

Prepoznavanje duvačkih instrumenata pomoću hromaprofila i neuralne mreže

Tatjana Miljković, Miloš Bjelić, Dragana Šumarac Pavlović, Goran Kvašček

Elektrotehnički fakultet, Univerzitet u Beogradu

Beograd, Srbija

mt185029p@student.etf.rs; bjelic@etf.rs; dsumarac@etf.rs; kvascev@etf.rs

Sažetak— U procesima automatskog prepoznavanja instrumenata koriste se različita obeležja zasnovana na analizi spektralnog sadržaja audio zapisa. U ovom radu prikazan je hroma profil tonova kao novo obeležje, izdvojeno iz audio zapisa muzičkog sadržaja, na osnovu kojeg se vrši prepoznavanje duvačkih instrumenata pomoću neuralne mreže. Duvački instrumenti koji su bili od interesa u ovom istraživanju su: klarinet, flauta i oboa. Navedeni instrumenti odabrani su tako da pokrivaju širok opseg varijacija u pogledu načina pobuđivanja, oblika cevi i strukture rezonantnih modova. Metodologija karakterizacije duvačkih instrumenata bazirana na hroma profilu tonova pokazala je očuvanje osnovnih razlika u karakteristikama instrumenata. Hroma profil predstavlja prikaz relativnih odnosa energije na pojedinim tonovima unutar oktave. Posmatrani su hroma profili tonova podeljeni na 12 i 1200 podopsega. Korišćenjem hroma profila tonova kao ulaznih parametara neuralne mreže ostvareni su visoki procenti prepoznavanja instrumenta.

Ključne riječi— muzički instrumenti; audio obeležja; neuralna mreža; prepoznavanje; hromatogram; hromaprofil;

I. UVOD

Evolucijom digitalnih tehnologija pojavila se mogućnost za razvoj različitih metodologija kojim se mogu analizirati audio zapisi muzičkog sadržaja. Takođe, razvitkom tehnologije javila se potreba za različitim multimedijalnim aplikacijama koje su imale za cilj da izvršavaju automatsku transkripciju nota, klasifikaciju muzičkih žanrova, identifikaciju pevača kao i prepoznavanje muzičkih instrumenata. Stečena znanja o osobinama i građi zvuka stvorila su osnov za dalja istraživanja audio zapisa muzičkog sadržaja, ali takođe i za određivanje atributa zvuka muzičkih instrumenata, pomoću savremenih softverskih alata. Kao reference pri istraživanju muzičkih signala najčešće se uzimaju u obzir nekoliko osnovnih parametara koji ih opisuju i to: visina tona, melodija, harmonija i ritam [1].

Iako postoje raznovrsni muzički sadržaji njihova osnova je ista i čine je tonovi. Ton predstavlja zvuk koji ima izraženu tonsku visinu. Spektar tona je diskretan i sastavljen od harmonijskog niza frekvencija koji se u muzičkoj literaturi naziva alikvotni niz [2]. Harmonijski niz frekvencija kao takav se prema osnovnom tonu odnosi kao $1:2:3:4:n$, gde je n određena granična frekvencija. Iako je intervalski odnos harmonika prema osnovnom tonu uvek jednak, kod različitih

muzičkih instrumenata razlikuje se relativna jačina pojedinih alikvota. Razlike u relativnoj jačini pojedinih alikvota formiraju anvelopu spektra tona koja predstavlja jedan od faktora kojim se definiše boja tona.

Proces percepcije tonske visine kod čoveka formira se na logaritamskoj frekvencijskoj osi. Proces koji se odvija u čulu sluha dovodi do preslikavanja harmonijskog niza frekvencija na logaritamsku frekvencijsku osu. Pri preslikavanju spektralnih komponenti tona na hroma krug pozicije nekih komponenti odgovaraju određenim tonovima unutar jedne oktave. Analizom pozicija harmonika određenog tona na hroma krugu nakon preslikavanja pokazano je da se ukupna energija tona dominantno nalazi na osnovnoj frekvenciji tona i na frekvencijama koje odgovaraju intervalima velike terce i čiste kvinte [1]. Raspodela energije jednog tona unutar hroma kruga predstavlja pogodan podatak za prepoznavanje različitih karakteristika kojima se mogu opisati muzički signali. Takođe, zbog izražene harmonijske strukture muzički ton se može jednostavno matematički modelovati. U literaturi je pokazano da takav model predstavlja jedan od glavnih ulaznih parametara algoritama za automatsko prepoznavanje atributa muzičkih instrumenata [3].

U cilju automatskog prepoznavanja muzičkih instrumenata razvijen je veliki broj algoritama. Iako je broj algoritama značajan, velika većina se zasniva na izdvajanju karakterističnih obeležja. Prvobitna istraživanja u pravcu izdvajanja obeležja muzičkih audio zapisa zasnivala su se na tehnikama izdvajanja govornih obeležja kao što su: LPC (*Linear Prediction Coefficients*) i MFCC (*Mel-Frequency Cepstral Coefficients*) koeficijenti [4]. Marques i Moreno su pomoću ovih obeležja uspeali da ostvare klasifikaciju od oko 70% muzičkih audio zapisa [5]. Dalja istraživanja su podlegla drugačijem pristupu. U cilju stvaranja što robusnijeg sistema za prepoznavanje instrumenata, Eronen i Klapuri, razvili su sistem u kojem izdvajaju preko 20 različitih obeležja [6]. Koristili su kako spektralna tako i vremenska obeležja, među kojima su različite varijacije MFCC koeficijenata, spektralni centroid, “udarna”(attack) i “opadajuća”(decay) faza tona i druge. Uprkos većem broju kako vremenskih tako i spektralnih obeležja rezultati klasifikacije instrumenata iznosili su oko 80% uspešnosti [7].

Kako bi se povećao procenat uspešnog prepoznavanja instrumenata izvršeno je istraživanje sa ciljem pronalaska

novih obeležja muzičkih signala koji bi se koristili kao ulazni parametri veštačkih neuralnih mreža. U ovom radu prikazana je metodologija za analizu audio sadržaja muzičkih signala koja se naziva hromatogram. Kada se govori o hromatogramu kao alatu za analizu muzičkih signala podrazumeva se da je na osnovu hromatograma moguće izvršiti analizu raspodele energije u vremenu datog audio zapisa, kao i ustanovljavanje energetskog bilansa na očekivanim parcijalima analiziranih tonova. Na osnovu hromatograma muzičkih signala dobijeni su hroma profili datih signala, gde hroma profil predstavlja prikaz raspodele energije signala u vremenu. Obrađivani su hroma profili duvačkih instrumenata: klarineta, flaute i oboe. Analizom hroma profila pomenutih instrumenata uočeni su različiti bilansi energije za različite instrumente. Na osnovu uočenih razlika u hroma profilima instrumenata pokazalo se da hroma profil predstavlja adekvatno obeležje jednog instrumenta, kao i povoljan ulazni parametar veštačke neuralne mreže. Kao izlazni produkt veštačke neuralne mreže ostvareni su visoki procenti prepoznavanja duvačkih muzičkih instrumenata. Cilj ovog rada je da pokaže da hroma profil predstavlja adekvatno obeležje na osnovu kog se može obučavati veštačka neuralna mreža. Takođe, ovako trenirana mreža postiže veoma visoke procenete uspešnog prepoznavanja instrumenata koji pripadaju istoj familiji, odnosno duvačkim instrumentima.

Rad je organizovan kako sledi. U drugom poglavlju prikazani su hromatogram i hromaprofil kao metodologije za analizu audio signala. Takođe prikazane su osnovne osobine duvačkih instrumenata koji su analizirani. U narednom poglavlju prikazani su eksperimentalni rezultati i diskusija dobijenih rezultata. Na kraju dat je zaključak o mogućnosti korišćenja hroma profila kao obeležja na osnovu kog se može obučavati neuralna mreža.

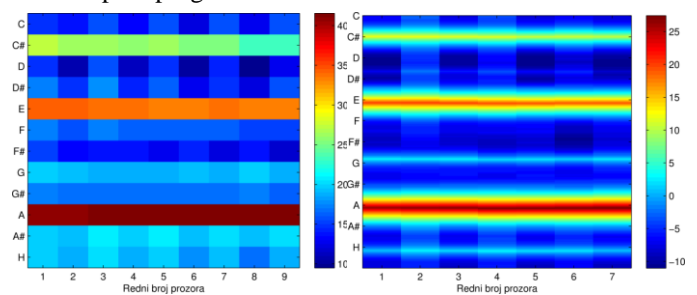
II. METODOLOGIJA

A. Hromatogram i Hroma profil

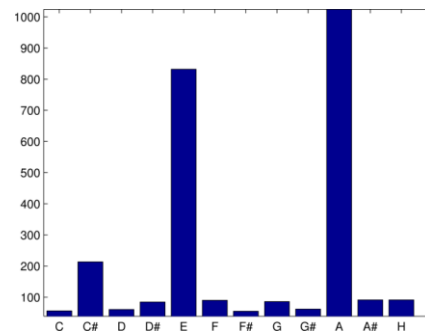
U procesima automatskog prepoznavanja sadržaja audio signala koriste se različita obeležja zasnovana na analizi spektralnog sadržaja koja svode kompletan spektar signala na opseg jedne oktave. Jedan od alata na osnovu kojeg se izdvajaju obeležja spektralnog sadržaja signala je hromatogram. Hromatogram predstavlja sredstvo za analizu raspodele energije signala unutar određenog broja intervala jedne oktave. Kako je akcenat istraživanja na muzici zapadne kulture sa temperovanim skalama posmatrana je raspodela energije signala unutar 12 intervala koji odgovaraju polutonovima unutar jedne oktave. U tom slučaju hromatogram predstavlja vremensko-frekvencijski prikaz signala gde je kompletan spektar signala transponovan na jednu izabranu oktavu koja je izdvojena na 12 opsega.

Izračunavanje hroma karakteristika jednog muzičkog signala se svodi na množenje hroma matrice i spektrograma datog signala, čime se dobija nova matrica koja se naziva hromatogram [8]. Spektrogram signala predstavlja matricu S dimenzija $N \times M$ dobijenu pomoću kratkovremene Furijeove transformacije ($STFT$). Redovi matrice S odgovaraju svakoj

od N frekvencija i imaju indekse k , $k \in [0, N-1]$. Broj kolona matrice S odgovara broju prozora u vremenu M . Hroma matrica C je dimenzija $K \times N$, gde K predstavlja broj opsega u okviru jedne oktave [9]. Vrednost parametra K je proizvoljna, ali se uglavnom uzimaju vrednosti 12 i 1200. Vrednost 12 predstavlja broj polutonova koji čine temperovanu skalu, dok se vrednost 1200 dodeljuje u slučaju kada je poželjno da se postigne rezolucija od jednog centa. Podela na cente podrazumeva da je u okviru jedne oktave definisano 1200 intervala čija je širina $2^{1/1200}$. Redovi matrice C dobijaju se sumiranjem vektora c svih oktava koji odgovaraju traženoj noti. Na Slici 1a) prikazan je izgled hromatograma sa 12 podopsega za ton A4 odsviran na oboi. Ukoliko je pri analizi sadržaja audio signala neophodno imati uvid u tačnost intonacije svakog tona bolje je koristiti prikaz hromatograma sa rezolucijom jednog centa, odnosno sa 1200 podopsega unutar oktave. Na Slici 1b) prikazan je izgled hromatograma sa 1200 podopsega za ton A4 odsviran na oboi.



Slika 1. Izgled hromatograma za ton A4 odsviran na oboi sa: a) 12 i b) 1200 podopsega



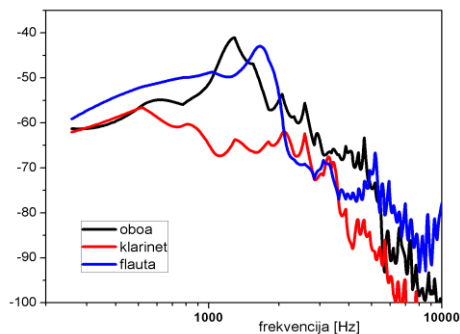
Slika 2. Hroma profil sa 12 podopsega tona A4 odsviranog na oboi

Pored hromatograma, na osnovu kog se može posmatrati raspodela energije signala unutar oktave, postoji još jedno sredstvo za analizu raspodele energije i naziva se hroma profil [10]. Hroma profil predstavlja jedan vid hromatograma. Naime, sumiranjem hromatograma za sve vremenske prozore audio signala dobija se hroma profil datog audio signala. Potreba za pojavom hroma profila audio signala javila se usled želje za preglednijim prikazom relativnih odnosa energije na pojedinim tonovima unutar oktave. Na Slici 2 prikazan je hroma profil sa 12 podopsega tona A4 odsviranog na oboi.

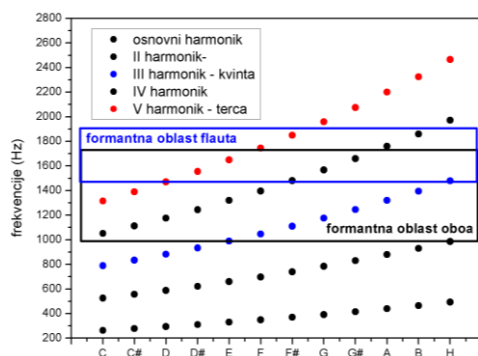
B. Obvojnica spektra drvenih duvačkih instrumenata

Kod tonova iste tonske visine koji su proizvedeni na različitim instrumentima uočavaju se određene razlike kako u vremenskoj obvojnici signala, tako i u obvojnici spektra. U

procesu nastanka zvuka kod drvenih duvačkih instrumenata mnogo faktora utiče na konačnu vremensku i spektralnu strukturu zvuka, kao i opseg koji instrument može da reprodukuje. Najznačajniji faktori su način pobude, poprečni presek i završetak cevi instrumenta [10]. Kao posledica svih tih faktora u ukupnom spektru zvuka vidljive su razlike koje se ispoljavaju u manje ili više istaknutim formantnim oblastima, gde formantna oblast definiše odnos energije na pojedinim intervalima u hroma profilu.



Slika 3. Usrednjeni spektar svih tonova u okviru celog opsega instrumenata



Slika 4. Frekvencije prvih 5 parcijala tonova iz opsega od C4 do H4

Na osnovu hromatske skale odsvirane u kompletnom opsegu, koji pokrivaju posmatrani instrumenti, klarinet, flauta i oboja, izvršena je estimacija spektra zvuka koji daju pojedinačni instrumenti [10]. Rezultat je prikazan na Slici 3. Spektar zvuka klarineta nema izražene formantne oblasti i ima ujednačen spektar u širem frekvencijskom opsegu u kome se nalaze osnovne frekvencije tonova koje klarinet može da proizvede, kao i viši harmonici tonova iz opsega od C4 do H4. Spektar zvuka oboe karakteriše pojava izražene formantne oblasti u opsegu od 1000 Hz do 1700 Hz sa maksimumom na 1300 Hz, dok je kod flaute taj maksimum pomeren na nešto više frekvencije i nalazi se u oblasti između 1450 Hz i 1900 Hz. Rečeno je da formantna oblast definiše odnos energije na pojedinim intervalima u hroma profilu. Na Slici 4 označene su frekvencije prvih 5 parcijala tonova iz opsega C4-H4. Takođe, su označene zone u kojima se nalaze istaknute formantne oblasti kod flaute i oboe. Plavim i crnim kvadratom uokvireni su frekvencijski opsezi koji odgovaraju formantnim oblastima vidljivim u spektru flaute i oboe. U okviru prvih pet parcijala tri odgovaraju osnovnom tonu, dok treći parcijal odgovara čistoj kvinti i peti parcijal velikoj terci osnovnog tona. Sa slike se može uočiti da se u formantnoj oblasti oboe

dominantno nalaze parcijale trećeg harmonika, koje odgovaraju čistoj kvinti, dok se u formantnoj oblasti flaute nalaze parcijali petog harmonika koji odgovaraju velikoj terci.

C. Neuralna mreža

Veštačke neuralne mreže (*Artificial Neural Networks*) predstavljaju grupu računarskih modela. Izvorno, neuralne mreže inspirisane su biološkim strukturama, posebno strukturom mozga zbog njegove sposobnosti da rešava kompleksne probleme [11]. Veštačke neuralne mreže dobile su ime zahvaljujući njihovoj strukturi koju čini pregršt procesorskih jedinica, koje su poznate kao neuroni organizovani po slojevima. Svaki od slojeva povezan je sa prethodnim i sledećim slojem i na taj način formira se mreža sačinjena od povezanih neurona. Postoje tri tipa slojeva neuralne mreže: ulazni, skriveni i izlazni. Ulazni sloj predstavlja vezu sa polaznim parametrima neuralne mreže. Skriveni slojevi su zaduženi za transformaciju ulaznih podataka u prihvatljive informacije za izlazni sloj. U zavisnosti od broja skrivenih slojeva može se meriti sofisticiranost informacija koje su predviđene za izlazni sloj [12].

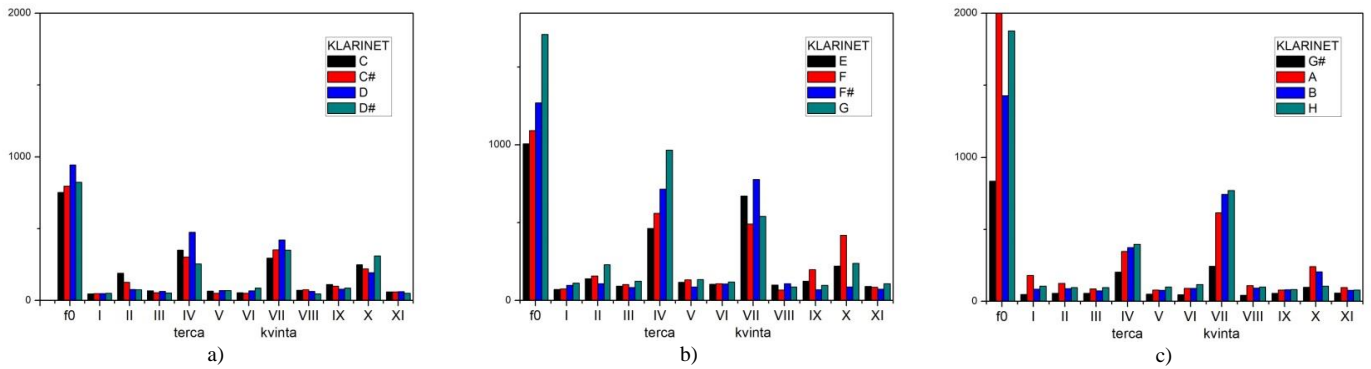
III. EKSPERIMENTALNI REZULTATI I DISKUSIJA

A. Baza snimaka

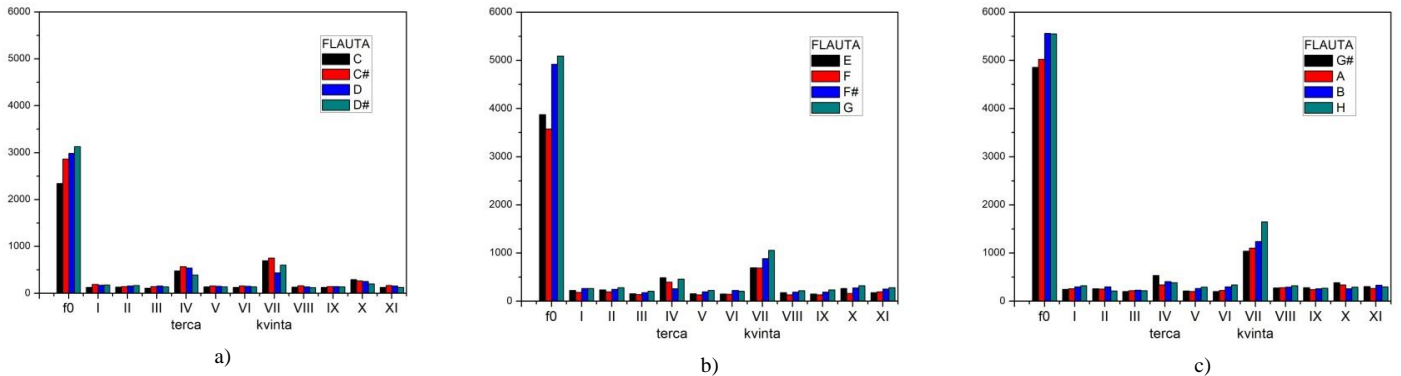
Za potrebe ovog rada napravljena je baza snimaka koja je korišćena za treniranje neuralne mreže za prepoznavanje duvačkih instrumenata. Bazu čine snimci, ukupno njih 216, gde je 48 snimaka preuzeto sa internet stranice [13], dok je preostalih 168 snimaka snimljeno u Laboratoriji za akustiku Elektrotehničkog fakulteta. Audio snimci koji čine bazu odsvirani su na: klarinetu, flauti i oboi. Snimci koji se nalaze u bazi su izolovani, tj. ne predstavljaju polifoni audio zapis, te tako predstavljaju idealan slučaj za prepoznavanje i izdvajanje obeležja audio zapisa muzičkog sadržaja. Audio snimci koji su preuzeti sa internet stranice predstavljaju tonove hromatske lestvice celokupnog registra posmatranih instrumenata. Takođe, snimci se odlikuju različitim trajanjem: 1 minuta, 0.5 i 0.25 sekundi. Odsvirani su snimci u tri različita dinamička manira: forte, piano i pianissimo. Ostatak snimaka koji čine bazu snimani su u Laboratoriji za akustiku elektrotehničkog fakulteta pomoću sledeće aparature: mikrofona *MINI SPL* [14], zvučne kartice *Steinberg* [15] i računara. Takođe, i snimljeni audio sadržaj predstavlja sve tonove hromatske lestvice koje pomenuti instrumenti mogu da proizvedu. Ovakvo formirana baza je značajna jer se može koristiti i u daljim istraživanjima.

B. Hroma profil duvačkih instrumenata

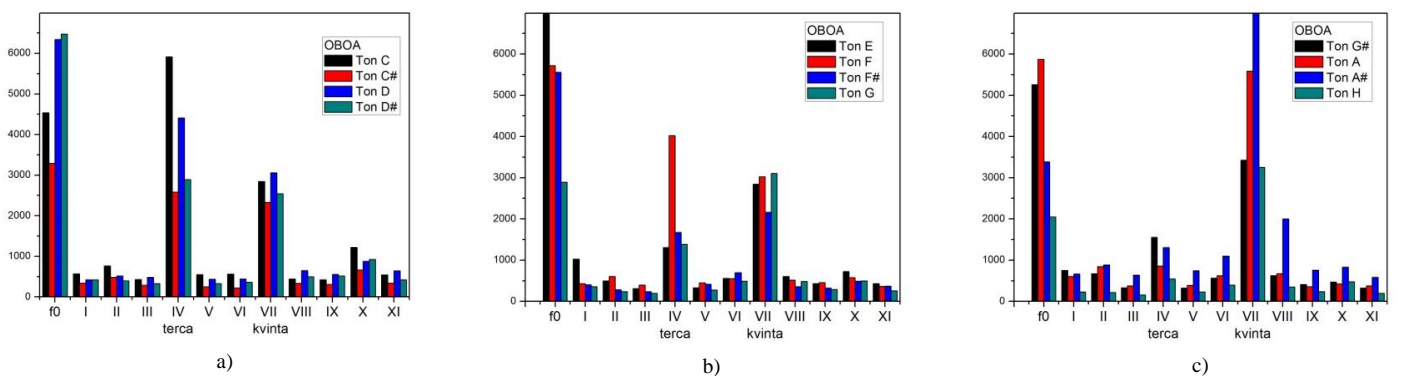
Uzmajući u obzir razlike u načinu nastanka tona na pojedinim posmatranim instrumentima, analizirani su hroma profili tonova odsviranih u kompletnom opsegu sa ciljem da se ustanove njihove međusobne razlike koje bi mogle da posluže kao osnov za prepoznavanje instrumenata u zadacima iz oblasti automatske analize muzičkog sadržaja. Analizom su obuhvaćeni tonovi hromatskih lestvica u celom opsegu koji posmatrani instrumenti mogu da proizvedu.



Slika 5. Hroma profili tonova iz opsega C4-H4 odsviranih na klarinetu



Slika 6. Hroma profili tonova iz opsega C4-H4 odsviranih na flauti

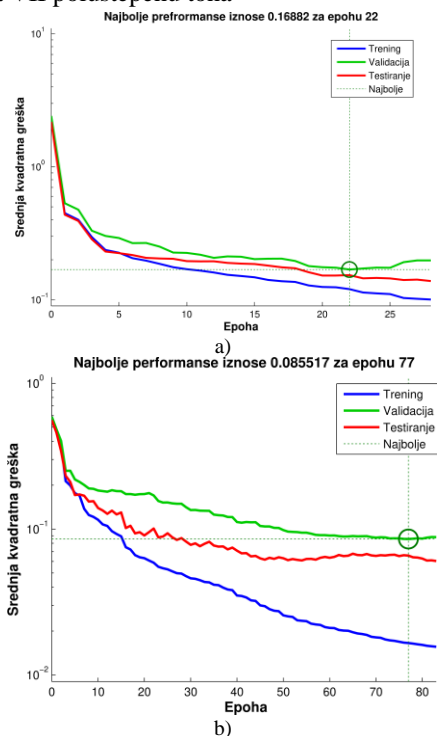


Slika 7. Hroma profili tonova iz opsega C4-H4 odsviranih na oboi

Hromatske lestvice odsvirane su bez dodatne intonacije, tako da je analizi prethodila provera tačnosti intonacije svih tonova kao uslov za primenu hromatske analize. U radu su detaljno analizirani tonovi iz oktave koja obuhvata tonove od C4 do H4 jer su u toj oktavi bile u svim slučajevima najtačnije intonacije. Na Slikama 5, 6 i 7 prikazani su hroma profili za 12 tonova odsviranih na klarinetu, flauti i oboi, respektivno. Zarad preglednosti i lakšeg uočavanja raspodele energije slike su podeljene na tri dela: a), b) i c). Na osnovu Slike 5. i hroma profila klarineta može se uočiti da kod klarineta energetski ne dominira osnovni ton, već se javlja relativno velika energija na IV i VII polustepenu. Takođe, kod hroma profila klarineta karakteristična je relativna izraženost energije na X polustepenu. Ovakva raspodela energije može se objasniti

obvojnicom spektra klarineta u kome postoji određen nedostatak energije u zoni prvih parcijala za tonove iz opsega četvrtre oktave, ali i relativno ujednačen spektar na višim frekvencijama kao posledica uticaja piska. Na Slici 6. prikazani su hroma profili tonova flaute. Na osnovu slike se može zaključiti da hroma profili tonova flaute imaju najujednačeniju strukturu unutar cele posmatrane oktave. Formantna oblast kod flaute je na nešto višim frekvencijama, oko 1700 Hz, i kao takva ona ima manje uticaja na odnos energije prvih značajnih parcijala tonova iz opsega četvrtre oktave. Za razliku od hroma profila klarineta, kod hroma profila flaute može se uočiti veoma izražena energija na osnovnom tonu koja dominira u odnosu na energiju na IV i VII polustepenu.

Na Slici 7. prikazani su hroma profili tonova oboe. Sa slike se može zaključiti da zbog postojanja izražene formantne oblasti na frekvenciji od oko 1200 Hz postoji različita raspodela energije unutar pojedinih intervala. Razlika u raspodeli energije kod tonova oboe pripisuje se činjenici da se unutar formantne oblasti instrumenta nalazi veliki broj parcijala posmatranih tonova. Kod tonova C4, C#4, D4, D#4, F4, koji čine donji tetrahord hromatske lestvice, veoma je ujednačena energija na osnovnom tonu, IV i VII polustepenu. Nasuprot, tome kod druge grupe tonova G4, A4, A#4, H4, koji čine gornji tetrahord hromatske lestvice, počinje da dominira energija na VII polustepenu tona

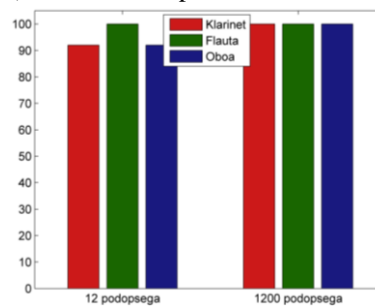


Slika 8. Prikaz srednje kvadratne greške za trening, validacioni i test podskup za a) 12 i b) 1200 ulaznih parametara

C. Rezultati prepoznavanja duvačkih instrumenata

Kao klasifikator instrumenata korišćena je MLP (*Multilayer Perceptron*) *feed forward* neuralna mreža sa propagacijom unazad. Struktura mreže se sastoji od ulaznog sloja koji je povezan sa početnim parametrima neuralne mreže, odnosno sa hroma profilima instrumenata. U zavisnosti od toga da li su kao ulazni parametri korišćeni hroma profili tonova sa 12 ili 1200 podsega, broj ulaznih parametara neuralne mreže bio je 12, odnosno 1200. Broj neurona koji je korišćen u skrivenom sloju je 40. Izlaz neuralne mreže predstavlja proces dekodovanja koji karakteriše koji je instrument prepoznat. Takođe, neuralna mreža je potpuno povezana, odnosno svaki neuron je povezan sa svakim neuronom iz sledećeg sloja bilo direktno ili indirektno. Skup podataka koji se koristi za prepoznavanje podeljen je na podskupove za trening, validaciju i test. Podskup za trening čini 70% skupa podataka, dok su podskupovi za validaciju i test jednaki i iznose 15% skupa podataka. U procesu treniranja neuralne mreže korišćen je skalirani konjugovani gradijentni

algoritam (*Scaled Conjugate Gradient Algorithm*). Kao aktivaciona funkcija neurona korišćena je tangent sigmoid funkcija (*Tansig function*) [16]. U slučaju kada su ulazni parametri hroma profili tonova sa 12 podsega broj epoha je 27, a vrednost *learning* parametra 0.08. Dok za slučaj kada su ulazni parametri hroma profili tonova sa 1200 podsega broj epoha je 83, a vrednost *learning* parametra 0.1. Pri treniranju mreže nije korišćena regularizacija. Na slici 8. prikazane su performanse obučavane mreže, gde je prikazana zavisnost srednje kvadratne greške za sva tri podskupa (trening, validacioni i test) u odnosu na epohu.



Slika 9. Rezultat prepoznavanja duvačkih instrumenata na osnovu hroma profila od 12 i 1200 podsega

Na Slici 9 prikazan je rezultat prepoznavanja duvačkih instrumenata na osnovu hroma profila sa 12 i 1200 podsega koji su predstavljali ulazne parametre neuralne mreže. Kada su kao ulazni parametri neuralne mreže uzeti hroma profili tonova sa 12 podsega procenti prepoznavanja za klarinet i oboe iznose 92%, dok procenti prepoznavanja za flautu iznose 100%. Činjenica da je flauta prepoznatna sa stopostotnim učinkom odlikuje se upravo njenim ujednačenim hroma profilom i time što se on odlikuje dominantnom energijom na osnovnom tonu dok je procenat energije na ostalim parcijalima znatno niži u odnosu na energiju na osnovnom tonu. U slučaju klarineta i oboe procenti prepoznavanja su nešto niži, zbog hroma profila tonova iz viših oktava. Naime, tonovi koji pripadaju višim oktavama ovih instrumenata ne poseduju drastične razlike u hroma profilima, već imaju sličan bilans energije na osnovnom tonu i frekvencijama koje odgovaraju IV i VII polustepenu hromatske lestvice. Takođe, na osnovu usrednjenog spektra u okviru celog frekvencijskog opsega instrumenata uočava se da su obvojnice spektara klarineta i oboe veoma bliske na visokim frekvencijama, što predstavlja potvrdu male razlike hroma profila ovih instrumenata na visokim frekvencijama.

Za slučaj kada se kao ulazni parametar neuralne mreže koristi hroma profil tonova sa 1200 podsega procenti prepoznavanja za sva tri instrumenta je maksimalan i iznosi 100%. Računanjem hroma profil tonova sa 1200 podsega unutar oktave postignuta je rezolucija od jednog centa koja doprinosi uvid u tačnost intonacije svakog tona. Te na osnovu tako računatih hroma profila moguće je bilo uočiti veoma male razlike u hroma profilima klarineta i oboe za tonove iz viših oktava, koji su predstavljali prepreku za maksimalni učinak neuralne mreže pri prepoznavanju instrumenta. Iako je procenat prepoznavanja instrumenata za slučaj kada ulazne parametre predstavljaju hroma profili tonova sa 1200

podopsega idealan, ovakvi ulazni parametri neuralne mreže zahtevaju više vremena za obradu. Stoga, u pogledu vremenske kompleksnosti neuralna mreža kod koje su ulazni parametri hroma profili tonova sa 12 podopsega ima prednost za korišćenje u aplikacijama koje rade u realnom vremenu. Time se postiže ušteda u vremenu na štetu nešto nižem ostvarenom procentu uspešnog prepoznavanja instrumenta od 92%.

IV. ZAKLJUČAK

U ovom radu prikazana je upotreba hroma profila tonova kao novog obeležja izdvojenog iz audio zapisa muzičkog sadržaja duvačkih instrumenata. Takva obeležja korišćena su za prepoznavanje duvačkih instrumenata pomoću neuralnih mreža. Duvački instrumenti koji su prepoznavani su: klarinet, flauta i oboa. Ovi instrumenti odabrani su tako da pokrivaju širok opseg varijacija u pogledu načina pobuđivanja, oblika cevi i strukture rezonantnih modova. Predložena metodologija karakterizacije instrumenata bazirana na hroma profilu tonova pokazala je očuvanje osnovnih razlika u karakteristikama pojedinih instrumenata. Prikazane razlike između instrumenata pružaju dobru osnovu za upotrebu hroma profila kao obeležja na osnovu kojih neuralna mreža vrši prepoznavanje instrumenta. Ostvareni su visoki procenti prepoznavanja instrumenta za slučaj kada su kao ulazni parametri neuralne mreže uzeti hroma profili tonova sa 12 i 1200 opsega. Iako prikazana neuralna mreža ostvaruje stopostotni učinak u prepoznavanju instrumenata za slučaj kada su ulazni parametri predstavljali hroma profile tonova sa 1200 podopsega, uočeno je da takvi parametri zahtevaju više vremena za njihovu obradu. Te u smislu vremenske kompleksnosti rada neuralne mreže povoljnije je kao ulazne parametre koristiti hroma profile tonova sa 12 podopsega, gde je rezultat ovako trenirane mreže veoma dobar i iznosi 92% uspešnog prepoznavanja za slučaj klarineta i oboe i 100% uspešnog prepoznavanja u slučaju flaute. Cilj rada bio je ispitati mogućnost upotrebe hroma profila tonova duvačkih instrumenata kao obeležja audio zapisa muzičkog sadržaja na osnovu kog je moguće izvršiti prepoznavanje instrumenta. Pokazano je kroz ostvarene visoke procente prepoznavanja duvačkih instrumenata da je hroma profil tonova adekvatno obeležje koje se može koristiti u daljim istraživanjima. Tako bi se u nekim od budućih istraživanja uzeli u obzir hroma profili ne samo monofonih nego i polifonih audio zapisa.

LITERATURA

- [1] Meinard Mller, "Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications", 1st edition, Springer Publishing Company, Incorporated, 2015, ISBN:3319219448.
- [2] R. N. Shepard, "Circularity in judgements of relative pitch", J. Acoust. Soc. Amer., vol. 36, pp. 2346-2353, 1964.
- [3] C. Harte and M. Sandler, "Automatic chord identification using a quantised chromagram", In Proceedings of the 118th AES Convention, Barcelona, Spain, 2005.
- [4] D. L. Ishwar, K. Sethi, N. Dimitrova, T. McGeec, "Classification of general audio data for content-based retrieval", Pattern Recognition Letters, vol. 22(5), pp. 533-544, 2001.

- [5] J. Marques, P.J. Moreno, "A study of musical instrument classification using gaussian mixture models and support vector machines", Cambridge Research Laboratory Technical Report Series CRL,4, 1999.
- [6] A. Eronen, A. Klapuri, "Musical instrument recognition using cepstral coefficients and temporal features", In ICASSP'00, Vol. 2, pp. II753-II756, 2000.
- [7] N. Segal, "Automatic Musical Instrument Recognition Using Convolutional Neural Networks", MSc in Digital Signal Processing Project Report, 2016.
- [8] M. Muller, D. W. Ellis, A. Klapuri, G. Richard, "Signal Processing for Music Analysis", Journal of latex class files, Vol. 6, No. 1, January 2007.
- [9] M. Slavković-Ivić, D. Šumarac Pavlović, "Analiza muzičkih tonova pomoću hromatograma", ETRAN, Zlatibor, jun 2016, Zbornik radova, ISBN: 978-86-7466-618-0.
- [10] Tatjana Miljković, Miloš Bjelić, Dragana Šumarac Pavlović, Jovana Protić, Hromatogram tonova duvačkih instrumenata, ETRAN, Palić, Zbornik radova 88-93, ISBN: 978-86-7466-752-1, 2018.
- [11] I.A. Basheer, M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application", Journal of microbiological methods, 43(1), pp.3-31, 2000.
- [12] S. Haykin, "Neural Network A Competitive Foundation", 2th ed. Prentice Hall International, 1990.
- [13] http://www.philharmonia.co.uk/explore/sound_samples/clarinet
- [14] Tehnička dokumentacija proizvođača, dostupno na mreži: <http://www.nti-audio.com/Portals/0/data/en/MiniSPL-Measurement-Microphone-Product-Data.pdf>, pristupano 20.1.2019.
- [15] Tehnička dokumentacija proizvođača, dostupno na mreži: http://download.steinberg.net/downloads_hardware/UR22/UR22_documentation/UR22_OperationManual_en.pdf, pristupano 20.1.2019.
- [16] S. Masood, S. Gupta and S. Khan, "Novel approach for musical instrument identification using neural network.", Annual IEEE India Conference (INDICON), New Delhi, pp. 1-5, doi: 10.1109/INDICON.2015.7443497, 2015.

ABSTRACT

In the processes of automatic recognition of the content of audio signals, various features based on the analysis of spectral content are used. In this paper are presented the chroma tones profile as a new feature separated from music content on the basis of which recognition of wind instruments by neural network is performed. Wind instruments that were of interest to this research are: clarinet, flute and oboe. These instruments have been selected to cover a wide range of variations in terms of excitation, tube shape and resonance modes structure. The methodology for the characterization of wind instruments based on chroma tone profiles has shown the preservation of the basic differences in the characteristics of the instruments. The chroma profile represents relative energy ratios on individual tones within octave. Chroma tone profiles that were observed in this paper were divided into 12 and 1200 ranges. Using chroma tone profiles as the input parameters of the neural network high recognition percentages of the instrument were achieved.

WIND INSTRUMENTS RECOGNITION USING CHROMA PROFILE AND NEURAL NETWORK

Tatjana Miljković, Miloš Bjelić, Dragana Šumarac Pavlović,
Goran Kvašćev