



## TURINYS

<b>Žymenys ir santrumpos</b> . . . . .	<b>8</b>
<b>1. Įvadas. Užduoties analizė</b> . . . . .	<b>9</b>
<b>2. Analogiškų sistemų apžvalga</b> . . . . .	<b>10</b>
<b>3. Naudojamų metodų apžvalga</b> . . . . .	<b>13</b>
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



## Žymenys ir santrumpos

- 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 nuosekioji jungtis (angl. *Universal Serial Bus*).
- COM - RS232 nuoseklusis prievadas.



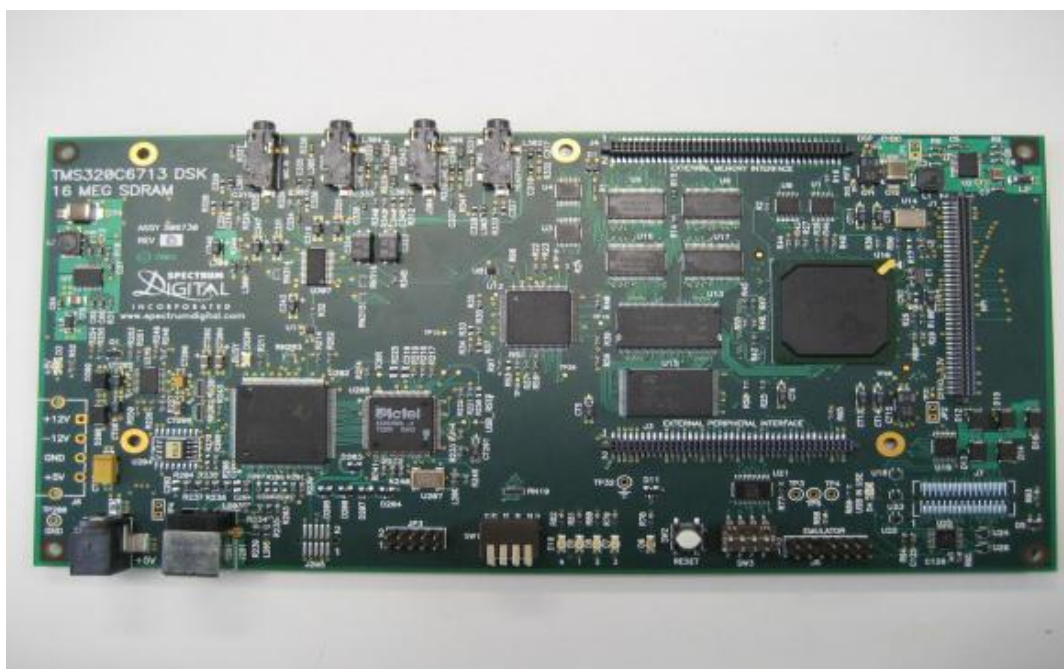
## 1. Įvadas. Užduoties analizė

Baigiamasis bakalaurinio baigiamojo darbo tema: „Virtuali garso realybės sistema“. Šiuo darbu siekiama ištirti galimybes bei sukurti realiuoju laiku veikiančią sistemą gebančią generuoti binauralinį garsą priklausomai nuo vartotojo galvos orientacijos ir virtualios aplinkos parametrų. Baigiamojo bakalaurinio darbo sistemos vartotojo sąsajos 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. Vartotojo sąsajos programinė įranga dirba Windows 7 operacinėje sistemoje, nes senesnės versijos (*XP*, *Vista*) jau yra nebepalaikomos Microsoft korporacijos. *USB* lizdo versija negali būti mažesnė nei 1.1, 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* prievada. Išlaikoma 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 gyvavimo laiką be įkrovimo 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 matmenys:  $50 \times 100 \times 100$  mm.

## 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 nusakyti lėktuvų trimatę poziciją. Pati *sensor fusion* technologija taip pat jau gana ilgai naudojama aviacijos srityje, norint kur kas tiksliau nusakyti poziciją. Kalbant apie binauralinio garso generavimą realiu laiku, paieškos rezultatai stipriai sumažėja. Tokio tipo projektų pasaulyje yra tik du. Projekto tikslai – 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 pasiekiantis 1800 milijonų instrukcijų per sekundę spartą. Taip pat ši plokštė gali apdoroti aukštos kokybės 24 bitų skaitmeminį stereo garsą, turi 512 kB *Flash* bei 16 MB *SDRAM* atmintis. *JTAG*<sup>1</sup> prieinamas per *USB* jungtį.

Kaip teigia gaminio kūrėjai, jie generuoja binauralinį garsą naudojant specialų stereo filtrą suprogramuotą minėtoje plokštėje. Programinė įranga naudoja nuo galvos priklausančias impulso atsako funkcijas 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

<sup>1</sup>JTAG - Programavimo ir testavimo jungtis.

testavimo akimirką demonstruoja 2..2 pav.



**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 „*Aural 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 „*Aural*“ 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 įsikišimo - 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.

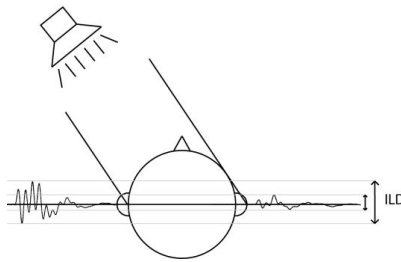
#### 3.2. Galvos šešėlis ir pečių aidas

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.







**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 sekančią 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  $\mu$ s, kai garso šaltinis buvo padėtas lygiai 90° kampu vertikalios 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  $\mu$ 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.