



TURINYS

Žymenys ir santrumpos	8
1. Įvadas. Užduoties analizė	9
2. Analogiškų sistemų apžvalga	11
3. Naudojamų metodų apžvalga	13
3.1. HRTF funkcijos	13
3.2. Galvos šešėlis ir pečių aidas	13
3.3. Tarpausinis laiko skirtumas	14
3.4. Tarpausinis garso lygių skirtumas	14
3.5. HRTF metodo panaudojimas	15
3.6. Duplekso teorija	15
3.7. Kambario impulsinis atsakas	19
3.8. Skyriaus apibendrinimas	20
4. Sistemos kūrimas	21
4.1. Sistemos struktūrinė schema	21
4.2. Jutiklių duomenų sintezės filtras	22
4.3. Garso apdorojimo mechanizmas	24
4.4. Galvos sekimo įrenginys	26



Žymenys ir santrumpos

- MARG - Magnetinis laukas, kampinis pagreitis ir laisvojo kritimo pagreičio vektorius (angl. *Magnetic, Angular Rate, and Gravity*)
- HRTF - su galva susijusios perdavimo funkcijos (angl. *Head Related Transfer Function*).
- ATF - anatominė perdavimo funkcija (angl. *Anatomical Transfer Function*).
- FFTF - (angl. *Field-Free Transfer Function*).
- RIR - kambario impulsinis atsakas (angl. *Room Impulse Response*).
- HRIR - (angl. *Head Related Impulse Response*).
- ITD - tarpausinis laiko skirtumas (angl. *Interaural Time Difference*).
- ILD - tarpausinis garso lygių skirtumas (angl. *Initial Level Difference*).
- USB - universalioji nuosekloji jungtis (angl. *Universal Serial Bus*).
- COM - RS232 nuoseklusis prievadas.



1. Įvadas. Užduoties analizė

Darbo tema yra „virtualios garso realybės sistema“. Šiuo metu virtualio ir papildytos realybės programinė įranga yra toli pažengusi, tačiau garso generacijai taip ir toliau naudojamas stereo garsas, kuriuo nėra įmanoma tiksliai atkurti aplinkos garsų. Šiuo darbu siekiama ištirti galimybes sukurti realiuoju laiku veikiančią sistemą gebančią generuoti binauralinį garsą priklausomai nuo vartotojo galvos orientacijos ir virtualios aplinkos akustinių parametrų. Sistema yra suskirstyta į tris pagrindines dalis: galvos sekimo įrenginys - kuris seka vartotojo galvos orientaciją; programa asmeniniame kompiuteryje - kuri siunčia virtualios aplinkos duomenis ir garso šaltinio skleidžiamą garsą; garso apdorojimo įrenginys - kuris apjungia informaciją gautą iš galvos sekimo įrenginio bei asmeninio kompiuterio ir sugeneruoja binauralinį garsą. Tokia architektūra buvo pasirinkta dėl modulinės struktūros, be to perkėlus garso generavimą į atskirą įrenginį, atlaisvinamas asmeninio kompiuterio centrinis procesorius, ko pasekoje virtualiai aplinkai sukurti lieka daugiau asmeninio kompiuterio resursų. Pradiniai bandymai bus atlikti *Matlab* terpėje, kas leis įsitikinti sistemos įgyvendinimo galimybėmis, bei duos progą patobulinti naudojamus procesus. Virtualios garso realybės tema buvo pasirinkta dėl savo perspektyvumo, nors binauralinis garsas pradėtas tirti pakankamai senai, realiai veikiančių įrenginių nėra sukurta. Tema yra aktuali žaidimų ir medicinos pramonei. Didžiausia sistemos problema slypi HRTF (angl. *Head Related Transfer Functions*) panaudojime, kadangi HRTF deriniai turi būti pritaikyti individualiai kiekvienam klausytojui. Tačiau panaudojus papildomus procesus garso generavimo eigoje, pamėginta sumažinti HRTF neigiamą įtaką. Darbo eigoje bus sukurta pilnai veikianti sistema gebanti kuo tiksliau atkurti virtualią akustinę aplinką, programinė įranga bus parašyta garso apdorojimo, galvos orientacijos įrenginiams, bei asmeniniam kompiuteriui. Taip pat bus suprojektuoti abu naudojami įrenginiai. Baigiamojo bakalaurnio darbo sistemos asmeninio kompiuterio programinei įrangai reikalingi mažiausiai 20MB diskinio kaupiklio dalies talpos. Toks atminties kiekis reikalingas norint vartotojo sąsają padaryti patrauklesnę vartotojui, t.y. grafinės iliustracijos, garsai ir kt. Asmeniniame kompiuteryje programinė įranga dirba Windows 7 operacinėje sistemoje, nes senesnių operacinių sistemų versijų (XP, Vista) Microsoft korporacijos palaikymas bus nutraukiamas artimiausiu metu. Garso apdorojimo įrenginyje bus naudojama ne mažesnė nei 1.1 USB lizdo versija, nes sistema suprojektuota darbui su šia ar aukštesnėmis versijomis. Duomenų perdavimą taipogi būtų galima realizuoti ir kitomis sąsajomis, tokiomis kaip COM, bet tokiu atveju greitis būtų nepakankamas ir sistema įgautų netoleruotiną vėlinimą. Be to, binauralinio garso generavimo sistema yra šiuolaikinė ir labiau taikintina nešiojamiems kompiuteriams, kuriuose nebediegiami RS232 prievadai. Išlieka galimybė naudoti RS232 prievadą, bet tokiu atveju tektų papildomai turėti iš USB į COM keitiklio kabelį, kuris taip pat naudoja USB jungtį. Todėl geriausias pasirinkimas spartos atžvilgiu rinktis virtualų COM prievadą per USB jungtį. Siekiant išlaikyti kuriamos sistemos kuo ilgesnį autonominio veikimo laiką sistema toleruoja maksimalų 0,5 A srovės suvartojimą



(tiek galvos orientacijos nustatymo įrenginys tiek garso apdorojimo įrenginys), be to 0,5 A tai didžiausia srovė kurią pagal specifikaciją gali atiduoti USB 1.1 ir USB 2.0 prievada. Norint išlaikyti sistemos mobilumą, kurio didžioji dalis priklauso nuo galvos sekimo įrenginio dydžio, buvo pasirinkti pakankamai nedideli pastarojo įrenginio maksimalūs matmenys: $50 \times 100 \times 100$ mm. Sistemos veikimas bus vertinamas sugeneruoto binauralinio garso tikslumu (palyginus su įrašytu binauraliniu garsu), bei autonominio veikimo trukme.

2. Analogiškų sistemų apžvalga

Šiame skyriuje bus apžvelgtos analogiškos sistemos: buvę prototipai ir šiuo metu esami rinkoje produktai. Apibendrintai bus aptarti nagrinėjamų sistemų teigiamos ir neigiamos savybės.

Galvos sekimo įrenginio kūrimas nėra naujiena. Tokie įrenginiai naudojami norint nustatyti lėktuvų trimatę poziciją. Pati *sensor fusion* technologija taip pat jau gana ilgai naudojama aviacijos srityje, norint kur kas tiksliau nustatyti poziciją. Kalbant apie binauralinio garso generavimą realiu laiku, paieškos rezultatai stipriai sumažėja. Tokio tipo projektų pasaulyje yra tik du. Projekto tikslas – binauralinio garso generavimas realiu laiku. Vienas iš projektų naudojo „Texas instruments“ pagamintą plokštę - C6713 DSK kuri pavaizduota 2.1 pav.



2.1 pav. C6713 DSK spausdintinė plokštė.

Šios plokštės mikrovaldiklis dirba 225 MHz dažniu ir pasiekia 1800 milijonų instrukcijų per sekundę spartą. Taip pat ši plokštė gali atkurti aukštos kokybės 24 bitų skaitmeminį stereo garsą, turi 512 kB *Flash* bei 16 MB *SDRAM* atminties. *JTAG*¹ prieinamas per *USB* jungtį.

Kaip teigia gaminio kūrėjai, jie generuoja binauralinį garsą naudojant specialų stereo filtrą suprogramuotą minėtoje plokštėje. Programinė įranga naudoja *HRTF* sugeneruotas „*Floridos tarptautinio universiteto*“ asmenų (nepaminėta kieno). Dviejų kanalų išėjimui panaudotos 12 *HRTF* funkcijų, kurios taip pat nebuvo sugeneruotos projekto autorių.

Kūrėjų teigimu, rezultatai jų stipriai nestebina – geriausias efektas jaučiamas tik garsui esant už galvos, nes panaudotos *HRTF* funkcijos nėra pritaikytos jų ausims. Šio įrenginio testavimo akimirką demonstruoja 2.2 pav.

¹JTAG - Programavimo ir testavimo jungtis.



2..2 pav. C6713 DSK plokštės patikra.

Kompiuterinis binauralinio garso generavimas yra įmanomas, bet reikalauja didelių kompiuterio resursų, taip atimdamas galią iš centrinio procesoriaus, todėl kuriamos išorinės garso plokštės, kurios nepriklausomai nuo kompiuterio procesoriaus gali generuoti binauralinį garsą naudojant *HRTF* funkcijas.

Pirmoji garso rinkoje pasirodęs produktas, galėjęs pasiūlyti trimatį pozicinį garso generavimą, buvo 1998 išleista „*Aureal Vortex via the A3D API*“ garso plokštė. Deja nėra galimybės jos išbandyti, bet pagal vartotojų atsiliepimus ši plokštė sugebėdavo gana gerai suformuoti trimačio garso pojūtį. Kadangi įmonė „*Aureal*“ pasitraukė iš rinkos, tai jos poziciją perėmė „*Creative Inc.*“ firma. Gerą garso kokybę bei neblogą trimačio garso efektą gali pasiūlyti tik labai brangios pastarosios įmonės garso plokštės – „*X-Fi*“.

Šiuo metu binauralinio garso generavimas realiu laiku priklausomas nuo vartotojo galvos pozicijos neturi jokių analogų. Kaip ir buvo minėta, atskirai abi dalys yra daugiau ar mažiau naudojamos pasaulinėje rinkoje, bet būtent šių dviejų sričių sujungimas yra visiškai unikalus, todėl analogiškų sistemų apžvalga čia yra neįmanoma.

Kaip buvo minėta apžvalgoje sistema pasiūlytame pavidale yra unikali ir perspektyvi. Šiuo metu rinkoje nėra įrenginių gebančių sugeneruoti binauralinį garsą priklausomai nuo klausytojo galvos orientacijos. Rinkoje esantys analogai suteikia programuotojams galimybę įgyvendinti šį efektą naudojantis papildomais programiniais įrankiais, tačiau įrenginių gebančių tai padaryti savaime, be papildomo programuotojų isikišimo - nėra.



3. Naudojamų metodų apžvalga

Šiame skyriuje bus apžvelgti metodai kuriais remiantis buvo sukurti matematiniai modeliai ir realios sistemos prototipas.

3.1. HRTF funkcijos

HRTF (angl. *Head Related Transfer Function*) dar vadinamos *ATF* (angl. *Anatomical Time Difference*) tai funkcijos nusakančios atsaką kurį kuris charakterizuoja, kaip žmogaus ausis priima garsą sklindantį iš tam tikro taško erdvėje. Panaudojus *HRTF* porą, skirta kiekvienai ausiai, įmanoma sintezuoti binauralinį garsą taip sukuriant erdvės pojūtį. Žmogaus smegenys apskaičiuoja garso šaltinio poziciją erdvėje naudojantis: 1) garso užuominomis atkeliavusiomis į kiekvieną ausį (angl. *monoaural cues*), 2) abiejų ausų garso užuominų palyginimu (angl. *difference cues* arba *binaural cues*). Pagrindinės garso užuominos yra:

1. Tarpausinis laiko skirtumas (angl. *Interaural Time Difference*)
2. Tarpausinis garso lygių skirtumas (angl. *Interaural Level Difference*)
3. Spektriniai skirtumai (angl. *Spectral cues* arba *Pinae effect*)
4. Galvos šešėlis (angl. *Head Shadow*)
5. Pečių aidas (angl. *Shoulder Echo*)
6. Ankstyvieji atspindžiai (angl. *Early Reflections*)
7. Atgarsiai (angl. *Reverberation*)

HRTF aprašo tik dalį iš aukščiau pateiktų užuominų: TLS ir TGLS.

3.2. Galvos šešėlis ir pečių aidas

Galvos šešėlis arba dar kitaip vadinamas akustiniu šešėliu, tai sumažėjusios garso amplitudės regionas atsirandantis, nes galva sugeria arba sumažina amplitudę dalies ją pasiekusių garso bangų. Pasiekusi galvą garso banga gali keliauti aplink arba per galvą, kad pasiektų ausį esančią priešingoje pusėje. Tokiu būdu yra sukuriamas pakankamai geras pirminės lokalizacijos filtras. Ausį esančią šešėlyje garsas pasiekia maždaug 0,7 ms vėliau, ko pasekoje žmogaus smegenys sugeba nustatyti apytikslią garso sklidimo kryptį.

Pečių aidas - tai aidas atsirandantis garso bangom atsispindėjus nuo klausytojo pečių. Toks aidas tiesiogiai turi mažai lokalizacijos informacijos tačiau sudėjus kartu su galvos šešėliu filtras dar labiau patikslėja.

The diagram shows a top-down view of a human head. A speaker icon is positioned to the upper left, with lines representing sound waves traveling towards the head. The left ear is closer to the speaker than the right ear. Two lines from the speaker represent the sound path to each ear. A double-headed arrow between these two paths is labeled 'ITD', representing the time difference in sound arrival.

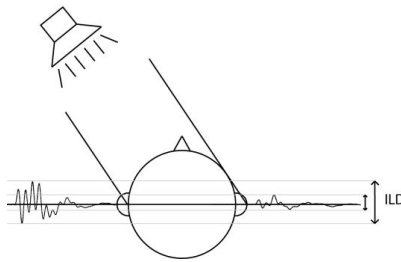
Tačiau išimtį sudaro garsas esantis tiesiai prieš arba už klausytojo. TLS taip pat įtakoja garso bangos fazės pakitimą. Jeigu fazių skirtumas didesnis nei 180 laipsnių, nustatyti kaip fazė yra pasislinkus tampa sunku. Žemuose dažniuose fazės skirtumas nusako pakankamai didelį laiko skirtumą ir yra naudingas orientacijos nustatymui. Aukšto dažnio signalams 180 laipsnių fazės poslinkis nusakys mažesnę laiko skirtumą. Tai reiškia kad ši savybė tampa mažiau naudinga garso bangoms kurių dažnis priklausomai nuo kampo viršija 700-1500 Hz.

3.4. Tarpausinis garso lygių skirtumas

TGLS nusako garso stiprumo skirtumą tarp dviejų taškų (ausų). Šis reiškiny pasireiškia kai garso banga iš dalies yra blokuota galvos. Taip pat kaip ir TLS, TGLS negali tiksliai nusakyti skirtumo tarp garso esančio tiesiai prieš arba už klausytojo. TGLS efektas taipogi kaip ir TLS turi dažninę priklausomybę. Žemuose dažniuose dėl difrakcijos tik maža garso bangos dalis yra blokuota galvos, ko pasėkoje garso intensyvumo skirtumas pasidaro labai mažas.

Tačiau aukštų dažnių ruože difrakcija pasireiškia žymiai mažiau kas įtakoja žymiai didesni garso lygių skirtumą (3.2 paveikslas). Kai kuriais atvejais garso lygių skirtumas gali sudaryti iki 20 dB. Todėl TGLS yra dažniausiai naudojamas aukšto dažnio garso bangoms, o TLS žemo dažnio bangų kryptiai nustatyti.

Taipogi labai svarbi yra spektrinė dedamoji (angl. *pinnae effect*). Dėl kompleksinės ir asimetrinės ausies kriauklės formos garsas atsispindėjęs nuo ausies kriauklės spektriškai pasikeičia. Kartu su pečių atspindžio efektu tai formuoja filtrą turintį krypties priklausomybę ir todėl ne taip kaip TLS ir TGLS yra imanoma atskirti garsus sklindančius iš priekio ir iš galo.



3..2 pav. TGLS (angl. *ITD*) efektas

Dėl spektrinių pokyčių ši dedamoji geriausiai tinka atskirti plačiajuostį garsą - viršijantį 6 kHz dažnį.

3.5. HRTF metodo panaudojimas

Norint panaudoti šį metodą reikia turėti *HRTF* rinkinį visiems vertikalinių ir horizontalių kampų deriniams. Šiame darbe naudojami Masačiuseco technologijų instituto (angl. *Massachusetts Institute of Technology*) 1994 metais atlikto projekto „*KEMAR*“ metu gauti universalūs *HRTF* deriniai.

Turint baigtinio ilgio skaitmeninį garso signalą atliekama diskrečioji sąsuka (angl. *convolution*) pagal 3..1 formulę:

$$(f * g)[n] = \sum_{m=0}^{\infty} f[m] * g[n - m] = \sum_{m=0}^{\infty} f[n - m] * g[m] \quad (3..1)$$

čia f – vieno kanalo baigtinio ilgio daugianaris (angl. *polynomial*), g – atitinkamo kanalo *HRTF* funkcija (išreikšta daugianariu), n – daugianario narių skaičius.

Šioje formulėje atliekama dviejų daugianarių daugyba, rezultate gauto daugianario koeficientai atitinka originalaus daugianario (f) koeficientų seką po sąsukos, visi likę nežinomi koeficientai yra prilyginami nuliui kad nekiltų neapibrėžtumas. Šis veiksmas dar žinomas kaip dviejų daugianarių koeficientų Kauči daugyba (angl. *Cauchy product*).

Veiksmas atliekamas atskirai dešiniajam ir kairiajam kanalams naudojant tą patį garso signalą su atitinkamai dešiniajai arba kairiajai ausiai skirta *HRTF* funkcija.

3.6. Duplekso teorija

Duplekso teorija, pateikta Lordo Reilėjaus (Lord Rayleigh), pateikia paaiškinimus apie žmogaus galimybę lokalizuoti garsus pagal laiko skirtumą TLS tarp garsų, patenkančių į kiekvieną ausį ir taip pat skirtingo garso lygio TGLS patekimą į ausis. Bet iki šiol išlieka klausimas, ar šie skirtumai yra pastebimi.

Duplekso teorija teigia, kad TLS naudojamas žemo dažnio garso šaltinio lokalizacijai, o TGLS – aukšto dažnio. Kadangi natūralus garsas sudarytas ne tik iš žemo dažnio komponentų, bet ir aukšto dažnio, tai klausos sistema turės panaudoti abi metodikas (tiek TLS tiek

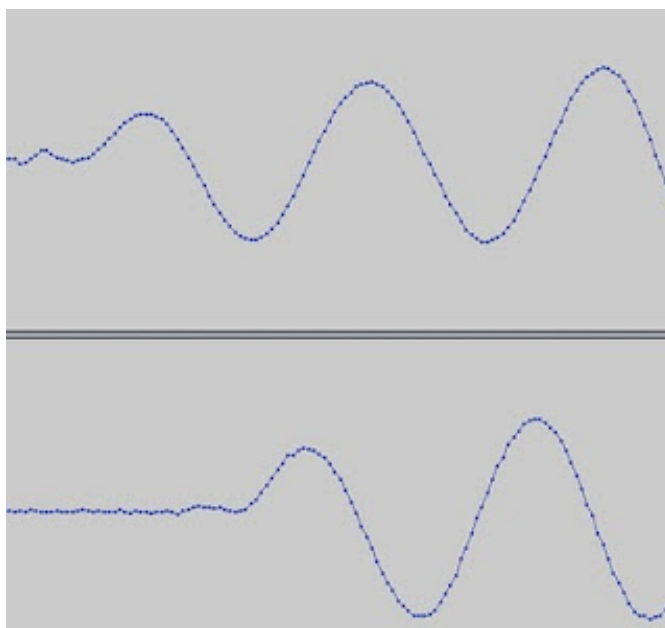
TGLS), norint tiksliai nustatyti garso šaltinio vietą. Šios dupleksinės sistemos rezultatas – ji gali taip pat generuoti taip vadinamus „garso lokalizacijos užuominų mainų“ arba „laiko intensyvumo mainų“ dirgiklius ausinėse, kur TLS, nukreiptas į kairiąją ausinių pusę, perstumtas per TGLS, kuris savo ruožtu nukreiptas į dešiniąją ausinių dalį, todėl garsas suvokiamas lyg sklįstų iš vidurio. Kaip ir visos teorijos, duplekso teorija taip pat susiduria su problemomis. Pastaroji negali iki galo paaiškinti kryptingo girdėjimo, o taip pat išlieka priekinio-galinio garso šaltinio padėties girdimumo problema. Taip pat ši teorija apima tik horizontaliąją garso šaltinio lokalizaciją aplink galvą. Dar vienas teorijos netikslumas – neatsižvelgimas į ausies kaušelio formą lokalizacijos paaiškinime.

1938 metais Vudvortas atliko eksperimentą, kuriame iš pilnavidurio rutulio buvo sumodeliuota žmogaus galva. Su šiuo modeliu buvo matuojami LS kaip kampo su vertikalia plokštuma funkcija skirtingiems dažniams. Šiame galvos modelyje tarpas tarp dviejų ausų buvo apie 22-23 cm. Pirminiai testo rezultatai parodė, kad didžiausias laiko tarpas per kurį garsas patenka iš vienos ausies į kitą buvo apie 660 μ s, kai garso šaltinis buvo padėtas lygiai 90° kampu vertikalaus plokštumos atžvilgiu. Šis laiko tarpas koreliuojasi su 1500 Hz dažnio garso bangos ilgiu. Rezultatai parodė, jog grojant garsui, kurio dažnis mažesnis nei 1500 Hz bangos ilgis yra didesnis negu minėtojo didžiausio laiko tarpo. Taigi rezultatai parodo, jog tarp garso bangų patenkančių į ausis atsiranda fazių skirtumas ir tai sukuria akustinės lokalizacijos užuominas. Garso dažniui artėjant prie 1500 Hz garso bangos ilgis yra panašus į natūralų garso vėlavimą. Dėl galvos dydžio ir tarpo tarp ausų yra sumažėjęs fazių skirtumas, taigi iš karto atsiranda lokalizacijos paklaida. Garso dažniui esant didesniui nei 1500 Hz garso bangos ilgis tampa mažesnis nei atstumas tarp ausų, tuomet atsiranda „galvos šešėlis“ ir GLS suteikia garso lokalizacijos užuominų.

Feddersen et al. taip pat atliko eksperimentus (1957 metais), kuriais jis matavo kaip kinta LS keičiant garsiakalbio kampą su vertikaliaja plokštuma aplink galvą esant skirtingiems dažniams. Bet jo eksperimentai skyrėsi nuo „Woodworth“ tuo, kad Feddersen‘as atliko tyrimus su realiai žmonėmis, o ne su galvos modeliais. Testo rezultatai sutapo su „Woodworth“ gautais rezultatais. Tyrimo rezultatai taip pat parodė, kad nėra jokio LS skirtumo tarp dviejų garso šaltinio pozicijų – kai garso šaltinis yra prie žmogaus galvą (90°) ir už galvos (180°). Tokie rezultatų paaiškinimas – garso šaltinio atstumas iki ausų yra vienodas abiem atvejais. Laiko skirtumas keičiasi tik garso šaltiniui keliaujant aplink galvą. Šiuo testu įrodyta, kad didžiausias laiko skirtumas – 660 μ s gaunamas garso šaltinį padėjus 90° kampu prie vienos ausies.

3.1 pav. matoma, jog kelio skirtumas nuo garso šaltinio iki kiekvienos ausies yra skirtingas. Garso kelias iki kairės ausies yra didesnis nei kelias iki dešinės ausies. Šis garso kelių skirtumas perkeltas į lauko ašį nusako tarpausinį laiko skirtumą (ang. *ITD*, trump. LS). LS yra dominuojanti užuomina dažniams iki 1.5 kHz (imtinai). Visiems kitiems dažniams didesniems nei 1.5 kHz dominuojanti užuomina yra tarpausinis garso lygio skirtumas (ang. *ILD*, trump. GLS). Tarpausinis garso lygio skirtumas atsiranda dėl taip vadinamojo galvos šešėlio. Jis at-

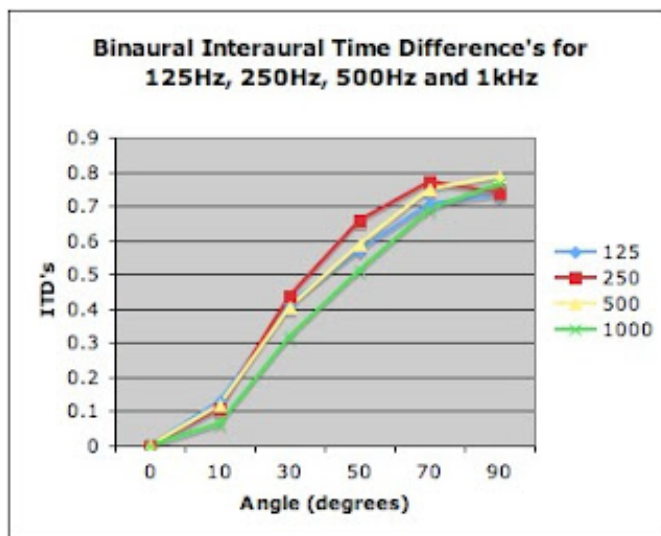
siranda tik prie aukštesnių dažnių, kuomet garso bangos nebesukelia difrakcijos aplink galvą, taip sukurdamos galvos šešėlį, kuris panašus į šešėlį susidarantį krentant šviesos bangoms link galvos. Norint įrodyti Lordo Reileigo teoriją, buvo atliktas eksperimentas, kuriame garso šaltinį ir manekeno galvą skyrė 1,75 m. Eksperimentas atliktas beaidėje aplinkoje. Jį atliko Levisas Skotas Diamondas (*Lewis Scott Diamond*). Naudojant skirtingo dažnio – 125 kHz, 250 kHz, 500 kHz ir 1000 kHz programa Csound sugeneruotus garsus, kai garso signalo šaltinis buvo padėtas skirtingais kampais – 0° , 10° , 30° , 50° , 70° , 90° , pavyko įrodyti šią teoriją.



3.3 pav. 100 kHz grynojo tono gauto manekeno galvoje oscilograma (viršuje – kairioji ausis, apačioje – dešinioji)

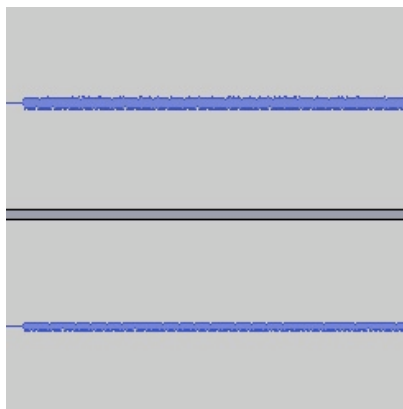
3.3 pav. matomi grafikai atvaizduoja manekeno galvoje esančių mikrofonų parodymus esant garso šaltinio skleidžiamam garsui, kurio dažnis siekė 1000 kHz ir atstumas iki jo – apie 1,75 m. Viršutinėje 3.3 pav. dalyje – kairiosios ausies mikrofono parodymai, apatinėje dalyje – dešinėsios. Iš 3.3 paveikslą matome, jog tarp abiejų mikrofonų parodymų susidaro fazių skirtumas, taip pat apatinėje paveikslėlio dalyje matomas atsiradęs vėlavimas lyginant su kairiosios ausies mikrofono grafiku. Pilno signalo tarpausinis laiko skirtumas siekė 0,735 ms, ir tai yra didžiausias pasiekiamas laiko skirtumas, nes manekeno galvos priekinė dalis buvo statmena garso signalo šaltiniui.

0,735 ms laiko skirtumas gali būti per didelis tarpausiniam laiko skirtumui, nes „maksimalus tarpausinis laiko skirtumas gali siekti tik 0,66 ms standartiniam žmogaus galvos dydžiui“. Taip teigiama 5-tame „Computational Auditory Scene Analysis“ knygos (išleista 2005 m) skyriuje pavadinimu „Binauralinio garso lokalizacija“. Knygos autoriai - DeLiang Wang ir Guy J. Brown. Kiti šaltiniai teigia, jog 90° kampo garso šaltinio padėtis prieš manekeno galvą gali sukurti 0,7 ms tarpausinį laiko skirtumą. Pastarasis straipsnis – „Binauralinė klausia“ autoriaus dr. Ruth



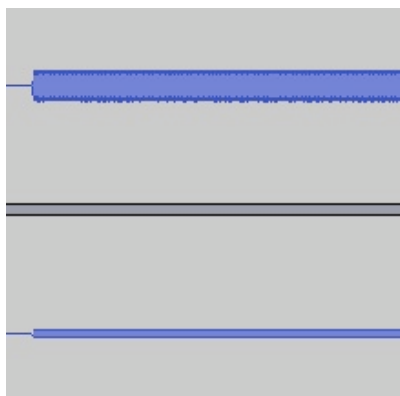
3.4 pav. Tarpausinio laiko skirtumo priklausomybė nuo garso kritimo kampo.

Y. Litovsky parašytas 2008 m. Šis straipsnis parodo, jog skirtumai tarp vidutinių žmogaus galvų dydžių sudaro maksimalaus tarpausinio laiko skirtumo vertės netikrumus. Jei ši vertė tampa per didelė, tai reiškia, kad mikrofonus reikia įdėti giliau į manekeno galvą, taip norint sumažinti atstumą tarp abiejų mikrofonų.



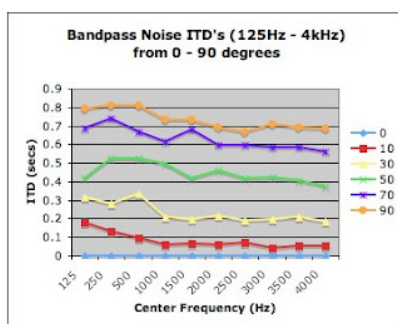
3..5 pav. Tarpausinis garso lygio skirtumas esant 125 kHz dažniui, kai garso šaltinio kampas manekeno galvos atžvilgiu siekia 90°

Iš 3.5 ir 3.6 pav. matome, jog lyginant du skirtingo dažnio garsus (125 kHz ir 1000 kHz), kai garso bangos kritimo kampas siekia 90° , susidaro tarpausinio garso lygių skirtumas. Kaip ir buvo minėta, esant aukštesniems dažniams nei 1,5 kHz, tarpausinis laiko skirtumas tampa labai mažas ir tuomet dominuoja tarpausinis garso lygių skirtumas. Fizikinis šio fenomeno paaiškinimas būtų toks, kad esant aukštesniems dažniams, garso bangos sutankėja ir jų galimybė persilenkti aplink galvą stipriai sumažėja. Esant skirtingiems objektų dydžiams ir skirtingiems garso bangų dažniams išlieka difrakcinė priklausomybė, bet šiuo atveju tai negalioja, nes naudojama standartinė vidutinio žmogaus galvos dydžio manekeno galva. Be abejo išlieka minimalūs



3.6 pav. Tarpausinis garso lygio skirtumas esant 1 MHz dažniui, kai garso šaltinio kampas manekeno galvos atžvilgiu siekia 90°

skirtumai, nes kiekvieno žmogaus galva turi savo unikalią formą ir dydį, o tai sukuria skirtingą šešėlinį dažnį, bet galima sakyti, jog šis skirtumas yra nykstamai mažas.



3.7 pav. Tarpausinio laiko skirtumo priklausomybė dažnio esant skirtingiems garso bangos kritimo kampams ($0^\circ - 90^\circ$)

3.7. Kambario impulsinis atsakas

Kambario Impulsinis Atsakas (sutrumpintai *KIA*, angl. *RIR* – *Room Impulse Response*) – tai yra laiko domeno signalas parodantis dažnių pasiskirstymą ir amplitudės mažėjimą. Naudojant *KIA* daroma prielaida jog kambario erdvė yra stabili ir linijinė, laikui bėgant nekintanti sistema.

KIA inkapsuliuoja sekančius parametrus:

1. Garso signalo vėlinimą (angl. *propagation delay*) – atstumas laike kurį garsas nukeliauja nuo garso šaltinio iki klausytojo.
2. Tiesioginį garsą (angl. *direct sound*) – tiesioginis garsas atitinką maksimumą kuris nusako trumpiausią atstumą nuo garso šaltinio iki klausytojo.
3. Ankstyvuosius atspindžius (angl. *early reflections*) – pirmos eilės (dažniausiai atspindį nuo grindų arba lubų), antrosios ir kitų eilių atspindžius.

4. Aido liekamąją dalį (angl. *reverberation tail*) – tai stochastinė aido dalis kurioje yra didelis atspindžių kiekis kurie nebegali būti suskirstyti į ankstyvuosius atspindžius.
5. Pradinio Vėlinimo Atotrūkis (PVA, angl. *ITDG – Initial Time Delay Gap*) – tai laikas tarp į klausytojo klausos sistemą atvykusio tiesioginio garso ir pirmojo ankstyvojo atspindžio (dažniausiai atspindėjusio ne nuo grindų). Dažnai norint išvengti papildomų skaičiavimų paprastas signalo vėlinimas yra panaudojamas PVA imituoti.

Šiame darbe yra panaudotas supaprastintas KIA modelis, darant prielaidą jog garso bangą sudaro baigtinis skaičius spindulių sklindančių tiesia linija visomis kryptimis. Ši prielaida yra klaidinga tačiau tokiu būdu yra ženkliai sumažinamas skaičiavimų kiekis kas savo ruožtu leidžia taikyti KIA modelį realiuoju laiku veikiančiame įrenginyje neturinčiame ypač didelių skaičiavimo galimybių. Priešingu atveju nepriėmus anksčiau minėtos prielaidos KIA generavimas tampa labai sudėtingas ir negali būti taikomas realaus laiko binauralinio garso generavimui.

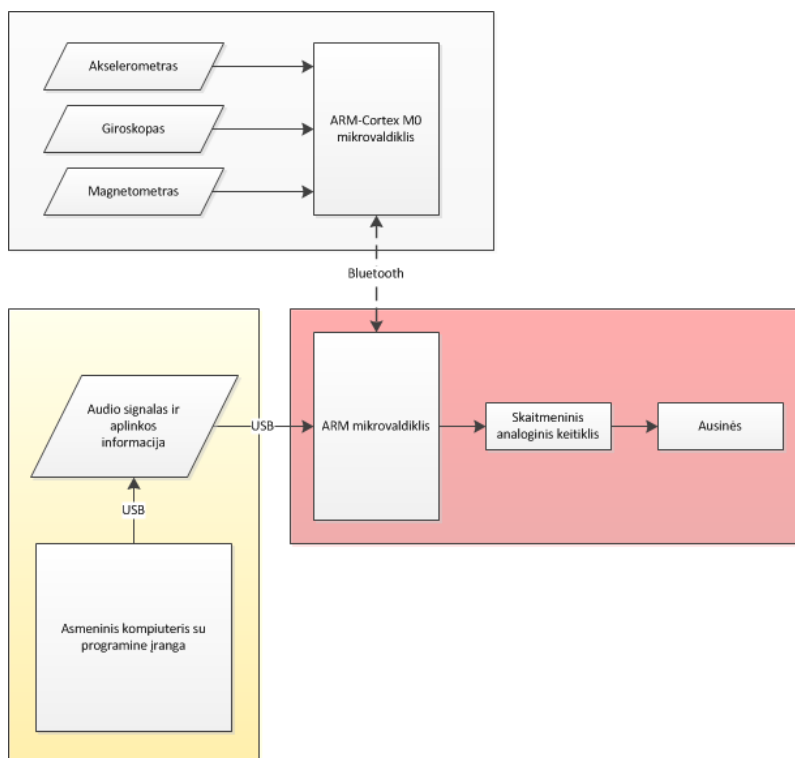
3.8. Skyriaus apibendrinimas

Šiame skyriuje buvo paminėti pagrindiniai metodai susiję su binauraliniu garsu. Minimaliam binauraliniam efektui gauti pakanka tik *HRTF* filtrų, bet kaip patvirtina atlikti bandymai - šis efektas nėra labai įtikinantis ir natūralus. Tačiau jei prieš panaudojant *HRTF* funkcijas garso signale yra užkoduojama kambario akustinė informacija, binauralinio garso efektas tampa panašesnis į realiam pasaulyje susidarančius garsus. Darbo metu sukurta garso apdorojimo sistema išnaudoja šiame skyriuje aprašytus metodus sukuriant pakankamai įtikinantį binauralinį efektą.

4. Sistemos kūrimas

Šiame skyriuje bus aptartas sistemos kūrimas, bendra struktūros apžvalga bei aprašyti naudojami algoritmai. Pagrindinis dėmesys bus skirtas galvos orientacijos įrenginio daviklių duomenų suliejimo algoritmui bei garso binauralinio garso generavimo metodų aprašymui.

4.1. Sistemos struktūrinė schema



4.1 pav. Struktūrinė schema

4.1 pav. pavaizduota baigiamojo bakalaurnio darbo struktūrinė schema. Sistemą sudaro trys dalys:

- Galvos sekimo įrenginys
- Garso apdorojimo įrenginys
- Asmeninis kompiuteris su specializuota programine įranga.

Galvos sekimo įrenginys skirtas sekti vartotojo galvos orientacijai erdvėje, tam kad, sistemai suteikti generuojamo garso priklausomybę nuo galvos orientacijos. Ši priklausomybė yra būtina norint kuo tiksliau sugeneruoti binauralinį garsą. Norint įgyvendinti nepriekaištingą galvos sekimo įrenginio veikimą, buvo suprojektuota ir pagaminta spausdintinė plokštė, atitinkanti fizinių išmatavimų ir elektros energijos sąnaudų reikalavimus. Taip pat parašytos dvi programinės įrangos – pačiam galvos orientacijos sekimo įrenginiui bei asmeniniam kompiuteriui.

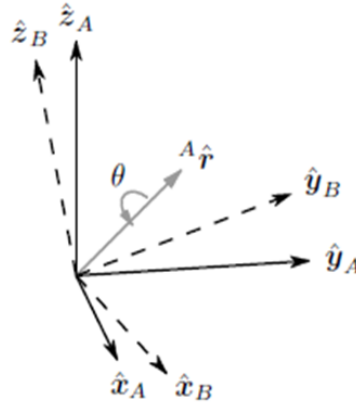
Sensor fusion algoritmas yra naudojami galvos orientacijos nuskaitymo įrenginyje. Galvos orientaciją erdvėje nuskaityti trys davikliai:

- triašis magnetometras – nustatyti kampą žemės magnetinių linijų ir magnetometro atžvilgiu.

- Visi šie trys davikliai turi savo stipriąsias ir silpnąsias puses, tačiau sensor fusion algoritmai padeda padidinti galutinių duomenų tikslumą. Dažniausiai naudojamas filtras – Kalmano, tačiau dėl didelio matematinio skaičiavimo kiekio (kuris neigiamai įtakoja bendrą mažos galios įterptinės sistemos greitaveiką) šiame darbe buvo pasirinktas paprastesnis Madgwick sensor fusion filtras. Palyginus su Kalmano filtru, Madgwicko filtras turi tik vieną žingsnį per kurį yra apskaičiuojami galutiniai duomenis, atsisakius nuspėjimo ir korekcijos žingsnių (kurie naudojami Kalmano filtre) sutaupomas laikas, ko pasekoje padidėja greitaveika. Tačiau atsisakius prieš tai minėtų dviejų žingsnių suprastėja filtro savybės, nes neatsižvelgiama į prieš tai buvusius duomenis.

22

$${}^A_B\hat{q} = \begin{bmatrix} q_1 & q_2 & q_3 & q_4 \end{bmatrix} = \begin{bmatrix} \cos\frac{\theta}{2} & -r_x\sin\frac{\theta}{2} & -r_y\sin\frac{\theta}{2} & -r_z\sin\frac{\theta}{2} \end{bmatrix} \quad (4.1)$$



4..2 pav. Dviejų rėmų orientacijos kitimas vienas kito atžvilgiu.

Paveikslas 4..2 nusako rėmo² (angl. *frame*) B orientaciją, jį pasukus kampą θ pagal $A_{\hat{r}}$ ašį, kai pradinė rėmo B orientacija buvo sutapusi su rėmo A orientacija.

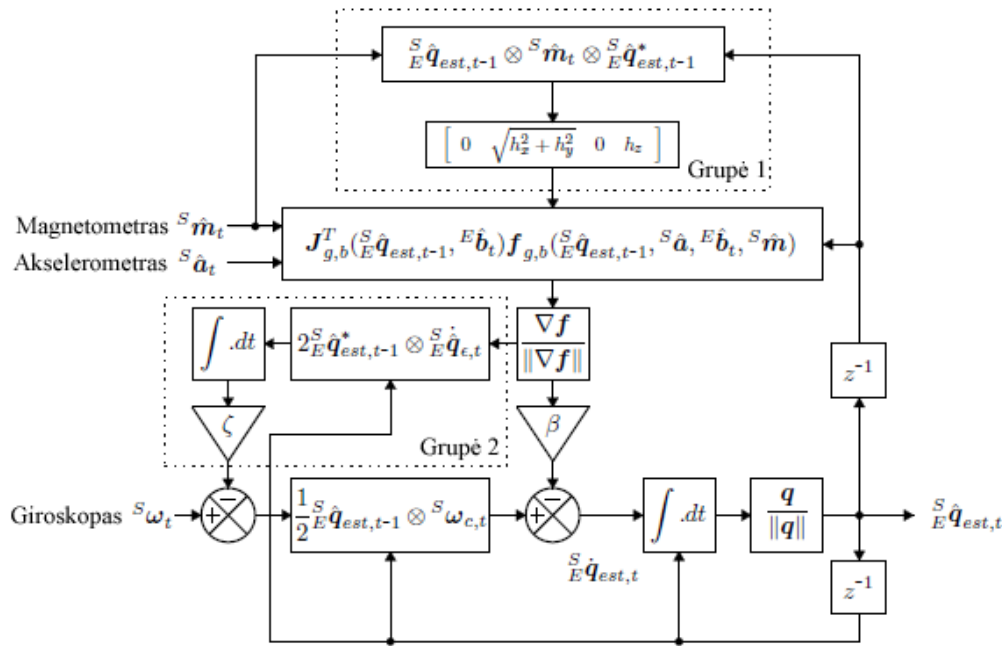
$${}^A_BR = \begin{bmatrix} 2q_w^2 - 1 + 2q_x^2 & 2(q_xq_y + q_wq_z) & 2(q_xq_z + q_wq_y) \\ 2(q_xq_y - q_wq_z) & 2q_w^2 - 1 + 2q_y^2 & 2(q_yq_z + q_wq_x) \\ 2(q_xq_z + q_wq_y) & 2(q_yq_z - q_wq_x) & 2q_w^2 - 1 + 2q_y^2 \end{bmatrix} \quad (4.2)$$

Taipogi kvaternioną galima išreikšti pasukimo matricos pavidalu 4..2, tai suteikia lankstumo skaičiuojant pasisukimą.

Matricinė kvaterniono išraiška labai patogi norint orientaciją paversti Eulerio sistemos kampais arba atliekant operacijas susijusias su kampais. Tačiau saugoti orientacijos informaciją tokiu pavidalu nėra patogu dėl anksčiau minėto ašių užsiblokavimo. Kaip matoma iš paveikslo 4..2 apskaičiuot kvaternionams reikalingi du rėmai, pirmasis atskaitos rėmas (angl. *reference frame*) antrasis - sistemos rėmas. Šiame darbe rėmu A (atskaitos rėmas) bus laikoma žemė, o rėmas B – galvos orientacijos įrenginys (sistemos rėmas).

Bendra filtro schema pavaizduota 4..3 pav. Kur ${}^S\hat{m}_t$ magnetometro duomenys, ${}^S\hat{a}_t$ – akcelerometro, ${}^S\hat{\omega}_t$ – giroskopo parodymai. 4..3 pav. parodo pilnąjį Madgwicko filtrą MARG (angl. *Magnetic, Angular Rate, and Gravity*) sistemai. Grupė 1 skirta pašalinti magnetometro magnetinius triukšmus (angl. *Magnetic distortion*), Grupė 2 – skirta pasinaudojant akcelerometro ir magnetometro duomenimis sumažinti giroskopo dreifą. Tokiu būdu filtras įgauna magnetometro triukšmo ir giroskopo dreifo kompensavimo savybes. Filtro rezultatas – ${}^S_E\hat{q}_{est,t}$ pilna objekto orientacija žemės atžvilgiu išreikšta kvaternionu.

²Geometrijoje objekto (linijos, plokštumos ar kietojo kūno) orientacija erdvėje yra dažnai pateikiama kaip vieno rėmo orientacija kito rėmo atžvilgiu, tokiu atveju antrasis rėmas vadinamas atskaitos rėmu (angl. *frame of reference*)

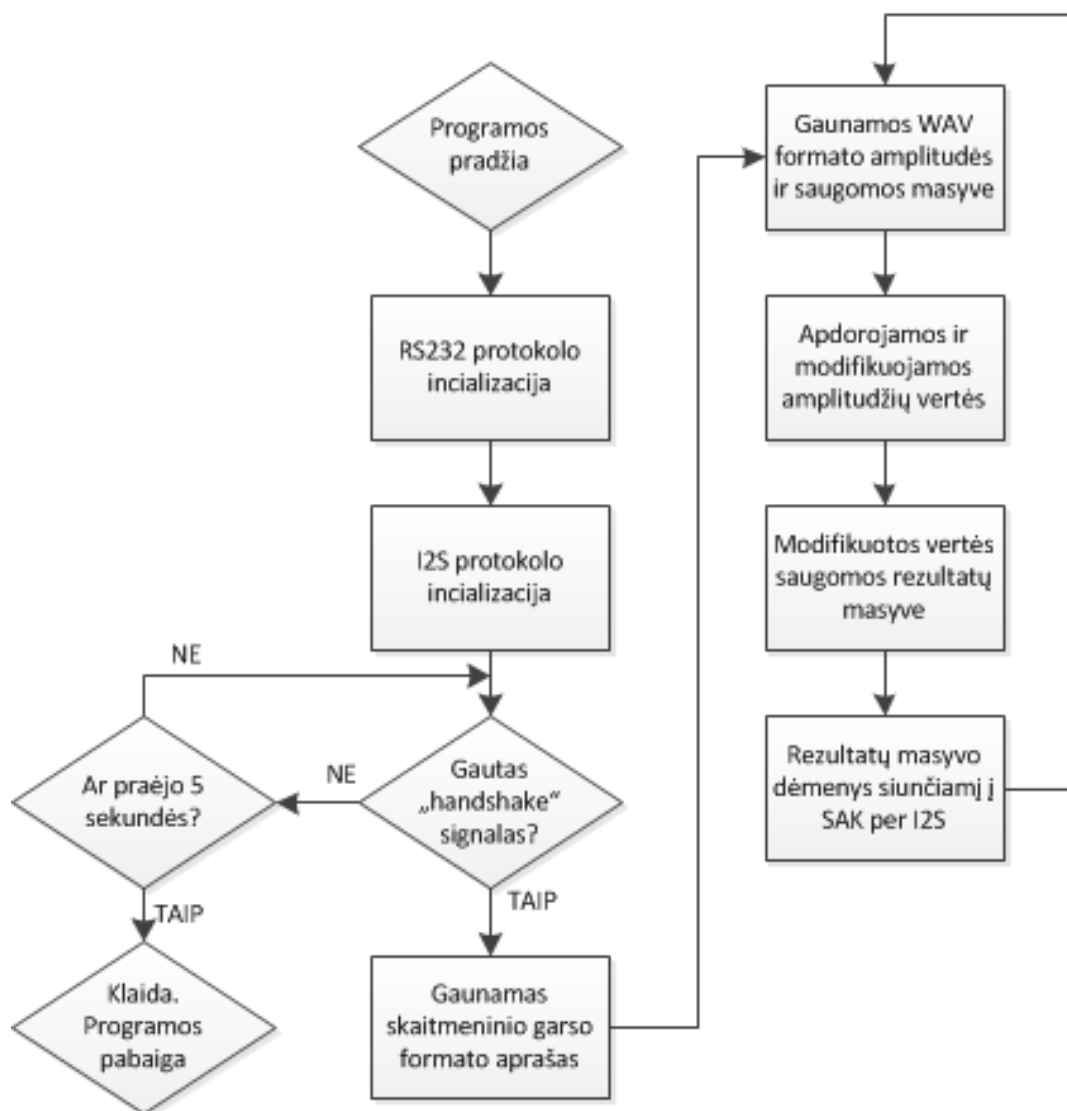


4.3 pav. Madgwicko filtro schema

4.3. Garso apdorojimo mechanizmas

Garso apdorojimo įrenginio programinę įrangą sudaro dvi dalys – programinė įranga asmeniniame kompiuteryje bei programinė įranga esanti pačiame garso apdorojimo įrenginyje. Norint įgyvendinti aparatinio lygmens garso išvestį, visų pirma, reikia skaitmeninio garso duomenų bei aparatinės įrangos skirtos skaitmeninį garsą paversti į analoginį (SAK – skaitmeninis analoginis keitiklis). Skaitmeninio garso informacija bei duomenys siunčiami naudojant *RS232* protokolą. Šių duomenų šaltinis – asmeninis kompiuteris su savo programine įranga. Skaitmeninio garso duomenų siuntimui į SAK naudojamas *I2S* protokolas. *I2S* – tai nuoseklus duomenų perdavimo protokolas, kuris buvo sukurtas sujungti skaitmeninio garso įrenginius tarpusavyje. Jis plačiai taikomas garso atkūrimo ir įrašymo pramonėje (*blue-ray*, *CD*, *DVD* grotuvuose ir kt.). Dažniausiai protokolas naudojamas sujungus tam tikro formato dekoderius ir skaitmeninius-analoginius keitiklius tarpusavyje. Protokolo fizinis lygmuo susideda iš trijų linijų – valdančioji bitų perdavimo dažnio linija, kanalo pasirinkimo linija bei duomenų perdavimo linija. Duomenys šiuo protokolu siunčiami sinchroniniu būdu. T.y. duomenys nuskaityti tik esant aukštam valdančiosios bitų perdavimo linijos įtampos lygiui. Kanalo parinkimas (kairiojo ar dešiniojo) vyksta keičiant kanalo parinkimo linijos įtampos lygį. Šis keitimas privalo būti atliktas prieš perduodant paskutinį esamojo kanalo duomenų bitą. 4.4 pav pavaizduotas garso apdorojimo įrenginio programos algoritmo blokinė diagrama.

Kaip matoma iš 4.4 pav., programos algoritmas prasideda nuo *RS232* bei *I2S* protokolų įgalinimo. *STM32F407-VGT*, kaip ir 8 bitų *AVR* mikrovaldiklis, turi aparatinio lygmens *RS232* protokolą. Pastarajame naudojamas 8N1 standartas – 8 duomenų perdavimo bitai, 1 start bi-



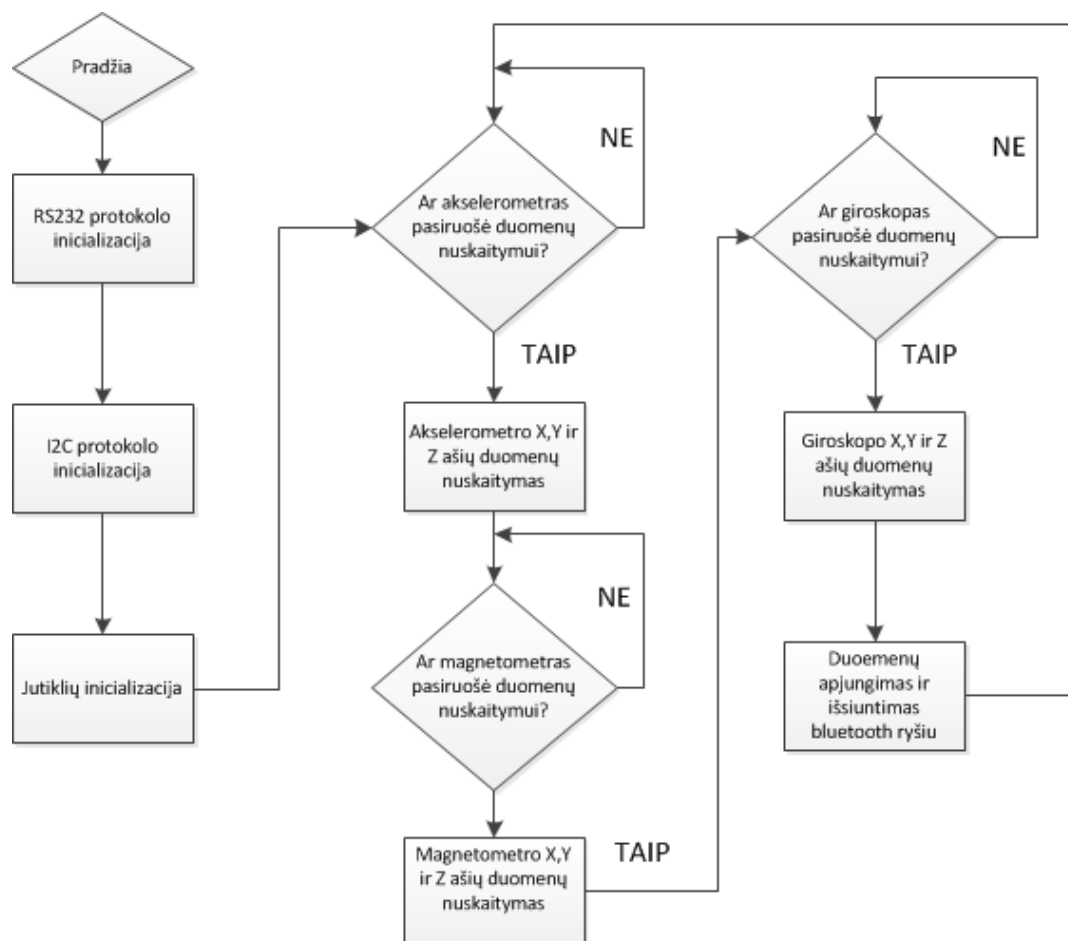
4.4 pav. Garso apdorojimo įrenginio programos algoritmo blokinė diagrama

tas, 1 stop bitas ir jokie lyginumo bito. Kaip ir buvo minėta anksčiau, skaitmeninio garso duomenų perdavimui naudojamas *I2S* protokolas. Pasirinktas 32 bitų mikrovaldiklis turi aparatinio lygmens *I2S* protokolą. Taigi, norint įgalinti protokolą, reikia įrašyti tam tikras vertes į atitinkamus registrus programiškai. Toliau programos algoritme seka „handshake“ signalas. Tai tam tikras sutartinis „pasisveikinimo“ signalas siunčiamas *RS232* protokolu asmeniniam kompiuteriui. Jeigu gaunamas teigiamas atsakymas, tuomet algoritmas tęsiamas. Jei negaunamas joks atsakymas, tikrinama ar praėjo 5 sekundės nuo „pasisveikinimo“ signalo išsiuntimo. Jei minėtasis laikas praėjo, tuomet registruojama programos algoritmo klaida ir programa yra užbaigiama. Jei 5 sekundės nepraėjo, tuomet toliau laukiamas atsakymas. Jei sulaukiamas „pasisveikinimo“ patvirtinimas, tuomet tęsiamas programos algoritmas. Tolimesnė programos eiga – gaunamas skaitmeninio garso duomenų aprašas. Pavyzdžiui – 16 bit duomenys vienam kanalui, 44,1 kHz diskretizavimo dažnis. Gavus šį aprašą, sukonfigūruojamos garso išvesties

funkcijos pagal gautą formatą. Toliau programa pereina prie begalinio ciklo, kuriame vyksta visos likusios operacijos galutiniam garso išvedimui į ausines. Visų pirma, per *RS232* sąsają gaunamos *WAV* failo amplitudžių vertės. Sekančiame žingsnyje vyksta visos apdorojimo bei modifikavimo, pagal esamą žmogaus galvos orientaciją, operacijos. Visos modifikuotos amplitudžių vertės paruošiamos galutiniam panaudojimui ir saugomos rezultatų masyve. Vertės iš rezultatų masyvo paeiliui siunčiamos į skaitmeninį-analoginį keitiklį *I2S* protokolui.

4.4. Galvos sekimo įrenginys

Šiame poskyryje aptariamas galvos sekimo įrenginio programos algoritmas. 4.5 paveiksle pavaizduota algoritmo blokinė schema.



4.5 pav. Galvos sekimo įrenginio programos algoritmo blokinė diagrama

Kaip matome iš 4.5 paveikslą, programos kodo pradžioje vyksta kintamųjų sukūrimas bei naudojamų įėjimų/išėjimų nustatymas. Kadangi naudojamų jutiklių duomenų perdavimo standartinis protokolas yra *I2C*, todėl reikia įgyvendinti *I2C* protokolo nustatymo mikrovaldiklyje operacijas. Pasirinktas mikrovaldiklis *Atmel Atmega8* savyje turi aparatinio lygio *I2C* valdymo bloką. Pagal specifikacinį *AVR* mikrovaldiklio aprašą, norint įgalinti *I2C* protokolą, reikia nustatyti duomenų perdavimo dažnį. Pasirinktas dažnis – 400 kHz. Tolimesnis programos kodas



– tai jutiklių pradinių parametrų nustatymas. Norint panaudoti *bluetooth* duomenų perdavimo ryšį, naudojamas *UART* protokolas. *Atmega8* mikrovaldiklis taip pat turi aparatinio lygio *UART* protokolo bloką. Taigi tinkamam pastarojo veikimui, reikia jį įgalinti programiškai kaip ir *I2C* protokolą. Pasirinkta duomenų perdavimo sparta – 57600 bodų per sekundę. Giroskopo *L3G4200D* pradiniai nustatymai:

- Energijos taupymo režimo išjungimas. T.y. jutiklio įjungimas Magnetometro *LSM303DLH* pradiniai nustatymai:
- Parenkamas duomenų atnaujinimo dažnis – 75 Hz
- Parenkamas $\pm 2,5$ Gauss jėjimo diapazonas
- Parenkamas nenutrūkstantis duomenų atnaujinimas

Akselerometro *LSM303DLH* pradiniai nustatymai:

- Parenkamas 1000 Hz duomenų atnaujinimo dažnis
- Išjungiamas energijos taupymo veikimo režimas.
- Išjungiamas vidinis duomenų filtras
- Nustatomas duomenų eiliškumas 8 bitų registruose

Atlikus jutiklių parametrų nustatymus, programos algoritmas pereina prie begalinio ciklo jutiklių duomenų nuskaitymui ir persiuntimui per *bluetooth* ryšį. Duomenys nuskaityti iš kiekvieno jutiklio šia eilės tvarka: akselerometro, magnetometro bei giroskopo duomenų nuskaitymas. Prieš kiekvieno jutiklio duomenų nuskaitymą yra laukiamas patvirtinimas, jog jutiklis yra pasirengęs perduoti duomenis. Todėl tikslaus duomenų atnaujinimo dažnio nustatyti neįmanoma, nes kiekvieną ciklo iteraciją jutiklių pasirengimas perduoti duomenis gali skirtis. Per kiekvieną begalinio ciklo iteraciją nuskaitytos visos 3 jutiklių ašys – *X, Y, Z*. Kadangi pastarieji yra padidinto tikslumo (12 bitų), tai nuskaityti duomenis 8 bitų mikrovaldiklio *I2C* protokolu, šiuos duomenis reikia apdoroti. Taip yra todėl, kad tiek jutiklių, tiek mikrovaldiklio registrai yra 8 bitų, bei *I2C* protokolas perduoda tik 8 bitus duomenų vienos užklausoje metu. Duomenų sujungimas šiuos atveju, tai tiesiog 8 aukščiausiųjų bitų perstūmimas į kairę lyginant su žemiausiais 8 bitais. Taip gaunamas 12 bitų tikslumas. Visi jutiklių duomenys po vienos ciklo iteracijos apjungiami į vieną string formato eilutę ir išsiunčiami *bluetooth* ryšiu, naudojant *UART* protokolą.