

ZBORNİK RADOVA

**5. MEĐUNARODNE ZNANSTVENO-STRUČNE KONFERENCIJE -
Istraživački dani Visoke policijske škole u Zagrebu
Unaprjeđivanje sigurnosne uloge policije primjenom novih
tehnologija i metoda
Zagreb, Hrvatska, 21. - 22. travnja 2016.**



PROCEEDINGS

**OF THE 5TH INTERNATIONAL SCIENTIFIC
AND PROFESSIONAL CONFERENCE
The Police College Research Days in Zagreb
*New Technologies and Methods Used for Improvement
of the Police Role in Security Matters*
Zagreb, Croatia, 21-22 April 2016**

**Visoka policijska škola u Zagrebu
Police College, Zagreb
Zagreb, 2016.**

Nakladnik/Publisher

Ministarstvo unutarnjih poslova Republike Hrvatske, Policijska akademija /
Ministry of the Interior of the Republic of Croatia, Police Academy

Za nakladnika/On behalf of the Publisher

Želimir Radmilović

Urednici/ Editors

Joško Vukosav

Ksenija Butorac

Joško Sindik

Lektorice/Language Consultants

Gabrijela Gorše

Antonija Rakuljić

Slava Rosandić

Priprema i tisak/Design and printing

Služba za razvoj policijskog obrazovanja i nakladničko-knjižničnu djelatnost

Service for Development of Police Education and Publishing

ISBN 978-953-161-299-9

**5TH INTERNATIONAL SCIENTIFIC AND PROFESSIONAL CONFERENCE
'POLICE COLLEGE RESEARCH DAYS IN ZAGREB'
New Technologies and Methods Used for Improvement of the Police Role
in Security Matters
ZAGREB, CROATIA
21-22 April 2016**

PROGRAMME COMMITTEE

Chair

Assist.Prof. Joško Vukosav, Dean of the Police College, Police Academy, MoI, Zagreb, Croatia

Members

Zvonimir Vnučec, General Police Director, Ministry of the Interior, Zagreb, Croatia

Krunoslav Borovec, PhD, Assistant General Police Director, Ministry of the Interior, Zagreb, Croatia

Želimir Radmilović, Head of the Police Academy, Ministry of the Interior, Zagreb, Croatia

Zvonimir Dujmović, PhD, Assistant Head of the Police Academy, Ministry of the Interior, Zagreb, Croatia

Full Prof. Kimmo Himberg, Police University College, Tampere, Finland

Full Prof. Thorsten Heyer, Police College, Wiesbaden, Germany

Full Prof. Gorazd Meško, Faculty of Criminal Justice and Security, University of Maribor, Ljubljana, Slovenia

Full Prof. Nihad Bunar, Stockholm University, Sweden

Full Prof. Oliver Baćanović, Faculty of Security, Skopje, Macedonia

Full Prof. Zoran Đurđević, The Academy of Criminalistics and Police Studies, Belgrade-Zemun, Serbia

Full Prof. Nedžad Korajlić, Faculty for Criminalistics, Criminology and Security Studies, University of Sarajevo, Bosnia and Herzegovina

Assoc. Prof. Eduardo Manuel Ferreira, Escola de Polícia Judiciária, Loures, Portugal

Assoc. Prof. Gabor Kovacz, National University for Public Service, Budapest, Hungary

Assoc. Prof. Želimir Kešetović, Faculty of Security Studies, University of Belgrade, Serbia

Assist. Prof. Mile Šikman, Department for Police Education, Ministry of Interior, Banja Luka, Republic Srpska

Robert Šumi, PhD, Center for Research and Social Skills, Police Academy, Ministry of the Interior, Ljubljana, Slovenia

Full Prof. Petar Veić, Faculty of Law, University of Rijeka, Croatia

Full Prof. Predrag Zarevski, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia

Full Prof. Ljiljana Mikšaj Todorović, Faculty of Education and Rehabilitation Sciences, University of Zagreb, Croatia

Full Prof. Duško Modly, Zagreb, Croatia

Assist. Prof. Predrag Pale, Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
Assist. Prof. Tihomir Katulić, Faculty of Law, University of Zagreb, Croatia
Assoc. Prof. Irena Cajner Mraović, Centre for Croatian Studies, University of Zagreb, Croatia
Assist. Prof. Anna-Maria Getoš Kalac, Faculty of Law, University of Zagreb, Croatia
Assist. Prof. Joško Sindik, Institute for Anthropological Research, Zagreb, Croatia
Assoc. Prof. Ksenija Butorac, Police College, Zagreb, Croatia
Assoc. Prof. Krunoslav Antoliš, Police College, Zagreb, Croatia
Assist. Prof. Željko Karas, Police College, Zagreb, Croatia
Assist. Prof. Josip Pavliček, Police College, Zagreb, Croatia
Stjepan Gluščić, PhD, Police College, Zagreb, Croatia
Ante Orlović, PhD, Police College, Zagreb, Croatia

ORGANISING COMMITTEE

Chair

Davor Štrk, MSc, Police College, Zagreb, Croatia

Members

Lana Milivojević, PhD, Police College, Zagreb, Croatia

Davorka Martinjak, PhD, Police College, Zagreb, Croatia

Ruža Karlović, PhD, Police College, Zagreb, Croatia

Davor Solomun, MSc, Police College, Zagreb, Croatia

Renata Odeljan, MSc, Police College, Zagreb, Croatia

Hrvoje Filipović, MSc, Police College, Zagreb, Croatia

Nikola Protrka, Police College, Zagreb, Croatia

Damir Maračić, Police College, Zagreb, Croatia

Izidora Radek, Police College, Zagreb, Croatia

Marina Heski, M.A., Police College, Zagreb, Croatia

Nikša Jelovčić, M.A., Police Academy, Zagreb, Croatia

Zlatko Kovač¹⁶⁵
Nikola Protrka¹⁶⁶

RAČUNALNA ANALIZA LJUDSKOG GLASA I GOVORA

Sažetak

U članku se govori o analizi ljudskog glasa (govora) baziranoj na dugotrajnom prosječnom spektru govora (Long-Term Average Spectrum – LTAS), sa aspekta forenzičke analize potrebne kod spornog i nespornog uzorka računalne forenzike odnosno vještačenja. Korištena su dva računalna programa, koji rade u sprezi – WaveSurfer i Catalina Forensic Audio Toolbox. Pri tom prvi program analizira audio zapis i rezultate sprema u obliku tekstualnih datoteka, a drugi program učitava te datoteke, statistički ih obrađuje i grafički prikazuje rezultate kroz niz dijagrama. Ukratko su opisana oba računalna programa, sa svojim mogućnostima i radnim prozorima. Kroz analizu jednog glasa (vokala a) prikazani su osnovni parametri ljudskog glasa (osnovna frekvencija, formanti i dugotrajni prosječni spektar). Zatim je provedena analiza dvaju različitih glasova (govor dviju ženskih osoba sličnih tonalnih karakteristika) i rezultati su komparirani kroz prikaz odnosnih dijagrama. Na kraju je provedena analiza glasa istog govornika, gdje mu je analizirana jedna i druga polovina tijeka govornog zapisa. Relevantni dijagrami su uspoređeni, kako bi se utvrdilo podudaranje krivulja na grafikonima ili utvrdilo moguće odstupanje (kao mjera variabiliteta u glasu i govoru iste osobe). Za navedene analize i komparacije, autori su postavili audio datoteke na internet adresu koje svaki zainteresirani čitatelj može preuzeti i naknadno analizirati.

Ključne riječi: vještačenje, govor, glas, analiza, LTAS.

1. UVOD

Ovim člankom prikazan je segment računalne audioforenzike ljudskog glasa i govora, baziran na dugotrajnom prosječnom spektru govora (LTAS). Prikazana dva softverska alata, *WaveSurfer* nastao na Kraljevskom institutu u Švedskoj¹⁶⁷ i *Catalina Forensic Audio Toolbox*¹⁶⁸ proizvođača Forensic Media Services., Ltd. iz Denvera, SAD, mogu koristiti audioekspertu – vještaku, u sprezi s mnoštvom drugih tehnologija, tehnika i koncepcija. Pošto su ovdje korišteni nekomercijalni programi, slobodno dostupni na internetu, korisno je ukazati

¹⁶⁵ Zlatko Kovač, voditelj programa specijalizacije, e-mail: zkovac@mup.hr

¹⁶⁶ Nikola Protrka, predavač na Visokoj policijskoj školi u Zagrebu, e-mail: nprotrka@fkz.hr

¹⁶⁷ Royal Institute of Technology (KTH), Department of Speech, Music and Hearing, Stockholm, <http://www.speech.kth.se/wavesurfer/> - datum pristupa 19. 2. 2016.

¹⁶⁸ Forensic Media Services., Ltd. Denver, CO, USA, <http://www.forensicav.ro/software.htm> - datum pristupa 19. 2. 2016.

na još jedan sličan program, također dostupan na internetu, a po namjeni sličan navedenima pod nazivom *Praat*¹⁶⁹, program širokih mogućnosti u području govorne analize i sinteze, nastao na Sveučilištu u Amsterdamu.

U smislu šireg razmatranja problematike forenzike govora, stručnjaci se danas koriste i slušnom metodom i spektrometrijskom metodom, a postoje i zagovornici potpune automatizacije analize govora, gdje bi računalo i softver davali odgovore oko identiteta glasa i govora.

Također se mnogo koriste rezultati iz druge, paralelne grane istraživanja glasa (engl. *Speaker Recognition* ili *Speaker Verification*) koja se razvijala za komercijalne svrhe – identifikaciju i autorizaciju osoba na temelju njihova glasa kao lozinke za ulaze u objekte (npr. banke, trezore, vojne objekte itd.) ili pristup različitim računima u elektroničkim transakcijama. Budući da se u ovim automatiziranim sustavima radi o potpuno strojnom procesu identifikacije, rezultat se izračunava pomoću statističkih algoritama i izražava kroz vjerojatnost poklapanja određenih parametara ispitivanog glasa s parametrima referentnih glasova iz glasovne baze.

2. OPIS RAČUNALNIH PROGRAMA WAVESURFER I CATALINA FORENSIC AUDIO TOOLBOX

Metoda analize ljudskog glasa i govora temeljena na dugotrajnom prosječnom spektru govornog signala (engl. Long-Term Average Spectrum – dalje u tekstu LTAS) je postupak kojim se dobiva prosječna vrijednost uzastopnih trenutačnih spektralnih analiza, koje se izračunavaju duž čitavog audiozapisa koji se analizira. Svaka trenutačna spektralna slika, koja se dobiva diskretnom Fourierovom transformacijom, odraz je fonetske karakteristike zvuka u danom trenutku. LTAS sumira ova trenutačna zvučna stanja preko cijelog ispitnog uzorka. LTAS ovisi o zajedničkom utjecaju analiziranog govora, okolnog šuma, šuma elektroničkih uređaja i frekvencijskoj karakteristici prijenosnog lanca.

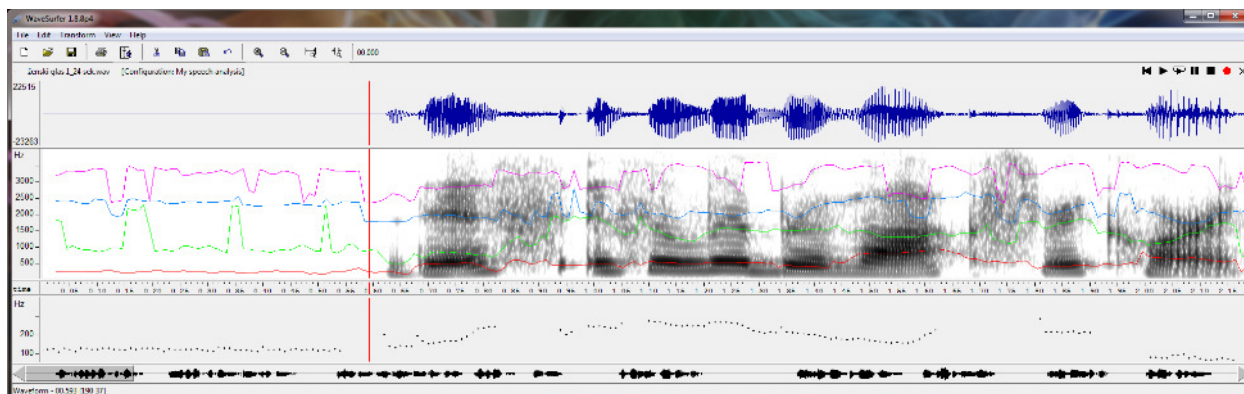
Za audioanalize danas postoji mnoštvo aplikacija, kako komercijalnih tako i slobodnih, dostupnih putem interneta. Važno je istaći da ovaj program svoje rezultate sprema u obliku tekstualnih datoteka. Ove datoteke učitava drugi program, *Catalina Forensic Audio Toolbox* (također slobodno dostupan na internetu), statistički ih obrađuje i rezultate prikazuje u obliku višestrukih dijagrama.

Audiozapis mora biti jednokanalni (mono), u *wav* zapisu i PCM formatu, uzorkovan frekvencijom 8 kHz, i rezolucije 16 bita. Ako je audiozapis u nekom drugom formatu, on se prije svega mora konvertirati u ovaj navedeni format. Ovo je potrebno zbog kompatibilnosti s *Catalina Forensic Audio Toolboxom*, koji će, u tom slučaju, najtočnije izvesti statističku obradu primljenih tekstualnih podataka izvezenih iz *WaveSurfera* i onda ih prikazati u obliku dijagrama. Nadalje, u svrhu kompatibilnosti ova dva programa, potrebno je prilagoditi i

¹⁶⁹ University of Amsterdam, Phonetic Sciences, <http://www.fon.hum.uva.nl/praat/> datum pristupa 19. 2. 2016.

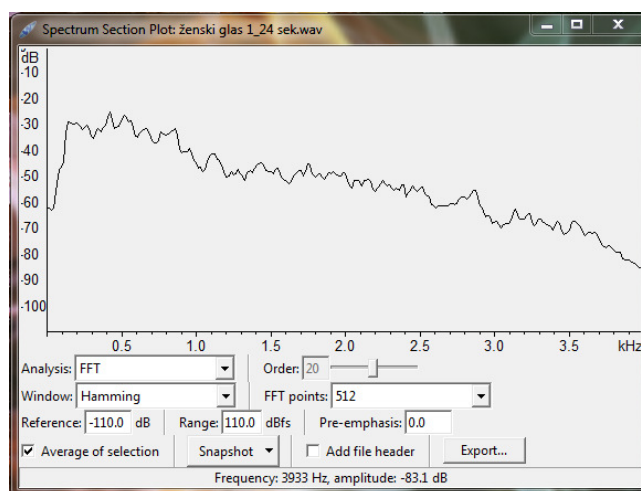
određene postavke u programu *WaveSurfer*, koje su navedene u *Catalina Forensic Audio Toolbox* korisničkom priručniku.

WaveSurfer program ima više inicijalnih konfiguracija, koje se, na početku upotrebe ovog alata, odabiru ovisno o poslovima koje primarno treba obaviti, kao npr. razne vrste transkripcije audiozapisa, analiza spektrograma, analiza govora, vremenska analiza audiosignala ili konfiguracija za demonstraciju mogućnosti aplikacije. Moguće je, također, postaviti razne parametre programa po svojoj volji i spremiti takve postavke kao svoju konfiguraciju za buduću uporabu. Na slici 1 prikazana je jedna od konfiguracija – za analizu govora (Speech analysis), s učitanim audiozapisom.



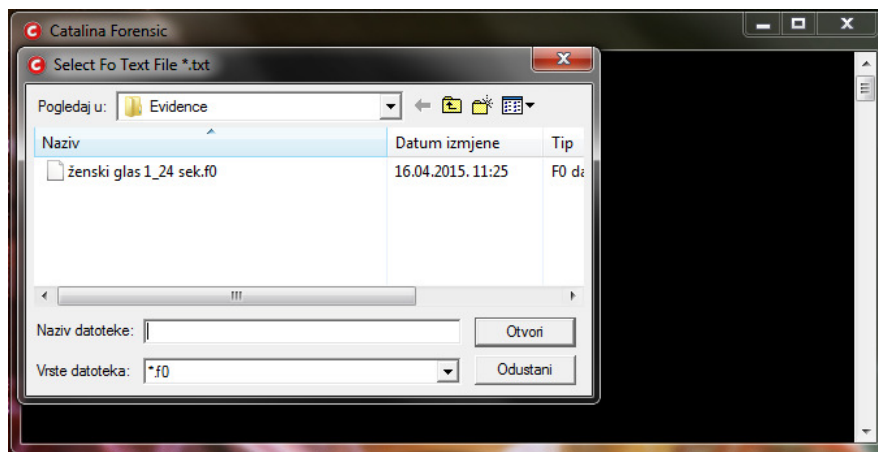
Slika 1: Radni prozor programa *WaveSurfer*

U gornjem redu radne površine aplikacije prikazan je početni dio analiziranog audiozapisa u vremenskoj domeni. Ukupni audiozapis prikazan je u donjem redu, a sivo označeni vremenski segment zapisa uvećan je u gornjem redu. Drugi red radne površine prikazuje spektrogram (u nijansama sivog) i četiri formanta prikazanih u boji (F1 – crveno, F2 – zeleno, F3 – plavo i F4 – ljubičasto). Treći red prikazuje osnovnu frekvenciju analiziranog glasa F0 (visina glasa, engl. *pitch*). Iz gore prikazanog prozora *WaveSurfera* može se otvoriti prozor za prikaz dugotrajnog prosječnog spektra – LTAS (slika 2).



Slika 2: *WaveSurfer* – LTAS

Kako je već navedeno, *WaveSurfer* stvara tekstualne datoteke rezultata svoje zvučne analize. Program kreira tri tekstualne datoteke oblika: *naziv_datoteke.fo* (za osnovnu frekvenciju glasa), *naziv_datoteke.frm* (za formante F1, F2, F3 i F4) i *naziv_datoteke.lts* (za dugotrajni prosječni spektar, LTAS). Prve dvije datoteke formiraju se iz osnovnog prozora *WaveSurfera* prikazano na slici 1, a treća datoteka formira se iz LTAS prozora vidljivo na slici 2, pritiskom na izbornik *Export...* Sve tri datoteke treba spremiti u direktorij *Evidence* programa *Catalina Forensic Audio Toolbox* i nakon toga pokrenuti program *Catalina Forensic Audio Toolbox*. Program je, zapravo, statistički softver za grafički prikaz podataka danih u tabelarnom obliku u tekstualnim datotekama koje generira *WaveSurfer* (slika 3).



Slika 3: Početni prozor programa *Catalina Forensic Audio Toolbox*

Program, odmah po učitavanju, traži da se učita prva od tri formirane datoteke s ekstenzijom *fo*. Nakon što se unese *fo* datoteka, program dalje radi automatski – sam pronalazi i učitava ostale dvije datoteke istog imena ali ekstenzija *frm* i *.lts* iz direktorija *Evidence*. Nakon obrade, *Catalina Forensic Audio Toolbox* stvara osam slika s dijagrama. Slike su u *.tif* formatu i spremaju se u direktorij *Plots* programa.

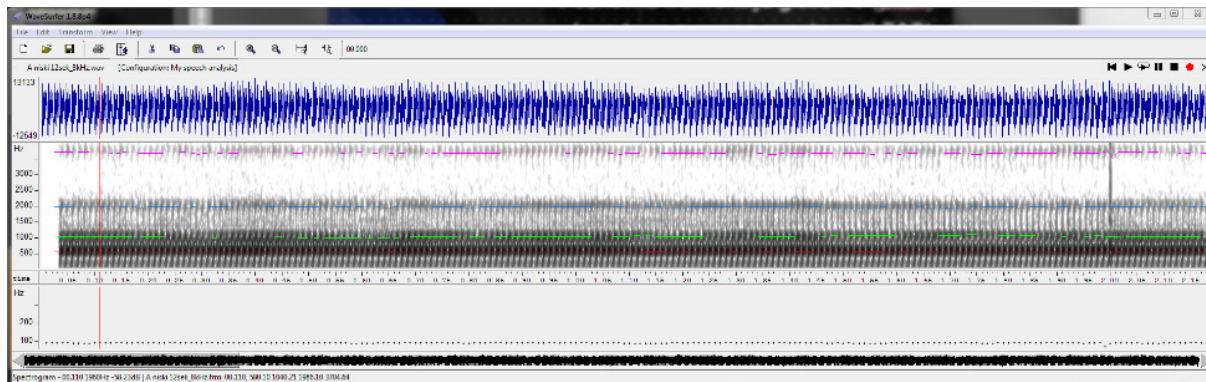
3. OSNOVNI PARAMETRI LJUDSKOG GLASA I GOVORA

Uporaba dijagrama za analizu audiozapisa pojedinih govornika bit će prikazana u nastavku, a prije toga, prikazat će se osnovni parametri ljudskog glasa – osnovna frekvencija glasa F0 (visina glasa), frekvencije formanta F1-F4 i dugotrajni prosječni spektar LTAS kroz primjer analize vokala *a*. Vokal je izabran, zato što vokali nose energiju glasa i imaju jasno određenu osnovnu frekvenciju F0 (što je u stvari frekvencija titranja glasnica, bez utjecaja ostalog vokalnog trakta) i jasno određene formante F1, F2, F3, F4 (naglašena frekvencijska područja glasa, kao rezultat rezonantnih karakteristika vokalnog trakta).

Sve korištene zvučne datoteke iz ovog članka postavljene su na internetskoj adresi <http://1drv.ms/1I3kHUS> i slobodne su za preuzimanje sa *Cloud* sustava *Microsoft OneDrive*.

Audiodatoteka glasa *a* - *A niski 12sek_8kHz.wav* je trajanja 12 sekundi, jer za analizu dugotrajnog prosječnog spektra (LTAS) *Catalina Forensic Audio Toolbox* zahtijeva duljinu

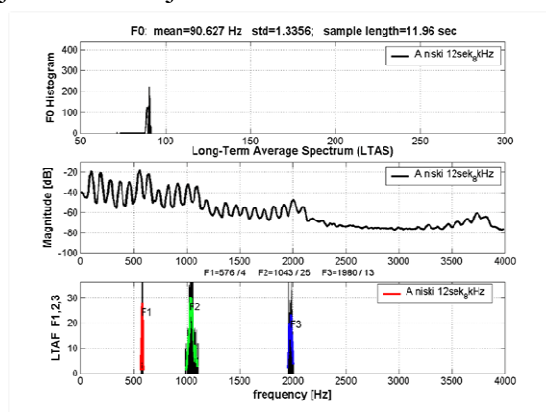
audiozapisa veću od 10 sekundi. Nakon učitavanja zvučne datoteke *A niski 12sek_8kHz.wav*¹⁷⁰ u *WaveSurfer*, program daje rezultat prikazan na slici 4.



Slika 4: Rezultat analize audiozapisa vokala *a*

Iz donjeg dijela slike može se iščitati osnovna frekvencija F_0 vokala *a* od približno 100 Hz. Sredina slike daje formante: $F_1 \approx 500$ Hz, $F_2 \approx 1000$ Hz, $F_3 \approx 2000$ Hz i $F_4 \approx 3700$ Hz. Formant F_4 neće se uzimati u obzir jer *Catalina Forensic Audio Toolbox* uzima i obrađuje samo 3 formanta (F_1 , F_2 , i F_3) što je u praksi najčešće dovoljno za analizu govora. Iz slike se vidi kako su osnovna frekvencija i formanti, u ovom slučaju, vremenski gotovo nepromjenjivi – ravne vodoravne linije duž vremenske osi. To, međutim, kako će se kasnije vidjeti, nije slučaj za govor. Prikazani rezultati samo su gruba orijentacija o vrijednosti parametara analiziranog glasa – točne vrijednosti dat će program *Catalina Forensic Audio Toolbox*.

Kako je prije rečeno, sada treba unutar *WaveSurfera*, formirati 3 tekstualne datoteke podataka (*niski 12sek_8kHz.fo*, *niski 12sek_8kHz.frm* i *niski 12sek_8kHz.lts*) koje će *Catalina Forensic Audio Toolbox* statistički obraditi i prikazati dijagramima. Detalji ovog postupka navedeni su u korisničkom priručniku i ovdje se neće pobliže objašnjavati. Formirane datoteke spremaju se u direktorij *Evidence* programa *Catalina Forensic Audio Toolbox*. Potom se aktivira program, učita se prva datoteka (*niski 12sek_8kHz.fo*) i dalje, nakon obrade, rezultati se u grafičkom obliku spremaju u direktorij *Plots* kao slike u .tif formatu – vidi sliku 5.



Slika 5: Dijagrami programa *Catalina Forensic Audio Toolbox* glasa *a*

¹⁷⁰ Internetska adresa Microsoft OneDrive <http://1drv.ms/1I3kHUS> (datoteka *A niski 12sek_8kHz.wav*), datum pristupa 19. 2. 2016.

Ovdje su, sada, uz iscrtane grafove, točno ispisane srednja vrijednost osnovne frekvencije vokala *a* F_0 u hercima s pripadajućom standardnom devijacijom ($F_0=90,627$ Hz, $std=1,3356$) i srednje vrijednosti formanata F_1 , F_2 i F_3 u hercima s pripadajućim standardnim devijacijama ($F_1=576/4$, $F_2=1043/25$, $F_3=1980/13$). Prikazan je i dijagram dugotrajnog prosječnog spektra (LTAS), na kojem se vide energetska nadvišenja na frekvencijama oko 500, 1000, 2000 i 3700 herca, što odgovara formantima.

Dijagrami sa slike 5 osnovni su rezultati audioanalize uz pomoć ova dva programa. U nastavku će biti pokazana šira primjena ovih alata u analizi dužeg govora ispitanika. Rezultantni dijagrami prikazat će se na osam slika koje su produkt statističke obrade ulaznih podataka uz pomoć programa *Catalina Forensic Audio Toolbox*.

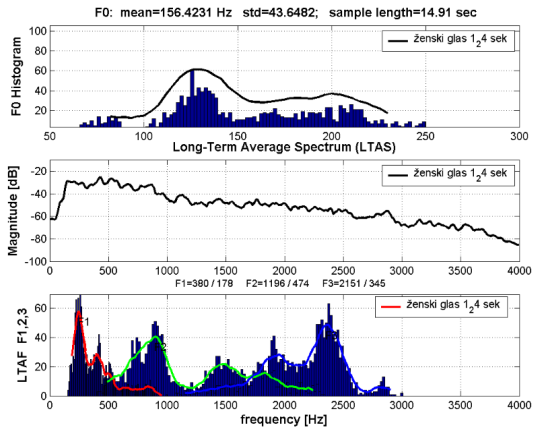
4. ANALIZA GLASA JEDNOG GOVORNIKA

Već je rečeno kako *Catalina Forensic Audio Toolbox* stvara ukupno osam slika s dijagramima. Pojašnjenja toga što se stvara na pojedinoj slici su sljedeća:

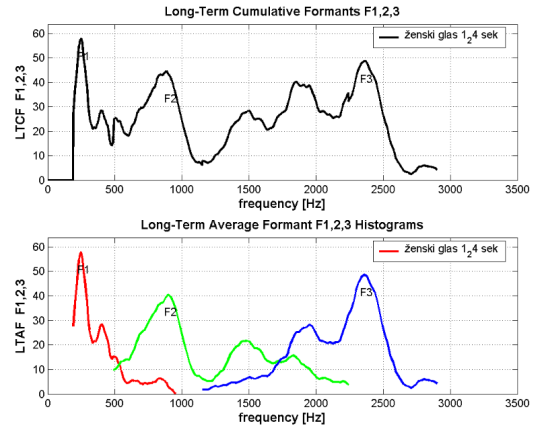
1. slika – 3 dijagrama; F_0 histogram, LTAS i LTAF (Long-Term Average Formants) histogram formanata F_1 , F_2 i F_3
2. slika – 2 dijagrama; kumulativni histogram formanata i pojedinačni histogram formanata F_1 , F_2 i F_3
3. slika – ukupni formantni prostor F_1 - F_2
4. slika – ukupni formantni prostor F_2 - F_3
5. slika – formantni prostor F_1 - F_2 vokala *a*, *e*, *i*, *o* (za engleski jezik)
6. slika – formantni prostor F_2 - F_3 vokala *a*, *e*, *i*, *o* (za engleski jezik)
7. slika – 2 dijagrama; Long-Term Average Spectrum-Fast Fourier Transform (LTAS-FFT) dijagram i LTAS histogram izvedeni su programom *Catalina Forensic Audio Toolbox*. LTAS-FFT dijagram identičan je dijagramu LTAS kojeg je iscrtao *WaveSurfer*.
8. slika – 3 dijagrama; LTAS-FFT, Long-Term Cumulative Formants F_1 , F_2 , F_3 i Long-Term Average Formant F_1 , F_2 , F_3 Histograms također su izvedeni programom *Catalina Forensic Audio Toolbox* i identični su dijagramima s podacima iz *WaveSurfera*.

Svi dijagrami bit će prikazani kao rezultat analize audiozapisa govora jedne ženske osobe – *ženski glas 1_24 sek.wav* i gore već navedenim slijedom.¹⁷¹ Ovu datoteku ćemo u opisu slika imenovati *Datoteka 1*. Postupak rada s programima upravo je opisan, pa sada slijede samo rezultati – *Catalina* dijagrami prikazani na osam slika – od slike 6 do slike 13.

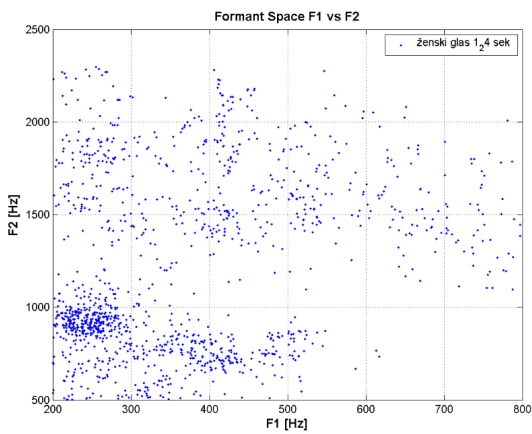
¹⁷¹ Internetska adresa Microsoft OneDrive <http://1drv.ms/1I3kHUS> (datoteka *ženski glas 1_24 sek.wav*), *Datoteka 1*, datum pristupa 19. 2. 2016.



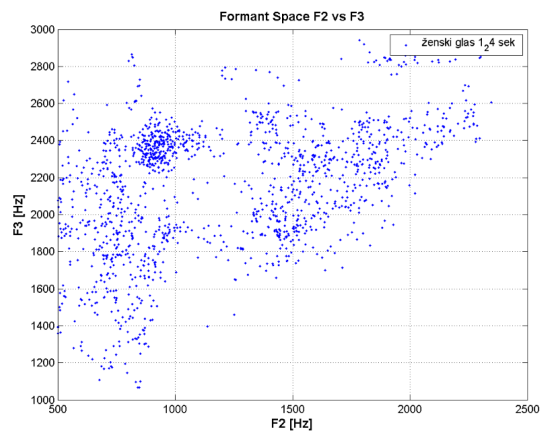
Slika 6: Datoteka 1 - dijagram 1



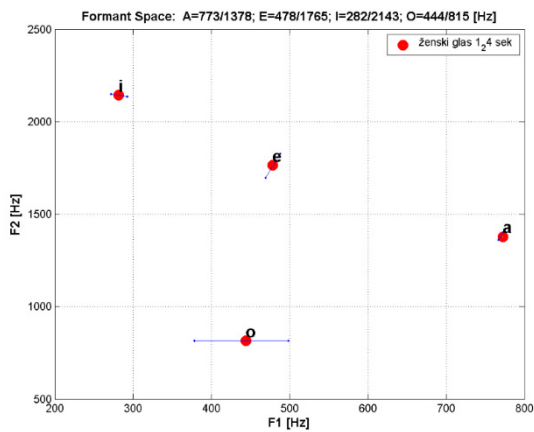
Slika 7: Datoteka 1 - dijagram 2



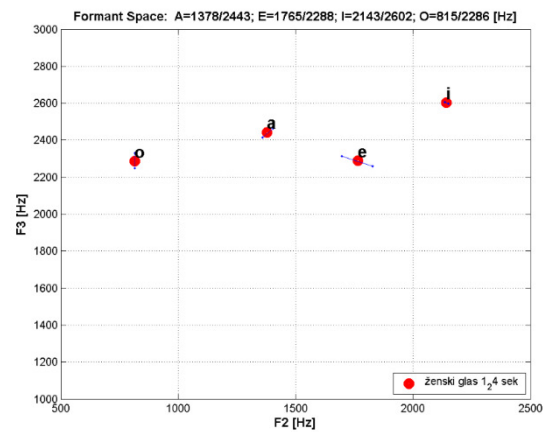
Slika 8: Datoteka 1 - dijagram 3



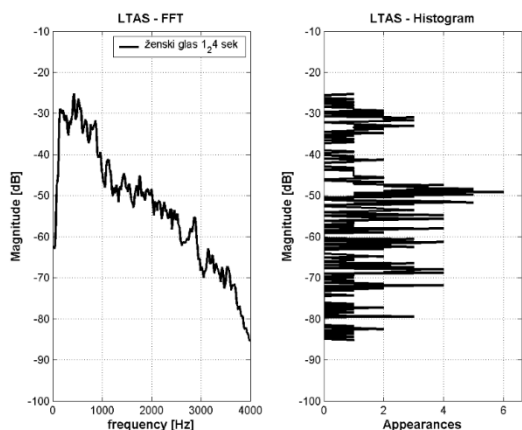
Slika 9: Datoteka 1 - dijagram 4



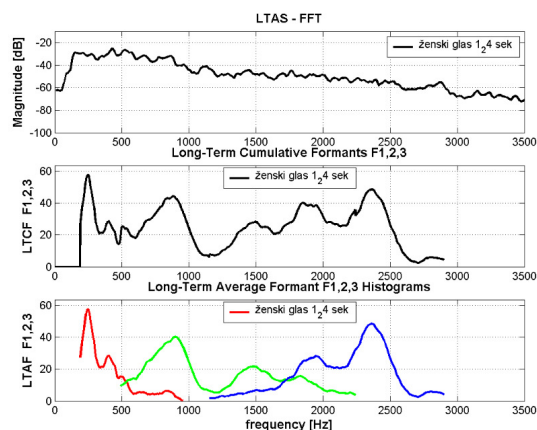
Slika 10: Datoteka 1 - dijagram 5



Slika 11: Datoteka 1 - dijagram 6



Slika 12: Datoteka 1 - dijagram 7



Slika 13: Datoteka 1 - dijagram 8

Komentari prikazanih dijagrama nam pojašnjavaju:

- Histogram LTAF na slici 7 – dolje jednak je onom na slici 6 – dolje, na kojem su prikazane samo vanjske konture histograma (bez ispuna ispod grafa).
- Histogram LTCF na slici 7 – gore grafički je zbroj pojedinačnih dijagrama F1, F2 i F3 sa donjeg grafa i čini jednu konturu (kumulativni ili zbrojeni formanti).
- Slike 8 i 9 prikazuju ukupni formantni prostor F1-F2 odnosno F2-F3.

Svaka točkica na dijagramu predstavlja jednu detektiranu frekvenciju formanata F1, F2 i F3 u kratkom vremenskom intervalu analize. Prozor analize pomiče se, zatim, na novi vremenski interval i tu se očitavaju formanti za sljedeću točkicu u grafu i tako dalje, do kraja audiozapisa.

Zanimljivo je uočiti vezu rezultata sa slike 8 i 9 i grafova na slici 7. Naime, najveća gustoća točkica na slici 8 je za formante F1 \approx 250 Hz i F2 \approx 900 Hz. Kad se, sada pogleda slika 7, vide se pikovi (najveći broj pojavljivanja) upravo na tim frekvencijama. Isto vrijedi i za sliku 9 – najveća gustoća točkica je za F2 \approx 900 Hz i F3 \approx 2300 Hz, što odgovara pikovima na slici 7.

- Slike 10 i 11 prikazuju položaj vokala *a*, *e*, *i*, *o* u formantnim prostorima F1-F2 i F2-F3. Naime, statistički su određene granice frekvencija formanata u hertzima, za pojedine vokale engleskog jezika. Vrijednosti su prikazane u tablici na slici 14 (*Catalina Forensic Audio Toolbox User's Manual*).

601	850	1100	1600	2200	2800	← vowel [a]
401	600	1500	2000	2100	2800	← vowel [e]
220	400	2000	2400	2400	2900	← vowel [i]
370	600	700	1200	2200	2600	← vowel [o]
↑	↑	↑	↑	↑	↑	
low	high	low	high	low	high	
limits	limits	limits				
for F1	for F2	for F3				

Slika 14: Frekvencije formanata vokala u engleskom jeziku

Vrijednosti sa slike 14 nalaze se u datoteci *formants.txt* direktorija *Catalina*, u kojem je instalirana aplikacija. Datoteka je promjenjiva i u nju se mogu upisati vrijednosti i za druge jezike ili vrijednosti za neka specifična istraživanja određenih skupina govornika – ispitanika. Na slikama 10 i 11 zatamnjeni kružići prikazuju prosječne vrijednosti formanata pojedinih vokala u cijelom audiozapisu, a u skladu s granicama iz tablice na slici 14. Male točkice oko kružića, povezane linijom, srednje su vrijednosti formanata određenog vokala, izračunate za prvu i za drugu polovicu trajanja ukupnog audiozapisa. Duljina linije ukazuje na varijabilnost vokala (a time i govora) iste osobe (ispitanika).

- Dijagrami na slikama 12 i 13 dobiveni su metodom brze Fourierove transformacije (FFT) unutar samog programa *Catalina Forensic Audio Toolbox*. Vidi se kako su ovi dijagrami identični onima izvedenim programom *WaveSurfer*.

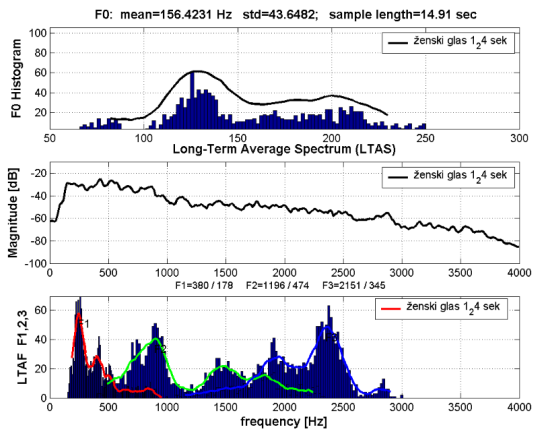
LTAS histogram na slici 12 prikazuje relativnu učestalost pojavljivanja (Appearances) određene energije govora (analogno glasnoći govora) za pojedinu frekvenciju govornog spektra (lijevi dijagram na slici 12, LTAS – FFT).

5. KOMPARACIJA GLASOVA DVOJE GOVORNIKA

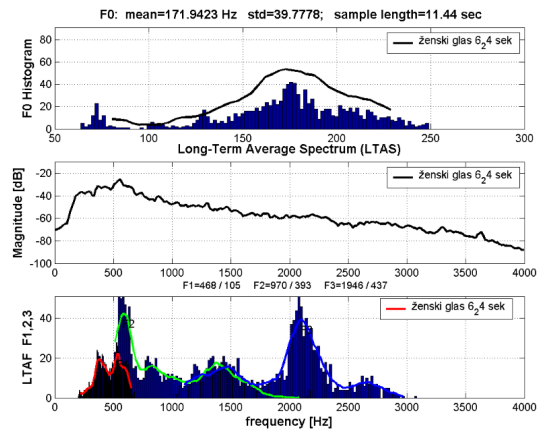
Pravi smisao forenzike ljudskog glasa ili govora dolazi u postupcima komparacije rezultata dviju analiza, gdje se, jednostavnom vizualnom usporedbom grafova, može zaključiti radi li se o glasovima jedne osobe na obje snimke, ili se radi o glasovima različitih osoba. Za demonstraciju će se usporediti govor dviju ženskih osoba, ženski glas 1, na audiozapisu *ženski glas 1_24 sek.wav* i ženski glas 2, na audiozapisu *ženski glas 6_24 sek.wav*. Datoteku *ženski glas 6_24 sek.wav* u opisu slika ćemo imenovati kao Datoteka 2.¹⁷²

Kako bi se komparacija, odnosno usporedba otežala, izabrani su slični ženski glasovi po boji, intenzitetu govora i govornoj dinamici (preporuča se poslušati navedene audiodatoteke radi slušne usporedbe). Rezultati za ženski glas 1 već su gore navedeni i objašnjeni, a za ženski glas 2 treba izvesti identični, već opisani postupak analize. Ponovno će se formirati osam slika s dijagramima za ženski glas 2 i ove će se slike međusobno usporediti. Rezultati su prikazani na slikama 15 – 22.

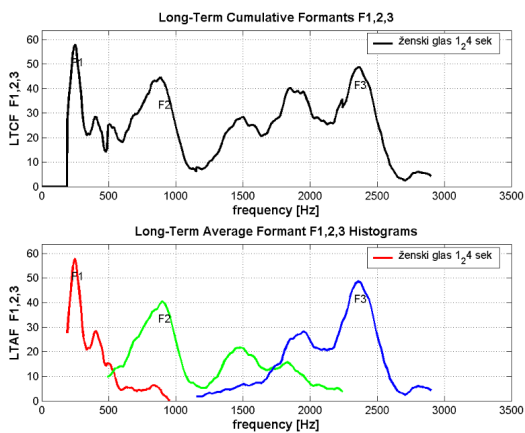
¹⁷² Internetska adresa Microsoft OneDrive <http://1drv.ms/1I3kHUS> (datoteka *ženski glas 6_24 sek.wav*), Datoteka 2, datum pristupa 19. 2. 2016.



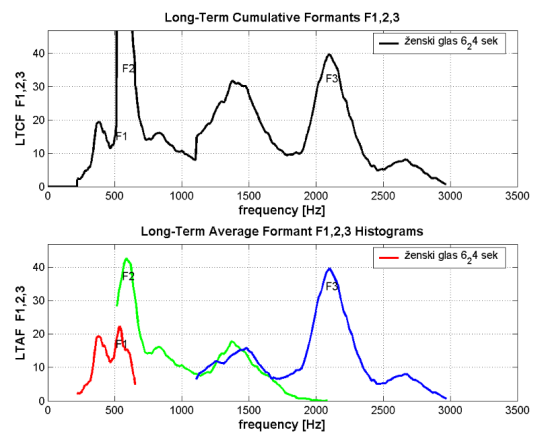
Slika 6: Datoteka 1 - dijagram 1



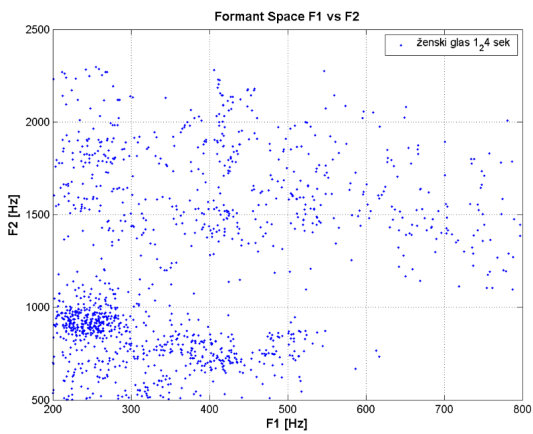
Slika 15: Datoteka 2 - dijagram 1



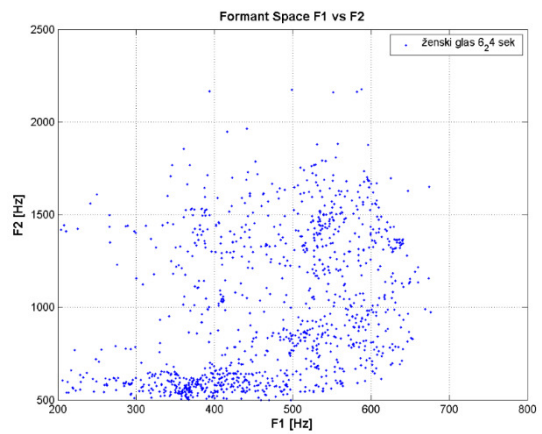
Slika 7: Datoteka 1 - dijagram 2



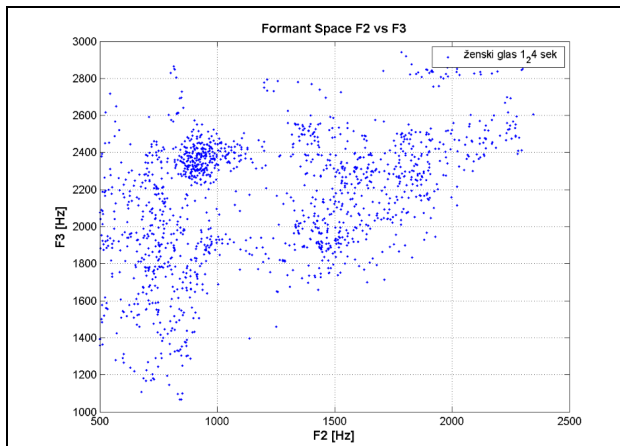
Slika 16: Datoteka 2 - dijagram 2



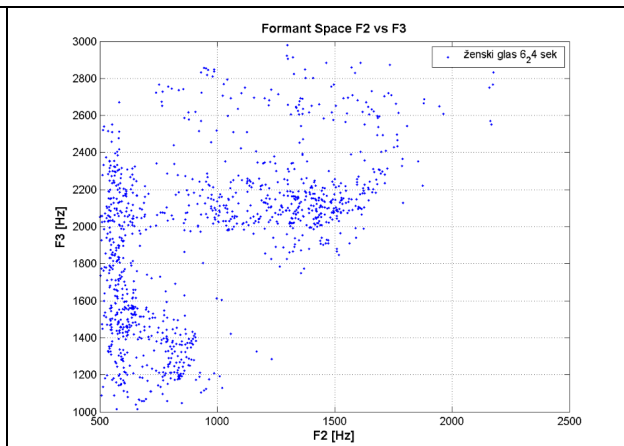
Slika 8: Datoteka 1 - dijagram 3



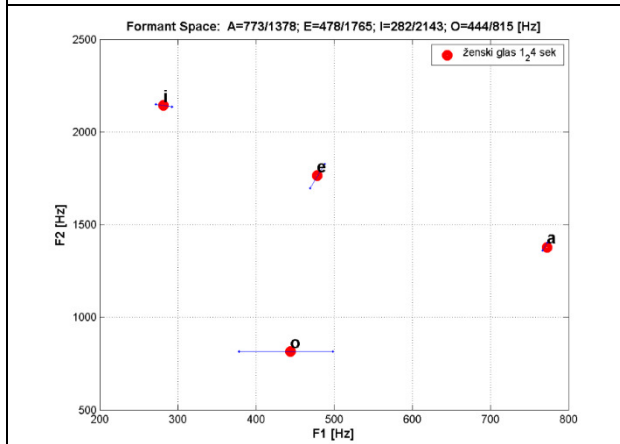
Slika 17: Datoteka 2 - dijagram 3



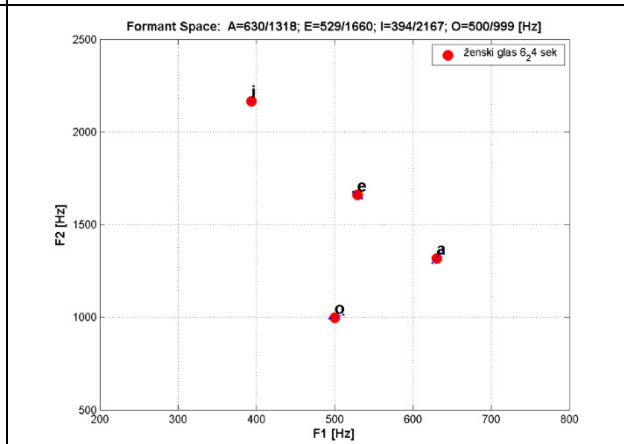
Slika 9: Datoteka 1 - dijagram 4



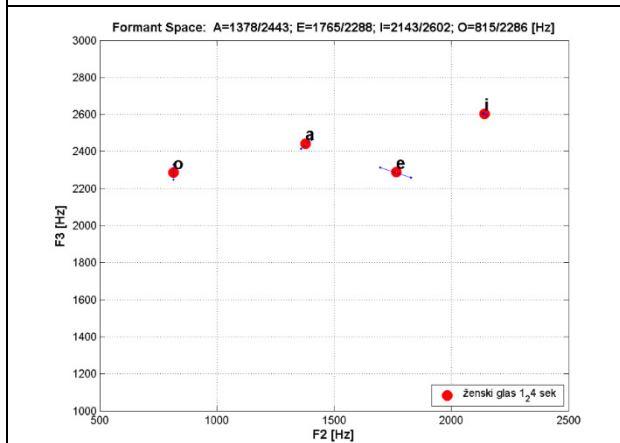
Slika 18: Datoteka 2 - dijagram 4



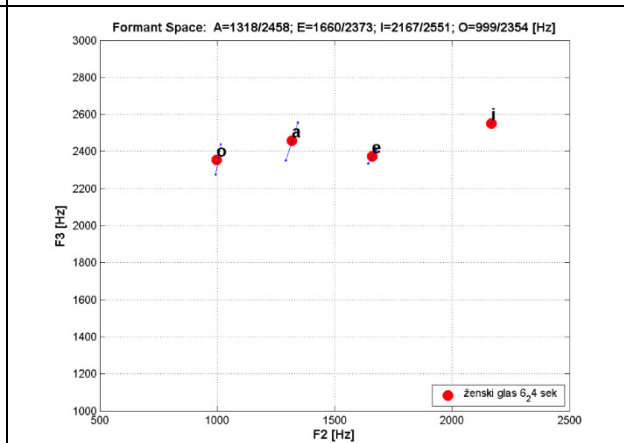
Slika 10: Datoteka 1 - dijagram 5



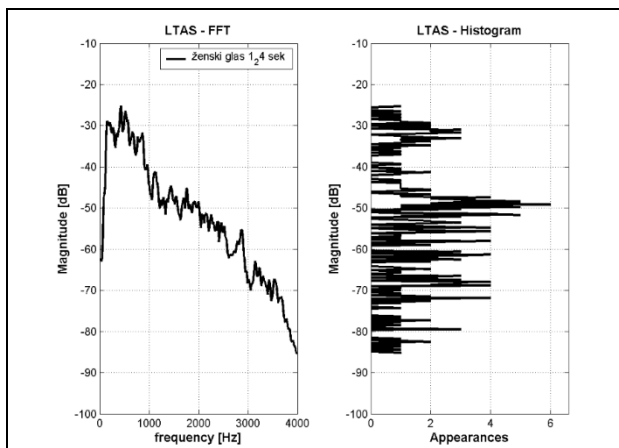
Slika 19: Datoteka 2 - dijagram 5



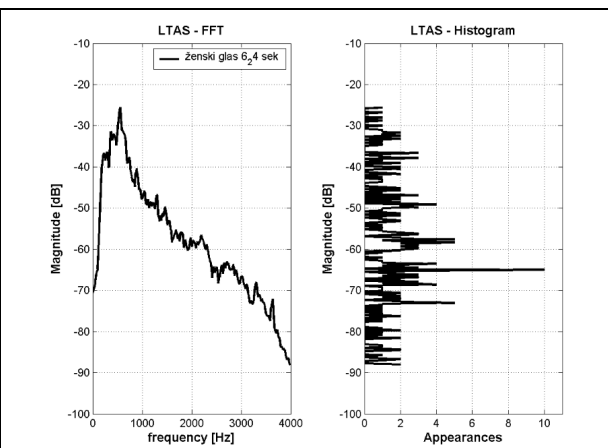
Slika 11: Datoteka 1 - dijagram 6



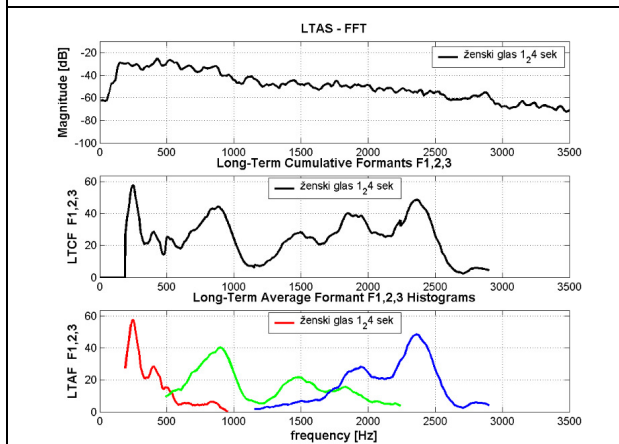
Slika 20: Datoteka 2 - dijagram 6



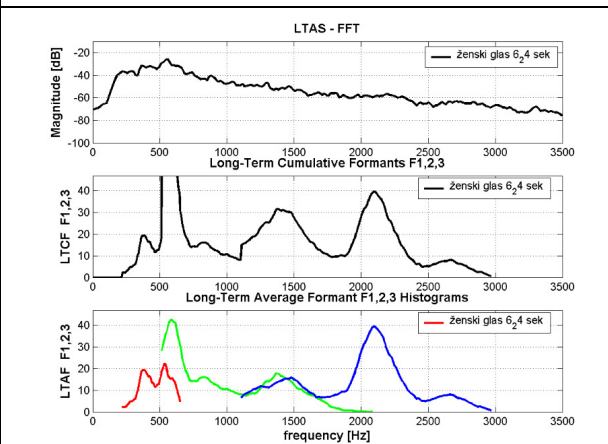
Slika 12: Datoteka 1 - dijagram 7



Slika 21: Datoteka 2 - dijagram 7



Slika 13: Datoteka 1 - dijagram 8



Slika 22: Datoteka 2 - dijagram 8

Usporedbu istovjetnih prikaza dijagrama Datoteke 1 i Datoteke 2 ne treba posebno komentirati. Izgledi i tijekovi grafova, te položaji vokala u formantnim prostorima, bitno se razlikuju, što upućuje na zaključak kako se radi o glasovima dviju različitih (ženskih) osoba.

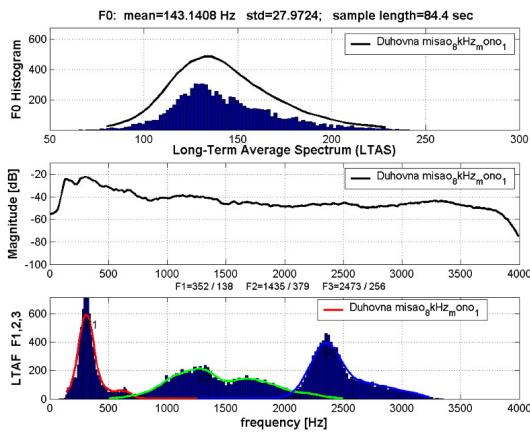
6. KOMPARACIJA GLASOVA ISTOG GOVORNIKA

U počecima audioforenzike, spektrogram nekog glasa nazivao se i *voiceprint*, po uzoru na forenziku papilarnih linija – *fingerprint*. Daljnjim proučavanjem ljudskog glasa i govora, došlo se, međutim, do spoznaje kako je ljudski glas sve samo ne tako stalan kao otisak prsta, daleko podložniji kojekakvim promjenama, pa ga se, strogo uzevši, ne može smatrati nikakvim postojanim "otiskom glasa". Jedan glas, riječ ili rečenica izgovoreni uzastopno, kod iste osobe, nisu potpuno jednaki po svojim zvučnim karakteristikama. Stoga, kod komparacije glasova u nekoj ekspertizi, uvijek se mora imati na umu i postojanje varijabiliteta unutar istog glasa.

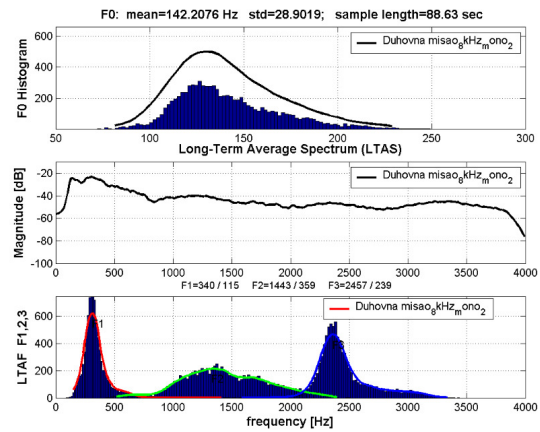
Kako izgledaju rezultati komparacije glasa istog govornika prikazat će se na sljedećem primjeru. Koristit će se audiozapis jedne radioemisije – Duhovna misao – govor muške osobe u trajanju od 4 minute i 10 sekundi. Audiozapis će se podijeliti na dva jednaka dijela –

datoteku *Duhovna misao_8kHz_mono_1.wav*, trajanja 2:04 min. i datoteku *Duhovna misao_8kHz_mono_2.wav*, trajanja 2:06 min. Datoteku *Duhovna misao_8kHz_mono_1.wav* u opisu slika ćemo imenovati Datoteka 3, a datoteku *Duhovna misao_8kHz_mono_2.wav* u opisu slika ćemo imenovati kao Datoteka 4.¹⁷³

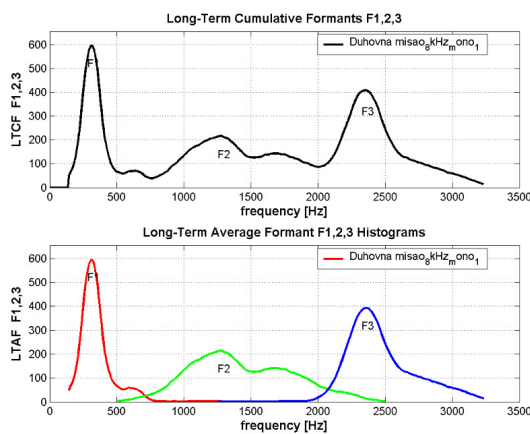
Na obje datoteke primijenit će se isti, gore već opisani postupak analize, a zatim će se pripadajući dijagrami usporediti. Rezultati su prikazani na slikama od 23 do 38.



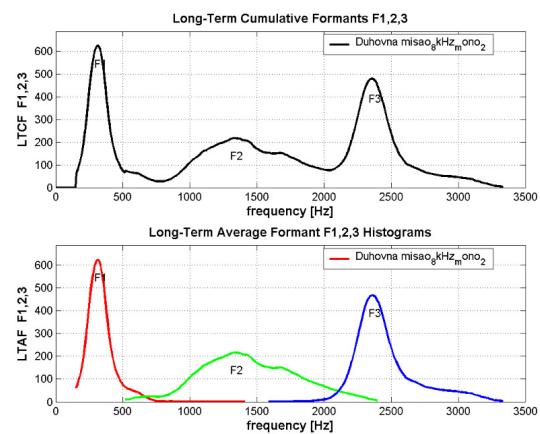
Slika 23: Datoteka 3 - dijagram 1



Slika 24: Datoteka 4 - dijagram 1

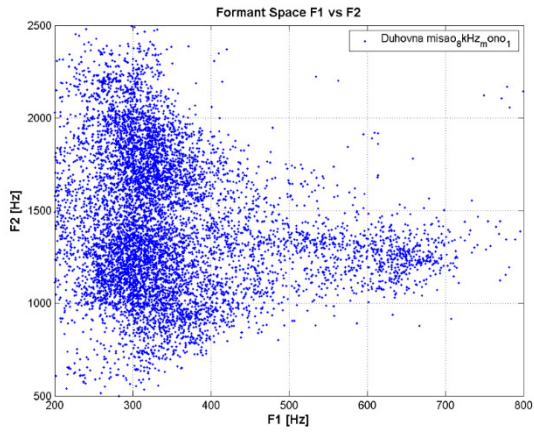


Slika 25: Datoteka 3 - dijagram 2

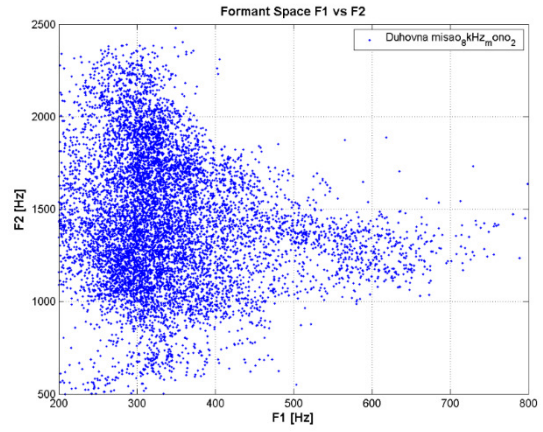


Slika 26: Datoteka 4 - dijagram 2

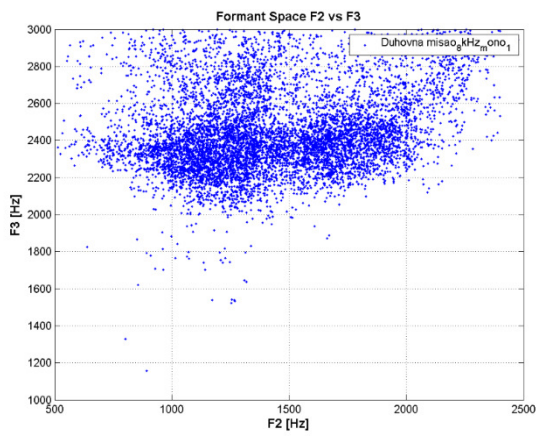
¹⁷³ Internetska adresa Microsoft OneDrive <http://1drv.ms/1I3kHUS> *Duhovna misao_8kHz_mono_1.wav* - Datoteka 3, *Duhovna misao_8kHz_mono_2.wav* - Datoteka 4, datum pristupa 19. 2. 2016.



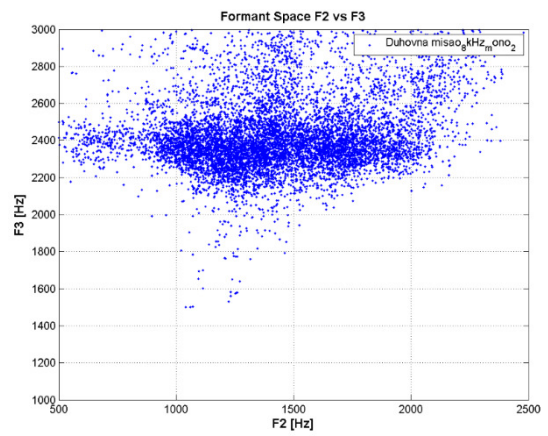
Slika 27: Datoteka 3 - dijagram 3



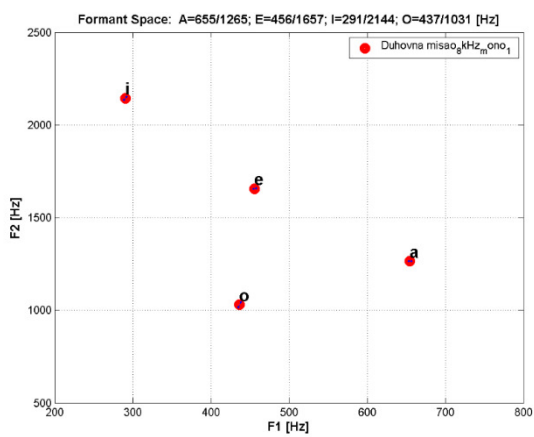
Slika 28: Datoteka 4 - dijagram 3



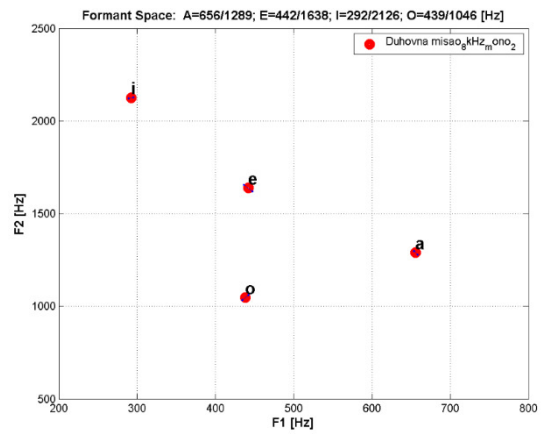
Slika 29: Datoteka 3 - dijagram 4



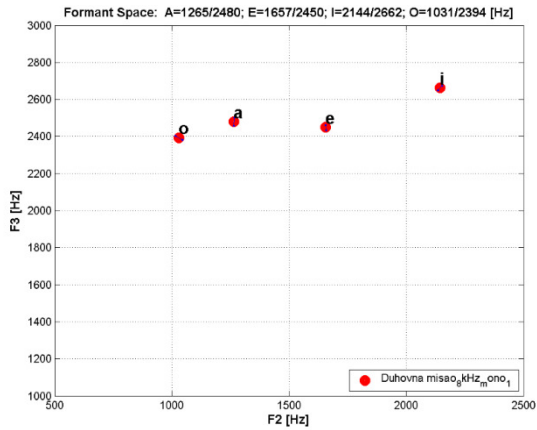
Slika 30: Datoteka 4 - dijagram 4



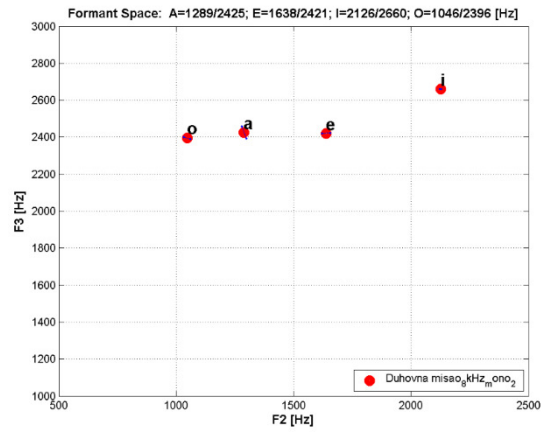
Slika 31: Datoteka 3 - dijagram 5



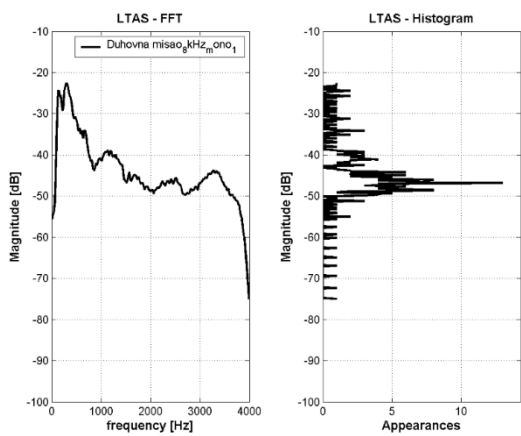
Slika 32: Datoteka 4 - dijagram 5



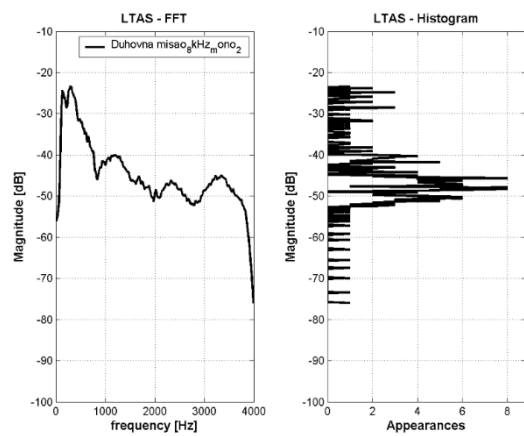
Slika 33: Datoteka 3 - dijagram 6



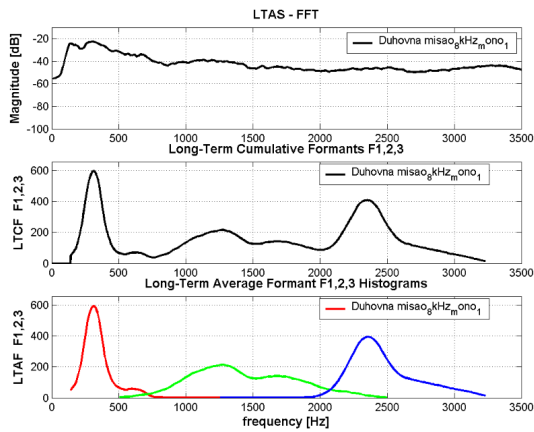
Slika 34: Datoteka 4 - dijagram 6



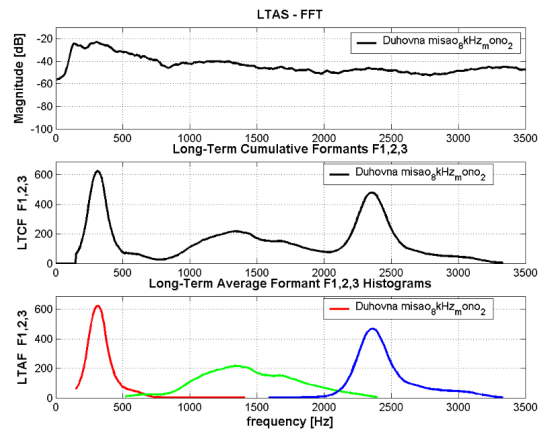
Slika 35: Datoteka 3 - dijagram 7



Slika 36: Datoteka 4 - dijagram 7



Slika 37: Datoteka 3 - dijagram 8



Slika 38: Datoteka 4 - dijagram 8

Međusobno pripadajući dijagrami pokazuju izvrsno poklapanje, što ukazuje na to kako je varijabilitet ovog glasa (govornika) minimalan.

7. ZAKLJUČAK

Iz prikazanog postupka analize govora, temeljene na dugotrajnom prosječnom spektru (LTAS), vidi se kako su rezultati prikazani kroz mnoštvo dijagrama. Ovim načinom omogućena je jednostavna komparacija dvaju ispitivanih zvučnih zapisa – relevantni dijagrami jednostavno se stavljaju jedan pored drugog, te se ocjenjuju tokovi krivulja, pronalaze vrhovi (maksimalne vrijednosti) na tim dijagramima, s pripadajućim frekvencijama i magnitudama. Lagano se uočavaju podudarnosti i/ili odstupanja, što može vidjeti i osoba koja nije ekspert u audioforenzici. Jednostavnom vizualnom usporedbom ocjenjuje se pripadaju li dva uzorka govora istoj osobi ili različitim osobama. Sadržaji govora dvaju audiozapisa ne moraju biti jednaki, jer LTAS postupak, u duljem vremenskom intervalu, usrednjuje frekvencijsku karakteristiku govornika. Međutim, ovi zapisi morali bi biti snimljeni s približno jednakim elektroničko prijenosnim karakteristikama kako bi se utjecaj varijabli na kvalitetu snimke što više smanjio.

Koliko je područje vještačenja i analize glasa i govora kompleksno i, kao dokaz na sudu, još uvijek nepouzđano ili samo u određenom stupnju pouzdano, govori i sudska praksa u svijetu. Identifikaciju osobe po glasu neki sudovi u SAD-u i drugim državama u svijetu prihvaćaju kao dokaz, neki uopće ne prihvaćaju, a neki prihvaćaju djelomično i uvjetovano, s određenim stupnjem dokazne snage. No, bez obzira na sudstvo, policija ovom forenzičnom disciplinom dobiva moćno oruđe u svojem svakodnevnom operativnom radu.

LITERATURA

Forensic Media Services, Ltd. Denver, CO, USA. <http://www.forensicav.ro/software.htm> (pristupljeno 19. 2. 2016.)

Royal Institute of Technology (KTH), Department of Speech, Music and Hearing, Stockholm. <http://www.speech.kth.se/wavesurfer/> (pristupljeno 19. 2. 2016.)

University of Amsterdam, Phonetic Sciences. <http://www.fon.hum.uva.nl/praat/> - (pristupljeno 19. 2. 2016.)

Zlatko Kovač
Nikola Protrka

COMPUTER ANALYSIS OF THE HUMAN VOICE AND SPEECH

Summary

This article discusses the analysis of the human voice and speech based on the long-term average spectrum of speech (Long-Term Average Spectrum - LTAS), in terms of forensic analysis needed at the disputed and undisputed pattern of computer forensics and expertise. Two computer programs working in conjunction have been used: WaveSurfer and Catalina Forensic Audio Toolbox. The first program analyses the audio track and the results stored in the form of text files, and the second program loads these files, processes them statistically and displays graphically the results through a series of diagrams. Authors briefly describe both of the computer programs, their capabilities and working windows through analysis of a voice ('a' vocal), the basic parameters of the human voice (fundamental frequency, formant and long-term average spectrum). Then an analysis of two different voices (speech of two female persons of similar tonal characteristics) has been made and the results compared through presentation of relevant diagrams. Finally, an analysis has been performed involving a voice of the same speaker, where both the first and second half of the voice record have been analysed. The relevant charts are also compared to determine matching curves on graphs or determine possible deviation (as a variability measure in voice and speech of the same person). For the above analysis and comparison, the authors have uploaded the audio files on the Internet and provided link to any interested reader for download and further analysis.

Key words: *expert analysis, speech, voice analysis, LTAS.*