

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Bojan Klemenc

**Vizualni model harmoničnih odnosov
med simultanimi toni v glasbi na
podlagi psihoakustičnega modela
percepcije zvoka**

MAGISTRSKO DELO

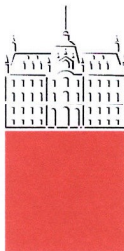
MAGISTRSKI PROGRAM
INFORMACIJSKI SISTEMI IN ODLOČANJE

MENTOR: prof. dr. Franc Solina

Ljubljana, 2016

Rezultati magistrskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov magistrskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil L^AT_EX.



Številka: 163-MAG-ISO/2016
Datum: 06. 04. 2016

Bojan KLEMENC, univ. dipl. inž. rač. in inf.

L j u b l j a n a

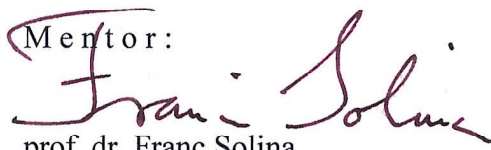
Fakulteta za računalništvo in informatiko Univerze v Ljubljani izdaja naslednjo magistrsko nalogo

Naslov naloge: **Vizualni model harmoničnih odnosov med simultanimi toni v glasbi na podlagi psihoakustičnega modela percepcije zvoka**


Visual model of harmonic relationships of simultaneous tones in music based on psychoacoustic model of sound perception

Tematika naloge:

Obstoječi modeli harmonične strukture glasbe opisujejo sorodnost med toni (lahko tudi akordi, tonalitetami itd.) na različne načine. Pogosto modelirajo sorodnost s pomočjo bližine (npr. evklidske razdalje) v prostorskem geometrijskem modelu. Če želimo zajeti vso kompleksnost glasbe, imajo takšni modeli veliko dimenzij, ki jih je težko vizualizirati in posledično zaradi svoje kompleksnosti niso v širši uporabi. Proučite obstoječe modele harmonične strukture glasbe, ki so grajeni na podlagi spoznanj psihoakustike, teorije glasbe, kognitivnih modelov in nevroznanosti. Zasnujte nov model harmoničnih odnosov glasbi z zmanjšanim številom dimenzij, ki bo temeljil na percepciji tonov in bo omogočal lažje razumevanje harmonične strukture glasbe oz. sočasno zvonečih tonov ter služil kot pomoč pri učenju glasbe. Model naj omogoča uporabo modela pri računalniški primerjavi, klasifikaciji ter iskanju in indeksiranju multimedijskih vsebin. Pri gradnji računalniškega modela upoštevajte področje psihoakustike in zaznavanja, kognitivne modele, spoznanja nevroznanosti, teorijo glasbe, obstoječe predlagane modele predstavitve odnosov med toni ter področja znanstvene in estetske vizualizacije podatkov. Model ovrednotite in primerjajte z obstoječimi modeli. Predlagajte konkretne primere uporabe vašega modela kot pomoč pri učenju glasbe.

Mentor:

prof. dr. Franc Solina



Dekan:

prof. dr. Nikolaj Zimic

Zahvaljujem se mentorju prof. dr. Francu Solini za pomoč v času po-diplomskega študija in za potrpljenje. Posebej se zahvaljujem Petru Ciuhi za začetno pobudo, za svetovanje in pomoč pri testiranju modela in izdelavi prototipa barvne klaviature. Doc. dr. Marku Jankovcu in Jerneju Vrščaju se zahvaljujem za pomoč pri izdelavi prototipa barvne digitalne klaviature. Pri izdelovanju modela in pisanju magistrske naloge sem uporabil mnogo prosto dostopnih odprtokodnih programskih paketov, zato se zahvaljujem vsem programerjem, ki so sodelovali pri njihovem nastajanju. Zahvaljujem se tudi vsem ostalim, ki so kakorkoli pripomogli pri nastajanju te magistrske naloge.

Kazalo

Povzetek

Abstract

1	Uvod	1
1.1	Motivacija	1
1.2	Cilji in metodologija	3
2	Pregled področja	5
2.1	Zvok in percepcija zvoka	5
2.2	Toni in percepcija tonov	10
2.3	Percepcija dveh in več sočasnih tonov	13
2.4	Predstavitev odnosov med toni	15
2.5	Vizualizacije glasbe	17
2.6	Vid in zaznavanje barv	19
3	Model	23
3.1	Barvni prostor	23
3.2	Tonski prostor in preslikava	36
3.3	Rezultati in razprava	49
4	Implementacija	57
4.1	Prototip fizične barvne klaviature	59
5	Zaključki in nadaljnje delo	65

Dodatki	67
A Ostale uporabljene formule	69
A.1 Izračun ΔE_{00}^*	69
A.2 Pretvorbe med CIE L*C*h° in CIE L*a*b*	70
A.3 Pretvorbe med CIE L*a*b* in CIE XYZ	71
A.4 Pretvorbe med CIE XYZ in sRGB	72
A.5 Pretvorbe med CIE XYZ in Adobe RGB	73
Literatura	75

Seznam uporabljenih kratic

kratica	pomen
CIE	fr. Commission internationale de l'éclairage (Mednarodna komisija za razsvetljavo)
EEG	elektroencefalografija
fMRI	angl. functional magnetic resonance imaging (funkcijsko magnetnoresonančno slikanje)
HSL	angl. hue, saturation, lightness (odtenek, nasičenost, svetlost)
HSV	angl. hue, saturation, value (odtenek, nasičenost, vrednost)
JND	angl. just noticeable difference (komaj zaznavna razlika)
MDS	angl. multidimensional scaling (večrazsežnostno lestvičenje)
MVU	angl. maximum variance unfolding (razvijanje z maksimalno varianco)
RMSE	angl. root mean square error (kvadratni koren povprečja kvadratov napake)
RGB	angl. red, green, blue (rdeča, zelena, modra)

Povzetek

Naslov: Vizualni model harmoničnih odnosov med simultanimi toni v glasbi na podlagi psihoakustičnega modela percepcije zvoka

Avtor: Bojan Klemenc

Marsikdo se je v svojih mladih letih usedel za klavir ali poprijel za kakšno drugo glasbilo in poskušal igrati. Vendar je po prvih nekaj neuspešnih poskusih odnehal. Naključno pritiskanje tipk z zelo majhno verjetnostjo postane ušesu prijazna glasba. Preproste melodije še gredo, ampak katere akorde lahko zaigramo ob melodiji, kdaj in zakaj, kaj je tonaliteta, kaj je na primer modulacija? To so vprašanja, na katera večina učencev dobi odgovor postopoma v okviru izobraževanja v glasbenih šolah. Kako približati teorijo glasbe tudi tistim, ki niso obiskovali glasbenih šol? Ali je možno glasbo predstaviti v vizualni obliki, tako da bi bila bolj razumljiva? Ali lahko, podobno kot za rdečo in oranžno barvo lahko hitro rečemo, da sta si podobni, predstavimo dva akorda vizualno tako, da lahko rečemo, da sta si podobna?

V magistrski nalogi bom predstavil nov model vizualne predstavitve harmoničnih odnosov v glasbi, ki temelji na percepciji tonov, omogoča lažjo analizo in razumevanje harmonične strukture glasbe ter lahko služi kot pomoč pri učenju glasbe. V okviru naloge sta bila razvita prototipa barvne klaviature in računalniške aplikacije za demonstracijo računskega modela.

Ključne besede: vizualizacija, glasba, percepcija tonov, psihoakustika, modeliranje zaznavanja, barvni prostori.

Abstract

Title: Visual model of harmonic relationships of simultaneous tones in music based on psychoacoustic model of sound perception

Author: Bojan Klemenc

It happens quite often that someone sits down in front of a piano or takes another instrument and tries to play for the first time, but eventually quits trying after first few unsuccessful attempts. Random presses of the piano keys can become pleasant music only with a very small probability. Easy melodies are manageable, but what chords can accompany the melody and when to play them, what is tonality, what is for example modulation? These are concepts that most students are taught during their education in music schools. How can we bring music theory closer to those, who have not attended a music school or have not received musical education? Can music be visualized to make it easier to understand? Can we, like we can see that red and orange are similar, visualize two chords in such a way that we can say that they are similar?

In this thesis I will present a novel model of visual representation of harmony in music. The model is based on tone perception, aimed to enable easier analysis of music, easier understanding of harmonic structure of music and can consequently complement traditional musical education. To demonstrate and test the model we developed a prototype application and integrated the model into a digital piano keyboard.

Keywords: visualization, music, tone perception, psychoacoustics, perception modelling, colour spaces.

Poglavje 1

Uvod

1.1 Motivacija

Marsikdo se je v svojih mladih letih usedel za klavir ali poprijel za kakšno drugo glasbilo in poskušal igrati. Vendar je po prvih nekaj neuspehlih poskusih odnehal. Naključno pritiskanje tipk z zelo majhno verjetnostjo zveni postane ušesu prijazna glasba. Preproste melodije še gredo, ampak katere akorde lahko zaigramo ob melodiji, kdaj in zakaj, kaj je tonaliteta, kaj je modulacija? To so vprašanja, na katere večina učencev dobi odgovor postopoma v okviru izobraževanja v glasbenih šolah. Kako približati teorijo glasbe tudi tistim, ki niso obiskovali glasbenih šol? Ali je možno glasbo predstaviti v vizualni obliki, tako da bi bila bolj razumljiva? Ali lahko, podobno kot za rdečo in oranžno barvo lahko hitro rečemo, da sta si podobni, predstavimo dva akorda vizualno tako, da lahko rečemo, da sta si podobna? Da bi lahko odgovorili na ta vprašanja, moramo najprej odgovoriti na vprašanje, kako deluje človeško zaznavanje in dojemanje glasbe, saj lahko na podlagi razumevanja zaznavanja in dojemanja gradimo modele, ki opisujejo njuno delovanje. Izhodiščni fizikalni dražljaji zunanjega sveta so relativno dobro definirani in lahko merljivi. Ti dražljaji potujejo po slušnem sistemu do možganov, pri čemer se ti dražljaji prenašajo, transformirajo, agregirajo, kar je težje neposredno merljivo. Na eni strani lahko opazujemo in interpretiramo rezul-

tate sistematičnih preizkusov človeških odzivov na posamezne dražljaje, z napredkom neinvazivnih tehnik, kot EEG (elektroencefalografija) in fMRI (funkcionalna magnetna resonanca) pa imamo vpogled tudi v delovanje posameznih struktur v možganih. Rezultate bolj invazivnih tehnik opazovanja, ki se običajno ne izvajajo na ljudeh, lahko na podlagi anatomskih podobnosti med ljudmi in živalmi uporabimo predvsem v nižjih kognitivnih stopnjah procesiranja zvoka. S pomočjo meritev gradimo modele in jih evalviramo na podlagi ujemanja z opazovanim. Da lahko najdemo ustrezne preslikave iz slušno-zaznavnega prostora v vizualno-zaznavni prostor, je na drugi strani potrebno razumeti delovanje vizualnega zaznavnega sistema, ki podobno kot slušni sistem vizualne dražljaje agregira, prenaša in transformira. Mehanizma zaznavanja se v obeh sistemih razlikujeta, vendar pa so v procesih obdelave dražljajev analogije, kot sta na primer metamerizem pri vidni zaznavi in zaznava višine tonov pri sluhu. Oba sta realizirana na precej različen način v svojih zaznavnih sistemih, tudi lokacija procesiranja dražljaja je, gledano po vrstnem redu, na različnih mestih, vendar pa je agregacijski učinek podoben. Takšne podobnosti so izhodišče za čim bolj intuitivno preslikavo iz slušnega v vizualni prostor, kjer poskušamo izkoristiti princip, da se podobnosti v enem sistemu slikajo, kolikor je možno, v podobnostih v drugem sistemu in da so te podobnosti intuitivne, oz. da je krivulja učenja, kaj je podobno in kaj ne, majhna.

V magistrski nalogi bom predstavil nov model vizualne predstavitve harmoničnih odnosov v glasbi, ki temelji na percepciji tonov, omogoča lažjo analizo in razumevanje harmonične strukture glasbe ter služi kot pomoč pri učenju glasbe. Barvni del modela se lahko uporabi samostojno ali pa integrira v druge vizualizacijske sisteme, oziroma jih dopolni, če te vizualne dimenzije še niso sami izkoristili, oziroma jo zamenja, če jo je mogoče menjati. V okviru naloge smo razvili računalniški prototip aplikacije za demonstracijo računskega modela, barvni del modela pa smo integrirali v električno klaviaturo kot demonstracijo možnosti nadgradnje obstoječih glasbil. Poleg analize in učenja [31] lahko takšna predstavitev služi tudi na področju multimedije –

danes, ko je skoraj vsa glasba v digitalni obliki in so te zbirke podatkov velike, je učinkovito preiskovanje in indeksiranje po vizualni obliki ali metapodatkih, ki opisujejo obliko, vse bolj pomembno. V teh primerih se lahko model uporabi za vizualne indikacije, na primer, spreminjanja tonalitet v skladbi, namesto uporabe alfanumeričnih metrik, saj omogoča večjo gostoto prikaza podatkov.

1.2 Cilji in metodologija

Želimo osnovati model predstavitve harmoničnih odnosov v glasbi, ki temelji na percepciji tonov in bi omogočal analizo harmonične strukture glasbe ter služil kot pomoč pri učenju glasbe. Obstoječi modeli sicer delno izpolnjujejo te naloge, vendar niso prešli v širšo uporabo. Eden od vzrokov je, da niso dovolj preprosti ali pa so neintuitivni za uporabo. Predlagani model predstavitve poskuša poenostaviti predstavitev, a še vedno ohraniti njeno uporabnost, hkrati pa se lahko uporablja komplementarno z združljivimi modeli. Če pogledamo z druge strani, lahko vizualni model uporabimo kot orodje za vizualno komponiranje, kjer lahko predstavitev pretvorimo v harmonične odnose med toni.

Pri gradnji modela smo poskušali najti odgovore na naslednja vprašanja:

- Kako definiramo preslikavo iz zvočnega prostora v vidni prostor tako, da bodo harmonični odnosi v glasbi čim bolj intuitivno predstavljeni v vidnem prostoru?
- Ali lahko na podlagi predlaganega modela razširimo obstoječa glasbila, da bodo vizualno prikazala harmonične odnose, ki jih lahko zaigramo?
- Kakšne so prednosti in pomanjkljivosti predlaganega modela v primerjavi z obstoječimi modeli? Ali lahko model integriramo v druge vizualizacije?

Gradnja modela je osnovana na spoznanjih s področja psihoakustike in zaznavanja, kognitivnih modelov, nevroznanosti in teorije glasbe. Zato si

bomo v poglavju 2 naprej pogledali osnove percepcije tonov od zunanjih dražljajev do obdelave v možganih. Podobno je za ustrezen model preslikave potrebno poznati področje vizualne zaznave, predvsem zaznavanje barv in njihovo predstavitev v barvnih prostorih, ki je predstavljena v razdelku 2.6. Na podlagi razumevanja slušnega zaznavanja lahko gradimo modele, ki opisujejo in simulirajo delovanje posameznega dela percepcijskega sistema ali percepcije kot celote; modeli se lahko osredotočijo tudi na posamezno komponento percepcije. Za vizualizacijo nas zanimajo predvsem modeli percepcije tonov in harmonske strukture glasbe. Pregled obstoječih modelov je v poglavju 2. Modeli harmonske strukture glasbe služijo kot osnova za preslikavo v vizualni prostor oz. vizualizacijo, zato sledi v razdelku 2.5 pregled relevantnih modelov preslikav iz slušnega prostora v vizualni prostor. Predlagan model preslikave iz slušnega v vidni prostor je predstavljen v poglavju 3. Poglavje je razdeljeno na tri dele: v prvem delu je potrebno definirati ustrezen barvni prostor, kamor bomo slikali iz tonskega prostora. V drugem delu je izpeljana preslikava. V tretjem delu pa so prikazani rezultati in razprava. Poglavje 4 predstavlja implementacijo modela kot prototipni vizualizacijski program in aplikacijo modela na električno klaviaturo. Na koncu v poglavju 5 sledi komentar rezultatov ter možnosti nadgradnje modela in dodatne aplikacije.

Poglavje 2

Pregled področja

Glavni gradniki glasbe so toni (in njihove višine), ki jih sestavljamo v časovnem zaporedju (horizontalna struktura glasbe – melodija) in lahko zvenijo sočasno (vertikalna struktura glasbe – harmonija), poleg tega pa glasbo gradijo še ritem, dinamika in barva zvoka (spekter zvena glasbil). V pričujoči magistrski nalogi se osredotočam predvsem na vertikalno strukturo glasbe oz. harmonijo. Osnova za razumevanje harmonije je zaznavanje tonov v človeškem slušnem sistemu. V nadaljevanju si bomo pogledali osnove percepcije tonov od zunanjih dražljajev do obdelave v možganih, ki služijo kot osnova za gradnjo modelov od „spodaj navzgor“, torej od dražljajev do zaznavanja, vključno z analizo fizioloških struktur, ki prisostvujejo pri prenosu zvoka oz. njegovi percepciji. Na drugi strani se lahko zaznavanja lotimo od „zgoraj navzdol“, kjer analiziramo rezultate sistematičnih preizkusov človeških odzivov na posamezne dražljaje in zgradimo kognitivne modele, ki ustrezno opisujejo delovanje celega sistema ali pa le dela sistema. Takšni modeli niso nujno neposredno vezani na fiziološke strukture.

2.1 Zvok in percepcija zvoka

Zvoka ne moremo obravnavati zgolj s fizikalnega vidika, saj se akustične meritve na vhodu v uho ne ujemajo vedno z zaznavo oz. interpretacijo v

možganih [29]. Že samo uho se pri prenašanju zvoka iz zunanjega proti notranjem ušesu obnaša nelinearno, na primer pri zaznavanju glasnosti v odvisnosti od frekvence [19] in zaznavanju tonov s seštevki in razlikami frekvenc [59].

Fizikalno je zvok val oziroma valovanje v snovi oziroma mediju, ki povzroči, da se delci snovi premikajo vzdolž smeri valovanja – longitudinalno oziroma vzdolžno valovanje ali pa prečno na smer valovanja – transverzalno oziroma prečno valovanje. Vrsta valovanja, hitrost valovanja in drugi atributi so odvisni od zgradbe oz. vezave atomov ali molekul v snovi. Zaradi tesne povezanosti med molekulami ali atomi je v trdninah možno vzdolžno in prečno valovanje, medtem ko je v tekočinah (kapljevinah¹ in plinih) zaradi nevezane narave atomov in molekul možno le vzdolžno valovanje.

Človeški slušni sistem je sestavljen iz ušesa in centralnih slušnih poti v možganih. Fizikalno se zvok preko zunanjega in srednjega ušesa prenese v notranje uho, kjer se tvorijo živčni impulzi, ki po centralnih slušnih poteh potujejo do slušne skorje velikih možganov. Pri prenosu iz zunanjega ušesa v notranje uho se zvok prenaša skozi več različnih medijev: v zunanjem ušesu je pri človeku medij tipično zrak (plin), preko katerega se valovanje prenese na koščice v srednjem ušesu (mehanski prenos), v notranjem ušesu pa je medij perilimfa (kapljevina), ki povzroči nihanje bazilarne membrane. Nihanje bazilarne membrane zaznajo notranje dlačnice, ki zaradi nihanja pošiljajo živčne impulze po slušnem živcu v možgane. Možgani interpretirajo te impulze kot percepcijo zvoka, ki jo v psihoakustiki pogosto imenujemo kar „zvok“. ANSI/ASA tako zvok v standardu ANSI/ASA S1.1-2013 za ameriško akustično terminologijo opredeli kot (a) Nihanje in pritisk, premik delcev, hitrost delcev itd., ki se propagirajo v mediju s pomočjo notranjih sil (npr. elastičnih, viskoznih); ali pa superpozicija takšnega propagiranega nihanja. (b) Slušna percepcija, ki jo sproži valovanje pod točko (a)²[1]. V nadaljevanju

¹z izjemo površinskih valov

²V izvorniku: (a) Oscillation in pressure, stress, particle displacement, particle velocity, etc., propagated in a medium with internal forces (e.g., elastic or viscous), or the superposition of such propagated oscillation. (b) Auditory sensation evoked by the oscillation

bomo v skladu z običajno rabo besedo zvok uporabljali za oba pomena besede, razen v primerih, kjer iz konteksta ne bo mogoče razbrati, kateri pojem je mišljen.

Kadar zvok obravnavamo kot fizikalni pojav, njegove lastnosti opisujemo z amplitudo, frekvenco, spektrom in trajanjem. V zaznavnem pomenu pa za zvok uporabljamo lastnosti kot glasnost, višina tona (angl. pitch), barva zvoka (angl. timbre) in dolžina tona. Poenostavljeno lahko predstavimo relacijo med fizikalnimi in zaznavnimi lastnostmi sledeče:

Preglednica 2.1: Poenostavljena predstavitev relacij med fizikalnimi in zaznavnimi lastnostmi zvoka.

fizikalno	→	zaznavno
amplituda	→	glasnost
frekvenca	→	višina tona
spekter	→	barva tona (glasbilo)
trajanje	→	dolžina tona

Vendar je percepcija precej bolj zapletena. Na primer, zaznamo lahko višino tona, ki ni prisotna v spektru, kot je v primeru manjkajočega osnovnega tona [55]. V nadaljevanju bomo bolj podrobno pogledali, kako zvok preide iz „fizikalne“ oblike v zaznavno.

Ko pripotuje zvok od vira do ušesa, naprej potuje preko uhlja in zunanjega ušesa preko sluhovoda proti srednjemu ušesu. Naloga uhlja je, da pomaga pri usmerjanju zvoka, prav tako določene frekvence ojača, kar je odvisno od oblike ušesa. Meja med zunanjim in srednjim ušesom je membrana, imenovana bobnič. Za bobničem je votlina srednjega ušesa, kjer je prav tako zrak. Bobnič je povezan na sistem treh koščic v srednjem ušesu: klavdice, nakovalce, stremence, ki služijo kot mehanični vzvod, preko katerega se prenese nihanje do t.i. ovalnega (preddvornega) okenca na polžu. Hkrati se zvok zaradi razmerja površine bobniča in ovalnega okenca ter razlike v dolžini klavdice described in (a).

in nakovalca ojača za približno 15-krat [15]. Ovalno okence je membrana, ki loči srednje uho od notranjega ušesa. Notranje uho je sestavljeno iz polža in krožnih lokov. Slednji so ravnotežni organ, za sluh pa je pomemben polž notranjega ušesa. Polž (kohlea, lat. cochlea) ima koščeno ogrodje, ki obdaja votlino v obliki vijačnice. Če bi vijačnico lahko raztegnili, bi bila votlina približno 30 mm dolg valj, ki je na začetku širši in se postopno zoži proti vrhu. Ovalno okence je na začetku (bazi) polža. Vzдолž po sredini polževe cevi je kožna vezivna cev (kohlearni duktus), napolnjena z endolimfo, ki je v prečnem prerezu približno trikotne oblike in je povezana na obe strani notranjosti polža in ga tako deli na dve stopnici. Zgornji del se imenuje preddvorni hodnik oz. stopnica (lat. scala vestibuli). Vanj pride zvok preko ovalnega okenca. Spodnji predel pa se končuje ob okroglem okencu proti bobnični votlini in ga zato imenujemo hodnik ali stopnica bobnične votline (lat. scala tympani). Oba prostora sta povezana na vrhu oz. koncu polža v helikotremi (lat. helicotrema) in sta napolnjena s perilimfo. V svojem zgornjem in spodnjem delu meji kohlearni duktus z dvema membranama proti hodnikoma. Pomembna je predvsem bazilarna membrana, ki ga razmejuje proti hodniku bobnične votline [15, 19].

Ko stremence zaniha ovalno okence, s tem zaniha tudi perilimfo v obeh hodnikih. Ker je perilimfa nestisljiva kapljevina, je prostornina, ki se pri pritisku stremenca na ovalno okence iztisne navznoter, enaka prostornini tekočine, ki posledično povzroči izbočenje okroglega okenca navzven. To valovanje v tekočini pa povzroči tudi valovanje bazilarne membrane. Valovanje bazilarne membrane zaznajo receptorske celice (notranje dlačnice), ki so na eni strani pritrjene na bazilarno membrano in se ob valovanju raztezajo in krčijo. Posledično ob raztezanju pošiljajo živčne impulze po slušnem živcu v možgansko deblo. Bazilarna membrana na različnih delih niha z različno amplitudo v odvisnosti od frekvence valovanja. Tako se nizki toni registrirajo blizu helikotreme, kjer je membrana širša, visoki pa na bazi, kjer je membrana ožja in bolj napeta. Krajša kot je valovna dolžina, oz. višja kot je frekvenca, bližje ovalnemu okencu bo zvok dosegel maksimalno

amplitudo. Odziv bazilarne membrane, oziroma vzdraženje notranjih dlačnic je zaradi tega tonotopično, kar pomeni, da lahko ločimo posamezne frekvenčne komponente kompleksnega zvoka, oziroma njihove amplitude. Vendar pa ločljivost frekvenčnega zaznavanja ni poljubno natančna in prisotnost drugih frekvenčnih komponent v odvisnosti od bližine opazovani frekvenci in v odvisnosti od amplitude povzroči maskiranje tonov – določenih frekvenc ob prisotnosti drugih frekvenc ne slišimo [15, 19].

Zaradi oblike zunanega ušesa in delovanja prenosnega mehanizma od zunanjega do notranjega ušesa, se izvorno zvočno valovanje od vhoda v uho do zaznave na bazilarni membrani tudi že delno spremeni. Predvsem se to odraža na ojačitvah nekaterih frekvenc, kar je odvisno tudi od oblike ušesa in se razlikuje od posameznika do posameznika [19], poleg tega pa se prenosni sistem ne obnaša popolnoma linearno in se tako ob nihanju na bazilarni membrani zaznajo tudi seštevki in razlike vhodnih frekvenc [59]. Poleg zaznavnih receptorskih celic (notranjih dlačnic), so na bazilarni membrani zunanje dlačnice, ki ojačajo nihanje pri določenih frekvencah in so del psihofizične povratne zanke možganov, ki omogoča ojačitev določenih frekvenc. S tem pa zunanje dlačnice povzročajo nelinearnosti v zaznavanju tonov na bazilarni membrani [15].

Iz notranjega ušesa gredo živčna vlakna v možgansko deblo, kjer potujejo preko več struktur (strukture so podvojene na levi in desni polovici možganov): kohlearno jedro (lat. *nucleus cochlearis*), zgornje olivarno jedro (lat. *nucleus olivaris superior*), spodnji kolikul (lat. *colliculus inferior*). Potem potuje signal preko slušnega dela talamusa (medialno genikulatno jedro, lat. *nucleus geniculatus medialis*) v medmožganih proti slušni skorji (avditornemu korteksu) v senčnem režnju (temporalnemu lobusu) velikih možganov. Živčna vlakna iz levega in desnega ušesa so že v možganskem deblu povezana na svojo (ipsilateralno) in na nasprotno (kontralateralno) stran. Poti iz leve in desne polovice se višje proti slušni skorji še večkrat povežejo v nasprotno polovico. Funkcionalni posnetki možganske aktivnosti kažejo, da se oblike tonotopične organizacije pojavljajo na praktično celotni poti od dlačnic v

notranjem ušesu do slušne skorje v velikih možganih [67].

2.2 Toni in percepcija tonov

Glavni gradniki glasbe so toni (in njihove višine). Glasbeni ton je sestavljen iz več čistih sinusoidnih komponent. Te komponente so tipično celoštevilski večkratniki najnižje frekvenčne komponente v spektru glasbenega tona. Celoštevilskim večkratnikom najnižje frekvence glasbenega tona pravimo tudi alikvotni toni, najnižji frekvenci pa osnovna frekvenca. Ko uho zazna frekvenčne komponente posameznega sestavljenega glasbenega tona, ki ga sestavljajo ton z osnovno frekvenco in alikvotnimi toni, možgani takšno kombinacijo frekvenčnih komponent interpretirajo kot "ton" določene višine. Višina tona je osnovana na tipično najnižji frekvenci sestavljenega tona. Celoten spekter, oziroma razporeditev in amplituda vseh tonskih komponent glasbenega tona pa določa „barvo zvoka“ (vrsto glasbila). V nasprotju s seštevki in razlikami frekvenc, ki so posledica nelinearnega odziva ušesa, se manjkajoči osnovni toni dodajo pri interpretaciji tonov v možganih [25]. Ideja o tonotopični zaznavi tonov, kjer višino tona določa osnovna frekvenca, je bila prevladujoča v 19. stoletju (Helmholtz, Ohm) [13], vendar ne more razložiti pojavov, kot je na primer manjkajoči osnovni ton (angl. missing fundamental): če glasbenemu tonu s tonsko komponento pri neki osnovni frekvenci in alikvotnimi toni odstranimo osnovno frekvenco, ostale višje frekvenčne komponente sestavljenega tona pa ostanejo, še vedno zaznamo ton iste višine [55]. Zgodovinsko gledano lahko zgodnje modele zaznavanja višine tonov delimo v skupino, ki je temeljila predvsem na tonotopičnosti, kjer se zaznavanje višine tona vrši na podlagi analize posameznih frekvenčnih komponent, in skupino, ki je temeljila na časovni analizi, oz. frekvenci pošiljanja impulzov preko slušnega živca. Obsežnejši zgodovinski pregled hipotez o percepciji tonov je opravil deCheveigné [13]. Iz tonotopičnih modelov izhajajo modeli, ki računajo višino tona na podlagi ujemanja vzorcev (začetnik deBoer). Iz modelov, ki temeljijo na časovni analizi signalov, pa izhajajo modeli, ki računajo tone na podlagi

samokorelacije signala (angl. autocorrelation; začetnik Licklider). Vendar pa imata obe skupini določene problematične primere, ki jih ne zajemata. Trenutno najboljši pristopi kombinirajo spoznanja obeh skupin, saj se kaže, da se pri analizi v slušnem sistemu uporabljata oba načina zaznave [5, 3].

Modeli, ki temeljijo na ujemanju vzorcev, izkoriščajo ugotovitev, da možgani rekonstruirajo manjkajoče dele nepopolnih vzorcev (kot na primer pri rekonstrukciji slike v slepi pegi očesa). Če iz spektra kompleksnega tona zbrisemo s pomočjo filtra osnovno frekvenco, bi možgani zapolnili izbrisano frekvenco, saj tonov naravnih izvorov, kjer manjka zgolj osnovna frekvenca, praktično ni. Predloga za višino posameznega tona, ki je sestavljena iz manjkajočega osnovnega tona in ostalih frekvenc, pa se precej dobro ujema že zaradi prisotnosti ostalih frekvenc. Prvi takšen model je predstavil De Boer [11], bolj znani modeli pa so modeli Goldsteina [22], Wightmana [74] in Terhardta [70]. Terhardt predpostavlja, da so predloge, na podlagi katerih se rekonstruira manjkajoča frekvenca, oziroma manjkajoči ton – t.i. virtualni ton oz. skupina kandidatov navideznih tonov, naučene. Podobno Goldsteinov model na podlagi verjetnosti določi najbolj verjetno višino tona, vendar pa ne poda mehanizma, kako se verjetnosti določijo. Problem modelov na podlagi tonotopičnih vzorcev je, da tudi toni, ki ne povzročijo odziva na neki fiksni točki na bazilarni membrani, povzročijo občutek tona z višino. Primer takega tona je zaporedje klikov (angl. click train) [61]. Pri modelih na podlagi samokorelacije se vzame krajši vzorec vhodnega zvoka se korelira (pomnoži) s samim seboj v različnih časovnih točkah, zmnožki pa se nato seštejejo. Če se vzorca v neki točki ujemata, je njun zmnožek velik in rezultat samokorelacije ima maksimume pri periodi ponavljanja signala. Obstajajo tudi modeli z odštevanjem in seštevanjem kvadratov signala, ki pa so ekvivalentni tistim z množenjem in seštevanjem [12]). Model je prvi predlagal Licklider [39], kasnejše dopolnjene modele pa so predlagali med drugimi Meddis in Hewitt [46], Moore [49], Patterson et al. [53]. Modeli izvajajo samokorelacijo na vzorcih poslanih impulzov po nevronih slušnega živca, kar je fiziološko pravilno; obstajajo pa tudi modeli, kot na primer Yostov [75], ki pri samokorelaciji

poenostavljeno operirajo neposredno na vhodnem zvočnem signalu v uho, kar pa nima fiziološke osnove.

Na posamezno notranjo dlačnico v polžu notranjega ušesa je povezanih od 10 do 30 živčnih vlaken. Kadar se dlačnica raztegne, se stereocilije (podaljški na zgornjem delu dlačne celice) premaknejo, nato se zaradi uhajanja kalijevih ionov iz endolimfe v dlačnico le-ta depolarizira, kar povzroči odprtje kalcijevih kanalov. Skozi kanale v dlačnico vstopi kalcij, ki sproži v spodnjem delu dlačnice izločanje živčnega prenašalca (nevrotansmitorja) v sinaptično špranjo (špranja med dlačnico in končičem živčnega vlakna), kar povzroči, da je verjetnost, da se sproži živčni impulz, večja. Ko se stereocilije premikajo v nasprotno smer, se živčni prenašalec odstrani iz špranje, zato je verjetnost, da se sproži impulz, majhna. Če je nihanje dovolj počasno, se živčni impulz tako proži fazno sinhrono z gibanjem. Zaradi „počasnosti“ sproščanja živčnega prenašalca ni nujno, da se impulz sproži vsako periodo, ampak lahko tudi nekaj period izpusti. Sinhronost pri večjih frekvencah, kot na primer nad 1500 Hz, je slaba, nad 5000 Hz pa je praktično ni [15], kar je problematično za razlago percepcije tonov pri višjih frekvencah, kjer je zaznavanje lahko tudi tonotopično. Dodaten problem samokorelacije je, da potrebuje zamik (in množenje), s pomočjo katerega signal zamaknemo, da ga lahko koreliramo, kar je problematično, saj zaenkrat še ne poznamo, katera fiziološka struktura v slušnih poteh možganskega debla bi lahko opravljala to nalogo [47]. Mehanizem, kako točno deluje zaznavanje višine tona še ni razjasnjen, se pa njegove učinke z različnimi modeli bolj ali manj natančno aproksimira. Področje zaznave tonov je zanimivo za ekstrakcijo tonov v polifonični glasbi [20], področje je primerno za uporabo tehnik strojnega učenja, saj obstajajo večje baze transkripcij, ki se lahko uporabljajo za učne množice in testiranje. V zadnjih letih vse bolj v ospredje prihajajo modeli z nevronskimi mrežami, kot na primer [2, 65], vendar je problem nevronskih mrež, da ne razlagajo delovanja. Transparenten model, ki simulira hierarhijo nevronov razvija Pesek s kolegi [54].

Za naš model je področje percepcije tonov zanimivo z več vidikov: na

osnovi zaznavanja so grajeni modeli sorodnosti tonov, tonskih razredov, akordov, tonaliteta, ki jih uporabljamo za grajenje modela preslikave. Na drugi strani naš model vizualizacije kot vhod prejema tone, bodisi neposredno kot zvočni zapis, kjer je potrebno simulirati delovanje slušnega sistema in izluščiti ustrezne informacije o harmonični strukturi, bodisi kot že izluščene tone z dodatnimi metapodatki, kjer uporabimo obstoječe metode za ekstrakcijo tonov iz zvočnega zapisa.

2.3 Percepcija dveh in več sočasnih tonov

Zaznavanje višine tona deluje med drugim na čistih sinusnih tonih (samo z osnovno frekvenco), na glasbenih tonih z osnovno frekvenco in alikvoti ter na tonih, kjer smo iz spektra zbrisali osnovno frekvenco, lahko pa tudi kakšne alikvotne tone. Tako se pojavi vprašanje, koliko frekvenčnih komponent lahko zberemo, preden začnemo zaznavati dva tona, oziroma tudi obratno, kaj če imamo dva sočasna glasbena tona, ki se jima določene frekvenčne komponente prekrivajo – ali v takšnem primeru možgani poskušajo interpretirati kombinacijo tonov kot en ton? Percepcija je odvisna od kombinacije frekvenčnih komponent in posledično lahko zaznamo več tonov ali pa tudi en sam ton [28, 48]. Če upoštevamo zlivanje posameznih frekvenčnih komponent sestavljenega tona, lahko kombinacijo obravnavamo kot neke vrste glasbeni ton, odvisno od stopnje zlivanja. Če na zlivanje gledamo z vidika Terhardtovega modela virtualnih tonov, možgani za sočasno kombinacijo tonov sestavijo množico kandidatov za navidezni ton, če je to en ton ali pa skupina tonov, ki so celoštevilski večkratniki enega od kandidatov, pogosto slišimo en ton. Kadar je kandidatov več, ne moremo nedvoumno določiti enega kandidata in slišimo več tonov, zven tona pa je lahko bolj disonanten. Idejo perceptualnega zlivanja tonov, oziroma poenostavljanje tonov v možganih lahko uporabimo kot osnovo za vizualizacijo odnosov med toni, kjer so toni, ki se perceptualno zlivajo, predstavljeni podobno [9, 31].

2.3.1 Percepcija konsonance in disonance

Harmonijo lahko opredelimo kot sočasno zvenenje več tonov. Kombinacija tonov lahko zveni konsonantno ali disonantno, oziroma je razpon med konsonanco in disonanco zvezen – tako določeni toni zvenijo bolj ali manj konsonantno/disonantno od ostalih. Pomen besed konsonanca in disonanca se je skozi zgodovino spreminjal in tako so določena sozvočja veljala za bolj disonantna, v kasnejših obdobjih pa bolj konsonantna [68]. Konsonanca oz. disonanca kot poslušalcu prijetno ali neprijetno zvenenje sočasnih tonov imenujemo glasbena konsonanca [69] in je zgodovinsko in kulturološko pogojena [6]. Plomp in Levelt [56] konsonanco in disonanco ožje opredelita kot hrapavost oz. grobost tona (angl. roughness), torej kako enakomerno se dva tona zlivata. Takšno ožjo opredelitev imenujemo senzorična disonanca [69]. Čeprav je senzorična disonanca dveh sočasnih čistih sinusnih tonov največja, če sta frekvenčno oddaljena četrtno širine kritičnega pasu in glasbeno konsonantni intervali, kot so kvinta in terca, ne izstopajo, saj funkcija disonance z oddaljenostjo večjo od četrtnine kritičnega pasu monotono pada, so glasbeni toni večinoma sestavljeni iz več frekvenčnih komponent. Če upoštevamo vse pare frekvenčnih komponent sestavljenih tonov, ugotovimo, da imamo vrhove senzorične konsonance pri glasbeno konsonantnih intervalih, kot so sozvočje, oktava, kvinta, kvarta, terca itd. [56]. Razlog za takšen vzorec je, da se lahko alikvotni toni, ki sestavljajo posamezne tone, prekrivajo. V tem primeru par alikvotov ne povzroča senzorične disonance. Lahko pa se ne prekrivajo in če so pari oddaljeni okrog širine kritičnega pasu ter je takih parov veliko, je skupna senzorična disonanca kombinacije glasbenih tonov večja. Če vzamemo dva tona, ki sta oddaljena za razdaljo pol kritičnega pasu in enega od njiju transponiramo v višjo oktavo (podvojimo osnovno frekvenco), zveni kombinacija manj „grob“, vendar pa še vedno ne konsonantno; ta disonanca je povezana z zlivanjem percepcije tonov [70], ki je predstavljena v razdelku 2.3.

Percepcija konsonance in disonance se ne pojavlja zgolj ob sočasno zvenceh tonih, ampak je tudi pogojena s toni, ki so zveneli pred zaigranim tonom. Na primer, kombinacija tonov C in G bo v kontekstu C-dura zvenela konsonantno,

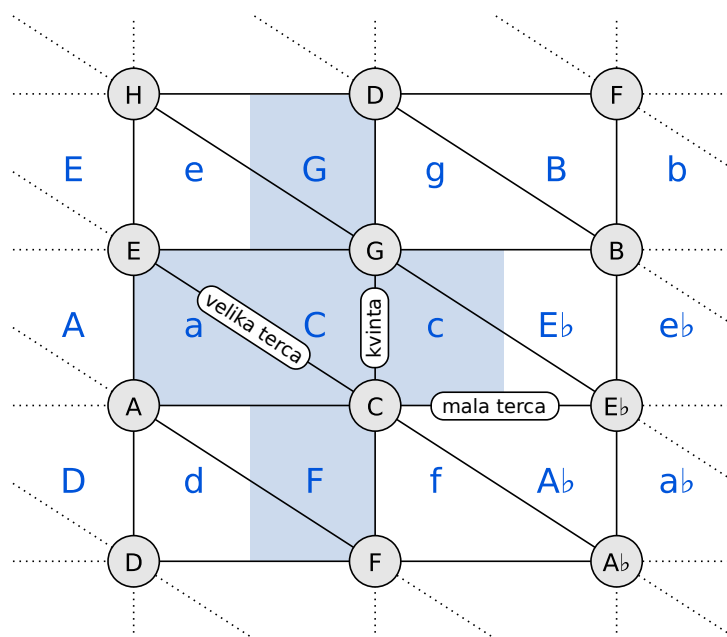
medtem ko bo v kontekstu Fis-dura zvenela disonantno. S raziskovanjem konsonance ali disonance testnih tonov in akordov, ki sledijo določenemu zaporedju akordov, oziroma so zaigrani v kontekstu neke tonalitete, se je obsežneje ukvarjala Carol Krumhansl [33, 34]. V seriji sistematičnih poskusov človeških odzivov na različna zaporedja tonov in akordov so poslušalcem najprej predvajali zaporedje tonov (kontekst), nato je sledil poskusni ton (angl. probe-tone) ali akord. Poslušalci so morali oceniti, kako dobro se ton ali akord ujema s kontekstom. Na podlagi rezultatov sta Krumhansl in Kessler s pomočjo večrazsežnostnega lestvičenja (MDS, angl. multidimensional scaling) dobila štiridimenzionalni model tonalitetnega prostora, kjer so posamezne tonalitete razvrščene glede na sorodnost – sorodne tonalitete so bližje skupaj, manj sorodne so bolj oddaljene [33].

2.4 Predstavitev odnosov med toni

Sorodnost med toni lahko obravnavamo z analizo sistematičnih preizkusov človeških odzivov na posamezne zvočne dražljaje, na drugi strani pa lahko na problem odnosov med toni, akordi in tonalitetami gledamo tudi z vidika glasbene teorije, čeprav je glasbena teorija v bistvu grajena na podlagi zaznavanja glasbe, vendar pa ne obravnava zvoka podrobneje iz psihoakustičnega vidika. Skozi zgodovino glasbe so se izoblikovali mnogi modeli, kako sistematično predstaviti sorodnost med toni in predvsem med tonalitetami (pregled zgodovine v [58]). Na Eulerjevem modelu mreže tonov (nem. Tonnetz, 1739) vsaj delno bazirajo številni novejši modeli, med drugim modeli Lewina [38], Gollina [23], Chew [8], Lerdahla [36, 37]. Mreža tonov predstavi prostor tonov v dveh dimenzijah, kjer posamezen ton predstavlja točko, njegovi sosedi pa so toni, ki so od tona oddaljeni za kvinto navzgor ali navzdol ter malo in veliko terco navzdol in navzgor. Enako velja za katerikoli ton v mreži tonov. Slika 2.1 prikazuje izsek tonske mreže. Če uporabimo enakomerno temperirano uglasitev pri sprehajanju v katerikoli smeri od nekega tona, prej ali slej (število korakov je odvisno od smeri) pridemo do istoimenskega

(enharmonično ekvivalentnega) tona – vzorec razporeditve tonov se v obeh dimenzijah ravnine ponavlja. Tako lahko model predstavimo v cilindrični obliki (ponavljanje v eni dimenziji združimo v plašč cilindra) ali v toroidni obliki (v krog združimo še ponavljanje v drugi dimenziji).

Na podoben način lahko poskušamo predstaviti odnose med regijami tonaliteta. Podobno kot Weber (1817) je Schönberg [62] v dveh dimenzijah izrisal „regije“, v katere se premakne harmonija tekom skladbe. Na primer, iz regije tonike se lahko premakne v sosednje regije dominante, subdominante, relativnega in paralelnega mola. Slika 2.1 prikazuje povezavo med modelom tonske mreže in Schönbergovimi regijami tonaliteta. Podobno kot pri tonski mreži opazimo ciklični vzorec kvintnega kroga. Z manjšimi modifikacijami modela in enharmonsko ekvivalenco se lahko model predstavi v obliki torusa medtonalitetnih razmerij [57]. Krumhansl in Kessler sta na podlagi psihoakustičnih poskusov, opisanih v razdelku 2.3, ugotovila, da lahko z upoštevanjem razdalje med tonalitetami dobimo razporeditev, ki je prikazana na sliki 3.16a, kar je ekvivalentno površini prej omenjenega torusa medtonalitetnih razmerij. Oba toroidna modela sta štiridimenzionalna, vendar ju prikazujemo v obliki tridimenzionalnega torusa, kar pomeni, da je pri razdaljah med toni in tonalitetami v torusu potrebno upoštevati vse štiri dimenzije, kar podrobneje predstavimo v razdelku 3.2 pri gradnji našega modela. Burgoyne in Saul na podlagi MVU (razvijanja z maksimalno varianco, angl. maximum variance unfolding) podatkov Krumhanslove in Kesslerjeve predlagata alternativno obliko tonalitetnega prostora [4], pri kateri sta predvsem paralelni dur in mol bolj oddaljena kot v Schönbergovem, so pa lokalne značilnosti boljše ohranjene. Harte je toroidni model razširil na 6 dimenzij za pomoč pri razpoznavanju akordov iz zvočnega zapisa glasbe [26]. Vendar nas za pomoč pri vizualizaciji predvsem zanimajo nižje dimenzionalni modeli. S posplošitvami oziroma večdimenzionalnim modelom tonskih mrež se je ukvarjal obsežnejše Tymoczko [72]. Na drugi strani pa nekateri modeli tudi dodatno poenostavljajo tonski prostor, kot na primer Gatzschejev model [18].



Slika 2.1: Izsek tonske mreže (obarvana s črno) okrog tona C v relaciji s Schönbergovimi regijami za C-dur (obarvane z modro). Povzeto po Harte [26].

2.5 Vizualizacije glasbe

Glasbo lahko vizualiziramo na mnogo načinov, tako da uporabimo prostorske in barvne dimenzije vida [30, 7]. Vizualiziramo lahko na primer melodijo, harmonijo, ritem, teksturo, strukturno hierarhijo in tudi kombinacije le-teh. Osredotočili se bomo predvsem na tiste vizualizacije, ki vizualizirajo harmonijo, pri čemer nas posebej zanimajo tiste, ki za vizualizacijo uporabljajo barvni prostor, prav tako pa tudi tiste, ki za vizualizacijo uporabljajo različno število prostorskih dimenzij. Modeli vizualizacije v prostorskih dimenzijah so z vidika naše vizualizacije zanimivi zaradi dveh razlogov: barvni prostor lahko predstavimo v do treh dimenzijah in lahko takšne modele neposredno uporabimo pri gradnji barvnega modela vizualizacije – v to skupino spadajo med drugim geometrijski modeli tonskega prostora in prostora tonalitet iz prejšnjega razdelka. Drugi razlog je, da lahko v vizualizacijske modele,

ki uporabljajo prostorske dimenzije tudi za vizualizacijo ostalih lastnosti glasbe, dodamo barvni model, če vizualizacija še ne uporablja barvnega prostora, oziroma nadomestijo obstoječe barvne modele, kjer je to smiselno. Pri vizualizacijah, ki uporabljajo barve za vizualizacijo harmonije, je dokaj pogosto barvanje tonalitet ali tonov po principu kvintnega kroga, na primer pri Sappovi vizualizaciji hierarhije tonalitet [60] ali Mardirossianovi vizualizaciji prehajanja med tonalitetami [44]. Uporablja pa se tudi barvanje po frekvenci [21, 16], ki se predvsem pogosto uporablja za barvanje tipk na klaviaturah. Vendar pa je takšen način barvanja precej manj primeren za prikaz harmonij, saj sosednost po frekvenci ne pomeni harmonične podobnosti med toni; je pa ta sistem bolj primeren za melodijo, kjer so intervali med toni tipično majhni (takšno barvanje tipk je delno redundantno, saj so tipke že urejene v vrstnem redu po velikosti). Toivianinenova vizualizacija [71] je primer, ki za prikaz harmonij uporablja mrežo medtonalitetnih razmerij in z barvo označuje jakost aktivacije percepcije posamezne tonalitete. Martorel in Gómez [45] uporabljata toroidni model medtonalitetnih razmerij, ki ga preslikata v barvni prostor CIE $L^*a^*b^*$ in tako dobita barve posameznih tonalitet in območij prehodov med tonalitetami. Več o modelu si bomo pogledali v poglavju 3, kjer ga bomo primerjali z našim modelom. Med modeli, ki harmonične odnose predstavljajo z barvami preko preslikave kvintnega kroga, praktično ni modelov, ki bi združevali barve podobnih tonov v eno barvo. Mandrino [43] uporablja tri osnovne barve (rdečo, zeleno in modro) za toniko, dominantno in subdominantno, ki jih potem kombinira pri sestavljanju akordov, vendar so možne kombinacije omejene. Modeli, ki preslikajo tonalitete preko zaporedja kvintnega kroga v barve, najbolj pogosto uporabljajo diskretne razrede za posamezno tonaliteto z redkimi izjemami, kot prej omenjeni model Martorella.

Naš model vizualizacije harmoničnih odnosov v glasbi omogoča združevanje kombinacij tonov na podlagi ideje perceptualnega združevanja tonov v slušnem sistemu [28, 48]. Pri združevanju se upošteva jakost tonov in harmonični kontekst [9, 32]. Tako lahko, poleg osnovne preslikave tonov v barve, pre-

slikamo v barve tudi vse kombinacije tonov in kombinacije tonov z različno jakostjo in v različnih kontekstih. Pri vizualizacijah, ki uporabljajo barve, se pojavljajo različni načini izkoriščanja barvnega prostora. Vizualizacije, ki uporabljajo osnovne barvne kategorije (rdeča, oranžna, rumena, zelena itd.), večinoma izberejo prototipične barvne predstavnike barvnih kategorij. Vizualizacije, ki uporabljajo zvezni barvni prostor, precej pogosto uporabljajo prostor RGB ali njegove cilindrične izvedbe, kot je na primer HSL. Vendar pa je uporaba takšnih prostorov zaradi njihove perceptualne neuniformnosti problematična. Zato si bomo v naslednjem razdelku pogledali, kako deluje zaznavanje barv, v poglavju 3 pa bomo predstavili ustrezen barvni prostor, ki ga lahko uporabimo za vizualizacijo.

2.6 Vid in zaznavanje barv

Barvo (določenega objekta) lahko natančno določimo s pomočjo merjenja spektralne porazdelitve valovnih dolžin svetlobe (oz. elektromagnetnega valovanja) v območju od $4,3 \cdot 10^{14}$ do $7,5 \cdot 10^{14}$ Hz, oziroma valovnih dolžinah od 400 do 700 nm [51]. Vendar nas pri vizualizaciji bolj zanima človeška percepcija svetlobe. V primerjavi s spektrometrom ima človeško oko na mrežnici dve vrsti čutnih (fotoreptorskih) celic – paličice in čepke. Paličice so občutljive predvsem na del spektra z viškom pri modro-zeleni svetlobi (498 nm) in omogočajo monokromatski (enobarvni) vid v slabih svetlobnih pogojih oz. majhnim številom fotonov, ki prispejo na mrežnico. Kadar je fotonov veliko, se paličice desenzibilizirajo in svetlobo primarno zaznavajo čepki. Čepki se delijo na tri podvrste, kjer je vsaka vrsta občutljiva na del spektra z viškom pri drugi valovni dolžini svetlobe. Delu spektra svetlobe, ki ga pokrivajo te tri vrste čutnih celic pravimo vidni spekter oz. vidna svetloba. Občutljivost posamezne vrste čepkov na različne valovne dolžine svetlobe nam omogoča razločevanje barv. Viški občutljivosti za posamezno vrsto čepkov so pri 564 nm (čepek, občutljiv na „rdečo“, oziroma daljše valovne dolžine – čepek L), 534 nm (čepek, občutljiv na „zeleno“, oziroma

srednje valovne dolžine – čepek M) in 420 nm (čepek, občutljiv na „modro“, oziroma krajše valovne dolžine – čepek S) [51]. Čeprav pogosto rečemo, da je čepek L občutljiv na rdečo, ima višek občutljivosti pri rumeno-zelenem delu spektra svetlobe; vendar pa je od vseh treh vrst čepkov še najbolj občutljiv za rdeči del spektra. Ključno za zaznavanje barve je ugotavljanje kontrastov oz. razlike stopnje vzdraženosti posamezne vrste čepkov v primerjavi z drugo vrsto čepkov [24]. Za barvni vid se ugotavljata dva kontrasta: rdeča (čepki L) – zelena (čepki M) in rumena (kombinacija signalov čepkov L in M) – modra (čepki S). Informacija o kontrastih se prenaša po vidnem živcu v možgane v vizualni korteks, kjer se na poti postopoma signali kontrastov posameznih delov slike z mrežnice sestavijo v celotno sliko vidnega polja. Torej možgani na podlagi kontrastov zaznajo posamezno barvo. V povprečju lahko človek tako razloči okrog $2,3 \cdot 10^6$ barv [40]. Če katera vrsta čepkov manjka ali pa se ne razlikuje dovolj od druge vrste, je lahko tak barvni prostor precej manjši. Na primer, če ne moremo ugotavljati kontrasta med „rdečo“ in „zeleno“ (popolna rdeče-zelena barvna slepota – protanopia, deuteranopia), je po ocenah možno razločevati okrog 10^4 barv [50].

Človeško oko ugotavlja zgolj kontraste med dražljaji treh različnih vrst čepkov, pri čemer ima vsak čepek svojo funkcijo občutljivosti, ki je odvisna od valovne dolžine. Posledica tega je, da se izgubi spektralna informacija izvirne svetlobe, oziroma lahko percepcijo iste barve dosežemo z različnimi spektri svetlobe, ki enako vzdražijo posamezno vrsto čepov (metamerizem). Tako lahko z dobro izbranimi valovnimi dolžinami svetlobe ustvarimo percepcijo večine možnih spektralnih barv [52], čeprav uporabljamo zgolj svetlobo treh valovnih dolžin. Da lahko barve ustrezno reproduciramo, je potrebno definirati barvni prostor, kjer lahko z ustreznim mešanjem barvnih komponent ustvarimo zelene barve [63].

Za opis barvnega prostora lahko uporabimo različne barvne sisteme. Pri računalniških monitorjih se uporablja sistem RGB (angl. red, green, blue), ker posamezna slikovna točka zasveti z „barvo“, ki je kombinacija komponent RGB (svetlobe treh valovnih dolžin: rdeče, zelene in modre). Problem

sistema je, da ni v linearnem razmerju s človeško percepcijo – sprememba posamezne vrednosti R, G ali B za isto vrednost $\pm\Delta$ ne pomeni iste spremembe v percepciji in je percepcija odvisna od začetne vrednosti posamezne komponente R, G ali B. Poleg tega je sistem RGB neintuitiven za uporabo, ker je s spreminjanjem komponent težje dobiti zeleno barvo.

Bolj intuitiven sistem za opis barvnega prostora je HSL (angl. hue, saturation, lightness), oziroma podobni sistemi (npr. HSV, angl. hue, saturation, value), kjer imamo komponente za odtenek barve, nasičenost barve in svetlost barve. Vendar pa so ti sistemi le predstavitev barvnega prostora RGB v cilindričnem koordinatnem sistemu in tako tudi ti sistemi niso linearni glede na človeško percepcijo. Med barvnimi prostori, ki linearno zrcalijo človeško percepcijo barv, se pogosto uporabljajo barvni prostori CIE (fr. Commission internationale de l'éclairage)³, predvsem CIE $L^*u^*v^*$ (CIELuv) in CIE $L^*a^*b^*$ (CIELab) [17]. Sistemi so nastali na podlagi meritev spektra svetlobe in človeške percepcije barve za določen spekter svetlobe, oziroma za določene spektralne komponente. Posledično so bili sistemi zasnovani tako, da čim bolj linearno odražajo percepcijo barvnega prostora. Čeprav je komponenta L^* v teh sistemih intuitivna (percepcija svetlosti), pa sta komponenti u^* in v^* ter a^* in b^* manj intuitivni. Zato obstaja še različica CIE $L^*C^*h^\circ$ (CIELCh), ki poleg z L^* , predstavlja barvni prostor s pomočjo odtenkov (h° , angl. hue) in nasičenosti oz. krome (C^* , angl. chroma) [42].

Prednost perceptualno uniformnega barvnega prostora je, da lahko definiramo razlike (razdalje) med barvami, kjer sta posamezni barvi, ki sta perceptualno enako oddaljeni, tudi enako oddaljeni v barvnem prostoru. Tako lahko določimo tudi najmanjšo razliko med barvama, ki jo oko pri človeku z normalnim vidom zazna.

V barvnem prostoru CIELab določajo barvo tri komponente L^* – svetlost (angl. lightness), a^* – razmerje med rdečo in zeleno, b^* – razmerje med modro in rumeno barvo. Razdalja med barvami je definirana kot evklidska razdalja ΔE_{76}^* v prostoru:

³<http://www.cie.co.at/>

$$\Delta E_{76}^* = \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}} \quad (2.1)$$

Ker prostor vseeno ni popolnoma perceptualno uniformen, je bila formula za izračun razdalje še dvakrat nekoliko popravljena. Trenutno aktualno različica [64] definira razdaljo ΔE_{00}^* kot:

$$\Delta E_{00}^* = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T \frac{\Delta C'}{k_C S_C} \frac{\Delta H'}{k_H S_H}} \quad (2.2)$$

Izračun razdalje vsebuje popravek R_T za rotacijo odtenka v problematičnem modrem delu okrog 275° , kompenzacije za nevtralno barvo (L' , C' , H') in kompenzacije za vsako od komponent (S_L , S_C , S_H).

Izračuni posameznih popravkov so podani v dodatku. Razlika ΔE je definirana v barvnem prostoru CIELab; če hočemo izračunati razliko med posameznimi barvami iz ostalih sistemov je potrebno barvo najprej pretvoriti v sistem CIELab. Formule za pretvorbe med sistemi so podane v dodatku.

Pri vizualizaciji nam uniformni barvni prostor omogoča preslikavo neke podatkovne dimenzije v barvno dimenzijo, pri čemer lahko tudi razdalje v podatkovni dimenziji prenesemo na razdalje v barvni dimenziji. V razdelku 3.1 pokazali razliko med uporabo uniformnega in neuniformnega barvnega prostora na primeru, ki ga bomo uporabili v naši vizualizaciji.

Poglavje 3

Model

Preden se lotimo preslikave, je potrebno ustrezno definirati barvni prostor. Poglavje je tako razdeljeno na tri dele: v prvem delu definiramo barvni prostor, kamor bomo slikali iz tonskega prostora. V drugem delu je izpeljana preslikava na podlagi modelov sorodnosti tonov in tonalitet. V tretjem delu pa so predstavljeni rezultati in razprava.

3.1 Barvni prostor

Lastnosti barvnega sistema želimo čim bolj izkoristiti pri vizualizaciji – koncepti, ki so v prostoru, kjer so definirani, oddaljeni, so lahko predstavljeni z oddaljenimi točkami v barvnem prostoru; koncepti, ki so si podobni, pa z bližnjimi. Če obstaja razlika med lastnostmi, želimo, da se tudi ta odraža v barvnem prostoru z ustrezno razdaljo, da je razlika zaznavna. Zato smo pri preslikavi tonsko-harmonskega prostora v barvni prostor uporabili sistem CIELCh¹.

CIELCh v primerjavi z CIELab uporablja namesto kartezičnega koordinatnega sistema cilindrični koordinatni sistem. L^* ostane nespremenjen, cilindrični koordinati C^* in h° pa lahko dobimo iz a^* in b^* s pretvorbo:

¹Pri CIELCh imamo dve različici, ena je izpeljana iz sistema CIELuv, druga iz CIELab. Mi bomo uporabljali slednjo.

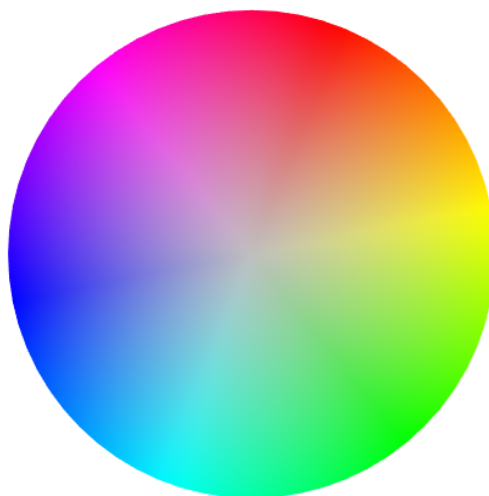
$$C^* = \sqrt{a^{*2} + b^{*2}}, \quad h^\circ = \arctan\left(\frac{b^*}{a^*}\right) \quad (3.1)$$

Vrednosti h° predstavljajo odtenke barve in jih v CIELCh podamo z vrednostmi kotov na intervalu $[0, 360)$. L^* predstavlja svetlost, kjer vrednost nič pomeni popolnoma temno, 100 difuzne svetle barve, popolnoma belo pa še pri višjih vrednostih. Podobno C^* predstavlja pri vrednosti 0 popolnoma nenasičeno barvo, nekje pri 100 že precej nasičeno, vendar pa so lahko vrednosti še višje. Vrednosti vseh treh komponent so decimalne številke.

Za prikaz barv na računalniškem zaslonu ali klaviaturi s pomočjo RGB-svetlečimi diodami je potrebno barve iz prostora CIELCh pretvoriti v prostor RGB (sRGB, Adobe RGB). Problem se pojavi, ker je barvni prostor CIE večji kot prostor sRGB ali Adobe RGB [52], tako da moramo pri preslikavi iz tonsko-harmonskega prostora v prostor CIELCh paziti, da kasneje pri pretvorbi v RGB izgube ne bodo prevelike. Iz prostora CIELCh moramo najprej barve pretvoriti v CIELab, nato CIELab v CIE XYZ, iz CIE XYZ v sRGB ali Adobe RGB. Pretvorbe so podane v dodatku.

Na drugi strani pa želimo uporabiti čim večji del prostora CIELCh, da bo razlika ΔE^* med barvami čim večja, saj lahko tako pri vizualizaciji prikažemo več različnih vrednosti atributov oz. je razlika med njimi bolj očitna. Če uporabljamo prostor HSV ali HSL, ki perceptualno ni uniformen, v nekaterih delih barvnega prostora dobimo razliko med barvami večjo, kot je v perceptualno uniformnem prostoru. Hkrati pa to pomeni, da je namesto razlike v odtenku verjetno prisotna še razlika v npr. svetlosti, posledično pa lahko uporabnik sklepa, da je razlika v svetlosti posledica spreminjanja še kakšnega drugega atributa, poleg tistega, ki smo ga preslikali na odtenek. V nekaterih predelih prostora pa je razdalja manjša kot v perceptualno uniformnem (slika 3.9).

Vzemimo za primer barvni prostor sRGB, v katerega bomo morali pretvoriti barve, preden jih bomo lahko prikazali na običajen računalniški zaslon ali s svetlečimi diodami na klaviaturi. Posamezno barvo v tem barvnem sistemu opišemo s tremi komponentami, ki predstavljajo rdečo, zeleno in



Slika 3.1: Barvni krog v prostoru HSL pri $S = 100$, $L_{HSL} = 50$.

modro. Vrednosti posamezne komponente ležijo na intervalu $[0, 1]$. Ko izrišujemo na zaslon, jih pogosto podamo v celoštevilski 8-bitni obliki, v tem primeru so celoštevilske vrednosti posamezne komponente na intervalu $[0, 255]$. Celoten barvni prostor RGB lahko predstavimo v obliki kocke, kjer oglišča kocke predstavljajo vrednosti za črno, rdečo, rumeno, zeleno, cian (sinjo), modro, magento (škrlatno) in belo. Vrednosti komponent za posamezno barvo so podane v preglednici 3.1. Da predstavimo barvni sistem RGB v cilindričnih koordinatah, lahko uporabimo sistema HSL in HSV (oziroma HSB). V barvnem sistemu HSL so barve podane kot odtenek – H (vrednosti na intervalu $[0, 360)$), nasičenost – S (vrednosti na intervalu $[0, 100]$), svetlost – L (vrednosti na intervalu $[0, 100]$). Vrednosti HSL za mejne točke sistema RGB so podane v preglednici 3.1. Vendar pa sistem RGB ni perceptualno uniformen, kar pomeni, da če na primer dve barvi spremenimo za vrednost $\Delta(R, G, B)$, perceptualna sprememba svetlosti ne bo nujno enaka. Da lahko izračunamo perceptualno razliko med dvema barvama, je potrebno najprej barvo iz sistema RGB (ali HSL) pretvoriti v sistem CIE Lab. Vrednosti $L^*a^*b^*$ za mejne točke sistema RGB so podane v preglednici 3.1.

Če izračunamo razdalje med posameznimi mejnimi vrednostmi sistema

Preglednica 3.1: Mejne vrednosti in središčna vrednost barvnega sistema RGB s pretvorbami v sisteme HSL, CIE $L^*a^*b^*$ in CIE $L^*C^*h^\circ$.

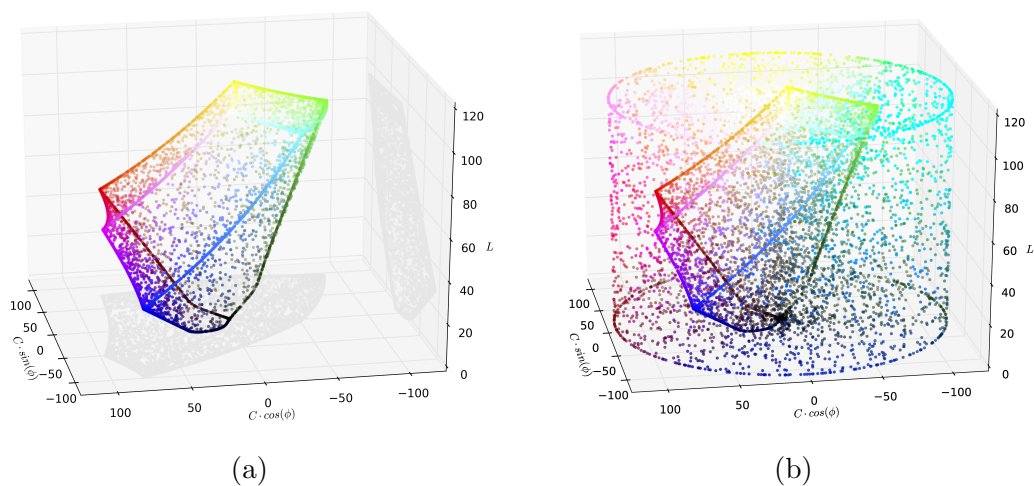
	RGB			HSL			CIE $L^*a^*b^*$			CIE $L^*C^*h^\circ$		
	R	G	B	H	S	L	L^*	a^*	b^*	L^*	C^*	h°
črna	0	0	0	0	0	0	0,00	0,00	0,00	0,00	0,00	360,00
rdeča	1	0	0	0	100	50	53,59	0,00	-0,01	53,23	104,58	40,00
rumena	1	1	0	60	100	50	100,00	0,01	-0,01	87,73	119,78	136,02
zelena	0	1	0	120	100	50	53,23	80,11	67,22	32,30	133,81	306,29
sinja	0	1	1	180	100	50	97,14	-21,56	94,48	97,14	96,91	102,85
modra	0	0	1	240	100	50	87,74	-86,18	83,18	60,32	115,57	328,23
magenta	1	0	1	300	100	50	91,12	-48,08	-14,14	91,12	50,12	196,39
bela	1	1	1	0	100	100	100,00	32,30	79,20	-107,86	0,02	296,81
siva	0,5	0,5	0,5	0	0	50	60,32	98,25	-60,84	53,39	0,00	-0,01

Preglednica 3.2: Razdalje ΔE_{00}^* med sosednjimi odtenki v sistemu HSL pri $S = 100$ in $L_{HSL} = 50$. V perceptualno uniformnem sistemu bi bile vrednosti ΔE_{00}^* med enako oddaljenimi odtenki enake.

$barva_1$	$barva_2$	ΔE_{00}^*	$ \Delta L^* $	$ \Delta C^* $	$ \Delta h^\circ $
rdeča	rumena	64,31	43,91	7,67	62,85
rumena	zelena	23,40	9,40	22,87	33,16
zelena	sinja	34,53	3,38	69,66	60,37
sinja	modra	66,46	58,81	83,70	109,90
modra	magenta	32,41	28,02	18,25	21,95
magenta	rdeča	42,59	7,09	10,99	71,77

sRGB, vidimo, da se perceptualne razlike ΔE_{00}^* med sosednjimi odtenki (sosednje ekvidistančne vrednosti H v sistemu HSL pri $S = 100$, $L_{HSL} = 50$) razlikujejo med seboj v nekaterih primerih skoraj za faktor dva. Na sliki 3.1 je prikazan presek barvnega prostora HSL pri konstantni vrednosti $L = 50$. Vidne so razlike med sosednjimi odtenki in razlike v prehajanju iz nasičene barve (obod kroga) proti nenasičeni (središče kroga, $HSL = (0, 0, 50)$). Odtenki na obodu imajo dokaj različne perceptualne svetlosti, $\overline{\Delta L^*} = 25, 10$, kar je četrtnina razlike med belo in črno; največja razlika pa je med modro in sinjo, kjer je razlika večja kot razlika med črno in sivo. V preglednici 3.2 so podane razlike med sosednjimi odtenki v sistemu HSL pri $S = 100$ in $L_{HSL} = 50$, v preglednici 3.3 pa razlike med barvami s HSL pri $S = 100$ in $L_{HSL} = 50$ ter belo, sivo in črno. Bela in črna iz sistema RGB/HSL sta oddaljeni $\Delta E_{00}^* = 100$, bela in siva $\Delta E_{00}^* = 33, 24$; siva in črna pa $\Delta E_{00}^* = 39, 93$. Iz podatkov je razvidno, da sistem RGB ni najbolj primeren za uporabo v vizualizaciji, saj so ob spremembi vrednosti ene komponente v sistemu RGB vrednosti v perceptualno uniformnem sistemu preveč spreminjajo. Sliki 3.2a in 3.2b prikazujeta barvni prostor sRGB izrisan v prostoru CIELCh. Opazna je deformacija kocke sRGB.

Za uporabo barve v vizualizaciji je potrebno uporabiti vizualno uniformen



Slika 3.2: Barvni prostor sRGB izrisan v koordinatnem sistemu prostora CIE $L^*C^*h^\circ$ (a) in barvni prostor sRGB skupaj barvnim prostorom CIE $L^*C^*h^\circ$ za $L^* \in [0, 110]$, $C^* \in [0, 120]$, $h^\circ \in [0, 360]$ (b).

Preglednica 3.3: Razdalje ΔE_{00}^* med odtenki v sistemu HSL pri $S = 100$ in $L_{HSL} = 50$.

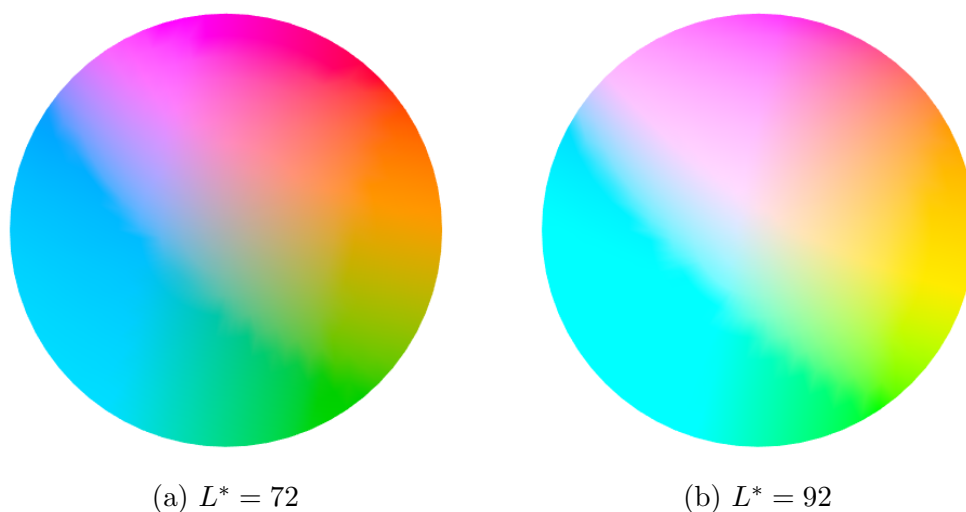
barva	črna	siva	bela
rdeča	50,41	31,20	45,82
rumena	101,20	43,97	30,54
zelena	87,87	41,70	33,27
sinja	90,53	37,79	25,70
modra	39,68	38,63	64,20
magenta	56,71	32,66	42,18

barvni prostor, pri čemer je potrebno poiskati ustrezno preslikavo med podatki in koordinatami v barvnem prostoru.

V primeru vizualizacije glasbe je zaradi krožne narave poenostavitve tonskega prostora primerna uporaba cilindričnega koordinatnega sistema pri preslikavi v barvni prostor. Analogno, kot smo dobili barvni krog s presekom barvnega prostora HSL pri $L_{HSL} = 0,5$, bi radi dobili perceptualno uniformen krog s pomočjo preseka barvnega prostora CIELCh. Vendar pa je potrebno upoštevati še sledeče omejitve:

- Ker je barvni sistem RGB podmnožica barvnega prostora CIELab, oziroma CIELCh, moramo izbrati takšno podmnožico barv iz prostora CIELCh, da pri preslikavi v prostor RGB ne bo prišlo do sovpadanja barv zaradi preslikave barv, ki so izven barvnega prostora RGB (sRGB, Adobe RGB), na robne ploskve tega prostora; oziroma, da bodo izgube zaradi tega čim manjše. Preslikovanje v barvni prostor RGB je potrebno zaradi izrisa na zaslon oziroma krmiljenja svetlečih RGB-diod.
- Prvo omejitev preprosto rešimo s tem, da vzamemo dovolj majhno podmnožico prostora, vendar pa je uporabnost v vizualizaciji slabša, ker imamo na voljo manj različnih barv, oziroma so le-te med seboj manj oddaljene. Pri normalnem človeškem vidu je spodnja meja še zaznavne razlike (JND, angl. just noticeable difference) med barvami pri $\Delta E_{76}^* \approx 2,3$ [41], kar pomeni, da morajo biti posamezne barve, ki jih želimo uporabiti v vizualizacijah, vsaj toliko oddaljene. Razdalja ΔE_{76}^* je za nekatere sploh bolj nasičene odtenke dala prevelike razlike, po popravku ΔE_{00}^* so razlike manjše, kar pomeni, da je JND po ΔE_{00}^* manjši. V nadaljevanju bomo uporabljali aktualno ΔE_{00}^* .
- Prostor naj bo še vedno čim bolj perceptualno uniformen.

Poleg omejitev imamo dodatne zahteve glede vsebovanosti določenih barv v barvnem prostoru: podmnožica prostora naj bi zajemala ali pa se vsaj v odtenku, svetlosti in nasičenosti čim bolj približala barvam, ki so prototipični predstavniki osnovnih jezikovnih barvnih kategorij (npr. prototipične rdeče,



Slika 3.3: Barvni krog, ki ga dobimo s presekom barvnega prostora CIE $L^*C^*h^\circ$ pri $L^* = 72$ in $L^* = 92$. Obod kroga je pri $C^* = 125$.

rumene, zelene itd.). Motivacija za to zahtevo je, da vizualizacija uporablja barve, za katere je populaciji uporabnikov splošni konsenz, kako so poimeno-vane. Prototipične barvne kategorije so pogojene z jezikom in se razlikujejo od jezika do jezika in od kulturnega prostora do kulturnega prostora, vendar je v povprečju pri nekaterih glavni prototipičnih kategorijah precej prekrivanja med jeziki [10].

Primer preseka podprostora CIELCh pri konstantni vrednosti L^* , nam dá barvni krog z uniformno razliko med odtenki in enotno svetlostjo barve. Slika 3.3 prikazuje barvni krog, ki ga dobimo pri presekih $L^* = 72$ in $L^* = 92$. Problem presekov je, da slabo zadostita zahtevi, da se približamo prototipičnim predstavnikom barvnih kategorij. Pri $L^* = 72$ so barve pri odtenkih od modre do rdeče v redu, medtem ko so v rumenem pretemne, ravno obratno je pri $L^* = 92$. Poleg tega pri $L^* = 92$ pri preslikavi v prostor sRGB izgubimo del barvnega prostora v sinjem in modrem delu spektra.

Izbiro podprostora lahko obravnavamo kot optimizacijski problem, kjer poskušamo minimizirati izgubo pri preslikavi v barvni prostor RGB, maksimizirati razliko med odtenki, minimizirati razliko med barvami v izbranem



Slika 3.4: Barvni krog v prostoru CIE $L^*C^*h^\circ$ po enačbi 3.2. Slika je kodirana v barvnem sistemu Adobe RGB.²

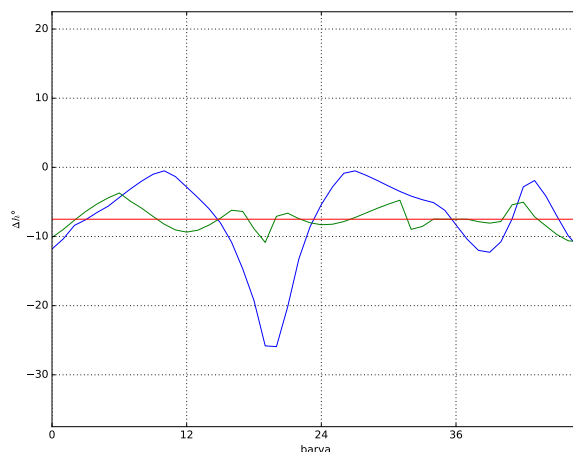
prostoru in prototipičnimi barvami ter minimizirati odstopanje od perceptualne uniformnosti zaradi morebitne deformacije prostora, da bi zadostili ostalim zahtevam.

Ena možna rešitev je, da barvni krog oz. vrednosti (L^* , C^* , h°) definiramo v odvisnosti od parametrov $r \in [0, 1]$ in $\varphi \in [0^\circ, 360^\circ)$:

$$\begin{aligned} L^* &= 85 + r \cdot (24 \cdot \cos(-120^\circ + \varphi) - 12) \\ C^* &= 100 \cdot r \\ h^\circ &= \varphi \end{aligned} \tag{3.2}$$

Zaradi omejitev barvnega razpona prostora (angl. gamut) sRGB, ki nas v zgornjih primerih omejuje predvsem v modro-sinjem delu spektra, bomo uporabili prostor Adobe RGB, ki ima širši barvni razpon, vendar pa je rezultate potrebno reproducirati na zaslonih, ki podpirajo barvni razpon Adobe RGB. Za zaslone, ki podpirajo samo barvni razpon sRGB, se uporabi prilagojena formula, ki dovoljuje večjo perceptualno neuniformnost.

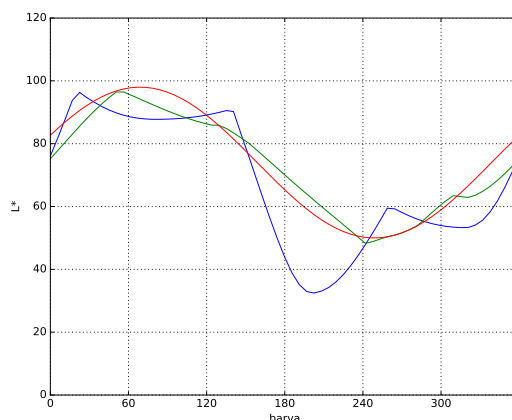
²Barvni prostor Adobe RGB ima večji barvni razpon (angl. gamut) kot prostor sRGB, ki ga uporablja večina monitorjev. Prikaz slike na monitorju, ki omogoča samo prikaz barv barvnega prostora sRGB povzroči, da so barve premalo nasičene.



Slika 3.5: Primerjava spreminjanja vrednosti odtenkov (h°) med 48-imi sosednjimi ekvidistančnimi barvami na obodu kroga CIE $L^*C^*h^\circ$ po enačbi 3.2 (rdeča), po preslikavi kroga v barvni prostor Adobe RGB (zelena) in krogom HSL pri $L_{HSL} = 0,5$, ki je prikazan v prostoru CIE $L^*C^*h^\circ$ (modra).

Slika 3.4 prikazuje barvni krog, ki smo ga dobili z zgoraj definiranimi transformacijami.

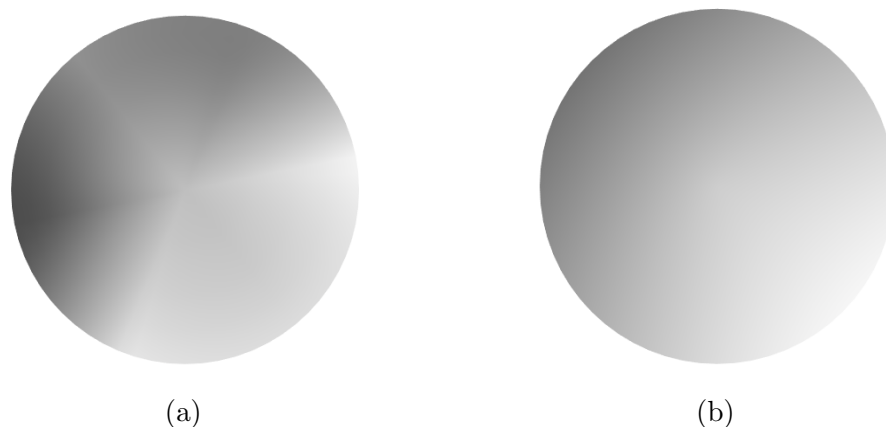
Grafi na slikah 3.5, 3.6 in 3.8 prikazujejo spreminjanje vrednosti Δh° , L^* in C^* v odvisnosti od h° na 48 ekvidistančnih vzorcih z oboda kroga. Vrednost Δh° je razlika med dvema sosednjima vzorcema z oboda. Z rdečo so predstavljene vrednosti na obodu kroga po enačbi 3.2 v barvnem prostoru CIELCh, z zeleno pa vrednosti po preslikavi v barvni prostor Adobe RGB. Z modro so izrisane vrednosti v krogu, ki smo ga definirali v prostoru HSL kot presek pri $L_{HSL} = 0,5$ in ga preslikali v prostor CIELCh. Vrednosti Δh° , L^* in C^* bi bile v primeru barvnega kroga pri $L^* = 72$ v prostoru CIELCh konstantne. Vendar pa pri istem L^* določene barve (npr. rdeča in rumena) ne spadajo v isto barvno kategorijo, kar v našem primeru ni sprejemljivo, zato je krog rotiran relativno na ravnino $L^* = 72$ in raztegnjen v smeri, ki je pravokotna na os rotacije, da barve po rotaciji ostanejo enako nasičene. Posledično imamo namesto konstantne vrednosti L^* na grafu 3.6 sinusoidno krivuljo. Pri



Slika 3.6: Primerjava spreminjanja svetlosti (L^*) sosednjih ekvidistančnih barv na obodu kroga CIE $L^*C^*h^\circ$ po enačbi 3.2 (rdeča), po preslikavi kroga v barvni prostor Adobe RGB (zeleni) in krogom HSL pri $L_{HSL} = 0,5$, ki je prikazan v prostoru CIE $L^*C^*h^\circ$ (modra).

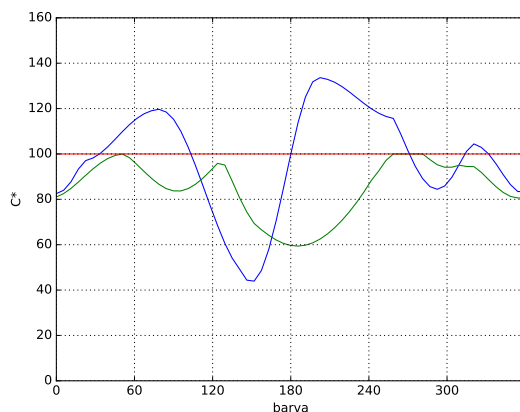
preslikavi v prostor Adobe RGB izgubimo del barvnega prostora CIE LCh , kar je razvidno iz grafov, vendar se krivulja precej boljše ($RMSE = 4,17$) in predvsem bolj enakomerno prilega referenčni rdeči krivulji kot krivulja, ki predstavlja obod kroga iz barvnega sistema HSL ($RMSE = 11,21$). Slika 3.7 prikazuje zgolj komponento L^* obeh krogov (po enačbi 3.2 in HSL), kjer je opazna uniformnost sistema CIE LCh tudi po preslikavi v prostor Adobe RGB po enačbi 3.2 in neuniformnost kroga iz sistema HSL.

Slika 3.9 prikazuje graf razlik ΔE_{00}^* med 48-imi sosednjimi ekvidistančnimi barvami na obodu barvnega kroga (barvni odtenki se med seboj razlikujejo za $\Delta H_{HSL} = 7,5^\circ$ oziroma $\Delta h^* = 7,5^\circ$, začeni z 0°); dodatno je z magento označena meja, kjer je še možna zaznava v razliki odtenka. Pri obodu kroga iz barvnega prostora HSL, ki je na grafu označen z modro, vidimo, da je v določenih predelih barvnega prostora razlika pri spremembi odtenka H za $7,5^\circ$ že nezaznavna (precej manjša od JND), v drugi predelih pa so spremembe odtenka precej bolj opazne. Razlika v perceptualno uniformnem prostoru bi morala biti med sosednjimi odtenki h^* enaka, vendar prostor ni

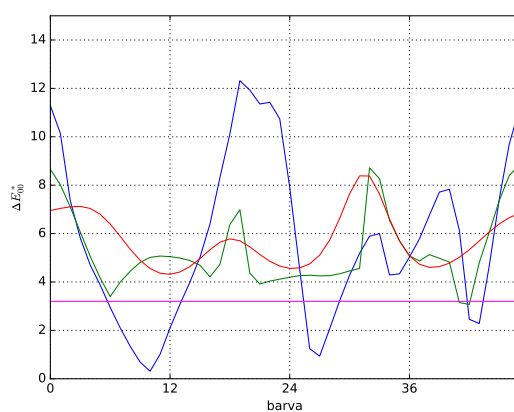


Slika 3.7: (a) Prikaz komponente L^* kroga iz HSL pri $L_{HSL} = 0,5$, prikazanega v prostoru CIE $L^*C^*h^\circ$ in (b) komponente L^* kroga po enačbi 3.2 po preslikavi v barvni prostor Adobe RGB, prikazanega v prostoru CIE $L^*C^*h^\circ$.

popolnoma uniformen (ravno zato je bilo določenih več popravkov izračuna razlike med barvami), je pa razpon nihanja razlik omejen. Razlike po preslikavi prostor Adobe RGB so sprejemljive (kvadratni koren povprečne kvadratne napake oddaljenosti od rdeče obarvane referenčne krivulje je $RMSE = 1,37$), večji skok je le pri modrih odtenkih kroga. Modro-zeleni del spektra je problematičen predvsem zaradi oblike barvnega prostora RGB, ki je pri sinji barvi precej ožji kot v drugih delih barvnega prostora CIELCh. Na obodu kroga lahko torej ločimo približno 48 odtenkov, prav tako vsebuje barve, ki so blizu prototipičnim predstavnikom barvnih kategorij. Barvne kategorije služijo za identifikacijo regije kroga, medtem ko vmesni odtenki določajo spremembe in odstopanje vizualiziranih vrednosti znotraj posamezne regije, kjer je predvsem pomembno prikazovanje spremembe in smeri spremembe.



Slika 3.8: Primerjava spreminjanja nasičenosti (C^*) sosednjih ekvidistančnih barv na obodu kroga CIE L*C*h° po enačbi 3.2 (rdeča), po preslikavi kroga v barvni prostor Adobe RGB (zeleni) in krogom HSL pri $L_{HSL} = 0,5$, ki je prikazan v prostoru CIE L*C*h° (modra).



Slika 3.9: Primerjava spreminjanja razdalje ΔE_{00}^* med 48-imi sosednjimi ekvidistančnimi barvami na obodu kroga CIE L*C*h° po enačbi 3.2 (rdeča), po preslikavi kroga v barvni prostor Adobe RGB (zeleni) in krogom HSL pri $L_{HSL} = 0,5$, ki je prikazan v prostoru CIE L*C*h° (modra). Z magento je označena meja, kjer je še možna zaznava v razliki odtenka.

3.2 Tonski prostor in preslikava

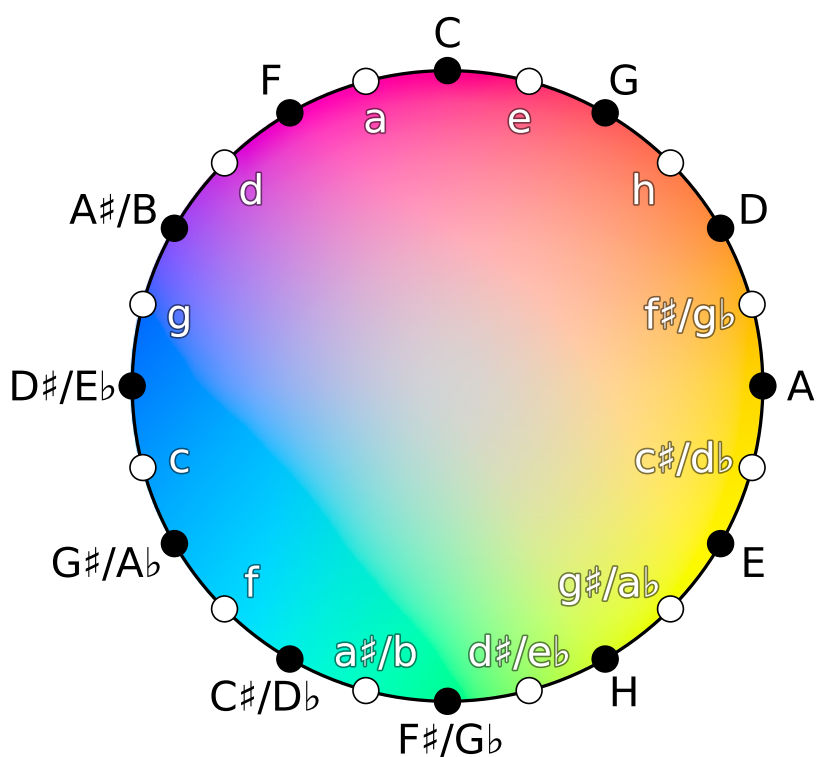
Predlagan model preslikave iz slušnega v vidni prostor poskuša izkoristiti princip, da se podobnosti v enem sistemu slikajo, kolikor je možno, v podobnostih v drugem sistemu. Možnih preslikav iz slušnega prostora v vidni prostor je veliko, npr. lahko bi se preslikala višina tonov v prostorsko dimenzijo vida v višinski osi tako, da bi bili visoki toni nižje, višji toni pa nižje na tej osi. Vendar pa takšna vizualizacija ni najbolj intuitivna, ker se mora uporabnik naučiti, kako deluje preslikava, oziroma, da preslikava deluje ravno nasprotno, kot smo navajeni, torej, da se visoko v enem prostoru preslika v visoko v drugem ter da je smer premikanja navzdol in navzgor v obeh sistemih obrnjena v isto smer. Cilj je, da pri preslikavi pazimo, da ni takšnih interferenc in da je posledično krivulja učenja, tam kjer je potrebno, čim manjša.

Jedro preslikave v vizualni prostor je preslikava harmoničnih odnosov med toni. Zaradi cikličnosti percepcijskega tonskega in tonalitetnega prostora, lahko predstavimo tonski prostor kot krožnico, krog, ploskev, (hiper)torus, oziroma eno-, dvo- ali večdimenzionalno, odvisno od tega, koliko relacij iz tonskega prostora smo pripravljene poenostaviti oziroma zanemariti pri poenostavljanju modela. Več kot tridimenzionalni modeli so pri vizualizaciji problematični zaradi neintuitivnosti. Na drugi strani pa je model z enodimenzionalno krožnico tako poenostavljen, da ne zajame dovolj relacij iz tonskega prostora in s tem je tudi omejena njegova uporabna vrednost. Še ena možnost preslikave v vizualni prostor je preslikava v barvni prostor. Barvni prostor lahko opišemo na različne načine in posamezno barvo lahko obravnavamo kot kombinacijo komponent. Nekaj barvnih sistemov smo predstavili v razdelku 3.1.

Vzemimo kvintni krog pri enakomerno temperirani uglasitvi, kjer so tonski razredi razporejeni ekvidistančno z razliko sosednjih tonov $\Delta\varphi_{k,k+1} = \frac{\pi}{6}$ na krožnici z $r_{ton} = 1$. Kot za posamezen tonski razred k dobimo $\varphi(k) = \frac{k\pi}{6}$, kjer je $k \in \{0, 1, 2, \dots, 10, 11\}$; $k = 0$ predstavlja ton C, $k = 1$ predstavlja ton G itn. v zaporedju kvint. Posamezen ton predstavimo kot točko na krožnici pri ustreznem tonskem razredu. Tonski prostor kvintnega kroga

preslikamo v barvni prostor, tako da v enačbi 3.2 za r vzamemo polmer kvintnega kroga $r_{ton} = 1$, za φ pa $\pm\varphi_{ton} + \varphi_{zamik}$, pri čemer je φ_{ton} kot, ki je prirejen tonskemu razredu. Smer tonov v kvintnem krogu je poljubna, zato lahko vzamemo pozitivne ali negativne vrednosti kotov $\pm\varphi_{ton}$, poleg tega je lahko krog poljubno rotiran, kar predstavlja φ_{zamik} . Barve v tem začetnem prirejanju nimajo psihološkega pomena, ampak predstavljajo način, kako prikazati razlike med toni, tonskimi razredi, sočasnimi toni, zato je začetno prirejanje poljubno. V klavirski implementaciji smo uporabili $\varphi_{ton} + \frac{\pi}{12}$, tako da se psihološko tople barve pokrivajo z belimi tipkami, hladne pa s črnimi. Slika 3.10 prikazuje barvni krog prirejen kvintnemu krogu (toni kvintnega kroga so označeni s črnimi točkami, kvintni in barvni krog sta v prikazu obrnjena in rotirana, da je ton C na vrhu in gredo kvinte v smeri urinega kazalca).

Kadar nek ton zveni, ga lahko na podlagi njegove pozicije v kvintnem krogu preslikamo v barvni prostor. Glede na to, da je v kvintnem krogu 12 tonskih razredov, imamo lahko sočasno preslikanih do 12 barv. Vendar želimo dimenzionalnost zmanjšati. Barvni krog omogoča, da predstavimo precej več vrednosti, kot je tonskih razredov v kvintnem krogu, zato bomo tone, ki so si med seboj podobni poenostavili tako, da bomo njihov prikaz združili. Najlažje združujemo tone, ki so si med seboj podobni. V kvintnem krogu sosednost predstavljata tonska razreda, ki sta najbolj sorodna opazovanemu tonskemu razredu (razen tonskega razreda samega). To je posledica ujemanja frekvenčnih komponent sočasno zvenečih tonov (več v razdelku 2.3). Kvintni krog je tudi interpretacija prve dimenzije modela štiridimenzionalnega modela, ki sta ga s pomočjo MDS dobila Krumhansl in Kessler [33] in tudi MVU [4] nad istimi podatki. Glede na to, da je v kvintnem krogu definirano 12 točk, v barvnem krogu pa je na obodu kroga neskončno možnih vrednosti, lahko izkoristimo prostor med preslikanimi točkami za prikaz kombinacij tonov. V praksi je sicer število točk na obodu kroga omejeno z razliko med sosednjimi točkami, le-ta mora biti večja od JND, ki pa je odvisna od tega, kako definiramo krog. Pri krogu, ki smo ga definirali v razdelku 3.1 imamo



Slika 3.10: Kvintno-terčni krog s prirejenim barvnim krogom. S črnimi črkami na zunanji strani kroga so označeni toni kvintnega kroga, z belimi črkami na notranji strani kroga pa so označene terce. Slika je kodirana v barvnem sistemu Adobe RGB.

v praksi na obodu več kot 48 vrednosti. Odtенок kombiniranega tona, ki je sestavljen iz dveh sosednjih tonov v kvintnem krogu lahko tako predstavimo kot odtenek $\bar{\varphi}$, ki je povprečje odtenkov obeh tonov (φ_1 in φ_2). V primeru, da je eden od tonov tišji od drugega, se vektorja proporcionalno utežita z utežima w_1 in w_2 :

$$\bar{\varphi} = \arg \left(\frac{w_1 e^{i\varphi_1} + w_2 e^{i\varphi_2}}{w_1 + w_2} \right) \quad (3.3)$$

Povprečni kot lahko izračunamo tudi s funkcijo $\arctan2$, katere izvajanje se je v knjižnicah na računalniških arhitekturah, na katerih smo implementirali model, izkazalo za hitrejše:

$$\bar{\varphi} = \arctan 2 \left(\frac{w_1 \sin \varphi_1 + w_2 \sin \varphi_2}{w_1 + w_2}, \frac{w_1 \cos \varphi_1 + w_2 \cos \varphi_2}{w_1 + w_2} \right) \quad (3.4)$$

Če vzamemo tona, ki sta oddaljena veliko sekundo (sosed soseda v kvintnem krogu), povprečje kotov sovpada s tonom, ki je med njima, kar bi v vizualizaciji pomenilo, da sta si ton G in kombinacija tonov C in D precej perceptualno podobna. Vendar pa je sočasna kombinacija dveh tonov, ki sta oddaljena veliko sekundo dokaj disonantna. Za predstavitev disonance lahko izkoristimo notranjost kroga, ki smo ga definirali v razdelku 3.1, kjer je središče kroga barva z vrednostjo krome $C = 0$, torej popolnoma nenasičena barva, obod kroga pa predstavljajo nasičene barve. Disonanco tako predstavimo z manj nasičenimi odtenki, konsonanco pa z nasičenimi. Kot smo omenili v razdelku 2.3 lahko disonanco obravnavamo v ožjem pomenu kot senzorično disonanco, lahko pa tudi kot glasbeno disonanco, ki predstavlja nestabilno stanje v danem kontekstu. Osnovna različica predstavljenega modela preslikave kot vhod pričakuje kromatični vektor dvanajstih tonskih razredov, kar pomeni, da se toni, ki so v istem tonskem razredu, vendar v različnih oktavah predstavijo enotno z eno komponento vektorja. Pri sestavljanju kromatičnega vektorja se v osnovni različici modela preslikave v posameznem tonskem razredu upošteva najglasnejši ton. Glasnosti tonov v vseh tonskih razredih se normalizirajo; najglasnejši ima utež 1, ostali glede na glasnost proporcionalno manjšo. Zaradi odsotnosti podatkov o oktavah tonov se v osnovni

različici modela ne uporablja senzorična disonanca. Štiridimenzionalni model Krumhansl in Kessler je zgrajen na podlagi testov konsonance in disonance oz. ujemanja tonov, ki sledijo posameznemu tonu ali več tonom. Točke v modelu so razdeljene tako, da so toni in tonalitete, ki so bolj oddaljeni, manj sorodni in zvenijo bolj disonantno kot tisti, ki so v modelu blizu skupaj. Ta lastnost velja tudi za kvintni krog, ki je prva dimenzija tega modela (po MDS [33] in tudi MVU [4]). Razdaljo med točkami v kvintnem krogu tako uporabimo za mero konsonance oz. disonance. Tona oddaljena za veliko sekundo sta v krogu bolj oddaljena kot tona, ki sta oddaljena za kvinto ali kvarto. Večja kot je razdalja med tonoma, manjša bo nasičenost barve, ki predstavlja kombinacijo. Disonanca kombinacije je odvisna tudi od glasnosti posameznega tona, zato pri izračunu disonance upoštevamo tudi uteži obeh tonov. Ton v kvintnem krogu smo predstavili kot točko, oziroma krajevni vektor, ki ga lahko obravnavamo tudi kot smeri vektor. Podobno, kot smo povprečen kót izračunali kot povprečje smeri vektorjev, disonanco izračunamo kot povprečno dolžino smernega vektorja:

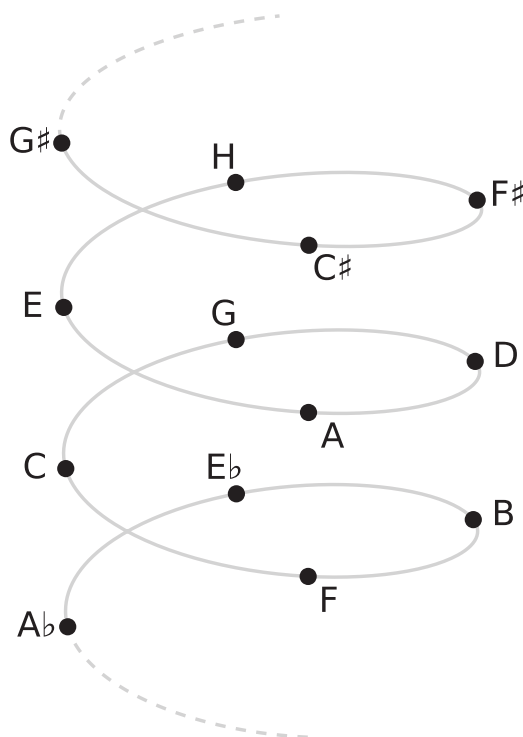
$$\bar{R} = \left| \frac{w_1 e^{i\varphi_1} + w_2 e^{i\varphi_2}}{w_1 + w_2} \right| \quad (3.5)$$

oziroma

$$\bar{R} = \sqrt{\left(\frac{w_1 \sin\varphi_1 + w_2 \sin\varphi_2}{w_1 + w_2} \right)^2 + \left(\frac{w_1 \cos\varphi_1 + w_2 \cos\varphi_2}{w_1 + w_2} \right)^2} \quad (3.6)$$

Daljši smerni vektor pomeni večjo konsonanco, krajši pa manjšo konsonanco oz. disonanco. Pri tonih, ki so oddaljeni veliko sekundo, je tako smerni vektor krajši in predstavlja bolj disonantno percepcijo kombinacije teh dveh tonov. Podoben primer je tudi pri tonih, ki so si nasproti v kvintnem krogu – tritonus, kjer je dolžina smernega vektorja pri enako uteženih tonih 0, njegova smer pa je v tem primeru nepomembna.

Problem se pojavi pri kombinacijah tonov, ki so oddaljeni malo ali veliko terco. V obeh primerih je dolžina smernega vektorja krajša kot pri intervalu velike sekunde, kar ni v skladu z rezultati percepcijskih poskusov (rezultati



Slika 3.11: Izsek spiralnega modela tonskega prostora. Kvinte ležijo na vijačnici, velika terca je en navoj višje ali nižje od posameznega tona.

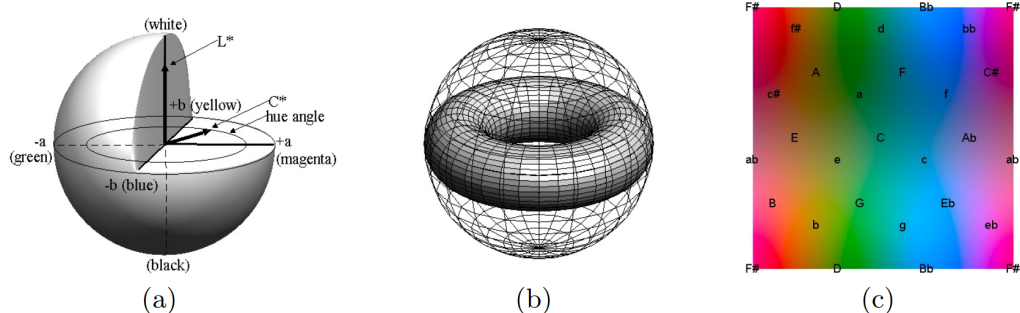
pecepcijskih poskusov so podrobneje predstavljeni v [33]). Glede na pogostost malih in velikih terc v glasbi je potrebno to percepcijo v vizualizaciji ustrezneje predstaviti. Kombinacije tonov, ki so oddaljeni malo ali veliko terco, so bolj konsonantne od kombinacij tonov, ki so oddaljeni veliko sekundo, ampak manj od takšnih, kjer so toni oddaljeni kvinto ali kvarto. Med geometrijskimi modeli tonskega prostora takšno relacijo zajemajo na primer tonske mreže (nem. Tonnetz), spiralni in toroidni modeli [58].

Slika 3.11 prikazuje spiralni oz. vijačnični model tonskega prostora, v različici, ki ga je predlagala Chew [8]. V modelu ležijo toni oz. tonski razredi v kvintnem zaporedju na vijačnici. Oddaljenosti med tonskimi razredi so razdalje v evklidskem prostoru med točkami, ki predstavljajo tonske razrede. Posamezen tonski razred je predstavljen kot točka v kartezičnem

koordinatnem sistemu, kjer so koordinate točke, ki predstavlja k -ti zaporedni element kvintnega zaporedja ($k \in \mathbb{Z}$), definirane kot:

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \begin{bmatrix} r \sin \frac{k\pi}{2} \\ r \cos \frac{k\pi}{2} \\ kh \end{bmatrix} \quad (3.7)$$

Da bi bile oddaljenosti med toni, kvintami, kvartami, velikimi in malimi tercami ter sekundami ustrezno urejene, je potrebno izbrati primerne vrednosti parametrov r in h . Pogojem zadostijo vsi r in h , kjer je $\sqrt{\frac{2}{15}} < \frac{h}{r} < \sqrt{\frac{2}{7}}$ [8]. Iz slike 3.11 je razvidno, da sta tona, ki sta v odnosu velike terce (npr. C in E) ali pa tona, ki sta v odnosu male terce (npr. C in A) v prostoru precej bližje, če vzamemo evklidsko razdaljo med njima, kot če gremo po vijačnici. Tona, ki sta oddaljena za veliko terco, imata v spiralnem modelu isti koordinati x in y , razlikujeta se v višini, oziroma koordinati z . Če imamo enakomerno temperirano uglasitev, se vijačnica po treh zavojih oz. dvanaajstih tonih vrne nazaj v isti tonski razred. Tako lahko os z spiralnega modela sklenemo v krog in dobimo torus (slika 3.13). Vendar je potrebno paziti, saj gre še vedno za dvodimenzionalni ciklični model – tonski prostor je definiran samo na površini torusa. Toni C, E in Gis, ki so na sliki 3.13 prikazani na od središča oddaljeni strani torusa, so od „središča“ enako oddaljeni kot toni G, Fis in B, ki so prikazani na središču bližnji strani cevi torusa. Torus lahko v kartezičnih koordinatah predstavimo v štirih dimenzijah, razdalje med toni pa so razdalje med točkami v štiridimenzionalnem prostoru. Za primer vzemimo 4 točke, ki predstavljajo tone T_C , T_E , T_G in T_{Fis} , v torusu so njihove koordinate oblike $(x, y, z, w) = (R \cdot \cos\varphi, R \cdot \sin\varphi, r \cdot \cos\theta, r \cdot \sin\theta)$, kjer je R polmer kvintnega kroga, φ kót tona v kvintnem krogu, r je polmer cevi zvitega spiralnega modela, θ pa je kót tona v spiralnem modelu. Če vzamemo v spiralnem modelu $h = 5$, potem lahko zaradi prej omenjene omejitve izberemo $r = 10$, za veliki polmer pa lahko izberemo $R = 10$ (zgoraj napisana omejitev razmerja $\frac{h}{r}$ sedaj ne velja več, ker smo valj sklenili in s tem tudi spremenili razdalje). V podanem primeru je evklidska razdalja med točkama

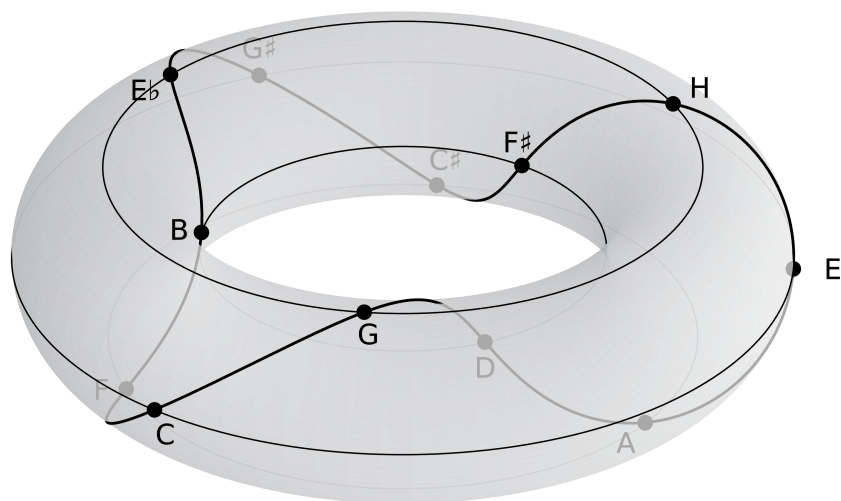


Slika 3.12: Preslikava toroidnega modela tonalitetnega prostora v barvni prostor CIE $L^*a^*b^*$, ki sta ga predlagala Martorell in Gómez [45]. Model pri preslikavi obravnava torus kot tridimenzionalno telo in ne kot štiridimenzionalno, posledično so razlike v barvah, glede na razlike v tonalitetah, popačene.

$d(T_C, T_E) = 17,321$, prav tako je razdalja $d(T_G, T_{Fis}) = 17,321$. Martorell in Gómez sta predlagala model vizualizacije, ki toroidni model tonalitetnega prostora postavi neposredno v barvni prostor CIE $L^*a^*b^*$ [45], vendar pa ima model problem ravno v tem, da torus obravnava kot tridimenzionalno telo in ne kot štiridimenzionalno, posledično so nekatere tonalitete predstavljene z barvami, ki so bližje skupaj ali pa bolj oddaljene, čeprav so v prostoru tonalitet enako oddaljene 3.12. Drug problem modela je, da prirejene barve niso najbolj intuitivne.

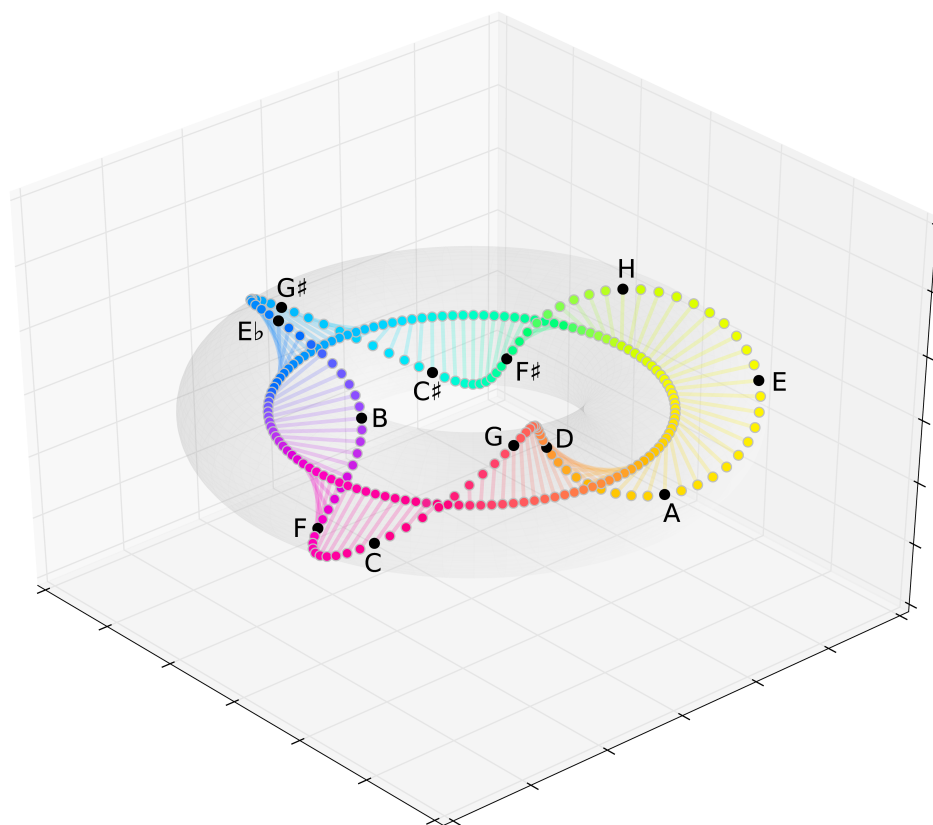
Če apliciramo torus, ki ga definira enačba 3.7 na naš model, se os z toroidnega modela pokrije s krožnico kvintnega kroga, pri čemer je potrebno upoštevati prej omenjeno lastnost, da so točke na površini cevi torusa enako oddaljene od središča torusa. Posledično posamezni točki na vijačnici priredimo barvo, ki ustreza barvi koordinate z , kar prikazuje slika 3.14.

V spiralnem modelu razdalja med točkami določa sorodnost oz. disonanco in konsonanco. Tritonske akorde pa lahko predstavimo z uteženo konveksno kombinacijo tonov, ki sestavljajo akord [8]. Podobno lahko predstavimo tudi ostale kombinacije tonov. Če se najprej omejimo na dva tona: vzemimo ton in njegovo kvinto; v vijačnici sta tona na sosednjih pozicijah. Če vzamemo povprečje obeh tonov z enakima utežema, dobimo pozicijo, ki leži



Slika 3.13: Predstavitev tonskega prostora v obliki torusa. Kvintna vijačnica je prikazana z odebeljeno črto.

blizu vijačnice in katere koordinata z je povprečje z -koordinat obeh tonov. Glede na prej definirano preslikavo v barvni prostor rezultat obarvamo z identičnim odtenkom kot v preprostejšem kvintnem modelu. Če pa vzamemo na primer veliko terco, dobimo točko, ki ima isto x in y -koordinato kot tona, ki sestavljata zven, višina z pa je povprečje koordinat z obeh točk. Po prej opredeljenem prirejanju, kjer lahko na podlagi koordinate z določimo odtenek, v tem primeru dobimo odtenek, ki je identičen odtenku, ki ga dobimo pri kombinaciji tona z njegovo veliko sekundo. Kar je isti problem, ki smo ga imeli pri preprostejšem kvintnem krogu, le da je sedaj barva kombinacije tona in njegove velike terce bolj nasičena, saj je razdalja med njima v spiralnem modelu manjša. Vendar pa iz percepcijskega vidika ta rezultat ni v redu, saj se kombinaciji precej bolj razlikujeta in glede na zahtevo, da podobnosti vizualiziramo s podobnostmi v barvnem prostoru, bi morala biti tudi razlika v odtenku. Ker je direktna pot od tona do njegove velike terce v spiralnem modelu bližnjica v primerjavi s potjo po vijačnici, bi morali ta hitrejši prehod



Slika 3.14: Vizualizacija združitve barvnega kroga in torusa tonskega prostora. Povezava med točkama za osi z in vijačnici, ki imata isto vrednost koordinate z , je prikazana s črto zaradi lažje predstave.

ustrezno vizualizirati. Ker smo prelivanje iz odtenka v odtenek že uporabili za zlivanje tonov na vijačnici, je ena možnost, da hitrejšo prelivanje med dvema barvama vizualiziramo s kombinacijo obeh barv v ustreznih razmerjih. S tem kombinacija tonov ni več prikazana zgolj z eno barvo. Vendar pa zaradi zahteve po čim večji enostavnosti modela in implementacije na klaviaturi, ki zaradi tehničnih omejitev zahteva eno barvo na kombinacijo tonov, želimo vseeno, da se rezultatom kombinacije tonov priredi samo ena barva. Dodatna dimenzionalna redukcija v tem primeru pomeni, da bomo pri vizualizaciji ne bomo zajeli določenih lastnosti toroidnega modela oziroma tonskega percepcijskega prostora, ki ga predstavlja.

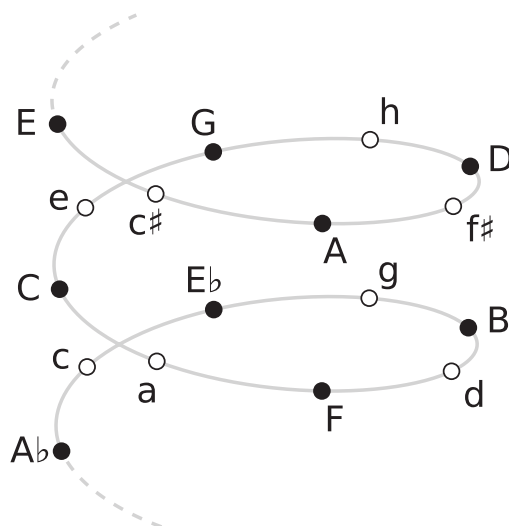
Gatzsche s kolegi je predlagal dekompozicijo štiridimenzionalnega modela Krumhansl-Kessler oziroma podobnih toroidnih modelov v dva kroga, kjer prvi primarno kaže relacije med tonalitetami, drugi pa znotraj tonalitete [18]. Prvi krog je baziran na modelu „terčne“ oz. terčno-kvintne vijačnice (slika 3.15), ki se lahko projicira na krožnico in tako dobimo terčno-kvintni krog, kot je na sliki 3.10.

Vzemimo spiralni model s kvintno vijačnico in definirajmo spiralni model s terčno vijačnico. Točke terčne vijačnice so definirane podobno kot točke kvintne vijačnice:

$$\begin{bmatrix} x'_m \\ y'_m \\ z'_m \end{bmatrix} = \begin{bmatrix} r \sin \frac{m\pi}{2} \\ r \cos \frac{m\pi}{2} \\ mh \end{bmatrix} \quad (3.8)$$

Razlika je v m , ki je zavzema vrednosti iz $M = \{ \dots -1, -0.5, 0, 0.5, 1, 1.5, 2 \dots \}$. Celoštevilske vrednosti m predstavljajo tone kvintnega kroga, točke $m = k + 0.5$ pa predstavljajo ton, ki je oddaljen veliko terco navzgor od tona k (oziroma malo terco navzdol od tona $k + 1$). Če za $m = 0$ vzamemo tonski razred C, potem $m = 0.5$ predstavlja e. Tonske razrede pri neceloštevilskih m bomo pisali z mali črkami, da jih ločujemo od točk kvintnega kroga. V modelih, v katerih so predstavljena medtonaliteta razmerja, pa bodo male črke predstavljale molovske tonalitete.

Tone iz kvintnega spiralnega modela bomo preslikali v terčno-kvintni



Slika 3.15: Dva zavoja kvintno-terčne vijačnice. Kvintni vijačnici smo dodali terce med sosednje pare kvint.

spiralni model, pri čemer se vsak ton kvintnega spiralnega modela preslika v oba istoimenska tona terčnega spiralnega modela. Ker je vijačnica del torusa, se bomo pri spiralnem modelu omejili na 3 zavoje vijačnice, oziroma na 12 tonskih razredov, zato bomo v spodnjih enačbah privzeli $k \bmod 12$ in $m \bmod 12$ za vsak k in m in za uteži vzeli maksimalne uteži posameznih tonskih razredov.

Za vsak ton k kvintnega kroga izračunamo utež $w_{vpliv}(k)$:

$$w_{vpliv}(k) = \max(w(k-3), w(k-4)) \quad (3.9)$$

Nato izračunamo $w'(m)$ za vsak ton m terčne vijačnice:

$$w'(m) = \begin{cases} w(m) \cdot (1 - w_{vpliv}(m)), & \text{če je } m \text{ celo število} \\ w(m+3 \cdot 5) \cdot w_{vpliv}(m+3 \cdot 5) & \text{drugače.} \end{cases} \quad (3.10)$$

Oziroma drugače povedano, porazdelimo utež tona k iz kvintne vijačnice na istoimenska tona m na terčni vijačnici.

Na koncu iz uteži tonov w' in kotov $\varphi(m) = \frac{m\pi}{6}$ izračunamo $\bar{\varphi}$ in \bar{R} , ki ju uporabimo za preslikavo v barvni prostor.

$$\bar{\varphi} = \arctan 2 \left(\frac{\sum_{m \in M} w'_m \sin \varphi_m}{\sum_{m \in M} w'_m}, \frac{\sum_{m \in M} w'_m \cos \varphi_m}{\sum_{m \in M} w'_m} \right) \quad (3.11)$$

$$\bar{R} = \sqrt{\left(\frac{\sum_{m \in M} w'_m \sin \varphi_m}{\sum_{m \in M} w'_m} \right)^2 + \left(\frac{\sum_{m \in M} w'_m \cos \varphi_m}{\sum_{m \in M} w'_m} \right)^2} \quad (3.12)$$

Opisana preslikava velja za kombinacijo sočasno zvenceh tonov v izolaciji. Vendar pa toni večino časa zvenijo v kontekstu drugih tonov. Že ko zaigramo prvi ton ali kombinacijo tonov, s tem začnemo ustvarjati kontekst. Kontekst lahko gledamo na različnih nivojih, lahko obravnavamo posamezne odseke skladbe, potujemo z drsečim oknom ali gledamo tonaliteto celotne skladbe. Kontekst predstavimo podobno, kot predstavljamo posamezne tone s smernim vektorjem. Na primer, ko določimo tonaliteto, nastavimo smer vektorja glede na tonaliteto različico toroidnega modela (torus medtonalitetnih razmerij predstavljen v razdelku 2.4), kjer vzamemo točko, ki predstavlja tonaliteto (točka leži v območju med dvema zavojema vijačnice) in jo projiciramo na vijačnico v središče območja, ki ga določajo trije toni durovega ali molovskega kvintakorda tonalitete. Vektor ustrezno utežimo in obravnavamo z vsemi ostalimi vektorji z eno razliko, vektor konteksta povzroči prestavljanje tonov v kvintno-terčnem modelu, vendar pa ostali toni ne povzročijo prestavljanja kontekstnega vektorja. Tako dobimo barvo tona v danem kontekstu tonalitete. Nato novo zaigrani ton prispeva k spreminjanju konteksta in se s tem upošteva pri naslednjem tonu.

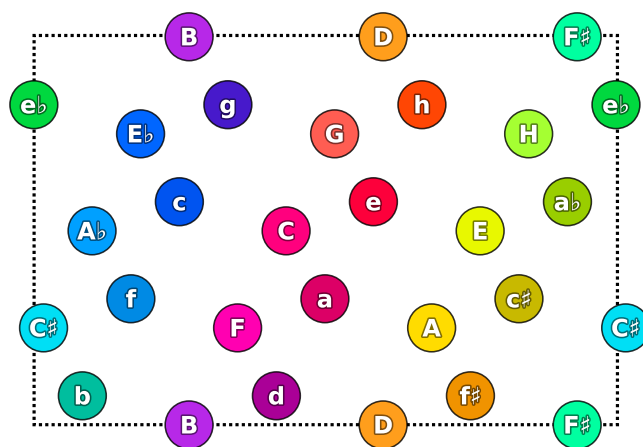
3.3 Rezultati in razprava

Po definirani preslikavi v barvni prostor preverimo, če je preslikava smiselna:

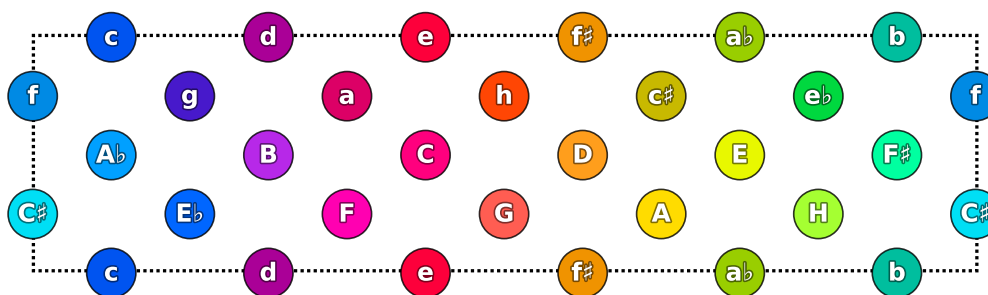
- v prostoru sorodnosti tonov, akordov ali tonalitet se morajo sorodnosti kazati s podobnimi barvami;
- razlika med toni od najbolj podobnega do najmanj podobnega mora biti v barvnem prostoru urejena enako, kot v slušnem prostoru glede na psihoakustične eksperimente;
- barve, ki predstavljajo sosednje kombinacije tonov, se morajo med seboj dovolj razlikovati, da je možna diferenciacija.

Slika 3.16 prikazuje prostor regij tonalitet v dveh dimenzijah (razprta površina torusa v ravnini). Razdalje med tonalitetami so predstavljene z evklidskimi razdaljami v ravnini. Prikazani sta dve različici torusa, ki ju dobimo z analizo podatkov poskusov Krumhansl-Kessler [33]. Slika 3.16a prikazuje površino torusa, ki ga dobimo z večrazsežnostnim lestvičenjem (MDS). Če za primer vzamemo C-dur, vidimo, da so perceptualno sosednje tonalitete, G-dur, F-dur, a-mol barvane s podobno barvo kot C-dur, vendar pa je večja razlika med C-durom in paralelnim molom c-mol. Razlika je posledica poenostavitve, kjer smo kombinacijo tonov vedno želeli predstaviti z eno barvo. Tako smo v spiralnem modelu projicirali rezultate, ki so bili med dvema zavojema vijačnice, na vijačnico. Ena možna rešitev problema je, da dovolimo barvanje z več barvami, oziroma takšne prehode obravnavamo kot posebno pravilo pri vizualizaciji, se pa takšni prehodi v povprečju vseeno ne pojavljajo tako zelo pogosto, da bi ovirali uporabnost modela. Slika 3.16b prikazuje površino torusa, ki ga dobimo z razvijanjem z maksimalno varianco (MVU) [4]. Metoda MVU boljše ohranja lokalno sosesko posameznih točk. V tem modelu relacija paralelnih durov in molov tako ni zajeta s sosednostjo na površini torusa, posledično tudi ni večjih diskontinuitet v barvi.

Druga lastnost, ki ji mora zadoščati model je, da so razdalje med toni pravilno urejene – v tem primeru obravnavamo razdalje med toni po poskusih



(a)



(b)

Slika 3.16: Razdalje med regijami tonaliteta na podlagi poskusov Krumhansl-Kessler [33], prikazane na raztegnjeni površini torusa, ki ga dobimo (a) z večrazsežnostnim lestvičenjem (MDS); in torusa, ki ga dobimo (b) z razvijanjem z maksimalno varianco (MVU) [4]. Črtkani pravokotnik označuje območje, kjer „zlepimo“ robove pravokotnika, da dobimo površino torusa. Robeve lepimo po istoimenskih regijah, ki ležijo na robu pravokotnika. Regijam v obeh prostorih smo priredili barve po modelu, predlaganem v razdelku 3.2. Pri (b) diskontinuitet v barvi na površini torusa ni, pri (a) pa je opazna diskontinuiteta med paralelnimi durmi in moli. Razdalje v modelih niso v istem merilu.

Preglednica 3.4: Razdalje ΔE_{00}^* med preslikanimi barvami posameznih tonov v kontekstu C-dura. Zadnji stolpec prikazuje povprečno oceno ujemanja posameznega tona v kontekstu C-dura po rezultatih poskusov Krumhanslove [34]. V poskusih se je točkovalo ujemanje z več točkami, neujemanje z manj, zato so rezultati v zadnjem stolpcu odšteti od maksimalnega ratinga.

ton	ΔE_{00}^* (1)	ΔE_{00}^* (2)	rating K-K
C	2,2	2,2	0,0
G	7,9	7,9	1,2
D	19,2	19,2	2,8
A	6,7	18,8	2,7
E	2,7	13,2	2,0
H	45,0	45,0	3,6
F \sharp /G \flat	39,4	39,4	3,7
C \sharp /D \flat	38,8	38,8	4,2
G \sharp /A \flat	34,3	34,3	4,0
D \sharp /E \flat	26,2	26,2	3,9
A \sharp /B	18,4	18,4	4,1
F	10,9	10,9	2,3

Krumhansl-Kessler kot ordinalno skalo, saj geometrijski modeli, na podlagi katerih smo izpeljali model preslikave barv, primarno obravnavajo razdalje med toni tako, da uporabljajo ordinalno skalo. Preglednica 3.16 prikazuje razdaljo preslikane barve ΔE_{00}^* (1) za posamezne tone v kontekstu C-dura in barvo C-dura. Problematična je razdalja med osnovno barvo C-dura in tonoma E ter A napram tonu G in F. Na podlagi perceptualnih poskusov bi ton E moral biti bolj oddaljen kot ton G, A pa bolj oddaljen kot tona G in F. To je posledica preslikave tonov v modelu terčne vijačnice, kjer je terca med obema kvintama (E med C in G). Brez preslikave je razdalja med C in E v primerjavi s C in G večja. Problem lahko rešimo z dodatnim „kaznovanjem“ preslikave terc. Ena možnost je, da se zmanjša utež tona, ki se je preslikal, kar bo vplivalo na komponento C^* pri preslikavi v barvni prostor. Druga možnost je, da izkoristimo barvni prostor nad ali pod ravnino barvnega kroga. Psihološko gledano se razlika med durom in molom pogosto opisuje kot razlika med svetlejšim, veselejšim durom in temnejšim, otožnejšim molom³. Preslikave tonov v terčnem krogu so bolj številčne pri molovih akordih kot pri durovih, iz tega razloga bi lahko ob preslikavi dodali utež, na podlagi katere bi preslikani ton predstavili negativno v smeri osi L^* . Dodajmo utež z_k , ki jo določimo:

$$z_m = \begin{cases} 0, & \text{če je } m \text{ celo število} \\ Z & \text{drugače.} \end{cases} \quad (3.13)$$

Povprečno utež \bar{z} izračunamo:

$$\bar{z} = \frac{\sum_{m \in M} w'_m z_m}{\sum_{m \in M} w'_m} \quad (3.14)$$

In dopolnimo L^* v enačbi 3.2:

$$L^* = 85 + r \cdot (24 \cdot \cos(-120 + \varphi) - 12) - r \cdot z \quad (3.15)$$

³Obstajajo pa tudi primeri vizualizacij, ki uporabljajo obratno prirejanje: temnejši duri, svetlejši moli, npr. [44].

Če izberemo za konstanto $Z = 20$, dobimo razdalje ΔE_{00}^* (2) v preglednici 3.4. S popravkom sta se razdalji med E in C-durom ter A in C-durom povečali in ustrezata profilu perceptualnih poskusov. Če pogledamo profile tonov v kontekstih drugih tonalitet, se vrstni red konsonantnih kombinacij ohranja, izjema je vrstni red kvarte in male terce, kjer se v nekaterih tonalitetah v vizualizaciji zamenja. Vzrok za variacijo je nepopolna uniformnost barvnega prostora CIE $L^*a^*b^*$ in zaradi rotacije barvnega kroga glede na ravnino $L^* = 50$. Vendar je v tem primeru razdalja posledica oddaljenosti tonov od barve tonalitete v dveh različnih dimenzijah: kvarta v odtenku, terca v svetlosti, zato menjava ni tako problematična. Zamenja se tudi vrstni red nekaterih disonantnih kombinacij, vendar so le-te manj nasičene in so razdalje med njimi precej manjše, posledično se razlike vizualno ne opazijo. Zaradi neuniformnosti in rotacije barvnega kroga razdalje tonov v tonalitetah variirajo v povprečju za 16,7 %, vendar je nihanje po zaporednih tonalitetah ciklično zaradi rotacije barvnega kroga – v tonalitetah v delih kroga z maksimalnim in minimalnim L^* so razdalje večje, v vmesnih manjše.

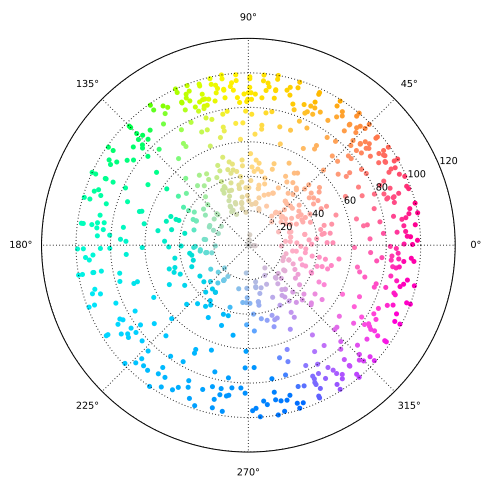
Slika 3.17 prikazuje sliko barvnega prostora s preslikavo posamičnih tonov in vseh kombinacij dveh, treh in štirih sočasnih tonov v barvni prostor (pogoste akorde in tudi precej manj pogostih akordov). Različica 1 je brez dodatnih uteži pri prestavljanju tonov v modelu terčne vijačnice, različica 2 prikazuje preslikave z utežmi. Najbolj disonantne kombinacije tonov so v središču (npr. tritonus, zvečani kvintakord, zmanjšani septakord z zmanjšano septimo), kar pomeni, da vizualno med njimi ne ločimo (razen, ko so zaigrani v kontekstu). Pri ostalih tonskih kombinacijah je pri različici 2 najbližji sosed v povprečju oddaljen $\overline{\Delta E_{00}^*} = 1,8$, s povprečno absolutno deviacijo od povprečja 0,7. Pri različici 1 je $\overline{\Delta E_{00}^*} = 1,3$. Model po zasnovi najbolj diferencira konsonantne kombinacije tonov, medtem ko so disonantne predstavljene z nenasičenimi barvami in se perceptualno tudi manj razlikujejo med seboj (se pa precej razlikujejo od konsonantnih kombinacij). Disonantne kombinacije so vse bolj pogoste, ko dodamo še več tonov. Ker se predstavitev določenih kombinacij, kot so tritonusi (kadar so zaigrani brez konteksta), preslikajo v isto točko,

zgubimo informacijo o tem, v katero smer lahko razrešimo disonanco. Podobno kot na kombinacije tonov brez konteksta, kombinacije tonov v kontekstu, kontekst (akordi, tonalitete) lahko gledamo na ta problem kot prikaz nivoja podrobnosti. Na strukturo lahko gledamo hierarhično, kjer se posamezni deli na nižjih nivojih sestavljajo v strukture na višjih nivojih. Če beležimo postopek združevanja na posameznih nivojih, lahko omogočimo v nekem trenutku vpogled, kateri deli so v kakšni meri prispevali do agregiranega rezultata na višjem nivoju.

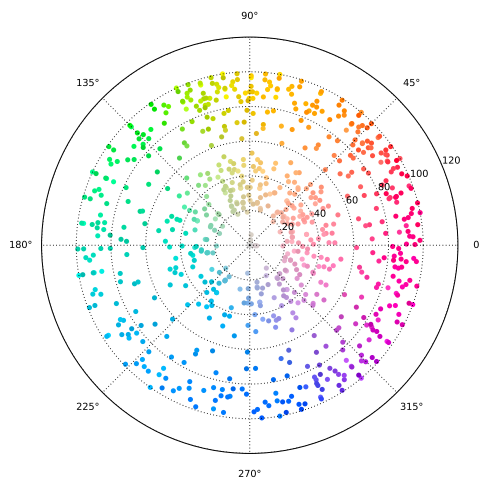
Model preslikave izkorišča barvni del človeške percepcije vida, zato se lahko vključi v obstoječe vizualizacije, ki uporabljajo ostale dimenzije, oziroma lahko tudi nadomesti barvno dimenzijo, če jo je v vizualizaciji možno menjati. Na primer, Sappov model prikaza hierarhične strukture tonalitete skladbe [60] že uporablja barvanje tonalitete po modelu preslikave kvintnega kroga v barvni krog, kjer molovske tonalitete pobarva s temnejšimi različicami durovih barv. V področjih, kjer so prehodi med regijami tonalitete, se lahko aplicira predlagan model, ki omogoča zvezno barvanje; prav tako se lahko uporabi tudi model z dvema barvama. Bergstromov model vizualizacije vzorcev akordov na tonski mreži (nem. Tonnetz) primarno ne uporablja barv. Ker model bazira na tonski mreži, ki je, če predpostavimo enakomerno temperirano uglasitev, predstavitev torusa v dvodimenzionalni ravnini s ponavljanjem, je aplikacija predlaganega modela precej neposredna. Modeli, ki omogočajo preiskovanje multimedijskih podatkovnih zbirk z netekstovnimi, kot na primer [35, 66], iz podatkovnih baz na podlagi podobnosti vizualno prikazujejo skladbe. Predlagana vizualizacija se lahko uporabi kot hitri globalni pregled spreminjanja harmonij v skladbi. Dokaj pogost način planarne in prostorske vizualizacije je izpeljava iz prikaza klavirskega avtomata, npr. [27, 14]; Malinowski⁴ v vizualizacijah med drugim uporablja model prirejanja barv tonom na podlagi zaporedja v kvintnem krogu; integracija modela z združevanjem tonov je tudi tukaj precej neposredna. Zato smo vizualizacijo na podlagi klavirskega avtomata uporabili kot osnovo za aplikacijo, s katero smo testirali model. Aplikacija je predstavljena v

⁴<http://www.musanim.com/>

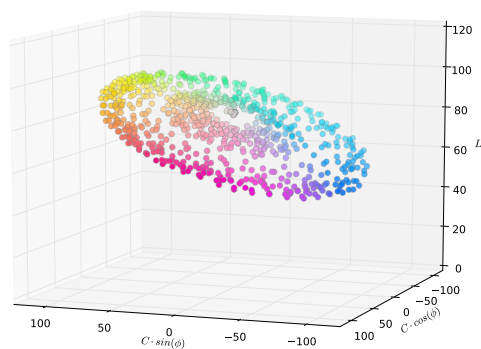
naslednjem poglavju. Še ena aplikacija modela je uporaba na klasičnih in modernih glasbilih, prvi prototip aplikacije na klaviaturi bomo prav tako predstavili v naslednjem poglavju.



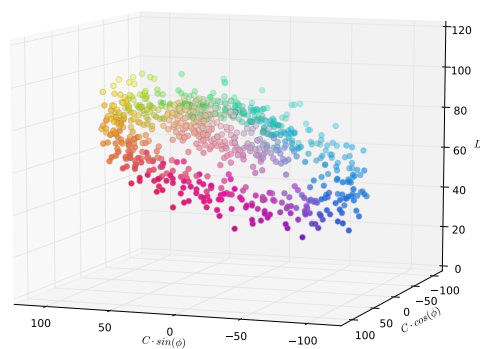
Različica 1 (od zgoraj)



Različica 2 (od zgoraj)



Različica 1 (s strani)



Različica 2 (s strani)

Slika 3.17: Prikaz porazdelitev preslikave posamičnih tonov in vseh kombinacij dveh, treh in štirih sočasnih tonov v barvni prostor. Različica 2 je model z dodatnimi utežmi pri prestavljanju tonov v modelu terčne vijačnice; različica 1 je brez uteži.

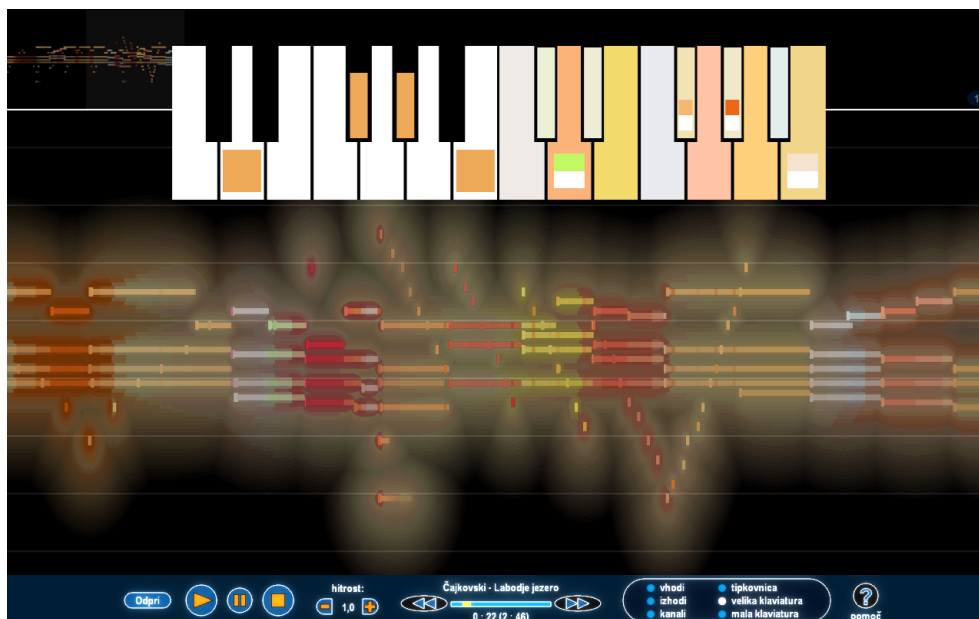
Poglavje 4

Implementacija

Za demonstracijo in testiranje modela preslikave harmoničnih odnosov med toni v barvni prostor smo implementirali aplikacijo za vizualizacijo glasbe, ki nadgrajuje vizualizacijo klavirskega avtomata z barvnim modelom (slika 4.1). V aplikaciji je možen prikaz klaviature, ki je nadgrajena z modelom barvanja harmonij. Za testiranje na fizičnem glasbilu smo predelali električno klaviaturo, podrobnosti predelave so predstavljene v razdelku 4.1.

Implementacija modela kot vhodni podatek ne pričakuje zvočnega zapisa, ampak že izluščene posamezne tone iz zvočnega zapisa, ki so zapisani v formatu MIDI. Za skladbe v zvočnem zapisu je v takšnem primeru potrebno uporabiti program, ki predprocesira zvok v MIDI-signal. Na podlagi amplitudne ovojnice posameznih glasbil in signalov `NoteOn` in `NoteOff` se izračuna trenutna glasnost tona. Poleg signalov za začetek in konec tona se upoštevajo tudi signali za podaljšano izzvenevanje `Sustain pedal`, signali za spreminjanje frekvence `Pitch bend` in še nekaj kontrolnih signalov. Vsak osveževalni cikel (v aplikaciji na osebem računalniku tipično 60-krat na sekundo) se na podlagi trenutno zvenečih tonov in zgodovine tonov (konteksta) izvrši izračun po modelu, ki je predstavljen v poglavju 3.2. Procesiranje tako poteka v živo, da je mogoče tudi igranje na klaviaturo in vizualizacija v živo.

Prikaz glasbe v obliki zapisa za klavirski avtomat preslika tone v dvodimenzionalni prostor. Os x predstavlja čas (ki narašča v desno), os y predstavlja



Slika 4.1: Zaslonski posnetek demonstracijske aplikacije, ki uporablja predlagani model preslikave harmoničnih odnosov med toni v barvni prostor. Na sliki je prikazana integracija barvnega modela v klaviaturo in nadgrajen zapis klavirskega avtomata s prikazom glasbenega konteksta.

višino tonov (višji toni so višje, nižji nižje, skala je linearna glede na percepcijo višine tonov). Ton je predstavljen kot pravokotnik s središčem na ustrezni višini, z dolžino, ki je proporcionalna trajanju tona in širino, da lahko tone pobarvamo. Ozadje izrisa je enobarvno (črno ali belo), toni pa so na podlago izrisani z različnimi stopnjami prosojnosti, so proporcionalno odvisne od glasnosti tona v danem trenutku. Kadar zveni več sočasnih tonov z isto višino, se vzame ton z največjo glasnostjo. Ozadje tonov se barva glede na barvo, ki ustreza trenutnemu kontekstu (tonaliteti).

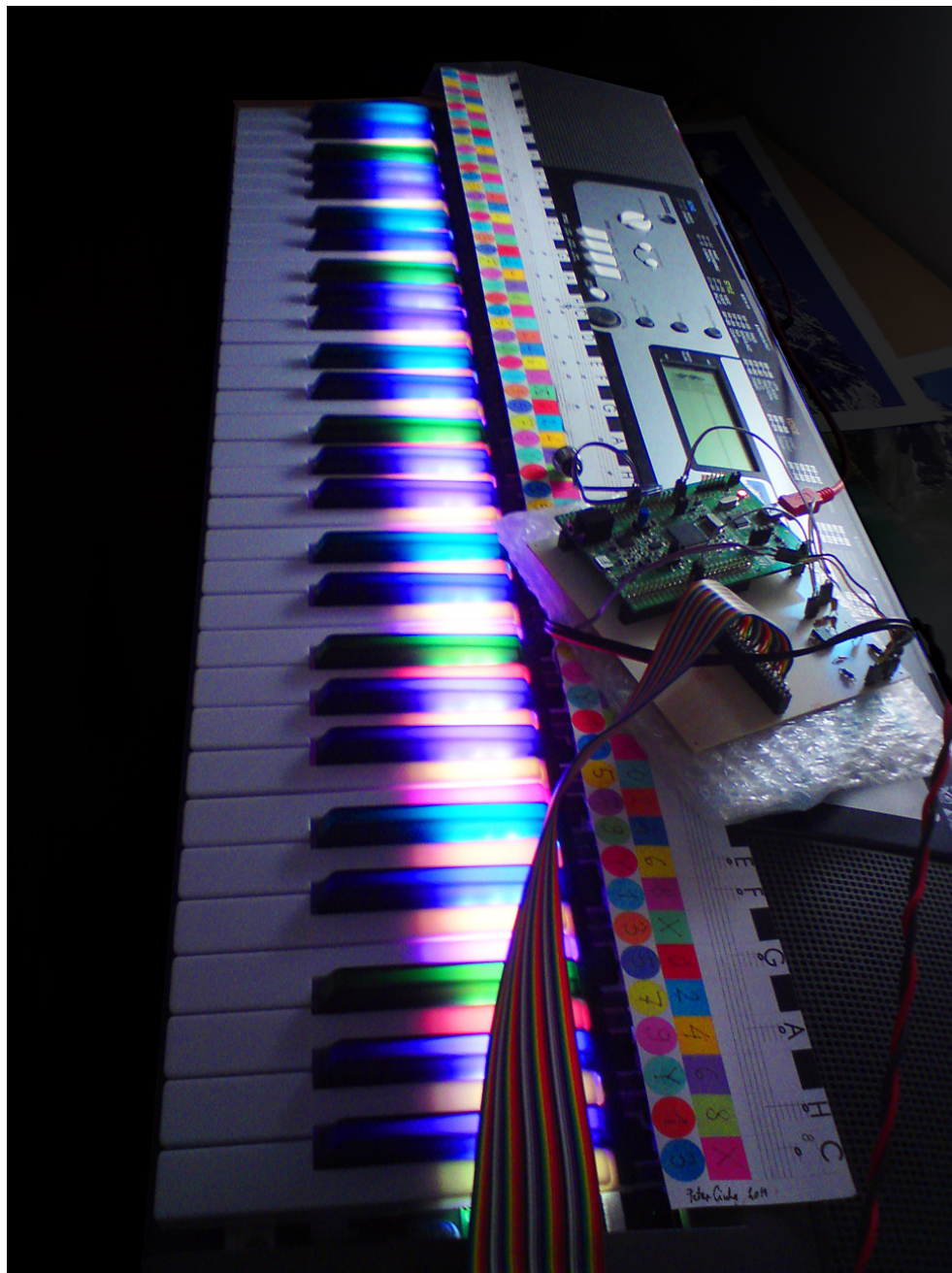
Drugi prikaz modela preslikave v aplikaciji je na klaviaturi, kjer so posamezne tipke pobarvane z barvami, ki jih dobimo, če pritisnemo posamezno tipko. Nato se barve tipk ustrezno prilagajajo glede na trenutno aktivne tipke oziroma zaigrane harmonije. Slabost klaviature na zaslonu je, da ni taktilnega odziva, zato smo model prenesli tudi na fizično klaviaturo [73].

4.1 Prototip fizične barvne klaviature

Za testiranje modela preslikave na fizičnem glasbilu smo prilagodili električno klaviaturo Yamaha EZ-200. Ta klaviatura ima tipke izdelane iz plastike, ki zadovoljivo prepušča svetlobo različnih valovnih dolžin, saj ima pod vsako tipko rdečo LED diodo, ki je namenjena osvetljevanju tipke (za vodeno igranje). V klaviaturo smo pod vsako tipko vgradili tribarvno (RGB) svetlečo diodo in mikrokrmilnik za obdelavo zajetih podatkov o tonih in krmiljenje svetlečih diod. Zajem, obdelava in krmiljenje morajo potekati v realnem času s čim manjšo latenco, zato smo uporabili mikrokrmilnik z enoto za računanje v plavajoči vejici. Krmilnik s pomočjo signalov preko MIDI-vmesnika zazna, katere tipke so pritisnjene in izračuna ustrezno barvo za posamezno tipko. Pri računanju barve smo se omejili na eno oktavo, oz. na dvanajst tipk. Tako lahko skupaj krmilimo vse svetleče diode pod tipkami iz istega tonskega razreda. Za implementacijo smo vzeli razvojno ploščico STM32FDISCOVERY z vgrajenim procesorjem ARM Cortex-M4, ki teče na frekvenci do 168 MHz in vsebuje enoto za operacije v plavajoči vejici. Program, ki teče na krmilniku, smo razdelili na 3 komponente: komponento za zajem signala, komponento za obdelavo podatkov in komponento za krmiljenje svetlečih diod.

Pri zajemu signala uporabljamo izhod MIDI, ki je standardno vgrajen v elektronske klaviature. Zanimajo nas predvsem signali `NoteOn`, `NoteOff` in krmilni signali kot so `Sustain pedal`, `All controllers off`, `All notes off`. Signale MIDI iz izhoda MIDI-out klaviature pripeljemo na pin Rx mikrokrmilnika preko optosklopnika (tako sta vezji električno ločeni).

Komponenta za obdelavo podatkov na podlagi zajetih MIDI-signalov in shranjenega stanja na vsakih 20 ms izračuna barve posameznih tonskih razredov. Za izračun barve se uporablja prilagojen algoritem, ki je predstavljen v poglavju 3. Pri MIDI-ju prejemo v najbolj osnovni različici bolj ali manj samo dogodke, ko uporabnik pritisne oz. spusti tipko, pritisne na pedal ali zamenja glasbilo, nato pa na primer sintetizator zvoka poskrbi za ustrezno predvajanje zvoka. Zato si mora komponenta za obdelavo podatkov zapomniti, kateri toni so bili pritisnjeni in na podlagi tega, kako se spreminja



Slika 4.2: Električna klaviatura Yamaha EZ-200 z vgrajenimi svetlečimi diodami, povezana na mikrokontrolnik STM32F407VGT6.

amplitudna ovojnica tona (in posledično percepcija tona) pri posameznem glasbilu, ustrezno utežiti tone pri izračunu barve. Poleg tega je potrebno upoštevati kratkoročni človeški spomin, ki povzroči, da že izzveneli toni tvorijo kontekst, v katerem interpretiramo nove tone, kar modeliramo s tem, da se izzveneli toni ustrezno uteženo upoštevajo pri izračunu.

Da bi dobili čim boljše kombinacije barv in zaradi pomanjkanja prostora pod tipkami, pridejo v poštev predvsem tribarvne (RGB) svetleče diode SMD (angl. surface mount device). Svetleča dioda RGB SMD je sestavljena iz treh diod (rdeče, zelene in modre) v enem ohišju, vsako pa lahko krmilimo neodvisno od drugih dveh. Tako lahko s kombiniranjem dobimo 7 različnih barv, oziroma še „črno“, ki jo predstavljajo ugasnjene diode. Ker potrebujemo več kombinacij, želimo da posamezne diode svetijo z različno jakostjo – to naredimo s pomočjo sekvence različno dolgih pulzov oz. modulacije širine pulza (angl. PWM). Komponenta za krmiljenje tako pretvori vrednost RGB za posamezen tonski razred v ustrezen signal PWM. Poleg tega je potrebno upoštevati tudi, da je svetlobni tok večji, če gori več diod hkrati, kar pomeni, da je potrebno pri določenih kombinacijah skrajšati dolžino pulza, da niso barve svetlejše od tistih, kjer je na primer aktivna samo eno svetlobna dioda. Za krmiljenje svetlečih diod potrebujemo 36 vhodno/izhodnih pinov, oziroma 12-krat po tri za vsako svetlečo diodo. Pod vsako tipko so tri RGB-svetleče diode – več jih zaradi fizičnih omejitev nismo mogli vgraditi – posledično so bele tipke osvetljene samo v zgornji polovici. Ker so svetleče diode pod belimi in črnimi tipkami, barva tipke vpliva na količino svetlobe posamezne valovne dolžine. Zato smo predhodno s pomočjo spektrometra umerili ustrezen tok za svetilne diode pod belimi in črnimi tipkami tako, da so perceptualno približno iste jakosti in barve za belo svetlobo.

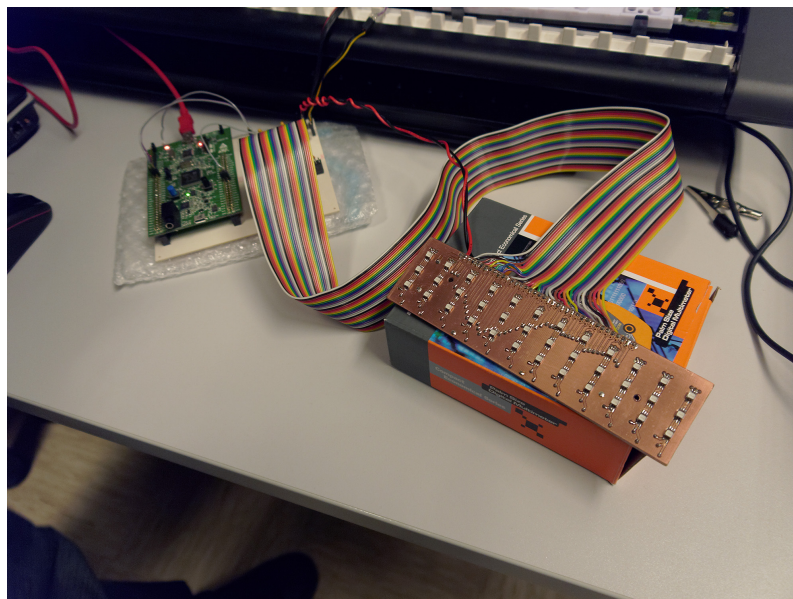
V primerjavi s predhodno programsko izvedbo klaviature je bila fizična klaviatura bolj „prijazna“ za testiranje zaradi taktilne povratne informacije. Ob pritisku na posamezno tipko se tipka pobarva z barvo določenega tona, vse ostale nepritisnjene tipke pa z barvo, ki predstavlja rezultat, če bi poleg pritisnjene tipke pritisnili tudi kakšno drugo tipko. Če pritisnemo dve tipki

oz. več tipk, se vse pobarvajo z barvo, ki predstavlja tonsko kombinacijo, preostale tipke pa se pobarvajo z barvo, ki bi bila rezultat pritiska še določene posamezne tipke ob vseh obstoječih pritisnjenih tipkah. Glede na vrsto glasbila in amplitudno ovojnico izzvenevanja se barve posodabljaajo z intervalom 20 ms.

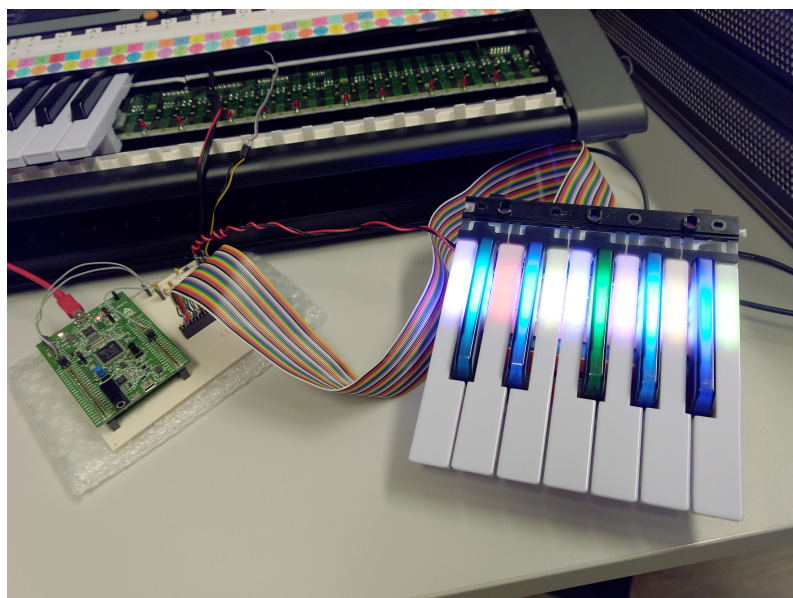
Pri testiranju klaviature se je pokazalo nekaj pomanjkljivosti: predvsem je začetna poenostavitev omejitve barvanja na eno oktavo precej omejujoča, ker ne pokaže razlike med toni v nižjih in višjih oktavah, ki še niso pritisnjeni, če nekje zveni en ton iz istega tonskega razreda. Poleg tega ne omogoča, da pokažemo kakšen bi bil rezultat, če bi posamezno tipko, ki je že pritisnjena, spustili. Dogodek spusta tipke ni tako pomemben kot pritisk tipke, ki je perceptualno precej izstopajoč dogodek, vendar pri tonih, ki ne izzvenevajo oz. izzvenevajo počasi vpliva na občutek konsonance in disonance.

Čeprav je s pomočjo umerjanja možno prilagajati barve referenčnim odtenkom, material tipk vpliva na obseg barvnega prostora, ki ga lahko prikažemo, kar pomeni, da barve polnih nasičenosti niso tako nasičene, kot bi morale biti v referenčnem barvnem modelu.

Pri neformalnem testiranju z uporabniki se je klaviatura izkazala zanimiva predvsem za začetnike, ki pri igranju gledajo precej na tipkovnico. Poleg potrebne nadgradnje, kjer bi se vse svetleče diode za vsako tipko krmililo ločeno, bi bilo potrebno dodati zgolj manjši zaslon (ali serijo diod), ki bi prikazovala trenutno barvo (namesto na pritisnjenih tipkah), tako da lahko na pritisnjenih tipkah prikažemo rezultate ob izzvenenju tona. Posamezna tipka vsebuje več svetlečih diod, tako bi lahko vsako tipko obarvali z dvema barvama, kot je predstavljeno v poglavju 3. Preizkusi s statičnim modelom dveh barv sicer kažejo, da uporabnik ne sme biti preobremenjen s količino informacij in lahko dodatne barve povzročijo prenasičenost, zato bi bilo potrebno preizkusiti selektivno dodajanje druge barve na tipke.



Slika 4.3: Mikrokontroler STM32F407VGT6 povezan s ploščo svetlečih SMD-diod.



Slika 4.4: Mikrokontroler STM32F407VGT6 povezan s ploščo svetlečih SMD-diod pod klaviaturo.

Poglavje 5

Zaključki in nadaljnje delo

V magistrski nalogi sem predstavil nov model vizualne predstavitve harmoničnih odnosov v glasbi, ki temelji na percepciji sorodnosti tonov oz. ideji perceptualnega združevanja sočasnih tonov. Na podlagi toroidnega modela tonskega prostora, oziroma toroidnega modela prostora medtonalitetnih razmerij, ki je osnovan na podlagi perceptualnih poskusov na eni strani in izpeljan iz glasbene teorije na drugi strani, slikamo tone v barvni prostor. Barvni prostor mora biti čim bolj uniformen, da lahko posamezne dimenzije iz tonskega prostora slikamo v druge dimenzije barvnega prostora. Vizualizacije pogosto uporabljajo barvni sistem RGB ali enega od opisov sistema RGB v cilindričnem koordinatnem sistemu, kot je na primer HSL. V razdelku 3.1 smo si pogledali, zakaj za naš primer vizualizacije tak sistem ni primeren. Za opis barvnega prostora smo uporabili sistem CIE $L^*a^*b^*$, ki je perceptualno dokaj uniformen, oziroma smo uporabili predstavitev prostora v cilindričnih koordinatah CIE $L^*C^*h^\circ$, ker je bolj intuitiven za navigacijo po komponentah. Za izhodišče preslikave smo vzeli preslikavo kvintnega kroga v barvni krog, ki je definiran kot krog znotraj prostora CIE $L^*C^*h^\circ$. Pri tem smo poskrbeli, da krog čim bolj zajame glavne barvne kategorije, ki se uporabljajo v naravnem jeziku, hkrati pa ostanejo razlike med „sosednjimi“ barvami čim bolj enakomerne. Poleg tega maksimiziramo ploščino kroga, da imamo čim večji prostor, kamor lahko slikamo iz tonskega prostora. Rezultate vizua-

lizacije prikažemo na računalniškem zaslonu, kjer se srečamo s problemom omejenega obsega barvnih prostorov RGB, predvsem prostora sRGB, ki se trenutno še vedno uporablja pri večini zaslonov. Manjši zaslonski barvni prostor predstavlja dodatno omejitev velikosti ciljnega barvnega prostora preslikave. Problem lahko rešujemo na dva načina: zmanjšamo barvni prostor preslikave (in zmanjšamo učinkovitost vizualizacije) ali pa prikazujemo na zaslonu z večjim barvnim razponom. Med zaslone z večjim barvnim prostorom se dokaj pogosto uporabljajo zaslone, ki podpirajo barvni prostor Adobe RGB. V razdelku 3.1 smo predstavili kompromisno rešitev, ki uporabi barvni sistem Adobe RGB, a še vedno izgubi nekaj barv pri preslikavi iz CIE L*C*h°. Za prikaz celotnega barvnega prostora vizualizacije bi morali vzeti zaslone z večjimi barvnimi razponi (npr. „Wide Gamut RGB“), temu primerno je prilagojena tudi demonstracijska aplikacija, ki že prikazuje barve glede na naložen barvni profil zaslona.

Poleg programa za vizualizacije harmoničnih odnosov v glasbi z barvo, ki nadgradi prikaz klavirskega avtomata z barvami (barva kombinacije tonov, barva kombinacije tonov v kontekstu in barva konteksta), smo implementirali model preslikave tonov v barve tudi na fizični digitalni klaviaturi, kjer so posamezne tipke osvetljene z svetlečimi-RGB diodami pod tipkami. Zaradi tehničnih omejitev smo vse tipke istega tonskega razreda krmilili skupaj, kar je se izkazalo za preveč omejujoče, saj želimo diferencirati med tipkami, ki so pritisnjene, in tipkami, ki niso. Ker so osvetljene tako bele kot črne tipke, smo svetleče diode umerili na belo barvo s pomočjo spektrometra. Drug problem je delna svetlobna propustnost črnih in belih tipk, ki povzroči, da barve niso tako nasičene, kot bi morale biti. V naslednjih različicah bi se tako lahko poskusil tudi projekcijski tip osvetljevanja tipk, ki je tudi primeren za vgradnjo na klaviature, ki nimajo prosojnih tipk. Učinkovitost osvetljevanja tipk z barvami, ki so odvisne od harmoničnih odnosov med toni, kot pripomoček za učenje, bi bilo potrebno evalvirati na večji skupini uporabnikov.

Osnovne barve tonov so prirejene na podlagi kvintnega kroga, kjer je posamezen ton predstavljen kot krajevni vektor pozicije v kvintnem krogu.

Ker kvintni krog ne zajema relacij med toni kot so terce, smo za predstavitev tonov uporabili toroidni model, v katerem je relacija kvint predstavljena s kvintno vijačnico. Model smo uporabili za predstavitev odnosov med toni in tonalitetami, pred preslikavo v vizualni prostor pa smo ga poenostavili – tone, ki so bili v terčnem razmerju, smo predstavili skupaj na terčno-kvintno vijačnico, da smo jih lahko predstavili samo z eno barvo. S tem smo izgubili del informacij, na primer odnos med paralelnim durom in molom. Poleg tega smo pokvarili vrstni red hierarhije tonov, kjer sta po psihoakustičnih poskusih tona pri intervalu kvinte bolj konsonantna kot tona pri intervalu terce. Problem smo rešili s tem, da smo prestavljanje terc dodatno „kaznovali“ z utežmi. Preden smo kombinacijo tonov lahko pretvorili v barvo, smo kombinacijo predstavili s smernim vektorjem, katerega smer in dolžina je uteženo povprečje smeri in dolžin sestavnih vektorjev. Predstavitev z eno barvo je omejujoča, zato bomo v nadaljevanju razvijali tudi metodo, ki omogoča barvanje z dvema barvama. Prav tako se iz disonantnih kombinacij ne vidi vedno, kako se jih lahko razreši. Ta problem lahko rešimo s pomočjo hierarhičnega prikazovanja različnih nivojev z različno stopnjo podrobnosti (manj ali več barvnih komponent).

Predlagani barvni model se lahko uporabi v drugih obstoječih sistemih vizualizacije, pri čemer lahko dopolnjuje vizualizacije, ki barve dimenzije še ne pokrivajo in pri katerih je prikaz harmoničnih odnosov smiseln. Lahko pa tudi nadomestijo obstoječo barvno shemo vizualizacij. Nekaj primerov smo podali v razdelku 3.3. Nadgradnjo modela za prikaz kombinacije tonov z dvema barvama in opisani model z eno barvo bi bilo potrebno integrirati v vizualizacije, da se lahko evalvira njihova učinkovitost.

Dodatek A

Ostale uporabljene formule

A.1 Izračun ΔE_{00}^*

Razlika ΔE_{00}^* je definirana v prostoru CIE L*a*b*.

$$\Delta E_{00}^* = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T \frac{\Delta C'}{k_C S_C} \frac{\Delta H'}{k_H S_H}}$$

$$\Delta L' = L_2^* - L_1^*$$

$$\bar{L} = \frac{L_1^* + L_2^*}{2} \quad \bar{C} = \frac{C_1^* + C_2^*}{2}$$

$$a'_1 = a_1^* + \frac{a_1^*}{2} \left(1 - \sqrt{\frac{\bar{C}^7}{\bar{C}^7 + 25^7}}\right) \quad a'_2 = a_2^* + \frac{a_2^*}{2} \left(1 - \sqrt{\frac{\bar{C}^7}{\bar{C}^7 + 25^7}}\right)$$

$$\bar{C}' = \frac{C'_1 + C'_2}{2} \text{ in } \Delta C' = C'_2 - C'_1$$

kjer je

$$C'_1 = \sqrt{a_1'^2 + b_1^{*2}} \quad C'_2 = \sqrt{a_2'^2 + b_2^{*2}}$$

$$h'_1 = \text{atan2}(b_1^*, a'_1) \text{ mod } 360^\circ, \quad h'_2 = \text{atan2}(b_2^*, a'_2) \text{ mod } 360^\circ$$

$$\Delta h' = \begin{cases} h'_2 - h'_1 & |h'_1 - h'_2| \leq 180^\circ \\ h'_2 - h'_1 + 360^\circ & |h'_1 - h'_2| > 180^\circ, h'_2 \leq h'_1 \\ h'_2 - h'_1 - 360^\circ & |h'_1 - h'_2| > 180^\circ, h'_2 > h'_1 \end{cases}$$

$$\Delta H' = 2\sqrt{C'_1 C'_2} \sin(\Delta h'/2)$$

$$\bar{H}' = \begin{cases} (h'_1 + h'_2)/2 & |h'_1 - h'_2| \leq 180^\circ \\ (h'_1 + h'_2 + 360^\circ)/2 & |h'_1 - h'_2| > 180^\circ, h'_1 + h'_2 < 360^\circ \\ (h'_1 + h'_2 - 360^\circ)/2 & |h'_1 - h'_2| > 180^\circ, h'_1 + h'_2 \geq 360^\circ \end{cases}$$

$$T = 1 - 0,17 \cos(\bar{H}' - 30^\circ) + 0,24 \cos(2\bar{H}') + \\ + 0,32 \cos(3\bar{H}' + 6^\circ) - 0,20 \cos(4\bar{H}' - 63^\circ)$$

$$S_L = 1 + \frac{0,015(\bar{L}-50)^2}{\sqrt{20+(\bar{L}-50)^2}} \quad S_C = 1 + 0,045 \cdot \bar{C}' \quad S_H = 1 + 0,015 \cdot \bar{C}''T$$

$$R_T = -2\sqrt{\frac{\bar{C}'^7}{\bar{C}'^7+25^7}} \sin \left[60^\circ \cdot e \left(-\left[\frac{\bar{H}'-275^\circ}{25^\circ} \right]^2 \right) \right]$$

A.2 Pretvorbe med CIE L*C*h° in CIE L*a*b*

Pretvorba iz CIE L*C*h° v CIE L*a*b*

Komponenta L^* ostane enaka.

$$a^* = C^* \cdot \cos(h^\circ)$$

$$b^* = C^* \cdot \sin(h^\circ)$$

Pretvorba iz CIE L*a*b* v CIE L*C*h°

Komponenta L^* ostane enaka.

$$C^* = \sqrt{a^{*2} + b^{*2}}$$

$$h^\circ = \arctan\left(\frac{b^*}{a^*}\right)$$

A.3 Pretvorbe med CIE L*a*b* in CIE XYZ

Pri pretvorbi se privzame bela referenčna točka pri standardnem iluminantu D65.

Pretvorba iz CIE L*a*b* v CIE XYZ

$$X = X_n f^{-1}\left(\frac{1}{116}(L^* + 16) + \frac{1}{500}a^*\right)$$

$$Y = Y_n f^{-1}\left(\frac{1}{116}(L^* + 16)\right)$$

$$Z = Z_n f^{-1}\left(\frac{1}{116}(L^* + 16) - \frac{1}{200}b^*\right)$$

kjer

$$f^{-1} = \begin{cases} t^3 & \text{če } t > \delta \\ 3\delta^2\left(t - \frac{4}{29}\right) & \text{drugače} \end{cases}$$

kjer je $\delta = \frac{6}{29}$

Pretvorba iz CIE XYZ v CIE L*a*b*

$$L^* = 116f(Y/Y_n) - 16$$

$$a^* = 500 [f(X/X_n) - f(Y/Y_n)]$$

$$b^* = 200 [f(Y/Y_n) - f(Z/Z_n)]$$

$$\text{kjer je } f(t) = \begin{cases} t^{1/3} & \text{if } t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29} & \text{otherwise} \end{cases}$$

Za X_n , Y_n , Z_n vzamemo

$$X_n = 95,047; Y_n = 100; Z_n = 108,883$$

A.4 Pretvorbe med CIE XYZ in sRGB

Pri pretvorbi se privzame bela referenčna točka pri standardnem iluminantu D65.

Pretvorba iz CIE XYZ v sRGB

$$\begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} = \begin{bmatrix} 3,2406 & -1,5372 & -0,4986 \\ -0,9689 & 1,8758 & 0,0415 \\ 0,0557 & -0,2040 & 1,0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

C v C_{sRGB} in C_{linear} nadomestimo z R , G in B za izračun posamezne komponente:

$$C_{sRGB} = \begin{cases} 12,92 \cdot C_{linear} & C_{linear} \leq 0,0031308 \\ (1+a)C_{linear}^{\frac{1}{2,4}} - a & C_{linear} > 0,0031308 \end{cases}$$

kjer je $a = 0,055$

Vrednosti R , G in B so na intervalu $[0, 1]$.

Pretvorba iz sRGB v CIE XYZ

C v C_{linear} in C_{sRGB} nadomestimo z R , G in B za izračun posamezne komponente. Vhodne vrednosti R , G in B so na intervalu $[0, 1]$.

$$C_{linear} = \begin{cases} \frac{C_{sRGB}}{12,92}, & C_{sRGB} \leq 0,04045 \\ \left(\frac{C_{sRGB}+a}{1+a}\right)^{2,4}, & C_{sRGB} > 0,04045 \end{cases}$$

kjer je $a = 0,055$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{bmatrix} \begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix}$$

A.5 Pretvorbe med CIE XYZ in Adobe RGB

Pretvorba iz CIE XYZ v Adobe RGB

$$\begin{bmatrix} R_{linear} \\ G_{linear} \\ B_{linear} \end{bmatrix} = \begin{bmatrix} 2,0414 & -0,5649 & -0,3447 \\ -0,9693 & 1,8760 & 0,0416 \\ 0,0134 & -0,1184 & 1,0154 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

C v $C_{AdobeRGB}$ in C_{linear} nadomestimo z R , G in B za izračun posamezne komponente:

$$C_{AdobeRGB} = C_{linear}^{\frac{1}{\gamma}}$$

kjer je $\gamma = 2,19921875$

Vrednosti R , G in B so na intervalu $[0, 1]$.

Pretvorba iz Adobe RGB v CIE XYZ

C v C_{linear} in $C_{AdobeRGB}$ nadomestimo z R , G in B za izračun posamezne komponente. Vhodne vrednosti R , G in B so na intervalu $[0, 1]$.

$$C_{linear} = C_{AdobeRGB}^{\gamma}$$

kjer je $\gamma = 2,19921875$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,57667 & 0,18556 & 0,18823 \\ 0,29734 & 0,62736 & 0,07529 \\ 0,02703 & 0,07069 & 0,99134 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Literatura

- [1] ANSI/ASA S1.1-2013 American National Standard Acoustical Terminology, 2013.
- [2] Sebastian Böck and Markus Schedl. Polyphonic piano note transcription with recurrent neural networks. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 121–124. IEEE, 2012.
- [3] D. Bendor, M. S. Osmani, and X. Wang. Dual-Pitch Processing Mechanisms in Primate Auditory Cortex. *Journal of Neuroscience*, 32(46):16149–16161, November 2012.
- [4] J. Ashley Burgoyne and Lawrence K. Saul. Visualization of Low Dimensional Structure in Tonal Pitch Space. 2005.
- [5] P. A. Cariani and B. Delgutte. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *Journal of Neurophysiology*, 76(3):1698–1716, September 1996.
- [6] Norman Cazden. The Definition of Consonance and Dissonance. *International Review of the Aesthetics and Sociology of Music*, 2:123–168, 1980.
- [7] Winnie Wing-Yi Chan. *A Report on Musical Structure Visualization*. PhD thesis, Hong Kong University of Science and Technology, Hong Kong, 2007.

-
- [8] Elaine Chew. *Towards a Mathematical Model for Tonality*. Ph.D. dissertation, Operations Research Center, MIT, Cambridge, MA, 2000.
- [9] Peter Ciuha, Bojan Klemenc, and Franc Solina. Visualization of concurrent tones in music with colours. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 1677–1680. ACM, 2010.
- [10] Richard S Cook, Paul Kay, and Terry Regier. The world color survey database. *Handbook of categorization in cognitive science*, pages 223–241, 2005.
- [11] Egbert De Boer. *On the "residue" in hearing*. PhD thesis, Uitgeverij Excelsior, 1956.
- [12] Alain De Cheveigné. Cancellation model of pitch perception. *The Journal of the Acoustical Society of America*, 103(3):1261–1271, 1998.
- [13] Alain De Cheveigné. Pitch perception models—a historical review. *CNRS-Ircam, Paris, France*, 2004.
- [14] Morwaread Farbood, Henry Kaufman, and Kevin Jennings. Composing with Hyperscore: an Intuitive Interface for Visualizing Musical Structure. In *Proceedings of the ICMC*, 2007.
- [15] Hugo Fastl and Eberhard Zwicker. *Psychoacoustics: Facts and Models*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.
- [16] Joyce Horn Fonteles, Maria Andréia Formico Rodrigues, and Victor Emanuel Dias Basso. Creating and evaluating a particle system for music visualization. *Journal of Visual Languages & Computing*, 24(6):472–482, December 2013.
- [17] Adrian Ford and Alan Roberts. Colour space conversions. *Westminster University, London*, 1998:1–31, 1998.

-
- [18] Gabriel Gatzsche, Markus Mehnert, David Gatzsche, and Karlheinz Brandenburg. A Symmetry Based Approach for Musical Tonality Analysis. In *Proceedings of the 8th International Conference on Music Information Retrieval, ISMIR 2007*, pages 207–210, 2007.
- [19] Stanley A Gelfand. *Hearing: an introduction to psychological and physiological acoustics*. Informa Healthcare, London, 2010.
- [20] David Gerhard, University of Regina, and Department of Computer Science. *Pitch extraction and fundamental frequency: history and current techniques*. Dept. of Computer Science, University of Regina, Regina, 2003. OCLC: 54005806.
- [21] Emilia Gómez and Jordi Bonada. Tonality visualization of polyphonic audio. In *Proceedings of International Computer Music Conference*, 2005.
- [22] Julius L. Goldstein. An Optimum Processor Theory for the Central Formation of the Pitch of Complex Tones. *The Journal of the Acoustical Society of America*, 54(1), 1973.
- [23] Edward Gollin. Some Aspects of Three-Dimensional "Tonnetze". *Journal of Music Theory*, 42(2):195, 1998.
- [24] Peter Gouras. Colour Vision. In John Wiley & Sons, Ltd, editor, *Encyclopedia of Life Sciences*. John Wiley & Sons, Ltd, Chichester, September 2007.
- [25] Joseph Hall and Robert W. Peters. Pitch for nonsimultaneous successive harmonics in quiet and noise. *Journal of The Acoustical Society of America*, 69(2):509–513, 1981.
- [26] Christopher Harte. *Towards automatic extraction of harmony information from music signals*. PhD thesis, Department of Electronic Engineering, Queen Mary, University of London, 2010.

-
- [27] Rumi Hiraga, Reiko Mizaki, and Issei Fujishiro. Performance visualization: a new challenge to music through visualization. In *The tenth ACM international conference on Multimedia*, pages 239–242. ACM, 2002.
- [28] A.J.M. Houtsma and John .G. Beerends. Pitch identification of simultaneous diotic and dichotic two-tone complexes. *Journal of the Acoustical Society of America*, 85(2):813–819, February 1989.
- [29] David M Howard and Jamie Angus. *Acoustics and psychoacoustics*. Focal, Amsterdam; London, 2009.
- [30] Eric Isaacson. What You See Is What You Get: On Visualizing Music. In *In: Proceedings of International Symposium in Music Information Retrieval*, 2005.
- [31] Bojan Klemenc, Peter Ciuha, and Franc Solina. Educational Possibilities of the Project Colour Visualization of Music. *Organizacija*, 44(3):67–75, May 2011.
- [32] Bojan Klemenc, Peter Ciuha, Lovro Šubelj, and Marko Bajec. Visual and aural : visualization of harmony in music with colour. *IPSI BGD Trans. Internet Res*, 7(1):48–53, January 2011.
- [33] Carol L. Krumhansl. *Cognitive Foundations of Musical Pitch*. Oxford University Press, USA, 1990.
- [34] Carol L. Krumhansl and Edward J. Kessler. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological review*, 89(4):334, 1982.
- [35] Paul Lamere and Douglas Eck. Using 3d Visualizations to Explore and Discover Music. In *Proceedings of the 8th International Conference on Music Information Retrieval, ISMIR 2007*, pages 173–174, 2007.
- [36] Fred Lerdahl. *Tonal pitch space*. Oxford University Press, 2004.

-
- [37] Fred Lerdahl and Carol L. Krumhansl. Modeling Tonal Tension. *Music Perception: An Interdisciplinary Journal*, 24(4):329–366, April 2007.
- [38] David Lewin. *Generalized Musical Intervals and Transformations*. Oxford University Press, April 2007. Google-Books-ID: aN924j8LyWcC.
- [39] J. C. R. Licklider. “Periodicity” pitch and “place” pitch. *The Journal of the Acoustical Society of America*, 26(5):945–945, 1954.
- [40] João Manuel Maciel Linhares, Paulo Daniel Pinto, and Sérgio Miguel Cardoso Nascimento. The number of discernible colors in natural scenes. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 25(12):2918–2924, December 2008.
- [41] M. Mahy, L. Eycken, and A. Oosterlinck. Evaluation of Uniform Color Spaces Developed after the Adoption of CIELAB and CIELUV. *Color Research & Application*, 19(2):105–121, April 1994.
- [42] Daniel Malacara. *Color Vision and Colorimetry: Theory and Applications, Second Edition*. SPIE, 1000 20th Street, Bellingham, WA 98227-0010 USA, July 2011.
- [43] Delfina Malandrino, Donato Pirozzi, Gianluca Zaccagnino, and Rocco Zaccagnino. A Color-based Visualization Approach to understand harmonic structures of Musical Compositions. In *19th International Conference on Information Visualisation Proceedings*, pages 56–61. IEEE, 2015.
- [44] Arpi Mardirossian and Elaine Chew. Visualizing Music: Tonal Progressions and Distributions. In *Proceedings of the 8th International Conference on Music Information Retrieval, ISMIR 2007*, pages 189–194, Vienna, 2007.
- [45] Agustín Martorell and Emilia Gómez. Two-dimensional visual inspection of pitch-space, many time-scales and tonal uncertainty over time. In *Mathematics and Computation in Music*, Paris, France, 2011. Springer.

-
- [46] Ray Meddis. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. *The Journal of the Acoustical Society of America*, 89(6):2866, 1991.
- [47] Ray Meddis and Lowell P. O'Mard. Virtual pitch in a computational physiological model. *The Journal of the Acoustical Society of America*, 120(6):3861, 2006.
- [48] Christophe Micheyl, Michael V. Keebler, and Andrew J. Oxenham. Pitch perception for mixtures of spectrally overlapping harmonic complex tones. *The Journal of the Acoustical Society of America*, 128(1):257, 2010.
- [49] Brian CJ Moore. *An introduction to the psychology of hearing*. Brill, 2012.
- [50] Jay Neitz, Joseph Carroll, and Maureen Neitz. Color vision: Almost Reason Enough for Having Eyes. *Optics & Photonics News*, 27, 2001.
- [51] Noboru Ohta and Alan R. Robertson. *Colorimetry: fundamentals and applications*. Wiley-IS&T series in imaging science and technology. J. Wiley, Chichester, West Sussex, England ; Hoboken, NJ, USA, 2005.
- [52] Danny Pascale. A review of rgb color spaces... from xyy to r'g'b'. *The BabelColor Company*, 7:8, 2003.
- [53] Roy D. Patterson, K. Robinson, J. Holdsworth, D. McKeown, C. Zhang, and M. Allerhand. Complex sounds and auditory images. *Auditory physiology and perception*, 83:429–446, 1992.
- [54] Matevz Pesek, Ales Leonardis, and Matija Marolt. A Compositional Hierarchical Model for Music Information Retrieval. In *ISMIR*, pages 131–136, 2014.
- [55] Christopher J. Plack, Richard R. Fay, Andrew J. Oxenham, and Arthur N. Popper. *Pitch: Neural Coding and Perception*. Springer Science, January 2005.

-
- [56] Reinier Plomp and Willem JM Levelt. Tonal consonance and critical bandwidth. *The journal of the Acoustical Society of America*, 38:548, 1965.
- [57] Hendrik Purwins. *Profiles of pitch classes circularity of relative pitch and key-experiments, models, computational music analysis, and perspectives*. PhD thesis, Technische Universität Berlin, Berlin, 2005.
- [58] Hendrik Purwins, Benjamin Blankertz, and Klaus Obermayer. Toroidal models in tonal theory and pitch-class analysis. *Computing in Musicology (Tonal Theory for the Digital Age)*, 15:73–98, 2007.
- [59] Luis Robles, Mario A. Ruggero, and Nola C. Rich. Two-tone distortion in the basilar membrane of the cochlea. *Nature*, 349(6308):413–414, January 1991.
- [60] Craig Stuart Sapp. Harmonic visualizations of tonal music. In *Proceedings of the International Computer Music Conference*, volume 1001, pages 423–430, 2001.
- [61] Jan Schnupp, Israel Nelken, and Andrew King. *Auditory neuroscience: Making sense of sound*. MIT press, 2011.
- [62] Arnold Schoenberg and Leonard Stein. *Structural functions of harmony*. WW Norton & Company, 1969.
- [63] Gaurav Sharma and Raja Bala. *Digital Color Imaging Handbook*. CRC Press, December 2002.
- [64] Gaurav Sharma, Wencheng Wu, and Edul N. Dalal. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1):21–30, 2005.
- [65] Siddharth Sigtia, Emmanouil Benetos, and Simon Dixon. An End-to-End Neural Network for Polyphonic Piano Music Transcription. *IEEE/ACM*

-
- Transactions on Audio, Speech, and Language Processing*, 24(5):927–939, 2016.
- [66] Aurea Soriano, Fernando Paulovich, Luis Gustavo Nonato, and Maria Cristina F. Oliveira. Visualization of music collections based on structural similarity. *University of Sao Paulo*, 2014.
- [67] T. M. Talavage. Tonotopic Organization in Human Auditory Cortex Revealed by Progressions of Frequency Sensitivity. *Journal of Neurophysiology*, 91(3):1282–1296, October 2003.
- [68] James Tanney. *A history of consonance and dissonance*. Excelsior, New York, 1988.
- [69] Ernst Terhardt. The Concept of Musical Consonance: A Link between Music and Psychoacoustics. *Music Perception: An Interdisciplinary Journal*, 1(3):276–295, 1984.
- [70] Ernst Terhardt, Gerhard Stoll, and Manfred Seewann. Pitch of complex signals according to virtual-pitch theory: Tests, examples, and predictions. *The Journal of the Acoustical Society of America*, 71(3):671–678, 1982.
- [71] P. Toiviainen. Visualization of Tonal Content in the Symbolic and Audio Domains. *Tonal theory for the digital age*, pages 187–198, 2007.
- [72] D. Tymoczko. The Generalized Tonnetz. *Journal of Music Theory*, 56(1):1–52, March 2012.
- [73] Jernej Vrščaj, Bojan Klemenc, Peter Ciuha, and Marko Jankovec. Elektronski sistem za vizualizacijo glasbe z barvnim osvetljevanjem klavirskih tipk. *Zbornik dvaindvajsete mednarodne Elektrotehniške in računalniške konference ERK 2013, 16.-18. september 2013, Portorož, Slovenija*, B:213–214, 2013. OCLC: 892523470.

-
- [74] F. L. Wightman. The pattern-transformation model of pitch. *Journal of the Acoustical Society of America*, 54(2):407–416, 1973.
- [75] W. A. Yost. Pitch strength of iterated rippled noise. *The Journal of the Acoustical Society of America*, 100(5):3329–3335, November 1996.