

NEGATIVNA KLASIFIKACIJA DIGITALNE INFORMACIJE BOJE GLASA

NEGATIVE CLASSIFICATION DIGITAL INFORMATION OF VOICE TONE

Tomislav JarmiĆ

Fakultet organizacije i informatike Varaždin, Sveučilište u Zagrebu, Varaždin

Sažetak

U tehničkom smislu ne postoji prepoznavanje osobe po boji glasa na razini čovjeka. Prepoznavanje glasa i prepoznavanje govora su pojmovi koji su sadržajno potpuno različiti. Svrha prepoznavanja govora je prepoznavanje riječi govornika, dok se prepoznavanjem glasa identificirao poznatu ili nepoznatu osobu. Boja glasa osobe [1] ključan je uzorak u glasu govornika. Može se promatrati kao subjektivni osjet glasnoće, visine i boje zvuka te njegove percepcije u prostoru. Cilj istraživanja je pronaći u zvučnom zapisu govornika karakteristične značajke kojima se može odrediti digitalna informacija boja glasa. Digitalnom informacijom boja glasa uz upotrebu informacijsko-komunikacijske tehnologije[2] prepoznajemo spol osobe, ciljanog govornika ili skupnu govornika. Namjera je do sada neiskorištene Walsh-Hadamardove koeficijente upotrijebiti za određivanje karakterističnih značajke. U procesu predprocesiranja izlučujemo karakteristične značajke za digitalnu informaciju boja glasa i to tako da se postupkom Kepstralne analize određuje osnovna frekvenciju F_0 za svaki zvučni zapis, temeljem koje se izračunavaju pripadajući harmonici. Dobivene uzorci klasificiraju se po spolu. U slijedećem koraku iz zvučnih zapisa izlučuju se Walsh-Hadamard-ovi koeficijenti, klasificiramo ih po spolu, te se uspoređuju rezultati klasifikacija. U praksi model se može koristiti za prepoznavanje unaprijed preddefinirane klasu govornika, odnosno za brzu eliminaciju ciljane skupine podataka.

Ključne riječi: digitalna informacija boja glasa, karakteristična značajka boja glasa, Walsh-Hadamard-ovi koeficijenti

Abstract

In a technical sense, does not exist a developed recognition of voice tone on a human level. Voice recognition and speech recognition are concepts that are completely different in content. The purpose of speech recognition is to recognize a speaker's speech, while recognizing a voice identifies the known person or an unknown person. The voice tone of a person is a key pattern in the voice of the speaker. It can be seen as a subjective sense of the volume, height and color of the sound and its perception in the space. The aim of the research is to find in the recorded voice of the speaker a characteristic feature that can determine the digital information of the of the voice tone. By using digital information of voice tone and using information-communication technology, we recognize a gender, a target speaker, or a group of speakers. Intention is to use Walsh-Hadamard's coefficients to determine the characteristic features. In the preprocessing process, we separate the characteristic features for digital information of the voice tone. With method of the Cepstral analysis determines the basic frequency F_0 for each recorded sound on which basisi the respective harmonics are calculated. The obtained samples are classified by gender. In the next step from the recorded sound, the Walsh-Hadamard coefficients are mark off, classified by gender, and the results of the classification are compared. In practice, the model is used to identify, or to quickly eliminate a target group of data.

Keywords: digital information of voice tone, improvement of voice tone characteristic, Walsh-hadamard coefficients

1. Uvod

1. Introduction

Predmet istraživanja rada je određivanje/unaprjeđenje karakterističnih značajki za digitalnu informaciju boje glasa. Negativna klasifikacija digitalne informacije boje glasa upotrebljava se u procesu brzog odabira ciljane skupine podataka. Današnji sustavi u stanju su identificirati ili autentificirati uzorak glasovnog zapisa što je dovoljno u kontekstu usporedbe 1:1 ili 1:N[3]. Kod procesa detekcije kome pripada taj uzorak, sustavi temeljeni na ovoj biometrijskoj karakteristici uglavnom su nemoćni. Digitalni uzorak boje glasa važan je za podsustav za predprocesiranje i proces klasifikacije. Primjena do sada u tu svrhu neiskorištenih Walsh-Hadamard-ovih koeficijenata[4] poboljšava proces klasifikacije glasovnog uzorka.

Različiti glasovi se razlikuju u visini tona, spektralnom obliku, količini energije i trajanju. Produkcijски i percepcijski organi za formiranje govora najefikasniji su na frekvencijama do otprilike 7 kHz. Glasaljke proizvode periodične zvučne titraje-zvučne glasove i aperiodične zvučne titraje-bezvučne glasove. Prosječna frekvencija titranja glasnica u govoru za odrasle muškarce je 120 Hz, za odrasle žene 220 Hz, a za djecu 300 Hz[5]. U pubertetu se muškarcima glasnice produže za oko 10 mm i glas im se po frekvenciji snizi za jednu oktavu, a ženama se produlje za 3-4 mm i glas im se snizi za oko trećinu oktave. Ove pretpostavke, koje razlikuju govornike, stvaraju preduvjete za provođenje klasifikacije osoba temeljem uzorka digitalne informacije boje glasa. Glas osobe ima stalnu i promjenjivu sastavnicu. Stalna sastavnica ovisi o organskim osobinama čovjeka, odnosno anatomsko-fiziološkim karakteristikama, a promjenjiva ovisi o izražajnoj mogućnosti samog čovjeka da upravlja rezonatorima i namješta ih s obzirom na željenu boju glasa. Taj perceptivni fenomen omogućuje razlikovanje govornika pa je moguće da ih upravo na temelju nje identificiramo. Boja glasa jedan je od ključnih čimbenika prepoznavanja ciljanog govornika. Osim što predstavlja govornika, boja glasa i naglasak-akcent[6], može predstavljati i veće skupine ljudi. Broj istraživanja unutar tog problemskog područja vrlo je malen.

Postojeći se rezultati uglavnom odnose na ispitivanja nekih osnovnih akustičkih značajki različitih populacija: osnovni laringealni ton ili fundamentalna frekvencija[7][8], intenzitet osnovnog laringealnog tona ili fundamentalne frekvencije[8], varijacija frekvencije osnovnog tona, a širih znanstvenih istraživanja temeljenih na proučavanju spektra složenog tona boje glasa nema.

Osobine koje opisuju boju glasa su [9]:

- Voluminoznost,
- Punoća,
- Zvonkost,
- Okruglost,
- Blještavost,
- Pucketavost,
- Nosnost,
- Kreštavost,
- Zaštitni glas, i
- Ugodan glas.

Sve osobine osim kreštavosti, zaštitnog glasa i ugodnog glasa koje su opisne, mogu se izraziti digitalnom informacijom pomoću osnovne frekvencije F_0 i pripadajućih harmonika $F_2 - F_n$ (Tablica 1). Svaka od osobina opisuje se osnovnom frekvencijom F_0 i harmonicima koji se nalaze u području spektra glasa specifičnom za svaku od osobina. Prosječna osnovna frekvencija F_0 kod muškaraca je oko 120 Hz, a kreće se u rasponu od 90 Hz do 210 Hz, što je specifično za svaku osobu posebno. Kod žena prosječna osnovna frekvencija F_0 je oko 220 Hz, a kreće se u rasponu od 139 Hz – 300 Hz. Broj harmonika koji opisuju karakteristične značajke za digitalnu informaciju boje glasa ovisi o vrijednosti osnovne frekvencije F_0 .

Boju glasa možemo odrediti sa osnovnom frekvencijom F_0 i pripadajućim harmonicima. Ljudski glas sadrži do 15 harmonika[10]. Osnovnu frekvenciju F_0 i harmonike izlučujemo iz govornog signala, te ih klasificiramo po spolu. To su poznati podaci i služe nam za kontrolu i uspoređivanje rezultata pri klasifikaciji Walsh-Hadamard-ovih koeficijenata.

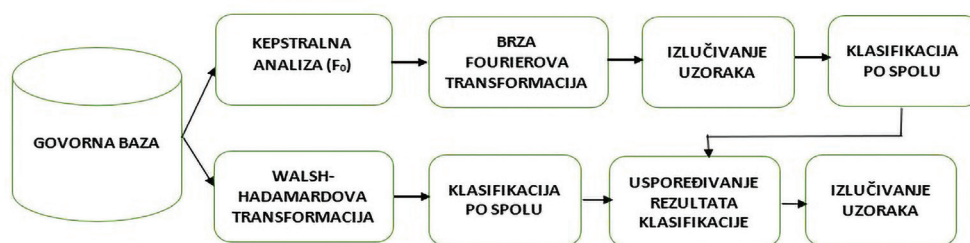
Tablica 1. Karakteristične značajke za digitalnu informaciju osobina kojim opisujemo boju glasa

Izvor: Osobine glasa i frekventijski spektar osobina glasa - Škarić, I.: Govorni dijelovi, III. dio. In: Babić, S. Et al: Povijesni pregled, glasovi i oblici hrvatskog jezika. Globus, Zagreb. 1991, Interpretacija digitalne informacije za muškarce i žene - Tomislav Jarmić

Table 1. Characteristic patterns for describing voice tone features.

Source: Features of voice and frequency spectrum of voice features - Škarić, I.: Speaking Parts, III. part. In: Babić, S. Et al: Historical overview, voices and forms of Croatian language. Globus, Zagreb. 1991 Interpretation for voice features in form of digital information for male and female - Tomislav Jarmic

Rbr.	Osobina glasa	Frekventijski spektar	Muškarci		Žene	
			F ₀ Hz	Harmonik	F ₀ Hz	Harmonik
1.	Volumioznost	ispod 800 Hz	120	F ₂ - F ₆	220	F ₂ - F ₃
2.	Punoća	400 Hz - 800 Hz	120	F ₃ - F ₆	220	F ₂ - F ₃
3.	Zvonkost	800 Hz - 2000 Hz	120	F ₆ - F ₁₆	220	F ₄ - F ₉
4.	Okruglost/ Grlenost	2000 Hz - 2500 Hz	120	F ₁₆ - F ₂₀	220	F ₉ - F ₁₁
5.	Blještavost	2500 Hz - 5000 Hz				
	Podzona 1 muškarci	2500 Hz - 3500 Hz	120	F ₂₀ - F ₂₉		
	Podzona 2 žene	3500 Hz - 5000 Hz			220	F ₁₁ - F ₂₂
6.	Pucketavost	5000 Hz - 10 000 Hz	120	F ₄₁ - F ₈₃	220	F ₂₂ - F ₄₅
7.	Nosnost	200 Hz - 400 Hz	120	F ₂ - F ₃	220	F ₂
		600 Hz - 1500 Hz	120	F ₅ - F ₁₂	220	F ₃ - F ₇
8.	Kreštavost	Opisne osobine				
9.	Zaštitni glas	Opisne osobine				
10.	Ugodan glas	Opisne osobine				



Slika 1 Dijagram toka podataka

Figure 1 Data Flow Diagram

Slika 1. prikazuje nam dijagram toka podataka i vidljivo je da upotreba Walsh-Hadamard-ovih koeficijenata skraćuje proces izlučivanja karakterističnih uzoraka za određivanje digitalne informacije boje glasa.

2. Govorna baza

2. Database of male and female voice recordings – „Speech“ database

Za provođenje eksperimenta snimljena je baza govornih zapisa. Govorna baza sadržava zvučne zapise različitih osoba, različitog spola i starosne dobi (M/Ž, od 21 – 60 godina). Zvučni zapis u trajanju od oko 40 sekundi sadrži tekst:

„Danas je putem interneta moguće pronaći veliku ponudu različitih detektora laži, uređaja i računalnih programa za procjenu stresa u glasu. Prilikom njihove nabave treba obratiti pažnju na barem dva kriterija da bi dobili što pouzdaniju opremu. Kod ljudi u otrovanom stanju stupanj trovanja i tip opojnog sredstva može biti povezan sa onim "što" i "kako" govore. Postoje dvije različite funkcionalne kategorije komunikacije: govor i motorička produkcija govora.“

Govorna baza snimljena je Vidicode Quatro-Octo snimačem koji se nalazi postavljen u Skupštinskoj dvorani Varaždinske županije i služi za profesionalno snimanje sjednica Skupštine Varaždinske županije.



Slika 2 Snimač razgovora Vidicode Quatro-Octo

Izvor: http://www.pelcom.hr/public_html/call_rec/silverLine.htm

Figure 2 Voice recorder Vidicode Quatro-Octo

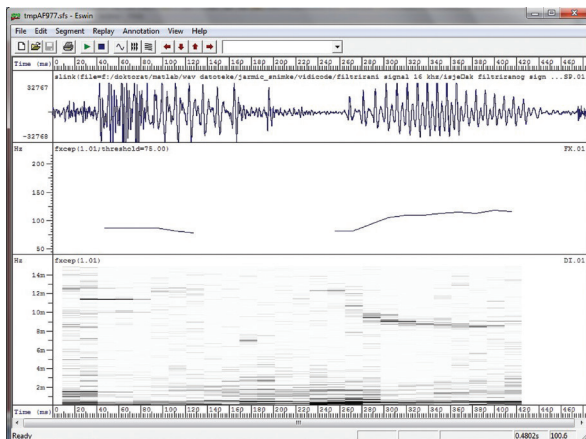
Source: http://www.pelcom.hr/public_html/call_rec/silverLine.htm

Uređaj ima sve potrebne profesionalne audio priključke kao i priključak na računalnu mrežu, te je vrlo jednostavno snimljeni zvučni zapis pohraniti na računalo. Zvučni zapis snima se pulsno kodnom modulacijom (PCM) s frekvencijom uzorkovanja f_s od 44100 Hz, duljinom riječi od 16 bita, u dvokanalnom zapisu (stereo) u *.wav formatu. U provođenju daljnje analize filtrira se zvučni zapis, te se time izbjegavaju smetnje koje mogu nastati utjecajem mrežnog napona. Također, ograničavamo područje frekventijskog spektra na 8 kHz, u kojem su percepcijski organi za formiranje govora najefikasniji.

3. Osnovna frekvencija F_0 i harmonici

3. Basic frequency F_0 and harmonics

Programskom podrškom SFSWin, pomoću funkcije „fxcep“ cepstralnom analizom određujemo osnovnu frekvenciju F_0 na isječku signala zvučnog zapisa u trajanju od 40 ms do 50 ms.



Slika 3 Određivanje osnovne ili fundamentalne frekvencije F_0 cepstralnom analizom – funkcija fxcep

Izvor: Programska podrška SFSWin

Figure 3 Determining of fundamental frequency F_0 with Cepstrum analysis - function fxcep, Source: Software SFSWin

Osnovna frekvencija F_0 je srednja vrijednost dobivenih osnovnih frekvencija f_0 i računa se prema izrazu (1).

$$F_0 = \frac{F_{0,1} + F_{0,2} + F_{0,3} + \dots + F_{0,N}}{N} \quad (1)$$

Uz osnovnu ili fundamentalnu frekvenciju F_0 određuju se pripadajući harmonici. Harmonici su višekratnici osnovne frekvencije F_0 , i računaju se prema izrazu (3).

$$a_1 \sin 2\pi f_1 t + a_2 \sin 2\pi f_2 t + a_3 \sin 2\pi f_3 t + \dots \quad (2)$$

$$\text{gdje je: } f_1 = 1f_0, f_2 = 2f_0, f_3 = 3f_0 \dots \quad (3)$$

Tablica 2. Vrijednosti osnovne ili fundamentalne frekvencije F_0 zvučnih zapisa glasova muških i ženskih osoba

Izvor: Podaci dobiveni eksperimentom

Table 2. Basic frequency values F_0 of recorded voices for male and female persons

Source: Data obtained by experiment

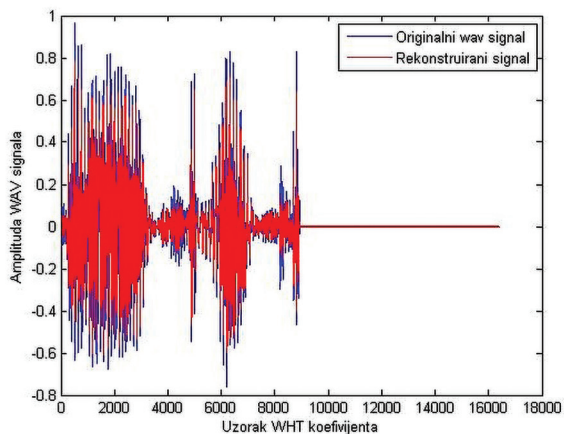
Muškarci	F_0 Hz
MRFSI1	114,00
MRFSI2	90,31
MRFSI3	107,47
MRFSI11	136,65
MVFSI11	206,68
MVFSI21	120,10
MVFSIDARKO12	100,86
MVFSIJURA2	138,17
MVFSIKELEMEN15	137,64
MVFSITEZACKI17	210,91
MVFSIVEDRAN1	125,38
MVFSIVINCEK6	149,48
MVFSIZVJEZDA7	119,22
MVFSIZVONIMIR5	163,10

Žene	F_0 Hz
ZVFSIURZUP4	222,40
ZRFIS1	175,20
ZRFIS2	172,23
ZRFIS3	139,78
ZRFIS4	159,83
ZVFIS10	252,20
ZVFIS17	207,72
ZVFIS18	235,21
ZVFIS19	203,79
ZVFIS20	218,20

4. Walsh-Hadamard-ovi koeficijenti

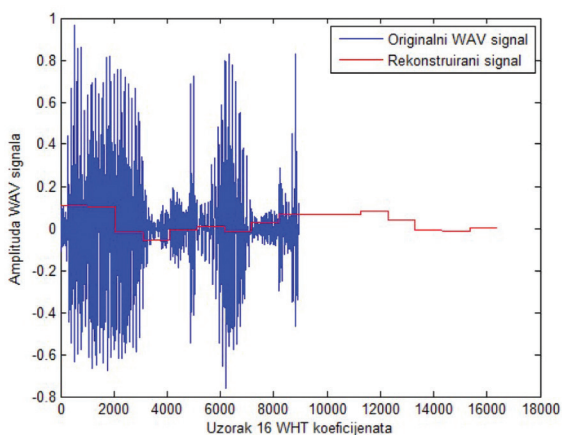
4. Walsh-Hadamard coefficients

Walsh-ove transformacije daju drugačije ali vrlo slične rezultate Fourier-ovoj transformaciji. Brza Walsh-Hadamard-ova transformacija daje spektralnu analizu isječka govornog signala koja se prikazuje koeficijentima. Koeficijenti sadrže karakteristične značajke i čine digitalnu informaciju boje glasa. Za rekonstrukcija WAV signala inverznom brzom Walsh-Hadamard-ovom transformacijom koristi se 2050, 16, 32 i 64 WHT koeficijenata.



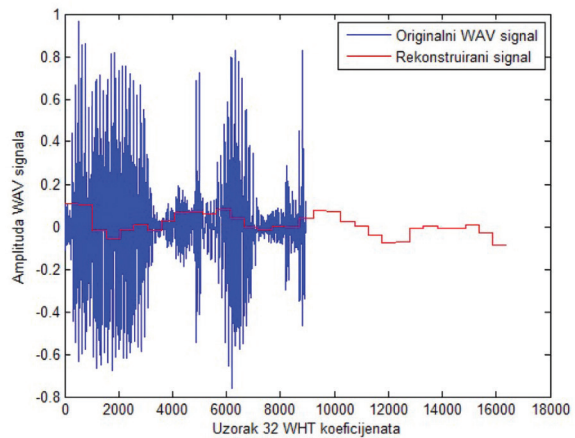
Slika 4 Rekonstrukcija originalnog signala (filtrirani isječak MVFISRKELEMEN16)
Izvor: Programska podrška „MATLAB R2013a“

Figure 4 Reconstruction of the original signal (filtered clip of male voice signal MVFISRKELEMEN16), Source: Software „MATLAB R2013a“



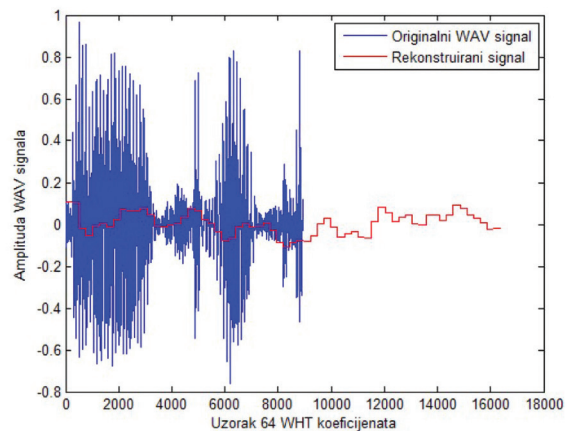
Slika 5 Rekonstrukcija originalnog signala sa 16 WHT koeficijenta; (filtrirani isječak MVFISRKELEMEN16)
Izvor: Programska podrška „MATLAB R2013a“

Figure 5 Reconstruction of the original signal with 16 WHT coefficients (filtered clip of male voice signal MVFISRKELEMEN16), Source: Software „MATLAB R2013a“



Slika 6 Rekonstrukcija originalnog signala sa 32 WHT koeficijenta (filtrirani isječak MVFISRKELEMEN16)
Izvor: Programska podrška „MATLAB R2013a“

Figure 6 Reconstruction of the original signal with 32 WHT coefficients (filtered clip of male voice signal MVFISRKELEMEN16), Source: Software „MATLAB R2013a“



Slika 7 Rekonstrukcija originalnog signala sa 64 WHT koeficijenta (filtrirani isječak MVFISRKELEMEN16)
Izvor: Programska podrška „MATLAB R2013a“

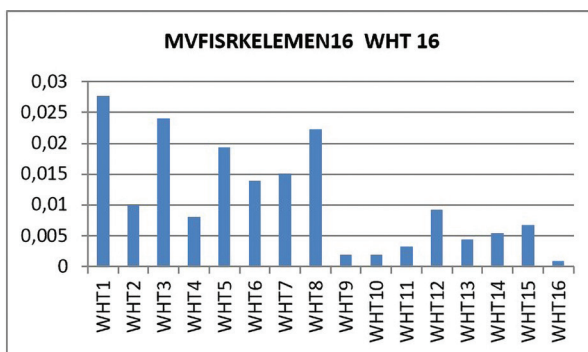
Figure 7 Reconstruction of the original signal with 64 WHT coefficients (filtered clip of male voice signal MVFISRKELEMEN16), Source: Software „MATLAB R2013a“

Na slici 4. prikazana je rekonstrukcija signala zvučnog zapisa inverznom Walsh-Hadamard-ovom transformacijom. Slike 5., 6., i 7. prikazuju rekonstrukciju signala zvučnog zapisa govora inverznom Walsh-Hadamard-ovom transformacijom sa 16, 32 i 64 koeficijenta. Originalni i rekonstruirani signal bitno se razlikuju.

Tablica 3. Vrijednosti WHT koeficijena (WHT 16) zvučnih zapisa slučajno odabranih muških osoba

Table 3. Values of WHT coefficient (WHT 16) for randomly chosen male persons

	MVFISRKELEMEN16	MRFSI2	MRFSI3	MRFSI11	MVFSI11	MVFSI21	MVFSIDARKO12
WHT1	0,0277	-0,0427	-0,0397	0,0276	-0,0162	0,0124	-0,0055
WHT2	-0,0100	0,0353	0,0226	-0,0067	0,0048	0,0000	-0,0009
WHT3	0,0241	-0,0292	-0,0105	0,0048	-0,0066	0,0028	0,0019
WHT4	-0,0081	0,0002	0,0011	0,0063	0,0005	0,0005	0,0027
WHT5	0,0194	-0,0101	0,0145	0,0029	0,0126	0,0019	-0,0119
WHT6	0,0139	0,0010	0,0008	-0,0006	0,0002	-0,0006	-0,0005
WHT7	0,0151	-0,0427	-0,0397	0,0276	-0,0162	0,0124	-0,0055
WHT8	0,0223	0,0353	0,0226	-0,0067	0,0048	0,0000	-0,0009
WHT9	0,0020	-0,0292	-0,0105	0,0048	-0,0066	0,0028	0,0019
WHT10	-0,0020	0,0002	0,0011	0,0063	0,0005	0,0005	0,0027
WHT11	0,0033	-0,0101	0,0145	0,0029	0,0126	0,0019	-0,0119
WHT12	0,0093	0,0010	0,0008	-0,0006	0,0002	-0,0006	-0,0005
WHT13	0,0044	-0,0427	-0,0397	0,0276	-0,0162	0,0124	-0,0055
WHT14	-0,0054	0,0353	0,0226	-0,0067	0,0048	0,0000	-0,0009
WHT15	-0,0067	-0,0292	-0,0105	0,0048	-0,0066	0,0028	0,0019
WHT16	-0,0009	0,0002	0,0011	0,0063	0,0005	0,0005	0,0027



Slika 8 Graf boje glasa muškog zapisa sa 16 WHT koeficijena (apsolutne vrijednosti)

Figure 8 Graph for recorded male voice tone of 16 WHAT coefficients (absolute values)

5. Klasifikacija

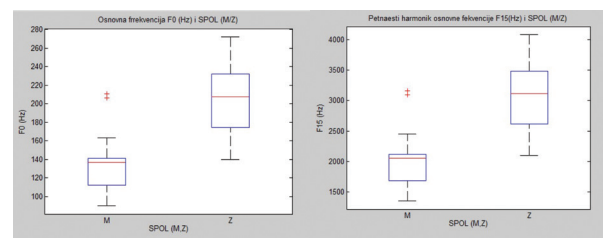
5. Classification

5.1. Kutijasti dijagram – Box-plot

5.1. Box Diagram - Box-plot

U procesu određivanja karakterističnih uzoraka za digitalnu informaciju boje glasa dobivene vrijednosti za osnovne frekvencije F_0 , harmonike $F_2 - F_n$ i Walsh-Hadamard-ove koeficijente WHT 16, 32 i 64 promatramo kroz Kutijasti dijagram (Box-plot). Izgled kutijastog dijagrama ukazuje na stupanj raspršenosti i asimetričnosti, te može pokazati vrijednosti koje odstupaju.

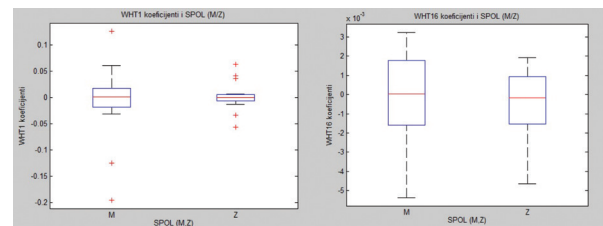
Na slici 9. prikazani su kutijasti dijagrami osnovnih frekvencija F_0 i harmonika F_{15} u odnosu na spol (M/Ž), dok su na slici 10. prikazani kutijasti dijagrami Walsh-Hadamard-ovih koeficijena WHT 1 i 16, u odnosu na spol (M/Ž).



Slika 9 Kutijasti dijagram osnovnih frekvencija F_0 i harmonika F_{15} u odnosu na spol (M/Ž)

Izvor: Programaska podrška „MATLAB R2013a“

Figure 9 Box-plot diagram of basic frequencies F_0 and harmonics F_{15} in relation to gender (male / female), Source: Software "MATLAB R2013a"



Slika 10 Kutijasti dijagram WHT koeficijena u odnosu na spol (M/Ž) (WHT1 i 16)

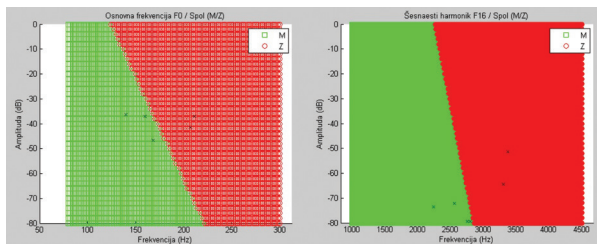
Izvor: Programaska podrška „MATLAB R2013a“

Figure 10 Box-plot diagram of WHT coefficients (WHT 1-16) in relation to gender (male / female), Source: Software "MATLAB R2013a"

Slika 9. pokazuje dobru distribuciju podatka u donosu na spol (M/Ž), odnosno jasno pokazuje odvajanje grupa podataka po spolu te vrlo mali dio podataka koji se nalazi izvan područja kutijastog dijagrama. Na slici 10. distribucija podataka pokazuje grupiranje podataka unutar određenog područja minimuma i maksimuma kutijastog dijagrama, a veći dio podataka nalazi se izvan područja kutijastog dijagrama.

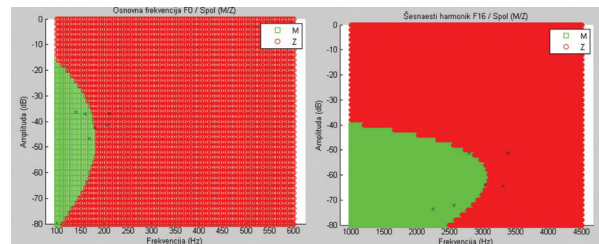
5.2. Linearna Diskriminacijska Analiza (LDA) i Kvadratna Diskriminacijska Analiza (QDA)
5.2. Linear Discrimination Analysis (LDA) and Quadratic Discrimination Analysis (QDA)

U daljnjem postupku, dobivene vrijednosti za osnovne frekvencije F_0 , harmonike $F_2 - F_n$ i Walsh-Hadamard-ovi koeficijenti klasificiraju se Linearnom diskriminacijskom analizom (LDA), Kvadratnom diskriminacijskom analizom (QDA), te Naivnim Bayes-ovim klasifikatorom s Gaussovom distribucijom i Kernel distribucijom.



Slika 11 LDA klasifikacija F_0 /SPOL (M/Ž), F_{16} /SPOL (M/Ž), x- krivo klasificirani podaci, Izvor: Programska podrška „MATLAB R2013a“

Figure 11 LDA classification F_0 /gender (male/female), F_{16} /gender (male/female), x- bad classified data Source: Software "MATLAB R2013a"



Slika 12 QDA klasifikacija F_0 /SPOL (M/Ž), F_{16} /SPOL (M/Ž), x- krivo klasificirani podaci Izvor: Programska podrška „MATLAB R2013a“

Figure 12 QDA classification F_0 /gender (male/female), F_{16} /gender (male/female), x- bad classified data Source: Software "MATLAB R2013a"

Tablica 4. Linearna Diskriminacijska Analiza (LDA) osnovne frekvencije F_0 i harmonika F_2-F_{16} , sa postotkom loše klasificiranih podataka

Table 4. Linear Discrimination Analysis (LDA) of the basic frequency F_0 and the harmonics F_2-F_{16} with the percentage of bad classified data

M=73 Ž=79		Linearna Diskriminacijska Analiza (LDA)																															
		Osnovne frekvencije F_0 i harmonika F_2-F_{16} (Hz) / Amplituda (dB) – M-Ž																															
BROJ UZORAKA		152		152		152		152		152		152		152		152		152															
REDISTRIBUCIJSKA POGREŠKA		0,153		0,1184		0,1711		0,1513		0,1513		0,1513		0,1447		0,1711		0,1513		0,1513		0,0921		0,1513		0,1513		0,1250		0,1711			
KONFIZIJSKA MATRICA		M		63 10		63 10		63 10		63 10		63 10		63 10		63 10		63 10		63 10		63 10		63 10		63 10		63 10		63 10			
		Ž		13 66		8 71		16 63		13 66		13 66		13 66		12 67		16 63		13 66		13 66		4 75		13 66		13 66		9 70		16 63	
LOŠE KLASIFIRIRANI UZORCI		23		18		26		23		23		23		22		26		23		23		23		14		23		23		19		26	
POSTOTAK LOŠE KLASIFIRIRANIH UZORAKA		%		15,13		11,84		17,11		15,13		15,13		15,13		14,17		17,11		15,13		15,13		9,21		15,13		15,13		12,50		17,11	

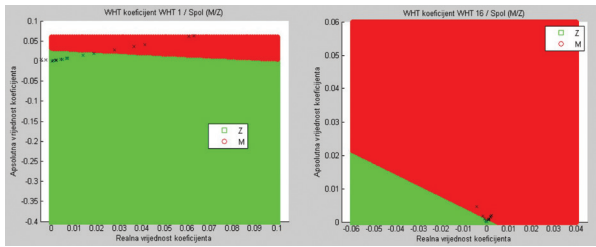
*KONFIZIJSKA MATRICA dijagonala matrice pokazuje dobru klasifikaciju

Tablica 5. Kvadratna Diskriminacijska Analiza (QDA) osnovne frekvencije F_0 i harmonika F_2-F_{16} , sa postotkom loše klasificiranih podataka

Table 5. Quadratic Discrimination Analysis (QDA) of the basic frequency F_0 and the harmonics F_2-F_{16} with the percentage of bad classified data

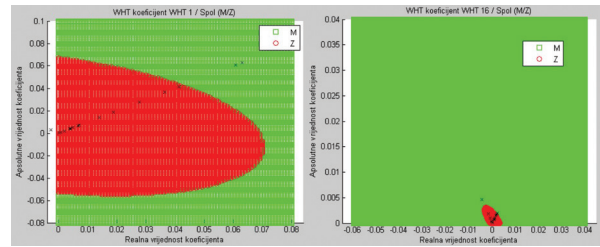
M=73 Ž=79		Kvadratna Diskriminacijska Analiza (QDA)																															
		Osnovne frekvencije F_0 i harmonika F_2-F_{16} (Hz) / Amplituda (dB) – M-Ž																															
BROJ UZORAKA		152		152		152		152		152		152		152		152		152															
REDISTRIBUCIJSKA POGREŠKA		0,1842		0,1645		0,1250		0,1974		0,1513		0,1645		0,1447		0,1513		0,1513		0,1974		0,0592		0,1974		0,1184		0,0921		0,1579			
KONFIZIJSKA MATRICA		M		63 10		63 10		59 14		59 14		59 14		63 10		58 15		59 14		63 10		68 5		63 10		63 10		63 10		63 10			
		Ž		18 61		15 64		9 70		16 63		9 70		11 68		12 67		8 71		9 70		20 59		13 66		4 75		20 59		8 71		4 75	
LOŠE KLASIFIRIRANI UZORCI		28		25		19		30		23		25		22		23		23		30		23		9		30		18		14		24	
POSTOTAK LOŠE KLASIFIRIRANIH UZORAKA		%		18,42		16,45		12,50		19,74		15,13		16,45		14,47		15,13		15,13		19,74		5,92		19,74		11,84		9,21		15,79	

*KONFIZIJSKA MATRICA dijagonala matrice pokazuje dobru klasifikaciju



Slika 13 LDA klasifikacija WHT koeficijenata za WHT1/SPOL (M/Ž), WHT16/SPOL (M/Ž), x- krivo klasificirani podaci
Izvor: Programska podrška „MATLAB R2013a“

Figure 13 LDA classification of WHT coefficients WHT 1 / gender (male/female), WHT16/gender (male/female), x- bad classified data Source: Software "MATLAB R2013a"



Slika 14 QDA klasifikacija WHT koeficijenata WHT1/SPOL (M/Ž), WHT16/SPOL (M/Ž), x- krivo klasificirani podaci
Izvor: Programska podrška „MATLAB R2013a“

Figure 14 QDA classification F0/gender (male/female), F16/gender (male/female), x- bad classified data Source: Software "MATLAB R2013a"

Tablica 6. Linearna Diskriminacijska Analiza (LDA) Walsh Hadamard-ovih koeficijenata WHT 1 – WHT 16 / SPOL (M-Ž)

Table 6. Linear Discrimination Analysis (LDA) of Walsh Hadamard coefficients WHT 1 - WHT 16 / gender (male/female)

M=73 Ž=79		Linearna Diskriminacijska Analiza (LDA)																															
		Walsh Hadamardovih koeficijenata WHT 1 – WHT 16 (realne : apsolutne vrijednosti) / SPOL (M-Ž)																															
BROJ UZORAKA		WHT1	WHT2	WHT3	WHT4	WHT5	WHT6	WHT7	WHT8	WHT9	WHT10	WHT11	WHT12	WHT13	WHT14	WHT15	WHT16																
REDISTRIBUCIJSKA POGREŠKA		0.4474	0.4408	0.3947	0.4211	0.4013	0.4737	0.4013	0.4868	0.4276	0.4737	0.3553	0.4408	0.3750	0.2895	0.3487	0.3947																
KONFIZIJSKA MATRICA	M	28	45	26	47	34	39	24	49	33	40	25	48	37	36	20	53	35	38	21	52	43	30	21	52	31	42	39	34	30	43	32	41
	Z	23	56	20	59	21	58	15	64	21	58	24	55	25	54	21	58	27	52	20	59	24	55	15	64	15	64	10	69	10	69	19	60
LOŠE KLASIFIRANI UZORCI		68	67	60	64	61	72	61	74	65	72	54	67	57	44	53	60																
POSTOTAK LOŠE KLASIFIRANIH UZORAKA		%	44,74	44,08	39,47	42,11	40,13	47,37	40,13	48,68	42,76	47,37	35,53	44,08	37,50	28,95	34,87	39,47															

*KONFIZIJSKA MATRICA dijagonala matrice pokazuje dobru klasifikaciju

Tablica 7. Kvadratna Diskriminacijska Analiza (QDA) Walsh Hadamard-ovih koeficijenata WHT1 – WHT16/ SPOL (M-Ž)

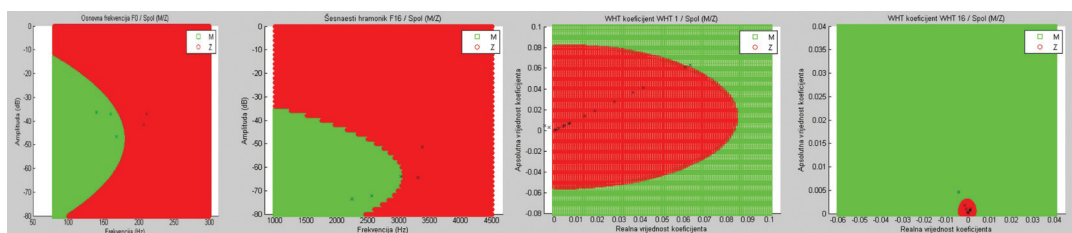
Table 7. Quadratic Discrimination Analysis (QDA) of Walsh Hadamard coefficients WHT 1 - WHT 16 / gender (male/female)

M=73 Ž=79		Kvadratna Diskriminacijska Analiza (QDA)																															
		Walsh Hadamardovih koeficijenata WHT 1 – WHT 16 (realne : apsolutne vrijednosti) / SPOL (M-Ž)																															
BROJ UZORAKA		WHT1	WHT2	WHT3	WHT4	WHT5	WHT6	WHT7	WHT8	WHT9	WHT10	WHT11	WHT12	WHT13	WHT14	WHT15	WHT16																
REDISTRIBUCIJSKA POGREŠKA		0.4276	0.4211	0.3750	0.4211	0.3947	0.4145	0.4342	0.4605	0.4408	0.4408	0.3553	0.4145	0.3487	0.2895	0.3553	0.3947																
KONFIZIJSKA MATRICA	M	17	56	25	48	30	43	14	59	24	49	22	51	26	47	17	56	26	47	21	52	33	40	20	53	30	43	39	34	29	44	32	41
	Z	9	70	16	63	14	65	5	74	11	68	12	67	19	60	14	65	20	59	15	64	14	65	10	69	10	69	10	69	10	69	19	60
LOŠE KLASIFIRANI UZORCI		65	64	57	64	60	63	66	70	67	67	54	63	53	44	54	60																
POSTOTAK LOŠE KLASIFIRANIH UZORAKA		%	42,76	42,11	37,50	42,11	39,47	41,45	43,42	46,05	44,08	44,08	35,53	41,45	34,87	28,95	35,53	39,47															

*KONFIZIJSKA MATRICA dijagonala matrice pokazuje dobru klasifikaciju

5.3. Naivni Bayesov Klasifikator sa Gausovom distribucijom

5.3. Naive Bayesian Classification with Gaussian Distribution



Slika 15 Klasifikacija F_0 i F_{16} /SPOL (M/Ž) i WHT 1 i 16 (M/Ž) (Naivnim Bayesovim Klasifikatorom sa Gausovom distribucijom, x- krivo klasificirani podaci; Izvor: Programska podrška „MATLAB R2013a“

Figure 15 Classification of basic frequencies F_0 and harmonics F_{16} by gender (male/female) and WHT coefficients WHT 1 and WHT 16 by gender (male/female) with Naive Bayesian Classification with Gaussian distribution, x- bad classified data Source: Software "MATLAB R2013a"

Tablica 8. Klasifikacija Naivnim Bayesovim Klasifikatorom sa Gausovom distribucijom osnovne frekvencije F_0 i harmonika F_2-F_{16} sa postotkom loše klasificiranih podataka

Table 8. Classification of basic frequency F_0 and harmonics F_2-F_{16} with Naive Bayesian Classifier with Gaussian distribution and percentage of bad classified data

M=73 Ž=79	Naivni Bayesov Klasifikator (Gausova distribucija)																
	Osnovne frekvencije F_0 i harmonika F_2-F_{16} (Hz) / Amplituda (dB) – M-Ž																
	F_0	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}	
BROJ UZORAKA	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	
REDISTRIBUCIJSKA POGREŠKA	0,1513	0,1645	0,1513	0,2237	0,1513	0,1645	0,1184	0,1842	0,1513	0,1974	0,1513	0,0592	0,1447	0,1513	0,0921	0,1382	
KROSVALIDACIJSKA POGREŠKA	0,1974	0,1579	0,1447	0,2237	0,1513	0,1645	0,1382	0,1776	0,1513	0,2039	0,1513	0,0592	0,1645	0,1711	0,0921	0,1513	
KONFIZIJSKA MATRICA	M Ž	63 10 13 66	63 10 15 64	63 10 13 66	59 14 20 59	59 14 9 70	59 14 11 68	63 10 8 71	58 15 13 66	59 14 9 70	63 10 20 59	63 10 13 66	68 5 4 75	63 10 12 67	63 10 13 66	63 10 4 75	63 10 11 68
LOŠE KLASIFIRANI UZORCI	23	25	23	34	23	25	18	28	23	30	23	9	22	23	14	21	
POSTOTAK LOŠE KLASIFIRANIH UZORAKA	%	15,13	16,45	15,13	22,37	15,13	16,45	11,84	18,42	15,13	19,74	15,13	5,92	14,47	15,13	9,21	13,82

*KONFIZIJSKA MATRICA dijagonala matrice pokazuje dobru klasifikaciju

Tablica 9. Klasifikacija Walsh Hadamard-ovih koeficijenata Naivnim Bayesovim Klasifikatorom s Gausovom distribucijom WHT1 -16 sa postotkom loše klasificiranih podataka

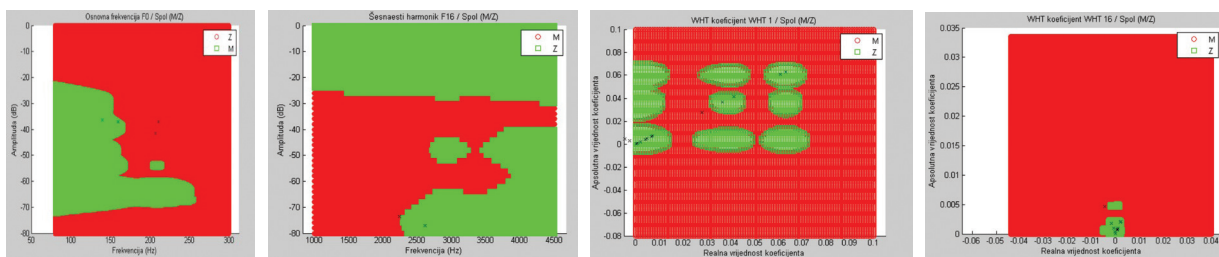
Table 9. Classification of Walsh Hadamard coefficients with Naive Bayesian Classification with Gaussian distribution WHT1 -16 and percentage of bad classified data

M=73 Ž=79	Naivni Bayesov Klasifikator (Gausova distribucija)																
	Walsh Hadamardovih koeficijenata WHT 1 – WHT 16 (realne : apsolutne vrijednosti) / SPOL (M-Ž)																
	WHT1	WHT2	WHT3	WHT4	WHT5	WHT6	WHT7	WHT8	WHT9	WHT10	WHT11	WHT12	WHT13	WHT14	WHT15	WHT16	
BROJ UZORAKA	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	152	
REDISTRIBUCIJSKA POGREŠKA	0,4605	0,4013	0,3750	0,4079	0,3882	0,4145	0,4342	0,4408	0,4211	0,4079	0,3289	0,4474	0,3750	0,3158	0,3487	0,3026	
KROSVALIDACIJSKA POGREŠKA	0,4539	0,4079	0,3750	0,4079	0,3882	0,4737	0,4408	0,4408	0,4211	0,4803	0,3355	0,4474	0,3750	0,3158	0,3553	0,3355	
KONFIZIJSKA MATRICA	M Ž	10 63 7 72	24 49 12 67	30 43 14 65	22 51 11 68	22 51 8 71	22 51 12 67	26 47 19 60	16 57 10 69	25 48 16 63	31 42 20 59	43 30 20 59	20 53 15 64	31 42 15 64	40 33 15 64	30 43 10 69	32 41 5 74
LOŠE KLASIFIRANI UZORCI	70	61	57	66	59	63	66	67	64	62	50	68	57	48	53	46	
POSTOTAK LOŠE KLASIFIRANIH UZORAKA	%	46,05	40,13	37,50	40,79	38,82	41,45	43,42	44,08	42,11	40,79	32,09	44,74	37,50	31,58	34,87	30,26

*KONFIZIJSKA MATRICA dijagonala matrice pokazuje dobru klasifikaciju

5.4. Naivni Bayesov Klasifikator sa Kernel distribucijom

5.4. Naive Bayesian Classification with Kernel Distribution



Slika 16. Klasifikacija F_0 i F_{16} /SPOL (M/Ž) i WHT 1 i 16 (M/Ž) Naivnim Bayesovim Klasifikatorom s „Kernel“ distribucijom, x- krivo klasificirani podaci

Izvor: Programaska podrška „MATLAB R2013a“

Figure 16 Classification of basic frequencies F_0 and harmonics F_{16} by gender (male/female) and WHT coefficients WHT 1 and WHT 16 by gender (male/female) with Naive Bayesian Classification with Kernel distribution, x- bad classified data Source: Software "MATLAB R2013a"

Tablica 10. Klasifikacija Naivnim Bayesovim Klasifikatorom sa Kernel distribucijom osnovne frekvencije F_0 i harmonika F_2 - F_{16} sa postotkom loše klasificiranih podataka

Table 10. Classification of basic frequency F_0 and harmonics F_2 - F_{16} with Naive Bayesian Classifier with Kernel distribution and percentage of bad classified data

M=73 Ž=79		Naivni Bayesov Klasifikator s Kernel distribucijom																														
		Osnovne frekvencije F_0 i harmonika F_2 - F_{16} (Hz) / Amplituda (dB) – M-Ž																														
		F0	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16															
BROJ UZORAKA		152	152	152	152	152	152	152	152	152	152	152	152	152	152	152																
REDISTRIBUCIJSKA POGREŠKA		0,1184	0,0855	0,0526	0,0921	0,1184	0,0754	0,1118	0,1118	0,0921	0,0921	0,1184	0,0789	0,0789	0,0526	0,1118	0,0526															
KROSVVALIDACIJSKA POGREŠKA		0,1579	0,1579	0,0526	0,0921	0,0921	0,0592	0,1118	0,1579	0,1250	0,0921	0,1184	0,0789	0,0789	0,0526	0,0987	0,0658															
KONFIZIJSKA MATRICA	M	63	10	68	5	69	4	59	14	59	14	69	4	64	9	63	10	59	14	68	5	64	9	68	5	63	10	69	4			
	Ž	8	71	8	71	4	75	0	79	4	75	7	72	8	71	7	72	0	79	9	70	4	75	7	72	3	78	3	76	7	72	4
LOŠE KLASIFIRANI UZORCI		18	13	8	14	18	11	17	17	14	14	18	12	12	8	17	8															
POSTOTAK LOŠE KLASIFIRANIH UZORAKA		%	11,84	8,55	5,26	9,21	11,84	7,54	11,18	11,18	9,21	9,21	11,84	7,99	7,99	5,26	11,18	5,26														

*KONFIZIJSKA MATRICA dijagonala matrice pokazuje dobru klasifikaciju

Tablica 11. Klasifikacija Walsh Hadamard-ovih koeficijenata Naivnim Bayesovim Klasifikatorom sa „Kernel“ distribucijom WHT1-16 sa postotkom loše klasificiranih podataka

Table 11. Classification of Walsh Hadamard coefficients with Naive Bayesian Classification with Kernel distribution WHT1 -16 and percentage of bad classified data

M=73 Ž=79		Naivni Bayesov Klasifikator s Kernel distribucijom																															
		Walsh Hadamardovih koeficijenata WHT 1 – WHT 16 (realne : apsolutne vrijednosti) / SPOL (M-Ž)																															
		WHT1	WHT2	WHT3	WHT4	WHT5	WHT6	WHT7	WHT8	WHT9	WHT10	WHT11	WHT12	WHT13	WHT14	WHT15	WHT16																
BROJ UZORAKA		152	152	152	152	152	152	152	152	152	152	152	152	152	152	152																	
REDISTRIBUCIJSKA POGREŠKA		0,3684	0,4013	0,3750	0,3816	0,3882	0,4079	0,3684	0,4013	0,4013	0,2961	0,2237	0,2895	0,3421	0,2697	0,3487	0,3750																
KROSVVALIDACIJSKA POGREŠKA		0,4145	0,4013	0,3750	0,4013	0,4276	0,4079	0,4145	0,4211	0,4211	0,3092	0,2500	0,3158	0,3421	0,2697	0,3553	0,3750																
KONFIZIJSKA MATRICA	M	25	48	23	60	30	43	29	44	22	51	30	43	35	38	23	50	23	50	28	45	53	20	29	44	31	42	42	31	30	43	21	52
	Ž	8	71	11	68	14	65	14	65	8	71	19	60	18	61	11	68	11	68	0	79	14	65	0	79	10	69	10	69	10	69	5	74
LOŠE KLASIFIRANI UZORCI		56	61	57	58	59	62	56	61	61	45	34	44	52	41	53	57																
POSTOTAK LOŠE KLASIFIRANIH UZORAKA		%	36,84	40,13	37,50	38,16	38,82	40,79	36,84	40,13	40,13	29,61	22,37	28,95	34,21	26,97	34,87	37,50															

*KONFIZIJSKA MATRICA dijagonala matrice pokazuje dobru klasifikaciju

Tablica 12. Usporedni rezultati klasifikacije uzoraka osnovne frekvencije F_0 , harmonika F_2 - F_{16} i Walsh-Hadamard-ovih koeficijenata Linearnom Diskriminacijskom Analizom, Kvadratnom Diskriminacijskom Analizom, Naivnim Bayesovim klasifikatorom sa Gausovom i Kernel distribucijom

Table 12. Comparative results for classification of samples of basic frequency F_0 , harmonics F_2 - F_{16} and Walsh-Hadamard coefficients with Linear Discrimination Analysis, Quadratic Discrimination Analysis, Naive Bayesian Classification with Gaussian and Kernel distribution

REZULTATI KLASIFIKACIJE KARAKTERISTIČNIH ZNAČAJKI				
KARAKTERISTIČNE ZNAČAJKE	LINEARNA DISKRIMINACIJSKA ANALIZA (LDA)	KVADRATNA DISKRIMINACIJSKA ANALIZA (QDA)	NAIVNI BAYESOV KLASIFIKATOR SA GAUSOVOM DISTRBUCIJOM	NAIVNI BAYESOV KLASIFIKATOR SA KERNEL DISTRIBUCIJOM
$F_0, F_2 - F_{16}$	85,30 %	84,95 %	85,03 %	90,98 %
F_0, F_2, F_{12}, F_{15}	87,83 %			
$F_0, F_1, F_{12}, F_{14}, F_{15}$		88,42 %		
$F_0, F_3, F_7, F_{12}, F_{15}, F_{16}$			88,16 %	
$F_0, F_2, F_3, F_4, F_9, F_{10}, F_{12}, F_{13}, F_{14}, F_{16}$				92,04 %
WHT 1-16	58,92 %	60,07 %	60,86 %	64,76 %
WHT10, 11, 12, 14			67,80 %	72,98 %
WHT 1-32	57,37 %	54,40 %	55,28 %	68,81 %
WHT4, 5, 6, 25,26, 28, 30				74,62%
WHT 1-64	54,53 %	55,54 %	54,31 %	71,50 %
WHT1, 8, 11,12, 14, 15, 18, 21, 22, 26, 28,29,32, 33, 35, 36, 37, 50, 52, 56,60, 64				79,05 %

6. Osvrt na dobivene rezultate istraživanja

6. Review of the research results

Metodama Linearne Diskriminacijske Analize, Kvadratne Diskriminacijske Analize, Naivnim Bayesovim kalsifikatorom sa Gausovom i Kernel distribucijom klasificirane su karakteristične značajke: osnovna frekvencija F_0 sa pripadajućim harmonicima i Walsh-Hadamard-ovi koeficijenti (Tablica 12.). Osnovna frekvencija F_0 sa pripadajućim harmonicima dobiva se u procesu predprocesiranja poznatim metodama. Rezultate njihove klasifikacije koristimo za uspoređivanje s rezultatima klasifikacije Walsh-Hadamard-ovih koeficijenta i izlučivanje karakterističnih značajki digitalne informacije boje glasa. Svaka od korištenih metoda klasifikacije daje nam različite rezultate i upućuje na određene karakteristične značajke. Klasifikacija osnovne frekvencije F_0 sa 15 pripadajućih harmonika po spolu (M/Ž) Linearnom Diskriminacijskom Analizom, Kvadratnom Diskriminacijskom Analizom i Naivnim Bayesovim kalsifikatorom sa Gausovom i Kernel distribucijom daje sljedeće rezultate: LDA 85,30%, QDA 84,95%, Naivni Bayesov kalsifikator sa Gausovom distribucijom 85,03% i Kernel distribucijom 90,98%. Klasifikacija Walsh-Hadamard-ovih koeficijenata po spolu (M/Ž) Linearnom Diskriminacijskom Analizom, Kvadratnom Diskriminacijskom Analizom i Naivnim Bayesovim kalsifikatorom sa Gausovom i Kernel distribucijom daje sljedeće rezultate: LDA WHT 1-16 58,92%, WHT 1-32 57,37%, WHT 1-64 54,53%, QDA WHT1-16 60,07%, WHT1-32 54,40%, WHT 1-64 55,54%, Naivni Bayesov kalsifikator sa Gausovom distribucijom WHT1-16 60,86%, WHT1-32 55,28%, WHT1-64 54,31% i Kernel distribucijom WHT1-16 64,76%, WHT1-32 68,81% WHT1-64 71,50%. Osnovnu frekvenciju F_0 sa pripadajućih 15 harmonika izlučujemo iz isječka signala zvučnog zapisa Brzom Fourierovom transformacijom sa 1024 uzorka. Kod Walsh-Hadamard-ove transformacije isječka zvučnog zapisa korištena su 16, 32 i 64 uzorka. Zbog malog broja korištenih uzoraka pri transformaciji isječka zvučnog zapisa, izlučivanih su Walsh-Hadamard-ovi koeficijenti koji imaju postotak pogreške klasifikacije po spolu (M/Ž) manji od 30%. Izlučene karakteristične značajke za Walsh-Hadamard-ovu transformaciju sa 16 uzoraka su WHT10, 11, 12,

14, sa 32 uzorka su WHT4, 5, 6, 25,26, 28, 30 i za 64 uzorka su WHT1, 8, 11,12, 14, 15, 18, 21, 22, 26, 28,29,32, 33, 35, 36, 37, 50, 52, 56,60, 64. Postotak klasifikacije izlučenih karakterističnih značajki Naivnim Bayesovim kalsifikatorom sa Kernel distribucijom je 72,98 % za koeficijent WHT10, 11, 12 i 14, 74,62% za koeficijente WHT4, 5, 6, 25,26, 28 i 30, te 79,05% za koeficijente WHT1, 8, 11,12, 14, 15, 18, 21, 22, 26, 28,29,32, 33, 35, 36, 37, 50, 52, 56, 60 i 64.

7. Zaključak

7. Conclusion

Cilj istraživanja je pronaći u zvučnom zapisu govornika karakteristične značajke kojima se može odrediti digitalna informacija boja glasa. Ljudski glas možemo promatrati kao subjektivni osjet glasnoće, visine i boje zvuka te njegove percepcije u prostoru. Sadrži do 15 harmonika. Proces izlučivanja karakterističnih značajki za određivanje digitalne informacije boje glasa sastoji se od određivanja osnovne frekvencije F_0 , izračuna pripadajućih 15 harmonika i rastavljanja signala na njegove osnovne sastavnice (Slika 2.). Za izlučivanje karakterističnih značajki korištena je brza Fourierova transformacija sa 1024 uzorka. Na istim uzorcima signala zvučnog zapisa napravljena je brza Walsh-Hadamard-ova transformacija sa 16, 32 i 64 uzorka. Dobiveni Walsh-Hadamard-ovi koeficijenti (16,32 i 64) inverznom transformacijom ne rekonstruiraju u cijelosti govorni signal (slike 5.,6.i 7.), no sadrže karakteristike govornog signala. Metoda klasifikacije po spolu (M/Ž) Naivnim Bayesovim kalsifikatorom sa Kernel distribucijom daje nam najviši postotak klasifikacije i to 92,04% za osnovnu frekvenciju F_0 i harmonike $F_2, F_3, F_4, F_9, F_{10}, F_{12}, F_{13}, F_{14}, F_{16}$ i 79,05% za Walsh-Hadamard-ove koeficijente WHT 1, 8, 11,12, 14, 15, 18, 21, 22, 26, 28,29,32, 33, 35, 36, 37, 50, 52, 56,60, 64 (Tablica 12.). Rezultati istraživanja pokazuju da se do sad neiskorišteni Walsh-Hadamard-ovi koeficijenti mogu upotrijebiti u procesu izlučivanja karakterističnih značajki za određivanje digitalne informacije boja glasa osobe i istovremeno skratiti proces (Slika 2.). Uzimajući u obzir broj od 64 uzorka korištenih kod brze Walsh-Hadamard-ove transformacije, postotak klasifikacije Walsh-Hadamard-ovih koeficijenata

po splu (M/Ž) od 79,05 % je zadovoljavajući, no za viši postotak klasifikacije govornog signala i određivanje karakterističnih Walsh-Hadamard-ovih koeficijenata koje sadrže bolje karakteristike govornog signala, potrebno je napraviti brzu Walsh-Hadamard-ovu transformaciju (FWT) govornog signala sa 512 ili 1024 uzorka i ponoviti postupak.

8. REFERENCE

8. REFERENCES

- [1] Kenneth Bozeman, The Pedagogic Use of Absolute Spectral Tone Color Theory, *Journal of Singing*, November/December 2017 Volume 74, No. 2, pp. 179–183 Copyright © 2017 National Association of Teachers of Singing,
- [2] MathRorks/Documentation/Classification, <https://www.mathworks.com/help/stats/examples/classification.html>,
- [3] Rakesh D, Jayasimhan N, Speaker recognition and authentication, *International Journal of Computer Science and Mobile Computing, IJCSMC*, Vol. 2, Issue. 5, May 2013, pg.402 – 407,
- [4] Dr. H B Kekre, Vaishali Kulkarni, Speaker Identification using Row Mean of DCT and Walsh Hadamard Transform, *International Journal on Computer Science and Engineering (IJCSE)*, ISSN : 0975-3397 Vol. 3 No. 3 Mar 2011 1295,
- [5] Petrinović, D.: Digitalna obrada govora (Interna zavodska skripta), Fakultet elektrotehnike i računarstva, Zavod za elektroničke sustave i obradbu informacija, Zagreb, 2002.,
- [6] Allison A. Johnson, Patrick F. Reidy, Jan R. Edwards: Quantifying robustness of the /t-/k/ contrast using a single, static spectral feature, Johnson et al.: *JASA Express Letters* <https://doi.org/10.1121/1.5049702> Published Online 10 August 2018,
- [7] A. H. Mehta and A. J. Oxenham: Fundamental-frequency discrimination based on temporal-envelope cues: Effects of bandwidth and interference *JASA Express Letters* <https://doi.org/10.1121/1.5079569> Published Online 21 November 2018,
- [8] Milan Vojnović, Uticaj klipovanja govornog signala na osnovnu frekvenciju glasa Telekomunikacioni forum TELFOR 2008 Srbija, Beograd, novembar 25.-27., 2008.,
- [9] Škarić, I.: Govorni dijelovi, III. dio. In: Babić, S. Et al: Povijesni pregled, glasovi i oblici hrvatskog jezika. Globus, Zagreb. 1991, str.291,
- [10] Mladen Heđever, Osnove fiziološke i govorne akustike, Skripta 2013., Edukacijsko-rehabilitacijski fakultet, Sveučilište Zagreb,
- [11] Lei Feng, Andrew J. Oxenham: Effects of spectral resolution on spectral contrast effects in cochlear-implant users, Lei Feng and Andrew J. Oxenham: *JASA Express Letters* <https://doi.org/10.1121/1.5042082> Published Online 14 June 2018,
- [12] Limin Hou, Li Xua: Role of short-time acoustic temporal fine structure cues in sentence recognition for normal-hearing listeners, Limin Hou and Li Xu: *JASA Express Letters* <https://doi.org/10.1121/1.5024817> Published Online 20 February 2018
- [13] Kimberly G. Smitha, Daniel Fogerty: Speech recognition error patterns for steady-state noise and interrupted speech, *J Acoust Soc Am*. 2017 Sep; 142(3): EL306–EL312. Published online 2017 Sep 22. doi: 10.1121/1.5003916] PMID: PMC5724738 PMID: 2896405
- [14] Man-Wai MAK and Hon-Bill YU: A Study of Voice Activity Detection Techniques for NIST Speaker Recognition Evaluations, Preprint submitted to *Computer Speech and Language* May 19, 2013,
- [15] Moe Pwint and Farook Sattar, Speech/Nonspeech Detection Using Minimal Walsh Basis Functions, Received 1 November 2005; Revised 30 May 2006; Accepted 12 June 2006, Hindawi Publishing Corporation *EURASIP Journal on Audio, Speech, and Music Processing* Volume 2007, Article ID 39546, 9 pages doi:10.1155/2007/39546

AUTOR · AUTHOR

Tomislav Jarmić

Pomoćnik je pročelnika za komunalno gospodarstvo i civilnu zaštitu Varaždinske županije, mr.sc. informacijskih znanosti.

Korespondencija

tomislav.jarmic@vz.t-com.hr