

НАСТАВНО-НАУЧНОМ ВЕЋУ

Предмет: Реферат о урађеној докторској дисертацији кандидата Бранка Марковића.

Одлуком бр. 955/3 од 13.12.2017. године, именовани смо за чланове Комисије за преглед, оцену и одбрану докторске дисертације кандидата Бранка Марковића под насловом

“Анализа обележја у говорном сигналу за потребе препознавања мултимодалног говора“

После прегледа достављене дисертације и других пратећих материјала и разговора са кандидатом, Комисија је сачинила следећи

РЕФЕРАТ

1. УВОД

1.1. Хронологија одобравања и израде дисертације

Бранко Марковић је пријавио тему за израду докторске дисертације 7.5.2015. године. Комисија за студије трећег степена је на својој седници дана 12.5.2015. године разматрала поднету пријаву теме докторске дисертације и свој предлог о оцени подобности теме и кандидата упутила Наставно-научном већу на усвајање. Наставно-научно веће је на својој седници дана 30.6.2015. (Одлука бр. 955/1) године именовало Комисију за оцену услова и прихватање теме ове докторске дисертације. Наставно-научно веће на седници одржаној дана 10.7.2015. године усвојило је поднети извештај Комисије (Одлука бр. 955/2). Веће научних области техничких наука својом одлуком број 61206-3775/2-15 од 15.9.2015. године дало је сагласност на предлог теме докторске дисертације под насловом: “Анализа обележја у говорном сигналу за потребе препознавања мултимодалног говора“.

Кандидат је дана 31.10.2017. године предао урађену докторску дисертацију на преглед и оцену. Комисија за студије трећег степена на својој седници одржаној дана 28.11.2017. године потврдила је испуњеност потребних услова за подношење предлога Наставно-научном већу Електротехничког факултета за формирање Комисије за преглед и оцену докторске дисертације. На основу тога Наставно-научно веће Факултета својом одлуком број 955/3 од 13.12.2017. године именовало је Комисију за преглед и оцену докторске дисертације у саставу који је наведен на крају овог извештаја.

1.2. Научна област дисертације

Ова докторска дисертација припада области Техничких наука – Електротехника, ужа научна област Техничка акустика. За ментора је одређен др Миомир Мијић, редовни професор. Он је

изабран у звање редовног професора за област Техничке акустике и сви његови публиковани радови који га квалификују за ментора су из те области.

1.3. Биографски подаци о кандидату

Бранко Марковић је рођен 3. јануара 1966. године у Сивчини, општина Ивањица. Основну школу је завршио у Слатини, општина Чачак, а гимназију у Чачку.

На Електротехничком факултету у Београду дипломирао је на смеру Телекомуникације 1992. године са просечном оценом 9.13, а на дипломском 10. Исте године уписао је постдипломске студије на смеру „Дигитални пренос информација“. На магистарским студијама положио је све испите са оценом 10. Након прекида од 10 година које је провео на раду у иностранству (Канада и САД) одбранио је 2004. године магистарску тезу под називом „Функција ‘Позивање говором’ у мобилној телефонији“, а ментор је био проф. др Слободан Јовичић.

Бранко Марковић је радну каријеру започео у Институту „Михајло Пупин“ у Београду где је у периоду 1992-1993. био ангажован на развоју мрежних уређаја мултиплекса и модема. Од 1993. до 2002. године радио је у Канади и САД-у у областима софтвера за телекомуникационе компаније: Bell Canada, Architel и Nortel Networks, на пројектима као што су Video Conferencing Manager, NYNEX OLP Project, AIMS, ASAP, Telezone, OMS и други.

Почев од 2004. године до данас Бранко Марковић ради као предавач на Високој школи техничких струковних студија у Чачки где је ангажован за групу предмета „Информациони системи и технологије“. Креирао је одређен број нових предмета као што су Рачунарске мреже, Интернет технологије (Основне студије), Софтверски алати, Мултимедијалне комуникације и Вишеслојна софтверска архитектура (Специјалитичке студије). Био је ментор већег броја дипломских и специјалистичких радова. Аутор је уџбеника/скрипти за одређен број предмета. Као аутор и коаутора објавио је више од 50 стручних и научних радова везаних за ове области.

2. ОПИС ДИСЕРТАЦИЈЕ

2.1. Садржај дисертације

Докторска дисертација Бранка Марковића је написана на 148 страна са 86 слика и 65 табела. На почетку се налазе насловна страна, кратак резиме на српском и енглеском језику, садржај као и списак слика и табела. Текст рада састоји се од седам поглавља, а на крају се налази преглед коришћене литературе. Поглавља су организована по следећем редоследу: 1 Увод, 2 Говор и врсте говора, 3 Векторска обележја, 4 Whi-Spe говорна база, 5 Поређење говорних узорака, 6 Експериментална мерења и 7 Закључак. На самом крају дисертације, као додатак, налазе се прилози: резултати на бази НММ метода, биографија и изјаве аутора.

2.2. Кратак приказ појединачних поглавља

Материја изложена у докторској дисертацији може се логички поделити на два дела. Први део се састоји од четири поглавља у којима је, после увода, изложена одговарајућа теорија неопходна за разумевање приказа самосталног рада кандидата и добијених резултата. То су у тези поглавља 2, 3, 4 и 5. У њима су описани говор и његове основне карактеристике, начин издвајања одговарајућих говорних обележја, коришћени метод за поређење говорних узорака као и говорна база која је снимљена за намену истраживања мултимодалног говора. Други део дисертације представља експериментална мерења за једанаест одабраних векторских обележја са одговарајућом анализом. То је уједно и најобимнији део дисертације. На крају се

налази закључак у коме су истакнути постигнути резултати, диприноси ове дисертације као и могући даљи правци истраживања.

У другом поглавља рада описана је природа говорног сигнала, а такође и различите врсте/модалитети говора. Детаљно су описани процеси генерисања и пријема говора, као и органи који у томе учествују. Истакнуте су основне кратеристике говора у временском и спектралном домену као и проблеми који настају услед утицаја маскирања и шума. Мултимодалном говору је посвећена посебна пажња са акцентом на шапат и нормални говор. Дато је и детаљно објашњење како настаје шапат (са аспекта функције вокалног тракта) као и акустичка и прецептивна обележја шапата, са посебним акцентом на померање првих форманата при артикулацији вокала, разликовања „доброг“ и „лошег“ шапата, утицај здравственог стања говорника на продукцију говора итд. Дат је и преглед досадашњих резултата који се односи на истраживања у овој области, а посебно на релацијама нормалан говор-шапат.

Треће поглавље се бави акустичким обележјима која су употребљена за репрезентовање говорних узорака нормалног говора и шапата. Коришћена су следећа обележја: LPCC (Linear Prediction Cepstral Coefficients), LFCC (Linear Frequency Cepstral Coefficients), MFCC (Mel Frequency Cepstral Coefficients), GFCC (Gammatone Filterbank Cepstral Coefficients), PLPCC (Perceptual Linear Prediction Cepstral Coefficients), RASTACC (RelAtive SpecTrA Cepstral Coefficients). На ова векторска обележја је примењена нормализација, а на поједина, где је то било могуће, примењен је и нелинеарни Teager Energy (TE) оператор. Употребом овог оператора добила су се и нова обележја типа TELFCC, TEMFCC, TEGFCC, TEPLPCC и TERASTACC. Утицаји нормализације и TE оператора су од кључне важности у овим истраживањима, па је у раду њима посвећена посебна пажња. За свако од једанаест акустичких обележја изложено је теоретски и приказано детаљно (блок дијаграмима) како се добијају одговарајући вектори који их репрезентују.

Четврто поглавље садржи детаљан опис говорне базе „Whi-Spe“ која је креирана за потребе ових истраживања и у којој се налазе одабране речи српског језика. Најпре је описан дизајн базе са објашњењем како је она конципирана, описана је структура говорника који су учествовали у снимању и форма и нотација како су одговарајући узорци складиштени. Потом је објашњен поступак снимања и карактеристике употребљене опреме. Приказани су и проблеми који су уочени током снимања, различити ефекти и грешке приликом аквизиције узорака. На крају су дати елементи говорне базе који представљају три подскупа, и то: речи основних боја, речи одабраних бројева и акустички балансиране речи, тј. дат је речник који је коришћен. Приказана је и одговарајућа IPA (International Phonetic Alphabet) нотација овог речника.

Пето поглавља, последње у првом делу рада, садржи детаљан приказ метода које су коришћене за поређење говорних узорака. Овај метод се базира на динамичком програмирању, и његово спровођење подразумева креирање вектора састављених од кепстралних коефицијената који репрезентују говорне узорке. Описан је начин добијања мере за поређење вектора као и векторска дистанца. Детаљно је објашњено поређење скупова вектора у процесу тражења оптималне стазе у 2D равни и на бази тога проналажење најбољег решења при препознавању. Метод који то омогућава је DTW (Dynamic Time Warping). Он је врло ефикасан за поређење говорних узорака када је број узорака ограничен и када је систем за аутоматско препознавање говора (АПГ) „зависан од говорника“. Представљена је детаљна анализа одговарајућег алгорита динамичког програмирања за добијање оптималне стазе као и начин имплементације DTW-а. Објашњена су и ограничења која се приликом употребе DTW-а могу применити.

Други део тезе у коме су приказани резултати истраживања кандидата почиње шестим поглављем, које представља главни део дисертације са добијеним резултатима. Бројне табеле

са детаљним приказом резултата које прате излагање издвојене су у прилозима на крају дисертације и са овим поглављем чине целину.

У шестом поглављу су на почетку објашњени типови вектора који се састоје од одговарајућих коефицијената. Коришћена су четири типа ових вектора и то:

- вектор од 12 кепстралних коефицијената без нормализације,
- вектор од 12 кепстралних коефицијената са нормализацијом,
- вектор од 24 коефицијента (12 кепстралних плус 12 делта) са нормализацијом и
- вектор од 36 коефицијената (12 кепстралних, 12 делта и 12 делта-делта) са нормализацијом.

У наставку поглавља су детаљно приказани резултати препознавања за следеће врсте векторских обележја: LPCC, LFCC, TELFCC, MFCC, TEMFCC, GFCC, TEGFCC, PLPCC, TEPLPCC, RASTACC и TERAACC. Резултати су дати за сваки од четири типа вектора и за сваки од четири сценарија и то: нормалан/нормалан, шапат/шапат, нормалан/шапат и шапат/нормалан. Примена нормализације и побољшање препознавања за неусаглашене сценарије посебно је истакнута. Наглашен је утицај Teager-овог енергетског оператора и показано је како се његова примена одражава на успешност препознавања. Сви резултати су дати табеларно и дијаграмима. Урађена је t-статистичка анализа са прорачуном средњих вредности и одговарајућих интервала поузданости. На крају овог поглавља извршена је глобална компарација свих векторских обележја и на бази коришћене нормализације и нелинеарног TE оператора изведени су одговарајући закључци.

Закључак даје сумарне резултате који су остварени овим радом, као и могућности њихове примене у системима за препознавање мултимодалног говора. Такође су назначени и правци за будућа истраживања у циљу добијања робуснијег система за препознавање ове врсте говора.

3. ОЦЕНА ДИСЕРТАЦИЈЕ

3.1. Савременост и оригиналност

Рад на овој докторској дисертацији иницирала је чињеница да, осим система за препознавање свакодневног (нормалног) говора који се детаљно истражује задњих педесетак година, фокус савремених истраживања се померио и на друге врсте говора, са посебним акцентом на шапат. Шапат се по својим акустичким особинама разликује од осталих врста говора (то јест нормалног, тихог, полугласног, викања, итд.). Чињенице да код шапата долази до померња фреквенција одређених форманата вокала, да је он по енергији често на нивоу шума, да се јавља турбулентно кретање ваздуха у одређеном делу вокалног тракта током његовог генерисања и слично разлог су потребе за додатним истраживањима. С друге стране, шапат се у људској комуникацији разуме са високим процентом препознавања. Због тога постоји изазов како одредити акустичка обележја која ће омогућити да и машина успешно препознаје шапат, који су то елементи на које треба обратити највише пажње и која врста додатних елемената предобраде (као што је нормализација, примена нелинеарних оператора и слично) може увећати успешност препознавања. Оригиналност теме је заснована на чињеници да је у фокусу читав спектар различитих акустичких обележја са различитим типовима вектора и при различитим сценаријима, њихово поређење као и креирање нових обележја, а све са циљем да се постигне већа успешности препознавања мултимодалног говора.

3.2. Осврт на референтну и коришћену литературу

Током рада на дисертацији кандидат је користио обимну литературу из области која је обрађена. Списак референци дат на крају дисертације садржи 111 наслова. У оквиру тог

броја налази се врло широк опсег публикација који укључује књиге, часописе, зборнике са домаћих и међународних конференција и web линкове. Велики део публикација је новијег датума, мада су укључени и сви најзначајнији историјски, може се рећи референтни наслови да би се употпунила слика о обрађиваној материји. У списку референци налазе се и радови у којима је кандидат аутор или коаутор, а који садрже оригиналне резултате директно произашле из рада на дисертацији.

3.3. Опис и адекватност примењених научних метода

Област и тема којом се бави дисертација условила је да се она заснива на следећим научним методама:

- развој алгоритама за добијање статичких и динамичких кепстралних коефицијената за класична обележја LPC, LFCC, MFCC, GFCC, PLP, RASTA са и без нормализације;
- развој алгоритама за добијање статичких и динамичких кепстралних коефицијената за обележја са нелинеарним Teager енергетским оператором и то TELFCC, TEMFCC, TGFCC, TEPLPCC и TERASTACC са и без нормализације;
- примена добијених алгоритама на говорну базу Whi-Spe, при чему је акценат био на четири сценарија говора: нормалан/нормалан, шапат/шапат, нормалан/шапат и шапат/нормалан;
- експериментално одређивање успешности препознавања за сва обележја коришћењем стандардног метода за препознавање (DTW);
- статистичка анализа добијених резултата коришћењем нивоа поузданости од 95% и на бази ње одређивање најповољнијих обележја.

На основу примењених метода формирана је градацијска представа успешности акустичких обележја за све типове вектора и све сценарије и одређене су најуспешније.

3.4. Применљивост остварених резултата

Резултати који су приказани у овој докторској дисертацији могу имати примену у више области. С једне стране, добијени резултати помажу да се изврши класификација широког спектра векторских обележја и на бази тога омогући адекватан увид и примена у креирању система за препознавање мултимодалног говора, а са друге стране дати су алгоритми за креирања нових векторских обележја (TEPLPCC и TERASTACC). Једно од њих (TERASTACC) је омогућило знатно повећање успешности препознавања говора, посебно код неусаглашених сценарија. Експериментално је доказано да комбинација нормализације и нелинеарног Teager енергетског оператора представља услов за креирање успешног система за препознавање мултимодалног говора.

3.5. Оцена достигнутих способности кандидата за самостални научни рад

Кандидат Бранко Марковић је кроз одабир и систематичан преглед актуелне литературе, кроз самостални експериментални рад на анализи мултимодалног говора, кроз теоријску анализу утицаја различитих акустичких обележја репрезентованих одговарајућим векторима, кроз снимање и формирање говорне базе али и кроз друге сегменте свог експерименталног и теоријског рада показао висок степен самосталности у истраживачком раду. Томе треба додати и чињеницу да је највећи део радова на међународним и националним скуповима самостално презентовао.

4. ОСТВАРЕНИ НАУЧНИ ДОПРИНОС

4.1. Приказ остварених научних доприноса

У овој докторској тези остварени су следећи доприноси који представљају унапређење постојећег знања у области препознавања мултимодалног говора:

- Креиран је велики број векторских обележја (њих једанаест) и извршена статистичка анализа успешности препознавања мултимодалног говора за свако од њих.
- Размотрени су различити сценарији (усаглашени и неусаглашени) и различити типови вектора (састављени од кепстралних, делта и делта-делта кепстарлних коефицијената) и њихов утицај на успешност препознавања.
- Доказано је (теоретски и експериментално) да нормализација директно утиче на препознавање мултимодалног говора са посебним акцентом на неусаглашене сценарије. Применом нормализације успешност препознавања је знатно побољшана.
- Показано је да нелинерани Teager енергетски оператор (који успешно апроксимира турбулентно кретања ваздуха у вокалном тракту током шаптања) даје добре резултате и омогућава побољшање препознавања мултимодалног говора са посебним акцентом на шапат.
- Примена нелинеарног Teager енергетског оператора омогућила је добијање нових векторских обележја. Предложени су нови алгоритми и добијена нова векторска обележја као што су TEPLPCC и TERASTACC.
- Показано је да са новим векторским обележјима (посебно са TERASTACC) резултати успешности препознавања добијају значајно побољшање.
- Експериментално је доказано да комбинација нормализације и TE оператора даје добре резултате. Дата је препорука да за препознавање шапата ова комбинација буде заступљена у одговарајућим системима.
- На бази поређење свих векторских обележја по свим сценаријима и по свим типовима вектора предложена је градијентска скала векторских обележја. Она представља основу за избор најпогоднијих обележја за препознавање мултимодалног говора.

4.2. Критичка анализа резултата истраживања

Набројани научни доприноси значајни су у домену препознавања мултимодалног говора кроз неколико аспеката:

- верификовање потребе коришћења нормализације приликом препознавања мултимодалног говора, посебно код неусаглашених сценарија (нормалан/шапат и шапат/нормалан);
- показивање оправданости коришћења нелинеарног Teager енергетског оператора за све случајеве препознавања говора, а посебно за шапат;
- увођење нових алгоритама (као што су TEPLPCC и TERASTACC) за издававање акустичких обележја и експериментални доказ да они омогућавају успешније препознавање, посебно за неусаглашене сценарије;
- анализу утицаја статичких и динамичких кепстралних обележја на препознавање говора;

- креирање градацијске скале по свим поменутиим векторским обележјима и на тај начин добијање јасније слике која обележја користити у одговарајућим сценаријима.

Успешност ових истраживања верификована је и кроз радове објављене у часописима, међународним и домаћим конференцијама.

4.3. Верификација научних доприноса

Научни доприноси докторске дисертације Бранка Марковића верификовани су у следећим радовима (наведени по М категоријама према Правилнику Министарства просвете и науке Србије):

Категорија М23:

1. **Branko R. Marković**, Jovan Galić, Miomir Mijić, “*Application of Teager Energy Operator on Linear and Mel Scales for Whispered Speech Recognition*”, Archives of Acoustics, 2018, Vol. 43, No. 1, ISSN 2300-262X, pp. 1-7 (in press) (IF=0.816)

Рад је прихваћен и у процесу је штампања

2. **Branko Marković**, Jovan Galić, Đorđe Grozdić, Slobodan T. Jovičić and Miomir Mijić, „*Whispered Speech Recognition Based on Gammatone Filter Cepstral Coefficients*“, Journal of Communication Technology and Electronics, 2017, Vol. 62, No. 11, ISSN 1555-6557, pp. 1255-1261 (IF=0.446)

doi:10.1134/S1064226917110134

3. Đorđe Grozdić, Slobodan Jovičić, Dragana Šumarac Pavlović, Jovan Galić, **Branko Marković**, “*Comparison of Cepstral Normalization Techniques in Whispered Speech Recognition*”, Advances in Electrical and Computer Engineering, 2017, Vol. 17, No. 1, ISSN 1582-7445, pp. 21-26 (IF=0.459)

doi: 10.4316/AECE.2017.01004

Категорија М14:

4. Đorđe T. Grozdić, Slobodan T. Jovičić, Jovan Galić and **Branko Marković**: “*Experiments in whisper recognition using neural networks*”, In: S.T. Jovičić, M. Subotić, M. Sovilj (Eds.): VERBAL COMMUNICATION QUALITY, Interdisciplinary Research, II; CUŽA, Belgrade, ISBN 978-86-81879-46-7, pp. 91-110, 2013.

Категорија М33:

5. **Branko R. Marković**, Goran Stevanović, Slobodan T. Jovičić, Milomir Mijić, Jovan Galić and *Đorđe T. Grozdić*, „*Recognition of Normal and Whispered Speech Based on RASTA Filtering and DTW Algorithm*”, Proceedings of 4th International Conference of Electrical, Electronic and Computer Engineering, IcETRAN 2017, ISBN 978-86-7466-692-0, AKI2.8.1-4, Kladovo, Serbia, June 05-08.
6. **Branko R. Marković**, Slobodan T. Jovičić, Miomir Mijić, Jovan Galić and Đorđe T. Grozdić, „*Recognition of Whispered Speech Based on PLP Features and DTW Algorithm*”, Proceedings of 3rd International Conference of Electrical, Electronic and Computer Engineering, IcETRAN 2016, ISBN 978-86-7466-618-0, pp. AKI1.3.1-5, Zlatibor, Serbia, June 2016,

7. **Branko R. Marković**, Slobodan T. Jovičić, Jovan Galić and Đorđe T. Grozdić, „*Recognition of the Multimodal Speech Based on GFCC Features*”, Proceedings of 2nd International Conference of Electrical, Electronic and Computer Engineering, IcETRAN 2015, ISBN 978-86-80509-71-6, pp. AKI1.3.1-5, Silver Lake, Serbia, June 8-11, 2015.
8. Đorđe Grozdić, Slobodan Jovičić, Jovan Galić, **Branko Marković**: „*Application of inverse filtering in enhancement of whisper recognition*“, Zbornik radova konferencije NEUREL 2014, ISBN 978-1-4799-5887-0, str. 157-161, Beograd, Srbija, 2014.
9. Jovan Galić, Slobodan T. Jovičić, Đorđe Grozdić, **Branko Marković**: „*Constrained Lexicon Speaker Dependent Recognition of Whispered Speech*“, Zbornik radova konferencije INDEL 2014, ISBN 978-99955-46-22-9, pp. 180-184, Banja Luka, Republika Srpska, 2014.
10. Jovan Galić, Slobodan T. Jovičić, Đorđe Grozdić and **Branko Marković**, „*HTK-Based Recognition of Whispered Speech*”, A. Ronzhin et al. (Eds.): SPECOM 2014, LNAI 8773, ISBN 978-3-319-11581-8, pp. 251–258, 2014, Springer International Publishing Switzerland 2014.
11. **Branko R. Marković** and Đorđe T. Grozdić, „*The LPCC-DTW Analysis for Whispered Speech Recognition*”, Proceedings of 1st International Conference of Electrical, Electronic and Computer Engineering, IcETRAN 2014, ISBN 978-86-80509-70-9, pp. AKI1.1.1-4, Vrnjačka Banja, Serbia, June 2-5, 2014.
12. Đorđe Grozdić, **Branko Marković**, Jovan Galić, Slobodan Jovičić, Darko Furundžić, „*Neural Network-Based recognition of whispered speech*”, Proceedings Speech and Language 2013, 4th International Conference on Fundamental and Applied Aspects of Speech and Language, ISBN 978-86-81879-44-3, pp. 223-229, Belgrade, October 25-26, 2013.
13. Jovan Galić, Slobodan Jovičić, Đorđe Grozdić, **Branko Marković**, „*The influence of feature vector selection on performance of automatic recognition of whispered speech*”, Proceedings Speech and Language 2013, 4th International Conference on Fundamental and Applied Aspects of Speech and Language, ISBN 978-86-81879-44-3, pp. 258-264, Belgrade, October 25-26, 2013.
14. **Branko Marković**, Jovan Galić, Đorđe Grozdić, Slobodan T. Jovičić, „*Application of DTW method for whispered speech recognition*”, Proceedings Speech and Language 2013, 4th International Conference on Fundamental and Applied Aspects of Speech and Language, ISBN 978-86-81879-44-3, pp. 308-315, Belgrade, October 25-26, 2013.
15. **Branko Marković**, Slobodan T. Jovičić, Jovan Galić and Đorđe Grozdić, „*Whispered Speech Database: Design, Processing and Application*“, I. Habernal and V. Matousek (Eds.): TSD 2013, LNAI 8082, pp. 591-598, 2013, Springer-Varlag Berlin Heidelberg 2013. ISBN 978-3-642-40584-6

Kategorija M53:

16. Đorđe T. Grozdić, **Branko Marković**, Jovan Galić, Slobodan T. Jovičić, „*Application of Neural Networks in Whispered Speech Recognition*“, TELFOR Journal, Vol. 5, No. 2, 2013, ISSN 1821-3251, pp. 103-106.
17. **Branko Marković**, "Upotreba skrivenih Markovljevih modela u prepoznavanju govora", Nauka Tehnika Bezbednost, Beograd, ISSN 0353-5517, br. 2(2002), str. 3-23.

Категорија М63:

18. Đorđe Grozdić, Dragana Šumarac Pavlović, Slobodan Jovičić, Jovan Galić, **Branko Marković** „*Komparacija tehnika normalizacije kepstralnih koeficijenata u automatskom prepoznavanju šapata*”, Zbornik radova 59. Konferencije ETRAN, ETRAN 2015, ISBN 978-86-80509-72-3, str. AK11.8.1-5, Srebrno jezero, Srbija 8-11. jun, 2015.
19. Jovan Galić, Milan Popović, **Branko Marković**, Đorđe T. Grozdić, Slobodan T. Jovičić, „*Primjena skrivenih Markovljevih modela u prepoznavanju govora u šapatu*“, Zbornik radova konferencije INFOTEH 2013, ISBN 978 99955-763-1-8, str. 387-390, Jahorina, Republika Srpska, 2013.
20. Đorđe Grozdić, **Branko Marković**, Jovan Galić, Slobodan Jovičić, „*Primena neuralnih mreža u prepoznavanju govora u šapatu*“, Zbornik radova 20tog Telekomunikacionog foruma TELFOR 2012, ISBN 978-1-4673-2982-8 str. 728-731, Beograd, 20-22. novembra 2012.
21. **Branko Marković**, Goran Luković, „*WiseWave 1.4 – softver za analizu uticaja prozorovanja na prepoznavanje izolovano izgovorenih reči*“, Zbornik radova 56. Konferencije za ETRAN, ISBN 978-86-80509-67-9, str. AK1.6-1-4, Zlatibor, 11-14. juna 2012.
22. Rajko Mitrović, Ivan Živanović, Ivan Radeljić, **Branko Marković**, „*Korišćenje asinhronog dinamičkog programiranja u rešavanju problema optimalne staze*“, Tehnika i informatika u obrazovanju, TIO 2012, 4. internacionalna konferencija, ISBN 978-86-7776-138-7, Zbornik, str. 193-199, Tehnički fakultet, Čačak, 1-3. juna 2012.
23. Gordana Marković, **Branko Marković**, „*Vizuelni DTW kao nastavno sredstvo za poređenje govornih uzoraka*“, Tehnika i informatika u obrazovanju, TIO '08, ISBN 978-86-7776-062-5, str. 409-415, Tehnički fakultet, Čačak, 9-11. maja.
24. **Бранко Р. Марковић**, „*WiseEdit 1.0 – софтверски пакет за аквизицију говорних узорака у облику wave фајлова*“, Zbornik radova XLVI za ETRAN, ISBN 86-80509-54-X, tom II, str. 388-391, Budva, 5-10. juna 2005.
25. **Branko R. Marković**, „*WiseDigit – Softver za prepoznavanje vezano izgovorenih cifara u uslovima šuma*“, Zbornik radova DOGS 2004, ASR.3, Sombor, 9-10. septembra 2004.

5. ЗАКЉУЧАК И ПРЕДЛОГ

На основу чињеница изложених у овом извештају Комисија је закључила да докторска дисертација Бранка Марковића под насловом “Анализа обележја у говорном сигналу за потребе препознавања мултимодалног говора“ испуњава све формалне и суштинске услове предвиђене Законом о високом образовању и прописима Универзитета у Београду и Електротехничког факултета.

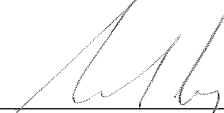
У дисертацији су истраживачки обрађене области анализе акустичких обележја коришћењем једанаест врста векторских обележја. Резултати истраживања су дали одговоре на постављене хипотезе и омогућили формирање градијентског погледа на могућности коришћења сваког од обележја по различитим сценаријима и за различите типове вектора. Дошло се и до нових решења за алгоритме предобраде која омогућавају знатно побољшање у процесу препознавања мултимодалног говора, а посебно када су у питању неусаглашени сценарији типа нормалан/шапат и шапат/нормалан. Један од главних закључака и доприноса

ове тезе је да се за препознавање ове врсте говора користе заједно нормализација и нелинеарни Teager енергетски оператор, чиме се достиже оптимум препознавања.

На основу свега изложеног Комисија предлаже Наставно-научном већу Електротехничког факултета да се докторска дисертација под називом “Анализа обележја у говорном сигналу за потребе препознавања мултимодалног говора“ кандидата Бранка Марковића прихвати, изложи на увид јавности и упуту на коначно усвајање Већу научних области техничких наука Универзитета у Београду.

У Београду, 24.1.2018. године

ЧЛАНОВИ КОМИСИЈЕ



др Миомир Мијић, редовни професор,
Универзитет у Београду – Електротехнички факултет



др Драгана Шумарац Павловић, ванредни професор
Универзитет у Београду – Електротехнички факултет



др Мишко Суботић, научни сарадник,
Центар за унапређење животних активности, Београд