем).

Перформансе GS метода прате се кроз релацију тачности препознавања и смањења броја оних Гаусових компоненти за које се израчунавања врше директно. Перформансе система препознавања у смислу његове тачности изражене су у термину стандардне грешке препознавања (*Word Error Rate - WER*). Смањење потребног броја операција изражено је у следећој форми:

$$CF = \frac{G_{new} + R_{comp}}{G_{full}} \quad (9.1)$$

Термини G_{new} и G_{full} представљају просечан број гаусијана израчунат по једном фрејму у систему који користи GS и у полазном систему респективно. R_{comp} представља број израчунавања неопходан за одређивање логаритамских вероватноћа хипер-гаусијана, у циљу доношења одлуке о томе да ли ће компоненте одређене Гаусове смеше бити егзактно евалуиране или не. Овај термин може се апроксимирати односом $R_{comp} \approx \frac{G_{new}}{n_{avr}}$, при чему је n_{avr} просечан број полазних гаусијана по једном кластеру, те се користи следећа апроксимација:

$$CF \approx \frac{G_{new}}{G_{full}} + \frac{1}{n_{avr}} \quad (9.2)$$

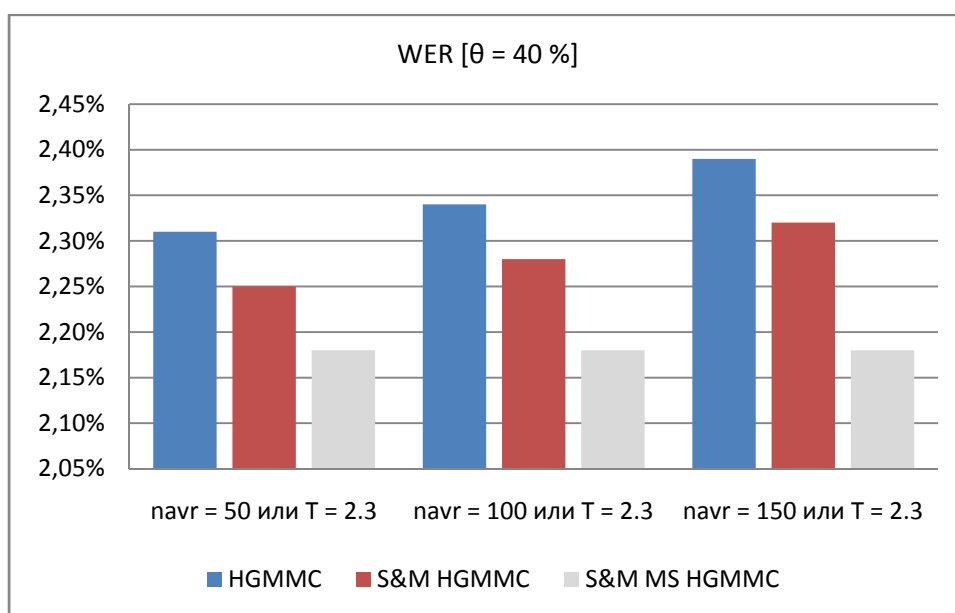
У фази декодовања, односно препознавања у GS систему, за сваки улазни вектор опсервација одређују се логаритамске вероватноће свих хипер-гаусијана. За предефинисани проценат θ [%] хипер-гаусијана са највећим логаритамским вероватноћама Гаусове компоненте које припадају поменутиим кластерима процењују се директно за дату улазну опсервацију,

док се за остале компоненте логаритамске вероватноће апроксимирају логаритамским вероватноћама припадајућих хипер-гаусијана, одређеним за дату улазну опсервацију.

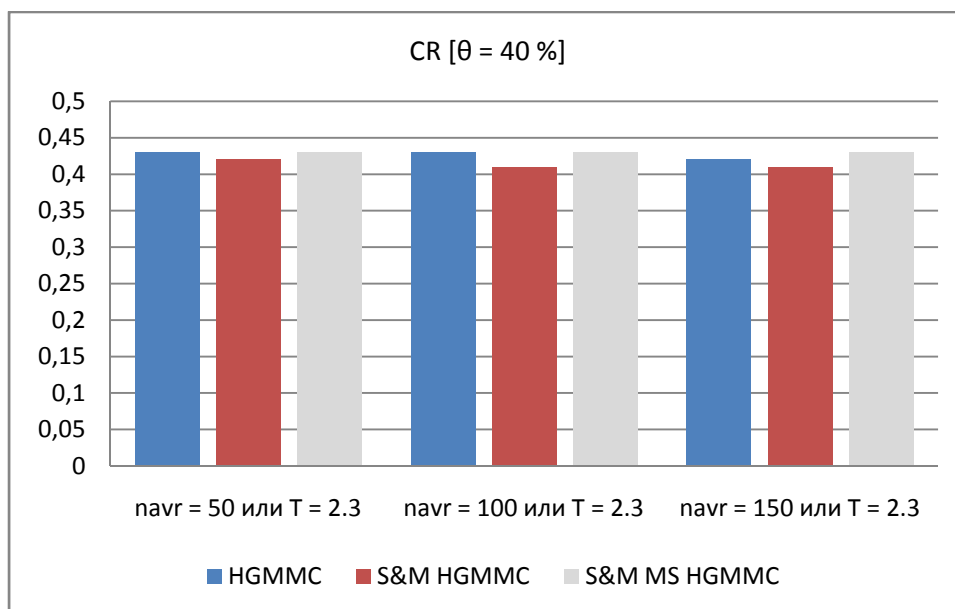
У Табели 9.4 и на Графиконима 9.13 до 9.18 упоредо су приказане перформансе полазног система, GS система који користи референтни HGMMC алгоритам, као и GS система који користи S&M MS HGMMC алгоритам. У свим експериментима у оквиру реалног CSR система примењена је само LBVA апроксимација KL дивергенце (5.14), с обзиром да би примена MCSA апроксимације у систему који користи тако велики број гаусијана захтевала превелики број рачунских операција, без обзира на чињеницу да се процес кластеровања одвија *off-line*. Може се запазити да S&M HGMMC алгоритам остварује боље резултате у смислу односа између тачности препознавања (WER) и рачунске комплексности у односу на базни HGMMC алгоритам, за све вредности n_{avr} и самим тим за сваки предефинисани број кластера $m = \left\lfloor \frac{K}{n_{avr}} \right\rfloor$ тестиран у оквиру поменутих експеримената. Такође се може приметити да се коришћењем S&M MS HGMMC алгоритма остварују бољи резултати у смислу поменутог односа, уз одговарајући одабир прага T (параметар је одређен емпиријски и постављен на $T = 2.3$), у односу на друге приказане методе, за сваку вредност θ коришћену у експериментима.

ТАБЕЛА 9.4 РЕЗУЛТАТИ ПРИМЕНЕ ПРЕДЛОЖЕНИХ S&M HGMMC И S&M MS HGMMC АЛГОРИТАМА, РЕФЕРЕНТНОГ HGMMC АЛГОРИТМА И ПОЛАЗНОГ СИСТЕМА У ТЕРМИНИМА WER И CF ФАКТОРА, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА.

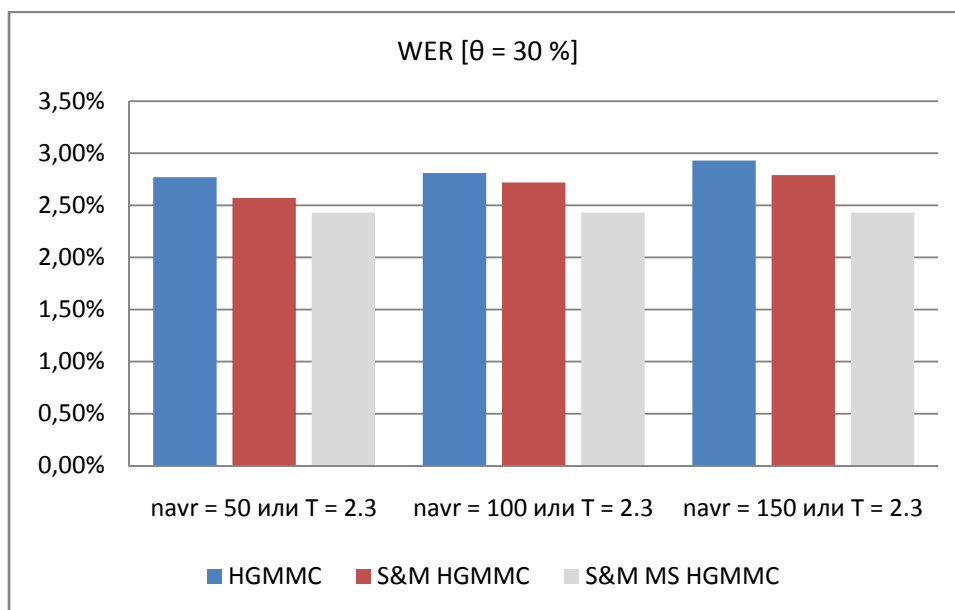
Перформансе система	n_{avr}	T	m	θ [%]	WER[%]	CF
Полазни систем	-	-	-	-	2.12	1.0
HGMMC	50	-	591	40	2.31	0.43
	100	-	295	40	2.34	0.43
	150	-	197	40	2.39	0.42
	50	-	591	30	2.77	0.34
	100	-	295	30	2.81	0.33
	150	-	197	30	2.93	0.32
	50	-	591	20	3.37	0.23
	100	-	295	20	3.42	0.23
	150	-	197	20	3.63	0.22
S&M HGMMC	50	-	591	40	2.25	0.42
	100	-	295	40	2.28	0.41
	150	-	197	40	2.32	0.41
	50	-	591	30	2.57	0.33
	100	-	295	30	2.72	0.32
	150	-	197	30	2.79	0.32
	50	-	591	20	3.15	0.23
	100	-	295	20	3.19	0.22
	150	-	197	20	3.23	0.21
S&M MS HGMMC	-	2.3	583	40	2.18	0.43
	-	2.3	583	30	2.43	0.32
	-	2.3	583	20	3.08	0.22



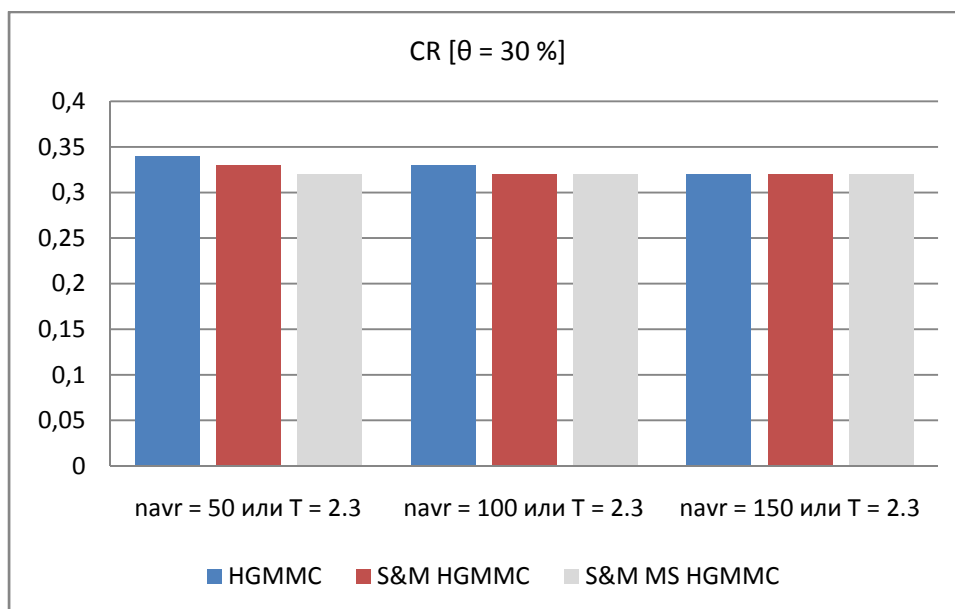
ГРАФИКОН 9.13 WER ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [θ = 40 %].



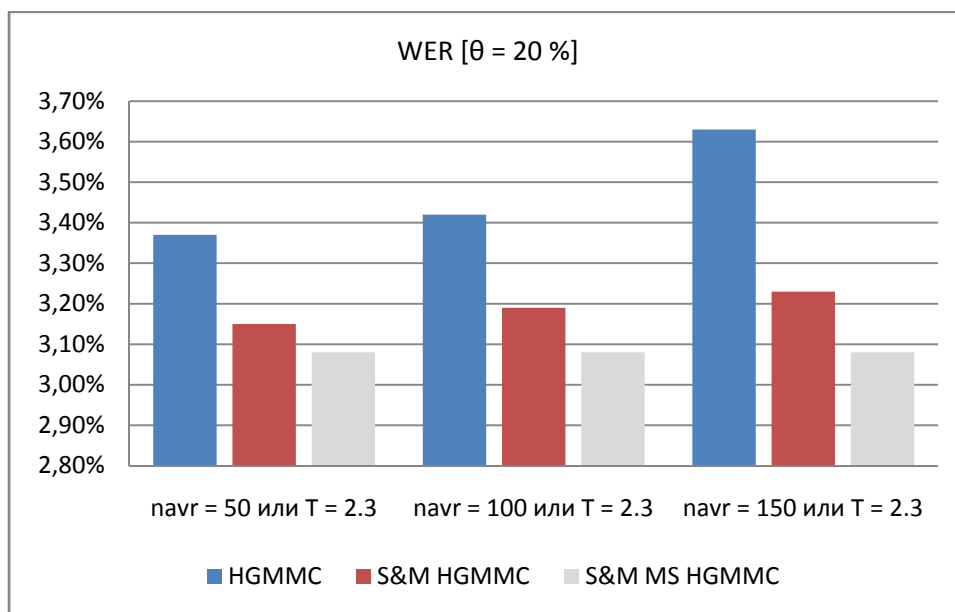
ГРАФИКОН 9.14 CF ФАКТОР ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [θ = 40 %].



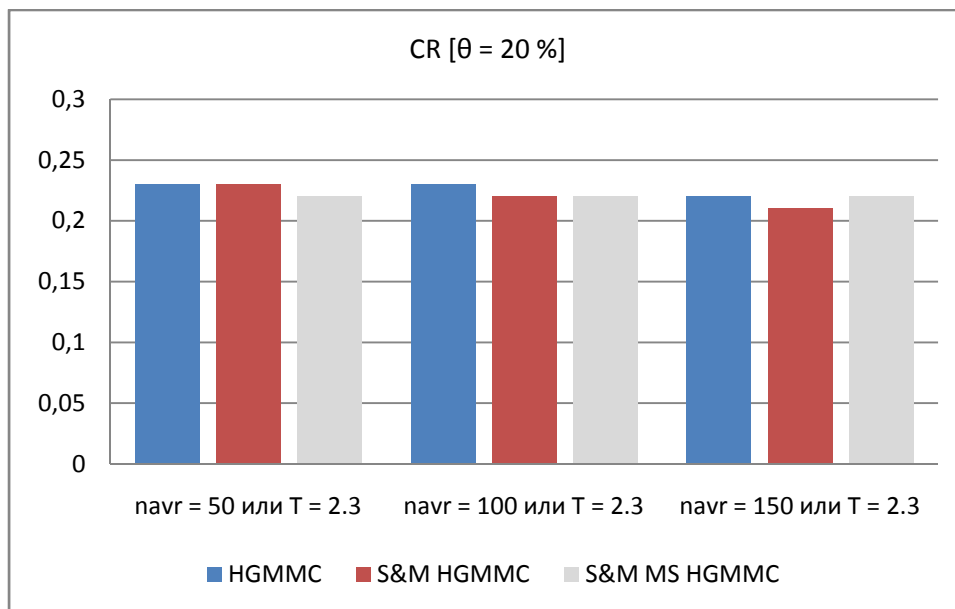
ГРАФИКОН 9.15 WER ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 30\%$].



ГРАФИКОН 9.16 CF ФАКТОР ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 30\%$].



ГРАФИКОН 9.17 WER ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 20\%$].



ГРАФИКОН 9.18 CF ФАКТОР ЗА РЕФЕРЕНТНИ HGMMC АЛГОРИТАМ И ПРЕДЛОЖЕНЕ S&M HGMMC И S&M MS HGMMC АЛГОРИТМЕ, ПОСМАТРАНО У GS ЗАДАТКУ, У ОКВИРУ РЕАЛНОГ CSR СИСТЕМА [$\theta = 20\%$].

Потребно је нагласити да су времена извршавања за предложене S&M HGMMC и S&M MS HGMMC алгоритме знатно већа у односу на референтни HGMMC алгоритам. Међутим, ово ни на који начин не утиче на примену предложених алгоритама, с обзиром да се оба ова алгоритма користе у процесу *off-line* обуке, при чему не постоји зависност брзине препознавања и потребног броја рачунских операција за поменуте алгоритме.

9.3. Перформансе алгоритма селекције гаусијана

У овом одељку приказани су експериментални резултати остварени применом метода описаних у Поглављу 7, тј. избором одговарајуће комбинације параметара алгоритма селекције гаусијана [14]. CSR НММ систем описан је у одељку 8.1, док је говорна база описана у одељку 8.2. И у овом случају, тренирање система извршено је применом ТВС алгоритма. Систем користи 10020 физичких стања, као и 37525 гаусијана. Груписање гаусијана извршено је у 6 група, при чему су границе интервала постављене као $\tau_{vec} = [2 \ 2.8 \ 3.92 \ 5.48 \ 7.68]$. Границе су одабране тако да једначина $\tau(i+1) = c\tau(i)$ буде задовољена (у овом случају, $c = 1.4$). То је урађено да би се на тај начин испратио ред величине сопствених вредности, које су мултипликативне природе.

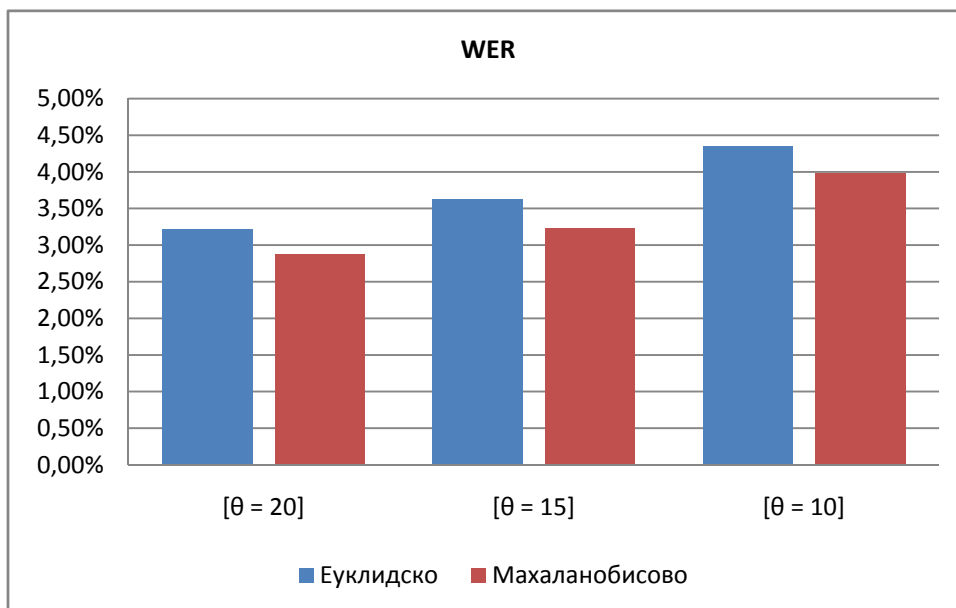
Уместо апсолутне вредности прага, као и у претходном случају, користи се предефинисани проценат најбољих хипер-гаусијана, при чему се број хипер-гаусијана одређује променом параметра n_{avr} , у циљу добијања најповољнијег односа између тачности и рачунске сложености у CSR задатку. У том циљу, експерименти су извођени уз промену предефинисаног броја кластера, типа растојања коришћеног при евалуацији хипер-гаусијана, броја хипер-гаусијана чији се придружени гаусијани евалуирају директно, као одговарајућег типа кластерована у смислу степена преклапања између појединачних кластера.

Упоредни приказ перформанси GS система који користи еуклидско, односно, Махаланобисово растојање у процесу евалуације хипер-гаусијана, приказан је у Табели 9.5 и на Графиконима 9.19 и 9.20. Просечан број полазних Гаусових компоненти постављен је на фиксну вредност ($n_{avr} = 50$), при чему се варира број директно евалуираних хипер-гаусијана. Систем

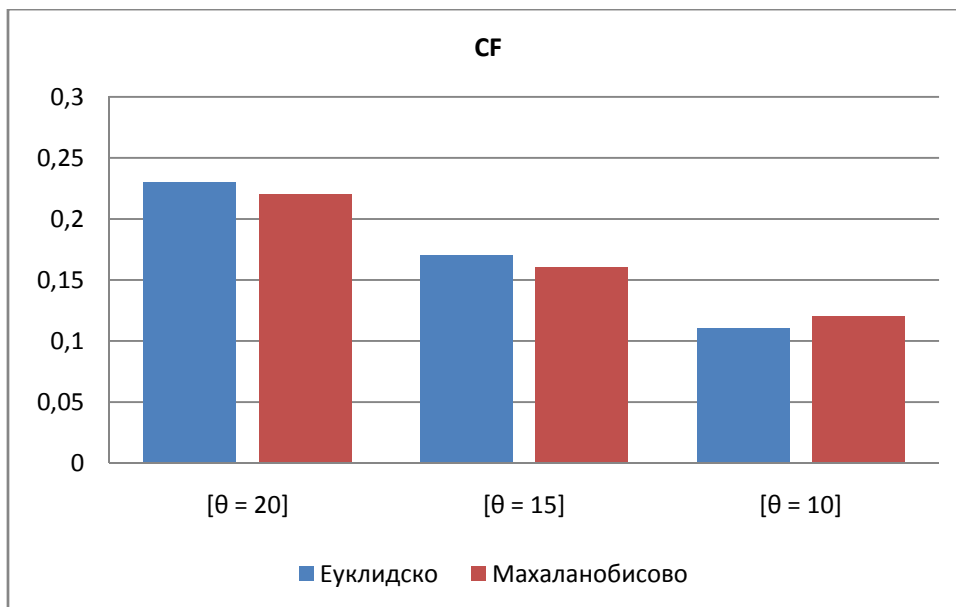
користи KL дивергенцу у процесу кластеровања. На основу резултата приказаних у табели, могуће је закључити да примена Махаланобисовог растојања при декодовању даје боље резултате у смислу односа тачности и рачунске сложености у задатку препознавања, односно, у смислу односа WER и CF фактора, од примене еуклидског растојања или логаритамске веродостојности.

ТАБЕЛА 9.5 УПОРЕДНИ ПРИКАЗ ПЕРФОРМАНСИ АЛГОРИТМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И CF ФАКТОРА ПРИ ПРОМЕНИ ТИПА РАСТОЈАЊА КОРИШЋЕНОГ ПРИ ДЕКОДОВАЊУ.

Перформансе система	n_{avr}	Тип растојања	θ [%]	WER[%]	CF
GS 1	50	еуклидско	20	3.21	0.23
GS 1	50	еуклидско	15	3.62	0.17
GS 1	50	еуклидско	10	4.35	0.11
GS 2	50	Махаланобисово	20	2.87	0.22
GS 2	50	Махаланобисово	15	3.23	0.16
GS 2	50	Махаланобисово	10	3.98	0.12



ГРАФИКОН 9.19 WER У ЗАДАТКУ СЕЛЕКЦИЈЕ ГАУСИЈАНА ПРИ ПРОМЕНИ ТИПА РАСТОЈАЊА КОРИШЋЕНОГ ПРИ ДЕКОДОВАЊУ ($n_{avr} = 50$).

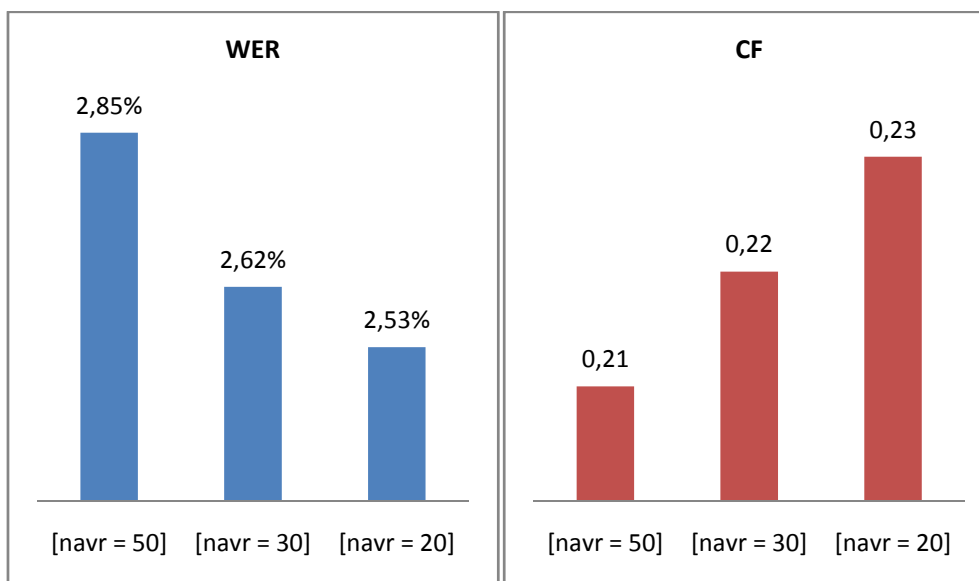


ГРАФИКОН 9.20 CF ФАКТОР У ЗАДАТКУ СЕЛЕКЦИЈЕ ГАУСИЈАНА ПРИ ПРОМЕНИ ТИПА РАСТОЈАЊА КОРИШЋЕНОГ ПРИ ДЕКОДОВАЊУ ($n_{avr} = 50$).

У Табели 9.6 и на Графикону 9.21 приказане су перформансе GS система, који користи процедуру кластеровања са малим степеном преклапања између новонасталих кластера (GS 3). Систем користи KL дивергенцу у процесу кластеровања, као и Махаланобисово растојање у процесу декодовања, с обзиром да се на тај начин остварује најповољнији однос између WER и CF фактора. Систем је тестиран применом различитог просечног броја гаусијана по једном кластеру (n_{avr}). Приметно је да ниже вредности броја n_{avr} (тј. већи број кластера) дају боље резултате у смислу односа тачности и рачунске сложености, у задатку препознавања.

ТАБЕЛА 9.6 ПЕРФОРМАНСЕ АЛГОРИТМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И CF ФАКТОРА ПРИ ПРОМЕНИ ПРОСЕЧНОГ БРОЈА ГАУСИЈАНА ПО ЈЕДНОМ КЛАСТЕРУ. СИСТЕМ КОРИСТИ МАХАЛАНОБИСОВО РАСТОЈАЊЕ У ПРОЦЕСУ ЕВАЛУАЦИЈЕ.

Перформансе система	n_{avr}	Тип растојања	θ [%]	WER[%]	CF
GS 3	50	Махаланобисово	20	2.85	0.21
GS 3	30	Махаланобисово	20	2.62	0.22
GS 3	20	Махаланобисово	20	2.53	0.23

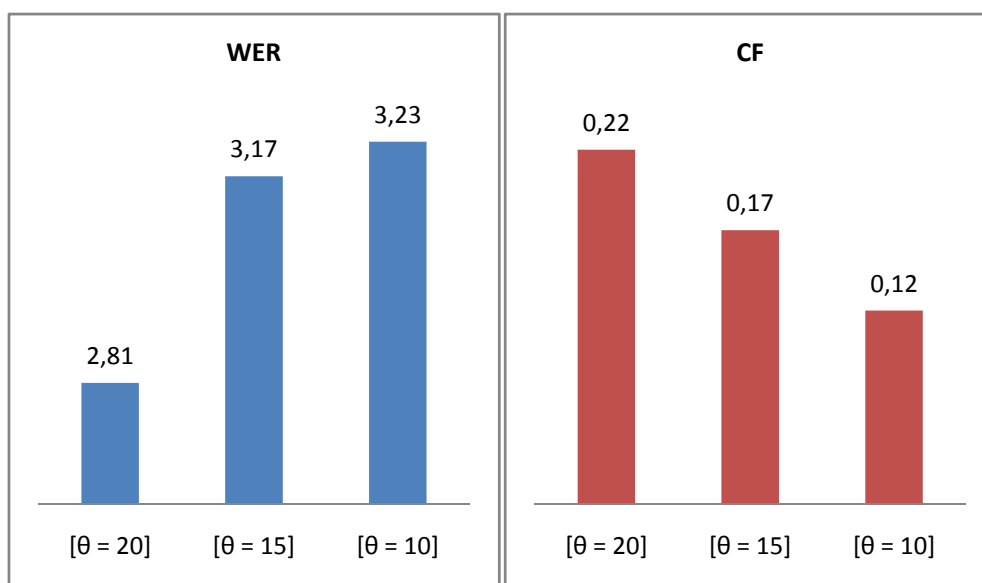


ГРАФИКОН 9.21 УПОРЕДНИ ПРИКАЗ ПЕРФОРМАНСИ АЛГОРИТМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И CF ФАКТОРА ПРИ ПРОМЕНИ ПРОСЕЧНОГ БРОЈА ГАУСИЈАНА ПО ЈЕДНОМ КЛАСТЕРУ. СИСТЕМ КОРИСТИ МАХАЛАНОБИСОВО РАСТОЈАЊЕ У ПРОЦЕСУ ЕВАЛУАЦИЈЕ.

Перформансе GS система у оквиру којег се врши евалуација само једног кластера по фрејму (GS 3), при чему кластери заузимају значајан део простора обележја (између 10 % и 20 %) приказане су у Табели 9.7 и на Графикону 9.22. Јасно је да приступ у коме кластери заузимају мањи део простора обележја даје боље резултате, јер такав приступ обезбеђује виши ниво прилагођавања локалним расподелама. Приступ у коме кластери заузимају велики део простора обележја истовремено значи да је велики број оних гаусијана који се налазе далеко од центра одговарајућег хипер-гаусијана, а самим тим је и однос WER и CF фактора у оквиру CSR задатка лошији.

ТАБЕЛА 9.7 УПОРЕДНИ ПРИКАЗ ПЕРФОРМАНСИ АЛГОРИТМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И CF ФАКТОРА. ВРШИ СЕ ЕВАЛУАЦИЈА САМО ЈЕДНОГ КЛАСТЕРА ПО ФРЕЈМУ, ПРИ ЧЕМУ КЛАСТЕРИ ЗАУЗИМАЈУ ЗНАЧАЈАН ДЕО ПРОСТОРА ОБЕЛЕЖЈА. СИСТЕМ КОРИСТИ МАХАЛАНОВИСОВО РАСТОЈАЊЕ У ПРОЦЕСУ ЕВАЛУАЦИЈЕ.

Перформансе система	n_{avr}	Тип растојања	θ [%]	WER[%]	CF
GS 4	50	Махалановисово	20	2.81	0.22
GS 4	50	Махалановисово	15	3.17	0.17
GS 4	50	Махалановисово	10	3.23	0.12



ГРАФИКОН 9.22 УПОРЕДНИ ПРИКАЗ ПЕРФОРМАНСИ АЛГОРИТМА СЕЛЕКЦИЈЕ ГАУСИЈАНА У ТЕРМИНИМА WER И CF ФАКТОРА. ВРШИ СЕ ЕВАЛУАЦИЈА САМО ЈЕДНОГ КЛАСТЕРА ПО ФРЕЈМУ, ПРИ ЧЕМУ КЛАСТЕРИ ЗАУЗИМАЈУ ЗНАЧАЈАН ДЕО ПРОСТОРА ОБЕЛЕЖЈА. СИСТЕМ КОРИСТИ МАХАЛАНОВИСОВО РАСТОЈАЊЕ У ПРОЦЕСУ ЕВАЛУАЦИЈЕ.

Поглавље 10

Закључак

У овом раду предложен је нови алгоритам поделе и спајања за хијерархијско кластеровање модела Гаусових смеша, који тежи побољшању резултата оствареног иницијалном констелацијом, односно, додатном приближавању локалног минимума глобалном минимуму функције цене [15]. Алгоритам се иницијализује локално оптималним параметрима, израчунатим применом референтног HGMMC алгоритма, сличног k-means алгоритму. Приближавање глобалном минимуму функције цене остварује се итеративним процесом поделе и спајања кластера Гаусових компоненти, при чему је полазна констелација остварена применом референтног HGMMC алгоритма. У циљу побољшања односа тачности и рачунске сложености, у предложени алгоритам укључен је и метод селекције модела у оквиру задатка селекције гаусијана унутар реалног процеса препознавања. Предложени метод тестиран је како над синтетичким подацима, тако и у оквиру реалног CSR система.

Просечна вредност, као и стандардна девијација функције цене (4.2) остварене у експериментима над вештачки генерисаним подацима, уз довољно широк опсег полазних параметара, као што су димензије центроида, број гаусијана и разноврсност њихових параметара, пружају увид у понашање предложених алгоритама у оквиру реалног задатка препознавања. На основу резултата експеримената приказаних у Табелама 9.1 и 9.2, јасно је да предложени алгоритми резултују смањењем просечне вредности, као и просечне стандардне девијације функције цене (4.2) у великој већини случајева (Табела 9.3). Укупно просечно смањење веће је од 15%, како за просечну вредност, тако и за просечну стандардну девијацију функције цене, у поређењу са референтним HGMMC алгоритмом.

Експериментални резултати приказани у Табели 9.4 указују на видно боље перформансе предложених S&M HGMMC и S&M MS HGMMC алгоритама у односу на референтни HGMMC алгоритам, у случају GS задатка, у оквиру реалног CSR система. Ефикасност примене процедуре

селекције модела у оквиру предложених метода сукцесивне поделе и спајања експериментално је потврђена. Додатна побољшања остварена су применом Махаланобисовог растојања у процесу евалуације хипергаусијана, као и кластеровања са малим степеном преклапања између тако формираних кластера.

Поред примене предложеног алгоритма хијерархијског кластеровања модела Гаусових смеша у оквиру задатка селекције гаусијана унутар реалног CSR система, резултате изложене у оквиру докторске дисертације могуће је применити у и оквиру других система који захтевају рад у реалном времену, у областима попут компјутерске визије, роботике, у оквиру система за препознавање и идентификацију, разноврсних система комуникације, као и у разним другим областима.

Поглавље 11

Додаци

11.1. Додатак А: Приказ HGMMC алгоритма

- Иницијализација:

- За предефинисано n_{avr} и укупан број полазних Гаусових компоненти k , број кластера одредити по формули

$$m = \left\lfloor \frac{K}{n_{avr}} \right\rfloor.$$

- На случај селектовати m различитих центроида μ_j из скупа од укупно k центроида и сваком центроиду доделити јединичне коваријансне матрице $\Sigma_j^{(0)} = I$.
Извести једну итерацију алгоритма прерасподеле и реестимације на основу формула (4.12) и (4.13). Добија се $(\hat{\pi}^{(0)}, \hat{g}^{(0)})$.

- Кластеровање:

За предефинисано $\varepsilon > 0$:

-
- Почети са $(\hat{\pi}^{(0)}, \hat{g}^{(0)})$ и итеративно примељивати алгоритам прерасподеле и реестимације на основу формула (4.12) и (4.13), све док се у некој p -тој итерацији не добије $|d(f, \hat{g}^{(p)}, \hat{\pi}^{(p)}) - d(f, \hat{g}^{(p-1)}, \hat{\pi}^{(p-1)})| < \varepsilon$.
 - $(\pi^*, g^*) = (\pi^{(p)}, g^{(p)})$ усваја се као коначно решење.

11.2. Додатак Б: Приказ S&M HGMMC алгоритма

– Иницијализација:

- За предефинисани број кластера $m = \lfloor \frac{K}{n_{avr}} \rfloor$ и фиксно $\varepsilon > 0$, итеративно примењивати алгоритам прерасподеле и реестимације, до конвергенције, при чему се добија иницијално решење $(\hat{\pi}^{(0)}, \hat{g}^{(0)}) = (\hat{\pi}, \hat{g})$.

– Кластеровање:

Почети са иницијалним решењем $(\hat{\pi}^{(0)}, \hat{g}^{(0)})$ и понављати доле описани поступак, све док се у некој p -тој итерацији не добије $d(f, \hat{g}^{(p)}, \hat{\pi}^{(p)}) \geq d(f, \hat{g}^{(p-1)}, \hat{\pi}^{(p-1)})$.

- Нека је $(\hat{\pi}, \hat{g}) \leftarrow (\hat{\pi}^{(p)}, \hat{g}^{(p)})$. Селектовати кандидате за операцију спајања $\hat{\pi}_{\hat{k}_1}, \hat{\pi}_{\hat{k}_2} \in \hat{\pi}$ на основу критеријума

(5.7). Повезати кластере \hat{k}_1 и \hat{k}_2 у јединствени кластер \hat{k} ,

тј. $\hat{\pi}_{\hat{k}} = \hat{\pi}_{\hat{k}_1} \cup \hat{\pi}_{\hat{k}_2}$

- Извршити корак оптимизације: одредити $\hat{g}_{\hat{k}}$ обједињавањем смеше (5,9), користећи естимацију описану у (4.9).
- Одабрати кандидате $\hat{\pi}_j = \hat{\pi} \setminus \{\hat{\pi}_{\hat{k}_1}, \hat{\pi}_{\hat{k}_2}\}$ за извођење операције поделе коришћењем критеријума (5.10), и одговарајуће апроксимације ((5.11) или (5.14)).
- Извршити корак оптимизације: иницијално одредити $(\hat{\beta}_{j_1}^{(0)}, \hat{\beta}_{j_2}^{(0)})$, $(\hat{\mu}_{j_1}^{(0)}, \hat{\Sigma}_{j_1}^{(0)})$ и $(\hat{\mu}_{j_2}^{(0)}, \hat{\Sigma}_{j_2}^{(0)})$, користећи (5.24) до (5.27). Извести једну операцију поновног груписања да би се добио одговарајући пар кластера $(\hat{\pi}_{j_1}^{(0)}, \hat{\pi}_{j_2}^{(0)})$. Извршавати алгоритам поновног груписања и придруживања над $(\hat{\pi}_{j_1}^{(0)}, \hat{\pi}_{j_2}^{(0)})$, све до конвергенције у некој S -тој итерацији, користећи предефинисано $\varepsilon > 0$, при чему се добија $(\hat{\pi}_{j_1}, \hat{\pi}_{j_2}) = (\hat{\pi}_{j_1}^{(S)}, \hat{\pi}_{j_2}^{(S)})$, $(\hat{g}_{j_1}, \hat{g}_{j_2}) = (\hat{g}_{j_1}^{(S)}, \hat{g}_{j_2}^{(S)})$ и $(\hat{\beta}_{j_1}, \hat{\beta}_{j_2}) = (\hat{\beta}_{j_1}^{(S)}, \hat{\beta}_{j_2}^{(S)})$.

- Извршити једну итерацију алгоритма поновног груписања и придруживања применом формула (4.12) и (4.13), над $(\tilde{\pi} \setminus \{\hat{\pi}_{\hat{k}_1}, \hat{\pi}_{\hat{k}_2}, \hat{\pi}_j\}) \cup \{\hat{\pi}_{\hat{k}}, \hat{\pi}_{j_1}, \hat{\pi}_{j_2}\}$, све до конвергенције у зависности од $\varepsilon > 0$, при чему се добија $(\hat{\pi}^{(p)}, \hat{g}^{(p)})$.

Усвојити коначно решење $(\pi^*, g^*) = (\pi^{(p-1)}, g^{(p-1)})$.

11.3. Додатак В: Приказ S&M MS HGMMC алгоритма

– Иницијализација:

- Почети са иницијалним бројем кластера $m^{(0)} = 2$ и итеративно примењивати алгоритам поновног груписања и придруживања, при чему се добија $(\hat{\pi}_0^{(0)}, \hat{g}_0^{(0)})$.

– Кластеровање:

Примењивати доле описану процедуру, све док се не добије

$$d(f, \hat{g}_0^{(K)}, \hat{\pi}_0^{(K)}) \geq d(f, \hat{g}_0^{(K-1)}, \hat{\pi}_0^{(K-1)}) - T \text{ у некој } K\text{-тој}$$

итерацији, за предефинисани праг $T > 0$.

- Применити операције поделе и спајања: У k -том кораку алгоритма поделе и спајања добија се $(\hat{\pi}_{s,k}^{(p)}, \hat{g}_{s,k}^{(p)})$ и

$(\hat{\pi}_{m,k}^{(p)}, \hat{g}_{m,k}^{(p)})$ (p је индекс итерације). Уколико важи да је $d(f, \hat{g}_{m,k}^{(p)}, \hat{\pi}_{m,k}^{(p)}) < d(f, \hat{g}_{k-1}^{(p)}, \hat{\pi}_{k-1}^{(p)})$, при чему су $(\hat{\pi}_{k-1}^{(p)}, \hat{g}_{k-1}^{(p)})$ одређени у претходном, $k-1$ -ом кораку, усваја се $(\hat{\pi}_k^{(p)}, \hat{g}_k^{(p)}) \leftarrow (\hat{\pi}_{m,k}^{(p)}, \hat{g}_{m,k}^{(p)})$. У супротном случају, изводи се нова итерација алгоритма поделе и спајања, при чему се усваја $p \leftarrow p + 1$, $m^{(p+1)} \leftarrow m^{(p)} + 1$ и $(\hat{\pi}_0^{(p+1)}, \hat{g}_0^{(p+1)}) \leftarrow (\hat{\pi}_{s,k}^{(p)}, \hat{g}_{s,k}^{(p)})$.

Усвојити коначно решење $(\pi^*, g^*) = (\hat{\pi}_0^{(K)}, \hat{g}_0^{(K)})$.

Литература

- [1] Wang J (2007) Discriminative Gaussian mixtures for interactive image segmentation. Proc ICASSP, Honolulu, HI 1:l-601-l-604. doi: 10.1109/ICASSP.2007.365979
- [2] Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 77(2): 257-286. doi: 10.1109/5.18626
- [3] Reynolds DA, Rose RC (1995) Robust text-independent speaker identification using Gaussian mixture speaker models. IEEE Trans Speech Audio Process 3(1):72-83. doi: 10.1109/89.365379
- [4] Shin KS, Jeong Y-S, Jeong MK (2011) A two-leveled symbiotic evolutionary algorithm for clustering problems. Appl Intel, published online 08 July 2011. doi: 10.1007/s10489-011-0295-y
- [5] Bahrapour S, Moshiri B, Salahshoor K (2011) Weighted and constrained possibilistic C-means clustering for online fault detection and isolation. Appl Intel 35(2):269-284. doi: 10.1007/s10489-010-0219-2
- [6] Korkmaz EE (2010) Multi-objective genetic algorithms for grouping problems, Appl Intel 33(2):179-192. doi: 10.1007/s10489-008-0158-3
- [7] Goldberger J, Roweis S (2005) Hierarchical clustering of a mixture model. Adv Neural Inf Process Syst 17:505-512
- [8] Bocchieri E (1993) Vector quantization for efficient computation of continuous density likelihoods. Proc ICASSP, Minneapolis, MN 2:II-692-II-695. doi: 10.1109/ICASSP.1993.319405
- [9] Knill KM, Gales MJF, Young SJ (1996) Use of Gaussian selection in large vocabulary continuous speech recognition using HMMs. Proc ICSLP 1:470-473. doi: 10.1109/ICSLP.1996.607156
- [10] Simonin J, Delphin L, Damnati G (1998) Gaussian density tree structure in a multi-Gaussian HMM based speech recognition system. 5-th Int Conf Spok Lang Process, Sidney, Australia
- [11] Watanabe T, Shinoda K, Takagi K, Iso K-I (1995) High speed speech recognition using tree-structured probability density function. Proc ICASSP 1:556-559. doi: 10.1109/ICASSP.1995.479658

-
- [12] Janev M, Pekar D, Jakovljevic N, Delic V (2010) Eigenvalues driven Gaussian selection in continuous speech recognition using HMM's with full covariance matrices. *Appl Intel* 33(2):107-116. doi: 10.1007/s10489-008-0152-9
- [13] Shinoda K, Lee C-H (2001) A structural Bayes approach to speaker adaptation. *IEEE Trans Speech Audio Process* 9(3):276-287. doi: 10.1109/89.906001
- [14] Pekar D, Janev M, Jakovljevic N, Popovic B, Delic V (2011) Improving the Performance of Gaussian Selection Algorithm, *Proc SPECOM 2011*, 89-95, Kazan, Russia. ISBN: 978-5-88983-395-6
- [15] Popovic B, Janev M, Pekar D, Jakovljevic N, Gnjatovic M, Secujski M, Delic V (2012) A Novel Split-and-Merge Algorithm for Hierarchical Clustering of Gaussian Mixture Models, *Appl Intel*, Online First, 11.01.2012. doi: 10.1007/s10489-011-0333-9
- [16] Linde Y, Buzo A, Gray R (1980) An algorithm for vector quantizer design. *IEEE Trans Commun* 26(1):84-95. doi: 10.1109/TCOM.1980.1094577
- [17] McCrosky J (2008) A new measure for clustering model selection. Master thesis, University of Waterloo, Waterloo, Ontario, Canada
- [18] Axelrod S, Goel V, Gopinaht RA, Olsen PA, Visweswariah K (2005) Subspace constrained Gaussian mixture models for speech recognition, *IEEE Trans Speech Audio Process* vol. 13(6):1144-1160. doi: 10.1109/TSA.2005.851965
- [19] Dharanipragada S, Visweswariah K (2006) Gaussian mixture models with covariances or precisions in shared multiple subspaces, *IEEE Trans Audio Speech Lang Process* 14(4):1255-1266. doi: 10.1109/TSA.2005.860835
- [20] Olsen PA, Gopinaht RA (2004) Modeling inverse covariance matrices by basis expansion, *IEEE Trans Speech Audio Process* 12(1):37-46. doi: 10.1109/TSA.2003.819943
- [21] Sun J, Kaban A (2008) A fast algorithm for robust mixtures in the presence of measurements errors, *IEEE Trans Neural Netw* 21(8):1206-1220. doi: 10.1109/TNN.2010.2048219
- [22] Verbeek JJ, Nunnink JRJ, Vlassis N (2006) Accelerated EM-based clustering of large data sets, *Data Min Knowl Disc* 13:291–307. doi: 10.1007/s10618-005-0033-3
- [23] Moore AW (1999) A very fast EM-based mixture model clustering using multiresolution kd-trees, *Adv Neural Inf Process Syst*, MIT Press Cambridge, MA, USA 11:543-549. ISBN: 0-262-11245-0
- [24] Ueda N, Nakano R, Ghahramani Z, Hinton GE (2000) Split and merge EM algorithm for improving Gaussian mixture density estimates. *J VLSI Signal Process Syst* 26(1/2):133-140. doi: 10.1023/A:1008155703044
-

-
- [25] Zhang Z, Chen C, Sun J, Chan KL (2003) EM algorithms for Gaussian mixtures with split-and-merge operation. *Pattern Recognit* 36(9):1973-1983. doi: 10.1016/S0031-3203(03)00059-1
- [26] Hershey JR, Olsen PA (2007) Approximating the Kullback Leibler divergence between Gaussian mixture models. *Proc ICASSP, Honolulu, HI 4:IV-317-IV-320*. doi: 10.1109/ICASSP.2007.366913
- [27] Montgomery DC, Runger GC (2010) *Applied Statistic and Probability for Engineers*, 5th Edition. John Wiley & Sons. ISBN-10: 0470053046, ISBN-13: 978-0470053041
- [28] Hageman LA, Young DM (2004) *Applied Iterative Methods (Dover Books on Mathematics)*. Dover Publications. ISBN-10: 048643477X, ISBN-13: 978-0486434773
- [29] Duda RO, Hart PE, Stork DG (2000) *Pattern Classification (2nd Edition)*. Wiley-Interscience. ISBN-10: 0471056693 ISBN-13: 978-0471056690
- [30] Blekas K, Lagaris IE (2007) Split-Merge Incremental LEarning (SMILE) of Mixture Models. *Lecture Notes in Computer Science (LNCS)*, No. 4669, 291-300. ISSN: 0302-9743
- [31] Kullback S (1997) *Information Theory and Statistic (Dover Books on Mathematics)*. Dover Publications. ISBN-10: 0486696847 ISBN-13: 978-0486696843
- [32] Webb AR (1999) *Statistical Pattern Recognition*. Defence Evaluation and Research Agency, Arnold, UK
- [33] Knill KM, Gales MJF, Young SJ (1999) State-Based Gaussian Selection in Large Vocabulary Continuous Speech Recognition Using HMMs. *IEEE Trans Speech Audio Process*, 7(2):152-161. doi: 10.1109/89.748120
- [34] Delic V (2007) A review of R&D of speech technologies in Serbian and their applications in western Balkan countries. Keynote lecture at 12th SPECOM (Speech and Computer), Moscow, Russia, pp 64-83
- [35] Delic V, Secujski M, Jakovljevic N, Janev M, Obradovic R, Pekar D (2010) *Speech Technologies for Serbian and Kindred South Slavic Languages*. *Adv Speech Recognit*, Noam Shabtai (Ed.), InTech. ISBN: 978-953-307-097-1
- [36] Kannan A, Ostendorf N, Rohlicek JR (1994) Maximum likelihood clustering of Gaussian mixtures for speech recognition. *IEEE Trans Speech Audio Process* 2(3):453-455. doi: 10.1109/89.294362
- [37] Young SJ, Odell JJ, Woodland PC (1994) Tree-based state tying for high accuracy state modeling. *Proc ARPA Workshop Hum Lang Technol*, pp 307-312. doi: 10.3115/1075812.1075885
-

- [38] Jakovljevic N (2005) Description of Training Procedure for AlfaNum Continuous Speech Recognition System. Proc EUROCON 2005, 2:1646-1649. ISBN: 1-4244-0049-X, doi: 10.1109/EURCON.2005.1630286
- [39] Jakovljevic N, Popovic B, Janev M, Pekar D (2011) The Impact of the Pitch on the Estimation of MFCC. Proc TELFOR 2011, pp 651-654, Belgrade, Serbia. ISBN: 978-1-4577-1499-3, doi: 10.1109/TELFOR.2011.6143631
- [40] European Telecommunications Standards Institute (2003) Speech Processing, Transmission and Quality Aspects (STQ); Distributed speech recognition; Front-end feature extraction algorithm; Compression algorithms. Technical standard ES 201 108, v1.1.3.