

TEMA T1.2: KVALITET VISOKOG OBRAZOVANJA (Digitalne tehnologije učenja, Projekat Ministarstva „Razvoj visokog obrazovanja“, ...)

Paper No.T1.2-1
08329

CENTRALNA AUDIO-BIBLIOTEKA UNIVERZITETA U NOVOM SADU

Vlado Delić¹, Dragiša Mišković², Siniša Suzić³, Tijana Delić⁴,
Branislav Popović⁵, Nikša Jakovljević⁶, Milan Sečujski⁷

^{1,2,3,4,5,6,7} Univerzitet u Novom Sadu, Fakultet tehničkih nauka, Novi Sad, Srbija

¹vlado.delic@uns.ac.rs

Kratak sadržaj: U radu je predstavljen projekat Centralna audio-biblioteka Univerziteta u Novom Sadu (CABUNS) čiji je cilj da se omogući automatizovano kreiranje audio-izdanja udžbenika, prezentacija i drugih nastavnih materijala pomoću nove tehnologije za sintetizovanje govora na osnovu teksta na srpskom jeziku. Opisani su arhitektura i funkcionalnosti realizovanog sistema, kako sa strane nastavnika i asistenata koji postavljaju nastavne materijale na CABUNS server, tako i studenata kao korisnika koji preuzimaju audio-izdanja i slušaju ih (i gledaju) preko svojih računara ili telefona. Na prezentaciji rada će biti prikazani primeri prvih audio-izdanja udžbenika i Power Point prezentacija iz predmeta Akustika i audio-tehnika. Analizirane su prednosti i nedostaci ove nove digitalne tehnologije učenja koja ima perspektive da značajno doprinese kvalitetu visokog obrazovanja, ali i svih drugih nivoa obrazovanja. Predstavljeni su i najnoviji rezultati u razvoju sintetizatora govora koji omogućuju konverziju karakteristika glasa govornika tako da će u narednom periodu biti omogućena produkcija audio-izdanja sa glasom nastavnika koji su napisali date udžbenike i/ili održali predavanja uz date prezentacije.

Glavne reči: Audio-biblioteka, audio-izdanja udžbenika, nove digitalne tehnologije učenja, sinteza govora na osnovu teksta, duboke neuronske mreže, konverzija govornika

CENTRAL AUDIO LIBRARY OF THE UNIVERSITY OF NOVI SAD

Abstract: The paper presents the project Central Audio-Library of the University of Novi Sad (CABUNS), aimed at automated creation of audio-editions of textbooks, presentation and other course material using the new technology of text-to-speech synthesis in the Serbian language. The paper describes the architecture and the features of the developed system, from the points of view of both teachers and assistants who upload course material to the CABUNS server as well as students who download audio-editions and listen to them (and view them) using their computers and mobile telephones. The examples of the first audio-editions of textbooks and PowerPoint presentations related to the course Acoustics and Audio Engineering are presented. The paper also analyzes the advantages and drawbacks of this new learning technology, which has a potential to greatly contribute to the quality of higher education, but also to education at other levels. The paper also presents the most recent results in the development of text-to-speech, enabling speaker conversion, which means that very soon it will be possible to produce an audio-edition in the voice of the lecturer who has actually written the textbook and/or delivered a lecture using the presentation.

Key Words: Audio-library, audio-editions of textbooks, new digital learning technologies, text-to-speech synthesis, deep neural networks, speaker conversion

1. UVOD

Sve brži tempo života i potreba za efikasnim iskorišćenjem vremenskih resursa doveli su do popularizacije audio knjiga. Ova tehnologija omogućava „čitanje“ knjige bez upotrebe ruku i očiju, npr. u javnom prevozu, uz odmor ili u šetnji. Iako su prvobitne audio knjige sadržavale samo književna dela, vremenom se javila potreba da se i stručna literatura učini dostupnom u audio formatu. Arhitektura jednog takvog sistema opisana je u radu [1].

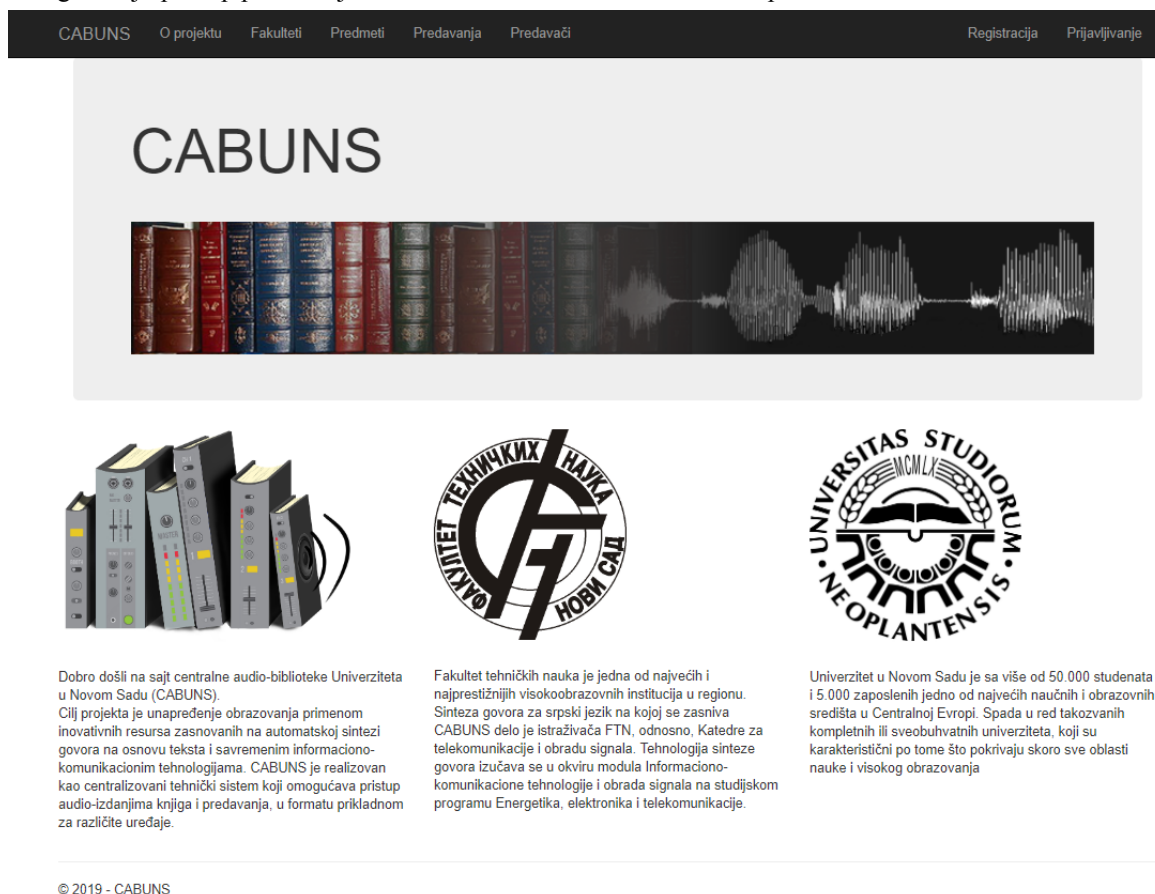
Proces snimanja kvalitetnih audio knjiga čiji sadržaj čitaju profesionalni govornici je vremenski i finansijski zahtevan. Stoga se korišćenje sinteze govora na osnovu teksta (TTS) nametnulo kao rešenje. Međutim, konkatentativni pristup sintezi, koji je dugo smatran za najbolji, ipak je imao relativno malu prijatnost slušanja generisanog govora zbog monotonosti intonacije i artefakata. Razvoj sinteze govora na osnovu teksta u poslednjih nekoliko godina omogućio je ne samo poboljšanje prirodosti sintetizovanog govora, nego i promenu glasa govornika kao i stila govora koristeći relativno male govorne baze za obuku modela [2-6].

Centralna audio biblioteka Univerziteta u Novom Sadu je projekat u kom se korišćenjem audio knjiga želi unaprediti i olakšati visoko obrazovanje. S obzirom da se nauka razvija velikom brzinom, za mnoge predmete na fakultetima je teško objaviti knjigu sa svim relevantnim informacijama, te se učenje često svodi na korišćenje prezentacija ili skripti koje često nisu dovoljno informativne. U okviru ovog projekta, profesorima će biti omogućeno da svoje prezentacije dopune tekstom koji i inače pričaju na predavanjima, te da imaju zvaničnu literaturu iz koje će studenti moći da uče, a koju je izuzetno jednostavno izmeniti u svakom trenutku.

U narednom poglavlju opisan je CABUNS sistem i način njegovog korišćenja. Potom su analizirane njegove prednosti i mane, a na kraju je dat pregled zaključaka i pravci daljeg razvoja i unapređenja sistema.

2. OPIS SISTEMA CABUNS I NAČIN NJEGOVOG KORIŠĆENJA

Slika 1 prikazuje početnu stranu portala CABUNS. Pored osnovnih informacija o projektu, ponuđeni su linkovi koji omogućavaju pristup predavanjima u skladu sa izabranim fakultetom i predmetom.



Slika 5. Početna strana portala CABUNS.

U okviru strane **Predavanja** moguće je dodavanje novih sadržaja preko dva podržana formata iz paketa *Microsoft Office*: *PowerPoint* prezentacije i *Word* dokumenta. Slika 2 prikazuje izgled strane za postavljanje predavanja. Ova mogućnost je ograničena na registrovane korisnike – profesore Univerziteta u Novom Sadu.

Pošto obrada datoteka sa predavanjima može biti vremenski zahtevna, celokupan proces parsiranja i generisanja govora se ne obavlja u realnom vremenu, nego kao pozadinski proces. U okviru ovog procesa se vrši:

- parsiranje dokumenta kako bi se izdvojile strane i tekst koji će se slati na sintezu govora
- generisanje vizuelnog prikaza za različite uređaje (personalne i tablet računare, mobilne telefone).

Na kraju ovog procesa koji može da potraje od nekoliko minuta do nekoliko sati, profesor ili asistent koji je dodao novu nastavnu jedinicu dobija email od CABUNS servera da je audio-izdanje spremno za proveru. Nakon što nastavnik/asistent pregleda audio-izdanje (Slika 3), ako je zadovoljan omogućuje pristup ovlašćenim studentima. Ukoliko nije zadovoljan ili želi nešto da doda ili promeni, uvek mu je na raspolaganju mogućnost da postavi revidiranu verziju predavanja na server, i uskoro će biti generisana ažurna verzija audio-izdanja tih predavanja.

U radu [1] preliminarno je predstavljena primenjena arhitektura i detalji vezani za kompletnu obradu datoteka sa predavanjima. Važno je napomenuti da se kod prezentacija (*PowerPoint*), audio-zapis generiše na osnovu sadržaja beleški dodatih svakom slajdu (*notes* ispod slajda). Kod dokumenata (*Word*) se izdvaja celokupni tekst na pojedinim stranicama dokumenta i na osnovu njega se sintetizuje govor. U oba slučaja, vizuelni sadržaj se formira konverzijom datoteka u pdf formu radi prikaza na raznim uređajima uz, naravno, mogućnost preslušavanja datog teksta – preslušavanje date stranice dokumenta se pokreće preko standardnog grafičkog prikaza plejera.

Jedan od parametara koji se unosi prilikom postavljanja predavanja je i identifikator jezika. Sistem trenutno omogućava generisanje govora na osnovu teksta samo za srpski, hrvatski i engleski jezik. Sintetizator za srpski se zasniva na postupcima parametarske sinteze govora primenom neuronskih mreža [2]. Ovo omogućava da audio-sadržaji u okviru predavanja sadrže govor koji je prijatniji za slušanje, sa promenljivom intonacijom u cilju naglašavanja pojedinih delova, sa različitim bojama glasa i drugim karakteristikama prirodnog govora. Najnoviji rezultati istraživanja omogućuju konverziju sintetizovanog glasa tako da on može da liči na glas profesora koji je postavio data predavanja i to je moguće napraviti na osnovu njegovog/njenog govornog uzorka od samo 10-tak minuta [3-5].

Predavanje

Predavač:

Naziv nastavne jedinice:

Predmet:

Jezik:

Ishodi obrazovanja:

Studenti će biti u stanju da objasne specifičnosti zvuka u zatvorenom prostoru, da izmere vreme reverberacije i proračunaju apsorpciju prostorije, kao i potrebnu snagu ozvučenja za željeni nivo zvuka. Moći će da ocene akustičke karakteristike i predlože akustičku obradu prostorije.

Datum postavljanja:

Fajl:

[Povratak](#)

© 2019 - CABUNS

Slika 6. Izgled strane za postavljanje predavanja. Podržani su formati PowerPoint i Word.

Akustika - analogije

2eb04f49_page_7.pdf 1 / 1

ekvivalentna električna šema rezonatora prikazana na slici 5.7b. Ovo je rezonantno kolo čija je frekvencija rezonanse data relacijom:

$$f_r = \frac{1}{2 \cdot \pi \cdot m_a \cdot C_a}$$

Akustička impedansa kola je minimalna na rezonansi i iznosi $Z_0 = R_0$.

a) Slika 5.8. Filtar nepropusnih opsega a) izgled akustičkog sistema, b) ekvivalentna šema

Jedan uprošćen oblik filtra nepropusnika opsega frekvencija sa rezonantnim kolom u paralelnoj grani prikazan je na slici 5.8a. Protoci akustičkih masa m_{a1} i m_{a2} su q_1 i q_2 . Protok komore i ulazne cevi u komoru je isti i jednak je razlici ulaznog i izlaznog protoka, što znači da iznosi $q_1 - q_2$. Idući sleva nadesno ulazni protok se deli na protok komore sa ulaznom cev, $q_1 - q_2$, i izlazni protok q_2 . Ako se zna da kapacitivnost komore mora jednim krajem biti uzemljena onda je lako doći do ekvivalentne električne šeme prikazane na slici 5.8b. Ovakav filter se koristi kod ventilacionih sistema kada u kanalima treba oslabiti osnovnu frekvenciju buke ventilatora.

1.7 Primeri mehaničkih sistema

U tabeli 5.2 prikazani su prosti slučajevi mehaničkih sistema takve konfiguracije da svi elementi sistema imaju istu brzinu ili su izloženi dejstvu iste sile. Kod složenijih mehaničkih sistema postupak oko dobijanja ekvivalentne šeme nije uvek jednostavan. Pri tome treba voditi računa da zbir brzina u svim čvorovima kola bude jednak nuli, kao i da se kolo na kraju zatvori.

Slika 7. Primer jedne strane postavljenog predavanja. U donjem delu strane se nalaze još komande za kontrolu audio reprodukcije i komande za navigaciju unutar predavanja.

3. PREDNOSTI I MANE CABUNS-A

Novo vreme traži da nastavni materijali balansiraju između raspoloživog vremena i narastajućeg obima znanja. Ovo je posebno izraženo u inženjerskim disciplinama jer se tempo tehnološkog razvoja značajno ubrzava. Srećom, nove tehnologije omogućuju i razvoj novih metoda predavanja i učenja. U ovom radu je reč o primeni ICT i TTS tehnologija za automatizovano generisanje audio-izdanja nastavnih materijala.

Audio-izdanja omogućuju studentima da preslušavaju lekcije čak i dok se odmaraju, šetaju, putuju. Pa i kada aktivno uče, audio-biblioteka im omogućuje da preslušavanjem umesto čitanja iskoriste vreme za razgledanje detalja na ilustracijama umesto da gledaju u tekst dok čitaju. Za osobe koje inače sporije čitaju, ili im misli često odlutaju, audio-biblioteka održava tempo čitanja i nivo koncentracije. Audio-knjige su izuzetno važne i za osobe sa invaliditetom (oštećenjem vida, disleksijom ili fizičkim hendikepom koji otežava korišćenje knjige ili računara). Klasična produkcija audio-izdanja oslanja se na govornike koji čitaju tekstove, što je vremenski zahtevno i podložno greškama. Mnogo je efikasnije kreirati audio-izdanje automatski, pomoću sintetizatora govora.

U nedostatke CABUNS koncepta treba ubrojati nekoliko problema za koja se rešenja još uvek traže. Naime, TTS pretvara u govor samo napisani tekst, ali ne i slike, formule i tabele. Ovo je naravno problem samo kod kompletnih audio-izdanja koja se preslušavaju bez gledanja u ekran uređaja. Za ove probleme je za sada predviđeno da nastavnici sastave tekstualni opis slika, formula i tabela – onako kako ih opisuju dok ih prikazuju u toku predavanja. Npr. niko ne čita celu tabelu, red po red, nego izdvoji nekoliko detalja i govori o njima.

Audio-biblioteka pruža studentima novi način učenja koji će za mnoge studente biti motivacija, a nekim osobama sa invaliditetom će praktično omogućiti da lakše ostvare pravo na ravnopravnost u obrazovanju. Sa druge strane, nastavnicima se pruža mogućnost da studente lakše navedu na korišćenje adekvatnih nastavnih materijala. Studenti će tamo moći da uče iz ispravnih materijala koji su pod kontrolom nastavnika, a ne iz nećijih pogrešnih beleški što je čest slučaj danas. Dodatno, ako sintetizovani govor jako liči na glas nastavnika – to će bolje podsećati studenta na ono što je slušao na predavanjima; naravno, ovo poslednje je samo pretpostavka koju ćemo u narednom periodu imati priliku da istražimo zahvaljujući rezultatima u razvoju ICT i TTS za srpski čiji su akteri koautori ovog rada.

4. ZAKLJUČAK

U radu je predstavljena Centralna audio-biblioteka UNS, koja studentima pruža mogućnost pristupa audio izdanjima udžbenika i knjiga kao i zvukom obogaćenim prezentacijama sa predavanja. Kreiranje audio-knjiga je do sada podrazumevalo angažovanje ljudi koji bi čitali date tekstove, dok se u okviru CABUNS audio sadržaj generiše automatski na osnovu teksta, što u značajnoj meri ubrzava i pojeftinjuje proces njihovog kreiranja. Primera radi, CABUNS platforma omogućava da sadržaj prezentacija koje se koriste na predavanjima bude obogaćen odgovarajućim objašnjenjima, na osnovu teksta koji se nalazi u beleškama za svaki od slajdova. Savremene TTS tehnologije pored prirodosti i prijatnosti za slušanje, omogućavaju sintezu glasom samog predavača, što omogućuje da se bez velikih ulaganja u kvalitetnu audio i video opremu kreiraju multimedijalna predavanja slična onima koja se nude na *on-line* kursovima prestižnih svetskih univerziteta. Interfejs prema platformi je intuitivan i pregledan, te je korišćenje CABUNS platforme jednostavno i svodi se na podizanje knjiga u .docx formatu, odnosno prezentacija u .pptx formatu na sajt uz popunjavanje odgovarajućih polja od strane nastavnika, dok je preuzimanje audio i multimedijalnog materijala sa sajta standardno.

Osnovna prednost predloženog koncepta ogleda se u mogućnosti automatskog generisanja audio-izdanja udžbenika, naspram klasičnog snimanja glasa govornika, što je vremenski zahtevno i podložno greškama. CABUNS nastavnicima obezbeđuje efikasniji način za generisanje audio-materijala, studentima znatno veću dostupnost i kraći put pribavljanja proverene literature, a njegova posebna prednost ogleda se u mogućnosti pristupa osoba sa invaliditetom. Naredni koraci u istraživanju i razvoju predstavljenog koncepta obuhvataju šire institucionalno umrežavanje zainteresovanih strana, mogućnost konverzije glasa u odgovarajući glas govornika primenom kratke sekvence i poslednjih rezultata istraživanja, kao i praktično rešavanje problema prikaza slika, formula i tabela.

5. LITERATURA

- [1] Dragiša Mišković, Milan Gnjatović, Nikša Jakovljević, Vlado Delić, *Realizacija audio-biblioteke Univerziteta u Novom Sadu*, 11. konf. Digitalna obrada govora i slike (DOGS), Novi Sad, Srbija, 22-25.11.2017., pp. 53-56
- [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, ISSN: 1821-3251, DOI: 10.5937/telfor1701032D, pp. 32-37
- [3] Tijana Delić, Siniša Suzić, Milan Sečujski, Darko Pekar *Rapid Development of New TTS Voices by Neural Network Adaptation*, 17th Int. Symposium INFOTEH-JAHORINA, B&H, March 21-23, 2018, pp. 1-6
- [4] Siniša Suzić, Tijana Delić, Stevan Ostrogonac, Simona Đurić, Darko Pekar *Style-Code Method for Multi-Style Parametric Text-To-Speech Synthesis*, SPIIRAS Proceedings, 5(60), 2018. pp. 216-240, DOI: 10.15622/sp.60.8
- [5] 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 Int. Conf. on Electromechanics and Robotics "Zavalishin's Readings", ER(ZR)-2018, St. Petersburg, Russia, DOI:10.1051/mateconf/201816103005, pp. 1-6