

Analiza i implementacija Gold-Rabiner algoritma za određivanje osnovne frekvencije ljudskog govora na Spartan3 FPGA razvojnoj ploči

Ovaj rad prikazuje implementaciju Gold-Rabiner algoritma za određivanje osnovne frekvencije ljudskog glasa u VHDL-u, jeziku za opis hardvera. Algoritam je testiran na Xilinx Spartan3 FPGA razvojnoj ploči. Motivi za specifičnu hardversku implementaciju dolaze iz činjenica da Gold-Rabiner algoritam ne zahteva operacije množenja i deljenja i da je veoma pogodan za paralelizaciju. Dodatno, u Matlab-u je izvršeno poređenje Gold-Rabiner algoritma sa algoritmom baziranim na autokorelaciji.

Uvod

Pič perioda (eng. pitch period) je osnovna perioda govornog signala, to jest osnovni period oscilacije glasnih žica. Pič frekvencija je osnovna učestanost govornog signala i predstavlja recipročnu vrednost pič periode. Pič perioda osobe se ne menja značajno u odraslom dobu pa se stoga može koristiti kao parametar za prepoznavanje govornika. Opseg pič frekvencija za muškarce je 85-160 Hz, dok je za žene 165-225 Hz (Titze 1994).

Pič period estimator je uređaj koji na osnovu ulaznog govornog signala izračunava pič periodu. Određivanje pič periode može se izvršiti u vremenskom i u frekvencijskom domenu. U frekvencijskom domenu dobre rezultate daje meto-

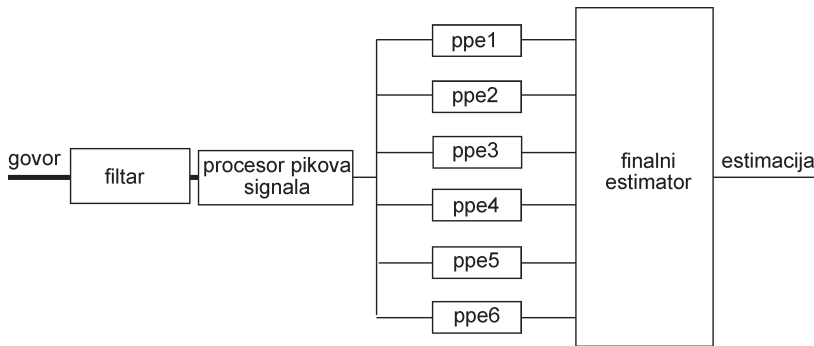
da bazirana na izračunavanju autokorelacije govornog signala. U vremenskom domenu metoda koju su predložili Gold-Rabiner (Gold i Rabiner 1969) predstavlja računski efikasnu metodu za estimaciju, bazirajući se na principu kombinacije više jednostavnih estimatora. Najčešća primena estimatora pič periode se nalazi u vokoderima i sistemima za identifikaciju i verifikaciju govornika.

Field Programmable Gate Array (FPGA) je integrisano kolo koje se može programirati nekim od HDL jezika (Hardware Description Language). Sam FPGA čip se sastoji od lookup (eng. lookup) tabela, multipleksera i logičkih gejtova (eng. gate) koji se mogu međusobno povezivati, čime formiraju gradivne blokove za složenije digitalne sisteme. Glavnu primenu ima u procesu projektovanja i proizvodnje Application Specific Integrated Circuit (ASIC) čipova, gde se koristi za testiranje komponenti koje se razvijaju. Korišćenjem FPGA u ovom procesu smanjuju se troškovi razvoja sistema, jer se može menjati hardverska konfiguracija sistema u toku rada. Karakteristična arhitektura ovih čipova omogućava paralelno izvršavanje više procesa. Iz tog razloga je za implementaciju Gold-Rabinerove metode određivanja preko paralelnog procesiranja, izabran FPGA kao razvojni sistem.

U ovom radu upoređene su metoda paralelnog procesiranja kao računski efikasnog predstavnika algoritama za određivanje pič periode, sa autokorelacionim metodom. Takođe, izvršena je implementacija Gold-Rabiner algoritma u VHDL-u i urađena analiza performansi i resursa potrebnih za izvršavanje algoritma na FPGA čipu.

Lazar Mitrović (1996), Beograd, Rudo 2, učenik 2. razreda Matematičke gimnazije u Beogradu

MENTOR: Marija Janković, Elektrotehnički fakultet Univerziteta u Beogradu



Slika 1.
Blok šema
Gold-Rabinerove
metode

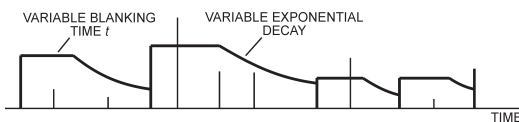
Figure 1.
Block diagram of
Gold-Rabiner's method

Metod

Šema sistema za određivanje pič periode predloženog u radu Golda i Rabinera (Gold i Rabiner 1969) prikazana je na slici 1.

Signal se odabira učestalošću od $F_s = 24$ kHz. U originalnom radu (Gold i Rabiner 1969), govorni signal se propušta kroz filtar propusnik opsega učestanosti (engl. bandpass filter) sa graničnim vrednostima 100 Hz i 800 Hz. Filtrirani signal se zatim obrađuje u procesoru pikova signala, koji za izlaz ima šest impulsnih nizova (u daljem tekstu m-funkcija) definisanih u originalnom radu (Gold i Rabiner 1969). M-funkcije se generišu izdvajanjem i kombinovanjem minimuma i maksimuma odbiraka signala, pošto se na osnovu njih mogu izdvojiti informacije o periodičnosti.

Formirane m-funkcije se obrađuju u zasebnim pič period estimatorima (slika 2). Amplituda prvog nenultog odbirka se pamti, a zatim se čeka da prođe vreme tokom kojeg sistem ne vrši estimaciju (eng. variable blanking time, u daljem tekstu VBT). Ovim se sprečava registrovanje pič periode koja je manja od moguće (Titze 1994: 118). Nakon isteka VBT, računa se broj odbiraka do presecanja eksponencijalne funkcije (variable exponential decay) sa sledećim pikom. Eskpo-



Slika 2. Princip rada pič estimatora

Figure 2. Pitch estimator scheme

nencijalna funkcija je oblika $e^{-\lambda t}$, gde je λ konstanta određena sa ciljem da ograniči maksimalnu vrednost procene pič periode. Sabiranjem VBT i vremena do presecanja sa sledećim pikom dobija se procena pič periode.

Autokorelaciona metoda se sastoji od tri koraka: klipovanja signala, primene autokorelacije nad tako obrađenim signalom i traženja maksimuma autokorelacione funkcije u oblasti gde su očekivane vrednosti pič frekvencije.

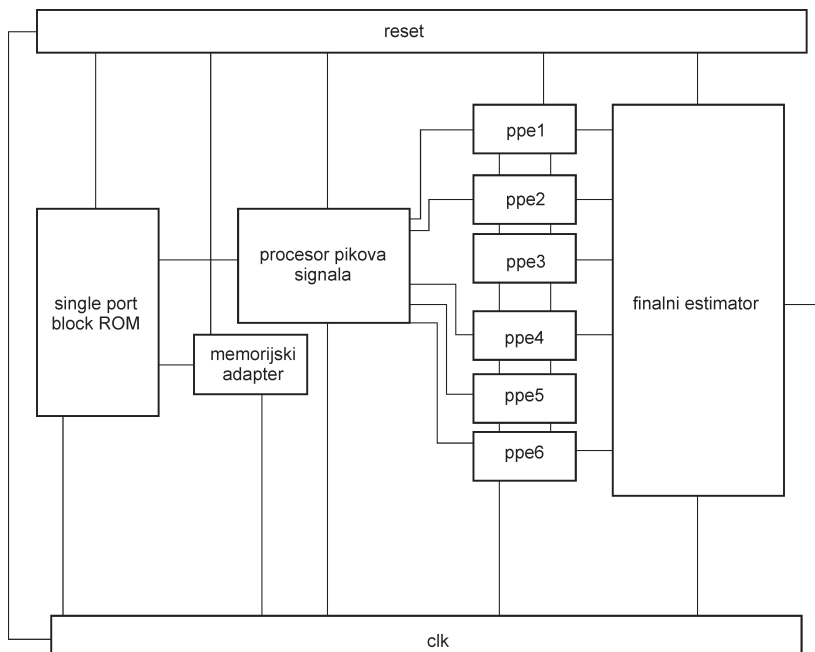
Klipovanje je process u kojem se ostavljaju odbirci signala koji su po apsolutnoj vrednosti iznad određenog praga, dok se vrednosti manje od praga izjednačavaju sa nulom. Time se postiže smanjenje uticaja šuma koji može dovesti do pogrešne estimacije. Takođe, klipovanjem se ubrzava izračunavanje autokorelacije, sa malim uticajem na ishod datog metoda.

Nakon izračunavanja autokorelacije klipovanog signala, pronalazi se maksimum autokorelacione funkcije u oblasti u kojoj se očekuju vrednosti pič frekvencije.

Snimljeni su govorni signali tri osobe (dve ženskog i jedna muškog pola) pri izgovaranju glasova A i O, po tri puta. Snimanje je vršeno funkcijom wavrecord u programskom paketu Matlab.

Estimacija je vršena bez filtriranja. Za svaki signal pitch frekvencija je određivana autokorelacionom metodom, implementiranom u Matlabu, metodom paralelnog procesiranja implementiranoj u Matlabu i verzijom ovog algoritma implementiranog na FPGA Xilinx XC36500 čipu.

Zbog ograničene blok memorije za smeštanje ulaznih podataka na FPGA čipu, pri estimaciji je korišćeno 10000 odbiraka.



Slika 3.
Blok dijagram
implementacije
Gold-Rabiner
algoritma u VHDL-u

Figure 3.
Gold-Rabiner
algorithm in VHDL

Autokorelaciona metoda koristi trinivovsko klipovanje, sa koeficijentom klipovanja 0.6.

Implementacija Gold-Rabinerove metode u VHDL-u izvršena je na osnovu šeme na slici 3. Svaki blok predstavlja jedan entitet. Razlikuje se po tipu podataka – Matlab koristi podatke sa pokretnim zarezom, dok FPGA radi sa podacima sa fiksnim zarezom.

Govorni signal se skladišti u jednoportnu blok ROM memoriju. Memorijski adapter u svakom taktu šalje memorijskom bloku adresu podatka za čitanje. Pročitani odbirak signala se iz memorije dovodi na procesor pikova signala, koji na osnovu sačuvanih prošlih stanja i novog odbirka formira m funkcije. Njihove trenutne vrednosti se prenose na 6 pič period estimatora. Koeficijenti eksponencijalne su implementirani kao niz konstanti u VHDL programu. Kako bi bilo moguće izračunavanje vrednosti eksponencijalne funkcije i pič frekvencije, korišćen je paket `fixed_pkg_c.vhdl`, koji obezbeđuje manipulaciju brojevima sa fiksnim zarezom.

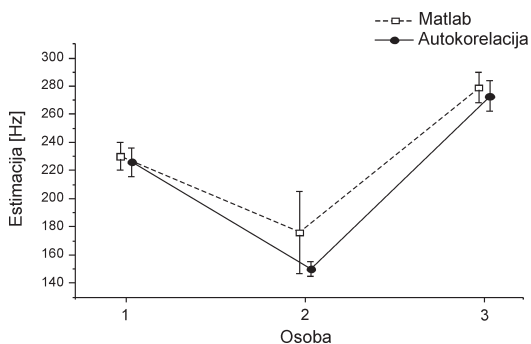
Dobijene procene pič perioda se skladište u tabeli sa 36 elemenata – poslednja tri izlaza svih estimatora i zbirovima te tri procene (prvom + drugom, drugom + trećom i sve tri). Finalna estimacija se dobija izborom one koja ima najviše koincidencija sa elementima tablice. Koinci-

dencijom se smatra vrednost koja se nalazi u ± 300 ms.

Testiranje je obavljeno na Spartan 3E razvojnoj ploči sa Xilinx XC36500 FPGA čipom.

Rezultati

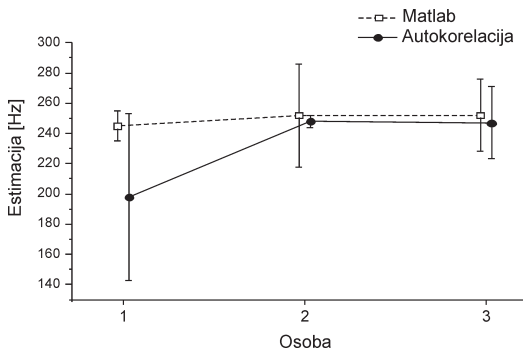
Graficik na slici 4 prikazuju srednje vrednosti estimacija pič frekvencija za tri osobe, pri izgovaranju glasa A.



Slika 4. Estimacije pič frekvencije paralelnim procesiranjem i autokorelacijom sa odgovarajućim standardnim devijacijama za glas A

Figure 4. Pitch period estimation by parallel processing and autocorelation for vowel A

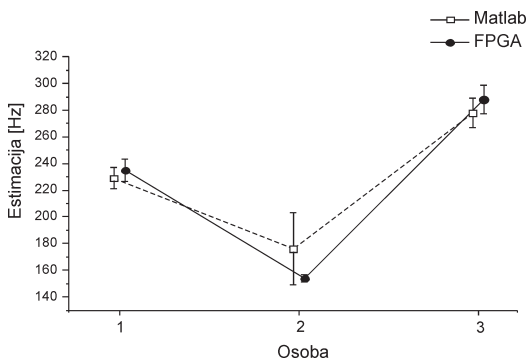
Grafik na slici 5 prikazuju srednje vrednosti estimacija pič frekvencije za tri osobe sa odgovarajućim standardnim devijacijama, pri izgovaranju glasa O.



Slika 5. Estimacije pič frekvencije paralelnim procesiranjem i autokorelacijom sa odgovarajućim standardnim devijacijama za glas O

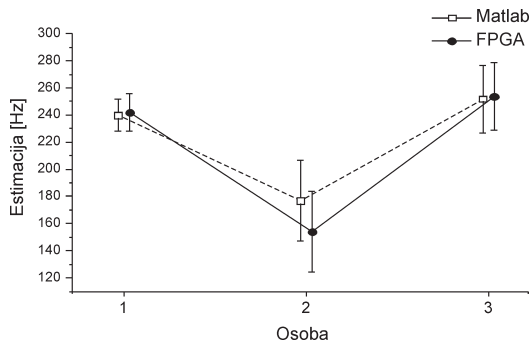
Figure 5. Pitch period estimation by parallel processing and autocorrelation for vowel O

Na grafiku na slici 6 prikazane su estimacije pič frekvencije dobijene implementacijom metoda paralelnog procesiranja u Matlab programskom paketu i na FPGA čipu. Snimljeni glas je glas A.



Slika 6. Estimacije pič frekvencije paralelnim procesiranjem implementiranim u Matlab-u i na FPGA čipu za glas A

Figure 6. Pitch period estimation by parallel processing implemented in Matlab and FPGA for vowel A



Slika 7. Estimacije pič frekvencije paralelnim procesiranjem implementiranim u Matlab-u i na FPGA čipu za glas O

Figure 7. Pitch period estimation by parallel processing implemented in Matlab and FPGA for vowel O

Na grafiku sa slike 7 prikazane su estimacije pič frekvencije dobijene implementacijom metoda paralelnog procesiranja u Matlab programskom paketu i na FPGA čipu. Snimljeni glas je glas O.

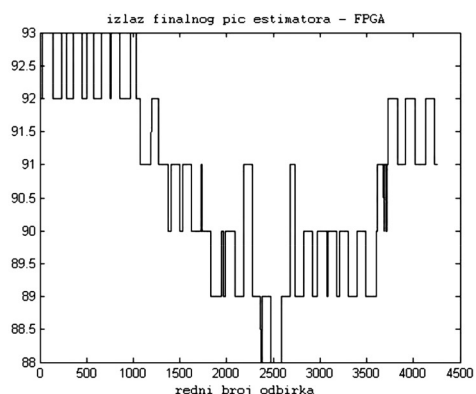
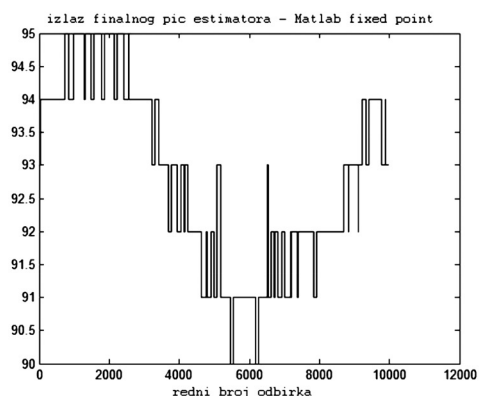
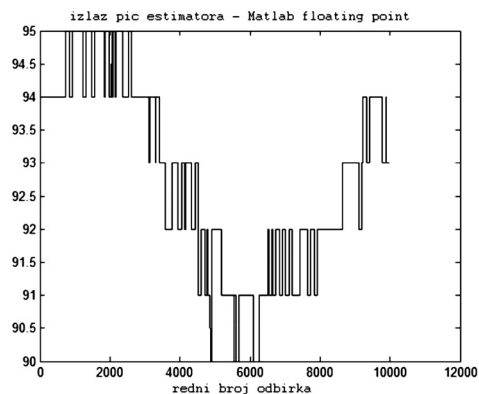
Na osnovu rezultata prikazanih na slikama 4 i 5, nije moguće zaključiti da postoji značajna razlika u preciznosti između autokorelacione i Gold-Rabiner metode

Radi provere funkcionalnosti implementacije Gold-Rabiner algoritma na FPGA čipu, podaci u Matlabu su predstavljeni u formatu sa fiksnim zarezom. Grafici na slici 8 prikazuju izlaz finalnog pič estimatora za implementaciju u Matlabu sa pokretnim zarezom, za Matlab implementaciju sa fiksnim zarezom i za FPGA implementaciju, respektivno. Estimacija je izražena u broju odbiraka.

Diskusija

Teorijski, autokorelaciona metoda, kao metoda u frekvencijskom domenu, bi trebalo da daje preciznije rezultate. Mogući razlozi za nedovoljnu preciznost su nedostatak filtriranja signala pre obrade, kao i mali skup odbiraka na kojima se vrši obrada.

Rezultati za snimke sa slovom O (slika 5) nisu zadovoljavajući jer je i kod autokorelacione i kod Gold-Rabiner metode prisutna velika devi-



Slika 8. Izlaz finalnog pič estimatora za implementacije GR algoritma u Matlabu – floating point, Matlab – fixed point i FPGA, respektivno

Figure 8. Final estimations by Gold-Rabiner implementation in matlab with floating point, Matlab with fixed point and FPGA, respectively

jacija (maksimalna za autokorelacionu metodu - 59.7 Hz, a za Gold-Rabinerovu 37.3 Hz).

Posmatrajući grafike na slikama 6 i 7 uočava se da postoji odstupanje u proceni pič periode. Do manjih odstupanja između FPGA i Matlab implementacije Gold-Rabinerove metode dolazi iz više faktora. Matlab implementacija radi sa brojevima sa pokretnim zarezom, dok FPGA radi sa fiksnim zarezom i preciznošću od 8 decimala. Takođe, koeficijenti eksponencijalne funkcije i memorisanih odbiraka koji su zapamćeni kao konstante u FPGA implementaciji dovode do nagomilavanja greške. Posmatrajući standardne devijacije ovih merenja, uočava se da su obe implementacije slične preciznosti.

Kako bi dodatno proverili funkcionalnost implementacije Gold-Rabiner algoritma na FPGA čipu, kreirana je i implementacija u Matlab programskom paketu koja obrađuje odbirke u istom formatu, kao i FPGA. Poređenjem prvog i drugog grafika na slici 8 uočava se da promena u preciznosti podataka dovodi do promene u estimaciji pič periode. Međutim, postoji i razlika između drugog i trećeg grafika na istoj slici. Ova razlika se može objasniti drugačijim okruženjem u kojem se algoritam, izvršava, različitom arhitekturom jedinica koje izvršavaju aritmetičke i logičke operacije. Razlika u samo vrednosti estimacije najverovatnije potiče od drugačijeg indeksiranja, koje može rezultovati konstantnim pomerajem signala po y osi.

Zaključak

Pokazano je da predložena FPGA implementacija Gold-Rabinerove metode daje estimacije pič periode u zadovoljavajućem opsegu. Zbog malog broja merenja nije moguće uporediti rezultate Gold-Rabinerove metode i metode na bazi autokorelacije.

Na osnovu obavljenih merenja može se zaključiti da je za računanje pič periode glas A pogodniji od glasa O.

Lirteratura

Gold B., Rabiner L. 1969. Parallel Processing Techniques for Estimating Pitch Periods of Speech in the Time Domain. *The Journal of the Acoustical Society of America*, **46**: 442.

Titze I. R. 1994. *Principles of Voice Production*. Prentice Hall

Wakerly J. F. 2000. *Digital Design: Principles and Practices*. Prentice-Hall

Lazar Mitrović

Analysis and Implementation of the Gold-Rabiner Algorithm for Pitch Period Estimation of Human Speech Using Spartan3 FPGA Development Board

This paper shows the implementation of the Gold-Rabiner algorithm for pitch period estimation of human speech in the hardware description language VHDL. The algorithm was tested on the Spartan3 FPGA development board. The motivation for doing specific hardware implementation of this algorithm was based on two facts: the algorithm does not require multiplication and division and it has very good parallelization capabilities. In addition, Matlab implementation was tested versus the autocorrelation based pitch period estimation algorithm. Results showed that implementation of this algorithm is possible but accuracy could not be calculated due to lack of measurements. 