

BALSIŲ SPEKTRO RYŠIO SU AKUSTINIO VAMZDŽIO PARAMETRAIS TYRIMAS

Dainius Balbonas, Gintautas Daunys

Šiaulių universitetas, Technologijos fakultetas

Įvadas

Kalbos spektro tyrinėjimai susiję su šiuolaikiniais žmogaus ir kompiuterizuotų sistemų sąsajos veikimo principais, kurių vienas iš pagrindų – kalbos sintezė ir atpažinimas. Kalbos atpažinimas apima daugelį sričių: signalų technologijų, akustikos, intonacijos atpažinimo, komunikacijos ir informacijos teoriją, lingvistiką, psichologiją, kompiuterių mokslą, fiziologiją [1].

Atliekamas darbas siejamas su kalbos sintezės problemomis ir kartu su galimybėmis panaudoti elektronines priemones joms spęsti. Teksto pavidalu saugoma informacija, taikant sintezę, gali būti perskaitoma balsu. Aktualiausios kalbos sintezės vartojimo sritys: priemonės neįgaliesiems, kai kurie medicininiai taikymai, mokymo procesų tobulinimas, darbas su kompiuterizuotomis sistemomis (aptarnavimo srityse) ir telekomunikacinės paslaugos.

Kalbos sintezė neįgaliesiems (regėjimo ir motoriniai sutrikimai) leistų naudoti kompiuterizuotomis darbo vietomis ir įvairiomis informacinėmis sistemomis. Mokymo procese kalbos sintezę būtų galima panaudoti kalbos mokymuisi, logopedijai. Kalbos sintezės teikiamais privalumais plačiausiai pasinaudotų aptarnavimo ir telekomunikacinių paslaugų vartotojai – būtų galima pagerinti klientų aptarnavimą, komunikaciją virtualioje erdvėje pakelti į aukštesnį lygį. Balso technologijos jau senokai priskiriamos strateginei naujų telekomunikacinių paslaugų kūrimo kryptčiai [2].

Balso technologijos (kalbos sintezė ir atpažinimas) glaudžiai siejasi su kalbos signalų spektro tyrimais. Gerai žinoma, kad žmonių klausos sistemoje analizuojamos kalbos garsinio signalo spektrinės dedamosios, tuo tarpu laikinių charakteristikų įtaka nėra tokia akivaizdi ir, atrodo, kad joms tenka pagalbinis vaidmuo kalbos aiškume. Kalbos spektro

tyrimai išsiplėtė išaugus kompiuterių skaičiavimo pajėgumui. Signalų tyrimai atliekami skaitmeninio signalų apdorojimo metodais, plačiai taikomais elektronikoje [3–5].

Šiame darbe naudojamos nuosekliai sujungtos T tipo elektrinės grandinės kalbos trakto filtrui sukurti. Kalbos laikinis signalas gautas netiesiniu diferencialinių lygčių integravimo metodu. Toliau kalbos laikinis signalas taikytas formantų analizei. Tyrimo rezultatai lyginti su formantais, gautais iš kalbos trakto modeliavimo keturpoliu.

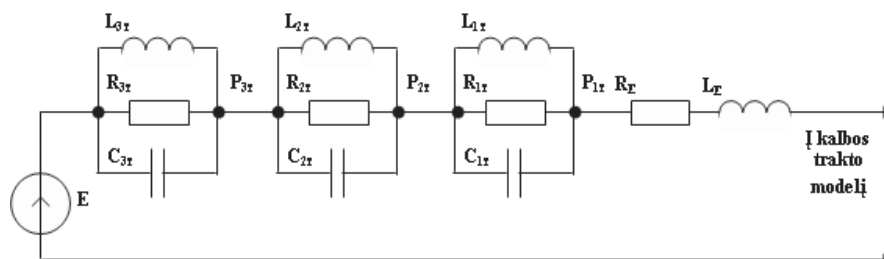
Darbo tikslas – elektroninėmis priemonėmis eksperimentiškai nustatyti balsių spektro ryšį su akustinio vamzdžio parametrais.

Uždaviniai: sukurti balsių spektro modeliavimo metodus, panaudojant elektrines grandines, atlikti balsių modeliavimą.

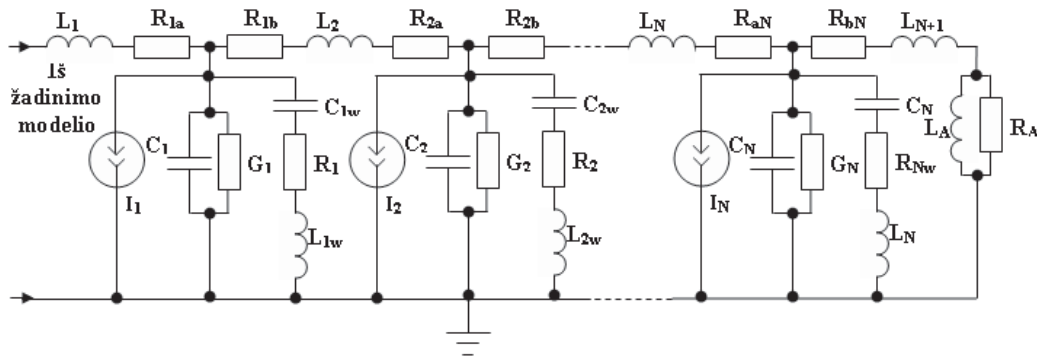
Tyrimo metodai

Pilnas elektroninis akustinio vamzdžio modelis sudarytas iš žadinimo dalies ir kalbos trakto (filtras). Žadinimo modelis parodytas 1 paveiksle, kalbos trakto modelis – 2 paveiksle.

Vienas iš jo praktinių įgyvendinimų, naudojant Matlab programos paketą, aprašytas G. Childers knygoje [6]. Tačiau pateiktą kodą sunku pritaikyti akustinio vamzdžio modeliui tobulinti. 2 paveiksle įtampos šaltinis žadinimo pradžioje modeliuoja atmosferos slėgį, srovės šaltiniai leidžia įskaityti papildomą oro srautą, atsirandantį dėl kalbos padargų judesių. RLC grandinė, įjungta lygiagrečiai talpai, leidžia įskaityti kalbos trakto sienelių svyravimus. Kandangi modelyje naudotos T tipo grandinės, todėl 2 paveiksle nurodyti induktyvumai $L_2 \dots L_N$ yra bendri abiem T tipo grandinėm. Atliekant skaičiavimus, šie induktyvumai dalijami į du atskirus, pvz., L_2 dalijamas į L_{2a} ir L_{2b} ir t. t.



1 pav. Žadinimo modelis

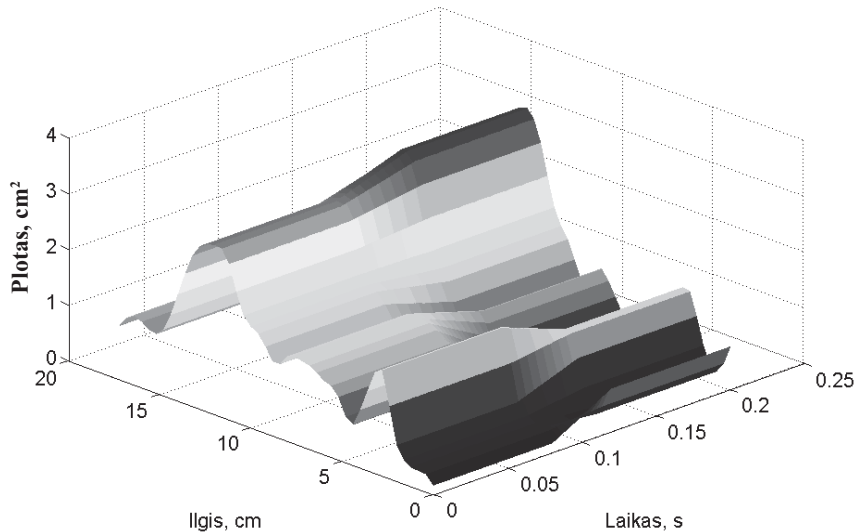


2 pav. Akustinio vamzdžio pilno signalo modelis

Keičiant T segmentų parametrus, galima keisti akustinio vamzdžio segmento skerspjūvio plotus ir taip imituoti perėjimą iš vieno balsio akustinio modelio į kito balsio akustinį modelį. Skerspjūvio plotų perėjimas tarp trumpo [i] ir ilgo [i] pavaizduotas 3 paveiksle, kuriame aiškiai matosi ilgojo [i] skerspjūvio ploto padidėjimas prieš kalbos trakto pabaigą, palyginus su trumpuoju [i], ir ploto sumažėjimas kalbos trakto viduryje. Skerspjūvio plotai iš modelio gaunami pasinaudojant L, C ir apkrovos varžos sąryšiais su skerspjūvio plotu. Sąryšiai pateikiami žemiau.

Ryšį tarp oro tūrio greičio (elektrinės srovės analogas I) ir oro slėgio (įtampos analogas U) galima paaiškinti per impedansą Z, pasitelkiant varžos,

talpumo ir induktyvumo sąvokas. Kadangi oro masė yra inertiška ir kartu spūdi, tai inertiškumą galima prilyginti induktyvumui L, spūdumą prilyginti talpai C. Vamzdžio, kurio skersmuo S, atkarpoje induktyvumas gali būti aprašomas $\rho\Delta l/S$ (ρ – oro tankis, Δl – segmento ilgis), o atitinkamai talpumas aprašomas $S\Delta l/\rho c^2$ (c – balso greitis ore), apkrovos varža $4\pi\mu\Delta l/S^2$ (μ – klampumas). Taigi, akustinį vamzdį galima modeliuoti panaudojant diskretinius elektronikos elementus [7]. Toks modelis nėra tikslus, kadangi vamzdžio skerspjūvis kinta, todėl keičiasi elementų parametrai pagal vamzdžio ilgį. Atvaizduojant diskretiniais elementais, reikia priimti, kad skerspjūvis tam tikroje vamzdžio atkarpoje nekinta.



3 pav. Skerspjūvio plotų kitimas laike priklausomai nuo kalbos trakto vietos, pereinant nuo [i.] į [i.]

Matematinis akustinio vamzdžio modelis sudarytas pritaikius būsenos kintamųjų metodą, kur būsenos kintamieji yra įtampa talpoje (atitinka garso slėgį) ir srovė induktyvume (atitinka oro tūrio greitį) [8]. Kiekvienam segmentui sudaromos 4 pirmos eilės diferencialinės lygtys. Lygčių koeficientai, esantys dešinėje lygties pusėje, yra priklausomi nuo laiko. Koeficientai keičiasi palengva dėl lėto kalbos trakto kitimo.

Pritaikius Kirchofo dėsnius, galima užrašyti įtampų ir srovių lygtis 2 paveiksle pateiktam akustinio vamzdžio modeliui.

Induktyvumo įtampą galima išreikšti kaip induktyvumo srovės išvestinę, padauginantą iš induktyvumo; talpumo srovę – kaip talpumo įtampos išvestinę, padauginantą iš talpos. Pasinaudojus srovių ir įtampų sąryšiais, iš įtampos ir srovių lygčių galima užrašyti diferencialines įtampų ir srovių lygtis:

$$\begin{aligned}
 L_1 \frac{di_{L_1}}{dt} &= E - u_{R_g} - u_{R_{1a}} - u_{C_1}, \\
 L_{1w} \frac{di_{L_{1w}}}{dt} &= u_{C_1} - u_{C_{1w}} - u_{R_{1w}}, \\
 L_2 \frac{di_{L_2}}{dt} &= u_{C_1} - u_{R_{1b}} - u_{R_{2a}} - u_{C_2}, \\
 L_{2w} \frac{di_{L_{2w}}}{dt} &= u_{C_2} - u_{C_{2w}} - u_{R_{2w}}, \\
 &\dots\dots\dots \\
 L_3 \frac{di_{L_3}}{dt} &= u_{C_2} - u_{R_{2b}} - u_{R_{3a}} - u_{C_3}, \\
 &\dots\dots\dots \\
 L_N \frac{di_{L_N}}{dt} &= u_{C_N} - u_{R_{bN}} - U_A, \\
 L_{Nw} \frac{di_{L_{Nw}}}{dt} &= u_{C_N} - u_{C_{Nw}} - u_{R_{Nw}}, \\
 C_1 \frac{du_{C_1}}{dt} &= i_{L_1} - i_{G_1} - i_{L_2} - i_{L_{1w}} - i_1, \\
 C_{1w} \frac{du_{C_{1w}}}{dt} &= i_{L_{1w}}, \\
 C_2 \frac{du_{C_2}}{dt} &= i_{L_2} - i_{G_2} - i_{L_3} - i_{L_{2w}} - i_2, \\
 C_{2w} \frac{du_{C_{2w}}}{dt} &= i_{L_{2w}}, \\
 &\dots\dots\dots \\
 C_N \frac{du_{C_N}}{dt} &= i_{L_N} - i_{G_N} - i_{L_{N+1}} - i_N, \\
 C_{Nw} \frac{du_{C_{Nw}}}{dt} &= i_{L_{Nw}}, \\
 \frac{U_A}{R + R_{bN}} &= i_{L_{N+1}},
 \end{aligned} \tag{1}$$

Lygtyse išvestinės pakeičiamos baigtiniais pokyčiais ir atliekamas netiesioginis integravimas. Tuo atveju ne tik išvestinės pakeičiamos baigtiniais pokyčiais, kartu vidurkinamos ir būsenos kintamųjų vertės tarp gretimų laiko momentų. Gaunama tiesinių lygčių sistema, kurią pertvarkius gaunama tiesinės lygčių sistemos matrica ir laisvųjų narių stulpelis.

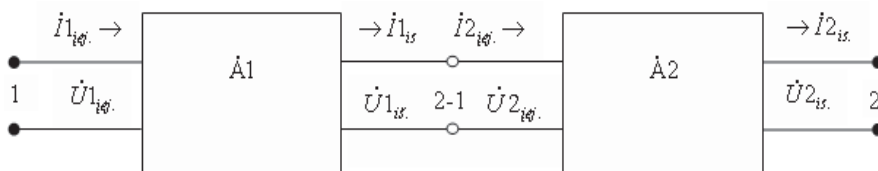
Svarbiausias iš lygčių sistemos sprendinių yra priešpaskutinis nežinomasis. Šis sprendinys nusako išėjimo srovės stiprį, kuris atitinka oro srautą kalbos trakto išėjime, o jo pirma išvestinė pagal laiką yra proporcinga mikrofону užregistruotam kalbos signalui.

Pasirinkus žadinimo signalą ir kalbos trakto skerspjūvių priklausomybę nuo laiko bei pritaikius

aukščiau aprašytą būsenos kintamųjų metodą, galima atkurti kalbos signalą. Iš laikinio signalo gaunamas dažnio-amplitudės signalas, taikant TP metodą.

Kalbos trakto amplitudės-dažnio charakteristika iš karto nesunkiai gali būti gaunama panaudojant keturpolius. Kalbos trakto modeliui aprašyti tinka laipsniškai sujungti keturpoliai. Laipsniškai sujungti keturpoliai atitinka T tipo liniją – vienas keturpolis atstoja vieną T linijos segmentą.

Keturpolio savybės nusakomos matricomis (2 x 2), kurios yra savitas keturpolio vaizdas. Laipsniškas jungimas aprašomas A parametrais. Dviejų keturpolių laipsniškas jungimas pavaizduotas 4 paveiksle [9].



4 pav. Laipsniškas dviejų keturpolių jungimas

Šiuo atveju vieno keturpolio išėjimas sujungtas su kito keturpolio įėjimu. Keturpolių sujungimo vietoje abiejų keturpolių srovių kryptis sutampa.

Abiejų keturpolių matricinė išraiška:

$$\begin{pmatrix} \dot{U}_{1_{iej.}} \\ \dot{I}_{1_{iej.}} \end{pmatrix} = \mathbf{A1} \cdot \begin{pmatrix} \dot{U}_{1_{is.}} \\ \dot{I}_{1_{is.}} \end{pmatrix}, \quad \begin{pmatrix} \dot{U}_{2_{iej.}} \\ \dot{I}_{2_{iej.}} \end{pmatrix} = \mathbf{A2} \cdot \begin{pmatrix} \dot{U}_{2_{is.}} \\ \dot{I}_{2_{is.}} \end{pmatrix}. \tag{2}$$

Kadangi $\dot{U}_{1_{is.}} = \dot{U}_{2_{iej.}}$ ir $\dot{I}_{1_{is.}} = \dot{I}_{2_{iej.}}$, todėl (4) esančias lygtis galima užrašyti taip:

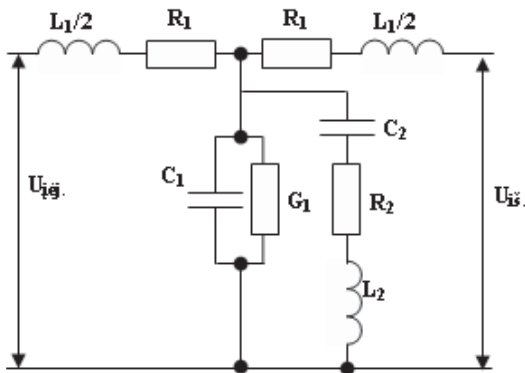
$$\begin{pmatrix} \dot{U}_{1_{iej.}} \\ \dot{I}_{1_{iej.}} \end{pmatrix} = \mathbf{A1} \cdot \mathbf{A2} \cdot \begin{pmatrix} \dot{U}_{2_{is.}} \\ \dot{I}_{2_{is.}} \end{pmatrix} = \mathbf{A} \cdot \begin{pmatrix} \dot{U}_{2_{is.}} \\ \dot{I}_{2_{is.}} \end{pmatrix}. \tag{3}$$

Taigi, laipsniškai sujungtų keturpolių A parametų matrica lygi abiejų keturpolių matricų sandaugai.

Išskleidę (3) lygybę, gausime:

$$\begin{aligned}
 \dot{A}_{11} &= \dot{A}_{11} \cdot \dot{A}_{21} + \dot{A}_{12} \cdot \dot{A}_{21}; \\
 \dot{A}_{12} &= \dot{A}_{11} \cdot \dot{A}_{22} + \dot{A}_{12} \cdot \dot{A}_{22}; \\
 \dot{A}_{21} &= \dot{A}_{21} \cdot \dot{A}_{11} + \dot{A}_{12} \cdot \dot{A}_{21}; \\
 \dot{A}_{22} &= \dot{A}_{21} \cdot \dot{A}_{12} + \dot{A}_{22} \cdot \dot{A}_{22}.
 \end{aligned} \tag{4}$$

Ankščiau paminėta T tipo liniją galima išreikšti elektrine schema, pavaizduota 3 paveiksle.



5 pav. Vieno akustinio vamzdžio segmento elektrinė schema

Šiuo atveju L ir C apskaičiuojami pagal formules:

$$L = \frac{\rho}{S} \cdot \Delta x, \quad C = \frac{S}{\rho c^2} \Delta x, \quad (5)$$

čia ρ – oro tankis, S – vamzdžio segmento skerspjūvis, Δx – vamzdžio sekcijos ilgis, c – garso greitis ore. A parametrai 2 paveiksle pavaizduotos grandinės atveju:

$$a_{11} = 1 + \frac{(R_1 + 0,5j\omega L_1)(Z_n + Z_l)}{Z_n Z_l},$$

$$a_{12} = R_1 + j\omega L_1 + R_2 + \frac{(R_1 + 0,5j\omega L_1)(R_2 + 0,5j\omega L_1)(Z_n + Z_l)}{Z_n Z_l},$$

$$a_{21} = \frac{Z_n + Z_l}{Z_n Z_l}, \quad a_{22} = 1 + \frac{(R_2 + 0,5j\omega L_1)(Z_n + Z_l)}{Z_n Z_l}. \quad (6)$$

Lentelė. Modeliuotų stacionarių balsių formantai ir jų standartiniai nuokrypiai

Garsas	$F_1, \text{ Hz}$	$F_2, \text{ Hz}$	$F_3, \text{ Hz}$	$F_4, \text{ Hz}$	$F_5, \text{ Hz}$
[a]	509,2 ± 8,1	1293,6 ± 3,3	2437,7 ± 3,6	3446,4 ± 17,6	4374,0 ± 2,4
[a:]	668,4 ± 1,3	1081,3 ± 7,6	2378,9 ± 8,5	3398,9 ± 26,1	3971,6 ± 4,0
[e]	487,6 ± 2,3	1396,7 ± 11,8	2293,3 ± 5,6	3503,1 ± 1,5	4068,0 ± 20,0
[e:]	595,9 ± 8,6	1471,3 ± 5,0	2281,8 ± 9,3	3358,5 ± 16,0	4251,3 ± 8,6
[i]	398,9 ± 5,7	1551,6 ± 9,9	2297,6 ± 3,5	3500,5 ± 8,2	4319,4 ± 25,0
[i:]	405,1 ± 5,9	1588,0 ± 1,5	2354,4 ± 3,1	3617,7 ± 6,9	4496,4 ± 15,6
[o:]	578,4 ± 2,1	934,8 ± 1,6	2365,6 ± 18,9	3315,8 ± 28,5	3690,1 ± 2,8
[o]	401,7 ± 5,4	1000,1 ± 1,0	2414,2 ± 24,8	3399,9 ± 7,9	4191,6 ± 107,5
[u:]	378,3 ± 2,0	1128,5 ± 8,2	2221,6 ± 19,8	3609,3 ± 45,1	4100,7 ± 2,8
[u]	294,9 ± 5,3	891,0 ± 3,8	2151,4 ± 35,8	3522,6 ± 145,8	3970,6 ± 49,3

4–8 paveiksluose pateikiami minėtų penkių balsių laikinės ir dažnio amplitudės modeliavimo rezultatai, kai modeliuojamas perėjimas iš trumpo į ilgą balsį. Formantų kitimo išsitiesinės linijos gautos

$$\text{Čia } Z_l = \frac{1}{j\omega C_1 G_1 \left(\frac{1}{j\omega C_1} + \frac{1}{G_1} \right)} \text{ ir } Z_n = R_2 + j\omega L_2 + \frac{1}{j\omega C_2}.$$

Skaičiuojant akustinio vamzdžio dažninę charakteristiką, reikia įvertinti apkrovos impedanso įtaką. Sakykime, kad apkrovos impedansą sudaro lygiagrečiai sujungta varža R ir induktyvumas L_r , kurių reikšmės surandamos iš lygčių A parametru matricių.

$$R = \frac{128\rho c}{9\pi^2 S}, \quad L_r = \frac{8\rho}{3\pi\sqrt{\pi S}}. \quad (7)$$

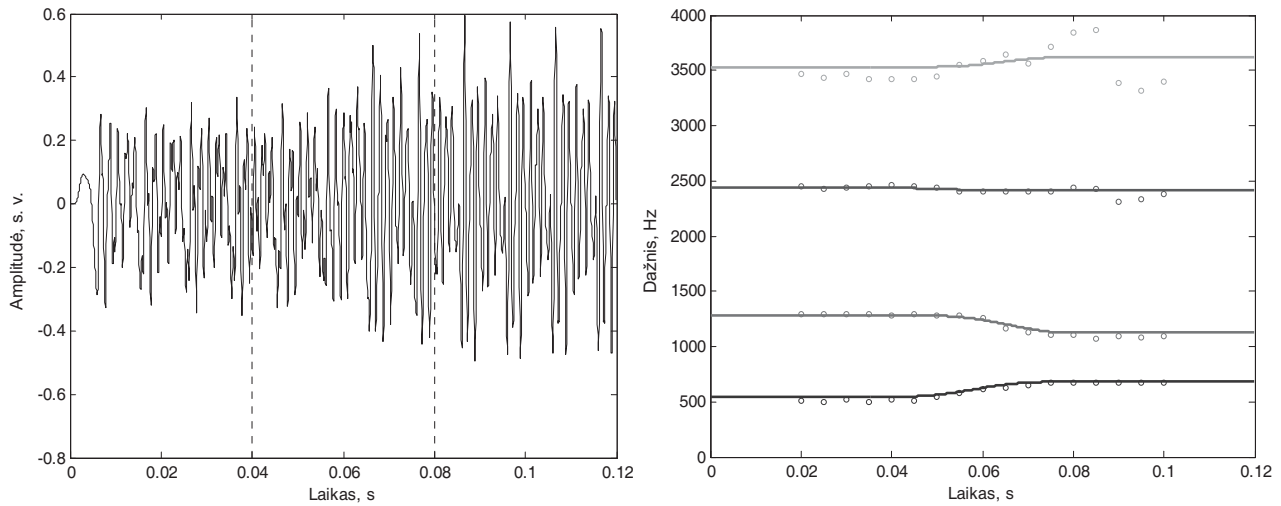
Aprašyta metodika įgalina greitai suskaičiuoti kalbos trakto filtro dažninę charakteristiką, aprėpiant daug akustinio vamzdžio segmentų.

Tyrimo rezultatai

Pritaikius aprašytą elektroninio akustinio vamzdžio modelį diferencialinių lygčių sistemoms integruoti, gauti įvairių balsių kalbos signalai. Rezultatai pateikiami 4–8 paveiksluose ir lentelėje.

Lentelėje pateikiami 5 sumodeliuotų stacionarių balsių pirmi penki formantai ir jų standartiniai nuokrypiai. Standartiniai nuokrypiai gauti taikant lango funkciją (formantų dažniai apskaičiuojami siaurame lange, paskui langas perstumiamas ir vėl skaičiuojami formantų dažniai).

vartojant modeliavimo keturpolius. Skrituliukais pažymėtos formantų vertės nustatytos iš simuliuotų laikinių signalų, pritaikius Praat programoje tiesinės prognozės analizę.

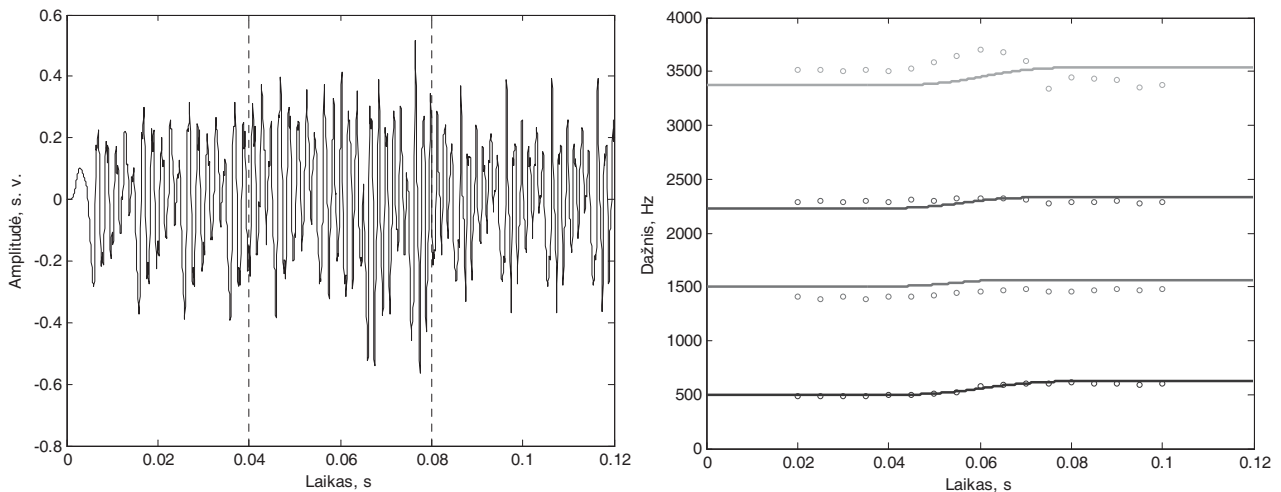


6 pav. Sumodeliuotas perėjimas iš [a] į [a:] garsą (kairėje) ir to perėjimo formantai (dešinėje, nuo apačios į viršų F_1, F_2, F_3, F_4)

Kaip matyti iš 4 paveikslų, esant perėjimui iš [a] į [a:], matomi amplitudės pokyčiai laikinėje diagramoje ir formantų dažnio pokyčiai dažninėje diagramoje. Perėjimo pradžia (0,04 s) ir pabaiga (0,08 s) laikinėje diagramoje pažymėtos punktyrinėmis linijomis. Šioje laiko atkarpoje matyti, kaip keičiasi F_1 ir F_2 formantų dažniai. Šiuo atveju [a] F_2 formanto dažnis mažėja, o F_1 didėja. Būtent, pirmieji du formantai yra reikšmingiausi, kiti formantai mažiau veikia garso kokybę. Kitų garsų paveiksluose matyti

panaši situacija. Visais atvejais pereinant nuo trumpo į ilgą garsą, pastebimi signalo amplitudės pokyčiai, o perėjimo laikinėje atkarpoje – formantų pokyčiai.

Iš 5 paveikslų matyti, kad, pereinant iš [e] į [e:], formantų F_1 ir F_2 dažniai didėja. Tačiau skiriasi antro formanto dažniai, gauti skirtingais metodais. Pastebimas nežymus F_3 dažnio, nustatyto keturpolių metodu, didėjimas. Tuo tarpu Praat programa formanto dažnio didėjimo neduoda.



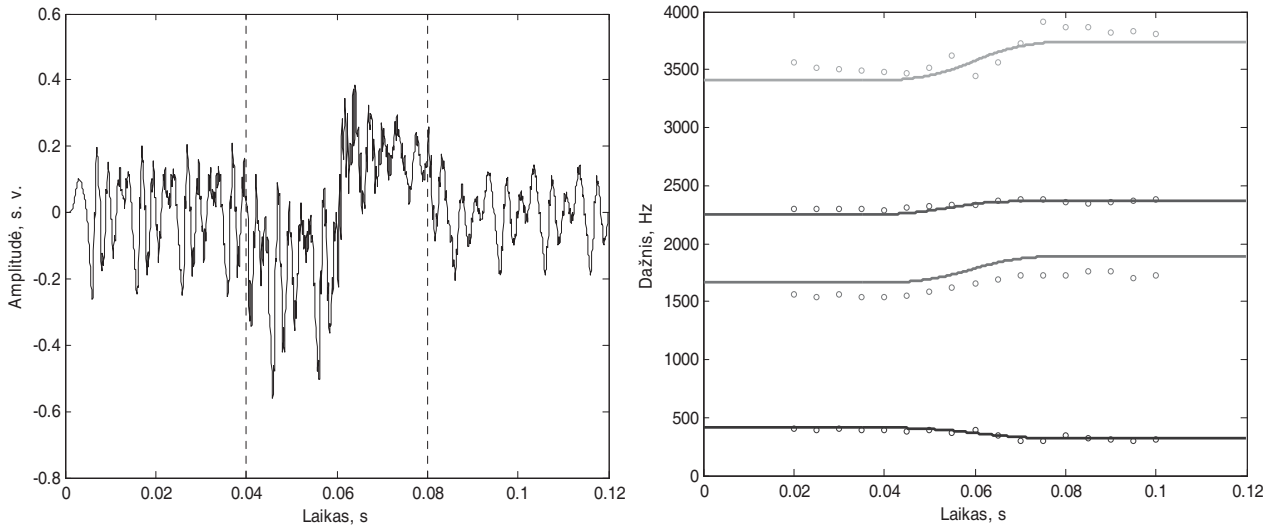
7 pav. Sumodeliuotas perėjimas iš [e] į [e:] garsą (kairėje) ir to perėjimo formantai (dešinėje, nuo apačios į viršų F_1, F_2, F_3, F_4)

Pereinant nuo [i] į [i:], pastebimas pirmo formanto dažnio mažėjimas, o antro formanto – didėjimas (6 pav.). Vėl skiriasi F_2 dažniai, nustatyti skirtingais metodais. F_3 formanto dažnis beveik nekinta.

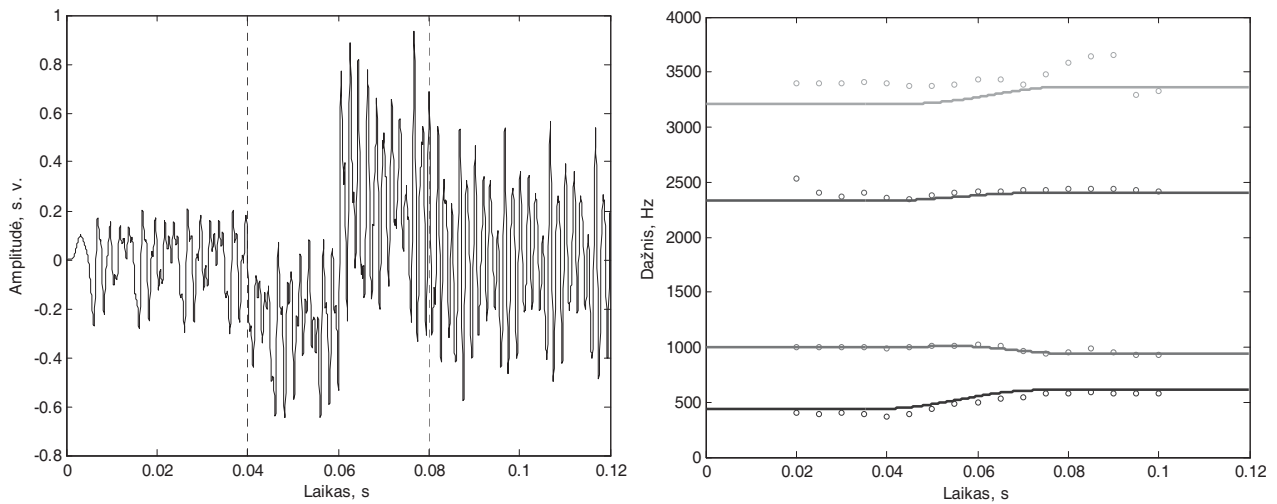
7 paveiksle atvaizduoti garsų [o] ir [o:] duomenys rodo ryškų amplitudės šuolį ir didelį formanto F_1 didėjimą, F_2 formanto dažnio sumažėjimą perėjimo

pabaigoje ir F_3 padidėjimą perėjimo pradžioje.

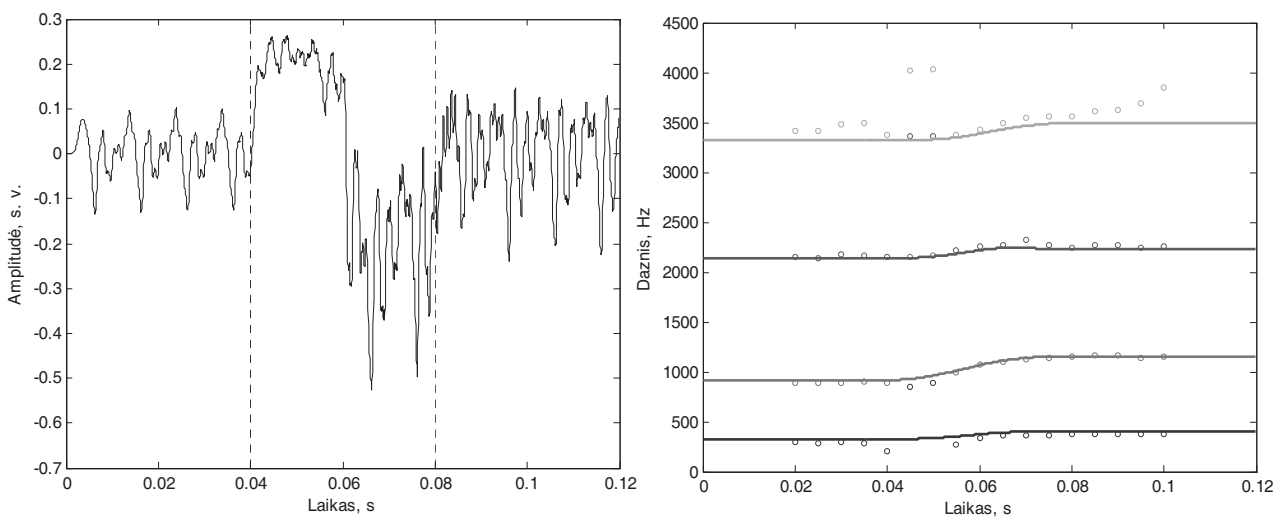
Garso [u] ir [u:] duomenys pateikti 8 paveiksle taip pat rodo formantų dažnio ir amplitudės kitimą. Iš 8 paveikslų matyti, kad Praat programa neteisingai nustatė formantus garso kitimo momentais. Formanto F_1 dažnis buvo per daug sumažintas ir atmestas. Todėl visos formantės tapo blogai sunumeruotos.



8 pav. Sumodeliuotas perėjimas iš [i] į [i:] garsą (kairėje) ir to perėjimo formantai (dešinėje, nuo apačios į viršų F_1, F_2, F_3, F_4)



9 pav. Sumodeliuotas perėjimas iš [o] į [o:] garsą (kairėje) ir to perėjimo formantai (dešinėje, nuo apačios į viršų F_1, F_2, F_3, F_4)



10 pav. Sumodeliuotas perėjimas iš [u] į [u:] garsą (kairėje) ir to perėjimo formantai (dešinėje, nuo apačios į viršų F_1, F_2, F_3, F_4)

Šie tyrimo rezultatai parodė, kad intensyvumas priklauso nuo perėjimo iš trumpo į ilgą garsą, vyksta kartu su kalbos trakto skerspjūvio pokyčiu.

Išvados

1. Formantų dažnių kitimo modeliavimas, atliktas dviem metodais, parodė, kad kalbos signalas tiksliau modeliuojamas laiko srityje, bet, taikant TP metodą, sunku iš laiko signalo surasti tikslius formantų dažnius.
2. Nuosekliai sujungtų keturpolių metodas, kurį vartojant tiesiogiai galima gauti formantų dažnius, parodė geresnius rezultatus.

Literatūra

1. Rabienr L., Juang B.-H., 1993, *Fundamentals of speech recognition*. Prentice-Hall International.
2. Rudžionis A., Ratkevičius K., Rudžionis V. ir kt., Projekto „Lietuvių kalbos bazinių elementų diegimas kompiuterinėse sistemose“ skyriaus „Balso technolo-

gijų taikymo lietuvių kalbai analizė ir perspektyvinių veiklos krypčių pagrindimas“ ataskaita. <http://www.likit.lt/frames/balzo_tech/balsotech_st.htm>.

3. Gold B., Morgan N., 2000, *Speech and audio signal processing*. John Wiley & Sons.
4. O' Shaughnessy D., 2000, *Speech Communications. Human and Machine*. New York: IEEE PRESS.
5. Deller J. R., Hansen J. H. L., Proakis J. G., 2000, *Discrete-Time Processing of Speech Signals*. New York: IEEE.
6. Childers D. G., 2000, *Speech processing and synthesis toolboxes*. John Wiley & Sons.
7. Maeda Sh., 1982, A digital simulation method of the vocal-tract system, *Speech Communication*. Nr. 1. P. 199–229.
8. Balbonas D., Daunys G., 2009, Kalbos trakto modeliavimas elektroninėmis grandinėmis. *Virtualūs instrumentai biomedicinoje 2009*. Klaipėda: Klaipėdos universiteto leidykla. P. 109–115.
9. Balbonas D., Daunys G., 2008, Lietuvių kalbos balsių spektro modeliavimas akustiniu vamzdžiu, *Biomedicininė inžinerija 2008*. Kaunas: Technologija. P. 249–252.

INVESTIGATION INTO RELATIONSHIP BETWEEN SPECTRUM OF VOWELS AND PARAMETERS OF ACOUSTIC TUBE

Dainius Balbonas, Gintautas Daunys

Summary

This research is related to problems of speech synthesis and possibilities to use electronic circuits for solving these problems.

Voice technologies (speech synthesis and recognition) are closely related to speech signal spectrum analysis. It is well known that the human auditory system is analyzed with an acoustic speech signal spectral components, while the temporal characteristics influence is not so clear, and it seems that they play an auxiliary role in speech. The continuous increase in computer calculation capacity expanded the studies of language spectrum. Studies of signals are being carried out by using digital signal processing methods that are now widely used in electronics.

A vocal tract can be described using the electric circuits of T-type connected in series. Its frequency response was obtained by two ways: simulation of the signal and later calculation of the spectrum, and calculation of frequency response of serially connected two-ports. In the first case, a mathematical model of the acoustic tube was built by using a state variable approach, where the state variables are capacitor voltages (equivalent to sound pressure) and the inductance currents (equivalent to air volume speed). An obtained system of linear differential equations was integrated by indirect method. In the second case, two-port parameters were used to obtain all system frequency response. During vowel pronunciation the areas of the vocal tract cross-section are changing over time. In both methods formant trajectories were extracted from the vocal tract frequency response. The results of both methods were compared. The trajectories of formants are more stable when they are obtained using a second approach.

Keywords: vocal tract, formants, spectrum of vowels.

BALSIŲ SPEKTRO RYŠIO SU AKUSTINIO VAMZDŽIO PARAMETRAIS TYRIMAS*Dainius Balbonas, Gintautas Daunys***Santrauka**

Atliekamas darbas siejamas su kalbos sintezės problemomis ir kartu su galimybėmis pritaikyti elektronines priemones šioms problemoms spęsti.

Balso technologijos (kalbos sintezė ir atpažinimas) glaudžiai susiję su kalbos signalų spektro tyrimais. Gerai žinoma, kad žmonių klausos sistemoje analizuojamos kalbos garsinio signalo spektrinės dedamosios, tuo tarpu laikinių charakteristikų įtaka nėra tokia akivaizdi ir, atrodo, kad joms tenka pagalbinis vaidmuo kalbos aiškume. Kalbos spektro tyrimai išsiplėtė išaugus kompiuterių skaičiavimo pajėgumui. Signalų tyrimai atliekami skaitmeninio signalų apdorojimo metodais, kurie dabar plačiai vartojami elektronikoje.

Kalbos traktas gali būti atvaizduotas, panaudojant nuosekliai sujungtas T tipo elektrines grandines. Dažninių charakteristikų modeliavimui taikyti du būdai: modeliuojant signalą bei skaičiuojant jo spektrą ir skaičiuojant nuosekliai sujungtų keturpolių dažninę charakteristiką. Pirmuoju atveju sudarytas matematinis akustinio vamzdžio modelis pritaikius būsenos kintamųjų metodą, kai būsenos kintamieji yra kondensatorių įtampos (atitinka garso slėgius) ir induktyvumų srovės (atitinka oro tūrio greičius). Gauta tiesinių diferencialinių lygčių sistema integruojama netiesioginiu metodu. Antruoju atveju skaičiuojami visų keturpolių a parametrai ir iš jų gaunama visos sistemos dažninė charakteristika. Tariant balses, kinta kalbos trakto skerspjūviai, todėl elementų parametrai kinta laike. Abiem metodais gavus kalbos trakto dažnines charakteristikas, buvo nustatyti formantai. Abiejų metodų rezultatai lyginti tarpusavyje. Modeliuojant antruoju metodu, gaunamos pastovesnės formantų kitimo kreivės.

Prasminiai žodžiai: kalbos traktas, formantai, balsių spektrinės charakteristikos.

Įteikta 2009-09-02