

# Metoda za segmentaciju muzičkog signala na osnovu matrice samosličnosti

Marija Ratković, Matija Marijan, Tatjana Miljković, Miloš Bjelić

**Apstrakt**—Rad se bavi problemom automatske segmentacije i analizom strukture muzičkih dela na bazi izdvojenih obeležja iz audio zapisa dela. Složenost problema segmentacije muzičkih dela posledica je više elemenata po kojima se delovi razlikuju, a to mogu biti melodija, harmonija, ritam, i boja. Kao osnova za segmentaciju u radu je korišćena matrica samosličnosti formirana na bazi izabranih obeležja signala. Kao obeležja korišćena su MFCC (eng. *Mel-Frequency Cepstral Coefficients*) i hroma obeležja koja prate promene boje i harmonijske strukture. Predložen je postupak segmentacije matrice samosličnosti zasnovan na jednostavnim metodama za morfološku obradu slike. Algoritam dinamičkog usklađivanja u vremenu (eng. *Dynamic Time Warping*, DTW) korišćen je kao alat za procenu stepena sličnosti segmentiranih delova. Predložena procedura primenjena je na 2 dela klasične muzike.

**Ključne reči**—DTW algoritam; hroma obeležja; matrica samosličnosti; MFCC obeležja; muzička struktura.

## I. UVOD

Osnovu svakog muzičkog dela čine pojedinačni muzički događaji, pojedinačne note koji se mogu okarakterisati sa nekoliko atributa kao što su njihov intenzitet, visina, boja, trajanje, i vremenski intervali između pojedinačnih nota [1]. Čitava struktura muzičkog dela se na osnovnom nivou karakteriše melodijom, harmonijom i ritmičkom strukturom. Kombinacijom pojedinačnih događaja koji se sukcesivno pojavljuju u vremenu gradi se melodijska struktura dela. Istovremena pojava više tonova definiše harmonijsku strukturu dela, a trajanje, broj, i način ponavljanja pojedinačnih događaja određuje ritmičku strukturu.

Na sledećem nivou muzička dela sastoje od manjih celina koje su na određeni način međusobno organizovana u odgovarajuće hijerarhijske strukture unutar kojih imaju specifične funkcije [1]. Struktura muzičkog dela se razlikuje kod klasične i savremene muzike, ali se u svakoj kategoriji prepoznaju delovi koji se nalaze u međusobnom odnosu i ostvaruju određene funkcije [1]. U popularnoj muzici to su najčešće uvod, refren i strofe koji se ponavljaju uzastopno [1, 2]. U klasičnoj muzici to su najčešće ekspozicija, razvoj i recapitulacija, kod forme sonata [3]. U vreme baroka,

Marija Ratković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (e-mail: [rm223039m@student.etf.bg.ac.rs](mailto:rm223039m@student.etf.bg.ac.rs)),

Matija Marijan – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (e-mail: [matijamarijan99@gmail.com](mailto:matijamarijan99@gmail.com)),

Tatjana Miljković – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (e-mail: [tm@etf.rs](mailto:tm@etf.rs)),

Miloš Bjelić – Elektrotehnički fakultet, Univerzitet u Beogradu, Bulevar kralja Aleksandra 73, 11120 Beograd, Srbija (e-mail: [bjelic@etf.rs](mailto:bjelic@etf.rs)).

popularna je bila *rondo* struktura, kod koje se nekoliko različitih tema smenjuju oko glavne teme [4]. Sa izuzetkom nekih avangardnih kompozicija i žanrova, muziku generalno karakterišu struktura i repeticija [5]. Izdvajanje segmenata različite strukture i funkcije se u klasičnoj analizi muzičkih dela vrši manuelno.

Automatska analiza strukture muzičkog dela ima za cilj prepoznavanje i izdvajanje vremenskih segmenata određenih karakteristika na bazi talasnog oblika signala koji je predstavljan odgovarajućim obeležjima. Takođe, analiza i segmentacija strukture muzičkog dela može biti korisna pri vizuelizaciji strukture kompleksnih muzičkih dela [6].

Najčešće se karakterišu tri različita pristupa segmentaciji muzičkog dela, a to su segmentacija zasnovana na detekciji ponavljajućih delova, segmentacija zasnovana na detekciji novina, odnosno tranzicija između različitih delova, i na kraju segmentacija zasnovana na detekciji homogenosti, odnosno detekciji delova koji su međusobno konzistentni prema nekoj muzičkoj osobini [1]. Konkretni kriterijum koji se može usvojiti za segmentaciju može biti konzistentnost boja i detekcija prisustva različitih instrumenata i harmonija, i analiza promena u tempu, dinamici, ili tonalitetu.

Zavisno od kriterijuma za segmentaciju koriste se različita obeležja koja odražavaju neke od pomenutih karakteristika. Za segmentaciju muzičkih dela najčešće se koriste spektralna MFCC (eng. *Mel-Frequency Cepstral Coefficients*) obeležja, hroma obeležja, i obeležja zasnovana na analizi tempa [1].

Neki od algoritama dostupnih u literaturi su zasnovani na analizi semantičke strukture i tekstualnih informacija kod vokalne muzike [7], na kombinaciji melodijskih, ritmičkih, i strukturalnih obeležja muzičkog dela [8], na metodama dubokog učenja za segmentaciju i labeliranje segmenata muzičkih dela [9], i druge. Metode za segmentaciju zasnovane na analizi matrice samosličnosti se najčešće bave analizom matrice kao slike i njenom obradom. Neke od metoda za obradu i analizu matrice samosličnosti obuhvataju filtriranje specifičnim kernelima za dijagonalnu korelaciju [10, 11], gausovskim kernelima za filtriranje unapred po dijagonali uz klasterizaciju pomoću singularne dekompozicije [12], isticanje izraženo sličnih segmenata filtrima za usrednjavanje [13], posmatranje standardne devijacije segmenata matrice samosličnosti [14], klasterizaciju na segmente pomoću skrivenih Markovljevih modela [15], i druge.

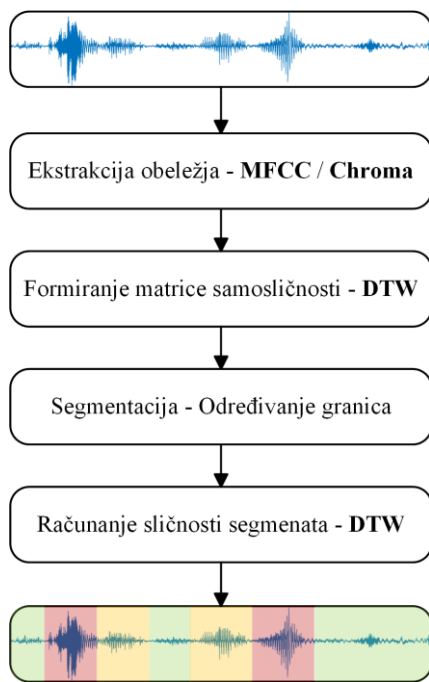
Segmentacija delova vrši se na bazi matrice samosličnosti formirane na bazi izabranih obeležja, za koje je pokazano da mogu biti koristan alat pri analizi audio signala [13]. Korišćena su MFCC i hroma obeležja. Algoritam za segmentaciju je zasnovan na jednostavnim metodama za

obradu slike, sa ciljem da se izdvoje delovi na kojima se može primetiti izražena sličnost, i na osnovu njih, odrediti granice između segmenata muzičkog dela, te se ovaj algoritam može okarakterisati kao *repetition-based* algoritam za segmentaciju. Međusobni odnosi segmenata se definišu primenom DTW (eng. *Dynamic Time Warping*) algoritma na podmatrice određene granicama dobijenim iz segmentacije.

U drugom poglavlju je detaljno opisana implementacija algoritma za analizu strukture muzičkog dela. Primeri rezultata segmentacije matrica samosličnosti za dva primera muzičkih dela su prikazani u poglavlju III, zajedno sa diskusijom rešenja. Konačno, u poglavlju IV su dati zaključci o uspešnosti opisanog algoritma, o važnosti izabranih obeležja i njihovim razlikama, kao i razmatranja o budućem razvoju algoritma.

## II. METODOLOGIJA ANALIZE

Predložena metoda za analizu strukture muzičkog dela u ovom radu je zasnovana na ekstrakciji obeležja iz audio zapisa, nakon čega se formira matrica samosličnosti uz DTW algoritam. Za podelu muzičkog dela na delove se koriste metode za obradu i segmentaciju slike, odnosno matrice samosličnosti. Na podeljenim segmentima se zatim uz pomoć DTW algoritma analizira međusobna sličnost i vizuelno se vrši procena efikasnosti podele na segmente. Algoritam za analizu i segmentaciju muzičkih signala je realizovan u okviru *MATLAB* softverskog okruženja [16]. Na Sl. 1 je prikazan blok dijagram izvršavanja algoritma za analizu strukture muzičkog dela.



Sl. 1. Blok šema korišćenog algoritma.

### A. Muzička obeležja

Kako bi muzički podaci mogli kvantitativno da se porede i računarski analiziraju, prvi korak u postupcima obrade muzičkih signala predstavlja izdvajanje odgovarajućih obeležja relevantnih za analizu. U ovom radu su za analizu

korišćena i hroma obeležja, i kepralni koeficijenti mel frekvencijske skale (MFCC).

Hroma obeležja govore o udelu i promeni različitih muzičkih tonova raspodeljenih po klasama u okviru nekog muzičkog signala. Svaka klasa tonova za hroma obeležja predstavlja sumiranu snagu muzičkog signala za jedan ton kroz sve frekvencijske opsege tog tona, odnosno kroz različite oktave sadržane u tom muzičkom signalu. Za segmentaciju muzičkog dela hroma obeležja mogu biti od koristi za izdvajanje granica u kojim dolazi do promene tonaliteta [17].

Boja tona je posebno svojstvo nekog zvuka koje omogućava ljudima da razaznaju muzičke instrumente u nekom audio signalu [17]. Spektralna MFCC obeležja, zasnovana na transformaciji spektra snage muzičkog signala, se često koriste u svrhu razlikovanja boje zvuka i izdvajanja informacija iz muzičkih signala. Čovekovo čulo sluha je osetljivije na nižim nego na višim frekvencijama [18]. Raspon frekvencija iz čulnog opsega čoveka se ne deli po linearnoj skali već po Mel frekvencijskoj skali. Ova skala bolje aproksimira odziv čovekovog čula sluha, zbog čega se često koriste za obeležja pri prepoznavanju govora, ali i za klasifikaciju različitih žanrova, merenje sličnosti, i detekciju promene u boji muzičkih signala.

### B. DTW algoritam

Cilj DTW algoritma je određivanje optimalnog vremenskog poravnanja između dve sekvence koje se porede, i on se može koristiti za sinhronizaciju dve različite interpretacije ili izvedbe iste muzičke sekvence. Takođe, kao što je prikazano u ovom radu, mogu se koristiti za ocenjivanje sličnosti dve muzičke sekvence. Sekvence koje se porede su diskretna obeležja muzičkih signala. Za poređenje dve sekvence obeležja potrebno je definisati lokalnu funkciju cena,  $c(x, y)$ . Ova funkcija će predstavljati meru sličnosti, element  $x$  jedne sekvence i element  $y$  druge sekvence. Ako su  $x$  i  $y$  slični cena je mala, a ukoliko se razlikuju cena je veća. Obeležja koja čine sekvence su višedimenzionalni vektori koji opisuju muzičku sekvencu. U ovom radu, cena,  $c(x, y)$ , između dva elementa,  $x$  i  $y$ , je definisana kao kosinusna distanca prikazana u izrazu:

$$c(x, y) = 1 - \frac{\langle x | y \rangle}{\|x\| \cdot \|y\|} . \quad (1)$$

Za sve parove elemenata  $x$  i  $y$  iz celokupnih sekvenci obeležja  $X$  i  $Y$  određuju se cene koje definišu matricu cena kao u izrazu:

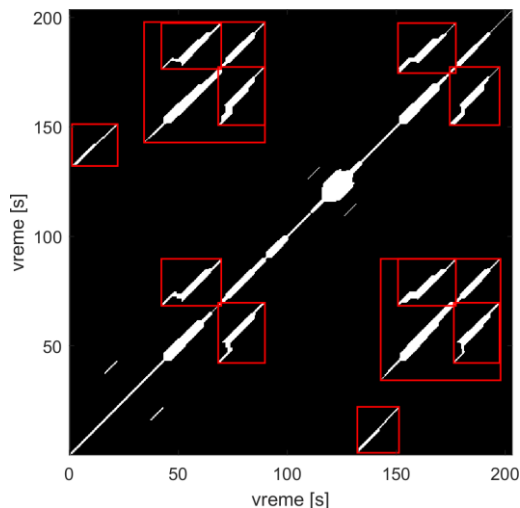
$$C(n, m) = c(x_n, y_m) , \quad (2)$$

gde  $n$  i  $m$  predstavljaju redne brojeve elementa sekvenci  $X$  i  $Y$ . Cilj je naći vremensko poravnanje između sekvenci na osnovu matrice cena. Poravnanje je definisano putanjom, odnosno nizom polja matrice cena, tako da ukupna cena cele putanje bude najmanja moguća. Za pronalaženje optimalne putanje koristi se algoritam baziran na dinamičkom programiranju. Potrebno je definisati akumuliranu matricu cena  $D$ . Svaka vrednost  $D(n, m)$  označava akumuliranu cenu optimalne putanje počevši od početnih indeksa do koordinata  $(n, m)$ . Ukupna cena sličnosti dve sekvence je  $D(N, M)$ , gde su  $N$  i  $M$  dimenzije matrice cena odnosno dužine sekvenci obeležja koje se porede. U ovom radu DTW algoritam je primenjen na

matrici samosličnosti, odnosno  $X$  i  $Y$  sekvence obeležja su iste. Ova matrica cena se formira da bi se uočili segmenti unutar muzičke sekvence koji su međusobno slični. Na osnovu grafičkog prikaza matrice cena vrši se segmentacija, odnosno definisanje granica segmenata unutar sekvence koji će se međusobno porediti.

### C. Segmentacija

Predložena metrika u ovom radu za detekciju i izdvajanje granica između različitih segmenata u strukturi muzičkog dela je međusobna sličnost tih segmenata. Granice između segmenata se analiziraju i određuju metodama za obradu slike, posmatrajući matricu samosličnosti. Postupak automatske segmentacije muzičkog dela, na osnovu matrice sličnosti započinje normalizacijom vrednosti matrice samosličnosti. Kako bi se izdvojile vrednosti matrice samosličnosti koje imaju značajno izraženu sličnost, vrši se binarizacija vrednosti u odnosu na unapred izabran prag intenziteta, koji je moguće podešavati. Za dalju morfološku obradu slike koristi se istidijagonalni strukturni element, oblika jedinične matrice, sa dimenzijama koje se unapred biraju, i mogu se podešavati. Redom se nad binarizovanom matricom sličnosti vrše morfološke operacije zatvaranja, otvaranja slike, popunjavanja rupa, i ponovnog zatvaranja slike [19]. Zatim se korišćenjem ugrađenih funkcija izdvajaju svi povezani objekti sa slike. Od svih izdvojenih objekata, zadržavaju se oni čije dimenzije odgovaraju unapred zadatom opsegu. Na Sl. 2 je prikazan primer izdvajanja objekata, gde su crvenom bojom su označene granice koje opisuju izdvojene objekte. One predstavljaju preliminarne granice između segmenata izabranog muzičkog dela. Kako bi se eliminisale granice koje mogu delovati neprirodno blisko, izračunava se srednja vrednost između svake dve bliske granice i tako izračunate granice se čuvaju kao finalne vrednosti granica između segmenata. Minimalna dozvoljena udaljenost dve granice je unapred zadata i može se podešavati.



Sl. 2. Primer izdvojenih objekata sa slike.

Nakon izvršene segmentacije, uz pomoć grafičkog prikaza matrice sličnosti i predloženih granica između segmenata dobijenih obradom slike, moguće je ručno podesiti, dodati ili ukloniti neke od granica, kao i promeniti parametre za segmentaciju, ukoliko one nisu precizno odabrane.

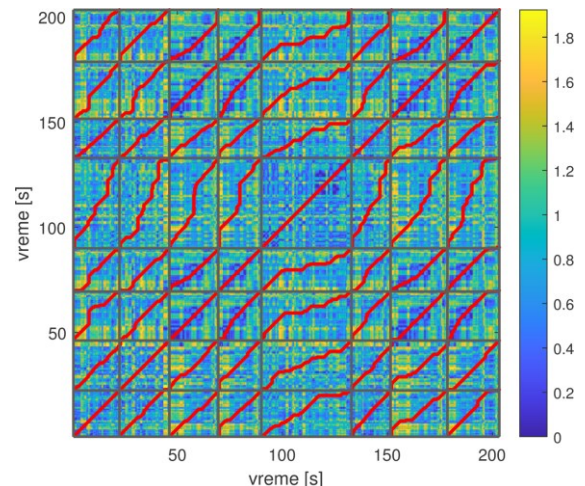
### D. Sličnost segmenata

Na osnovu dobijenih granica određeni su segmenti muzičkog dela koji se međusobno poredi. Matrica samosličnosti je podeljena na manje matrice za koje se računaju cene koje će biti pokazatelj da su određeni segmenti unutar muzičkog dela sličniji u odnosu na druge segmente. Za sve novodefinisane matrice je određena cena optimalnih putanja primenom DTW algoritma. Cene su normalizovane dužinom optimalne putanje, da bi mogle međusobno da se poredi. Ove cene predstavljaju meru sličnosti bilo koja dva segmenta muzičkog dela. Cena matrice segmenata koji su slični će biti manja u odnosu na cene matrice segmenata koji se razlikuju. Na osnovu ovih cena se može zaključiti koji segmenti muzičkog dela su međusobno slični.

U zavisnosti od vrste obeležja koja se analiziraju određuju se različite granice segmenata. Neki segmenti koji se izdvoje primenom segmentacije na matricu cena MFCC obeležja se ne izdvoje primenom segmentacije na matricu cena hroma obeležja i obrnuto. U ovom radu je predloženo kombinovanje granica dobijenih korišćenjem oba obeležja. Takođe, pokazano je da je primenom granica dobijenih za jednu vrstu obeležja na matrice cena za drugu moguće primetiti približno iste odnose cena segmenata za obe vrste obeležja. Koliko će se granice dobijene za obe vrste obeležja razlikovati zavisi od samog muzičkog dela koje se analizira.

## III. REZULTATI I DISKUSIJA

Predloženi algoritam je pokazan na dva primera klasičnih muzičkih dela. To su Mađarska igra br. 5 Johanesa Bramsa i Valcer br. 2 Dmitrija Šostakoviča. Kroz ova dva primera se mogu uočiti drugačije karakteristike algoritma. Za oba primera muzičkog dela prvo je potrebno odrediti MFCC i hroma obeležja.



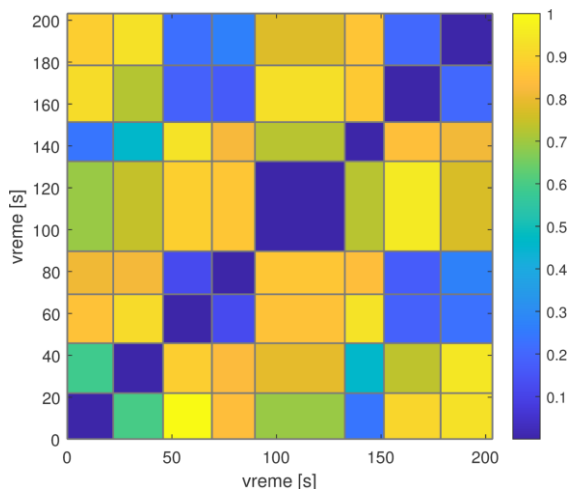
Sl. 3. Matrica samosličnosti za MFCC obeležja sa definisanim granicama i optimalnim putanjama podeljenih segmenata.

### A. Primer 1. Mađarska igra br. 5 – Johanes Brams

Na osnovu analize matrice samosličnosti, i uočavanjem jasnije izraženih sličnih segmenata, odlučeno je da će algoritam za segmentaciju biti primenjen na matrici samosličnosti MFCC obeležja. Sličniji delovi imaju manju vrednost cene i prikazani su tamnijom plavom bojom. Na

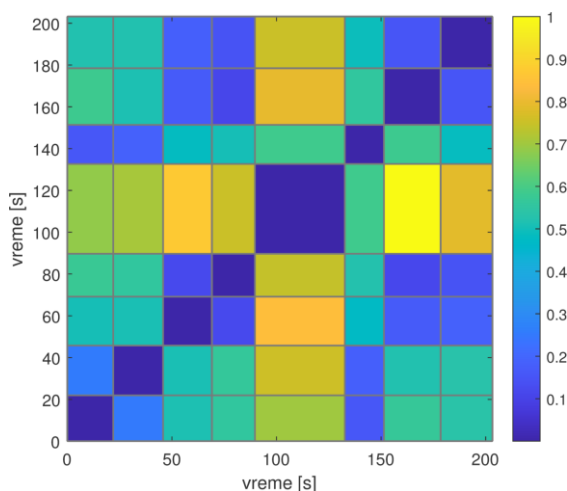


osnovu vrednosti sličnih delova koje su očitane sa grafika odlučeno je koji će biti prag intenziteta za izdvajanje objekata na slici. Na Sl. 3 su sivom bojom prikazane granice koje su određene i usvojene kao adekvatne za ovaj primer.



Sl. 4. Matrica cena međusobnih odnosa segmenata, granice izračunate na matrici MFCC obeležja i cene računute na matrici MFCC obeležja.

Na Sl. 4 su grafički predstavljene normalizovane cene koje označavaju međusobne odnose izdvojenih segmenata. Plavom bojom su prikazana polja koja pokazuju da su segmenti koji se porede međusobno slični, njihova cena je manja. Jasno se vide razlike u bojama kod međusobno sličnih segmenata i međusobno različitih segmenata. Polja koja se protežu po glavnoj dijagonali imaju cenu 0 jer se radi o matrici samosličnosti i porede se isti segmenti. Posmatranjem Sl. 4 se vidi da su prvi i drugi segment slični sa šestim segmentom. Takođe se vidi da su treći, četvrti, sedmi i osmi segment međusobno slični.



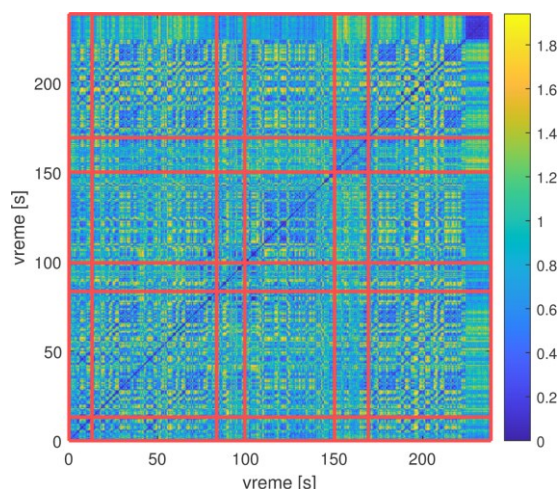
Sl. 5. Matrica cena međusobnih odnosa segmenata, granice izračunate na matrici MFCC obeležja, a cene računute na matrici hroma obeležja.

Kroz ovaj primer se mogu posmatrati i rezultati kada se granice segmenata dobijene iz matrice obeležja MFCC koeficijentata primene na matricu samosličnosti hroma obeležja za računanje normalizovane cena polja. Ovi rezultati su prikazani na Sl. 5 Granica između cena segmenata koji su slični i cena onih segmenata koji se razlikuju je osetljivija, nije toliko jasna razlika u bojama kao na Sl. 4. Ipak se vidi da su polja koja su prikazana plavom bojom na Sl. 4 prikazana

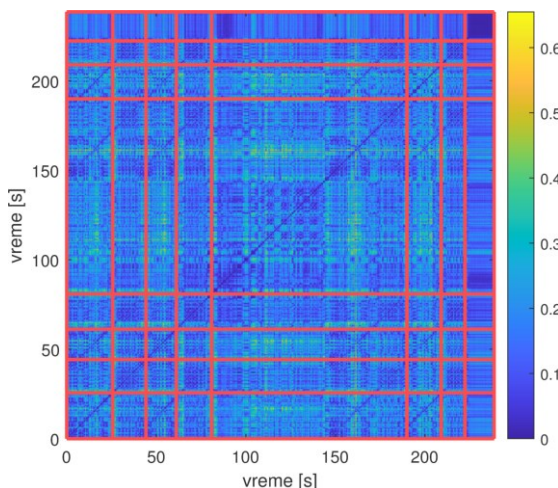
plavom bojom i na Sl. 5, odnosno za obe vrste obeležja se izdvajaju isti segmenti kao slični. Polje koje predstavlja sličnost prvog i drugog segmenta na Sl. 4 nije jasno definisano kao polje koje predstavlja slične segmente, dok je na Sl. 5 jasno izdvojeno kao polje koje ukazuje na sličnost ova dva segmenta.

#### B. Primer 2. Valcer br. 2 – Dmitrij Šostakovič

Kao i u prethodnom primeru i za ovu muzičku sekvencu su određena hroma i MFCC obeležja i formirane su matrice samosličnosti primenom DTW algoritma za obe vrste obeležja. Primenom algoritma za segmentaciju na ove matrice samosličnosti se ne dobijaju iste vrednosti granica segmenata. Što ukazuje da vrsta obeležja kojim se opisuje muzičko delo ima uticaj na podelu na segmente. Na Sl. 6 i Sl. 7 su prikazane matrice samosličnosti sa određenim granicama segmenata za MFCC obeležja i hroma obeležja.



Sl. 6. Matrica samosličnosti za MFCC obeležja, sa definisanim granicama.

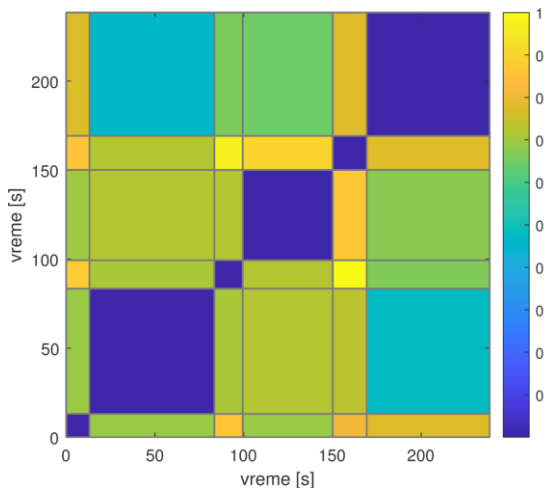


Sl. 7. Matrica samosličnosti za hroma obeležja, sa definisanim granicama.

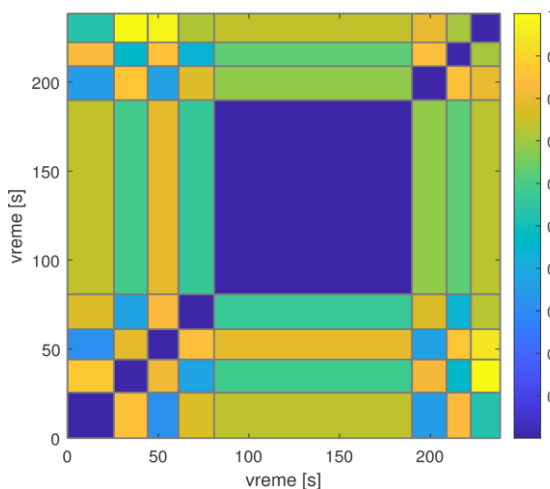
Segmenti koji su odvojeni analizom hroma obeležja se ne izdvajaju i nisu jasno definisani na matrici samosličnosti kada se analiziraju MFCC obeležja. Deo oko 200 s koji se izdvaja na matrici hroma obeležja se delimično primećuje na matrici MFCC obeležja. Takođe ovaj slučaj važi i za deo oko 50 s. Dok deo oko 100 s i 150 s koji se izdvaja na MFCC matrici se ne izdvaja na hroma matrici, a uočava se i na hroma matrici kao segment koji bi trebalo odvojiti. Ovakvom analizom je

zaključeno da bi se kombinovanjem granica dobijenih iz matrica obe vrste obeležja dobila najbolja segmentacija za ovaj muzički signal.

Na Sl. 8 i Sl. 9 su prikazane cene međusobnih odnosa segmenata muzičkog dela za MFCC i hroma obeležja na osnovu prethodno definisanih granica. Manji delovi oko 200 s koji su prikazani plavom bojom na Sl. 9 su sadržani u velikom segmentu MFCC matrice. Paralelnim posmatranjem ove dve matrice cena možemo zaključiti da je segment od oko 20 s do oko 80 s sličan sa krajnjim delom ovog signala od oko 180 s do kraja.

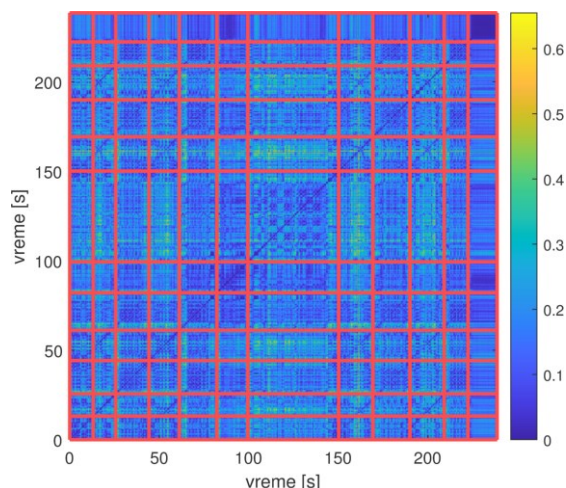


Sl. 8. Matrica cena međusobnih odnosa segmenata, MFCC obeležja.

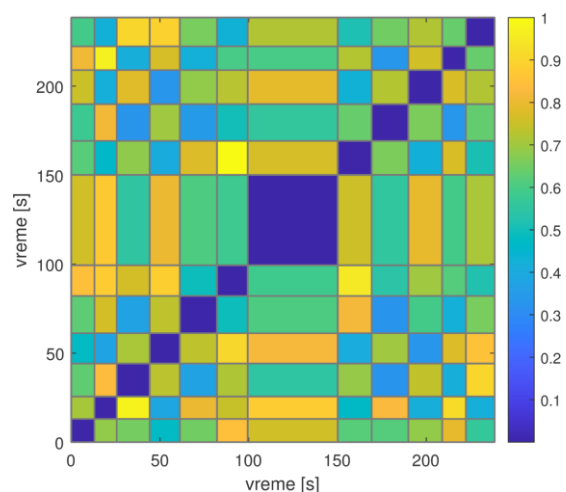


Sl. 9. Matrica cena međusobnih odnosa segmenata, hroma obeležja.

Poređenjem granica koje su prikazane na Sl. 7 i granica prikazanim na Sl. 10 se vidi da su ovim kombinovanjem dodate granice za koje se može reći da su nedostajale na Sl. 7. Zaključuje se da algoritam za segmentaciju ne može dovoljno dobro da izdvoji segmente na osnovu samo jedne vrste obeležja. Kombinovanjem granica za obe vrste obeležja se izdvajaju segmenti koji se posmatranjem matrice obeležja mogu definisati kao zasebni segmenti.



Sl. 10. Matrica samosličnosti za hroma obeležja, sa kombinovanim granicama.



Sl. 11. Matrica cena međusobnih odnosa segmenata, hroma obeležja za kombinovane granice.

Na Sl. 11 su prikazane cene međusobnih odnosa segmenata za novodefinisane granice. Vidi se da su delovi koji su bili plavom bojom definisani kao slični na Sl. 9 takođe definisani kao slični i na Sl. 11. Uz to se odvajaju novi manji segmenti kao međusobno slični koji nisu bili tako definisani samo sa hroma granicama. Ako se posmatra deo između 180 s i 220 s na Sl. 9 i Sl. 11, vidi se da na Sl. 8 u tom delu nema segmenata koji su označeni plavom bojom, dok na Sl. 11 ima sličnih delova označenih plavom bojom. Takođe ovakva situacija se javlja i sa segmentima između 150 s i 200 s kada se porede sa segmentima između 20 s i 80 s. Kombinovanjem granica se dobijaju manji segmenti i preciznije se mogu odrediti odnosi između njih. Kada su segmenti manji pronalaze se dodatne međusobne sličnosti nekih segmenata.

#### IV. ZAKLJUČAK

U radu je predložena efikasna metoda za segmentaciju muzičkog dela, izdvajanjem i analizom MFCC i hroma obeležja, i njihovom obradom korišćenjem metoda za obradu slike, i DTW algoritmom. Ovakva segmentacija može poslužiti kao pomoćni alat pri statističkoj analizi muzičkih

dela, bez potrebe da se delo preslušava, i ručno deli na segmente. Takođe, za korišćenje ovog algoritma i primenu na nekim muzičkim delima, nije neophodno nikakvo predznanje o samom delu, žanru kom pripada, ili bilo kakvoj teoriji muzike koja stoji iza njega.

Kroz ove primere možemo zapaziti sličnosti i razlike u rezultatima u zavisnosti od obeležja kojim se opisuje muzička sekvenca. Takođe se uviđa i veza između dve vrste obeležja. U prvom primeru pokazano je da se isti segmenti definišu kao slični bez obzira na vrstu obeležja. Tačnije, primenom granica određenih iz matrice samosličnosti MFCC obeležja na matricu hroma obeležja, i obrnuto, se dobijaju isti rezultati, odnosno isti segmenti se izdvajaju kao slični. Primenom metode na drugom primeru se uviđa da se dobijaju drugačiji rezultati u zavisnosti od vrste obeležja – granice koje su određene iz matrice samosličnosti za MFCC i za hroma obeležja nisu iste. Kroz ovaj primer je pokazana mogućnost kombinovanja granica za jednu i drugu vrstu obeležja. Zaključeno je da se ovakvim kombinovanjem, kada se rezultati segmentacije razlikuju za različita obeležja, može izvršiti preciznija segmentacija – definišu se manji segmenti i javljaju se nove informacije o međusobnoj sličnosti novodefinisanih segmenata.

U nastavku istraživanja predložena metoda za segmentaciju muzičkog dela se može primeniti na matricu sličnosti dve različite interpretacije, ili izvedbe istog muzičkog dela. Na ovaj način bi se omogućilo automatsko vremensko usklađivanje dela koja su interpretirana na drugačiji način. Takođe, jedan od smerova za nastavak istraživanja može biti analiziranje sličnosti određenih delova dva različita muzička signala, sa potencijalnom primenom u analiziranju i otkrivanju muzičkih plagijata. Korišćenjem predloženog algoritma bi se mogli automatski uočiti obrasci koji se pojavljuju u određenim žanrovima

#### ZAHVALNICA

Ovaj rad je finansijski podržan od strane Ministarstva nauke, tehnološkog razvoja i inovacija Republike Srbije, pod brojem ugovora 451-03-47/2023-01/200103.

#### LITERATURA

- [1] J. Paulus, M. Müller, A. Klapuri, "State of the Art Report: Audio-Based Music Structure Analysis", *Ismir*, pp. 625-636, Aug, 2010.
- [2] A. Hennion, "The production of success: an anti-musicology of the pop song", *Popular Music*, 3, pp. 159-193, Jan, 1983.
- [3] J. Adrian, "The Ternary-Sonata Form", *Journal of Music Theory*, pp. 57-80, 1990.
- [4] C. A. Williams, "The Rondo Form, as it is found in the Works of Mozart and Beethoven", *Proceedings of the Musical Association*, 17, pp. 95-112, 1890.
- [5] J. Foote, "Visualizing music and audio using self-similarity", *Proceedings of the seventh ACM international conference of Multimedia*, 1, pp. 77-80, Oct, 1999.
- [6] W. Y. Chan, H. Qu, W.H. Mak, "Visualizing the semantic structure in classical music works", *IEEE transactions on visualization and computer graphics*, 16(1), pp. 161-173, Jun, 2009.

- [7] H. T. Cheng, Y. H. Yang, Y. C. Lin, H. H. Chen, "Multimodal structure segmentation and analysis of music using audio and textual information", *2009 IEEE International Symposium on Circuits and Systems*, pp. 1677-1680, May, 2009.
- [8] M. Marolt, "A Mid-level Melody-based Representation for Calculating Audio Similarity" *ISMIR*, pp. 280-285, Oct, 2006.
- [9] J. C. Wang, Y. N. Hung, J. B. Smith, "To catch a chorus, verse, intro, or anything else: Analyzing a song with structural functions", *ICASSP 2022-2022*, pp. 416-420, May, 2022.
- [10] J. Foote, "Automatic audio segmentation using a measure of audio novelty", *ICME 2000*, 1, pp. 452-455, Jul, 2000.
- [11] K. Lee, M. Cremer, "Segmentation-Based Lyrics-Audio Alignment using Dynamic Programming", *ISMIR*, pp. 395-400, 2008.
- [12] J. T. Foote, M. L. Cooper, "Media segmentation using self-similarity decomposition", *Storage and Retrieval for Media Databases 2003*, 5021, pp. 167-175. Jan, 2003.
- [13] M. Müller, F. Kurth, "Enhancing similarity matrices for music audio analysis", *2006 IEEE International Conference on Acoustics, Speech, and Signal Processing Proceedings*, 5, pp. V-V, May, 2006.
- [14] S. Jun, E. Hwang, "Music segmentation and summarization based on self-similarity matrix", *Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication*, pp. 1-4, Jan, 2013.
- [15] M. Levy, M. Sandler, "Structural segmentation of musical audio by constrained clustering", *IEEE transactions on audio, speech, and language processing*, 16(2), pp. 318-326, Jan, 2008.
- [16] *The MathWorks Inc.* (2022). MATLAB version: 9.13.9 (R2022b), Natick, Massachusetts.
- [17] S. Jun, S. Rho, E. Hwang, "Music structure analysis using self-similarity matrix and two-stage categorization", *Multimedia Tools and Applications*, 74, pp. 287-302, Jan, 2015.
- [18] M. Mijić, "Čulo sluha kao audio prijemnik", in *Audio sistemi*, ch. 2, sec. 2, pp. 56-66, Beograd, Srbija, Akademski misao, 2011.
- [19] R. C. Gonzalez, R. E. Woods, "Morphological Image Processing", in *Digital Image Processing*, ch. 9, sec. 3, pp. 635-639, New Jersey, USA: Pearson Prentice hall, 2008.

#### ABSTRACT

This paper deals with the problem of automated segmentation and structure analysis of musical pieces on the basis of extracted features from audio recordings of the piece. The complexity of musical piece segmentation problems is a consequence of the multiple elements by which the segments differ, and those can be melody, harmony, rhythm, and timbre. Self-similarity matrices (SSM), based on selected features are used as a cornerstone for segmentation. The features used are MFCC (Mel-Frequency Cepstral Coefficients) and chroma features which follow changes in timbre and harmonic structure. A self-similarity matrix segmentation procedure based on straightforward methods for morphological image processing is proposed. The Dynamic Time Warping algorithm (DTW) was used as a tool for evaluating the degree of similarity of the segments. The proposed procedure was applied to 2 pieces of classical music.

#### Musical signal segmentation method based on self-similarity matrices

Marija Ratković, Matija Marijan, Tatjana Miljković, Miloš Bjelić