

KAUNO TECHNOLOGIJOS UNIVERSITETAS

INFORMATIKOS FAKULTETAS

MULTIMEDIJOS INŽINERIJOS KATEDRA

Urtė Radzickaite

**„Hard - Lost“ ir „DUET“ šaltinių išskyrimo
algoritmu, taikant juos garso analizės srityje,
palyginimas**

Magistro darbas

Darbo vadovas

doc. A. Ostreika

KAUNO TECHNOLOGIJOS UNIVERSITETAS

INFORMATIKOS FAKULTETAS

MULTIMEDIJOS INŽINERIJOS KATEDRA

Urtė Radzickaitė

**„Hard - Lost“ ir „DUET“ šaltinių išskyrimo
algoritmų, taikant juos garso analizės srityje,
palyginimas**

Magistro darbas

Recenzentas

Dr. T. Blažauskas

2007-05-28

Vadovas

doc. A. Ostreika

2007-05-28

Atliko

IFM-1/2 gr. stud.

Urtė Radzickaitė

2007-05-15

Kaunas, 2007

TURINYS

1. ĮVADAS	5
2. GARSO ŠALTINIŲ IŠSKYRIMO METODŲ ANALIZĖ	6
2.1. Garso šaltinių išskyrimo algoritmai fundamentaliųjų dažnių analizės kontekste.....	7
2.2. Fundamentaliųjų dažnių metodikos.....	8
2.2.1. Girdimosios aplinkos analizė.....	8
2.2.2. Klausos modelių pagrįsta garso šaltinių analizė.....	10
2.2.3. „Lentos“ („Blackboard“) principu pagrįsta garsų srauto skaidymo sistema.....	11
2.2.4. Signalų modeliais pagrįsta tikimybinė analizė.....	13
2.2.5. Prie duomenų prisitaikančios sistemos.....	14
2.3. Fundamentaliųjų dažnių metodikų palyginimas.....	15
2.4. Garso šaltinių išskyrimo metodai	16
2.4.1. Aklas šaltinių išskyrimas	16
2.4.2. Retasis programavimas, neneigiamų matricių faktorizavimas	17
2.4.3. Sinusoidinis modeliavimas, garso srauto aukščių analizė.....	17
3. GARSO ŠALTINIŲ IŠSKYRIMO SISTEMOS PROJEKTAS.....	18
3.1. „SSS“ sistemai keliami reikalavimai	18
3.1.1. „SSS“ sistemos funkcinių reikalavimų sąrašas	20
3.1.2. „SSS“ sistemos nefunkcinių reikalavimų sąrašas.....	20
3.2. „SSS“ sistemos architektūra	21
3.2.1. „SSS“ sistemos ribos – panaudos atvejų modelis.....	21
3.2.2. „SSS“ sistemos statinė ir dinaminė struktūra	23
3.3. „SSS“ sistemos projekto apibendrinimas	27
4. „SSS“ SISTEMOS KOKYBĖS TYRIMAS	29
5. „DUET“ IR „HARD-LOST“ ALGORITMŲ TYRIMAS	31
5.1. „DUET“ algoritmas	31
5.1.1. „DUET“ algoritmo koncepcija	31
5.1.2. „DUET“ algoritmo aprašymas.....	32
5.2. „Hard - Lost“ algoritmas	33
5.2.1. „Hard - Lost“ algoritmo koncepcija	34
5.2.2. „Hard-Lost“ algoritmo aprašymas.....	35
5.3. „DUET“ ir „Hard- Lost“ algoritmų palyginimas	36
5.4. Eksperimentų rezultatai	37
6. IŠVADOS	41
7. LITERATŪRA	44
8. TERMINŲ IR SANTRUMPŲ ŽODYNAS.....	46
9. PRIEDAI.....	48
9.1. Algoritmų, reikalaujančių šaltinių išretinimo apribojimo palyginimas, pagal [22]	48
9.2. „SSS“ sistemos funkcinis aprašymas	48
9.2.1. Trumpas „SSS“ sistemos aprašymas	48
9.2.2. „SSS“ sistemos paskirtis.....	48
9.2.3. „SSS“ sistemos vartotojai	49
9.2.4. „SSS“ sistemos funkcijos	49
9.2.5. „SSS“ sistemos vaizdas	50

SUMMARY

This work presents two algorithms – “DUET” by Jorjine et al. and “Hard - Lost” by Pearlmutter et al. comparison. Concern areas are algorithm usage and development perspectives in non-scientific, user – oriented field. Broad field analysis was done from methodics analysis to specific algorithm analysis. Choise of algorithm comparison was made, considering reported algorithm acurracy and study popularity.

Work presents comparison of two data dependent system methodics algorithms. Algorithm acurracy and work speed were compared. Also, conclusions were drawn about algorithm usage in practical field.

“Hard - Lost” algorithm was implemented in a broader system, called “SSS” – “Sound (source) separation system”. System was created for algorithm usage for practical purposes in non-scientific field. Work introduces potetial expanding of the system and algorithm itself for further development in user – oriented areas.

1. ĮVADAS

Šaltinių išskyrimas yra sparčiai besivystanti tyrimų kryptis. Taip yra todėl, kad taikymų sritis šaltinių išskyrimui yra plati. Pagal [2] šaltinių išskyrimas gali būti naudojamas medicinoje – širdies ritmų ir smegenų virpesių analizei, vaizdų išryškimui ir suspaudimui, ekonometrijoje – poveikio finansiniams duomenims esminių veiksnių radimui, telekomunikacijoje – kelių vartotojų signalų atskyrimui.

Viena iš dažniausių šaltinių išskyrimo taikymo sričių yra garso šaltinių analizė skaitmeninių signalų apdorojimo procese. Standartinis garso šaltinių išskyrimo pavyzdys yra „kokteilių vakarėlio“ problema:

Įsivaizduokime keletą žmonių vienu metu kalbančių tam tikroje erdvėje („kokteilių vakarėlyje“). Klausytojas bando išskirti atskiro žmogaus kalbą. Žmogaus smegenys garso šaltinio išskyrimo darbą atlieka gana gerai, tačiau ši problema skaitmeninių signalų apdorojimo srityje yra sudėtinga. Šiuo metu neegzistuoja nė vieno matematinio metodo, galinčio atskirti du realaus pasaulio garso šaltinius 100% tikslumu. Egzistuoja įvairūs problemos sprendimo būdai, besiskiriantys savo sudėtingumu, sprendimo sistemos suteikiamais apribojimais, taikymo kryptimis.

Vienas iš labiausiai paplitusių garso išskyrimo metodų yra aklas šaltinių išskyrimas (BSS). Šis metodas yra plačiai naudojamas ir tiriamas dėl savo taikymo universalumo. Taip pat, skirtingai nuo kitų garso šaltinių išskyrimo metodų, jis nereikalauja išankstinės informacijos apie garso šaltinius. Šios metodo pagrindinis trūkumas yra tas, kad norint naudoti jį realiems, nesintezuotiems signalams, reikia apibrėžti apribojimus šiems signalams.

Darbo tikslas yra susisteminti egzistuojančius duomenis apie garsų analizę bei konkrečiai garso šaltinių išskyrimą. Kadangi yra atliekama daug individualių tyrimų garsų analizės kryptimi, taip pat yra siekiama numatyti, kuri tyrimų kryptis yra perspektyviausia ir turėtų vystytis toliau. Kad tai būtų atlikta, darbe apžvelgtos skirtingos garsų analizės metodikos, jas lyginant tarpusavyje. Gilesnei analizei pasirinkta „prie duomenų prisitaikančios sistemos“ metodika. Tiriama šios metodikos naudojami metodai ir išsirinktas geriausias rezultatus duodantis metodas – „aklas šaltinių išskyrimas“. Darbe tiriama du šio metodo įgyvendinimo būdai – „DUET“ ir „Hard - Lost“ algoritmai, siekiant nustatyti, kuris iš jų yra perspektyvesnis.

2. GARSO ŠALTINIŲ IŠSKYRIMO METODŲ ANALIZĖ

Norint atskirti garso šaltinius iš garsų srauto ir panaudoti juos garsų analizės srityje visų pirma reikia remtis garso šaltinių atskyrimo metodika. Taip yra todėl, kad garso šaltinių išskyrimas nėra triviali užduotis ir neegzistuoja bendras algoritmas, leidžiantis išskirti garso šaltinius su patenkinama žmogaus ausiai kokybe.

Šiuo metu egzistuojantys garso šaltinių atskyrimo sprendimai yra labai netobuli, ypač kai yra kalbama apie realaus pasaulio garsus. Dauguma sprendimų yra apriboti atskiriamų garso šaltinių skaičiumi, garsų diapazonu, kilme. Dažnai norint išskirti garso šaltinius yra daromos įvairios apribojančios prielaidos. Pavyzdžiui: dažnai turi iš anksto būti žinomas šaltinių skaičius; garso šaltiniai turi būti „išretinti“ – vienu metu aktyvus būti tik vienas šaltinis; yra išskiriami tik tam tikro ilgio garsų srautai ir kita.

Kiekviena garsų srauto skaidymo metodika turi savo stipriąsias ir silpnąsias puses. Dažnai jose naudojami algoritmai bei technikos sutampa arba yra panašūs. Todėl yra sudėtinga pilnai atskirti metodikas vieną nuo kitos.

Egzistuoja trys pagrindinės garsų srauto išskaidymo, susijusio su muzikine sritimi tyrimo grupės:

- Muzikinio metro atpažinimo tyrimai [4];
- Muzikologinė analizė [5];
- Fundamentaliųjų dažnių analizė.

Pirmoji grupė yra skirta atpažinti ir analizuoti pasikartojančius elementus garsų sraute – muzikinį metrą. Antroje grupėje yra pritaikomi turimų muzikinių duomenų ir taisyklių rinkiniai, kad būtų galima interpretuoti garsų srautą kaip muziką ir šiuo būdu patikslinti išskaidytą garsų srautą.

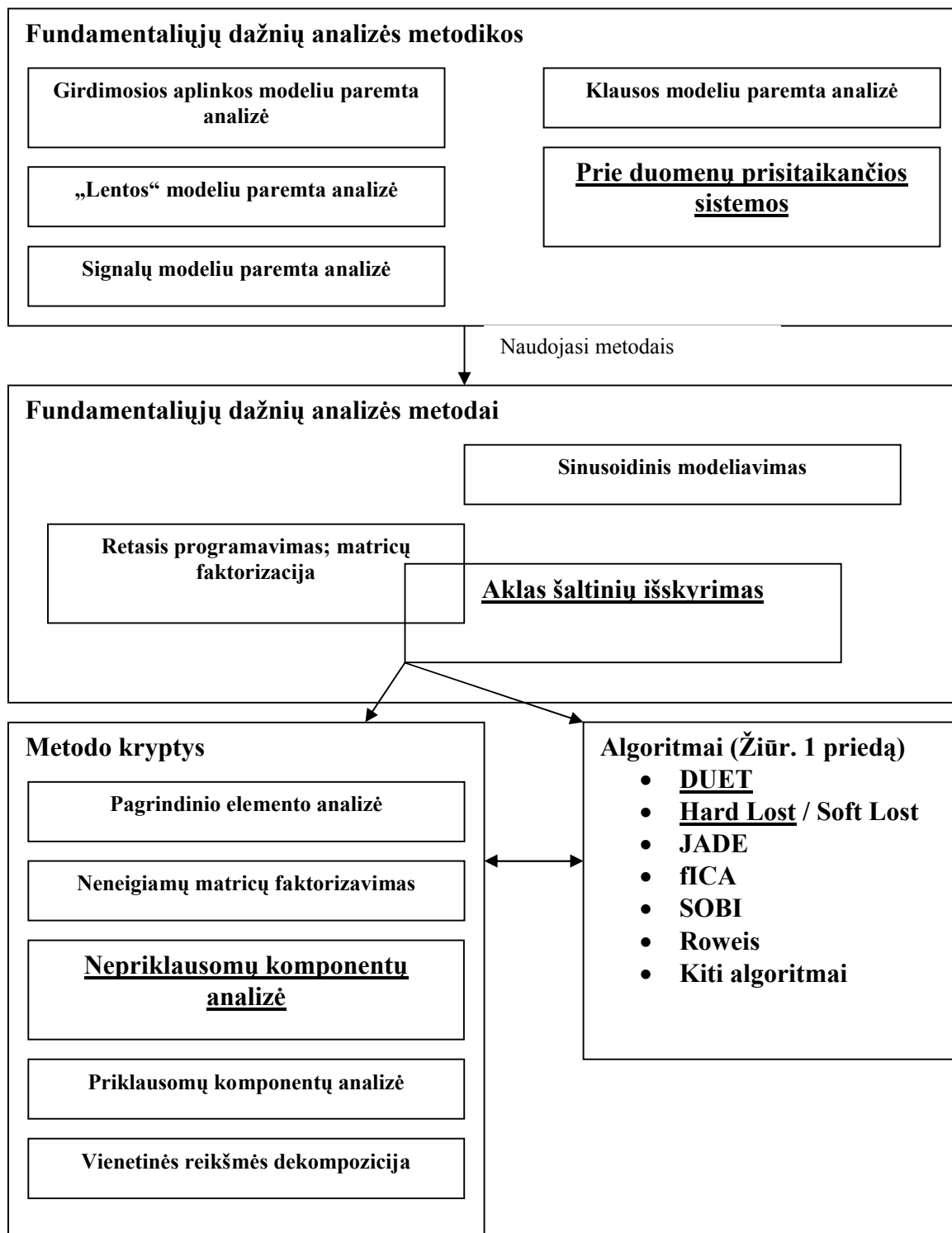
Tačiau pagrindinė tyrimų kryptis yra fundamentaliųjų dažnių analizė, nes tai yra garsų srauto skaidymo į atskirus šaltinius pagrindas. Atlikus fundamentaliųjų dažnių analizę yra gaunami garsų srauto skaidymo rezultatai – garso šaltiniai arba individualūs garsai (priklausomai nuo tyrimo krypties).

Fundamentaliųjų dažnių analizė gali būti apibrėžta keliais skirtingais modeliais. Metodikos nagrinėja skirtingus garso šaltinių išskyrimo atvejus arba srauto skaidymą skirtingais aspektais. Šių metodikų naudojamos technikos garso šaltinių išskyrimui neretai yra panašios. Todėl galima sakyti, kad viena metodika tam tikrais atvejais papildo kitą. Galimas garso šaltinių skaidymo metodikų ir jų

naudojamų technikų apibendrinimas yra pateiktas skyriuje „2.1 Garso šaltinių išskyrimo algoritmai fundamentaliųjų dažnių analizės kontekste“.

2.1. Garso šaltinių išskyrimo algoritmai fundamentaliųjų dažnių analizės kontekste

Kaip minėta anksčiau, garsų srauto analizė yra sudėtingas procesas, reikalaujantis modelių, kurie yra paremti konkrečiais metodais. Darbe yra lyginami du garso srauto skaidymo algoritmai „Hard Lost“ ir „DUET“. Jie priklauso aklo šaltinių išskyrimo metodų grupei, kuri savo ruožtu yra viena iš duomenimis paremtų sistemų. Žemiau yra pavaizduota fundamentaliųjų dažnių analizės hierarchija, kurioje pažymėti darbe analizuojami algoritmai. Skyriuose „2.2 *Fundamentaliųjų dažnių metodikos*“ ir „2.4 *Garso šaltinių išskyrimo metodai*“ apžvelgiama darbo sritis. Pačių algoritmų aprašymai yra pateikiami tiriamojoje dalyje, skyriuose „5.1 „DUET“ algoritmas“ ir „5.2 „Hard - Lost“ algoritmas“.



1 pav. Fundamentaliųjų dažnių analizės hierarchija

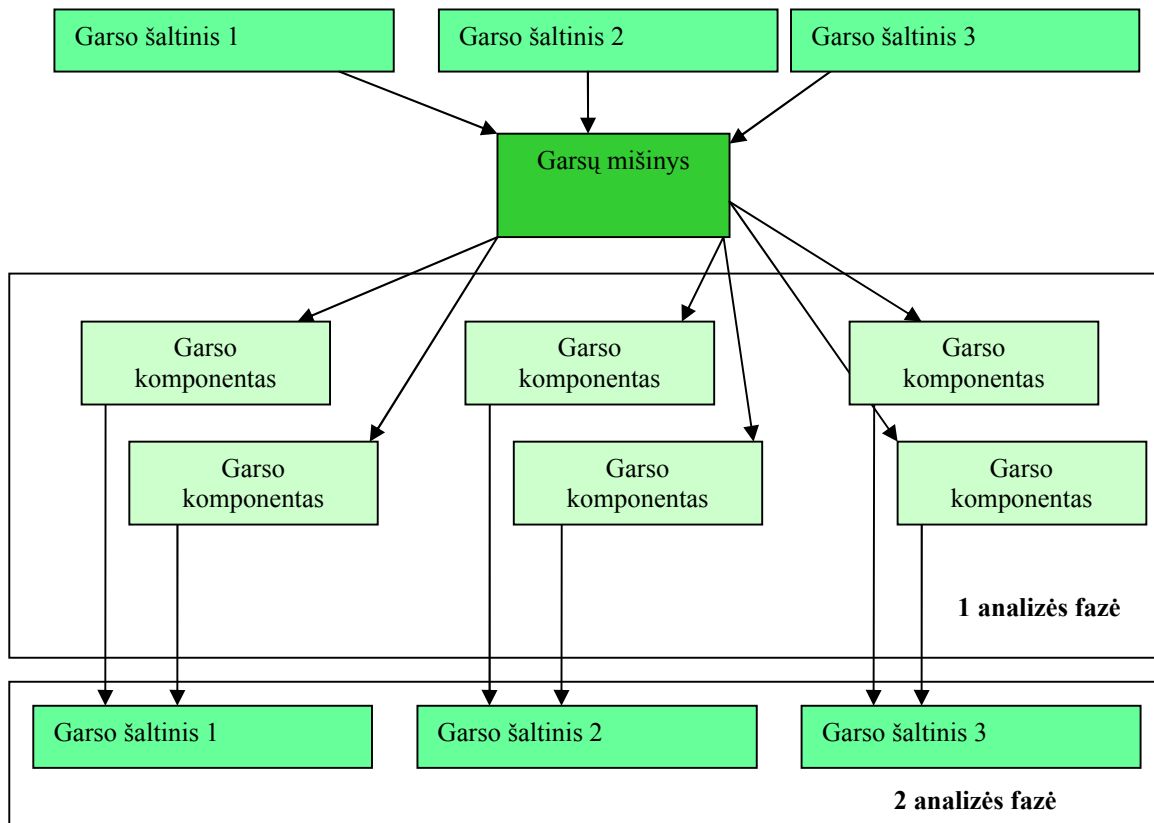
2.2. Fundamentaliųjų dažnių metodikos

2.2.1. Girdimosios aplinkos analizė

Girdimosios aplinkos analizė (ASA) – procesas, kurio metu klausos sistema priima garsų mišinį iš sudėtingos natūralios aplinkos ir suskirsto į garsinius vienetus, kur manoma, kad kiekvienas vienetas yra greičiausiai kilęs iš vieno garso šaltinio. [6]

Natūrali girdimoji aplinka ypatinga tuo, kad ji susideda iš *įvykių* garsų sekos mišinio, sukkelto kelių konkuruojančių garsų. Ši įvykių seka pasiekia žmogaus ausį su skirtinga akustine energija.

Girdimosios aplinkos analizė remiasi psichoakustiniais metodais. Čia yra analizuojami metodai, kuriais žmogaus smegenys suvokia įeinantį garso srautą, kaip skirtingų dažnio dedamųjų visumą, bei šių dedamųjų grupavimą į atkirus šaltinius, atsirandančius laike. Apibendrinus [6], [7] apžvelgtus darbus, galima nubraižyti abstraktų girdimosios aplinkos analizės modelį:



2 pav. Girdimosios aplinkos analizės modelis

Skaičiuojamoji girdimosios aplinkos analizė (CASA) remiasi akustinės informacijos iš fizinės aplinkos bei dažnai pasikartojančių atskirų įvykių joje interpretacijos skaičiuojamąja analize.

Žmonės tyrinėję šią sistemą: A. Bregman, A. D. Sterian, D. Mellinger, J. Eggink, G. J. Brown, K. Kashino ir kiti.

Svarbiausi tyrimai:

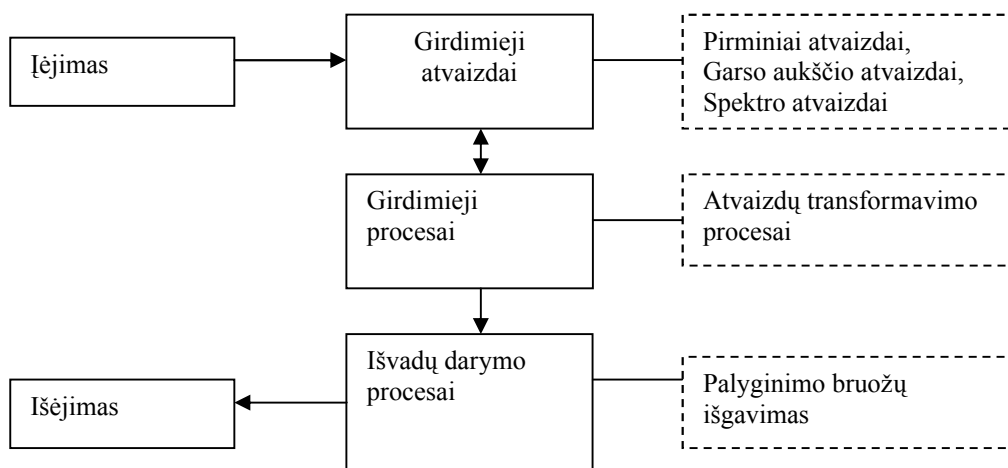
Girdimosios aplinkos analizės metodai labai patobulėjo bei buvo pradėti aktyviai taikyti garsų srauto analizėje, kai Bregman nurodė komponentų bruožus, pagal kuriuos reikėtų skaidyti garso šaltinius: laiko-dažnio panašumas, harmonijos dažnių sąryšiai, sinchroniški pasikeitimai bei erdvinis panašumas.

Vienas iš svarbiausių muzikos analizės krypties darbų sekusių po to – Kashino laiko – dažnio „plokštumų suspaudimo“ („pinching planes“) metodas, kuris išgauna besitęsiančius dažnio komponentus iš tiriamų duomenų. Garso srautų analizei buvo panaudoti tembro modeliai (kiekvienos garso šaltinio identifikavimui), tonų pavyzdžiai (persidengiančių dažnio komponentų atskyrimui). Buvo atlikta akordų analizė (skaičiuojama tikimybė naiti būti akorde), o akordų perėjimo tikimybės koduojamos trigramomis (panaudotos Markovo grandinės). Integracijai („iš apačios į viršų“ analizei, laikiniam sujungimui bei „iš viršaus – žemyn“ apdorojimui) naudojami Bajeso tikimybiniai tinklai.

Kitas šio tipo metodas, aprašytas Sterian naudojo Kalmano filtrus, kad atskirtų sinusines daleles bei atvaizdavo garsų grupavimo į šaltinius taisykles kaip tikimybės funkcijas, kurios įvertino atrastų komponentų tikėtinumą, kai yra duotas hipotetinis grupavimas. (Pagal [9])

2.2.2. Klausos modeliu paremta garso šaltinių analizė

Kitas galimas garsų srauto skaidymo į šaltinius modelis yra klausos modeliu paremta garso šaltinių analizė. Pagal [9] klausos modelis gali būti vaizduojamas taip:



3 pav. Klausos modelis, pritaikytas garso šaltinių analizei

Girdimieji atvaizdai atspindi realaus garso atvaizdavimo bruožus. Girdimojo atvaizdo turinys atspindi bruožus, susijusius su garso šaltinio signalais. Įėjimo signalai yra pernešami neuronų masyvo. Žiūrint iš kompiuterinio modeliavimo pusės, girdimasis atvaizdas yra surikiuotų skaičių masyvas, kuris atspindi nervinę aktyvaciją. Ji modeliuojama, skaičiuojant neuronų sužadavimo tikimybę per tam tikrą laiko intervalą.

Atvaizdų transformavimo procesas paverčia garsą į atvaizdus. Skirtingos atminties sistemos turi skirtingus atvaizdavimo būdus. Modeliuojant trumpalaikę ir ilgalaikę atmintį, atvaizdai išskiriami į trumpalaikius bei ilgalaikius. Procesas yra lyginamas su žmogaus fiziologija.

Išvados apibūdinamos kaip išvestinė informacija. Jos gali būti lyginamos su žmogaus elgesio reakcija.

Išvadų darymo procesai – tai procesai, kurie lygina, tikrina bei išgauna tam tikrus bruožus iš muzikinių atvaizdų.

Bendrai klausos modelių paremti tyrimai gali apimti kelis klausos modeliavimo aspektus. Modeliai gali būti sensoriniai, percepciniai, pažintiniai/kognityvieji.

Svarbiausi tyrimai:

Vienas iš pagrindinių metodų, kuriais remiasi šio tipo tyrimai, yra Meddis ir Hewitt „vienetinio garso aukščio modelis“ aprašytas [10]. Jie modeliavo periferines klausos dalis. Nors į kelių garsų išskyrimą šiame tyrime nebuvo atsižvelgta, metodas tapo pagrindu kitiems šio tipo tyrimams.

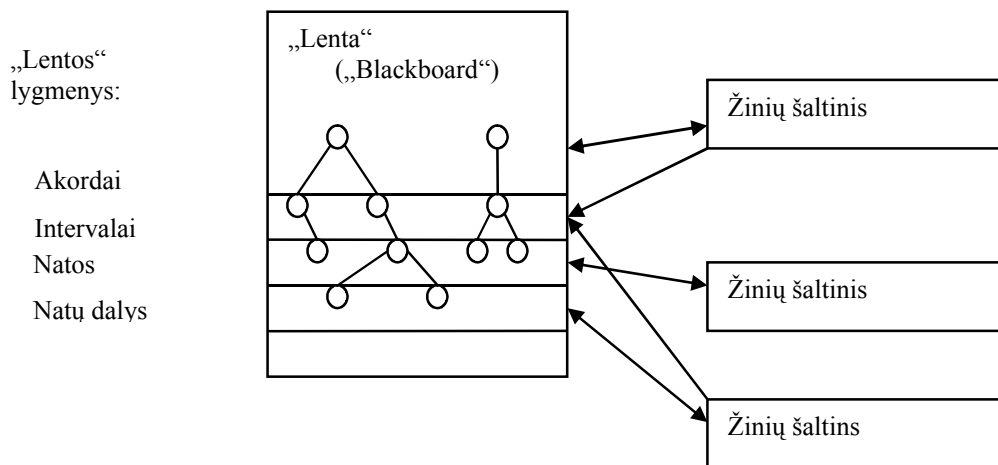
Cheveigne ir Kawahara pritaikė modelį daugelio dažnių išskyrimui. Jie pasiūlė metodą, kur po garso aukštumo nustatymo sekė identifikuoto garso anuliavimo fazė iš bendro garsų srauto, o tolimesnis aukštumo nustatymas taikomas likusiems signalams. Jie taip pat pasiūlė daug sąnaudų reikalaujantį garsų sekos atskyrimo modelį. Į laikinį tęstinumą tarp garsų srauto kadru (dalių, į kurias garsų srautas skirstomas norint jį analizuoti) neatsižvelgta. Nors tyrimui buvo taikomi dirbtiniai duomenys, jis buvo gana sėkmingas [11].

Tolonen bei Karjalainen išvystė skaičiavimams pritaikytą „vienetinio aukščio modelio“ versiją bei pritaikė kelių dažnių garso srautų analizei. Įeinančio garso srauto spektras buvo išlyginamas naudojant atvirkštinį „iškreiptąjį linijinį tikimybių“ filtrą. Garso aukščio skaičiavimams naudojami tik du kanalai, vietoj originaliame „vienetinio aukščio modelyje“ naudojamų 40-120, tačiau skaičiavimo charakteristikos nuo to beveik nenukentėjo. Metodas optimizuotas pasinaudojant greita Furjė transformacija. Kelių aukščių nustatymas buvo pasiektas sub-harmonikų anuliavimu modelio išėjime. Galutinėje pasiūlytoje „pagerintoje santraukos autokoreliacijos funkcijoje“ visus dažnius galima išrinkti be pasikartojančio aukštumo nustatymo bei garsų anuliavimo. Pasak [12] metodas pasiekė gana tikslių rezultatų.

2.2.3. „Lentos“ („Blackboard“) principu paremta garsų srauto skaidymo sistema

„Lentos“ sistemos modelis – struktūrizuotas oportunistinio problemų sprendimo dalinis atvejis. Jis apibrėžia organizacinę problemų sprendimo struktūrą, kurioje yra aiškūs tarpiniai bei galutiniai problemos sprendimų žingsniai bei egzistuoja bendra sprendimų priėmimo erdvė [13].

Apibendrinta „lentos“ sistemos architektūra, pritaikyta muzikinei sričiai:



4 pav. „Lentos“ modelis, pritaikytas garso šaltinių analizei

Sistema susideda iš atskirų sutvarkytų dalių – žinių šaltinių – su skirtingais žodynais, algoritmais bei euristinėmis taisyklėmis, skirtomis problemų sprendimui. Bendra sritis („lenta“) sudaryta iš problemos sprendimo informacijos dalių, paimtų iš atskirų žinių šaltinių, kur informacija gali būti susieta su keliais informacijos lygmenimis.

Žinių šaltiniai sąlygoja pasikeitimą „lentoje“, o ši savo ruožtu keičia informaciją, esančią kituose žinių šaltiniuose. Taip, besikeičiant informacija, per kelis žingsnius prieinama iki galutinio problemos sprendimo.

Oportunistiniai samprotavimai sistemai pritaikomi dinamiškai, po vieną žingsnį, apibendrinant dalinius sprendimus į vieną.

Žmonės, tyrinėję šio tipo garsų srauto analizės sistemą: J. Bello, G. Monti, M. Sandler, D. Ellis, D. Godsmark, G. Brown, S. Hainsworth, K. Kashino, K. Nakadai, T. Kinoshita, H. Tanaka, K. Martin, A. Klapuri bei kiti.

Svarbiausi tyrimai:

Vienas iš pirmųjų metodų, pasiūlytų Martin'o, analizavo fortepijono muziką. Šioje sistemoje garso įvedimo modelis tarnavo kaip sąsaja su „lenta“, kurioje buvo žinios apie fizinį garso išgavimą, muzikos tonų valdymo taisykles bei „šiukšlių surinkimo“ euristicas. Dažnių analizė rėmėsi išskaidymu į kadrus, pridėdant papildomą žodyno informaciją, padedančią kurti smulkiusių dalelių – natų – hipotezes.

Kitas modelis, pasiūlytas Godsmark'o ir Brown'o buvo konkrečiai pritaikytas palengvinti skirtingų garsinių organizavimo principų apjungimą. Programa kūrė „synchronines gijas“, kurių kiekviena atvaizdavo dominuojantį laiko-dažnio komponentą. Jos buvo sujungiamos į garso įvykius, išgaunant kiekvienos gijos savybes (tam pritaikant Bregmano primityvios organizacijos principus). Toliau garsiniai įvykiai grupuojami į garso šaltinius, apskaičiuojant aukštumo bei tembro artimumą tarp garsų. Muzikinio metro informacija naudojama nuspėti, kada įvyks garsinis įvykis. Modelis

sugebėjo išskirti melodijos linijas iš daugiabalsės muzikos. Garso srautų atskyrimo tikslumas nebuvo metodo pagrindinis tikslas. Plačiau apie tyrimą galima skaityti [13],[14].

Daug darbų šioje srityje atliko J. Bello, G. Monti ir M. Sandler, kurie tyrė šią sistemą nuo 2000 metų. Jie rėmėsi Martin'o darbais, todėl jų sistema turi panašius trūkumus kaip ir Martin'o sistema. Tačiau jie labai patobulino „lentos“ sistemą naudodami daug skirtingų technikų (įtrauktas „iš viršaus į apačią“ bei „iš apačios į viršų“ apdorojimas, neuroniniais tinklais paremtas akordų atpažinimo įrankis, miglotoji išvadų darymo sistema, naudojama išankstinis muzikos instrumentų žinynas bei pridėtas psichoakustinės informacijos atributų išplėtimas).[15]

2.2.4. Signalų modeliais paremta tikimybinė analizė

Dažnių išskyrimo problemą įmanoma aprašyti signalų modeliais, kuriems turi būti nustatyti parametrai.

Bendras signalų modelis aprašytas matematiškai (gali būti išplėstas), pagal [14]:

$$y_t = \left\{ \sum_{k=1}^K \sum_{m=1}^{M_k} a_{k,m} \cos[m\omega_k t] + b_{k,m} \sin[m\omega_k t] \right\} + v_t$$

K – vienalaikių signalų skaičius; M_k – garsą k sudarančių dalelių skaičius; ω_k – fundamentalus garso k dažnis; $a_{k,m}$, $b_{k,m}$ – garso dalelės amplitudė bei fazė; v_t – liekamasis triukšmas.

Į visus šiuos parametrus turi būti atsižvelgta, sekant garso srautą bei remiantis išankstinėmis parametru išsidėstymo žiniomis, kurios dažniausiai apskaičiuojamos pagal tikimybinę Bajeso teoremą. Kadangi reikalaujama daug išankstinių žinių apie garso šaltinius, pats modelis yra sudėtingas ir sudėtingai pritaikomas praktikoje.

Svarbiausi tyrimai:

Davy ir Godsill [17] detalizavo šį signalų modelį įtraukdami laike besikeičiančių signalų amplitudžių, neidealios harmonijos bei nebaltojo liekamojo triukšmo skaičiavimus. Išankstiniai parametru pasiskirstymai buvo kruopščiai pasirinkti. Įeinantis signalas išskaidomas į dalis, kuriose nėra garso dalelių perėjimų. Tada nustatyti kiekvieno segmento signalų modelio parametrai laiko erdvėje. Pagrindinė modelio problema slypi skaičiavimuose. Kiekvieno realaus signalo parametru erdvė yra milžiniška, o signalų apjungimas negali būti tiesiogiai išreikštas. Kad būtų sumažinti skaičiavimai, parametru erdvės analizei naudotas Markovo grandinių Monte Carlo išeigos atrankos euristinis metodas. Sistema reikalauja daug skaičiavimų.

Vienas iš sėkmingiausių tos srities metodų yra Goto pasiūlytas metodas [18], kuris modeliuoja trumpalaikį signalo spektrą, atvaizduojamą kaip garso aukščio modelių mišinys. Kiekvienas garso aukščio modelis susideda iš konkretaus skaičiaus harmoninių (susietų tarpusavyje) komponentų,

kurie atvaizduojami dažnių spektre Gauso pasiskirstymu. Goto išvedė „tikėtumo maksimizavimo“ algoritmą, kuris pastoviai atnaujiną kiekvieno garso aukščio modelio svorio koeficientus bei jų harmonines amplitudes, kas savo ruožtu veda į maksimalų išankstinį garso aukščio numatymą. Į laikinį tęstinumą atsižvelgta, stebint dažnio dedamųjų svorius atliekant algoritmą. Šis modelis sėkmingai pritaikytas melodijų bei bosinių partijų išskyrimui iš kompaktinių diskų įrašų.

2.2.5. Prie duomenų prisitaikančios sistemos

Skirtingai nuo modeliais paremtų sistemų („lentos“, girdimosios aplinkos analizės ir t.t..) prie duomenų prisitaikančios sistemos remiasi tuo, kad jose nėra nei parametrinio modelio, nei kitos išankstinės informacijos apie garso šaltinius. Garso šaltinių signalai yra nustatomi iš duomenų. Realus laiko signalams tokie nepriklausomų komponentų analizės būdai nėra pilnai tinkami, nes iš gaunamos informacijos be apribojimų šiuo metu nėra įmanoma atskirti muzikos garsus. Tačiau, garso šaltiniams pritaikius apribojimus, prie duomenų prisitaikančios technologijos gali būti panaudojamos realiais muzikos analizės atvejais. Tokie apribojimai, pavyzdžiui gali būti: šaltinių nepriklausomumas bei negausumas, kas reikš, kad šaltiniai didžiąją laiko dalį yra neaktyvūs.

Svarbiausi tyrimai:

Retojo programavimo paradigmai Virtanen [19] pridėjo laiko tęstinumo apribojimą. Jis naudojo signalų modelį, kuris atvaizdavo nulinės fazės galios įėjimo spektrogramą, kaip linijinę kintančių laike statinių šaltinių spektrų sumą. Algoritme atsižvelgta į klaidų spektrą, atstatymo klaidas, duomenų išsklaidymą bei laikinį tęstinumą. Algoritmas, taikomas išskirti toną turinčius muzikos instrumentus bei mušamuosius instrumentus iš realaus garso srauto.

Mark Plumbley, daugiausiai tyrinėjantis BSS, naudojant ICA, pasiūlė „neprižiūrimo mokymosi“ sistemą, paremtą garsinių spektrogramų išretintuoju programavimu. Daugeliu bruožų ji yra panaši į Virtanen sistemą. Plumbley priėjo išvados, kad sistema gali būti pritaikyta tam tikro tipo muzikai (net ir polifoninei) analizuoti, nenaudojant išankstinių duomenų. Tačiau sistema susidūrė su problemomis, iš kurių svarbiausios – garso signalų praradimas bei sunki analizė, jei pamatuotas išankstinis šaltinių pasiskirstymas nesutampa su tikruoju.

2.3. Fundamentaliųjų dažnių metodikų palyginimas

1 lentelė fundamentaliųjų dažnių metodikų tikslumo ir sudėtingumo palyginimas

Metodų grupės pavadinimas	Išankstiniai duomenys	Papildomi apribojimai	Sudėtingumas bei skaičiavimai	Analizės tikslumas*
Suvokiamojo pobūdžio dažnio dedamųjų grupavimas	Muzikinių tonų informacija, akordų informacija	Naudojami, norint supaprastinti skaičiavimus	Skaičiavimai sudėtingi, jų daug. Dažniausiai remiasi BSS, ICA metodais.	70-80%
Klausos modeliavimas	Informacija apie klausos modelio parametrus (aukščiausia/žemiausia girdimumo riba ir kt.)	Signalas periodinis, signalui dideli apribojimai. Dažniai – žemi ir vidutiniai.	Vidutiniškas (labai supaprastintas pagrindinis metodų grupės modelis) Vidutiniški	62-100% priklausomai nuo duomenų
Blackboard“ („lentos“) architektūra	Muzikinių taisyklių žinios, norimas hierarchijos lygių skaičius	Garso šaltinių skaičiaus apribojimai	Sudėtinga muzikos dalių integracija/valdymas, pasikartojantis procesas, pati struktūra intuityvi Vidutiniški (programavimas remiasi „jei-tai“ tipo struktūra)	65% - 98% priklausomai nuo duomenų
Signalų modeliais paremti tikimybiniai modeliai	Reikalingi išsamūs išankstiniai duomenys parametrų. Atskiri instrumentų tonai turi būti žinomi.	Nuo parametrų kiekio priklauso sudėtingumas.	Labai didelis Daug skaičiavimų (parametrinės lygtys sudėtingos, reikalaujančios daug kompiuterio laiko)	80-90%
Prie duomenų prisitaikančios sistemos	Nereikalingi, kartais naudojami, norint patikslinti įvedimo signalą	Galimi dažniausi apribojimai: duomenų negausumas, ortogonalumas, nepriklausomumas neneigiamumas; laikinis muzikos srautų tęstinumas	Didelis sudėtingumas (reikia skaičiuoti tikimybinės funkcijas, naudoti matematinius modelius) Sudėtingi skaičiavimai	50-80%

* tikslumas matuojamas skirtingais metodais

Kaip matoma iš metodų grupių palyginimo, nėra vieno konkretaus modelio, visiškai teisingai atskiriančio garso srautą. Signalų modeliai tiksliausiai aprašo analizuojamus garsus, tačiau jie tinka tik tyrimams, nes realiam garso srautui apdoroti jie reikalauja per daug skaičiavimų. „Lentos“ architektūra privalo naudoti kitus modelius (dažniausiai perceptinius), kad būtų pasiekti geri rezultatai. Prie duomenų prisitaikančios sistemos tinka analizei, nes jos, nors ir turi apribojimų

įvedamam garso srautui, joms pateikiamų duomenų universalumas yra vienas iš pagrindinių programos reikalavimų. Klausos modeliavimas yra sąlyginai neištirta sritis, tačiau sistemos savybės galima panaudoti pradinės informacijos apie garso srautą sukonkretinimui. Taip pat tinkamas būtų girdimosios aplinkos analizės modelis, nes jis turi labai platų taikymo sričių sąrašą bei yra nedideli apribojimai įeinančiam garso srautui.

2.4. Garso šaltinių išskyrimo metodai

Garsų srauto skaidymo į atskirus garso šaltinius sprendimai yra paremti matematiniais metodais. Skyriuje apžvelgiami populiariausi egzistuojantys garso šaltinių atskyrimo metodai (metodų grupės). Konkretūs matematiniai algoritmai, taikomi garso šaltinių atskyrimui šiame skyriuje nebus nagrinėjami.

Apžvelgiami matematiniai metodai aprašo atskiras dažnių analizės metodikų dalis.

2.4.1. Aklas šaltinių išskyrimas

Aklas šaltinių išskyrimas (BSS), besiremiantis atskirų komponentų analize (ICA) yra labiausiai paplitęs garso srauto skaidymo metodas. Jis dažnai yra taikomas ne tik muzikos analizei, bet ir balso atpažinimui, telekomunikacijose, medicinoje. Tai yra universaliausiais metodas šaltinių išskyrimui, nes jis neturi tokių didelių apribojimų garso srautui, kaip kiti metodai.

Pagrindinė aklo šaltinių išskyrimo metodo prielaida yra ta, kad šaltiniai nėra dubliuojami. Turint vieną signalų (garsų) rinkinį yra bandoma išgauti kitą signalų rinkinį taip, kad gautų šaltinių vienodumas ir reguliarumas būtų maksimalus, o vienodumas tarp skirtingų šaltinių būtų minimalus. T.y. kad šaltiniai būtų kuo labiau statistiškai nepriklausomi.

Aklo šaltinių išskyrimo metodikos kryptys pagal [1] yra šios:

- Pagrindinio elemento analizė
- Vienetinės reikšmės dekompozicija
- Nepriklausomų komponentų analizė
- Priklausomų komponentų analizė
- Neneigiamų matricių faktorizavimas

Taip pat šis metodas yra pilniausiai matematiškai pagrįstas. Šis metodas remiasi principu atskirti „n“ skaičių šaltinių iš „m“ skaičiaus garso srautų. Metodas vadinamas gerai apibrėžtu, kai stebimų garso srautų skaičius viršija išskiriamų šaltinių skaičių. Šiuo atveju pakanka informacijos, kad būtų galima iš stebimų garso srautų išskirti garso šaltinius. Tačiau kai stebimų garso srautų skaičius yra mažesnis už garso šaltinių skaičių (o muzikos įrašuose dažniausiai stebėjimas/užrašymas yra vienas), reikalingi papildomi matematiniai metodai, modeliuojantys išankstinę informaciją apie

garsą. Dažniausiai naudojami Bajeso teorema, EM algoritmu, MCMC modeliavimu. Pažymėtinas yra DUET algoritmas ir jo papildymas DASSS.

DUET sistema pirmą kartą aprašyta Scott Rickard et al. Tai dviejų kanalų (dešiniojo ir kairiojo – palyginimui su stereo garsu) „n“ garso šaltinių atskyrimo sistema, kur „n“>2. Ši sritis priklauso Bajeso dviejų kanalų modeliavimo problemai. Metodas reikalauja, kad šaltiniai būtų išretinti. Egzistuoja pristatyti patobulinimai šiam metodui – tai DASSS sistema, kuri pagerina kanalų vėlinimo bei skaičiavimo sistemos nustatymą. (Pagal [19],[20])

2.4.2. Retasis programavimas, neneigiamų matricių faktorizavimas

Retasis programavimas garso srauto atvaizdavimui naudojant neneigiamų reikšmių matricas, aprašo garso srauto apdorojimo būdus, kai srautas yra „išretintas“. Tai yra, kai signalų modeliui reikalingų duomenų kiekis nėra pilnas ir reikia rasti mažiausią stebimų duomenų sąrašą, kuris pilnai aprašytų sistemą. Kaip rašoma [21], retasis programavimas, kaip ir ICA bei BSS, yra „prie duomenų prisitaikanti“ technika, kuri nereikalauja informacijos apie duomenų srautą.

Retasis programavimas yra naudojamas kaip BSS metodo dalis. Kiti šio metodo panaudojimo atvejai yra šablonų atpažinimas, neuroninių procesų modeliavimas. Jis yra naudingas signalų apdorojimui ir konkrečiai garso srauto skaidymui, nes tai būdas, kurį galima pritaikyti įvairių rūšių duomenims ir kuris nereikalauja sudėtingos matematinės struktūros.

2.4.3. Sinusoidinis modeliavimas, garso srauto aukščių analizė

Sinusoidinis modeliavimas – modeliavimas, paremtas garso srauto skaidymu į sinusinius signalus. Šis būdas remiasi prielaida, kad įeinantį signalą galima išskirstyti į begalinį skaičių skirtingo dažnio bei fazės komponentų, kurie būtų sinuso funkcijos formos. Egzistuojantys šios srities tyrimai atsižvelgia į garsų nevienalaikiškumą, tačiau tokiam modeliavimui yra taikomi griežti apribojimai (pvz.: fundamentalieji dažniai negali sutapti, reikalingi stabilūs periodiniai signalai). Sinusoidinis modeliavimas muzikos srityje yra dažniausiai taikomas harmoninių garsų išskyrimui (konkrečiai – muzikinių akordų išskaidymui) trumpais laiko periodais.

3. GARSO ŠALTINIŲ IŠSKYRIMO SISTEMOS PROJEKTAS

Viena iš aklo šaltinių išskyrimo metodo krypčių yra nepriklausomų komponentų analizė. Projekte realizuotas vienas iš garso srauto analizės metodų – „Hard Lost“, paremtas nepriklausomų komponentų analize. Tai – pagrindinė sukurtos „SSS“ sistemos dalis, kuri buvo papildyta pagalbinais moduliais, leidžiančiais sistemą naudoti muzikinėje arba garsų analizės srityje.

„SSS“ sistema leidžia, įrašius įrašą, arba pateikus jį vienu iš sistemos atpažįstamų garso failų formatų („wav“, „mp3“, „mpeg“), konvertuoti failą į reikalingą formatą (nesuspaustą Windows garso failų formatą – „PCM wav“), analizuoti garso failą, naudojant garso signalų apdorojimo technologijas bei išskirti jame garso šaltinius, pavaizduoti rezultatus atskirais garso takeliais, saugoti bei eksportuoti balsų išskyrimo rezultatus. Papildomos sistemos funkcijos yra sistemos atributų keitimas bei vartotojo pagalba.

Sukurta „SSS“ sistema orientuota į:

- Pageidautinų garso srauto analizę ir garso šaltinių išskyrimą;
- Galimą muzikos transkripcijos programų išplėtimą, tobulinimą;
- Muzikinės klausos treniravimo veiklą.

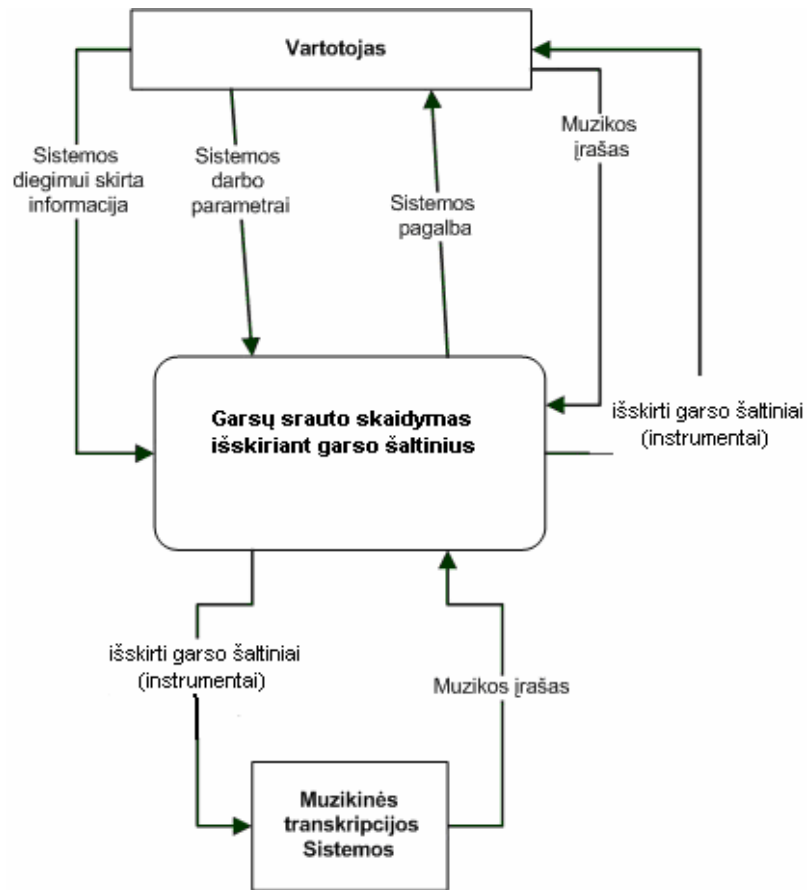
Tolimesniuose skyriuose yra apibendrintai vaizduojamas ir aprašomas „SSS“ sistemos projektas. Detalų projekto aprašymą galima rasti „SSS“ sistemos dokumentuose [24], [25].

3.1. „SSS“ sistemai keliami reikalavimai

Skyriuje apžvelgiami sistemai kelti reikalavimai. Vaizduojama sistema iš vartotojo pusės. Trumpai aprašoma sistemos aplinka, kuri yra pateikiama sistemos konteksto diagrama. Toliau yra pateikiami funkciniai ir nefunkciniai „SSS“ sistemos reikalavimų sąrašai. Akcentuojami tie reikalavimai, kurie yra tiriami darbe.

„SSS“ Sistemos konteksto diagrama vaizduoja „SSS“ sistemos garsų srauto skaidymo veiklos sąryšį su kitomis sistemomis – vartotoju ir muzikinės transkripcijos sistemomis. Čia nubrėžiamos veiklos, kuriomis išorinės sistemos siejasi su nagrinėjama veikla.

„SSS“ sistemos konteksto diagrama (5 pav.):



5 pav. „SSS“ sistemos konteksto diagrama

2 lentelė „SSS“ sistemos konteksto diagramos veiklų aprašymas

Eil. Nr.	Įvykio pavadinimas	Įeinantys/išeinantys informacijos srautai
1	Vartotojas diegia sistemą	Sistemos diegimui skirta informacija (in)
2	Vartotojas nustato įrašymo, techninių įrenginių bei kitus parametrus	Sistemos darbo parametrai (in)
3	Vartotojas gauna informaciją apie sistemos darbą, jam rūpimus klausimus	Sistemos pagalba (out)
4	Vartotojas įrašo bei pateikia garso įrašą, skirtą garso analizei. (Taip pat muzikinė transkripcijos sistema pateikia duomenis garso šaltinių išskiriami.)	Garso įrašas(in)
5	Vartotojas arba muzikinės transkripcijos sistemos pareikalauja garso šaltinių išskyrimo daugiabalsėje muzikoje rezultatų	Išskirti garso šaltiniai (out)

3.1.1. „SSS“ sistemos funkcinių reikalavimų sąrašas

Sistemos funkcinių reikalavimų sąrašas nurodo „SSS“ sistemos galimybes. Čia pateikiamos sistemos atliekamos funkcijos. Plačiau apie sistemos funkcinius reikalavimus galima skaityti [24].

- Sistema turi išskirti bent du garso šaltinius vienbalsėje muzikoje
- Sistema privalo turėti įdiegimo programą
- Sistema privalo turėti diegimo dokumentus
- Sistema privalo turėti realiaje laike veikiančią muzikos įrašymą
- Sistema privalo turėti aplinkos parametrų keitimo galimybę
- Sistema privalo turėti muzikos klausymo galimybę
- Sistema privalo turėti garso failo pateikimo garso šaltinių išskyrimui galimybę
- Sistema privalo turėti garso failų konvertavimo galimybę
- Sistema privalo turėti išskirtų garso šaltinių atskiro klausymo galimybę

3.1.2. „SSS“ sistemos nefunkcinių reikalavimų sąrašas

Sistemos nefunkcinių reikalavimų sąrašas nurodo nefunkcinius sistemos atributus. Du iš šių atributų yra tiriami, lyginant „SSS“ sistemą su „DUET“ algoritmu. Plačiau apie sistemos funkcinius reikalavimus galima skaityti [24].

- Sistemos standartiškumas
- Grafinė muzikos vaizdavimo sąsaja
- Naudojimo patogumas
- Sėkmingas sistemos panaudojimas be apmokymo naudotis sistema
- Išplečiamumas
- Veikimo greitis
- Garsų srauto analizės tikslumas
- Lengva programos priežiūra

Du pagrindiniai nefunkciniai reikalavimai keliami sistemai yra sistemos garsų srauto analizės tikslumas bei sistemos vartojimo patogumas. Šie sistemos reikalavimai yra analizuojami tiriamojoje darbo dalyje. Naudotas sistemos algoritmas lyginamas su etaloniniu laikomu „DUET“ algoritmu tikslumo bei sistemos greičio, kuris yra tiesiogiai susijęs su vartojimo patogumu, aspektais.

Tikslumas – sistema turi veikti kaip įmanoma tiksliau, kad galėtų konkuruoti su egzistuojančiais sprendimais. Taip pat šiame darbe buvo svarbu išsiaiškinti, kurios garso šaltinių išskyrimo metodikos yra daugiausiai žadančios ir turi tendencijų tobulėti ateityje. Šiam reikalavimui

išpildyti po metodikų analizės realizavimui bei tyrimui pasirinkta prie duomenų prisitaikančios sistemos metodika, naudojant vieną iš aklo šaltinių išskyrimo algoritmų.

Naudojimo patogumas – vartotojo neturi apsunkinti darbo aplinka, darbas su sistema turi būti lengvai išmokstamas, garso srauto analizės procesas turi taupyti laiką. Šis reikalavimas nurodo, kad reikia atsižvelgti į sistemos darbo greitį. Sistemos pagrindinė funkcija – garsų analizė – yra laikui imlus procesas. Dauguma sprendimų negali būti pritaikyti realiose situacijose būtent dėl per ilgai užtrunkančios garsų srauto analizės ir garso šaltinių išskyrimo.

3.2. „SSS“ sistemos architektūra

3.2.1. „SSS“ sistemos ribos – panaudos atvejų modelis

Sistemos panaudos atvejų modelis parodo, kaip „SSS“ sistema gali būti naudojama. Šiame modelyje akcentuojamas garso analizės panaudojimo atvejis. Jis yra pagrindinis sistemai. Kiti sistemos panaudojimo atvejai yra skirti sistemos pritaikymui patogiam vartotojo darbui.

„SSS“ sistemos panaudos atvejų ir aktorių sąrašai

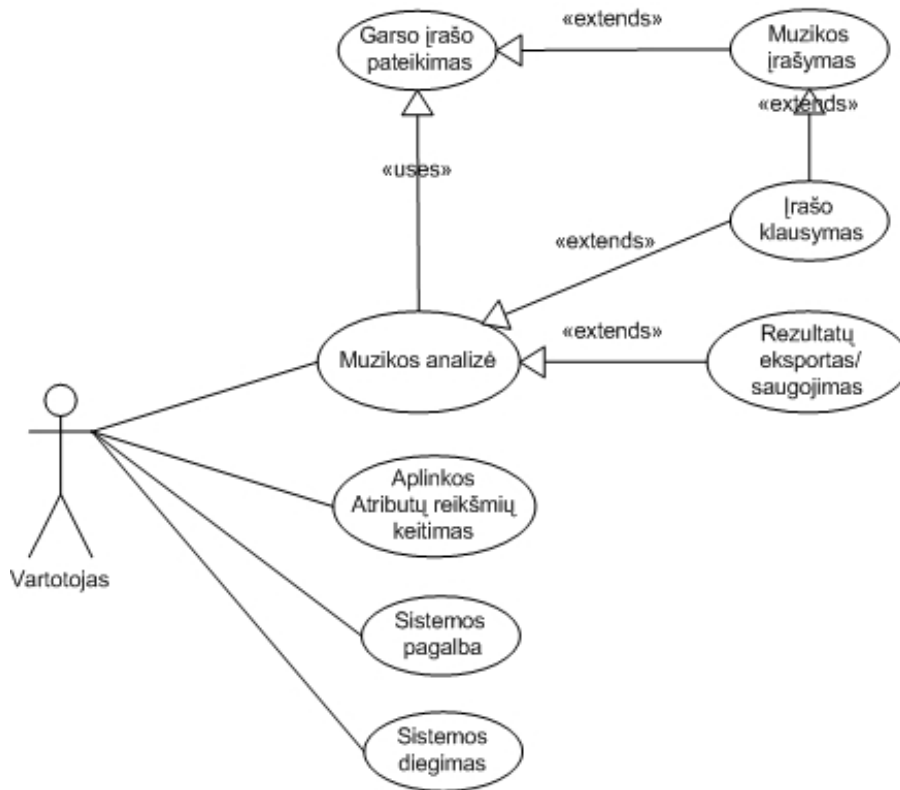
Aktoriai:

Sistemos vartotojas

Panaudos atvejai:

1. Garso įrašo pateikimas
2. Muzikos įrašymas
3. Įrašo klausymas
4. Muzikos (garsų srautų) analizė
5. Rezultatų eksportas/saugojimas
6. Aplinkos atributų reikšmių keitimas
7. Sistemos pagalba
8. Sistemos diegimas

„SSS“ sistemos panaudos atvejų diagrama



6 pav. „SSS“ sistemos panaudojimo atvejų diagrama

Pateikiama garsų srauto (muzikos) analizės panaudos atvejo specifikacija:

- *Tikslas ir laukiami rezultatai:* Atlikti pagrindinį programos darbą – išskirti garso šaltinius daugiabalsėje solinėje muzikoje. Rezultatas –kuo tiksliau išanalizuotas signalas, atvaizduotas ekrane.
- *Dalyvaujantys aktoriai:* vartotojas.
- *Ryšiai su kitais panaudojimo atvejais:*
 - Garso įrašo pateikimas
 - Įrašo klausymas
 - Rezultatų eksportas/saugojimas
- *Nefunkciniai reikalavimai:* Programa muzikos analizę turi atlikti kaip įmanoma greičiau (turi neviršyti 30 sekundžių, vienos minutės ilgio įrašui). Taip pat turi būti kaip įmanoma daugiau automatizuota.
- *Prieš-sąlygos:*
 - Programa turi būti įjungta.

- Turi būti pasirinktas įrašas, skirtas muzikos analizei.
- *Sužadavimo sąlyga*: Paspaudžiamas mygtukas/pasirenkamas meniu punktas muzikos analizei.
- *Po-sąlyga*: išanalizuota muzika, atvaizduota ekrane.
- *Pagrindinis scenarijus*: Vartotojas, pasirinkęs muzikos įrašą arba įrašęs savo, paspaudžia muzikos analizės mygtuką, pasirenka muzikos analizės parametrus, sulaukia, kol procesas bus atliktas (tą indikuoja proceso progreso matuoklis matomas ekrane). Vartotojas gali toliau tęsti darbą (saugoti projektą...).
- *Alternatyvūs scenarijai*:
 - Vartotojas pasirenka muzikos analizę, nepasirinkęs muzikos įrašo – programa prašo pasirinkti įrašą.
 - Vartotojas viduryje proceso nusprendžia nutraukti analizės procesą (paspausdamas mygtuką nutraukti).

Plačiau apie sistemos sritį bei kitus sistemos panaudojimo atvejus skaityti [24].

3.2.2. „SSS“ sistemos statinė ir dinaminė struktūra

Šiame skyriuje vaizduojama „SSS“ sistemos statinė ir dinaminė struktūra.

Statinės struktūros vaizdavimui pateikiamas sistemos išskaidymas į loginius paketus ir juos sudarančias klases.

Funkcinis išskaidymas į paketus leidžia atskirus programos modulius pakartotinai panaudoti kitose (muzikos transkripcijos) sistemose. Pasirinkta sistemos architektūra leidžia nesunkiai praplėsti sistemą naujomis funkcijomis, nekeičiant visos sistemos architektūros bei esamų ryšių tarp paketų/klasių.

Programos bendro valdymo modulis užtikrina sistemos vientisumą. Keičiamų sistemos parametrų galimybės leidžia pritaikyti sistemą skirtingiems vartotojų poreikiams.

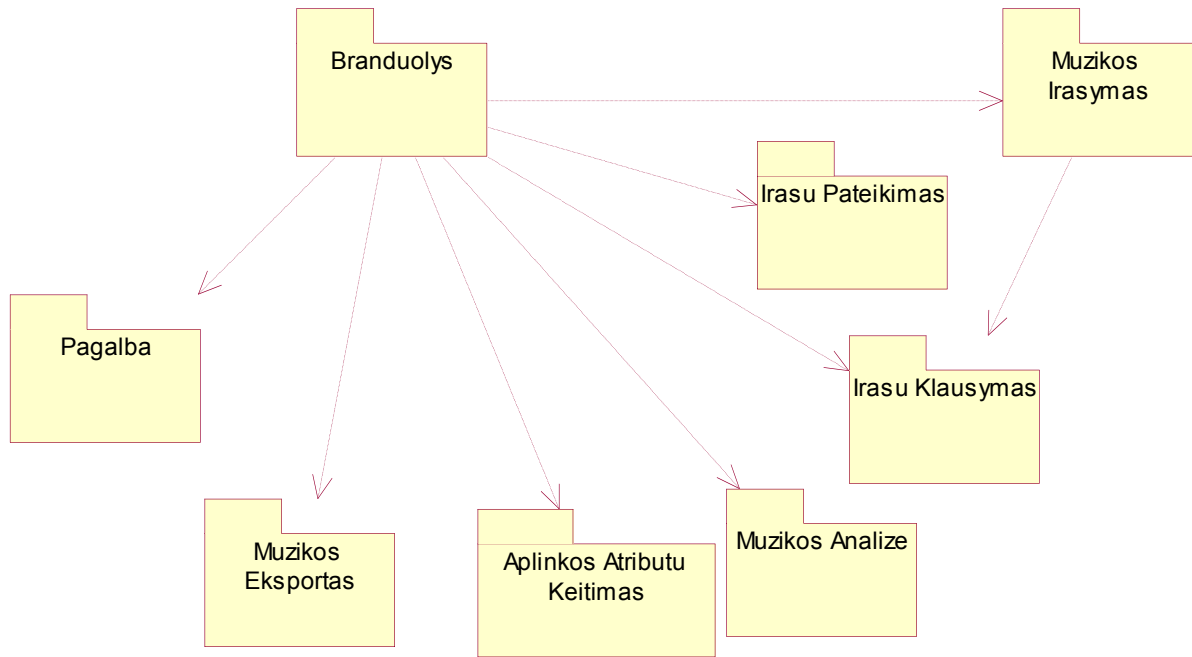
Taip pat pateikiamos „SSS“ sistemos garsų analizės modulio diagramos ir pagrindinio sistemos objekto – garso failo būsenų diagrama. Šios diagramos vaizduoja sistemos dinaminę dalį, susijusią su garso srauto analize.

Sistemos architektūros aprašymas nėra susietas su konkrečia programavimo kalba. Projektuojant sistemą buvo naudojamos Rational Rose priemonės, kad būtų galima palengvinti sistemos programavimą Visual C++ kalba, nes Rational Rose turi automatinio kodo generavimo šioje kalboje galimybę.

Sistemos architektūra užtikrina sistemos nepriklausomumą nuo kitų sistemų (išskyrus Windows operacinę sistemą).

Pateikiamas statinis sistemą sudarančių paketų vaizdas ir „SSS“ sistemos diagramos, skirtos garsų analizės moduliui.

„SSS“ sistemos statinė struktūra – sistemą sudarančios dalys



7 pav. „SSS“ sistemos paketų diagrama

Sistema susideda iš tokių paketų:

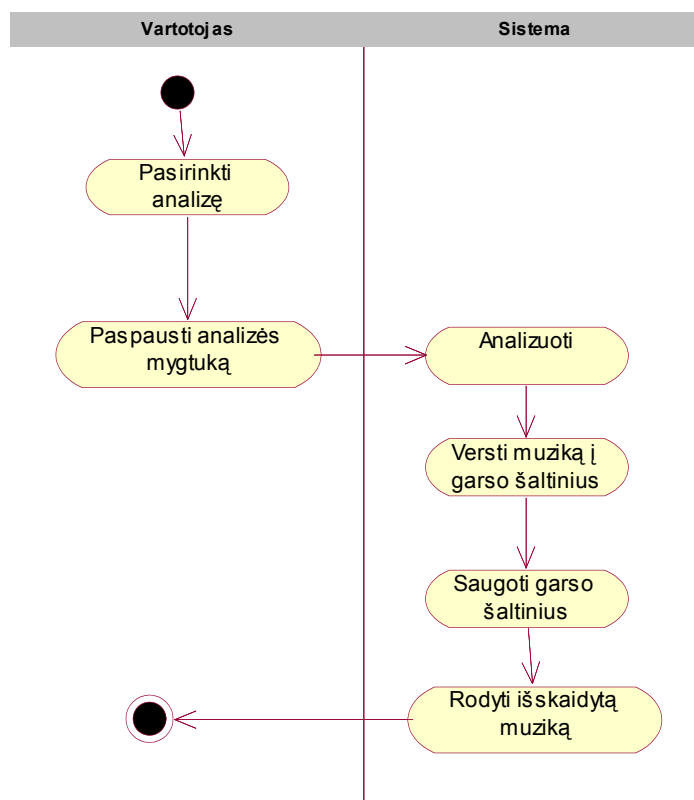
- Branduolys – paketas, valdantis bendrus programos duomenis bei objektus. Šio paketo klasės valdo kitų paketų sąsajų kūrimą. Paketas sąveikauja su visais kitais paketais: Muzikos įrašymas, Įrašų pateikimas, Įrašų klausymas, Muzikos analizė, Muzikos eksportas, Aplinkos atributų keitimas, Pagalba.
- Muzikos įrašymas – paketas, skirtas realaus laiko muzikos įrašymui, naudojant prie kompiuterio prijungtą mikrofoną. Paketas atsakingas už muzikos sąsajos vaizdavimą, įrašymo valdymą, darbinio įrašo saugojimą bei įrašymo parametrų pritaikymą.
- Įrašų pateikimas – paketas, skirtas vartotojo pateikiamiems muzikos įrašams apdoroti bei įkelti į programą. Paketas yra atsakingas už įrašų pateikimo valdymą: konvertavimą į darbinį formatą, tikrinimą, ar galima įkelti įrašą, bei pati įrašo įkėlimą.
- Įrašų klausymas – paketas, skirtas muzikos įrašų klausymo valdymui. Pakete valdomas įrašų perklausymas: įkeltas įrašas klausomas, sustabdomas, vaizduojamas grafiškai.
- Muzikos analizė – paketas, skirtas garsų srauto analizei bei garsų srauto skaidymui į atskirus garso šaltinius.
- Muzikos eksportas – paketas, atsakingas už darbinio įrašo bei išskirtų balsų konvertavimą į vartotojo pageidaujamą muzikos formatą bei jų saugojimą vartotojo kompiuteryje.
- Aplinkos atributų keitimas – paketas, skirtas sistemos aplinkos pritaikymui skirtingiems vartotojams. Paketas valdo aplinkos atributų reikšmių keitimą bei saugojimą.

- Pagalba - paketas, skirtas suteikti vartotojui pagalbą, susijusią su naudojama sistema. Pagalba susideda iš programinės įrangos aprašymo, temų žinyno, pradžiamokslis (kaip pradėti dirbti su sistema). Paketas valdo šios informacijos vaizdavimą, paiešką pagal temas.

„SSS“ sistemos dinaminis vaizdas – garsų analizės modulio veiklos diagrama

Garsų srauto analizė atliekama automatiškai, vartotojui pasirinkus garsų srauto analizės parametrus. Analizė pradedama, vartotojui paspaudus garsų analizės mygtuką. Sistema atlieka garso failo nuskaitymo, vertimo į dažnių erdvę, garso analizės iš skaidymo ir vertimo atgal į laiko erdvę veiksmus. Po atliktų veiksmų yra pateikiamas išanalizuotas garso failas vartotojui.

Atlikus garsų analizę yra galimybės kitoms veikloms – išskirtų garso šaltinių klausymui kartu ir atskirai; šaltinių saugojimui ir vertimui į kitus garso failų formatus.

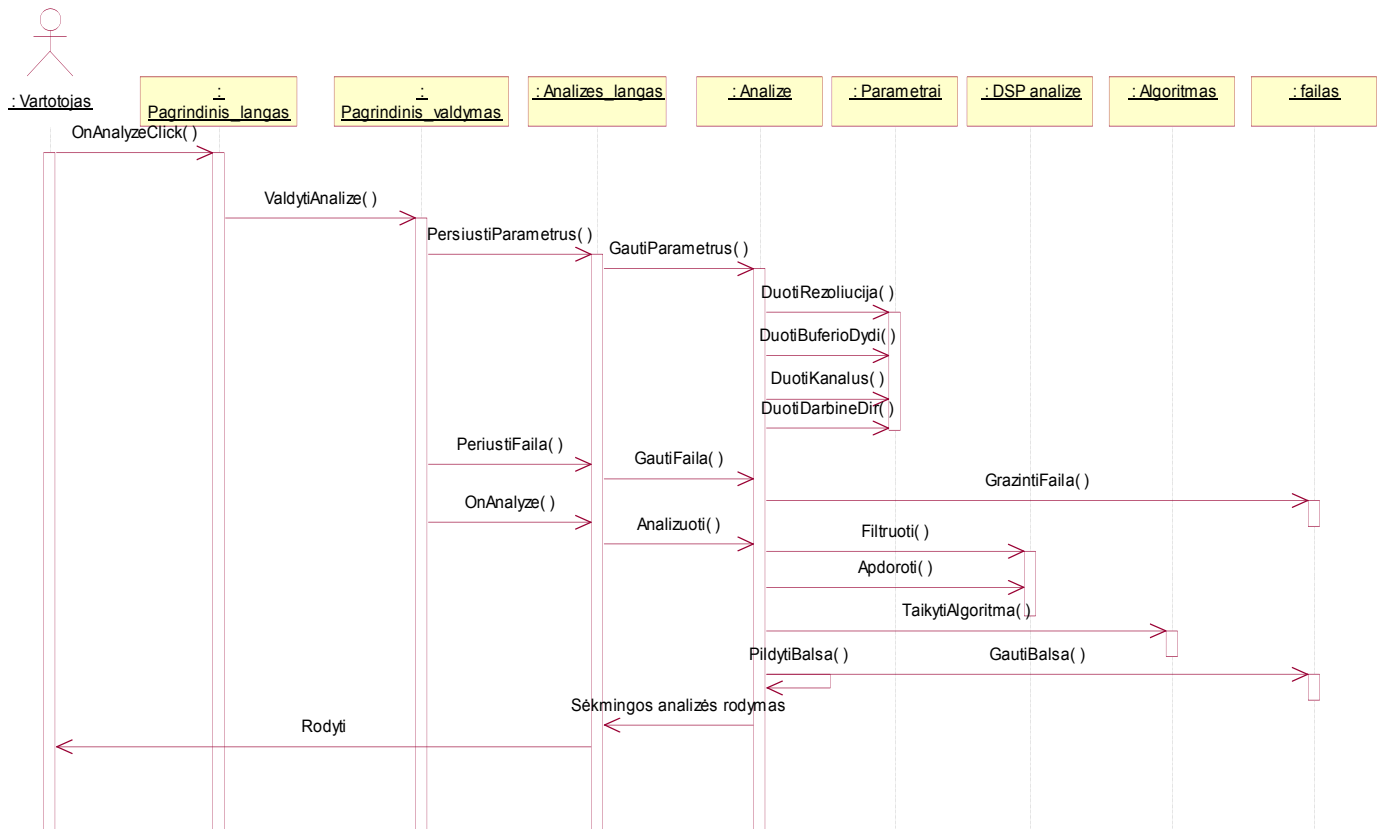


8 pav. Garsų analizės modulio veiklos diagrama

„SSS“ sistemos dinaminis vaizdas – garsų analizės modulio sekų diagrama

Garsų analizės modulio sekų diagramoje yra vaizduojamas sistemoje atliekamų veiksmų vaizdas, vartotojui pasirinkus garsų analizės veiklą. Diagramoje vaizduojama vartotojo sąveika su sistema garsų analizės metu. Kadangi garsų srauto analizė yra automatizuotas procesas, vartotojas čia dalyvauja tik kaip proceso iniciatorius.

Sekų diagramoje matomas sistemos architektūros išskaidymo į tris lygius vaizdas: vartotojas su sistema bendrauja per sąsajas, kurios veiksmus perduoda valdymo klasėms. Valdymo klasės valdo procedūras ir kitų sąsajų iškvietimą. Taip pat valdymo klasės duomenis įrašo į duomenims skirtą klasę – pagrindinį programos objektą – garso failą.

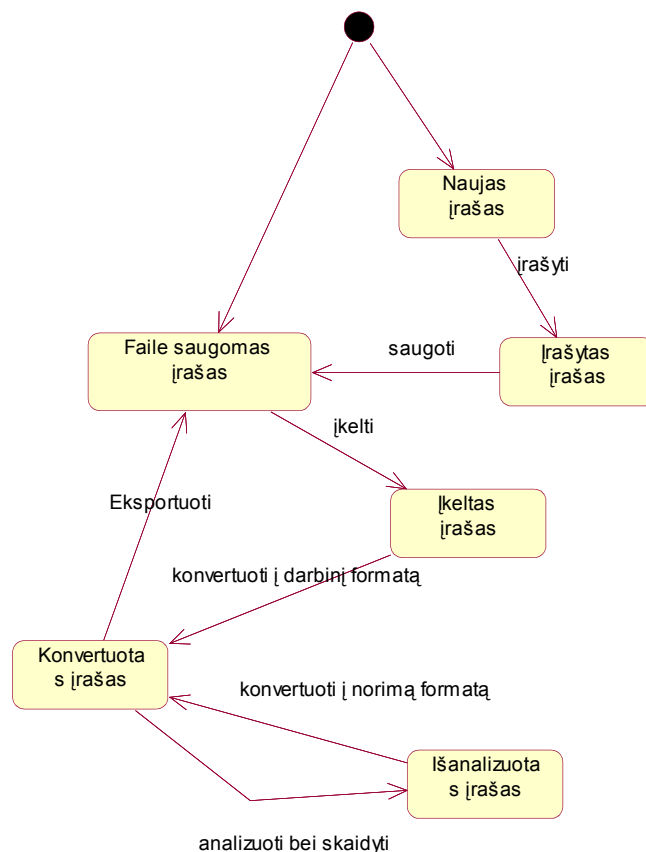


9 pav. Garsų analizės modulio sekų diagrama

„SSS“ sistemos dinaminis vaizdas – objekto „garso failas“ būsenų diagrama

Pagrindinis „SSS“ sistemos naudojamas objektas yra garso failas. Be šio objekto sistemos galimos veiklos yra tik sistemos pagalba bei aplinkos atributų reikšmių keitimas.

Garso failas įsijungus programą yra pateikiamas analizei dviem būdais: įrašant naują garso failą arba pateikiant seniau išsaugotus garso failus. Pateikiant garso failą yra reikalinga, kad jis būtų nesuspausto Windows PCM „wav“ formato. Todėl yra atliekamas failo konvertavimas į darbinį formatą. Tokios būsenos failą galima analizuoti. Tačiau, kaip matoma iš diagramos, konvertuotą garso failą taip pat galima išsaugoti neatlikus garso srauto analizės. Tai dar vienas sistemos pritaikymo būdas, nesusijęs su pagrindine „SSS“ sistemos atliekama veikla. Išanalizuotas įrašas taip pat gali būti konvertuotas į pageidaujimą failų formatą bei išsaugotas.



10 pav. Garso failo objekto būsenų diagrama

3.3. „SSS“ sistemos projekto apibendrinimas

Sukurta sistema leidžia vartotojui patogiai valdyti automatizuotą garso šaltinių išskyrimo procesą (detalesniam susipažinimui su sukurta sistema skaityti priedą “SSS’ sistemos funkcinis aprašymas”). „SSS“ sistema nuo kitų skiriasi tuo, kad kitos garso šaltinių išskyrimo programos yra

daugiau tiriamojo, o ne taikomojo pobūdžio. Jos nepritaikytos bendram naudojimui be išankstinių žinių apie sistemą ir reikalauja specialaus pasirengimo naudoti sistemą (pavyzdžiui, reikalauja pateikti tik tam tikro tipo failus, įdiegti papildomas programas, rezultatai nėra išskirti garso failai, o tik faktas apie išskyrimą). Tuo tarpu „SSS“ sistema net nepatyrusiam vartotojui leidžia išbandyti garso šaltinių išskyrimą iš garso srauto ir rezultatus panaudoti vėlesnei garsų analizei.

Sistema suskaidyta į modulius taip, kad kiekvienas modulis yra nepriklausoma dalis, kurią galima panaudoti atskirai. Tai yra aktualu norint sistemą praplėsti naujomis galimybėmis arba panaudoti kitose egzistuojančiose ar kuriamose sistemose. Vienas iš galimų dalies sistemos panaudojimo atvejų yra garso šaltinių išskyrimo modulio pritaikymas muzikinės transkripcijos sistemose. Taip pat dėl naudojamo algoritmo specifikos, garso srauto analizės modulį galima naudoti kitose srityse – medicinoje ar telekomunikacijoje.

4. „SSS“ SISTEMOS KOKYBĖS TYRIMAS

Tiriant sukurtos sistemos kokybę buvo atliktas sistemos modulių bei sistemos integracijos testavimas. Taip pat sistema buvo įdiegta ir jos veikimas testuotas 5 vartotojų kompiuteriuose.

Po sistemos naudojimo vartotojai užpildė anketas, skirtas sistemos nefunkcinių reikalavimų įvertinimui. Buvo registruotas apsimokymo dirbti su sistema laikas (50% sėkmingas pasinaudojimas sistema be apsimokymo, kitiems atvejams teko pasinaudoti sistemos pagalba), sistemos diegimo/šalinimo lengvumas (100% sėkmingas sistemos diegimas ir pašalinimas, kai vartotojas turėjo darbo su Windows sistema pagrindus).

Išsiaiškinta ir užrašyta vartotojų nuomonė apie sistemos paprastumą (standartiškumą) bei patogų valdymą (60% vartotojų laiko sistemą patogiai vartoti).

Buvo atlikta sistemos programinio kodo analizė, pasitelkiant Mats Henricson ir Erik Nyquist taisykles ir rekomendacijas C++ kodo rašymui. Plačiau su taisyklėmis susipažinti galima [27].

Buvo tikrinami tokie kodo rašymo aspektai:

- Vardai
- Stilius
 - Klasės
 - Funkcijos
 - Sąlygos ir ciklo sakiniai
 - Rodyklės
 - Konstantos
 - Atminties valdymas
- Kodo portatyvumas
- Bendros taisyklės

Sistemoje kodo rašymo požiūriu rasta klaidų. Daugiausia klaidų rasta atminties valdymo dalyje ir kodo portatyvumo užtikrinime. Mažiausiai klaidų buvo aptikta vardų naudojimo kode ir kodo rašymo stiliaus dalyse.

Rastos klaidos gali atsiliiepti sistemos kokybei, nes garsų srauto analizė yra procesas, reikalaujantis daug resursų, o neefektyviai išnaudojant atmintį, šis procesas gali tapti nevaldomas. Galimi klaidų taisymo būdai būtų:

- Egzistuojančioje sistemoje nustatyti vienodą atminties valdymo būdą. Programoje naudojamos išorinių matematinių bibliotekų funkcijos turi savo atminties valdymą. Šis atminties valdymas nėra suvienodintas.

- Perrašyti sistemą programavimo kalba, kuri turi įdiegtą automatinį atminties valdymą.

„SSS“ sistema buvo palyginta su artimiausiomis galimomis programomis – muzikinės transkripcijos sistemomis. Lyginama ne su garso šaltinių išskyrimo programomis, nes šių programų komercinių arba laisvo platinimo, pritaikytų vartotojų darbui nėra. Palyginimas darytas pagal programinės įrangos kokybės vertinimo standartą ISO/IEC 9126, atmetant punktą „Tikslumas“, nes šis punktas nėra tinkamas šių sistemų lyginimui ir yra aprašytas prie „Hard - Lost“ ir „DUET“ algoritmų analizės dalyje.

Vertinimo metrikos yra apibrėžiamos skalėje: „Nepatenkinamas“ – „Patentkinamas“ – „Geras“ – „Puikus“ vertinimas konkrečiam kriterijui.

3 lentelė Bendras sistemos vertinimas pagal ISO/IEC 9126, lyginant su muzikos transkripcijos sistemomis

Charakteristikos Nr.	Parametras	Transkripcijos programa „IntelliScore Polyphonic Wav To Midi Converter“ vertinimas	Transkripcijos programa „Akoff music composer“ vertinimas	„SSS“projekto vertinimas
1. Funkcionalumas	Tinkamumas	Patentkinamas	Patentkinamas	Geras
	Sąsajos su kitomis programomis	Yra	Nėra	Nėra, gali būti pritaikoma kaip komponentas
	Suderinamumas	Geras	Geras	Geras
2. Patikimumas	Branda	Geras	Geras	Patentkinamas
	Klaidų tolerancija	Geras	Geras	Geras
	Atstatymas po klaidos	Nėra	Nėra	Nėra
3. Panaudojamumas	Suvokiamumas	Patentkinamas	Puikus	Puikus
	Išmokstamumas	Geras	Geras	Geras
	Valdymas	Patentkinamas	Geras	Patentkinamas
4. Efektyvumas	Laiko (negalima palyginti)	Geras (transkripcijos srityje)	Geras (transkripcijos srityje)	Patentkinamas (balsų išskyrimo srityje)
	Resursų	Geras	Patentkinamas	Patentkinamas
5. Palaikomumas	Analizuojamumas	Neaišku, programa komercinė	Neaišku, programa komercinė	Geras
	Keičiamumas	Neaišku, programa komercinė	Neaišku, programa komercinė	Geras
	Stabilumas	Neaišku, programa komercinė	Neaišku, programa komercinė	Patentkinamas

	Testuojamumas	Neaišku, programa komercinė	Neaišku, programa komercinė	Geras
6. Portatyvumas	Pritaikymas kitose aplinkose	Patenkinamas	Patenkinamas	Patenkinamas
	Įdiegimo lengvumas	Puikus	Puikus	Puikus
	Portatyvumo standartai	Geras	Patenkinamas	Patenkinamas
	Pakeičiamumas	Patenkinamas	Geras	Geras

5. „DUET“ IR „HARD-LOST“ ALGORITMŲ TYRIMAS

„DUET“ ir „Hard – Lost“ algoritmai yra vieno metodo – aklo altinių išskyrimo – skirtingi įgyvendinimai. Jie remiasi skirtingomis koncepcijomis. „DUET“ algoritmas skaičiuoja garso slopimą ir vėlinimą. „Hard - Lost“ algoritmas iš turimų duomenų, naudodamas atsiktiniu būdu sugeneruotus galimus sprendimo variantus, bando juos patikslinti ir atstatyti numanomą garso šaltinį.

Abu algoritmai sugeba išskirti daugiau nei du garso šaltinius iš dviejų garso srautų. Tačiau „DUET“ nekeistas algoritmas visada naudoja tik du garso srautus. Todėl darbe tiriamas tik dviejų garso srautų rezultatų palyginimas. Buvo atlikti bandymai „Hard - Lost“ algoritmui, naudojant daugiau garso srautų, tačiau šie rezultatai lyginant neįtraukti.

5.1. „DUET“ algoritmas

„DUET“ algoritmas, pasiūlytas A. Jourjine, Rickard S. et al. buvo pirmasis algoritmas naudojantis, pakeisto (pablogėjusio) garso srauto atstatymą, remiantis garso amplitudžių ir vėlinimo palyginimu. Šis algoritmas yra iš nedaugelio, sugebančių efektyviai atskirti garso šaltinius, turint tik du garso srautus. Algoritmas yra aktyviai tiriamas ir tobulinamas. Yra pasiūlyta jo praplėtimo galimybė DASS. Taip pat yra pasiūlyta algoritmo praplėtimo galimybė, kad būtų naudojami daugiau nei du garso srautai.

5.1.1. „DUET“ algoritmo koncepcija

Bendrai garso šaltinių maišymo koncepcija „DUET“ atveju atrodo taip:

$$x_1(t) = \sum_1^N s_j(t) \quad (1)$$

$$x_2(t) = \sum_1^N a_j s_j(t - \delta_j) \quad (2)$$

$x_j(t)$ – garso srautas laiko momentu t .

N – garso šaltinių skaičius

$s_j(t)$ – garso šaltinis laiko momentu t

a – garso šaltinio amplitudė (slopinimas)

δ - garso vėlinimo koeficientas

Atlikus greitą Furjė transformaciją (FFT) yra gaunamos formulės išraiškos dažnių erdvėje. Turint omenyje, kad laiko momentu “t” aktyvus vienas dažnio vienetas, iš lygčių sistemos galima suskaičiuoti amplitudę bei vėlinimą:

$$a_i = \left| \frac{x_2(\omega)}{x_1(\omega)} \right| \quad (3)$$

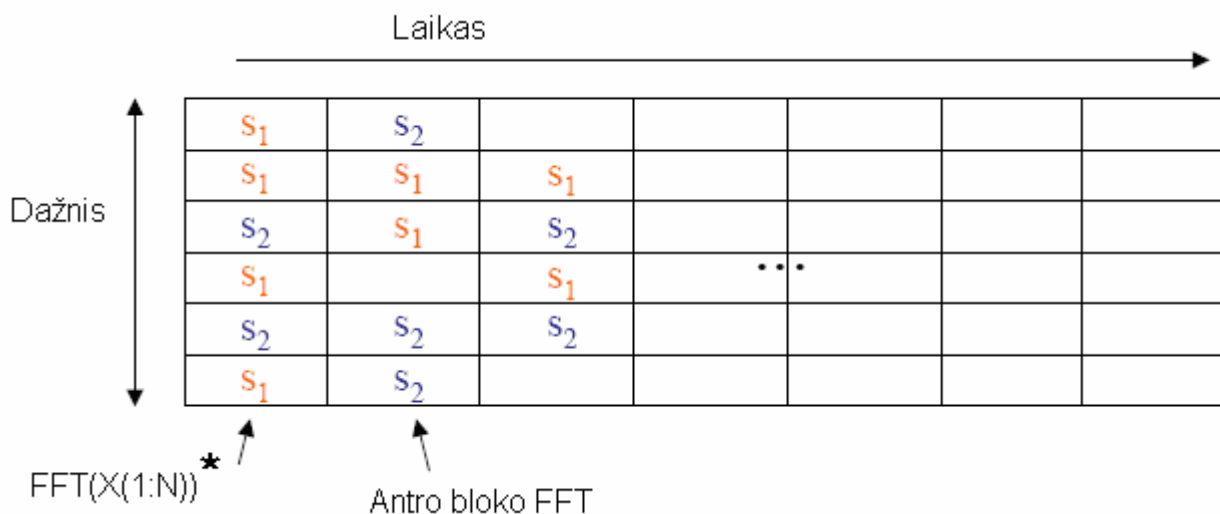
$$\delta_i = \frac{1}{\omega} \angle \left(\frac{x_1(\omega)}{x_2(\omega)} \right) \quad (4)$$

Algoritmas nėra universalus. Šio algoritmo taikymui reikalingi apribojimai:

- Vienu metu aktyvus tik vienas garso šaltinis pasirinktame dažnio diapazono vienetė
- Standartinis metodas naudoja du garso srautus
- Aplinka neturinti aido

5.1.2. „DUET“ algoritmo aprašymas

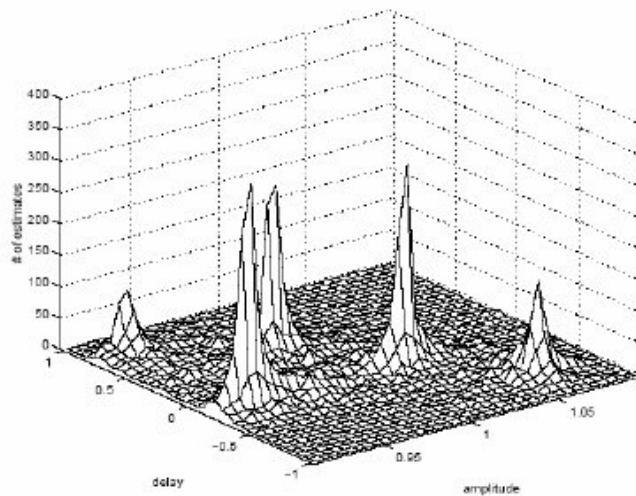
1. Garsų srautas suskaidomas į M blokų laiko skalėje.
2. Analizuojamas vienas blokas. Jis suskaidomas į N atskirų dažnio diapazono vienetų arba celių, kuriuose pagal apribojimą gali egzistuoti tik vienas garsas vienu metu (pavyzdyje s1 – garsas 1 arba s2 – garsas 2).



5 pav. „DUET“ algoritmo analizės blokų sudarymas

* FFT – atliekama vieno bloko greita Furjė transformacija

3. Nustatomas kiekvienos celės vėlinimas bei garso amplitudė.
4. Suformuojama dvimatė vėlinimo – amplitudės histograma, kuri nurodo pasikartojančio vėlinimo – amplitudės reikšmių skaičių kiekvienai celei N.



6 pav. „DUET“ algoritmo amplitudžių – vėlinimo histograma iš dviejų garso srautų, kuriuos kiekvieną sudaro tie patys 5 garso šaltiniai.

5. Kiekvienas dažnio komponentas yra priskiriamas garso šaltinio grupei pagal tai, kokia histogramos maksimumo reikšmė yra artimiausia šiam komponentui.
6. Norint atstatyti vieną garso šaltinį reikia palikti tik vieno garso šaltinio grupės dažnio komponentų reikšmes, kitų garso šaltinių grupių elementų reikšmes paverčiant 0.
7. Vienai garso šaltinio grupei atliekama atvirkštinė Furjė transformacija.
8. 6-7 žingsniai atliekami su visomis garso šaltinių grupėmis.

5.2. „Hard - Lost“ algoritmas

„Hard - Lost“ algoritmas, pasiūlytas Paul O’Grady ir Barak Pearlmutter [3] gali būti pritaikytas garso šaltinių išskyrimui, papildžius originalų algoritmą vertimo į dažnių erdvę ir iš dažnių erdvės dalimi. Taip pat šį algoritmą galima taikyti kitų šaltinių, pavyzdžiui, kelių sumaišytų paveikslų atskyrimui.

Algoritmas naudoja kelis šaltinių srautus. Konkrečiai garso šaltinių išskyrimui yra reikalingi bent du garso srautai. Algoritmas nenaudoja išankstinės informacijos apie šaltinius. Algoritmo naudojami parametrai yra tik apsimokymo greitis ir paketo, į kuriuos yra skaidomas bendras garso srautas atliekant Furjė transformaciją, dydis.

Algoritmas pasižymi tuo, kad juo galima skaidyti garso srautą į šaltinius, turint mažiau garso srautų nei juos sudaro garso šaltinių.

5.2.1. „Hard - Lost“ algoritmo koncepcija

Garso šaltinių išskyrimo problema yra vaizduojama kaip kryptinių linijų radimo problema duotoje duomenų aibėje.

Turėdami M garso šaltinių (pvz., keli muzikos instrumentai) ir N garso srautų (pvz., mikrofonai, išdėstyti kambaryje, įrašantys bendrą garsą) galime formuluoti tiesinio garsų išskyrimo uždavinį.

Egzistuoja N sensorių, „įrašančių“ $X = (x(1) | \dots | x(T))$ garso srautų, kurie yra tiesinis mišinys iš M garso šaltinių $S = (s(1) | \dots | s(T))$. Šaltiniai S mūsų nestebimame procese yra paverčiami garso srautais X . Procesas yra charakterizuojamas $N \times M$ maišymo matrica A , kuri yra nežinoma. Garso srautų įrašymas nėra visiškai tikslus. Jį iškreipia triukšmai ε .

$$x(t) = As(t) + \varepsilon(t). (5)$$

Jei įsivaizduosime, kad $N=M$, o triukšmas apytiksliai lygus nuliui, galime rasti atstatymo matricą W :

$$\hat{s}(t) = Wx(t).$$

Čia $W=A^{-1}$, o $\hat{s}(t)$ - garso šaltiniai laiko momentais t .

Nežinomieji yra maišymo matrica A bei garso šaltiniai S . Paprastai rasti šių dviejų nežinomųjų yra neįmanoma. Norint rasti maišymo matricą, reikia daryti apribojimą, kad vienu trumpu laiko momentu tik vienas šaltinis yra aktyvus. Šis apribojimas yra dažnai priimamas aklo šaltinių išskyrimo algoritmuose susijusiose su garsų išskyrimu, kai garso duomenys iš laiko erdvės yra verčiami į dažnių erdvę ir įgauna išskaidytą vaizdavimą.

Jei duota maišymo matrica

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} (6)$$

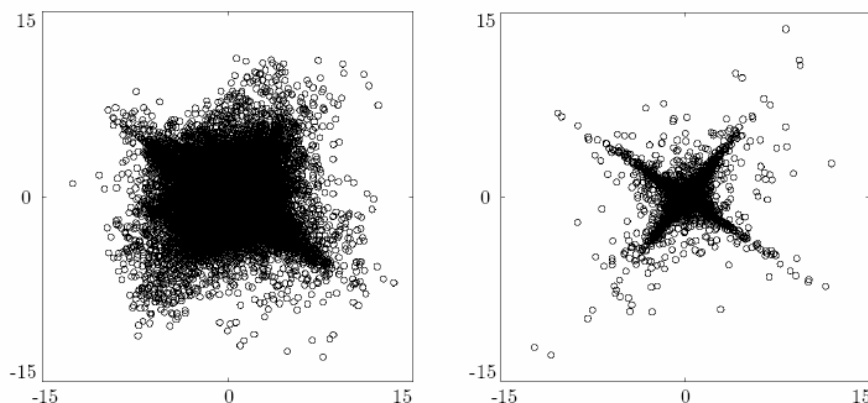
ir tik vienas šaltinis yra aktyvus vienu metu (pvz., s_1), sumaišyti garso srautai bus

$$x(t) = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} s_1(t) (7)$$

$x_1(t)$ ir $x_2(t)$ turės tą patį centrą ir skirtingas kryptis, kurias nusakys $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$.

Kai trumpais laiko momentais aktyvus tik vienas šaltinis, bendras vaizdas bus sudarytas iš tiesių, kurios yra prasiplėtusios dėl triukšmų ir kartkartėmis pasitaikančių veiksmų, vykstančių tuo pačiu metu. Šių tiesių kryptys atitinka A matricos stulpelius, todėl, jei yra įmanoma atstatyti iš turimų

duomenų tieses, galima sukonstruoti ir maišymo matricą A . Kai matrica A yra sukonstruota, galima išskirti garso šaltinius iš turimo garsų srauto.



7 pav. Dviejų ($N=2$) garso srautų X duomenų pavyzdys (kairėje) laiko erdvėje ir (dešinėje) išskaidytoje transformuotoje erdvėje, kurioje duomenys gauti, atlikus greitą Furjė transformaciją

5.2.2. „Hard-Lost“ algoritmo aprašymas

(pilną algoritmo aprašymą galima rasti [3].)

1. $N \times T$ duomenų matrica $X(t)$, sudaryta iš N garso srautų yra paverčiama į išskaidytą dažnių erdvę.
2. Randama maišymo matrica A (modifikuotas k -tųjų vidurkių algoritmas):
 - a. Atsitiktiniu būdu išdėstyti M (kiek turi būti šaltinių) linijų krypties vektorių
 - b. Priskirti kiekvieną nuskaitytų ir paverstų į išskaidytą atvaizdavimą duomenų dalį artimiausiai linijai
 - c. Perskaičiuoti linijų krypties vektorius naudojant stochastinį gradientų algoritmą (rekursinė funkcija)
 - d. Iš vektorių suformuoti maišymo matricą A .
3. Tiesių orientacijai (matricos A stulpeliai) yra priskiriami išmatuoti ir paversti į išskaidytą vaizdavimą duomenys. Jei šaltinių ir garso srautų skaičius sutampa duomenys priskiriami tiesiogiai. Jei garso šaltinių yra daugiau nei garso srautų, yra naudojama tiesinė funkcijų sistema apskaičiuoti tikėtinausius koeficientus naudojamus vietoj išmatuotų duomenų:

$$\text{Minimizuoti } \|c_j\|, \text{ taip kad būtų tenkinama sąlyga } Ac_j = d_j$$

Kur c_j – gautieji koeficientai; d_j – išmatuoti duomenys paversti į dažnių erdvę

4. Gaunama $M \times T$ $S(w)$ matrica, kurios eilutėse yra išskirti šaltiniai. Jie paverčiami atgal į laiko erdvę.

5.3. „DUET“ ir „Hard- Lost“ algoritmų palyginimas

Buvo tiriamas „DUET“ ir „Hard - Lost“ algoritmų vykdymo tikslumas bei greitis.

Vykdymo tikslumas lygintas išskirtų garso šaltinių iš garsų srauto garso – triukšmo santykio (SNR) koeficientu, matuojamu decibelais.

Vykdymo greitis lygintas tiesiogiai – matuojant algoritmų vykdymo trukmę pateiktiems duomenims.

Naudotas 950 MHz kompiuteris, turintis 512 RAM atminties. Kaip greičio tyrimo ir tikslumo tyrimo programa naudota „Matlab Release 12“ versija.

Tyrimui pasirinkti 5 realūs garso įrašai. Šie įrašai buvo maišomi tarpusavyje, sintezuojant garso srautus. Garsų sintezavimui naudota programa „Cool Edit Pro“. Reikia pažymėti, kad tokiu būdu sintezuojant garso srautus, garso šaltinis nebūtinai turi būti vienas garsą skleidžiantis objektas. Šiuo atveju garso šaltiniu yra laikomas vienas garso įrašas.

Garso šaltiniai – įrašai:

- Muzikos įrašas iš kasetės – daug garso šaltinių (šalt.1)
- Muzikos įrašas iš kasetės – daug garso šaltinių (šalt. 2)
- Pianino muzika, įrašyta mikrofonu (šalt. 3)
- Daina, įrašyta mikrofonu (šalt. 4)
- Dviejų žmonių pokalbis (šalt. 5)

Taip pat buvo atliktas palyginimas su garso įrašais iš [28]. Šiuo atveju tirti 5 iš TIMIT duomenų bazės paimti sintezuoti įrašai bei realiai sumaišyti (naudojant du mikrofonus) Tee-Won-Lee įrašai.

Visi įrašai buvo sutrumpinti iki 10 sekundžių. Įrašai pateikti „SSS“ sistemai ir „DUET“ sistemai (įgyvendinta MATLAB aplinkoje). Gauti rezultato įrašai palyginti su tikrais šaltiniais pagal formulę, pateiktą [3]:

$$SNR_i = 20 \log_{10} \frac{\|s_i\|}{\|\hat{s}_i - s_i\|} \quad (8)$$

Čia:

SNR – išskirtas garso – triukšmo santykis

s_i – originalus garso šaltinis

\hat{s}_i - išskirtas garso šaltinis

Išmatuotas garso triukšmo santykis įvairiems garso šaltinių srautų variantams yra pateikiamas vienuoliktoje lentelėje.

Garso šaltinių išskyrimo greitis buvo tiriamas tokiems duomenims:

- Du 10 sekundžių garso srautai, kuriuos sudaro du garso šaltiniai (šalt. 1, šalt. 2)
- Du 10 sekundžių garso srautai, kuriuos sudaro trys garso šaltiniai (šalt. 1, šalt. 2, šalt. 3)
- Du 10 sekundžių garso srautai, kuriuos sudaro keturi garso šaltiniai (šalt. 1, šalt. 2, šalt. 3, šalt. 4)
- Du 10 sekundžių garso srautai, kuriuos sudaro keturi garso šaltiniai (šalt. 1, šalt. 2, šalt. 3, šalt. 4, šalt. 5)
- Trys 10 sekundžių garso srautai, kuriuos sudaro du garso šaltiniai (šalt. 3, šalt. 4) (tikrinta tik „Hard – Lost“ algoritmui, nes yra netinkamos sąlygos „DUET“ algoritmo veikimui)

Buvo matuojamas algoritmų atlikimo laikas nuo analizės algoritmo paleidimo pradžios iki rezultatų (garso failų) gavimo. Tai buvo atliekama su vienu garso mišiniu po 10 kartų ir buvo skaičiuojamas rezultatų vidurkis.

Pasirinkti duomenys leidžia ištirti visus aklo šaltinių išskyrimo algoritmo atvejus:

- Kai duomenys yra gerai apibrėžti – garso srautų yra daugiau nei garso šaltinių
- Kai duomenys yra pakankamai apibrėžti – kai garso šaltinių yra tiek kiek ir garso srautų
- Kai duomenys yra nepakankamai apibrėžti – kai garso šaltinių yra daugiau nei garso srautų.

Kadangi „DUET“ algoritmo apribojimas yra du įeinantys garso srautai, tiriant gerai apibrėžtų duomenų atvejį jo nebuvo galima naudoti. Tyrimas atliktas tik su „Hard - Lost“ algoritmu.

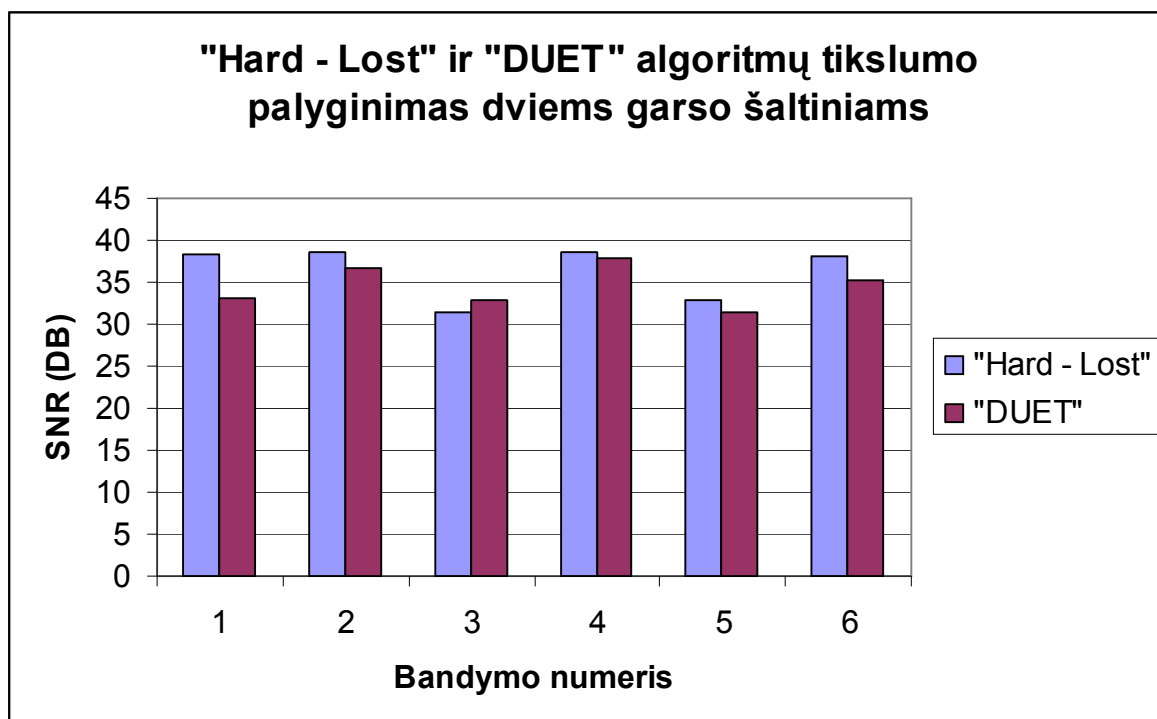
5.4. Eksperimentų rezultatai

Algoritmų „Hard - Lost“ ir „DUET“ tikslumo palyginimo rezultatai, kai garso srautų skaičius 2, o šaltinių skaičius keičiasi nuo 2 iki 5:

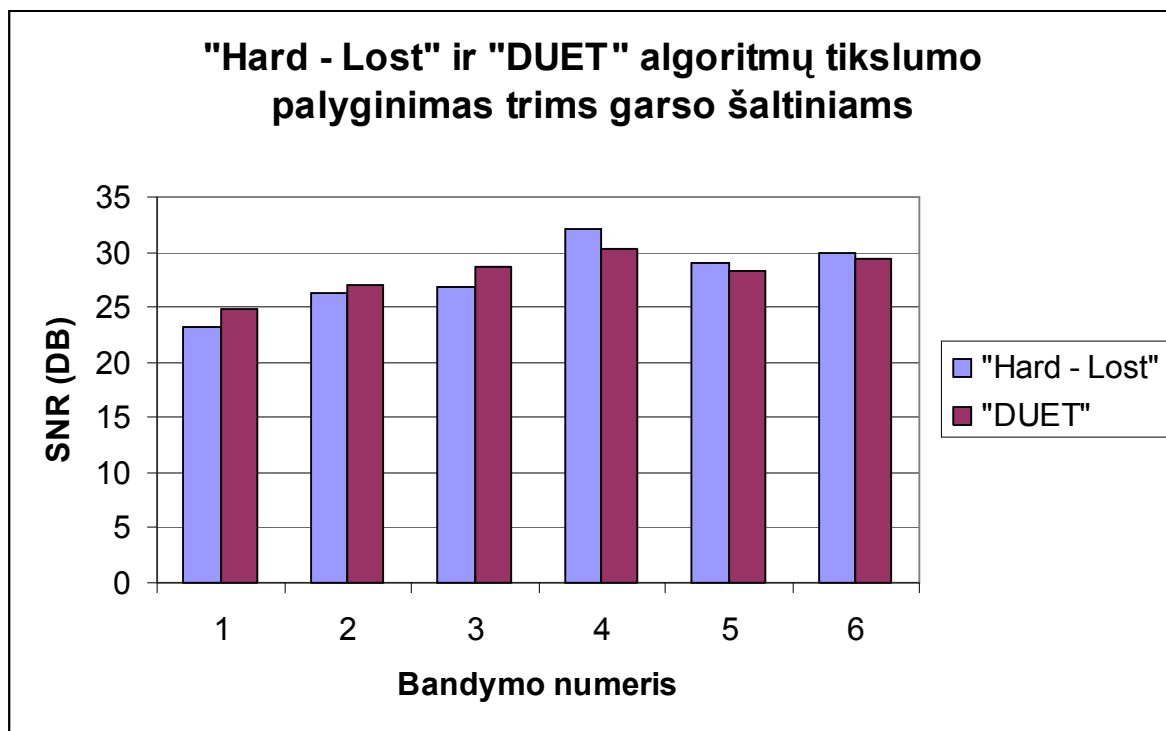
4 lentelė “DUET” ir “Hard - Lost” garso šaltinių išskyrimo tikslumo palyginimas

Garso srautų skaičius	Garso šaltinių skaičius	Garso šaltiniai	Stulpeliai, žymi išskirtų šaltinių SNR (Db) Hard – Lost		Stulpeliai, žymi išskirtų šaltinių SNR (Db) DUET	
2	2	šalt.1 + šalt.2	39.56	37.20	35.05	31.21
2	2	šalt.3 + šalt.4	36.85	40.25	35.04	38.52
2	2	šalt. 1 + šalt.5	32.21	30.56	34.52	31.23

2	2	šalt.2 + šalt.3	38.26		39.05		37.65		37.89			
2	2	šalt.2 + šalt.4	29.86		35.98		28.95		33.68			
2	2	šalt.3 + šalt.5	38.52		37.55		35.64		35.05			
2	3	šalt.1 + šalt.2 + šalt.3	21.52	21.54	26.31		24.36	23.21	26.85			
2	3	šalt.1 + šalt.2 + šalt.4	23.43	25.85	29.63		25.03	26.05	30.21			
2	3	šalt.1 + šalt.2 + šalt.5	25.56	26.62	28.57		26.35	26.89	32.52			
2	3	šalt.1 + šalt.3 + šalt.5	30.68	35.45	30.21		31.02	30.56	29.16			
2	3	šalt.2 + šalt.4 + šalt.5	28.62	29.10	29.27		27.97	29.06	28.03			
2	3	šalt.3 + šalt.4 + šalt.5	29.53	29.59	30.52		28.26	29.87	29.80			
2	5	Iš TIMIT DB	20.13	19.45	19.60	17.63	20.07	21.30	19.43	19.42	18.67	20.13
2	2	Tee- Won – Lee (nesintezuoti šaltiniai)	9.02		5.35		10.50		8.64			



8 pav. "DUET" ir "Hard - Lost" garso šaltinių išskyrimo tikslumo palyginimas dviems gars šaltiniams ir dviems garso srautams

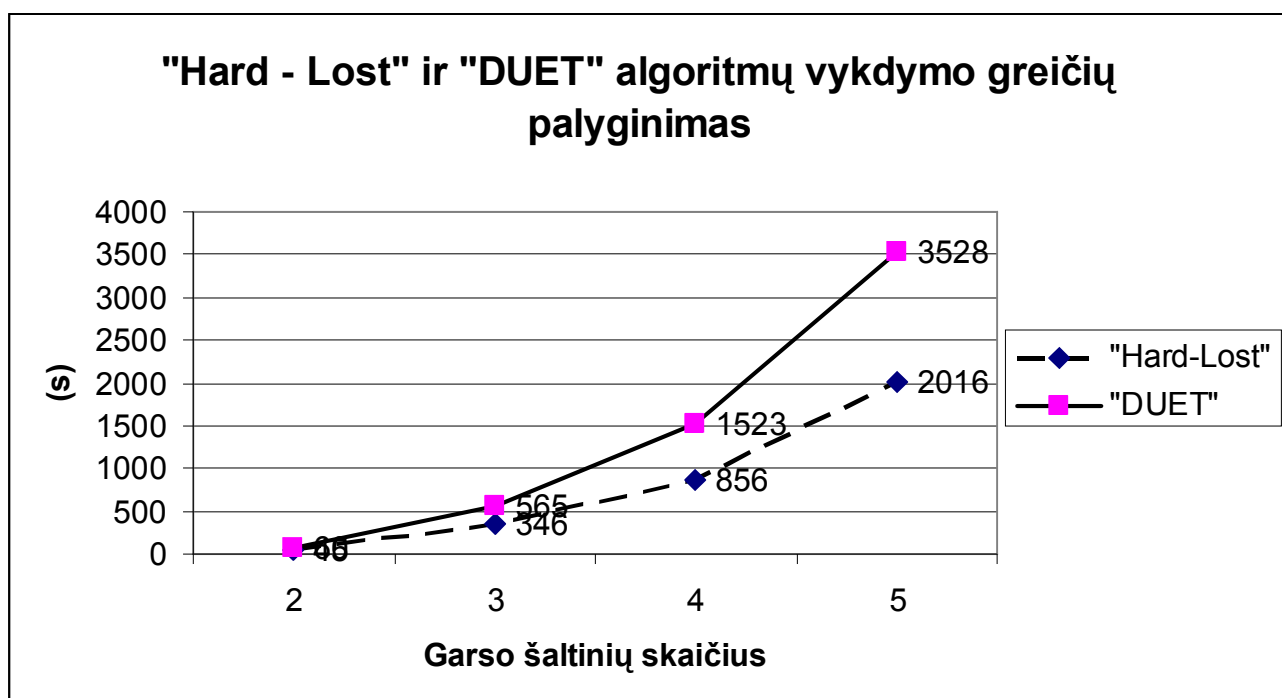


9 pav. "DUET" ir "Hard - Lost" garso šaltinių išskyrimo tikslumo palyginimas dviems gars šaltiniams ir dviems garso srautams

Algoritmų „Hard - Lost“ ir „DUET“ greičio palyginimo rezultatai, kai garso srautų skaičius 2, o šaltinių skaičius keičiasi nuo 2 iki 5:

5 lentelė "DUET" ir "Hard - Lost" garso šaltinių išskyrimo greičio palyginimas

Garso srautų skaičius	Garso šaltinių skaičius	Garso šaltiniai	Analizės atlikimo greitis (sekundėmis) Hard – Lost	Analizės atlikimo greitis (sekundėmis) DUET
2	2	šalt.1 – šalt.2	45	60
2	3	šalt.1 – šalt.3	346	565
2	4	šalt.1 – šalt.4	856	1523
2	5	šalt.1 – šalt.5	2016	3528
3	2	šalt.1 – šalt.2	52	-
3	3	šalt.1 – šalt.3	546	-



10 pav. "DUET" ir "Hard - Lost" garso šaltinių išskyrimo greičio palyginimas

Iš penktos ir šeštos lentelių bei aštuntos, devintos bei dešimtos diagramų matyti, kad:

- Abiejų algoritmų vykdymo tikslumas priklauso nuo išskiriamų šaltinių kiekio; kuo garso šaltinių skaičius yra didesnis, tuo blogesnis yra išskyrimo tikslumas;
- Lyginant absoliučiai, algoritmų tikslumas yra menkas. Pilnai atskirto garso šaltinio SNR = 100 Db. Tuo tarpu algoritmų SNR siekia 40 Db, kai duomenys yra gerai apibrėžti ir 30 Db, kai duomenys yra nepakankamai apibrėžti.
- Lyginant algoritmų tikslumą, galima pasakyti, kad konkretiems duomenims rezultatai yra panašūs. Sintezuotų duomenų atveju „Hard - Lost“ geriau atskiria garso šaltinius, tuo tarpu „DUET“ algoritmas geriau atskiria realaus pasaulio garsus nei „Hard - Lost“ algoritmas.
- Algoritmų vykdymo laikas priklauso nuo išskiriamų šaltinių skaičiaus. Geriausiai yra išskiriami du garso srautai, sudaryti iš dviejų garso šaltinių. Tačiau didėjant šaltinių skaičiui, ženkliai didėja ir išskyrimo laikas – garso šaltinių išskyrimas trunka iki valandos. Šiuo atveju „Hard - Lost“ algoritmas yra greitesnis.

Bendras pasirinkto "Hard - Lost" algoritmo įvertinimas, naudojant jį "SSS" programoje, pagal pasirinktus projekto kriterijus:

6 lentelė "Hard - Lost" garso šaltinių išskyrimo algoritmo įvertinimo apibendrinimas

Eil. Nr.	Parametras	„SSS“ projekto vertinimas
1.	Atpažintų instrumentų skaičius priklauso nuo naudojamų mikrofonų skaičiaus (kuo	Atpažįstami du ir daugiau garsų.

	<p>mikrofonų skaičius didesnis, tuo atpažinimo rezultatas tikslesnis). Įrašams naudojami 2 mikrofonai. Todėl minimalus atpažintų instrumentų skaičius turėtų būti 2.</p> <p>Dabartinės analogiškos sistemos atpažįsta 2-4 instrumentus.</p>	<p>Kokybė, atpažįstant daugiau nei du balsus blogėja.</p> <p>Įvertinimas: Patenkinamas</p>
4.	<p>Sintezuotų ir realių garso signalų SNR santykių palyginimas 14 vienodo ilgio garso mišinių.</p> <p>Rezultatai laikomi patenkinamais, jei sintezuotiems ir realiems duomenims algoritmo tikslumas siekia 40-50%.</p>	<p>30 – 40 dviems sintezuotiems garso šaltiniams</p> <p>20 – 30 realiems šaltiniams</p> <p>Lyginant su užsibrėžtais projekto tikslais – įvertinimas nepatenkinamas</p>

6. IŠVADOS

Darbo metu buvo atlikta garso šaltinių išskyrimo metodikų analizė. Buvo apsiribota tomis metodikomis, kurios yra naudojamos muzikinėje srityje. Buvo atliktas metodikų analizės palyginimas, pagal žinomiausius tų metodikų atliktus tyrimus, siekiant išsiaiškinti, kuri garso srauto analizės metodika yra labiausiai pažengusi vystymosi prasme ir turi tendenciją toliau būti tobulinama.

Išanalizavus garso šaltinių išskyrimo metodikomis paremtų tyrimų rezultatus bei šių metodikų nagrinėjimo intensyvumą buvo pasirinkta nuo duomenų priklausančios sistemos kūrimo metodika dėl savo pritaikymo universalumo ir egzistuojančių metodų skaičiaus. Buvo pasirinktas nuo duomenų priklausančios sistemos tyrimo metodas „aklas šaltinių išskyrimas“, kuris metodų grupėje yra vyraujantis.

Darbe buvo siekiama ištirti garsų analizės būdus, orientuojantis į jų panaudojimą realioje aplinkoje. Kad metodą ar algoritmą būtų galima naudoti realioje aplinkoje yra reikalinga, kad jis būtų efektyvus ir nereikalautų ilgų skaičiavimų. Taip pat algoritmas privalo kuo tiksliau išskirti garso šaltinius. Norint algoritmą panaudoti realioje aplinkoje taip pat yra reikalinga, kad algoritmas būtų pajėgus analizuoti realius duomenis, t.y., kad pateikiami duomenys galėtų būti kuo universalesni.

Pasirinkti analizuoti algoritmai nėra apriboti vien konkreto tipo duomenimis. Šiems algoritmams galima pateikti bet kokio ilgio garso įrašus. Taip pat yra galimybė teikti ne garso srauto duomenis. Tačiau, algoritmams yra taikomi kiti apribojimai:

- Garsų srauto išretinimo apribojimas – konkrečiame laiko tarpe yra aktyvus tik vienas šaltinis;

- Aplinkos apribojimas – aplinka be aido („DUET“ algoritmo sąlyga).

Tokie apribojimai yra neišvengiami garsų analizės srityje, nes analizuojami duomenys turi sudėtingą struktūrą. Juos gali sudaryti bet kokie aplinkos garsai (kurie patys savaime yra sudėtingi dėl jų įvairumo), visada egzistuoja aplinkos triukšmas, aidas, imtuvo nekokybiškumas ir kiti faktoriai, įtakoiantys analizuojamo garso kokybę.

Algoritmai pasirinkti pagal taikomų apribojimų kiekį ir analizės šaltiniuose nurodytos atlikimo kokybės santykį.

„Hard – Lost“ ir „DUET“ algoritmai priklauso „aklo šaltinių išskyrimo“ metodo grupei. Tačiau jie yra paremti skirtingomis koncepcijomis. Iš gautų rezultatų matome, kad jų garso šaltinių išskyrimo tikslumas yra panašus. „Hard – Lost“ algoritmas yra pranašesnis už standartinį „DUET“ algoritmą greičio prasme. Taip pat šis algoritmas nėra apribotas dviem garso srautais.

Abu algoritmai duoda gerus rezultatus, kai garso srautams yra naudojami dirbtinai sukurti duomenys. Tačiau, kai bandoma atlikti veiksmus su realiais duomenimis, algoritmai nėra efektyvūs. Jų garso šaltinių išskyrimo tikslumas siekia iki 30%. Todėl šie algoritmai nėra tinkami komerciniam naudojimui.

Šių garso šaltinių išskyrimo algoritmų pritaikymo galimybės yra mokslinėje veikloje, taip pat atliekant garsų analizę pažintiniais tikslais.

„SSS“ sistema, naudojanti „Hard – Lost“ algoritmą yra skirta garso analizei universitetinėje aplinkoje. Ją galima panaudoti keičiant patį algoritmą ar išplėsti pridedant galimus naujus sistemos komponentus.

Siūlomos „SSS“ sistemos keitimo kryptys:

- Kitų garso srauto analizės algoritmų įtraukimas į programą – kad būtų galima įvairiapusiškiau palyginti garso analizės rezultatus, į sistemą būtų galima įtraukti kitomis metodikomis paremtus algoritmus.
- Automatinis rezultato garso šaltinių skaičiaus nustatymas – galimas pačio algoritmo praplėtimas, naudojant muzikologinę analizę, atpažinti, kiek garso šaltinių egzistuoja konkrečiame garso sraute. Tai padarytų sistemą universalesnę.
- „Hard – Lost“ analizė ir tobulinimas, pasiremiant atlikto tyrimo rezultatais – algoritmą galima papildyti moduliais, leidžiančiais geriau atskirti realios aplinkos garsus. Tam galėtų būti taikomas triukšmų šalinimas, įeinančio garso srauto išankstinė analizė ir tikslesnis pervedimas į dažnių sritį ir iš jos.
- „Hard – Lost“ algoritmo specializavimas – algoritmą galima papildyti dalimis, skirtomis konkrečių garso srautų analizei. Šiuo metu algoritmas yra pritaikomas bet kokiam garso srautui. Yra neskiriama, ar garso šaltinis yra

žmogaus kalba, variklio burzgimas ar muzika. Algoritmą galima pritaikyti, skirti garso srautą pagal duotą žodyną – t.y. praplėsti algoritmą apmokymo sistema. Produktą galima būtų pritaikyti atpažinti mušamųjų, pučiamųjų instrumentų ar kitokią muziką.

- „Hard – Lost“ algoritmo pritaikymas ne garsų analizėje (telekomunikacijoje, medicinoje, vaizdo analizėje)
- Įrašymo modulis keliems mikrofonams – šiuo metu sukurtas sistemos įrašymo modulis, skirtas vienam mikrofonui. Norint įrašytą garsą įvesti į sistemą šiuo metu reikia prieš tai jį sintezuoti kita programa. Jei būtų galima vienu metu įrašyti kelis garso srautus (turint aparatūrinę įrangą) iš kelių mikrofonų, sistemą būtų galima sėkmingiau taikyti realios aplinkos garsų analizei.
- Sistemos garso srauto analizės algoritmo įdiegimas į muzikinės transkripcijos sistemas. Tai patobulintų muzikinės transkripcijos sistemų darbą, nes vienbalsės, arba vieno garso šaltinio muzikos analizė yra mažiau sudėtingas procesas, nei daugiabalsės muzikos analizė.

7. LITERATŪRA

- [1] Blind signal separation. Žiūrėta [2007 04 20], prieiga internete http://en.wikipedia.org/wiki/Blind_signal_separation
- [2] Vincent Yan Fu Tan. Blind audio source separation. Žiūrėta [2007 04 20], prieiga internete <http://web.mit.edu/vtan/www/masters.pdf>
- [3] Paul D. O'Grady, Barak A. Pearlmutter. Hard – LOST: Modified k-Means for Oriented lines Žiūrėta [2007 04 20], prieiga internete <http://www-bcl.cs.may.ie/~bap/papers/hard-LOST-ISSC2004.pdf>
- [4] Carmine Casciato, McGill University, Computer Music Seminar 2 (MUMT611) Music Information Acquisition, Preservation, and Retrieval seminaro medžiaga 2005 02 10. [Žiūrėta 2005 11 02], prieiga internete http://www.music.mcgill.ca/~ich/classes/mumt611_05/transcriptions.html
- [5] R. Ambrazevičius. „Psichoakustiniai darnos rekonstravimo metodai“ [Žiūrėta 2005 11 02], prieiga internete <http://www.ceeol.com/aspx/getdocument.aspx?logid=6&id=f36430bd-bbef-4f16-831f-df9bbe9c9cd4>
- [6] The auditory scene analysis problem .History of research Auditory Research Laboratory [Žiūrėta 2005 11 02], prieiga internete <http://www.psych.mcgill.ca/labs/auditory/introASA.html>
- [7] Kashino, Tanaka. (1993). „A sound source separation system with the ability of automatic tone modeling“. Proceedings of the International Computer Music Conference, 1993.
- [8] Andrew Sterian, Mary H. Simoni, Gregory H. Wakefield Model-Based Musical Transcription, 1999 International Computer Music Conference, Beijing, China [Žiūrėta 2005 11 06], prieiga internete <http://musen.engin.umich.edu/wp/papers/transcription.pdf>
- [9] Anssi P. Klapuri. Automatic transcription of music Proceedings of the Stockholm Music Acoustics Conference, August 6-9, 2003 (SMAC 03), Stockholm, Sweden [Žiūrėta 2005 11 03], prieiga internete http://www.cs.tut.fi/sgn/arg/klap/smac2003_klapuri.pdf
- [10] Meddis, R., Hewitt, M. J., “Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification,” J. Acoust. Soc. Am. 89(6), p.2866–2882, 1991.
- [11] de Cheveigné, A. and Kawahara, H., “Multiple period estimation and pitch perception model,” Speech Communication 27, pp. 175–185, 1999. [Žiūrėta 2005 11 06], prieiga internete <http://citeseer.ist.psu.edu/cache/papers/cs/12282/http:zSzzSzwww.ircam.frzSzequipeszSzpcmzSzcheveignzSzshzSzpszSzSpeechComm.pdf/decheveigne98multiple.pdf>
- [12] Tolonen, T. and Karjalainen, M., “A computationally efficient multipitch analysis model,” IEEE Trans. Speech Audio Processing, Vol. 8, No. 6, pp. 708-716, Nov. 2000. [Žiūrėta 2005 11 06], prieiga internete <http://ie.fing.edu.uy/investigacion/grupos/gmm/audio/papers/multipleF0/opr.pdf>
- [13] The blackboard model of problem solving and the evolution of blackboard architectures. *AI Magazine* 7 (2):38–53. [Žiūrėta 2005 11 02], prieiga internete <http://www.aai.org/Library/Magazine/Vol07/07-02/vol07-02.html>
- [14] Anssi P. Klapuri. Automatic transcription of music Proceedings of the Stockholm Music Acoustics Conference, August 6-9, 2003 (SMAC 03), Stockholm, Sweden [Žiūrėta 2005 11 03], prieiga internete http://www.cs.tut.fi/sgn/arg/klap/smac2003_klapuri.pdf
- [15] Godsmark, D. and Brown, G. J., “A blackboard architecture for computational auditory scene analysis,” Speech Communication 27, pp. 351–366, 1999. [Žiūrėta 2005 11 06], prieiga internete <http://www.dcs.shef.ac.uk/~guy/pdf/spcom99.pdf>
- [16] „Rebecca Fiebrink Blackboard Algorithms for Polyphonic Music Transcription Annotated Bibliography with Hyperlinks“, 2005 02 13. [Žiūrėta 2005 11 02], prieiga internete http://www.music.mcgill.ca/~ich/classes/mumt611_05/Presentations/2005/FiebrinkBlackboard.htm
- [17] Davy, M. and Godsill, S. J., “Bayesian harmonic models for musical signal analysis,” In J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith, editors, *Bayesian Statistics VII*, Oxford University Press, 2003. [Žiūrėta 2005 11 06], prieiga internete <http://www-sigproc.eng.cam.ac.uk/~sjg/talks/valencia%202002/talk.pdf>
http://www.idiap.ch/~paiement/references/to_read/music/feature_extraction/Davy_Bayes7_2002.pdf
- [18] Goto, M., “A predominant-F0 estimation method for realworld musical audio signals: MAP estimation for incorporating prior knowledge about F0s and tone models,” in Proc. Workshop on Consistent and reliable acoustic cues for sound analysis, Aalborg, Denmark, Sep. 2001. [Žiūrėta 2005 11 06], prieiga internete <http://staff.aist.go.jp/m.goto/PAPER/ICASSP2001goto.pdf>

- [19] Virtanen, T., "Sound source separation using sparse coding with temporal continuity objective," International Computer Music Conference, Singapore, 2003. [Žiūrėta 2005 11 06], prieiga internete <http://www.cs.tut.fi/sgn/arg/music/tuomasv/icmc2003.pdf>
- [20] Aaron S. Master (2003). DUET and DASSS Review [Žiūrėta 2005 11 10], prieiga internete <http://ccrma.stanford.edu/~asmaster/BTSM/master/node2.html>
- [21] Patrik O. Hoyer (2002). Non-negative sparse coding [Žiūrėta 2005 11 10], prieiga internete http://arxiv.org/PS_cache/cs/pdf/0202/0202009.pdf
- [22] Paul D. O'Grady, Barak A. Pearlmutter, Scott T. Rickard Survey of Sparse and Non-Sparse Methods in Source Separation [Žiūrėta 2007 05 01], prieiga internete <http://www.hamilton.ie/paul/bss-survey.pdf>
- [23] BALSŲ IŠSKYRIMO POLIFONINĖJE MUZIKOJE PROGRAMINĖS ĮRANGOS KŪRIMAS BEI TYRIMAS.
Literatūros apžvalga. Žiūrėta [2007 04 14], prieiga internete http://soften.ktu.lt/~s66811/magistras/siuntimui/radvickaite_spec.doc
- [24] BALSŲ IŠSKYRIMO POLIFONINĖJE MUZIKOJE PROGRAMINĖS ĮRANGOS KŪRIMAS BEI TYRIMAS.
Reikalavimų specifikacija. Žiūrėta [2007 04 14], prieiga internete http://soften.ktu.lt/~s66811/magistras/siuntimui/radvickaite_spec.doc
- [25] BALSŲ IŠSKYRIMO POLIFONINĖJE MUZIKOJE PROGRAMINĖS ĮRANGOS KŪRIMAS BEI TYRIMAS.
Sistemos analizė. Žiūrėta [2007 04 14], prieiga internete http://soften.ktu.lt/~s66811/magistras/siuntimui/radvickaite_arch.doc
- [26] BALSŲ IŠSKYRIMO POLIFONINĖJE MUZIKOJE PROGRAMINĖS ĮRANGOS KŪRIMAS BEI TYRIMAS.
Sistemos analizė. Žiūrėta [2007 04 14], prieiga internete http://soften.ktu.lt/~s66811/magistras/siuntimui/radvickaite_design.doc
- [27] Henricson M. and Nyquist E.. Programing in C++. Rules and recommendations. Žiūrėta [2006 09 10], prieiga internete <http://www.doc.ic.ac.uk/lab/cplus/c++.rules/>
- [28] Rickard S. Blind source separation. Internetinis puslapis, BSS tyrimo failai. Žiūrėta [2007 03 14], prieiga internete <http://eleceng.ucd.ie/~srickard/bss.html>

8. TERMINŲ IR SANTRUMPŲ ŽODYNAS

7 lentelė terminai

Terminas	Apibūdinimas
Garso šaltinių išskyrimas	Garso šaltinių, kurie vaizduojami
Garsų srautas	Sumaišyti garsai, sklindantys erdvėje kartu, kuriuos fiksuoja vienas garso imtuvas. Vieną garsų srautą galima įsivaizduoti, kaip vienu mikrofonu įrašytą, iš naujo klausomą kelių žmonių pokalbį.
Kompiuterizuota muzikos transkripcija	Muzikos, pateiktos garso failu, analizė bei užrašymas muzikos atlikimo komandomis (pavyzdžiui, natomis)
Muzikos analizė	Procesas, kurio metu yra atskiriami balsai – garso šaltiniai
Specifikacija	Sistemos funkcionalumo aprašymas formaliais metodais
(Rezultatų) eksportas	Išanalizuoto garso saugojimas vartotojui patogiu garso formatu

8 lentelė Santrumpos

Santrumpa – terminas	Pilnas santrumpos pavadinimas (jei egzistuoja)	Apibūdinimas
BSS	blind source separation	„aklas“ šaltinių išskyrimas. Matematinis metodas, skirtas šaltinių išskyrimui, neturint pilnos išskyrimui reikalingos informacijos.
CASA ASA	computational auditory scene analysis auditory scene analysis	(skaičiuojamoji) girdimosios aplinkos analizė – modelis, aprašantis garso srauto organizacijos būdą, paremtą žmogaus suvokiamąją aplinką.
DASSS	delay and scale subtraction scoring	metodas, dviejų įrašymo kanalų vėlinimo bei mastelio apskaičiavimui
DSP	Digital signal processing	skaitmeninis garso srauto apdorojimas.
DUET	Degenerate Unmixing Estimation Technique	algoritmas, dviejų įrašymo kanalų garso šaltinių išskyrimui
EM	expectation maximization	algoritmas, randantis maksimalaus tikėtimumo parametrų įverčius tikimybinuose modeliuose. Plačiau, http://en.wikipedia.org/wiki/Expectation-maximization_algorithm
ICA ISA	independent component analysis independent subspace analysis	atskirų komponentų/komponentų erdvės analizė. Skaičiuojamasis metodas signalų išskaidymui į adityvius sub-komponentus, paremtas ne Gauso komponentų statistine nepriklausomybe. Plačiau, http://en.wikipedia.org/wiki/Independent_component_analysis
LTAS	long term average spectrum	integrinio didelės trukmės garso signalo spektro analizė. Metodas, skirtas supaprastinti garso aukščių matavimui, kai intonavimas palyginti stabilus.

MCMC	Markov – chain Monte-Carlo	algoritmas, skirtas modeliuoti trūkstamus duomenis iš tikimybinės pasiskirstymo funkcijos. Plačiau, en.wikipedia.org/wiki/MCMC
MIDI	Musical Instrument Digital Interface	muzikinių instrumentų muzikinė sąsaja yra standartinis būdas aprašyti bei susieti muzikos natas bei jų parametrus tarp dviejų skaitmeninių įrenginių
MIR	music information retrieval	muzikos informacijos išskyrimas/analizė naudojant kompiuterį. Egzistuoja dvi tyrimų kryptys – muzikos išskyrimas iš garso srauto bei optinė muzikos analizė.
„mp3“, „wav“, „mpeg“, „wma“	Motin picture layer 3; Wave; Motin picture experts group; Windows media audio	muzikos formatų failų plėtiniai, naudojami sutrumpintai apibrėžti muzikos formatus.
PA	Use Case, Panaudos atvejis	panaudojimo atvejis
STFT	short time/term Furje transformation	trumpalaikė Furje transformacija. Furje transformacija, skirta išgauti besikeičiančio sinuso formos signalo dažnį bei fazę laikui bėgant.
UML	Unified Modeling Language	modeliavimo kalba, naudojama objektiškai orientuotame projektavime
Windows PCM wav	Windows pulse code simulation wave	Windows sistemos nesuspaustas garso failų formatas.
SNR	Sound To Noise Ratio	Garso - triukšmo lygis.
“SSS”	Sound source separation system	Darbe sukurtos sistemos pavadinimas

9. PRIEDAI

9.1. Algoritmų, reikalaujančių šaltinių išretinimo apribojimo palyginimas, pagal [22]

Table I: A Comparison of Various Source Separation Techniques (†denotes sparse methods)

Algorithm	#Sources (N)		#Sensors (M)			Mixing Model			Representation			Unmixing	
	M	$> M$	1	2	> 2	Inst.	Anechoic	Echoic	Time	Freq.	Signal Dict.	Parameter Est.	Separation
Wiener (1949) Filter		×	×			×	×	×		×		Power Spectrum Reshaping	Linear Filtering
Herauld and Jutten (1986)	×			×	×	×			×			Independence Maximisation	Adaptive Feedback Network
JADE (Cardoso and Soudoumiac, 1993)	×			×	×	×			×			Joint Cumulant Diagonalisation	Linear Transformation
ICA (Comon, 1994)	×			×	×	×			×			Mutual Information Minimisation	Linear Transformation
BS-InfoMax (Bell and Sejnowski, 1995)	×			×	×	×			×			Entropy Maximisation	Linear Transformation
Lambert (1995)	×			×	×			×	×			Multichannel Blind Deconvolution	Current Estimate Feedback
Lin et al. (1997) †	×	×		×		×			×			Hough Transform	Hard Assignment
SOBI (Belouchrani et al., 1997)	×			×	×	×			×			Joint Unitary Diagonalisation	Linear Transformation
fICA (Hyvärinen and Oja, 1997)	×			×	×	×			×			Non-Gaussianity Maximisation	Linear Transformation
Lee et al. (1999) †	×	×		×	×	×			×			Gradient Ascent Learning	Maximum a Posteriori
DUET (Jourjine et al., 2000) †	×	×		×		×	×			×		2D Histogram Clustering	Binary Time-Freq Masking
Bofill and Zibulevsky (2000) †	×	×		×	×	×				×		Potential Function Clustering	L_1 -Norm Minimisation
Zibulevsky and Pearlmuter (2001) †	×	×		×	×	×					×	MAP with Laplacian Prior	Joint Optimisation
Roweis (2001)		×	×			×				×		Hidden Markov Models	Spectral Masking and Filtering
Pearlmutter and Zador (2004) †		×	×					×	×			Known HRTF and Dictionary	L_1 -Norm Minimisation

9.2. „SSS“ sistemos funkcinis aprašymas

9.2.1. Trumpas „SSS“ sistemos aprašymas

„Sound Source Separation“ sistema, arba „SSS“ sistema – balsų išskyrimo daugiabalsėje muzikoje programa.

Programa atlieka polifoninio (daugiabalsio) muzikos įrašo analizę bei išskiria bendrame garso sraute atskirus „balsus“ – garso šaltinius.

Geriausiai sistemos darbą apibūdintų pavyzdys:

Grojant dviems muzikos instrumentams – smuikui ir fleitai yra įrašomi keli šio įrašo variantai vienu metu. Įrašai įkeliami į sistemą. Sistema išanalizuoja gautus duomenis (bendrą garso srautą) ir į skirtingus failus įrašo skirtingus instrumentus – viename girdimas smuiko grojimas, kitame – fleitos.

9.2.2. „SSS“ sistemos paskirtis

Produktas pateikiamas kaip atskira įdiegiama programa „Windows“ aplinkoje su vartotojo sąsaja. Taip pat programa gali būti naudojama, kaip muzikos transkripcijos (analizės bei užrašymo natomis) paprogramė vėlesniuose projektuose, todėl balsų analizės modulio biblioteka („dll“) pateikiama kartu su programine įranga.

Sistema skirta naudojimui universiteto veikloje plėsti studentų žinias apie muzikos technologijas, padėti efektyviau atlikti muzikos analizę. Taip pat programinė įranga gali būti įdiegta pageidaujančių programa suinteresuotų vartotojų kompiuteriuose. Programinė įranga bei su ja susijęs tyrimas gali būti panaudoti plėsti su šia kryptimi susijusią mokslinę veiklą.

Programinė įranga skirta:

- Pageidautinų melodijų išskyrimui bei analizei;
- Muzikos transkripcijos programų išplėtimui, tobulinimui;
- Muzikinės klausos treniravimui.

9.2.3. „SSS“ sistemos vartotojai

Neprofesionalus vartotojas (atsitiktinis sistemos vartotojas) – tai vartotojas, kuris nori išskirti balsus daugiabalsėje muzikoje. Nesigilina į įrašo analizės aspektus.

Profesionalus vartotojas (studentas, dėstytojas, muzikos analitikas, žmogus užrašantis muziką natomis) – vartotojas, kurį domina sistemos naudojimas balsų išskyrimui, išskirtos muzikos analizei.

9.2.4. „SSS“ sistemos funkcijos

Programų sistema leidžia, įrašius įrašą, arba pateikus jį vienu iš priimtinių failų formatų („wav“, „mp3“, kiti palaikomi garso failų formatai), konvertuoti failą į reikalingą formatą (nesuspaustą Windows garso failų formatą – „PCM wav“), analizuoti garso failą, naudojant garso signalų apdorojimo bei muzikologines technologijas bei išskirti jame balsus, pavaizduoti rezultatus atskiromis garso bangomis, saugoti bei eksportuoti balsų išskyrimo projektą. Papildomos sistemos funkcijos yra sistemos parametrų keitimas bei vartotojo pagalba.

Sistemos atliekamos funkcijos:

9. Muzikos įrašymas – garso failo įrašymas.
10. Įrašo ar kelių įrašų klausymas – pasirinktų garso failų klausymas atskirai/ vienu metu.
- 11. Muzikos analizė – garso šaltinių (“balsų”) išskyrimas iš bendro garso srauto.**
12. Rezultatų eksportas/saugojimas – garso failų konvertavimas, saugojimas „wav“, „mp3“, „wma“, „mpg“ failų formatais.
13. Sistemos atributų reikšmių keitimas – sistemos pritaikymas konkrečiau vartotojo darbui.
14. Sistemos pagalba – vartotojo vadovas, skirtas apmokymui dirbti su sistema ir spręsti iškilusias problemas.

9.2.5. „SSS“ sistemos vaizdas

