

Istraživanje programske podrške za diktiranje teksta

Brunčić, Dalibor

Master's thesis / Diplomski rad

2014

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:200:678235>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-01-10**

Repository / Repozitorij:

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
ELEKTROTEHNIČKI FAKULTET OSIJEK**

**ISTRAŽIVANJE PROGRAMSKE PODRŠKE ZA
DIKTIRANJE TEKSTA**

Diplomski rad

Dalibor Brunčić

Osijek, 2014.

SADRŽAJ

| | |
|---|-----------|
| 1. UVOD | 1 |
| 2. SUSTAVI ZA AUTOMATSKO RASPOZNAVANJE GOVORA | 3 |
| 2.1. Osnove ASR tehnologija..... | 3 |
| 2.2. Multidisciplinarnе kontribucije procesu strojne obrade i prepoznavanja govora | 10 |
| 2.3. Dijelovi i princip rada ASR sustava..... | 12 |
| 2.4. Glavne karakteristike ASR sustava..... | 17 |
| 2.5. Metrike za mjerenje efikasnosti ASR sustava..... | 19 |
| 2.6. Povijesni razvoj ASR tehnologija i njihovih mogućnosti | 26 |
| 3. KORIŠTENI ALGORITMI, TEHNIKE I METODE KOD AUTOMATSKOG RASPOZNAVANJA GOVORA..... | 32 |
| 3.1. DTW (eng. Dynamic Time Warping) | 32 |
| 3.2. Neuronske mreže..... | 33 |
| 3.3. Skriveni Markovljevi modeli | 34 |
| 3.4. Implementacije umjetne inteligencije | 36 |
| 4. MOGUĆNOSTI PROGRAMSKE PODRŠKE ZA DIKTIRANJE I UNOS TEKSTA..... | 39 |
| 4.1. Načini vrednovanja mogućnosti programske podrške | 39 |
| 4.2. Komercijalna programska podrška i njene mogućnosti | 41 |
| 4.3. Rezultati usporednih testova točnosti komercijalne programske podrške | 47 |
| 5. IZRADA PROGRAMA ZA RASPOZNAVANJE GOVORA I UNOS TEKSTA | 51 |
| 5.1. Općenito o izradi programa..... | 51 |
| 5.2. Opis razvojnog okruženja | 52 |
| 5.3. Koraci u izradi programa | 52 |
| 5.3.1. Izgradnja grafičkog sučelja..... | 52 |
| 5.3.2. Dodavanje reference na MS Speech API | 56 |

| | |
|---|------------|
| 5.3.3. Implementacija/izrada koda..... | 57 |
| 5.3.4. Testiranje ispravnosti programskog koda i validacija dizajna..... | 60 |
| 5.3.5. Postavljanje ulaza i podešavanje postavki..... | 63 |
| 5.3.6. Testiranje točnosti raspoznavanja govora..... | 65 |
| 5.3.7. Rezultati testiranja točnosti | 66 |
| 5.4. Povećavanje točnosti raspoznavanja govora (adaptacija) | 69 |
| 5.4.1. Treniranje Speech Recognition Engine-a | 69 |
| 5.4.2. Fino podešavanje postavki mikrofona u Windows operativnom sustavu | 71 |
| 5.4.3. Dodavanje funkcije za provjeru vjerojatnosti točno prepoznatog govora..... | 73 |
| 5.4.4. Ponovno testiranje točnosti raspoznavanja govora..... | 73 |
| 5.4.5. Rezultati testiranja točnosti adaptiranog modela..... | 75 |
| 6. ZAKLJUČAK..... | 89 |
| LITERATURA | 92 |
| POPIS KRATICA | 96 |
| SAŽETAK..... | 97 |
| ABSTRACT | 98 |
| ŽIVOTOPIS..... | 99 |
| PRILOZI | 100 |

SAŽETAK

Dan je uvod u osnove tehnologije za automatsko raspoznavanje govora. Među navedenim ubrajaju se: praktične primjene sustava za automatsko raspoznavanje govora; prepreke i izazovi pri dizajniranju kvalitetnog sustava; princip rada i uloga pojedinih dijelova sustava; glavne karakteristike i metrike ASR sustava; algoritmi, tehnike i modeli korišteni kod automatskog raspoznavanja govora; te povijesni pregled razvoja ASR tehnologija. Prikazane su neke nove tehnologije koje još nisu implementirane u postojeće komercijalne programe, a znatno povećavaju točnost prepoznavanja riječi u govoru. Istražene su mogućnosti današnjih komercijalnih programa i maksimalna točnost prepoznavanja koju mogu ostvariti, koja se pokazala vrlo visokom (preko 90%). Prikazan je postupak izrade programa za prepoznavanje govora, konfiguracije ulaza, te finog podešavanja svih potrebnih postavki operativnog sustava. Izrađeni program podvrgnut je treniranju i detaljnom testiranju, kako bi se ustanovilo koji sve elementi i uolikoj mjeri utječu na točnost prepoznavanja govora, te koja je najveća točnost prepoznavanja koju pritom može ostvariti. Doneseni su čvrsti zaključci i dana je osnova na temelju koje se može lakše razumjeti i promatrati daljnji razvoj tehnologija za automatsko raspoznavanje govora.

Ključne riječi: automatsko raspoznavanje govora, točnost prepoznavanja, pogreška u prepoznavanju, izrada aplikacije, testiranje sustava, povećavanje točnosti

ABSTRACT

Basics of speech recognition technology are introduced. Among the mentioned topics are: practical uses of speech recognition system; obstacles and challenges in quality system design; principle of operation and specific part roles; main characteristics and metrics of ASR system; algorithms, techniques and models used in automatic speech recognition; and historical overview of ASR technology development. This paper gives insight into some of the new technologies which have the ability of substantially increasing speech recognition accuracy, but have not yet been implemented into existing commercial programs. Abilities of today's commercial programs were explored, as well as maximum speech recognition accuracy they could achieve, which was very high by the way (over 90%). This paper also displays the procedure of making a speech recognition application, configuring input and fine tuning operating system settings. The application was trained and thoroughly tested in order to find out which elements affect recognition accuracy and in what measure, as well as what's the highest accuracy the application can accomplish. Solid conclusions are made and basis is obtained upon which one can easier understand and observe further speech recognition technology development.

Keywords: automatic speech recognition, recognition accuracy, word error rate, application programming, system testing, increasing accuracy