

6. Гутников В. С. Интегральная электроника в измерительных устройствах. Л., 1988.
 7. Хоровиц П., Хилл У. Искусство схемотехники. М., 1986. Т. 1—2.

SUMMARY

The article describes analysis amplitude-frequency characteristic and other performances and parameters of passive and active RC filter with the Micro-Cap software.

УДК 530.1

Дз. Дз. Ладзееў

СУЧАСНЫЯ МЕТАДЫ ІДЭНТЫФІКАЦЫІ ЧАЛАВЕКА ПА ГОЛАСЕ

У наш час значная ўвага надаецца распрацоўцы надзейных сістэм ідэнтыфікацыі асобы на аснове біямэтрычных прыкмет. У аснове такога роду падыходаў ляжыць ідэя распазнавання людзей па іх псіхалагічных ці паводзінскіх характарыстыках. Існуюць розныя тэхналогіі біямэтрыі: распазнаванне твараў і адбіткаў пальцаў, аналіз геаметрыі пальцаў, аналіз радужнай абалонкі вока, аналіз рысунка крывяносных сасудаў, аналіз подпісу, распазнаванне голасу. Значная цікавасць да метадаў на аснове біямэтрыі абумоўлена шэрагам пераваг, якія яны ўяўляюць у параўнанні са звычайнымі

спосабамі ідэнтыфікацыі [1].

Сярод біямэтрычных тэхналогій ідэнтыфікацыі сёння вялікая ўвага надаецца распрацоўцы сістэм на аснове аналізу моўных сігналаў чалавека. Моўны сігнал — складаная структура, у якой адлюстроўваюцца фактары рознай прыроды — семантычнай, лінгвістычнай, артыкуляцыйнай, акустычнай [2—3].

Існуе некалькі асноўных напрамкаў апрацоўкі моўных сігналаў, класіфікацыя якіх прадстаўлена на рыс. 1.

Інфармацыяй высокага ўзроўню ў моўным сігнале з’яўляюцца дыялект, акцэнт, манера размаўляць. Дадзеныя асаблівасці распазнаюцца і аналізуюцца чалавекам. Нізкаўзроўневая інфармацыя ўключае рытм, тон, частату, спектр, паласу прапускання, гэта значыць акустычныя характарыстыкі моўнага сігнала. Гэтыя асаблівасці выкарыстоўваюцца ў аўтаматычных сістэмах распазнавання.

Сучасны падыход да аўтаматычнага распазнавання дыктараў заключаецца ў пабудове стахастычнай мадэлі дыктара з выкарыстаннем параметраў, якія атрыманы з навучальнай моўнай паслядоўнасці. Задача ідэнтыфікацыі дыктара ў тым, што неабходна вызначыць, ці адпавядае голас гаворачага голасу пэўнага дыктара з базы N дыктараў. У агульным

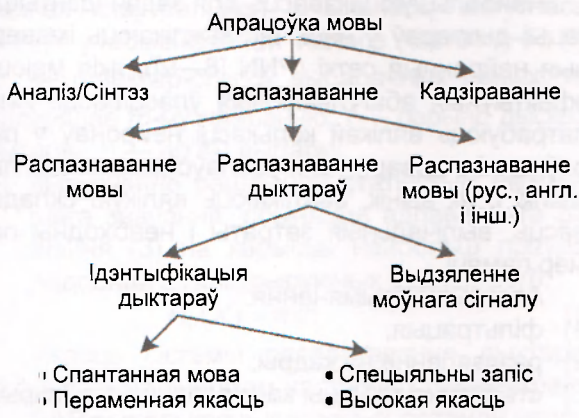
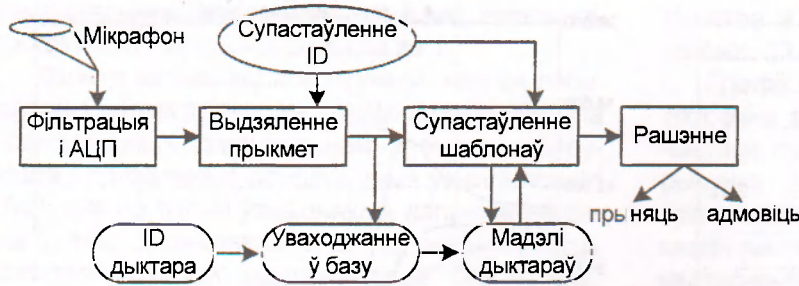


Рис. 1. Асноўныя задачы апрацоўкі мовы



Рыс. 2. Агульная схема сістэмы ідэнтыфікацыі дыктараў

выглядзе задача ідэнтыфікацыі складаецца з пяці этапаў:

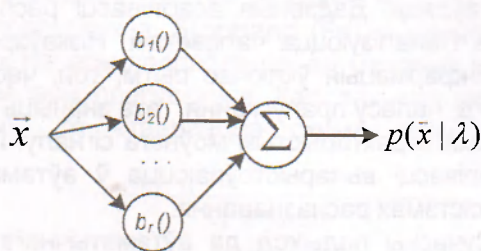
- 1) атрыманне моўнага сігнала ў лічбавым выглядзе;
- 2) выдзяленне прыкмет;
- 3) супастаўленне з эталонам;
- 4) прыняцце рашэння;
- 5) абнаўленне мадэлі дыктара.

Выдзяленне прыкмет вызначае кожны ўчастак моўнага сігнала ў пэўнае месца прасторы. Найбольш часта з гэтай мэтай выкарыстоўваюцца мелкепэстральныя каэфіцыенты (MFCC). Акрамя таго, папулярнымі прыкметамі з'яўляюцца становішча і траекторыі фармант, частоты асноўнага тону і інш. Дадзеная паслядоўнасць вектараў прыкмет x параўноўваецца з мадэллю дыктара. Вынікам параўнання з'яўляецца мера z для ўсёй сукупнасці вектараў. Па мерах і пры дапамозе метадаў класіфікацыі прымаецца рашэнне аб адпаведнасці ці неадпаведнасці дыктара эталону; метады дынамічнага змянення маштабу часу (DTW, шаблонныя мадэлі), маркаўскія мадэлі (HMM [4], GMM статычныя мадэлі), метады вектарнага квантавання (VQ [5—6], табліцы кадзіравання).

Асноўны метады класіфікацыі ў задачах распазнавання дыктараў GMM [7]. Гаўсаўская сумесь вектараў прыкмет \bar{x} — гэта ўзважаная сума M размеркаванняў:

$$p(\bar{x} | \lambda) = \sum_{i=1}^M p_i b_i(\bar{x}), \quad (1)$$

дзе \bar{x} — выпадковы D -мерны вектар.



Рыс. 3. Прынцып работы GMM

Кожнае кампанентнае размеркаванне можна прывесці да выгляду:

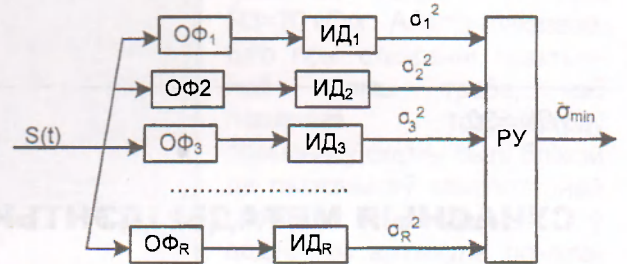
$$b_i(x) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\bar{x} - \bar{\mu}) \Sigma_i^{-1} (\bar{x} - \bar{\mu})\right\} \quad (2)$$

з матэматычным чаканнем $\bar{\mu}$ і каварыяцыйнай матрыцай Σ_i .

Вагавыя каэфіцыенты задавальняюць умове $\sum_{i=1}^M p_i = 1$.

У GMM ёсць дзве ўласцівасці, якія дазваляюць паспяхова выкарыстоўваць іх для задач распазнавання дыктараў. Першая звязана з ідэяй, што індывідуальныя

кампаненты-размеркаванні агульнага размеркавання могуць мадэліраваць пэўны набор акустычных класаў. Апошнія адлюстроўваюць агульныя асаблівасці галасавога тракту і могуць быць выкарыстаны для ідэнтыфікацыі гаворачага. Другая звязана з эмпірычным назіраннем, што лінейная камбінацыя базісных гаўсаўскіх функцый можа вылучыць вялікую колькасць розных размеркаванняў. GMM здольна фарміраваць гладкае прыбліжэнне адвольнага размеркавання (рыс. 4).



Рыс. 4. Схема выкарыстання нейронных сетак для ідэнтыфікацыі дыктараў

Для стварэння GMM неабходна для дадзенага дыктара і яго моўнай паслядоўнасці падбраць падыходзячыя GMM каэфіцыенты, гэта значыць — параметры GMM. Папулярным з'яўляецца метады найбольшай праўдападобнасці.

Наступным метадам для распазнавання дыктараў з'яўляюцца нейронныя сеткі. Пры выкарыстанні нейронных сетак звычайна прымяняецца двухэтапная апрацоўка. Першы этап — нейронная сетка для вызначэння групы дыктараў: мужчына, жанчына, бас, тэнар, сапрана. Другі крок дазваляе ідэнтыфікаваць дыктара ў групе.

Найбольшую цікавасць для задач ідэнтыфікацыі дыктараў у наш час выклікаюць імаверныя нейронныя сеткі (PNN [8—9]), якія маюць эфектыўныя абагульняючыя ўласцівасці. Яны патрабуюць вялікай колькасці нейронаў у параўнанні з адваротным распаўсюджваннем памылкі і, як вынік, выклікаюць вялікую складанасць, вылічальныя затраты і неабходны памер памяці.

Алгарытм прымянення:

- 1) фільтрацыя;
- 2) раздзяленне на кадры;
- 3) стварэнне табліцы кадзіравання з выкарыстаннем квантылізацыі вектараў (VQ);

- 4) першы слой апрацоўкі $a, 1 = \exp(-(| | / W_{11} - p | | b, 1)^2)$;
- 5) другі слой апрацоўкі $A = \text{complet}(LW_2, a1)$;
- 6) прыняцце рашэння.

Аб'яляючы фільтр — адзін з метадаў распазнавання дыктараў [10—11]. Спосабы карэляцыйнага ці спектральнага аналізу характарызуюцца недастатковай эфектыўнасцю з-за лішку даных у моўных сігналах, а таксама павышанай складанасцю алгарытмаў. Адным з радыкальных сродкаў да выкарыстання лішку можа служыць ідэя сціскання даных за кошт эканомнага кадзіравання. Найбольш перспектыўным кірункам у дадзеных даследаваннях з'яўляецца метадаў распазнавання моўных сігналаў, які заснаваны на тэарэтыка-інфармацыйным падыходзе і універсальнай метрыцы Кульбака-Лейблера. Гэты алгарытм атрымаў назву метадаў аб'яляючага фільтра. Яго прынцып рэалізуецца па схеме, адлюстраванай на рыс. 4, дзе:

ОФ_r — аб'яляючы фільтр, настроены на гэтую ацэнку спектральнай шчыльнасці магутнасці;

ИД_r — вымяральнік дысперсіі, які вызначае дысперсію нескампенсаванага лішку на выхадзе адпаведнага аб'яляючага фільтра;

РУ — рашаючае ўстройства, якое вызначае мінімальную дысперсію.

Для адрознення набору з *R* розных спектральных ацэнак мяркуецца наступная сукупнасць аперацый для назіранняў:

- 1) фарміраванне выбарачнай ацэнкі спектральнай шчыльнасці магутнасці;
- 2) вылічэнне набору рашаючых статыстык інтэгральнага выгляду (3):

$$H_v(X) = \frac{1}{2F} \int_{-F}^F \left[\frac{G_x(f)}{G_r(f)} + \ln G_r(f) \right] df, \quad (3)$$

дзе $G_x(f) = \frac{1}{M} \sum_{m=1}^M \frac{1}{2Fn} \left| \sum_{l=1}^n x_m(l) \exp(-2j\pi l f \tau) \right|^2$ —

выбарачная ацэнка шчыльнасці магутнасці з усярэднянасцю па *M* незалежных вектарных назіраннях; *G_r(f)* — канчатковы набор розных ацэнак спектральнай шчыльнасці магутнасці $r=1, R$ у абмежаванай паласе частот $\{-F, F\}$; \bar{x}_m — *n*-вектар-слупок выбарачных даных, які адпавядае *m*-му назіранню; τ — перыяд дыскрэтызацыі; *n* — аб'ём выбаркі;

- 3) вызначэнне рашаючай статыстыкі найменшага значэння, прыняцце адпаведнага рашэння (3) на карысць найбольш праўдападобнай з канкуруючых ацэнак *G_v(f)*:

$$H_v(X) = \min. \quad (4)$$

Якасць сістэмы распазнавання дыктараў ацэньваецца аналізам наступных параметраў:

FAR — колькасць дадзеных сістэмай няправільных станоўчых адказаў;

FRR — колькасць дадзеных сістэмай няправільных адмоўных адказаў;

EER — колькасць памылак, калі сістэма настроена такім чынам, што **FAR** = **FRR**.

Параўнальная характарыстыка магчымасцей некаторых мадэлей у задачах распазнавання дыктараў змешчана ў табліцы. Ацэнка праводзілася на базе YOHO [12], утрымлівае 108 мужчынскіх і 30 жаночых галасоў.

Табліца

Параўнанне метадаў распазнавання дыктараў

Мадэль дыктара	FAR (%)	FRR (%)	EER (%)
Гаўсаўскія сумесі	4,91	1,68	1,19
Эліптычныя базісныя функцыі	33,51	4,93	2,73
Імаверныя нейронныя сеткі	0,35	16,17	1,19

Пабудова рабасных сістэм распазнавання дыктараў ажыццяўляецца з выкарыстаннем больш надзейных сістэм параметраў і мер падабенства, перад апрацоўкай моўнага сігнала на ўваходзе сістэмы распазнавання або мадыфікацыяй (адаптацыяй) параметраў акустычных мадэлей гукаў.

Напрыклад, агульнапрынятым прыкладам павышэння дакладнасці распазнавання ў прысутнасці перашкод з'яўляецца дапаўненне сістэм кароткачасовых параметраў іх першымі і другімі вытворнымі па часе.

Папярэдняя апрацоўка моўнага сігнала ўключае метады яго нармалізацыі: адніманне кепсцральнага ці спектральнага сярэдняга, фільтрацыя часавых траекторый параметраў (RASTA-апрацоўка), метады нейронных сетак, а таксама калманаяўская фільтрацыя і розныя мадыфікацыі спектральнага аднімання для зніжэння ўплыву квазістацыянарных адытыўных перашкод [13—15].

Такім чынам, на сённяшні дзень існуе даволі вялікая колькасць метадаў і падыходаў да рашэння задач ідэнтыфікацыі дыктараў па голасе. Большасць з іх можна ўмоўна падзяліць на статычныя і шаблонныя. Найбольш папулярнымі з'яўляюцца нейронныя сеткі, скрытыя маркаўскія мадэлі, гаўсаўскія сумесі, метадаў квантылізацыі вектараў, метадаў аб'яляючага фільтра і інш.

У наш час тэндэнцыяй развіцця з'яўляецца стварэнне сістэм на аснове сукупнасцей некалькіх метадаў для атрымання больш надзейных вынікаў пры функцыянараванні ў рэальных умовах.

ЛІТАРАТУРА

- 1. Joseph P., Campbell Jr. Speaker Recognition: A Tutorial. <http://web.cse.msu.edu/~cse891/Sect601/textbook/8.pdf>
- 2. Gerik v.Graevenitz About Speaker Recognition

- Technology. http://www.speech.kth.se/~rolf/NGSLT/gsit_papers_2004/barisevicius_term_paper.pdf
3. Automatic recognition of speakers from their voices // Proc. IEEE. 1976. Vol. 64. P. 460—475.
 4. Koolwaj J. Fundamentals of HMM Based Speaker Verification. <http://www.ispeak.nl/start.html>
 5. A vector quantization approach to speaker recognition // AT&T Tech. J. 1987. V. 66. № 2. P. 14—26.
 6. Gray R. M. Vector Quantization // IEEE Acoustics, Speech and Signal Processing Magazine. 1984. № 4. P. 4—29.
 7. Reynolds D. A. and Rose R. C. Robust text-independent speaker identification using Gaussian mixture speaker models // IEEE Trans. on Speech and Audio Processing. 1995. № 3. P. 72—83.
 8. Claude A. Norton. Text Independent Speaker Verification Using Binary-Pair Partitioned Neural Networks // Old Dominion, 1995.
 9. Hartigan J. A. and Wong M. A. A k-means clustering algorithm // Applied Statistics. 1979. № 28. P. 100—108.
 10. Акатьев Д. Ю. Бочаров И. В. Распознавание дикторов по методу обеляющего фильтра. <http://zhumal.ape.relarn.ru/articles/2001/147.pdf>
 11. Atal B. S. Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification // J. Acoust. Soc. Amer. 1974. V. 55. № 6. P. 1304—1312.
 12. Campbell Jr. J. P. Testing with the YOHO CD-ROM voice verification corpus // ICASSP'95. 1995. P. 341—344.
 13. Basseville M. Distance measures for signal processing and pattern recognition // Signal Process. 1989. V. 18. P. 349—369.
 14. Марпл С. Л. Цифровой спектральный анализ и его приложения. М., 1990.
 15. Винцюк Т. К. Анализ, распознавание и смысловая интерпретация речевых сигналов. Киев, 1987.

SUMMARY

Automatic speaker recognition by means of modern statistical methodology and algorithms is presented. Considered Gaussian Mixture Models, Neural Networks, Whitewashing Filtering.

УДК 530.1

У. В. Станкевіч

АДЗІНАЯ КЛАСІФІКАЦЫЯ ДЭМАНСТРАЦЫЙНЫХ РЭСУРСАЎ ПА ФІЗІЦЫ

Дэманстрацыйны эксперымент з'яўляецца неад'емнай часткай працэсу выкладання. Вядома некалькі падыходаў да пабудовы класіфікацыі дэманстрацыйных эксперыментаў: па функцыянальным прызначэнні, ролі і месцы ў вучэбным працэсе, тыпе навучальнай установы і абсталяванні, якое прымяняецца на занятках, і г. д. Мэта артыкула — прапанаваць і ўкараніць распрацаваную намі адзіную класіфікацыю дэманстрацыйных рэсурсаў.

Класіфікацыя, якая найбольш выкарыстоўваецца, — па тыпе навучальнай установы, дзе дэманстрацыйныя эксперыменты прынята падзяляць на дзве групы: для сярэдняй (школа, ПТУ, тэхнікум) і вышэйшай школы. У метадычнай літаратуры замацавалася адпаведная назва: «школьны дэманстрацыйны эксперымент» і «лекцыйныя дэманстрацыі ў ВНУ». Між тым, адрозненне дэманстрацыйнага эксперыменту ў ВНУ ад школьнага не з'яўляецца прынцыповым, і ўсе адрозненні носяць хутчэй фармальны, чым змястоўны характар. Параўнанне шэрага дапаможнікаў па тэхніцы і методыцы дэманстрацыйнага эксперыменту ў школе і ВНУ паказвае, што змест іх пераклікаецца. Да таго ж, практыка сцвярджае, што шматлікія з дэман-

страцыйных эксперыментаў, якія праводзяцца ў школе, дэманструюцца ў ВНУ і наадварот.

На самай справе, ці можам мы назваць эксперыменты «Удар шароў», «Броўнаўскі рух» або «Дыфракцыя на шчыліне» «школьнымі» або «вузаўскімі», калі дэманстрацыя кожнага з іх прадугледжана як школьнай, так і праграмамі ВНУ, а дадзеныя доследы паспяхова ўспрымаюцца як навучэнцамі школ, так і студэнтамі ВНУ?

Такім чынам, прынцыповых адрозненняў паміж дэманстрацыйнымі эксперыментамі ў сярэдняй і вышэйшай школе, якія прадвызначалі б іх раздзяленне і адасабленне, няма. Ва ўсіх навучальных установах да іх прад'яўляюцца падобныя патрабаванні: нагляднасць прадстаўлення рэзультатаў дэманстрацыі, добрая бачнасць, кароткачасовасць, эстэтычнасць і інш.

Такім чынам, вядомы сёння парк дэманстрацыйных эксперыментаў мэтазгодна разглядаць як адзіную сістэму, улічваючы спецыфіку прымянення на розных ступенях фізічнай адукацыі. Асабліва неабходны такі падыход пры падрыхтоўцы настаўніка фізікі і можа быць узяты як метадалагічная аснова **выхаван-**