

Analiza govornog signala Goertzel algoritmom

Božo Tomas, Marko Obad

Sadržaj — U ovom radu opisan je Goertzel algoritam te implementacija Goertzel algoritma kod spektralne analize govornog signala. Ukazano je na mogućnosti primjene ovog algoritma kod spektralne analize govora a naslučuje se i mogućnost primjene ovog algoritma u mnogim konkretnim rješenjima za obradu govora u govornim tehnologijama.

Glavne reči — Goertzel algoritam, govorni signal, spektralna energija signala

I. UVOD

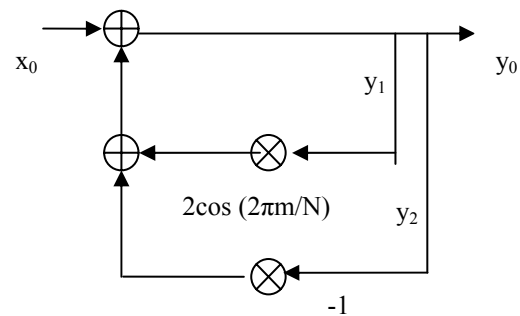
Implementacija sistema za analizu govora podrazumijeva frekvencijsku transformaciju govornog signala. DFT (*Diskretna Fourierova transformacija*) igra važnu ulogu kod implementacije algoritama u sisteme za digitalnu obradu signala. Najčešće korištena transformacija za spektralnu analizu govora je FFT (*Brza Fourierova transformacija*) koja je implementirana u mnoge specijalizirane programske pakete za spektralnu analizu govora. Korištenjem algoritma FFT značajno se skraćuje vrijeme računanja Fourierove transformacije. Ako je potrebno da se detektira jedan ili nekoliko tonova u audio signalu odnosno samo jedna ili par frekvencija, postoji puno brža metoda. Goertzel algoritam omogućava dekodiranje tona (frekvencije) uz znatno manje opterećenje procesora [1], [2], [3] u odnosu na brzu Furijeovu transformaciju (FFT).

Pošto je govor također složeni audio signal, vrijede isti zaključci, što znači da kod obrade digitalnog govornog signala možemo da koristimo iste tehnologije i algoritme, koje ćemo naravno da prilagodimo potrebama analize govornog signala. Evidentno je da za analizu digitalnog govornog signala možemo koristiti Goertzelov algoritam. Glavna prednost ovog algoritma je da su koeficijenti u jednačini za određenu frekvenciju fiksni, što proračun čini mnogo jednostavnijim. Kod analize govora ćemo da koristimo algoritam koji na izlazu, umjesto dekodiranih nivoa signala na željenim frekvencijama, daje spektralnu energiju analiziranog frekvencijskog opsega, tj. Goertzelov algoritam ćemo da koristimo za filtriranje. Kod određivanja spektralne energije digitalnog signala potrebno je da se pamte samo dvije prethodno izračunate vrijednosti, od zadnjeg koraka N, diferencne jednačine

Goertzelovog algoritma, koja se računa metodom korak po korak, dok je za detekciju potrebno po završetku filtriranja da se odredi spektar signala

II. OSNOVNI GOERTZEL

Osnovni Goertzel algoritam, koji postoji od 1958, izveden je iz diskretne Fourierove transformacije. Ovo je izuzetno efikasan algoritam kojim se iz bloka ulaznih podataka diskretnog vremenskog signala, na osnovu relacija između frekvencije uzorkovanja, dužine vremenskog bloka ulaznog signala i koeficijenta ugađanja (podešavanja) frekvencija, određuju frekvencijske vrijednosti željenog tj. ugođenog frekvencijskog opsega tog signala. Na sl. 1. je prikazan tok obrade osnovnog Goertzel algoritma za svaki diskretni trenutak uzorkovanog ulaznog signala [4].



Sl. 1. Tok obrade diskretnog signala Goertzel algoritmom

Algoritam prikazan na sl.1. je ustvari prezentacija slijedeće diferencne jednačine:

$$y(n) = x(n) + y(n-1) \times 2 \cos\left(\frac{2\pi m}{N}\right) - y(n-2) \quad (1)$$

Ako se umjesto $y(n)$ koristi oznaka y_0 , umjesto $y(n-1)$ oznaka y_1 , umjesto $y(n-2)$ oznaka y_2 , te $x(n)$ označi kao x_0 , dobije se jednostavniji i pregledniji zapis jednačine (1) u slijedećem obliku

$$y_0 = x_0 + y_1 \times 2 \cos\left(\frac{2\pi m}{N}\right) - y_2 \quad (1.a)$$

x_0 je je vrijednost tekućeg uzorka iz ADC konvertora, y_0 je vrijednost tekuće iteracije, y_1 je vrijednost zadnje (prethodne) iteracije, a y_2 je vrijednost predzadnje iteracije, m je koeficijent ugađanja odsječka frekvencijske domene, a N je veličina bloka uzoraka.

Uzorci se procesiraju uzorak po uzorak i učitavaju u jednačinu (1), tj. jednačina će da se rješava metodom korak po korak. Procesiranje se obavlja nad blokom

Božo Tomas, Fakultet strojarstva i računarstva Sveučilišta u Mostaru, BiH (telefon/faks: 00387 36 395 292; e-mail: bozo.tomas@ht.ba).

Marko Obad, doktorant na FER u Zagrebu

podataka duljine N. Nakon obrađenog bloka podataka N (kraja bloka), mogu da se odrede amplituda i faza ugođene frekvencije, pomoću memoriranih vrijednosti iteracije u trenutku N-1 (y_1) i trenutku N-2 (y_2). Također, ove vrijednosti će da se koriste i kod izračunavanja spektralne energije ugođenog frekvencijskog opsega analiziranog diskretnog vremenskog signala. Pogodnost Goertzel algoritam je što se iz vremenski procesiranog signala, jednostavno određuju frekvencijske vrijednosti signala u željenim podešenim frekvencijskim opsezima. Nakon što je blok podataka prošao kroz formulu (1) relativna amplituda može da se izračuna od sume kvadrata realne i imaginarne komponente.

Realna komponenta je opisana izrazom:

$$RC = y_1 - y_2 \times 2 \cos\left(\frac{2\pi m}{N}\right) \quad (2)$$

Imaginarna komponenta je opisana izrazom:

$$IC = y_2 \times 2 \sin\left(\frac{2\pi m}{N}\right) \quad (3)$$

Kvadrirana amplituda izračunava se izrazom:

$$\text{amplituda}^2 = RC^2 + IM^2 \quad (4)$$

Dovoljno je, kod detekcije tona, da se izračuna kvadratni korijen amplitude da se odredi relativna jačina signala dok se informacija o fazi može ignorisati.

Kod analize govornog signala u određenom frekvencijskom opsegu tj. Goertzelovoj ćeliji, koristi se slijedeća formula za određivanje spektralne energije frekvencije na koju će promatrana ćelija da bude podešena.

$$E_m = y_1^2 + y_2^2 - 2y_1y_2 \cos\left(\frac{2\pi m}{N}\right) \quad (5)$$

Izraz $2\cos(2\pi m/N)$ u formulama (1) i (5), ima ključnu ulogu u unaprijedom Goertzel algoritmu. U formuli (1) će svaka prethodno izračunata iteracija da se množi, u

tekućoj iteraciji, s $2\cos\left(\frac{2\pi m}{N}\right)$. Izraz $2\pi m/N$, kao i

njegov kosinus su konstante. Značajka Goertzel algoritma je da ovaj izraz kao i njegov kosinus mogu da se unaprijed izračunaju prije procesiranja bloka ulaznih podataka, tako da se izbjegne ponavljanje brojnih matematičkih operacija. N je broj uzoraka signala $x(n)$ tj. veličina bloka podataka a koeficijent ugađanja frekvencija m se određuje iz omjera veličine bloka i ugođenih frekvencija s brzinom uzorkovanja. Pošto je Goertzel izveden iz DFT, njime se praktično dijeli ukupni frekvencijski opseg na niz manjih transformacijskih podopsega kojima može da se ugađa frekvencija ali i širina opsega. Širina frekvencijskog opsega transformacije je:

Širina opsega transformacije = brzina uzorkovanja/N.

Na primjer, pretpostavimo brzinu uzorkovanja 8000 uzoraka u sekundi i $N=100$. Rezultujući opsezi transformacije su tada 80Hz (8000/100). Ovi diskretni opsezi se često nazivaju "binovi" ili Goertzel-ove ćelije. Ostajući na ovom konkretnom primjeru, svaki bin je širine 80Hz i počinje s binom 0. što znači da bin 0 ustvari sadrži DC komponentu signala. Dalje, frekvencijski opseg bina 1 je (80-160) Hz itd. Također, što je vrlo bitno, Goertzel-ove ćelije uopšte ne moraju biti u nizu, već je moguće po želji

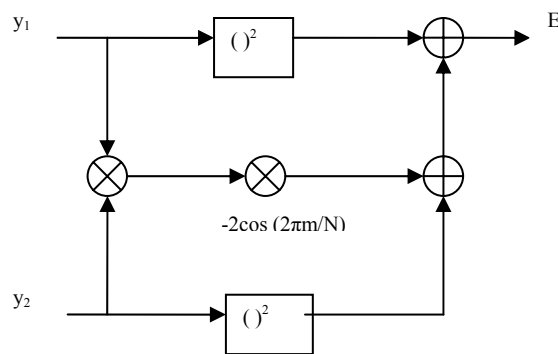
i potrebama analize, da se kombinuju koeficijenti ugađanja, da se proizvoljno odabiru nepovezani ili čak preklapajući frekvencijski segmenti (prozori) odnosno transformacijski opsezi.

Ovisno o relaciji između frekvencije koja se traži, brzine uzorkovanja i konstante N, nije uvijek moguće da se centrirana tražena frekvencija u centar bina. Kada tražena frekvencija nije precizno smješšana u bin, nešto od njegove energije procuri u sljedeći bin. Međutim, ako se govor analizira na frekvencijskim opsezima a ne samo na određenim frekvencijama onda to nije problem, ali ipak treba da se kreiraju Goertzelove ćelije s kojima će željeni frekvencijski opseg da se definira.

Ovaj fenomen curenja također ima i povratni efekt. Dio energije susjednih frekvencijskih sadržaja signala će da se detektira i na traženoj frekvenciji ako nije dobro ugođen bin za tu frekvenciju. Iako generalno ovo nije značajno, može da uzrokuje greške dekodiranja za višefrekventni signal gdje su frekvencije blizu jedna drugoj. Problem curenja može da se umanjí povećanjem konstante N, što smanjuje opseg transformacije (širinu bina).

III. SPEKTRALNA ENERGIJA BINA

Goertzelov algoritam je ustvari veoma visoko Q pojasno propusni filter kojeg karakterizira podesiv selektivan opseg. Goertzel se često implementira kao IIR filter drugog reda [1], [3]. Za analizu digitalnog govornog signala prikladan je sistem koji će na izlazu da daje spektralnu energiju analiziranog opsega signala. Takav sistem, prikazan na sl. 2., realizira se formulom (5) kojom se određuje spektralna energija frekvencije na koju je promatrana Goertzelova ćelija podešena.



Sl. 2. Određivanje spektralne energije ugođenih ćelija

Dakle, nakon kraja prolaska signala kroz formulu (1) po formuli (5) jednostavno izračunamo energiju signala analizirane ćelije (bina). Uvrštavanjem slijedeće vrijednosti m u formulu (1) te novodobivenih vrijednosti y_1 i y_2 , nakon izračuna formule (1), u formulu (5) dobivamo energiju slijedeće ćelije i tako redom. Ustvari, dovoljno jak procesor će istovremeno, nakon prolaska N uzoraka signala, da izračuna vrijednosti y_1 i y_2 za nekoliko ćelija kao i spektralne energije tih ćelija. Za određivanje energije svakog bina (ćelije) potrebno je da se pamte samo dvije prethodno izračunate posljednje vrijednosti formule (1) i da se optimizira ugađanje ćelija za potrebe određene analize. Kako bi Goertzelove ćelije podesili na željene

frekvencijske opsege potrebno je da se odrede početni parametri prije same realizacije.

Frekvencija uzorkovanja mora da bude barem dva puta veća od najveće frekvencije koja se analizira kako bi se spriječio aliasing. Nadalje potrebno je da se odredi broj uzoraka N nakon kojih će da se određuju spektralne energije pojedinih ćelija. Dakle, formula (1) se računa za svaki pojedini uzorak digitaliziranog govornog signala, od prvog do posljednjeg N-tog uzorka, dok se formula (5) za spektralnu energiju izračunava samo jednom nakon što se završi formula (1). Postupak opisan formulom (1) često se naziva filtarski dio spektralne analize Goertzel-ovog algoritma. Ugađanje ćelija možemo da radimo po formuli:

$$f_{\text{ćelije}} = \frac{mf_u}{N} \Rightarrow m = \frac{Nf_{\text{ćelije}}}{f_u} \quad (6)$$

gdje je f_u -frekvencija uzorkovanja, N -broj ulaznih uzoraka signala, m -koeficijent ugađanja odsječka frekvencijske domene, a $f_{\text{ćelije}}$ je frekvencija na koju se ugađa Goertzel-ova ćelija. Širina frekvencijskog opsega Goertzel-ove ćelije je f_u/N . Također za svaki bin odnosno odgovarajući koeficijent m možemo da odredimo

$$\text{konstante } k_m = 2 \cos\left(\frac{2\pi m}{N}\right).$$

Ove vrijednosti se odrede prije izvršavanje programa, formule (1), te memoriraju na određene memorijske lokacije kako bi se uštedjelo procesorsko vrijeme koje bi inače bilo potrebno za njihovo izračunavanje.

IV. PRAĆENJE FORMANTNE STRUKTURE GOERTZEL ALGORITMOM

Za analizu ćemo da koristimo govorni signal, vokal /a/, snimljen programom Sound Recorder s frekvencijom uzorkovanja $f_u=8000\text{Hz}$ te memoriran pod nazivom "Jako A". Ako Goertzelovu transformaciju, formulu (1), izvodimo na vremenskom odsječku snimljenog govornog signala trajanja 25ms, imati ćemo $N=200$ uzoraka govornog signala koje ćemo da procesiramo algoritmima opisanim formulom (1) i formulom (5). Kako je $f_u=8000\text{Hz}$ i $N=200$, frekvencijski opseg svake Goertzelove ćelije (bina) je: $B=f_u/N=40\text{ Hz}$.

Prvi formant ćemo da tražimo u frekvencijskom opsegu (400 – 720) Hz. Ovaj frekvencijski opseg dijelimo u slijedećih 8 binova. Frekvencijski opsezi naših osam binova kao i njihove centralne frekvencije na temelju kojih ćemo da računamo koeficijente ugađanja frekvencijske domene m, prikazani su u tabeli 1

TABELA 1: PODJELA FREKVENCIJSKOG OPSEGA NA BINOVE.

Bin	1	2	3	4	5	6	7	8
B (Hz)	400-440	440-480	480-520	520-560	560-600	600-640	640-680	680-720
$f_{\text{ćelije}}$	420	460	500	540	580	620	660	700

Sad ćemo za svaki bin da izračunamo koeficijente ugađanja frekvencijske domene m, koje ćemo potom da

uvrstimo u izraz $2\cos(2\pi m/N)$ te vrijednost tog izraza koja je konstanta, različita za svaki bin, memorisati na određene memorijske lokacije, kako bi procesor trošio manje vremena za izračun formule (1), pošto se ona izvodi u 200 koraka. To ćemo da ilustriramo za bin 1.:

$$f_{\text{ćelije}}^{(1)}=420 \Rightarrow m_1 = f_{\text{ćelije}}^{(1)} \frac{N}{f_u} = f_{\text{ćelije}}^{(1)} \frac{200}{8000} = \frac{f_{\text{ćelije}}^{(1)}}{40} = \frac{420}{40} = 10,5 \quad (7)$$

Koeficijent ugađanja frekvencijske domene, za k-ti bin m_k , kod frekvencije uzorkovanja $f_u=8000\text{Hz}$ i dužine bloka vremenskih uzoraka $N=200$ ćemo da izračunavamo po slijedećoj formuli:

$$f_{\text{ćelije}}^{(k)} = \text{definiran} \Rightarrow m_k = f_{\text{ćelije}}^{(k)} \frac{N}{f_u} = f_{\text{ćelije}}^{(k)} \frac{200}{8000} = \frac{f_{\text{ćelije}}^{(k)}}{40} \quad (8)$$

Uvrštavanjem dobivenih vrijednosti m_k u izraz $2\cos(2\pi m_k/N)$ dobivamo za svaki bin njegov koeficijent $k_{m(k)}$ za izračun formule (1). Prikaz svih vrijednosti dat je u tabeli 2.

Sad imamo sve potrebne ulazne konstante za izračun formula (1) i (5). Uzorci govornog signala se unose redom kako su uzorkovani u formulu (1), uzorak po uzorak, tj jednačina se izvodi korak po korak sve do posljednjeg koraka $N=200$. Izračunate vrijednosti jednačine u posljednja dva koraka $N=199$ i $N=198$ se memorišu te unose u jednačinu (5). Za svaki bin izračunavamo njegovu spektralnu energiju tako što koeficijente tog bina unosimo u jednačine (1) i (5). Ustvari, na taj način smo odredili spektralnu energiju na frekvenciji $f_{\text{ćelije}}$ u njenoj pripadajućoj Goertzel-ovoj ćeliji (binu).

TABELA 2: PRIKAZ PARAMETARA, KOEFICIJENTA I SPEKTRALNE ENERGIJE BINOVA.

Bin	2	3	4	5	6	7	8
$f_{\text{ćelije}}$	460	500	540	580	620	660	700
m_k	11,5	12,5	13,5	14,5	15,5	16,5	17,5
$k_{m(k)}$	1,87	1,84	1,82	1,79	1,76	1,73	1,70
E	4,60	69,4	74,4	48,7	144,9	416,1	58,1

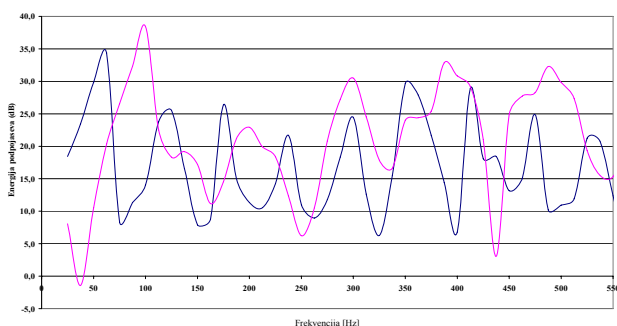
Isti postupak je i za ostale formante, koje bi tražili na višim frekvencijama. Također, ne moramo da uzimamo ćelije u nizu. Jednostavnim izborom konstanti mi možemo da računamo ćelije koje se preklapaju. Osim toga, ako nam je potrebna veća selektivnost binova možemo to da realiziramo povećanjem broja ulaznih uzoraka N uz nepromijenjenu frekvenciju uzorkovanja ili obrnuto.

Iz tabele 2 zaključujemo da je prvi formant vokala /a/ pozicioniran u binu 7 u frekvencijskom opsegu (640-680) Hz. Također, na temelju vrijednosti energija u binovima 6 i 8 zaključujemo da je prvi formant pozicioniran na početku bina 7, odnosno da je u opsegu frekvencija (640-660) Hz. Međutim, za većinu analiza nije uopšte potrebno da se određuju tačne frekvencijske pozicije formantata, nego je sasvim dovoljna ova podjela na binove,

pa se zaključci o emocionalnim stanjima govornika kao i prepoznavanju govora i govornika mogu da se donose temeljem poređenja među energijama odabranih binova.

Za preciznije određivanje frekvencije prvog formanta možemo da provedemo daljnji postupak Goertzel transformacije na binovima 6 i 7 tako što bi za taj frekvencijski opseg povećali broj uzoraka govornog signala npr $N=400$ pa bi tad frekvencijski opseg novih Goertzelovih ćelija bio: $B=f_u/N=8000/400=20\text{Hz}$. Analizirani opseg (620-680)Hz možemo da podijelimo na tri nova bina te za ta tri nova bina da izračunamo spektralne energije. Međutim, s $N=400$ tj. 50ms govornog signala izašli bi iz okvira kvazistacionarnosti, ali zato smo mogli s $N=200$ odabirom m -koeficijenta ugađanja odsječka frekvencijske domene, da analiziramo opsege s preklapanjem susjednih binova npr. 50% te da dobijemo spektralnu energiju na analiziranom opsegu frekvencije s distancom 20 Hz umjesto 40 Hz kao u tabeli 2 gdje su binovi u nizu bez preklapanja.

Kod određivanja harmonika potrebna je veća frekvencijska selektivnost što znači da trebamo uz nepromjenjenu frekvenciju uzorkovanja da uzmemo veći broj uzoraka govornog signala. Na sl. 3 su prikazani harmonici segmentiranog vokala /A/ iz riječi *mama* plavom bojom odnosno iz riječi *ma* crvenom bojom. Frekvencija uzorkovanja je 8 kHz, broj uzoraka Goertzelove transformacije $N=320$ pa je širina bina 25 Hz, a kako je preklapanje ćelija 50% to je distanca određivanja spektralne energije 12,5 Hz.



Sl. 3. Harmonici vokala /A/

Za odabrani broj uzoraka $N=320$ trajanje vremenskog okvira je 40 ms, što još uvijek zadovoljava uslove kvazistacionarnosti govora. Pored toga program omogućava da odaberemo od kojeg uzorka iz analizirane datoteke će početi očitavanje uzorka u formulu (1). U svim analizama se počinjalo od uzorka broj $k=50$ što omogućava analizu izdvojenog vokala u njegovoj finalnoj kvazistacionarnoj vremenskoj poziciji. Znači u program se upisuje zvučna datoteka, npr. „JakoA“, zatim koeficijente ugađanja frekvencija m , početni uzorak očitavanja k i broj uzoraka N koji je ujedno i broj iteracija Goertzelove transformacije tj. formule (1).

V. ZAKLJUČAK

Goertzelov algoritam je često primjenjivana tehnika kod realizacije digitalnih DTMF (Dual –Tone Multiple Frequency) prijemnika. DTMF signalizacija, tzv. tonsko biranje, izvodi se audio tonovima koje predajnik generira, a na prijemnoj strani je iste potrebno na neki način

dekodirati radi daljnje obrade. DTMF predajnik (koder) generira složeni audio signal sastavljen od dvije međusobno harmonijski neovisne frekvencije, kombiniranjem osam zadatih frekvencija, a prijemnik treba detektirati i dekodirati primljeni audio signal.

Svrha uvođenja Goertzelovog algoritma kod frekvencijske analize digitalnog govornog signala je brza obrada govornog signala, te određivanje spektralne energije iz digitalnog signala u željenim frekvencijskim ćelijama (binovima) na osnovu unaprijed određenih koeficijenata m . Zatim analizom izračunatih energija, po pojedinim binovima, možemo dobiti mnoštvo informacija koje govorni signal sadržava [5]. Također, pošto Goertzelovim algoritmom možemo, pored energetske vrijednosti govornog signala, da određujemo i harmonike segmentiranih vokala [6], [7], evidentno je kako je ovaj algoritam pogodan u aplikacijama kao što su prepoznavanje govora i govornika, prepoznavanje emocija i drugih nelingvističkih atributa govora. Za pretpostaviti je da će Goertzelov algoritam da bude implementiran u komunikacijskim modelima kod kojih će pored lingvističke poruke da bude uključen i emocionalni sloj izgovorene poruke [8].

Literatura

- [1] S. Bagchi and S.K. Mitra, “An efficient algorithm for DTMF decoding using the subband NDFT,” *Proc. IEEE Int. Symp. On Circuits and Systems*, May 1995, pp. 1936-1939
- [2] M.D. Felder, J.C. Mason and B.L. Evans, “Efficient ITU-Compliant Dual-Tone Multiple-Frequency Detection Using the Non-Uniform Discrete Fourier Transform”, *IEEE Signal Processing Letters*, vol.3, No.7, Jul. 1998, 160-163
- [3] P.Mock, “Add DTMF generation and decoding to DSP microprocessor designs”, *EDN*, vol.30 March 1985, 205-220
- [4] Eric Kiser, “Digital Decoding Simplified Sequential Exact-Frequency Goertzel Algorithm”, *CIRCUIT CELLAR*, Issue 182, September 2005, pp. 22-26
- [5] Božo Tomas “Raspoznavanje lingvističkih i nelingvističkih informacijskih sadržaja u govornom signalu”, [BIHTEL, BiH, 2006]
- [6] Božo Tomas, Mladen Maletić and Zlatko Raguž, “Influence of Emotions to Pitch Harmonics Parameters of Vowel /a/”, *ELMAR-2007*, 49th International Symposium focused on Mobile Multimedia, Zadar 12-14 septembar 2007, 151-154
- [7] Božo Tomas, Mladen Maletić and Zlatko Raguž, “Determination and Evaluation Pitch Harmonics Parameters Width Emotions Classification”, *SoftCOM 2007*, 15th International Conference on Software, Telecommunications and Computer Networks, Split-Dubrovnik, 2007,
- [8] Božo Tomas, Mladen Maletić and Marko Obad, “Recognition and Implementation of Emotions in Speech Communications Model”, 3rd Congress of the Alps Adria Acoustics Association, Graz-Austria, 27-28 September 2007,

ABSTRACT

This paper describes the Goertzel algorithm and its implementation with speech signal spectral analysis. It describes the method of formant analysis of vowel /a/ as well as the vowel /a/ harmonics analysis of Goertzel algorithm.

SPEECH SIGNAL ANALYSIS BY GOERTZEL ALGORITHM

Božo Tomas, Marko Obad