

## Fonemų klasifikavimas panaudojant garso ir vaizdo informaciją

**D. Balbonas, G. Daunys**

*Technologijos fakultetas, Šiaulių universitetas,*

*Vilniaus g. 141, LT-76353 Šiauliai, Lietuva, tel. +370 41 595830; el. p. dainius@tf.su.lt*

### Įvadas

Šnekos atpažinimo technologijos ateityje turėtų vaidinti svarbų vaidmenį žmogaus ir kompiuterio sąveikoje. Šiuo metu šnekai atpažinti daugiausia yra naudojamas vien tik garso signalas, kurio analizei taikomas paslėptųjų Markovo modelių metodas. Iš pasaulinės praktikos aiškėja, kad atpažinimo šiuo metodu tikslumas nebedidėja, nors ir dedama daug pastangų. Automatinį šnekos atpažinimą ypač veikia pašalinis triukšmas. Išeitis - ieškoti, kaip šneką atpažįsta žmonės, kurie tai daro kur kas tiksliau negu šiuolaikinės kompiuterinės sistemos.

Nagrinėjamos dvi galimybės, kaip būtų galima pagerinti šnekos atpažinimą:

- garso signalo intervalų klasifikavimas į fonemų grupes;
- vaizdo signalo analizė.

Diferencinius fonemų požymius siūloma naudoti todėl, kad tokios fonemų skirstymo schemas jau seniai yra naudojamos fonetikoje. Automatiniam šnekos atpažinimui pirmieji šiuos požymius pradėjo taikyti Masačusetso technologijos instituto mokslininkai [1,2]. D. Shipman ir W. Zue darbe [2] pateikiamas pavyzdys, kad angliškas žodis *splint* gali būti atpažintas net neatpažįstant priebalsinių fonemų /s/, /p/, /n/. Šis žodis 20 000 žodžių žodyne yra vienintelis, kuris tenkina šabloną: priebalsis – priebalsis – /l/ – balsė – nosinis garsas – sprogstamasis garsas. Fonetinių vienetų diskriminavimo lietuvių kalbai metodas yra tiriamas VU Kauno humanitariniame fakultete [3].

Gerai žinoma, kad žmonės su klausos negalia gali gana gerai išmokti suprasti šneką iš lūpų judesių. Todėl šnekai automatiškai atpažinti, burnos vaizdo informacija buvo pradėta kombinuoti su garso signalu. Mažiausias kalbos vienetas, išskiriamas vaizdo signale, buvo pavadintas vizema (angl. *viseme*). Vaizdinė informacija ypač naudinga triukšmingoje aplinkoje [4].

### Fonemų klasifikavimo požymiai

Fonemų skirstymo į stambias grupes metodika ir požymiai išsamiai yra aptarti D. Childers monografijoje [5].

Siūloma atlikti šiuos veiksmus:

1) skardžiųjų bei duslių garsų ir tylos intervalų nustatymas (angl. *voiced/unvoiced/silence detection*);

- 2) balsingųjų garsų (sonorantų) nustatymas;
- 3) balsių nustatymas;
- 4) skardžiųjų priebalsių nustatymas;
- 5) balso juostos nustatymas (angl. *voice bar detection*);
- 6) nosinių garsų nustatymas;
- 7) pusbalsių nustatymas;
- 8) skardžiųjų friktyvinių priebalsių nustatymas;
- 9) duslių sprogstamųjų ir friktyvinių priebalsių nustatymas.

Čia įgyvendinome 1-4 ir 7 veiksmus, pridėdami fonemų trukmės nustatymą. Nustatant daugumą požymių, yra naudojama garsumo funkcija  $V(i)$  [5] (angl. *volume function*). Indeksas  $i$  žymi intervalo numerį, kuriame apskaičiuota funkcija. Garsumo funkcija randama iš lygties:

$$V(i) = \frac{1}{N_i} \sqrt{\sum_{m=A}^B \left| H_i \left( e^{j\pi \frac{m}{M}} \right) \right|^2}, \quad (1)$$

čia  $N_i$  – atskaitų skaičius  $i$  intervale,  $A$  – juostinio filtro apatinę ribą atitinkančio dažnio indeksas,  $B$  – viršutinę ribą atitinkančio dažnio indeksas,  $M$  – maksimalus dažnio indeksas. Funkcija  $H(z)$  surandama iš lygties:

$$H_i(z) = \frac{G(i)}{a_0 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_N z^{-N}}, \quad (2)$$

čia  $a_j$  – koeficientai, gauti tiesinės prognozės (TP) analizės metu. Dydis  $G(i)$  surandamas iš lygties

$$G(i) = \sqrt{\sum r^2(n)}, \quad (3)$$

čia  $r(n)$  – TP analizės paklaidos.

**Skardžių/duslių/tylos intervalų nustatymas.** Iš pradžių signalas suskaidomas į vienodus 5ms intervalus ir atliekama kiekvieno intervalo tiesinės prognozės (TP) analizė. Šios analizės rezultatas yra liekana, kuri parodo balsaskylės uždarymo taškus, ir atspindžio koeficientai kiekvienam intervalui. Intervalas yra žymimas kaip skardus, jei atspindžio koeficientai yra didesni nei 0,2, ir duslus - jei mažesnis ar lygus 0,2. Balsaskylės uždarymo indeksai toliau naudojami antrą kartą atliekant TP analizę. Antrosios analizės metu duslaus garso ir tylos zonos atliekama asinchroninė analizė, kai nagrinėjama zona yra

suskaidoma į vienodo dydžio intervalus. Skardaus garso zonoje pasirenkama sinchroninė analizė, kai nagrinėjama zona yra suskaidoma į tam tikrą skaičių intervalų, todėl skirtingose skardaus garso zonoje intervalų ilgis skirtingas. Antroji TP analizė irgi atliekama kiekvienam intervalui. Duslaus garso zona nustatoma ten, kur balsaskylė atvira, skardaus garso zona – ten, kur balsaskylė keičia būseną. Antroji sinchroninė TP analizė leidžia pažymėti tranzitines sritis. Tranzitinės sritys yra tokios sritys, kai duslaus garso zona keičiasi į skardaus garso zoną ir atvirkščiai. Tranzitinės sritys išskirtos todėl, kad šiose vietose yra didelė tikimybė padaryti klaidų. Taip pažymėtos zonos priskiriamos prie skardžiųjų garsų zonos, nes šiose zonoje, ypač ten, kur skardžiojo garso zona pereina į dusliojo garso zoną, dėl spartaus energijos mažėjimo dažnai klaidingai gali būti priskirta prie duslaus garso zonos. Paskui reikia atskirti tylos ir duslumo zonas. Tam atvejui reikia apskaičiuoti garso funkciją ir lyginti su slenkstiniu triukšmo lygiu. Kai garso funkcija viršija slenkstinių triukšmo lygį ar yra jam lygi, nustatoma dusli zona, kai mažesnė, – nustatoma tylos zona. Po šio apdorojimo jau turime išskyrę tris skirtingas garso zonas: skardaus ir duslaus garso bei tylos.

**Balsingųjų garsų nustatymas.** Intervalas yra laikomas balsingu, kai jis yra skardus ir žemųjų dažnių garsumo funkcija gerokai didesnė už aukštųjų dažnių garsumo funkciją. Balsingųjų garsų grupei priklauso: balsiai, pusbalsiai, nosiniai priebalsiai, sprogstamųjų priebalsių pradžia. Balsingumui nustatyti naudojami du juostiniai filtrai (žemųjų ir aukštųjų dažnių srityse) ir surandami filtrų išėjimo signalų garsumai. Juostinio filtro žemųjų dažnių srityje pralaidumo juosta yra 98 Hz - 898 Hz, o juostinio filtro aukštesniųjų dažnių srityje – 3691 Hz – 5000 Hz [5]. Pirmosios reikšmės santykis su antrąja rodo balsingumo santykį  $R(i)$  (angl. *sonorant ratio*). Čia  $i$  – intervalo indeksas. Balsingumo santykis  $R(i)$  yra lyginamas su slenkstiniu skardumo lygiu  $T_{son}$ . Šitaip nustatomas balsingumo įvertis  $SS(i)$  (angl. *sonorant score*). Jei  $R(i) \geq T_{son}$  tai intervalas yra balsingas ir  $SS(i) = 1$ , priešingu atveju nebalisingas ir  $SS(i) = 0$ .

**Balsių nustatymas.** Balsiai nustatomi panašiai kaip ir balsingumas [5]. Keičiasi tik filtrų pralaidumo juostos. Jos atitinkamai yra 20 Hz – 996 Hz ir 1016 Hz – 5000 Hz. Tokiu pat principu apskaičiuojamas ir balsio santykis  $VWL(i)$  (angl. *vowel ratio*). Balsės įvertis  $VWLS(i)$  (angl. *vowel score*) yra apskaičiuojamas pagal formulę.

$$VWLS(i) = \left\{ \begin{array}{ll} 0, & \text{jei } VWL(i) \geq T_v \\ 1, & \text{jei } VWL(i) \leq T_a \\ \frac{T_v - VWL(i)}{T_v - T_a}, & \text{jei } T_a \leq VWL(i) \leq T_v \end{array} \right\}. \quad (4)$$

Dalies klaidų galima išvengti atlikus korekciją – intervalus laikyti neturinčiais balsių, kai intervalų trukmė yra trumpesnė nei 15 ms.

**Skardžiųjų priebalsių nustatymas.** Šiuo atveju naudojami juostiniai filtrai su ta pačia pralaidumo juosta, kaip ir nustatant balsius. Naudojami ir identiški slenkstiniai lygiai. Skardžiųjų priebalsių buvimo lygį  $VC(i)$  (angl.

*voiced consonant ratio*) galima nustatyti minėtų garsumo funkcijų santykio metodu. Skardžiųjų priebalsių buvimo zonos nustatomos priešingai negu balsio buvimo zonos, t.y. skardusis priebalsis yra ten, kur  $VC(i) \geq T_v$ , tada skardžiųjų priebalsių įvertis  $VCS(i)$  (angl. *voiced consonant score*) lygus vienetui, o kur  $VC(i) < T_v$ , nustatoma, kad skardžiojo priebalsio nėra ir  $VCS(i) = 0$ . Jeigu skardžiojo priebalsio buvimo lygis yra tarp abiejų slenkstinių lygių, tada tik su tam tikra tikimybe galima sakyti, kad tai yra skardusis priebalsis, ( $VCS(i)$  yra tarp 0 ir 1). Šiuo atveju reikia žiūrėti, prie kurio slenkstinio lygio arčiau yra skardžiojo priebalsio įvertis.

**Pusbalsių nustatymas.** Pusbalsių nustatymo algoritmas pagrįstas Espy ir Wilson metodu [6]. Pusbalsiams nustatyti naudojami du juostiniai filtrai. Pirmasis filtras – žemesniųjų dažnių srityje (20 Hz – 391 Hz), antrasis filtras – aukštesniųjų dažnių srityje (410 Hz – 977 Hz). Šiuo atveju iš pradžių nustatomas šniokštimo lygis  $MUR(i)$  (angl. *murmur ratio*), dalijant garsumo funkcijas vieną iš kitos. Nustatytas šniokštimo lygmuo lyginamas su viršutiniu ( $T_v$ ) ir apatiniu ( $T_a$ ) slenkščiu, kurie nustatomi eksperimentiškai. Jei  $MUR(i) \geq T_v$ , tai tame segmente yra šniokštimas, o šniokštimo įvertis  $MS(i)$  (angl. *murmur score*) lygus vienetui. Jei  $MUR(i) < T_a$ , tai tame segmente nėra šniokštimo, o  $MS(i) = 0$ . Kai  $MUR(i)$  yra tarp viršutinio ir apatinio slenkščio, tokiu atveju sunku tiksliai nustatyti, ar tame intervale yra šniokštimas, ar ne,  $MS(i)$  nustatomas tarp 0 ir 1. Toliau iš formulės galima nustatyti pusbalsio buvimo įvertį  $SVS(i)$  (angl. *semivowel score*):

$$SVS(i) = (1 - MS(i)) \cdot (1 - VBS(i)) \cdot VCS(i). \quad (5)$$

Intervaluose, kuriuose  $SVS(i) = 1$ , nustatomas pusbalsio buvimo, kuriuose  $SVS(i) = 0$ , nustatomas pusbalsio nebuvimas. Jeigu intervale  $SVS(i)$  yra tarp 0 ir 1, tada vienareikšmiškai sunku nustatyti pusbalsio buvimą. Patikslinant pusbalsio nustatymą naudojamos papildomos taisyklės [7].

Informacija apie lietuvių kalbos fonemų diferencinius požymius buvo paimta iš A. Girdenio [8] ir A. Pakerio [9].

## Rezultatai

Buvo įrašyta vieno kalbėtojo vyro 30 žodžių tarimo garso ir vaizdo informacija. Segmentuojant, buvo išskirta 150 fonemų ir jas atitinkančių vizemų.

Igyvendinus 2 dalyje aprašytus algoritmus (nepasiteisinus aprašytam skardžiųjų bei duslių ir tylos intervalų nustatymo algoritmui, informacija paimta iš Praat programos), fonemos buvo suskirstytos į šias grupes:

- balsiai ir dvibalsiai;
- pusbalsiai;
- skardieji sprogstamieji priebalsiai;
- skardieji friktyviniai priebalsiai;
- duslieji priebalsiai.

Iš rezultatų aiškėja, kad geriausiai iš balsių ir dvibalsių į savo grupę pateko balsiai „a“, „e“, „i“ ir dvibalsiai, į kuriuos įeina nors vienas iš šių balsių. Daugiausia kaip balsis klaidingai buvo suvokiamas pusbalsis „r“. Į balsius panašių savybių turi ir pusbalsis „l“. Tuo tarpu prie pusbalsių daugiausia buvo klaidingai priskirti balsiai „u“, „o“, „uo“. Skardžiuosius

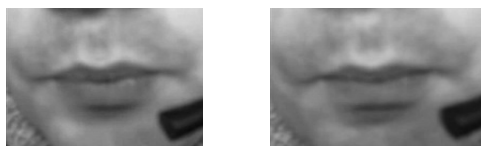
sprogstamuosius priebalsius „b“, „d“, „g“ daugeliu atveju leidžia teisingai klasifikuoti tik trumpesnė jų trukmė. Tiksliai kaip duslieji priebalsiai yra klasifikuojami garsai „s“, „š“.

Vizemų išskyrimas ir analizė buvo atlikti rankiniu būdu. 1 paveiksle pavaizduota balsių fonemų „o“ ir „a“ vizemos žodyje forma.



1 pav. Vizemos „a“ ir „o“ žodyje „forma“

Kaip matome iš 1 paveikslo, yra didelis vizualus skirtumas tarp šių vizemų, atstovaujančių skirtingoms fonemoms. Priešingai nei 1 paveiksle, tarp 2 paveiksle pavaizduotų vizemų, atstovaujančių skirtingoms fonemoms akivaizdus skirtumo nėra. Šiuo atveju taip yra todėl, kad abu garsai yra sprogstamieji ir lūpų padėtis yra visiškai vienoda.



2 pav. Vizema „b“ žodyje „bėga“ ir „p“ žodyje „Europa“

Taip pat galima pastebėti, kad tos pačios fonemos vizemos, esančios skirtingose pozicijose (padėtis atsižvelgiant į prieš ir po einančias fonemas), gali gerokai skirtis. Tokį konkretų pavyzdį galima pamatyti 3 paveiksle. Taip yra todėl, kad tos pačios fonemos vizema priklauso nuo toliau esančių fonemų.



3 pav. Vizemos „b“ žodyje „bėga“ ir „boikotas“

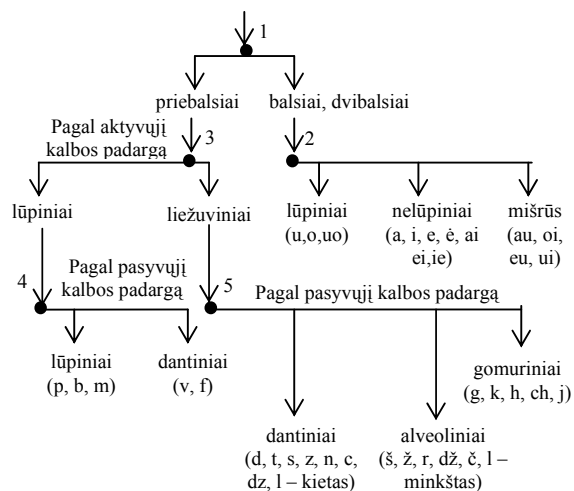
## Diskusija

Remiantis garsų artikuliacijos informacija ir preliminariais vizemų tyrimo rezultatais, galima pasiūlyti toki lietuvių kalbos vizemų grupavimą, koks pateiktas 1 lentelėje.

Remiantis 1 lentele, buvo nubraižytas vizemų medis (4 pav.)

1 lentelė. 8 vizemų grupės lietuvių kalboje

8 vizemų grupės	
priebalsiai	
Abilūpiniai	p, b, m
Dantiniai	d, t, s, z, n, dz, l–kietas
Gomuriniai	g, k, h, ch, j
Lūpiniai dantiniai	v, f
Alveoliniai	š, č, ž, r, dž, l–minkštas
balsiai	
Lūpiniai balsiai, dvibalsiai	o, u, uo
Nelūpiniai balsiai, dvibalsiai	a, i, e, é, ai, ei, ie
Mišrieji dvibalsiai	au, ui, eu, oi



4 pav. Vizemų klasifikacijos medis

1. Balsius ir priebalsius galima atskirti pagal balsaskylės atvirumą: tariant balsius burna būna plačiai atverta. Taip pat balsius nuo priebalsių galima atskirti pagal ilgumą, balsiai gerokai ilgesni už priebalsius. Ilgumą galima nustatyti iš garsinės informacijos.

2. Lūpinius balsius ir dvibalsius nuo nelūpinių galima atskirti pagal lūpų formą. Lūpinių balsių ir dvibalsių tarimo metu lūpos būna apvalios. „u“ nuo „o“ būtų galima atskirti pagal lūpų atvirumą, tariant „o“ lūpos būna labiau atvertos. Dvibalsį „uo“ galima atpažinti, kai iš pradžių užfiksuojamas balsis „u“, paskui burna yra plačiai atveriamą. Tariant nelūpinius balsius, burna pravira, bet lūpos neapvalios. Atskirai lengviau atskirti balsį „i“, nes tariant šį balsį burna yra mažiau atveriamą negu tariant kitus nelūpinius balsius. Jei burna būna atvira, kaip tariant „a“ ar „e“, o vėliau plyšys labai susiaurėja, galima teigti, kad tai dvibalsis „ai“ arba „ei“. Jei burna pražiota vienodai, ir iš pradžių matoma įtampa (lūpos ištemtos į šonus), tai gali būti dvibalsis „ie“. Dvibalsius nuo balsių galima atskirti pagal ilgumą, dvibalsiai yra ilgesni už balsius. Aptarsime mišriųjų dvibalsių atskyrimą. Jei iš pradžių burna būna atvira kaip tariant balsį „a“, o vėliau apvalėja, tada toje vietoje yra dvibalsis „au“ arba „eu“. Jei burna iš pradžių būna apvali, o vėliau išsitiesia ir prisičiaupia, tada tai bus dvibalsis „oi“ arba „ui“.

3. Pagal aktyvųjį kalbos padargą lūpinius priebalsius nuo liežuvinų priebalsių galima atskirti iš lūpų padėties. Tariant lūpinius priebalsius pradžioje lūpos būna sučiauptos ir įtemtos. Tuo tarpu tariant liežuvinus priebalsius, lūpos būna truputi pravertos, atpalaiduotos (beveik nejuda) ir be jokių įtampos požymių.

4. Pagal pasyvųjį kalbos padargą lūpiniai (pagal aktyvųjį kalbos padargą) priebalsiai yra skirstomi į lūpinius ir dantinius. Tokiu atveju pirmieji ir pagal aktyvųjį ir pagal pasyvųjį kalbos padargą yra lūpiniai, o antrieji yra lūpiniai ir dantiniai. Lūpinius nuo dantinių galima atskirti pagal tai, kad tariant lūpinius priebalsius lūpos yra stipriau sučiaupiamos, labiau įtemtos, nei tariant dantinius. Taip pat tariant lūpinius lūpos yra apvaulesnes nei tariant dantinius.

5. Pagal pasyvųjų kalbos padargą liežuviniai (pagal aktyvųjų kalbos padargą) priebalsiai yra skirstomi į dantinius, alveolinius ir gomurinius. Šiuo atveju, panaudojant vaizdinę informaciją, sunku atskirti šias priebalsių grupes, nes tarimo skirtumus lemia liežuvio padėtis (artikuliacija) burnoje: jis remiasi į priekinius dantis arba savo viduriu remiasi į gomurį, arba liežuvio galas remiasi gomurį. Kadangi liežuvio padėties kamera užfiksuoti negalima, tai, naudojant vizemas, beveik neįmanoma atskirti šių priebalsių tarpusavyje.

Pasinaudojant tokiu vizemų klasifikavimu ir garso signalo informacija, galima sudaryti sprendimų medį, kuris geriau klasifikuotų fonemas. Pagal tyrimo [3] duomenis blogai atpažįstama „m“ fonema (atpažinimo tikslumas 5%), nors 90% ji atpažįstama kaip pusbalsis. Tarp pusbalsių jis yra vienintelis, priklausantis abilūpinių vizemų grupei. Fonemos „b“ atpažinimo tikslumas 15,3%. Kaip sprogtamasis priebalsis jis atpažįstamas 52,5% tikslumu ir taip pat yra vienintelis priklausantis abilūpinių vizemų grupei. Mūsų tyrime „r“ pagal garsinę informaciją dažnai patenka prie balsių. Pagal burnos atvirumą ši fonema lengvai atskiriama nuo balsių.

### Išvados

Vien garsinio signalo šnekai atpažinti neužtenka, nes triukšmingoje aplinkoje signalas gali būti negrįžtamai iškraipytas. Todėl šnekos atpažinimą, reikėtų kombinuoti remiantis tiek garsine, tiek vaizdine informacija. Kaip parodė atlikta analizė, tada kai kurie garsai atpažįstami kur kas geriau nei vien pagal garsinę informaciją. Pasiūlyta vizemų klasifikavimo sistema. Vizemų atpažinimo automatizavimo darbus planuojame tęsti.

### Literatūros sąrašas

1. **Weinstein C. J., McCandless S.S., Mondschein L.F., Zue V.W.** A system for acoustic-phonetic analysis of continuous speech // IEEE Transactions on Acoustics, Speech, and Signal Processing. – 1975. – V. 23, No. 1. – P. 54 – 67.
2. **Shipman D., Zue V.W.** Properties of large lexicons: Implications for advanced isolated word recognition systems // IEEE Proceedings of the International Conference on Acoustics, Speech and Signal Processing. – 1982. – V. 7. – P. 546 – 549.
3. **Driaunys K., Rudžionis V., Žvinys P.** Hierarchinė fonemų struktūra grindžiamo LTDIGIT'S fonemų klasifikavimo tyrimas// Informacinės technologijos. – Kaunas: Technologija, 2005. – P. 283-288.
4. **Chen T., Rao R. R.** Audio-Visual Integration in Multimodal Communication // Proceedings of the IEEE. 1998. – V. 86, No 5. – P. 837 – 852.
5. **Childers D. G.** Speech processing and synthesis toolboxes. John Wiley & Sons, 2000. – 484 p.
6. **Espy Wilson C. Y.** A phonetically based semivowel recognition system // IEEE Proceedings of the International Conference on Acoustics, Speech and Signal Processing. – 1986. – P. 2775 – 2778.
7. **Daunys G., Balbonas D.** Garsų klasifikavimas panaudojant sprendimo medžius // Informacinės technologijos. – Kaunas: Technologija, 2005. – P. 277 – 282.
8. **Girdenis A.** Teoriniai lietuvių fonologijos pagrindai. Vilnius, Mokslo ir enciklopedijų leidybos institutas, 2003. – 388 p.
9. **Pakerys A.** Lietuvių bendrinės kalbos fonetika. Vilnius, Enciklopedija, 2003. – 244 p.

Pateikta spaudai 2005 03 15

**D. Balbonas, G. Daunys. Fonemų klasifikavimas panaudojant garso ir vaizdo informaciją // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2005. – Nr. 5(61). – P. 74 – 77.**

Nagrinėjama šnekos signalo analizė, panaudojant fonemų diferencinius požymius. Analizuojami parametrai, kurie reikalingi klasifikuojant šnekos intervalus į fonemų klases. Garso signalo analizei buvo panaudotas tiesinės prognozės metodas. Jo dėka buvo gautos garsumo funkcijos įvairioms dažnių juostoms. Aptikta, kad pusbalsiai „r“ ir „l“ pagal parametrus yra artimi balsiams „u“, „o“. Dėl to šnekos atpažinimui triukšmingoje aplinkoje vien garso informacijos neužtenka. Nagrinėjama galimybė panaudoti vaizdinę informaciją šnekos atpažinimui, sudarant vizemų sprendimo medžius ir naudojant juos kartu su fonemų diferenciniais požymiais. Pasiūlyta vizemų klasifikavimo sistema. Diskutuojama apie galimybę sukurti metodiką lietuvių kalbos šnekos atpažinimui, panaudojant tiek garsinę, tiek vaizdinę informaciją. Il. 3, bibl. 9 (lietuvių kalba; santraukos lietuvių, anglų ir rusų k.).

**D. Balbonas, G. Daunys. Classification of Phonemes using Sound and Visual Information // Electronics and Electrical Engineering. – Kaunas: Technologija, 2005. – No. 5(61). – P. 74–77.**

The paper concerns automatic speech recognition using differential features of phonemes. The methods to obtain parameters for phonemes classification are discussed. The analysis of application visual information for speech recognition is done. The system for classification Lithuanian language visemes was proposed. The possibilities about combining sound and visual information for speech recognition is discussed. Il. 3, bibl. 9 (in Lithuanian; summaries in Lithuanian, English and Russian).

**Д. Балбонас, Г. Даунис. Использование звуковой и визуальной информацией для классификации фонем // Электроника и электротехника. – Каунас: Технология, 2005. – № 5(61). – С. 74–77.**

Анализируются методы для получения нужных параметров для классификации речевых интервалов в классы фонем. Дискутируется возможность для опознания речи применить дифференциальные свойства фонем. Предложена схема для классификации визем литовского языка. Дискутируется о возможностях создать метод опознания литовского языка используя звуковую и визуальную информацию. Ил. 3, библи. 9 (на литовском языке; рефераты на литовском, английском и русском яз.).