

1.1.6. Математикаи ҳисоббарор  
1.1.6. Вычислительная математика  
1.1.6. The accounting mathematics

УДК – 519.2  
ББК – 30В6

**МОДЕЛИ МАТЕМАТИКИИ  
КОРКАРДИ ОВОЗ ВА ХОСИЯТҲОИ  
ИТТИЛООТИКУНОИ**

*Дадобоев Акмал Анварҷонович* - докторанти (PhD)  
кафедраи технологияҳои иттилоотию  
коммуникатсионӣ ва барномарезии ДДХБСТ  
(Тоҷикистон, Хучанд), e-mail:  
akmal\_dadoboev@mail.ru

**МАТЕМАТИЧЕСКАЯ МОДЕЛЬ  
РЕЧЕВОЙ ОБРАБОТКИ И  
КОММУНИКАТИВНЫХ СВОЙСТВ**

*Дадобаев Акмал Анварҷонович* - докторант (PhD)  
кафедры информационно-коммуникационных  
технологий и программирования ТГУПБП  
(Таджикистан, Худжанд), e-mail:  
akmal\_dadoboev@mail.ru

**MATHEMATICAL MODEL OF  
SPEECH PROCESSING AND  
COMMUNICATION PROPERTIES**

*Dadobaev Akmal Anvarjonovich* - Doctoral Student  
(PhD) of the Department of Information and  
Communication Technologies and Programming of the  
TSULBP (Tajikistan, Khujand), e-mail:  
akmal\_dadoboev@mail.ru

**Ключевые слова:** математическая модель, распознавание голоса, алгоритм MFCC, дискретное преобразование Фурье, гравитационные функции, мел-частота.

В данной статье представлены проблемы анализа голосовых данных, проблема распознавания голосовых данных, а также часто используемый для современных систем распознавания голоса алгоритм MFCC (Mel-Frequency Cepstral Coefficients).

**Вожаҳои калидӣ:** модели математикӣ, шинохтагирии овоз, алгоритми MFCC, табдилдиҳии дискретии Фурье, функсияҳои вазнинӣ, мел-басомад.

Дар мақолаи мазкур масъалаҳои моделиронии математикӣ оид ба таҳлили маълумотҳои овозӣ, масъалаи шинохтагирии маълумотҳои овозӣ ва инчунин дар бораи алгоритмаи коркарди компютерӣ дида шуда, барои системаҳои муосири шинохтагирифтани овози MFCC (Mel-Frequency Cepstral Coefficients-коэффициентҳои кепстралии мел-басомад) татбиқ карда мешавад.

**Keyword:** mathematical model, voice recognition, MFCC algorithm, discrete Fourier transform, gravity functions, mel-frequency.

In this article, the problems of voice data analysis, the problem of voice data recognition, as well as the algorithm that is often used for modern voice recognition systems, MFCC (Mel-Frequency Cepstral Coefficients) are presented.

Инсон дар даҳсолаҳои охир барои сохтани системаҳои таҳлил ва коркарди нутқ ҳаракат намуда, мақсади асосӣ ба омӯзиши сохт ва ҷабҳияти ҳолатҳои нутқу авзои шунавоии инсон, моделронии онҳо бо ёрии барномаҳои технологияи муосири компютерӣ ва дастгоҳҳои махсуси электронӣ шуруъ намудааст.

Пешрафти илм, техника ва технология бо тараққиёти воситаҳои муносибати одамон ва мошинҳои электронӣ зич алоқаманд буда, бо томиат гуфтан мумкин аст, ки маъмултараққиёти воситаи

байни инсон-мошин тавассути интерфейси маълумотӣ, яъне дар намуди маълумоти матнӣ таҳия мешаванд. Чунин усули ташкили муносибат ба имкониятҳои компютер асос карда шуда, иҷрои амалияи он дар компютер хело содда ва осон мебошад. Барои инсон, ки қобилияти қабули маълумоти овозӣ дорад, соддатарин воситаи муносибат – муносибати овозӣ ба шумор меравад. Коркарди системаҳои компютери таҳлили маълумоти овозӣ яке аз масъалаи актуалии технологияҳои телекоммуникатсионии муосир ба ҳисоб меравад. [1]

Масъалаи таҳлили маълумоти овозӣ дар 10 соли охир, яке аз масъалаи марказии рушди соҳаи технологияҳои иттилоотӣ маҳсуб меёбад. Якумин барномаҳои шинохтагрии маълумоти овозӣ ҳанӯз солҳои 90-уми асри гузашта пайдо шуда, имрӯз амалияи васеъро доранд ва дар замони муосир аз ҷумла, роботҳои компютерӣ бо қобилияти шинохтагрии маълумоти овозӣ хело маъмул ҳисобида мешаванд. [2]

Масъалаи шинохтагрии маълумоти овозӣ ба забони табиӣ алоқаманд аст. Имрӯз, ин масъала барои забони англисӣ зиёдтар, барои забонҳои дигар, дар мисоли, забонҳои русию қазоқӣ, хитойӣ немисӣ, ҳалли худро ёфтаанд. Мутаасифона, масъалаи шинохтагрии маълумоти овозӣ дар забони тоҷикӣ ва коркарди моделиронии компютерӣ то ҳол пурра ҳалли худро наёфтааст. Кӯшиши бештари муаллифон дар солҳои охир барои дарёфти таҳқиқотҳо дар ин самт натиҷаи назаррасро ба даст наовардааст. Бинобар ин, масъалаи гузашташуда дар ҳолати пурра ҳаллу фасл шудан, барои барандагони забони тоҷикӣ хеле актуалӣ мебошад [3].

Таҳия ва истифодаи модели математикӣ дар таҳқиқоти системаи компютери шинохтагрии маълумоти овозӣ аз марҳилаҳои муайян иборат буда, дар мақолаи мазкур низ роҳҳои амалигардонии он нишон дода шудааст. Масъалаи зиннаи аввалин дар ин самт, аз маълумотҳои умумии появии мавҷудбуда, кам намудани ҳаҷми онҳо ва дар охир бо назардошти хосиятҳои маълумотҳо равона шудааст, ки имконияти бо тезӣ коркард намуданро фароҳам меоварад. Дар зинаи дувум, миқдори маълумотҳои ҳосилшуда бояд хусусияти доимӣ дошта бошанд. Дар натиҷаи ҳалли чунин масъала, сохтори модели математикии шинохтагрии овоз тағйирнаёбанда мешавад. [2]

Мувофиқи талаботҳои умумӣ баҳодиҳии сифати кори модели математикии коркарди овозӣ бо назардошти барномаи компютерӣ ва ҷудо намудани хосиятҳои маълумотӣ, саҳеҳии системаи шинохтагрии овоз мебошад. Барои қонунгардонидани чунин хосиятҳо оид ба маълумоти сохташуда дар худ доштани қисми зиёди хислатҳои асосии садоҳои овозӣ зарур аст. Бо чунин назардошт аз тарафи олимони соҳа якҷанд моделҳои алгоритмӣ таҳия карда шудааст. Аз ҷумла LDA, MFCC ва ғ. Яке аз моделҳои алгоритмӣ, ки барои системаҳои муосири шинохтагрии овоз бештар истифода бурда мешавад, ин MFCC (Mel-Frequency Cepstral Coefficients- коэффисентҳои кепстралии мел-басомад) мебошад. Мел-басомад воҳиди ченаки психофизикии баландии овоз мебошад. Ин алгоритм аз як ҷанд марҳалаҳо иборат мебошад [4]:

Марҳилаи якум фишурдани пешакӣ маҳсуб ёфта, бо ёрии формулаи математикии зерин иҷро карда мешавад:

$$y_t = x_t - ax_{t-1}, \quad (1)$$

дар ин ҷо  $x_t$  – садои дохилшаванда,  $y_t$  – садои филтронидашуда,  $a$  – коэффисенти филтронӣ буда, чунин ҳудудро дарбар мегирад ( $a \in (0.9; 1)$ ). Одатан, коэффисенти филтронӣ қимати зеринро  $a = 0.97$  дошта, барои баробар намудани спектрӣ садо истифода бурда мешавад.

Марҳилаи дуюм, тақсим намудани садои овозӣ ба қисмҳои андозаашон аз 15 – 30 мс мебошад ва бо қадами доимӣ (одатан 10 мс) баробар аст. Марҳилаҳои навбатӣ барои ҳар як фрейм дар алоҳидагӣ иҷро карда мешаванд [1].

Марҳилаи навбатӣ, табдилдиҳии дискретии Фурье мебошад ва бо ёрии формулаи (2) табдилдиҳии дискретии Фурье иҷро карда мешавад:

$$Y_k = \sum_{t=0}^{T-1} w_t y_t e^{-\frac{2\pi i}{T} kt}, \quad k = 0, \frac{T}{2} \quad (2)$$

дар ин ҷо  $T$  – миқдори ададҳо дар фрейм,  $w_t$  – функсияи вазн.  $y_t$  барои паст кардани таъсирҳои гуногун дар канорҳо, ки дар натиҷаи тақсими садо ба қисмҳои андозаашон охиринок ҳосил мешаванд. Функсияи вазн  $w_t$  – функсияи Хэмминг номида мешавад.

Дар масъалаи шинохтагрии овоз одатан функсияи вазни Хэмминг истифода бурда мешавад, ки намуди зеринро дорад:

$$w_t^{hamm} = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi t}{T-1}\right), \quad t = 0, T-1. \quad (3)$$

Барои ташкили маҷмӯъ аз филтрҳои секунҷавӣ истифода менамоянд. Чунин секунҷаҳо дар шкалаи мел бо таври мунтазам ҷойгир мешаванд. Табдилдиҳии басомад ба шкалаи мел бо ёрии формулаи

$$B(f) = 1125 \cdot \ln \left( 1 + \frac{f}{700} \right) \quad (4)$$

ичро карда шуда, дар ҳолати баръакс бошад бо формулаи (5) ифода карда мешавад:

$$B^{-1}(b) = 700 \cdot \left( e^{\frac{b}{1125}} - 1 \right) \quad (5)$$

Барои бузургии  $m = 0, 1, \dots, T - 1$  филтрҳои секунҷавии  $H_m(k)$  аз рӯи формулаи

$$H_m(k) = \begin{cases} 0, & k < f(m-1), \\ \frac{k - f(m-1)}{f(m) - f(m-1)}, & f(m-1) \leq k \leq f(m) \\ \frac{f(m+1) - k}{f(m+1) - f(m)}, & f(m) \leq k \leq f(m+1) \\ 0, & k > f(m+1) \end{cases} \quad (6)$$

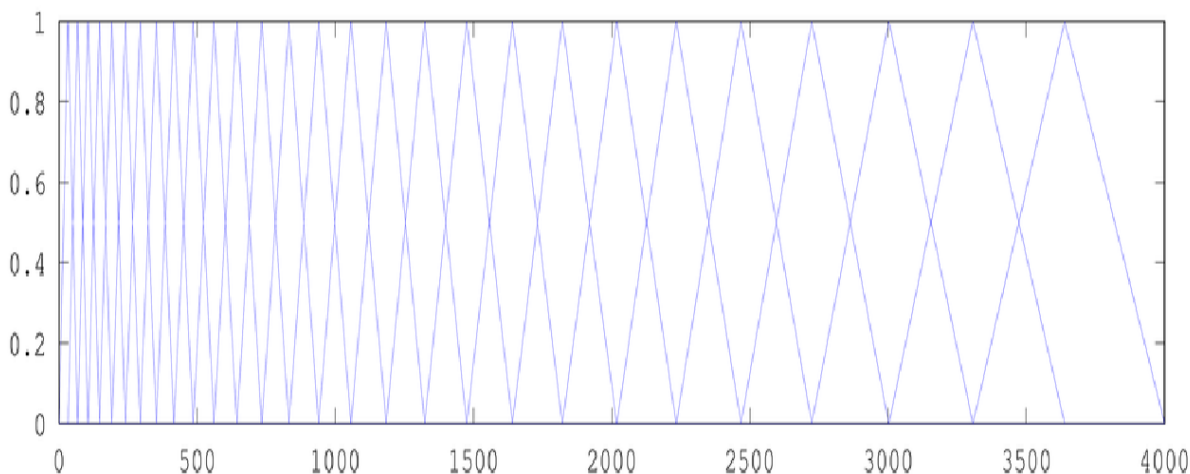
муайян карда мешавад, ки дар ин ҷо  $f(m-1)$ ,  $f(m)$  ва  $f(m+1)$  мувофиқан оғоз, байн ва охири филтрҳои секунҷавии филтри -ум буда, дар асоси формулаи

$$f(m) = \frac{T}{F_s} B^{-1} \left( B(f_{low}) + (m+1) \frac{B(f_{high}) - B(f_{low})}{M+1} \right), \quad (7)$$

ҳисоб карда мешавад. Дар ифодаи (7)  $f_{low}$  ва  $f_{high}$  мувофиқан, сарҳади поёнӣ ва болоии фосилаи басомади таҳлилшаванда (барои таҳлили овозҳои табиӣ, ки одам шунида метавонад ва баробар ба  $f_{low} = 16$  ва  $f_{high} = 20000$ ),  $F_s$ - басомади ба рақамгардонии садо аст.

Дар расми 1, намунаи 28 филтрҳои секунҷавии мел-басомад барои диапазони 0 – 4000 Хс дар асоси формулаи (6) тасвир карда шудааст.

Расми 1.



Расми 1. Намунаи филтрҳои секунҷавӣ

Чуноне ки аз расми 1 дида мешавад, барои ҳисоб намудани логарифми қувваи спектр вобастагии вақт бо басомад мавқеи асосиро дорад. Барои маҷмӯи филтрҳои секунҷавии сохташуда логарифми қувваи спектр бо ёрии формулаи зерин ҳисоб карда мешавад:

$$E_m = \ln \left( \sum_{k=0}^{\frac{T}{2}} |Y_k|^2 H_m(k) \right), \quad m = 0, 1, 2, \dots, M - 1 \quad (8)$$

дар ин ҷо  $|Y_k|^2$  – модули адади комплекси  $Y_k$  мебошад. [5]

1. Табдилдиҳии дискретии косинусӣ барои ҳисоб намудани қувваи логарифмии дар қадами пешина ҳисобшуда, бо ёрии формулаи (9) амалӣ карда мешавад:

$$c_n = \sum_{m=0}^{M-1} E_m \cos\left(\frac{\pi\left(m + \frac{1}{2}\right)n}{M}\right), n = 0, 1, 2, \dots, M - 1 \quad (9)$$

Таҷриба нишон медиҳад, ки натиҷаи MFCC барои якҷанд элементҳои аввалини вектори кепстралии  $c_n$  интиҳоб мешаванд. Одатан, 13 элементи аввалини ин вектор дар ҳалли масъалаи таҳқиқшаванда кифоя аст [5].

Моделҳои математикии тарҳрезшуда, дар мақолаи мазкур вобаста ба формулаҳои математикии ҳисобӣ аз он шаҳодат медиҳад, ки дар коркарди барномаи компютерӣ натиҷаҳои дақиқ ба даст оварда мешавад. Ҳамзамон дар иҷрои формулаҳои (6) ва (9) дидан мумкин аст, ки дарозии маълумотҳои барои коркарди оянда истифодашаванда ба миқдори катакҳои тарҳрезшуда баробар аст. Пас натиҷаи охири чунин аст, ки ба ҷои коркарди 400-500 адад, коркарди 10-30 адад кифоя аст, яъне чунин тарз кори модели ҳисоби дахҳо маротиба тез мешавад ва барои ягон миқдори муайяни катакҳои сохтори модули бетағйир мемунад.

#### АДАБИЁТ

1. У. А. Ли, Э. П. Нейбург, Т. Б. Мартин, Дж. Р. Уэлч, В. У. Зу, Р. М. Шварц, Дж. Е. Шуп, А. Р. Смит, М. Р. Самбур, Ф. Хейс-Роз, Г. Гудмэн, Р. Редди. Методы автоматического распознавания речи: В 2-х книгах. Пер. с англ. / Под ред. У. Ли. – М.: Мир, 1983. – Кн. 1. 328 с.
2. Ermilov A. V. Speech Technologies in human computer interactions // International Journal of Modern Manufacturing Technologies. 2013. Vol. 4. P. 52-57.
3. Larochelle H., Erhan D., Courville D. An Empirical Evaluation of Deep Architectures on Problems with Many Factors of Variation // International Conference on Machine Learning. 2007.
4. Анализатор качественных показателей передачи речи, DSLA II - Digital Speech Level Analyser [Электронный ресурс] / УП "ОПАТОВ" // Режим доступа: <http://www.Opatov.ru/content/view/125/64/lang, ru/>
5. Аграновский А.В. Теоретические аспекты алгоритмов обработки и классификации речевых сигналов / Аграновский А.В., Леднов Д.А. // Москва, «Радио и связь», 2004
6. Бабин Д.Н. О перспективах создания системы автоматического распознавания слитной устной русской речи / Бабин Д.Н., Мазуренко И.Л., Холоденко
7. А.В. //Интеллектуальные системы, т.8, вып. 1-4, 2004, с. 45-70
8. <http://www.frolov-lib.ru>
9. [https://en.wikipedia.org/wiki/Mel-frequency\\_cepstrum](https://en.wikipedia.org/wiki/Mel-frequency_cepstrum)
10. [https://en.wikipedia.org/wiki/Speech\\_recognition](https://en.wikipedia.org/wiki/Speech_recognition)

#### LITERATURA

1. W. A. Lee, E. P. Neuburg, T. B. Martin, J. R. Welch, W. W. Zu, R. M. Schwartz, J. E. Shoup, A. R. Smith, M. R. Sambur, F. Hayes-Rose, G. Goodman, R. Reddy. Methods of automatic speech recognition: In 2 books. Per. from English. / Ed. W. Lee. - M.: Mir, 1983. - Book. 1. 328 p.
2. Ermilov A. V. Speech Technologies in human computer interactions // International Journal of Modern Manufacturing Technologies. 2013. Vol. 4. P. 52-57.
3. Larochelle H., Erhan D., Courville D. An Empirical Evaluation of Deep Architectures on Problems with Many Factors of Variation // International Conference on Machine Learning. 2007.
4. Analyzer of quality indicators of speech transmission, DSLA II - Digital Speech Level Analyzer [Electronic resource] / UE "ОПАТОВ" // Access mode: <http://www.Opatov.ru/content/view/125/64/lang, ru/>
5. Agranovsky A.V. Theoretical aspects of algorithms for processing and classifying speech signals / Agranovsky A.V., Lednov D.A. // Moscow, "Radio and communication", 2004.
6. Babin D.N. On the prospects for creating a system for automatic recognition of continuous oral Russian speech / Babin D.N., Mazurenko I.L., Kholodenko
7. А.В. //Intelligent systems, v.8, no. 1-4, 2004, pp. 45-70
8. <http://www.frolov-lib.ru>
9. [https://en.wikipedia.org/wiki/Mel-frequency\\_cepstrum](https://en.wikipedia.org/wiki/Mel-frequency_cepstrum)
10. [https://en.wikipedia.org/wiki/Speech\\_recognition](https://en.wikipedia.org/wiki/Speech_recognition)