

## ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА

## ИЗВЕШТАЈ О ОЦЕНИ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ

-обавезна садржина- свака рубрика мора бити попуњена

(сви подаци уписују се у одговарајућу рубрику, а назив и место рубрике не могу се мењати или изоставити)

I ПОДАЦИ О КОМИСИЈИ
<p>1. Датум и орган који је именовao комисију Решењем бр. 012-199/52-2018 од 28.02.2019. године, на основу Одлуке Научно-наставног већа Факултета техничких наука, а у складу са Статутом Факултета техничких наука, декан Факултета техничких наука проф. др Раде Дорословачки, именовao је Комисију за оцену и одбрану докторске дисертације.</p> <p>2. Састав комисије са назнаком имена и презимена сваког члана, звања, назива уже научне области за коју је изабран у звање, датума избора у звање и назив факултета, установе у којој је члан комисије запослен:</p> <ol style="list-style-type: none"> <li>1. <b>др Милан Сечујски</b>, ванредни професор, ужа област Телекомуникације и обрада сигнала, изабран у звање 11.03.2016, Факултет техничких наука, Нови Сад.</li> <li>2. <b>др Жељен Трповски</b>, ванредни професор, ужа област Телекомуникације и обрада сигнала, изабран у звање 11.06.2014, Факултет техничких наука, Нови Сад.</li> <li>3. <b>др Татјана Грбић</b>, редовни професор, ужа област Теоријска и примењена математика, изабрана у звање 19.02.2019, Факултет техничких наука, Нови Сад.</li> <li>4. <b>др Зоран Перић</b>, редовни професор, ужа област Телекомуникације, изабран у звање 24.11.2010, Електронски факултет, Ниш.</li> <li>5. <b>др Никша Јаковљевић</b>, доцент, ужа област Телекомуникације и обрада сигнала, изабран у звање 11.10.2014, Факултет техничких наука, Нови Сад.</li> <li>6. <b>др Владо Делић</b>, редовни професор, ужа област Телекомуникације и обрада сигнала, изабран у звање 28.03.2013, Факултет техничких наука, Нови Сад.</li> </ol>
II ПОДАЦИ О КАНДИДАТУ
<ol style="list-style-type: none"> <li>1. Име, име једног родитеља, презиме: <b>Синиша, Бранко, Сузић</b></li> <li>2. Датум рођења, општина, држава: <b>04. 10. 1988, Загреб, Хрватска</b></li> <li>3. Назив факултета, назив студијског програма дипломских академских студија – мастер и стечени стручни назив: <b>Факултет техничких наука, Универзитет у Новом Саду, Енергетика, електроника и телекомуникације, Мастер инжењер електротехнике и рачунарства</b></li> <li>4. Година уписа на докторске студије и назив студијског програма докторских студија <b>2012. година, Енергетика, електроника и телекомуникације</b></li> <li>5. Назив факултета, назив магистарске тезе, научна област и датум одбране: /</li> <li>6. Научна област из које је стечено академско звање магистра наука: /</li> </ol>
III НАСЛОВ ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:
<b>Параметарска синтеза експресивног говора</b>

#### **IV ПРЕГЛЕД ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:**

Навести кратак садржај са знаком броја страна, поглавља, слика, шема, графикона и сл.

У докторској дисертацији под називом „Параметарска синтеза експресивног говора” кандидата Синише Сузића приказани су поступци за синтезу експресивног говора односно говора у одређеном стилу коришћењем параметарских приступа. Дисертација је написана на 96 страница (са пратећом документацијом дисертација садржи укупно 121 страницу). Садржај је подељен у 8 поглавља. Дисертација садржи 41 слику, 10 табела и 136 научних референци. На почетку тезе дати су: наслов, кључна документацијска информација на српском и на енглеском језику, сажетак на српском и на енглеском језику, садржај рада, попис слика, попис табела, списак скрећеница, као и захвалница.

Дисертација обухвата следећих 8 поглавља

1. Увод
2. Увод у синтезу експресивног говора
3. ХММ синтеза
4. ДНН синтеза
5. Синтеза експресивног говора
6. ДНН приступи за синтезу експресивног говора коришћењем мале количине експресивног материјала
7. Трансплантација стилова (емоција)
8. Закључак

#### **V ВРЕДНОВАЊЕ ПОЈЕДИНИХ ДЕЛОВА ДОКТОРСКЕ ДИСЕРТАЦИЈЕ:**

У **уводном поглављу** дефинисан је предмет истраживања. Наведена је дефиниција синтезе говора на основу текста и дати су неки примери употребе ове технологије. Назначена је потреба за генерисањем говора у различитим стиливима и емоцијама. У складу са овом потребом дефинисани су циљеви истраживања: развој приступа за синтезу експресивног говора са посебним освртом на синтезу базирану на употреби дубоких неуронских мрежа и развој поступка који омогућава трансплантацију говорног стила односно емоције. На крају овог поглавља наведен је садржај дисертације.

*Комисија сматра да су предмет и циљеви истраживања дефинисано јасно и концизно.*

У **другом поглављу** дат је преглед основних приступа у синтези говора. Набројани су и поступци који се користе у оцени квалитета синтетизованог говора. У другом делу овог поглавља наведено је неколико дефиниција појма експресивног говора из литературе и јасно дефинисано значење овог појма коришћено у дисертацији. Дат је и преглед досадашњих резултата у синтези експресивног говора.

*Комисија сматра да садржај овог поглавља омогућава разумевање теоријских основа истраживања. Наведене су предности и мане свих приступа и назначено због чега су у истраживању коришћени параметарски приступи. Опис експресивног говора је свеобухватан и појам је сагледан из више различитих аспеката. Литература коришћена у прегледу достигнућа у синтези експресивног говора је актуелна и релевантна.*

У **трећем поглављу** описана је синтеза говора коришћењем скривених Марковљевих модела. Прво је дат општи опис скривених Марковљевих модела, а потом су описане специфичности примене скривених Марковљевих модела у синтези говора. Обука модела коришћених у синтези на бази скривених Марковљевих модела описана је у последњем делу овог одељка.

*Комисија сматра да је примена скривених Марковљевих модела у синтези говора изложена на јасан и адекватан начин. Садржај овог поглавља неопходан је за свеобухватно разумевање параметарских приступа.*

У **четвртном поглављу** описана је синтеза говора коришћењем дубоких неуронских мрежа. Прво је дат опис функционисања дубоких неуронских мрежа, а потом је описана њихова примена у синтези говора. У последњем делу овог поглавља дато је поређење квалитета говора

синтетизованог коришћењем скривених Марковљевих модела и дубоких неуронских мрежа на примеру неутралног говора за српски језик. Објективним и субјективним тестовима показано је да се употребом синтетизатора на бази дубоких неуронских мрежа добија говор бољег квалитета.

*Комисија сматра да је примена дубоких неуронских мрежа у синтези говора изложена на јасан и адекватан начин. Заједно са садржајем претходног поглавља даје свеобухватан преглед параметарских приступа. Описани експерименти спроведени су коректно, а резултати интерпретирани објективно.*

У **петом поглављу** предложена је метода кодова стила за синтезу експресивног говора коришћењем параметарских приступа. Описана метода примењена је и у синтези на бази дубоких неуронских мрежа, као и у синтези на бази скривених Марковљевих модела. Дато је поређење изражености емоције у синтетизованом говору добијеном коришћењем ова два приступа. Поређење је дато за пример синтезе једног додатног стила поред неутралног. Показано је да се синтезом на бази дубоких неуронских мрежа постиже боља израженост емоције у генерисаном говору него у случају синтезе на бази скривених Марковљевих модела. Такође је показано да мала количина експресивног материјала за обуку не утиче на значајну деградацију квалитета експресивног говора у односу на говор генерисан у неутралном стилу. У последњем делу овог одељка дата је детаљнија анализа перформанси методе кодова стила у синтези на бази дубоких неуронских мрежа поређењем система који истовремено могу да генеришу више различитих стилова са системима који генеришу говор само у једном од датих стилова. Субјективним и објективним тестовима показано је да не постоји значајна разлика у квалитету добијеног говора коришћењем оба начина синтезе.

*Комисија сматра да је предложени приступ за синтезу експресивног говора описан на јасан и адекватан начин. Предложена метода успешно је примењена у синтези експресивног говора. Описани експерименти спроведени су коректно, а резултати интерпретирани објективно. Пошто је предност дубоких неуронских мрежа потврђена и у случају синтезе експресивног говора, оправдано је коришћење дубоких неуронских мрежа у даљим експериментима.*

У **шестом поглављу** предложене су још две нове методе за синтезу експресивног говора коришћењем дубоких неуронских мрежа: метода додатне обуке и метода базирана на архитектури са дељеним скривеним слојевима. Квалитет експресивног говора добијеног овим приступима поређен је са говором добијеним методом кодова стила на примеру синтезе три говорна стила. Субјективним и објективним тестовима показано је да два предложена приступа постижу слабији квалитет синтетизованог експресивног говора.

*Комисија сматра су два нова предложена приступа за синтезу експресивног говора описана на јасан и адекватан начин. Предложени приступи успешно су примењени у синтези експресивног говора. Описани експерименти спроведени су коректно, а резултати интерпретирани објективно. Иако се коришћењем две нове предложене методе не добија експресивни говор истог квалитета као у случају коришћења методе кодова стила, предложене методе могу бити значајне јер омогућују брже додавање новог стила у већ постојећи систем за синтезу говора.*

У **седмом поглављу** описана је нова метода за трансплантацију стилова базирана на архитектури са дељеним скривеним слојевима. Метода је тестирана на примеру трансплантације стила (емоције) између два говорника. Добијени резултати су поређени са једном референтном методом из литературе. Поређење је вршено и за тачност препознавања стилова као и за укупни квалитет реченица са трансплантираним стилем. Показано је да предложена метода постиже боље резултате у трансплантацији емоција од референтног система и у поређењу квалитета говора, као и у случају препознавања емоција. Квалитет трансплантираног говора је оцењен лошије него квалитет обученог говора. Тачност препознавања стила лошија је код трансплантираног говора (ово важи и за предложени и за референтни систем).

*Комисија сматра је предложени приступ за трансплантацију експресивног говора описан на јасан и адекватан начин. Описани експерименти спроведени су коректно, а резултати интерпретирани објективно. Поређењем са референтним приступом из литературе показана је*

успешност предложене методе.

У осмом поглављу сумирани су резултати истраживања и назначени су даљи правци истраживања.

Комисија сматра да су закључци донети на бази изложених резултата и да потврђују значај предложених приступа у синтези експресивног говора коришћењем параметарских метода. Предложени наредни кораци у истраживању су логични и мотивисани су закључцима до којих се дошло током тренутног истраживања.

**Литература** садржи 136 прегледно систематизованих библиографских навода.

Комисија сматра да је литература адекватно одабрана, правилно коришћена и да одговара теми дисертације.

## VI СПИСАК НАУЧНИХ И СТРУЧНИХ РАДОВА КОЈИ СУ ОБЈАВЉЕНИ ИЛИ ПРИХВАЋЕНИ ЗА ОБЈАВЉИВАЊЕ НА ОСНОВУ РЕЗУЛТАТА ИСТРАЖИВАЊА У ОКВИРУ РАДА НА ДОКТОРСКОЈ ДИСЕРТАЦИЈИ

Таксативно навести називе радова, где и када су објављени. Прво навести најмање један рад објављен или прихваћен за објављивање у часопису са ISI листе односно са листе министарства надлежног за науку када су у питању друштвено-хуманистичке науке или радове који могу заменити овај услов до 01. јануара 2012. године. У случају радова прихваћених за објављивање, таксативно навести називе радова, где и када ће бити објављени и приложити потврду о томе.

### Рад у међународном часопису (M23)

1. **Siniša Suzić**, Tijana Delić, Darko Pekar, Vlado Delić, Milan Sečujski, "Style Transplantation in Neural Network Based Speech Synthesis", *Acta Polytechnica Hungarica*, 2019, ISSN:1785-8860, accepted for publication

### Рад у часопису националног значаја (M52)

1. **Siniša Suzić**, Tijana Delić, Stevan Ostrogonac, Simona Đurić, Darko Pekar, "Style-Code Method for Multi-Style Parametric Text-To-Speech Synthesis", *SPIIRAS Proceedings*, Issue No. 5(60), 2018, pp. 216-240, ISSN 2078-9181 (print), ISSN 2078-9599 (online), DOI: 10.15622/sp.60.8
2. Tijana Delić, Milan Sečujski, **Siniša Suzić**, "A review of Serbian parametric speech synthesis based on deep neural networks", *Telfor Journal*, vol. 9, no. 1, 2017, pp. 32-37, ISSN: 1821-3251, DOI: 10.5937/telfor1701032D

### Саопштење са међународног скупа штампано у целини (M33)

1. **Siniša Suzić**, Tijana Delić, Vladimir Jovanović, Milan Sečujski, Darko Pekar, Vlado Delić, "A comparison of multi-style DNN-based TTS approaches using small datasets", *13th International Conference on Electromechanics and Robotics "Zavalishin's Readings (ER(ZR)-2018)*, Saint Petersburg, Russia, April 18-21, 2018, pages 6, DOI: 10.1051/mateconf/201816103005
2. Tijana Delić, **Siniša Suzić**, Milan Sečujski, Darko Pekar, "Rapid Development of New TTS Voices by Neural Network Adaptation", *17th International Symposium INFOTEH-JAHORINA*, Jahorina, Bosnia and Hercegovina, March 21-23, 2018, pp. 1-6, DOI: 10.1109/INFOTEH.2018.8345518
3. **Siniša Suzić**, Tijana Delić, Darko Pekar, Vladimir Ostojić, „Novel alignment method for DNN TTS training using HMM synthesis models“, *IEEE 15th International Symposium on Intelligent Systems and Informatics, SISY 2017*, Subotica, Serbia, September 14-16, 2017, pp. 271-276

### Саопштење са скупа националног значаја штампано у целини (M63)

1. Tijana Delić, **Siniša Suzić**, Stevan Ostrogonac, Simona Đurić, Darko Pekar, "Multi-style Statistical Parametric TTS", *Zbornik radova konferencije Digitalna obrada govora i slike (DOGS 2017)*, Novi Sad, Srbija, November 22-25, 2017, pp. 5-8, ISBN: 978-86-7892-993-9

## **VII ЗАКЉУЧЦИ ОДНОСНО РЕЗУЛТАТИ ИСТРАЖИВАЊА**

У дисертацији су приказане методе које у оквиру синтезе базиране на параметарским приступима омогућавају генерисање експресивног говора, односно говора у различитим стиливима. Показано је да се синтезом на бази дубоких неуронских мрежа постиже бољи квалитет говора, као и да је израженост емоције (стила) боља, него у случају синтезе на бази скривених Марковљевих модела. Симултано моделовање више стилова истовремено не доводи до деградације квалитета експресивног говора у односу на случај креирања синтетизатора са само једним говорним стилем (осим неутралног). Квалитет говора у генерисаним стиливима није једнак и зависи од особина говорног стила у бази за обуку.

Поред основног приступа са синтезу експресивног говора који користи кодове стила, у синтези на бази дубоких неуронских мрежа предложена су и два додатна приступа: додатна обука мреже и архитектура са дељеним скривеним слојевима. Квалитет експресивног говора добијеног овим приступима мањи је него квалитет говора добијен основним приступом, али имају одређене предности у брзини реализације.

Метода за трансплантацију емоција базирана на архитектури са дељеним скривеним слојевима постиже боље резултате у трансплантацији емоција у поређењу са референтним приступом из литературе који је базиран на додатним улазима, што представља један од најзначајнијих доприноса ове дисертације. Квалитет трансплантираног говорног стила свакако је мањи него квалитет говора генерисаног у стилу који постоји у бази.

Резултати препознавања емоција на оригиналним реченицама приближни су резултатима који се добијају у синтези за стилове који постоје у бази. Тачност препознавања транспланитраних стилова је мања.

Сви предложени приступи тестирани су коришћењем ограничене количине експресивног говора, чиме је показано да експресивни делови говорне базе могу бити значајно мањи него део са неутралним говором, а да квалитет синтетизованог експресивног говора не буде смањен, чиме је потврђена једна од основних хипотеза ове тезе и омогућује широку примену резултата ових истраживања.

## **VIII ОЦЕНА НАЧИНА ПРИКАЗА И ТУМАЧЕЊА РЕЗУЛТАТА ИСТРАЖИВАЊА**

Експлицитно навести позитивну или негативну оцену начина приказа и тумачења резултата истраживања.

На основу детаљног увида у садржај докторске дисертације комисија сматра да је кандидат јасно и прегледно приказао резултате истраживања. Сви добијени резултати представљени су помоћу одговарајућих табела и графикана. Интерпретација резултата је објективна и коректна.

Дисертација је проверена у софтверу за детекцију плагијаризма **iThenticate**. Извештај о подударности је показао да је дисертација оригинално ауторско дело кандидата (Similarity Index 3%).

У складу са наведеним, Комисија **ПОЗИТИВНО** оцењује начин на који су резултати приказани и тумачени.

## **IX КОНАЧНА ОЦЕНА ДОКТОРСKE ДИСЕРТАЦИЈЕ:**

Експлицитно навести да ли дисертација јесте или није написана у складу са наведеним образложењем, као и да ли она садржи или не садржи све битне елементе. Дати јасне, прецизне и концизне одговоре на 3. и 4. питање:

1.	<p>Да ли је дисертација написана у складу са образложењем наведеним у пријави теме</p> <p><i><b>ДА</b>, дисертација је у потпуности написана у складу са образложењем наведеним у пријави теме.</i></p>
2.	<p>Да ли дисертација садржи све битне елементе</p> <p><i><b>ДА</b>, дисертација садржи све битне елементе који се захтевају по Статуту Факултета техничких наука и Универзитета у Новом Саду, као и Закона о високом образовању.</i></p>
3.	<p>По чему је дисертација оригиналан допринос науци</p> <p><i>У дисертацији су предложене три методе за синтезу експресивног говора коришћењем параметарских приступа: метода кодова стила, метода додатне обуке мреже и метода са дељеним скривеним слојевима. Експериментално је показано да су сви предложени приступи успешни у изражавању стила у синтетизованом говору и направљено је њихово поређење.</i></p> <p><i>Такође је предложена и нова метода за поступке генерисања говорног стила који није присутан у бази једног говорника, али постоји у бази другог говорника – трансплантација стила. Ова метода заснована је на архитектури са дељеним скривеним слојевима и експериментално је показана њена предност над референтим поступком предложеним у литератури.</i></p>
4.	<p>Недостаци дисертације и њихов утицај на резултат истраживања</p> <p><i>У дисертацији <b>нису</b> уочени недостаци који би утицали на резултате истраживања</i></p>
<b>X</b>	<b>ПРЕДЛОГ:</b>
	На основу укупне оцене дисертације, комисија предлаже:
	да се докторска дисертација кандидата <b>Синише Сузића</b> под насловом „ <b>Параметарска синтеза експресивног говора</b> ” <b>прихвати</b> , а кандидату <b>одобри одбрана</b> докторске дисертације.

ПОТПИСИ ЧЛАНОВА КОМИСИЈЕ

У Новом Саду, 12. 04. 2018.

\_\_\_\_\_

др Милан Сечујски, ванредни професор  
Факултет техничких наука, Нови Сад, председник комисије.

\_\_\_\_\_

др Жељен Трповски, ванредни професор,  
Факултет техничких наука, Нови Сад, члан комисије.

---

др Татјана Грбић, ванредни професор,  
Факултет техничких наука, Нови Сад, члан комисије.

---

др Никша Јаковљевић, доцент,  
Факултет техничких наука, Нови Сад, члан комисије.

---

др Зоран Перић, редовни професор,  
Електронски факултет, Ниш, члан комисије.

---

др Владо Делић, редовни професор,  
Факултет техничких наука, Нови Сад, ментор.

НАПОМЕНА: Члан комисије који не жели да потпише извештај јер се не слаже са мишљењем већине чланова комисије, дужан је да унесе у извештај образложење односно разлоге због којих не жели да потпише извештај.