

# Nizkodimenzionalni model tonskega prostora

Bojan Klemenc<sup>1</sup>, Franc Solina<sup>1</sup>

<sup>1</sup>Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Večna pot 113, Ljubljana

bojan.klemenc@fri.uni-lj.si, franc.solina@fri.uni-lj.si

## Izvleček

Podrobnosti človeškega zaznavanja tonov oziroma glasbe kot celote, so še vedno odprto raziskovalno področje. Osredotočamo se na nizkodimenzionalne geometrijske modele, ki pomagajo človeku razumeti slušni prostor. Večini obstoječih modelov manjka intuitivna preslikava, ki ohranja zaznavne razdalje med tonskim prostorom, prostorom akordov in tonalitet. Zato predlagamo nov dopolnjen model preslikovanja tonov in tonalitet v vizualni prostor, ki v primerjavi z obstoječimi modeli boljše modelira zaznavne razdalje. Predlagani model je kvocientni prostor, ki smo ga izpeljali s pomočjo zmanjševanja števila dimenzij izvornih podatkov psihoakustičnih poskusov. Model ohranja zajeto varianco najboljših primerljivih modelov, pri čemer omogoča lažjo interpretacijo od obstoječih modelov.

**Ključne besede:** zaznavanje glasbe, tonski prostor, zmanjševanje števila dimenzij, vizualizacija

## Abstract

How human brain perceives and processes tones or music as whole is still an active area of research. We focus on low dimensional models that aid the human understanding of auditory space. Most existing models lack intuitive perceptual distance preserving mapping that would simultaneously represent relationships between tones, chords and tonalities. We propose an augmented model and mapping of tone and tonal region space to a visual space that uses improved perceptually uniform mapping from auditory space to visual space compared to existing solutions. The proposed quotient space model is based on existing models derived with dimensionality reduction methods from psychoacoustic data. The model preserves the captured variance of best comparable existing models while the dimensions of the model are easier to interpret compared to existing models.

**Keywords:** Music perception, pitch space, dimensionality reduction, visualization

## 1. UVOD

Podrobnosti človeškega zaznavanja tonov oziroma glasbe kot celote, so še vedno odprto raziskovalno področje. Raziskovalci na področju psihoakustike so v zadnjih desetletjih zgradili več zaznavnih modelov (Krumhansl and Cuddy, 2010), da bi razumeli značilnosti človeškega slušnega zaznavnega prostora, kot na primer zaznavanje harmonskih odnosov med toni, akordi in tonalitetami – kaj zaznamo kot konsonantno ali disonantno v različnih kontekstih. Te modele lahko uporabimo za analizo glasbe, pri učenju in ustvarjanju nove glasbe. Razdelimo jih v dve skupini: če so namenjeni človeku, modeli ne smejo biti preveč zapleteni, da olajšajo razumevanje; če pa so namenjeni računalniku, so lahko bolj zapleteni. Osredotočimo se na modele, ki pomagajo

človeku razumevati zaznavni tonski prostor. Človeku se lahko uči konceptov (npr. igranja glasbil) med drugim s ponotranjanjem množice pravil. Pri tem si lahko tudi pomagamo z vizualnimi modeli. Zato se osredotočamo na podskupino zaznavnih modelov tonskega prostora, ki jih je preprosto vizualizirati – geometrijske modele. Primerjali smo obstoječe modele in ugotovili, da mnogi modeli ne uporabljajo preslikav, ki bi v vizualnem modelu ohranjale slušno zaznavno razdaljo med toni, akordi in tonalitetami. Takšni neintuitivni modeli so vseeno uporabni, vendar imajo bolj strmo krivuljo učenja.

Predlagamo nov, dopolnjen model preslikave tonov in tonalitet v vizualni prostor, ki bolje ohranja zaznavne razdalje pri preslikavi iz slušnega v vizualni prostor. Pri izdelavi modela je bilo glavno vo-

dilo zgraditi model, ki lahko pomaga človeškemu razumevanju tonskega prostora. Model je zgrajen na podlagi obstoječih geometrijskih modelov, ki so izpeljani s pomočjo zmanjševanja števila dimenzij: večdimenzionalnega skaliranja (angl. multidimensional scaling – MDS) in razvijanja z maksimalno varianco (angl. maximum variance unfolding – MVU). Topološko imajo ti modeli obliko 2-torusa, kjer so tonalitete na površini torusa. Takšni nizkodimenzionalni modeli ponujajo dodatno vizualno orodje, s katerim lahko uporabniku olajšamo razumevanje, v kakšnih medsebojnih razmerjih iz harmonskega vidika so posamezni toni, akordi in tonalitete. Predpogoj za smotrnost uporabe takšnega modela pa je, da model zajame čim večjo varianco izvirnega večdimenzionalnega zaznavnega prostora.

Čeprav se topologija 2-torusa razlikuje od topologije barvnega prostora, lahko izbrane dimenzije 2-torusa preslikamo v barvni prostor. Zaznavna bližina v tonskem prostoru se tako preslika v zaznavno bližino v vizualnem oziroma barvnem prostoru. Posledično lahko preslikavo predlaganega modela v barvni prostor uporabimo za nadgradnjo obstoječih orodij za vizualizacijo glasbe ter za nadgradnjo glasbil, kjer si lahko uporabniki s pomočjo barv pomagajo pri orientaciji po predlaganem modelu. Predlagani model lahko tako pri učenju nadomesti preprostejše modele, kot na primer kvintni krog, ter se zaradi ohranjanja zaznavnih razdalj uporabi v obstoječih algoritmih za segmentacijo in prepoznavanje tonov, akordov in tonalitete.

## 2 OZADJE IN SORODNI MODELI

Zaznavni modeli so grajeni na osnovi človeške zaznave zvoka, zato za vhodne podatke ne uporabljajo fizičnega signala, ampak glasbene tone, ki že predstavljajo rezultat človeške zaznave zvoka. Na podlagi medsebojnih razmerij med toni lahko zgradimo model tonskega prostora. Eden od možnih pristopov je, da tone začnemo razporejati po prostoru glede na njihova medsebojna razmerja, kakor jih opisuje glasbena teorija (Tymoczko, 2012; Chew, 2000; Harter, 2010). Mnogi takšni modeli imajo probleme pri hkratnem modeliranju zaznavnih razdalj med toni, akordi in tonalitetami.

Drug možen pristop je, da za izhodišče vzamemo rezultate psihoakustičnih poskusov – zaznavne razdalje med toni, akordi in tonalitetami – ter jih postavimo v večdimenzionalen metrični prostor.

Psihoakustični poskusi Krumhanslove (Krumhansl and Cuddy, 2010; Krumhansl, 1990) so zajemali večje skupine ljudi, ki so ocenjevali, kako dobro nek ton sledi zaporedju tonov, akordov ali kakšnega drugega tonskega konteksta. Na podlagi teh poskusov so sestavili tonalitete profile. Tonalitetni profil za posamezno tonaliteto je vektor zaznavnih razdalj med tonaliteto in toni enakomerno temperirane 12-tonske lestvice. Tako dobimo večdimenzionalen metrični prostor, kjer evklidska razdalja med vektorji predstavlja razdaljo med tonalitetami. Nato z metodami za zmanjševanje števila dimenzij modele poenostavimo. Za ciljno število dimenzij 4 in z uporabo MDS, kot predlaga Krumhanslova (Krumhansl and Cuddy, 2010), dobimo 2-torus ( $T_2$ ) vloženo v  $R^4$ , kjer tonalitete ležijo na površini torusa.

Burgoyne in Saul (2005) sta predlagala uporabo MVU za zmanjšanje števila dimenzij. Podobno kot pri MDS, z uporabo MVU dobimo 2-torus, vendar z drugačno razporeditvijo tonalitete. Model MDS lahko dopolnimo tako, da vsebuje tudi tonske razrede (slika 1a) – tak model se ujema z nekaterimi starejšimi predlaganimi modeli (Purwins et al., 2007); medtem ko Burgoyne in Saul (2005) pri modelu MVU nista definirala mreže tonskih razredov.

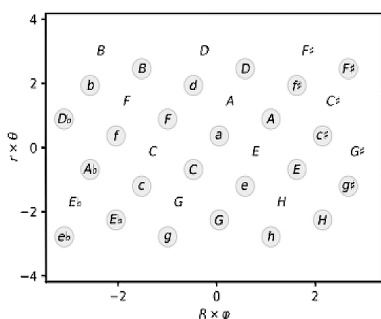
## 3 PREDLAGANI MODEL TONSKEGA PROSTORA IN REZULTATI

Predlagani model gradimo na obstoječih modelih, ki so izpeljani z zmanjševanjem števila dimenzij višje-dimenzionalnega metričnega prostora, kjer evklidska razdalja predstavlja zaznavno razdaljo. S ciljnim številom dimenzij 4 in uporabo MDS ali MVU dobimo 2-torus (kartezični produkt dveh krožnic  $T_2 = S_1 \times S_1$ ). Uporaba štirih dimenzij je kompromis med poenostavljanjem modela in ohranjanjem čim večje količine informacij.

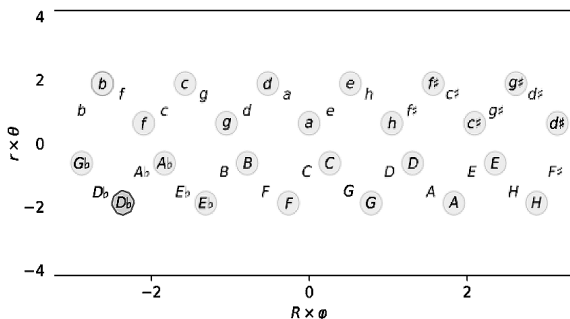
Modele smo evalvirali tako, da smo primerjali, kako dobro ohranjajo zaznavne razdalje proti izvornemu večdimenzionalnemu modelu. Za vsako tonaliteto smo v izvornem modelu izračunali tonalitetni profil – vektor razdalj tonalitete do ostalih tonalitete (T-T); podobno izračunamo razdalje posameznih tonov do tonalitete (t-T). Zanima nas, kako dobro model ohranja relativne razdalje tonalitetnih profilov, zato izračunamo Pearsonov korelacijski koeficient med vektorji razdalj izvornih tonalitetnih profilov in vektorji nizkodimenzionalnih modelov na podlagi MDS in MVU (model 1 in model 3 v tabeli 1). Podobno

izračunamo korelacijski koeficient za tonsko-tonalitetne profile. Pri obeh nizkodimenzionalnih modelih prvi dve dimenziji ( $R \times \varphi$  na slikah 1a in 1b) predstavljata dva prekrivajoča se kvintna kroga. Pri modelu MDS preostali dve dimenziji ( $r \times \theta$ ) ni tako preprosto interpretirati, medtem ko pri modelu MVU dobimo kontrast med duri in moli, kar je prednost modela

MVU. Po drugi strani pa je korelacija modela MDS z izvornim modelom boljša od modela MVU. Poleg tega je relativno preprosto dodati tonske razrede v model MDS (predlagana različica je prikazana na sliki 1a, pripadajoča korelacija z izvornimi profili pa je podana v tabeli 1 (model 2).



(a) Obstoječi model razporeditve tonalitet na podlagi večdimenzionalnega skaliranja (MDS) (Krumhansl, 1990) z dodanimi toni.



(b) Predlagani model na podlagi razvijanja z maksimalno varianco (MVU) je kvocientni prostor, kjer toni in tonalitate predstavljajo ekvivalenčne razrede. Vsak ton se v modelu vizualno pojavi dvakrat (enkrat v kontekstu dura – velike črke, enkrat v kontekstu mola – male črke), vendar obe pojavitvi predstavljata isti ekvivalenčni razred.

Slika 1: Primerjava razvitih površin 2-torusov obeh modelov tonalitetnega prostora – na podlagi MDS in na podlagi MVU z dodanimi ekvivalenčnimi razredi. Za lažje razločevanje so na sliki tonalitate obkrožene, toni pa so neobkroženi. Enharmonični ekvivalenti so zaradi boljše preglednosti označeni samo z enim od imen. Evklidske razdalje med toni in tonalitetami predstavljajo zaznavne razdalje.

Tabela 1: Primerjava tonsko-tonalitetnih modelov zgrajenih na podlagi psihoakustičnih podatkov. Primerjamo korelacijo vektorjev tonalitetnih profilov (T-T) in tonsko-tonalitetnih profilov (t-T) posameznih modelov z vektorji referenčnega modela. Model 2 smo dopolnili, da ga lahko primerjamo s predlaganimi modeloma 4 in 5.

modeli	korelacijski koeficienti	
	T-T	t-T
1. MDS → 2-torus (Krumhansl, 1990)	0.88	
2. MDS → 2-torus z dodatnimi toni (dopolnjeni model)	0.88	0.89
3. MVU → 2-torus (Burgoyne and Saul, 2005)	0.64	
4. MVU → 2-torus → kvocientni prostor nad $R^2$ (predlagani model)	0.80	0.79
5. MVU → 2-torus → kvocientni prostor nad $R^3$ (predlagani model)	0.87	0.90

Predlagamo nov model, osnovan na modelu MVU, ki združuje prednosti obeh obstoječih modelov. Če poskusimo dodati tone v model MVU, kakor smo dodali v MDS, je korelacija vektorja razdalj z vektorjem izvornih tonalitetnih profilov precej slabša. V predlagani model zato vsak ton dodamo dvakrat – enkrat za durov in enkrat za molov kontekst. Ker so vsi toni in tonalitate na površini torusa, lahko torus razvijemo v  $R^2$  (slika 1b) in definiramo ekvivalenčno relacijo, ki enači vse primerke iste tonalitete  $T$  v ekvivalenčni razred  $[T]$ .

Podobno definiramo ekvivalenčno relacijo za tone, ki vse primerke istega tona in njegovih enharmoničnih ekvivalentov enači v ekvivalenčni razred  $[t]$ . Dobimo kvocientni prostor  $R^2 / \sim_{[T,t]}$ .

Prednost tako opredeljenega modela je, da ga še vedno lahko vizualiziramo v  $R^2$  (slika 1b). Korelacija vektorja razdalj predlaganega modela (model 4 v tabeli 1) z vektorjem razdalj izvornih tonalitetnih profilov (stolpec T-T v tabeli 1) je višja od obstoječega modela MVU, prav tako je tudi korelacija z izvornimi tonsko-tonalitetnimi profili (stolpec t-T) visoka, ven-

dar manjša od modela MDS. Problem rešimo tako, da namesto nad  $R^2$  definiramo kvocientni prostor nad  $R^3$  (model 5) in tako dosežemo približno enako korelacijo z izvornimi profili kot model MDS, hkrati pa je dimenzije predlaganega modela lažje interpretirati kot dimenzije modela na podlagi MDS.

#### 4 ZAKLJUČEK

Predlagali smo nov model, ki je osnovan na nizkodimenzionalni vložitvi tonskega prostora na podlagi zmanjševanja števila dimenzij z metodo MVU. Obstoječi model na podlagi MVU vsebuje le tonalitete, predlagani model pa spremeni obstoječi model v kvocientni prostor, da lahko dodamo tonske razrede ter hkrati ohranjamo zaznavne razdalje glede na izvorne večdimenzionalne tonalitete profile. Korelacija predlaganega modela je primerljiva z najboljšim podobnim obstoječim modelom MDS, hkrati pa ima bolj enostavne interpretacije dimenzij iz vidika glasbene teorije in je posledično bolj primeren za človeško razumevanje tonskega prostora. V nadaljevanju bomo za testiranje predlaganega modela preslikali model v barvni prostor ter prenesli v obstoječa orodja za analizo glasbe in aplicirali na glasbila – npr. klavir s tipkami osvetljenimi z LED-lučkami, ki uporabniku

pomagajo najti lokacije tonov, akordov in tonalitete v računalniški vizualizaciji modela. Cilj predlaganega modela je pomagati človeškemu razumevanju slušnega tonskega prostora, kar bi lahko omogočilo lažje razumevanje glasbene teorije, hitrejše učenje igranja glasbil in ustvarjanja glasbe.

#### LITERATURA

- [1] Burgoyne, J. A. and Saul, L. K. (2005). Visualization of Low Dimensional Structure in Tonal Pitch Space. In *International Computer Music Conference Proceedings*, pages 243–246, Barcelona.
- [2] Chew, E. (2000). *Towards a Mathematical Model for Tonality*. Ph.D. dissertation, Operations Research Center, MIT, Cambridge, MA.
- [3] Harte, C. (2010). *Towards automatic extraction of harmony information from music signals*. PhD thesis, Department of Electronic Engineering, Queen Mary, University of London.
- [4] Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. Number no. 17 in Oxford psychology series. Oxford University Press, New York.
- [5] Krumhansl, C. L. and Cuddy, L. L. (2010). A Theory of Tonal Hierarchies in Music. In Riess Jones, M., Fay, R. R., and Popper, A. N., editors, *Music Perception*, volume 36, pages 51–87. Springer New York, New York, NY.
- [6] Purwins, H., Blankertz, B., and Obermayer, K. (2007). Toroidal models in tonal theory and pitch-class analysis. *Computing in Musicology (Tonal Theory for the Digital Age)*, 15:73–98.
- [7] Tymoczko, D. (2012). The Generalized Tonnetz. *Journal of Music Theory*, 56(1):1–52.

■  
**Bojan Klemenc** je asistent na Fakulteti za računalništvo in informatiko ter član Centra za jezikovne vire in tehnologije Univerze v Ljubljani. Magistriral je leta 2016 na Fakulteti za računalništvo in informatiko. Raziskovalno se ukvarja z modeliranjem zaznavanja glasbe, vizualizacijami in korpusno leksikografijo. Je soavtor polavtomatiziranega odzivnega digitalnega Slovarja sopomenk in Kolokacijskega slovarja sodobne slovenščine.

■  
**Franc Solina** je redni profesor računalništva in informatike na Fakulteti za računalništvo in informatiko Univerze v Ljubljani ter ustanovitelj Laboratorija za računalniški vid (1991). Diplomiral (1979) in magistriral (1982) je na Fakulteti za elektrotehniko Univerze v Ljubljani, doktoriral pa iz računalništva (1987) na University of Pennsylvania v ZDA. Raziskovalno se ukvarja s 3D modeliranjem na osnovi slik in z uporabo računalniškega vida na področju uporabniških vmesnikov ter umetniških instalacij. Je član IEEE, IAPR, ICOMOS, Zveze društev slovenskih likovnih umetnikov ter Evropske akademije znanosti in umetnosti.