

1.1.6.[01.01.07] МАТЕМАТИКАИ ҲИСОББАРОР  
1.1.6.[01.01.07] ВЫЧИСЛИТЕЛЬНАЯ МАТЕМАТИКА  
1.1.6.[01.01.07] COMPUTATIONAL MATHEMATICS

ТДУ 519.2

**АЛГОРИТМИ БОЗШИНОСИИ НУТҚ** *Дадобоев Акмал Анварҷоновиҷ, докторант (PhD)-и кафедраи технологияҳои иттилоотию коммуникатсионӣ ва барномарезии ДДҲБСТ (Тоҷикистон, Хуҷанд)*

**АЛГОРИТМ РАСПОЗНАВАНИЯ РЕЧИ** *Дадобаев Акмал Анварҷоновиҷ, докторант (PhD) кафедраи информационно-коммуникационных технологий и программирования ТГУПБП (Таджикистан, Худжанд)*

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

**Вожаҳои калидӣ:** латтишҳои овозӣ, амплитуда, фрейм, қиммати энтропия, нуқтаҳои дискретизатсия, сигнали овозӣ,  $N$

Дар мақола масъалаи бозшиносии нутқ ҳамчун муайянкунии функсияи  $SR:A \rightarrow B$ , аз рӯи ду хусусият – амплитуда ва басомад исбот карда шудааст. Сабаби ҷудокунии нутқ ба фреймҳо исбот карда шудааст. Фрейм дар намуди вектори  $x = (x_1, x_2, \dots, x_N)$ , ки дар ин ҷо  $N$  – андозаи фрейм аст, тавсиф шудааст. Инъикоси табдилёбии вектори андозааш  $N$ -ро ба вектори андозааш  $M$  муайян гардидааст.

**Ключевые слова:** голосовой всплеск, амплитуда, кадр, значение энтропии, точки дискретизации, голосовой сигнал,  $N$

В статье обосновывается задача распознавания речи как определение функции  $SR:A \rightarrow B$  по двум ее характеристикам – амплитуде и частоте. Доказана причина разделения речи на кадры, кадр описывается в виде вектора  $x = (x_1, x_2, \dots, x_N)$ , где  $N$  — размер кадра, отражения преобразования вектора размер  $N$  в вектор размера  $M$  определен.

**Key word:** voice splash, amplitude, frame, entropy value, discretization points, voice signal,  $N$

In the article, the problem of speech recognition is proved as the determination of the function  $SR:A \rightarrow B$ , based on its two characteristics - amplitude and frequency. The reason for dividing the speech into frames is proved, the frame is described in the form of a vector  $x = (x_1, x_2, \dots, x_N)$ , where  $N$  is the size of the frame, the reflection transformations of the vector of size  $N$  into the vector of size  $M$ , is defined.

Раванди бозшиносии нутқ (англ. Speech Recognition) ин табдилдиҳии сигнали нутқӣ ба маълумоти рақамӣ мебошад. Масъалаи бозшиносии нутқро мувофиқгузории элементи маҷмуи нишонаҳои сигнали овозӣ ба пайдарпайии калимаҳо ташкил медиҳад. Қайд бояд карда шавад, ки ин мувофиқгузорӣ дар асоси эҳтимолияти калонтарин амалӣ мегардад. Яъне, ба нишонҳои сигнали овозӣ ҳамон калима мувофиқ гузошта мешавад, ки эҳтимолияти калонтарини монандӣ дар асоси ягон меъёрро дорад.

Ҳамин тарик, масъалаи бозшиносии нутқро ҳамчун муайянкунии функсияи

**$SR:A \rightarrow B$ ,**

ки соҳаи муайяниаш (**A**) маҷмуи нишонаҳои сигнали овозӣ ва соҳаи қиматқабулкуниаш (**B**) пайдарпайии пешакӣ маълуми калимаҳо тавсиф додан мумкин аст.

Вобаста ба масъалаи бозшиносии нутқ, системаҳои гуногуни бозшиносии фарқ мекунад:

1. Системаҳои бозшиносии нутқ аз рӯи луғат. Пайдарпайии калимаҳо барои мувофиқгузорӣ дар чунин системаҳо аз миқдори охири калимаҳои забони табиӣ иборат мебошад. Вобаста ба миқдори калимаҳо дар пайдарпайии системаҳои бо луғати хурд ва калонро фарқ мекунад;

2. Системаҳои бозшиносии нутқ аз талаффузгар вобаста ё новобаста. Раванди ташкили нишонаҳои овозиро (элементҳои маҷмуи якуми масъалаи бозшиносии) вобаста ба сарчашмаи

маълумоти овозӣ ташкил намудан мумкин аст. Маҳз дар ҳамин қадам хусусияти вобаста будан ё набудани системаи бозшиносии нутқ аз шахсе, ки нутқи ӯ бозшиносӣ гирифта мешавад, тавсиф меёбад.

Маълум аст, ки нутқ пайдарпайии овозҳо аст. Овоз бошад дар навбати худ суперпозитсияи лаппишҳои овозӣ (мавҷҳо)-и басомадашон гуногун аст. Мавҷ одатан аз рӯи ду хусусияташ – амплитуда ва басомад муайян карда мешавад.

Барои нигоҳдории сигнал дар барандаҳои рақамӣ онро ба миқдори зиёди фосилаҳо тақсим карда ва ягон қимати миёна дар ин фосиларо барои нигоҳдорӣ қабул мекунад. Ба ҳамин тариқ, лаппишҳои механикӣ ба маҷмуи ададҳо табдил дода мешаванд, ки онҳоро тавассути компютер нигоҳ дошта, коркард ва интиқол кардан имконпазир аст.

Дар ин мақола мазкур маълумотҳои овозӣ аз файлҳои форматҳои wav гирифта мешаванд. Ин формат – яке аз форматҳои маҷмултарин буда, дар дилхоҳ платформаҳо дастрас аст. Масъалаи гирифтани маълумотҳои овозӣ аз таҳлили сохтори файли WAV иборат аст.

Дар қисмати сарлавҳаи файл маълумот оид ба битрейт, басомад, миқдори каналҳо, дарозии файл ва ғайра нигоҳ дошта мешавад. Дар қисми маълумот маҷмуи қиматҳои амплитуда, ки овозро тавсиф мекунад, нигоҳ дошта мешавад. Вобаста ба ин, гирифтани маълумот аз файл аз хондани сарлавҳа, дастрас намудани маълумот оид ба файл ва маҳдудиятҳо (масалан, мавҷудияти маҷмуи маълумот) ва нигоҳдории маҷмуи ададҳои овозро тавсифкунанда (қимати амплитудаҳо) дар массивҳо иборат аст.

Қадами аввалии коркард пас аз гирифтани маълумот аз файлоҷудкунии он ба фреймҳо ташкил медиҳад. Фрейм – ҷудокунии маълумоти овозӣ ба фосилаҳои кутухмуддати вақт аст. Ҷудокунии маълумот ба фреймҳо бошад, пайдарпай иҷро нашуда, болои ҳам хобанда иҷро карда мешаванд. Яъне, охири як фрейм ба аввали фрейми дигар мувофиқ аст.

Сабаби ҷудокунии нутқ ба фреймҳо аз он иборат аст, ки муқоисакунӣ дар асоси фреймҳо нисбат ба муқоисакунӣ дар ҳар як лаҳзаи вақт хеле осонтар аст. Болои ҳам хобидани аввал ва интиҳои фреймҳо имкон фарҳам меорад, ки раванди таҳлили фреймҳо суфта карда шавад.

Ҳамин тариқ, бо назардошти он, ки дарозии миёнаи калимаи талаффузшуда ба 500 миллисония баробар аст (дар асоси таҷрибаҳои гузаронидаи муаллиф) ҷудокунии маълумоти калимаи талаффузшуда ба фреймҳо ба  $\frac{500}{10 \cdot 0,5} = 100$  фрейм барои ҳар як калимаро медиҳад.

Пас аз ба фреймҳо ҷудокунии маълумоти овозӣ, муайянкунии фосилаҳои овозӣ, яъне муддати вақте, ки хомӯширо ифода мекунад. Табиист, ки чунин фосилаҳои овозӣ аз коркард ба мувофиқгузори бояд хорич карда шаванд.

Якчанд роҳи амалисозии раванди хоричкунии фосилаҳои хомӯшӣ мавҷуданд, ки аз муайянкунии қимати параметри махсус иборат мебошад, яъне дар ҳолати қимати амплитуда аз қимати параметр хурд будан он хомӯшӣ бозшиносӣ мешавад; дар акси ҳол, ҳамчун қисми калимаи талаффузшуда [1].

Муайян намудани қимати параметри мазкур бо тарзҳои зерин иҷро карда мешавад:

- параметр ҳамчун доимӣ қабул карда мешавад. Дар ин ҳолат, агар сигнали аввала ҳама вақт дар ҳолати якхела ва тарзи якхела ба даст оварда шуда бошад, он гоҳ, натиҷаи дилхоҳро ба даст овардан мумкин аст. Иваз шудани ҷойи сабт (истифодаи системаи бозшиносӣ) ва ё таҷвизоти сатбкунанда ба таҳлили нодурусти раванди бозшиносӣ оварда мерасонад;

- ба гурӯҳҳо ҷудо кардани қиматҳои сигнал ва ба таври ошкор ҷудо намудани маҷмуи қиматҳои, ки ба хомӯшӣ мувофиқ мебошанд. Чунин роҳ натиҷаҳои хубро медиҳад, агар фосилаҳои хомӯши муддати вақти ба қадри кофӣ калонро ташкил карда бошанд;

- таҳлили энтропия.

Дар мақолаи ҳозира, вобаста ба гузориши пурраи масъалаи бозшиносӣ, таҳлили энтропия истифода бурда шудааст. Барои ин, пеш аз ҳама қимати энтропия барои ҳар як фрейм ҳисоб карда мешавад, ки аз марҳилаҳои зерин иборат аст:

- бигзор сигнали коркардшаванда рақамгузори шуда бошад ва қиматҳои порчаи  $[-1; 1]$ -ро қабул кунад;
- қимати энтропия аз рӯи формулаи

$$E = \sum_{i=0}^{N-1} P_i \cdot \log_2 P_i$$

ҳисоб карда мешавад, ки дар ин ҷо  $P_i$  – қиматҳои нормализатсияшудаи сигнали овозӣ,  $N$  – миқдори нуқтаҳои дискретизатсия дар фрейм мебошанд.

Ҳамин тариқ қимати энтропия ҳисоб карда мешавад. Вале, энтропия танҳо яке аз хосиятҳои фреймро ташкил медиҳад. Бинобар ин барои аз фосилаи хомӯшӣ ҷудо кардани нутқ, зарурати муқоисакунии он бо ягон қимат пеш меояд. Дар назари аввал, ба сифати чунин қимат миёнаи арифметикӣ (ё ягон миёнаи дигар, масалан, геометрӣ ва ғайра)-и қиматҳои калонтарин ва хурдтарини энтропия барои ҳамаи фреймҳоро гирифтани мувофиқи мақсад аст. Дар асоси таҷрибаҳои гузаронидашуда чунин роҳи муайянкунии қимати худудӣ натиҷаҳои дилхоҳро нишон надодаанд.

Аз рӯи муайянкунии қимати энтропия чунин ҳулоса намудан мумкин аст, ки он нисбат ба ҳисобкунии қиматҳои миёнаи вобаста нест, бинобар ин, имконияти муайянкунии қимати он ҳамчун доимӣ мавҷуд аст (дар мақолаи мо он ба 0,1 баробар гирифта шудааст).

Ба ин нигоҳ нақарда муаммои муайянкунии хомӯшӣ дар нутқ ва хоричкунии он аз раванди бозшиносӣ ба итмом намерасад. Алалхусус, қимати энтропия метавонад дар миёнаи калимаҳои талаффузшаванда хурд гардида, дар натиҷаи мавҷуд будани овозҳои бегонаи муҳит қиматҳои калонро соҳиб гардад. Барои бартараф намудани муаммои яқум зарурати ворид намудани мафҳуми «масофаи хурдтарин байни калимаҳо» пеш меояд ва фреймҳои байни ҳам аз сабаби хурд гардидани қимати энтропия ҷудошуда (аз масофаи хурдтарини байни калимаҳо хурд буда) ба ҳамдигар наздик намудан лозим аст. Барои пешигирии муаммои дуҷум мафҳуми «дарозии хурдтарини калима» ворид карда мешавад [2].

Пас аз коркарди аввалаи нутқ маҷмуи фреймҳо, ки ба калимаи муайян мувофиқ аст, ба даст оварда мешавад. Барои муқоисаи маълумоти бадастовардашуда ва маҷмуи калимаҳои додашуда муайянкунии метрика зарур аст. Дар асоси метрикаи муайяншуда, масофаи байни векторҳои фреймҳои овозӣ ва векторҳои маҷмуи қиматҳои инъикоси масъалаи бозшиносӣ ҳисоб карда мешавад ва вобаста ба «наздикии» векторҳо мувофиқат муайян карда мешавад.

Ба сифати метрика, масалан, миёнаи квадратири гирифтани мумкин аст, ки намуди зеринро дорад [3]:

$$\rho = \sqrt{\frac{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}{n}}$$

ки дар ин ҷо  $n$  – миқдори қиматҳои қабулшуда дар маҷмуи фреймҳо,  $x_i, y_i, i = \overline{1, n}$  – мувофиқан компонентаҳои векторҳои соҳаи муайянӣ ва қиматҳои инъикоси масъалаи бозшиносӣ мебошанд. Вале, чунин метрика аз сабаби калон будани қимати  $n$  ва на барои ҳама гуна қимати  $i, i = \overline{1, n}$  мавҷуд будани маълумоти зарурии раванди бозшиносӣ натиҷаи дилхоҳ намедиҳад.

Барои баланд бардоштани сатҳи бозшиносии нутқ (маълумоти овозӣ) зарурати хорич намудани маълумотҳое, ки ба бозшиносии нутқ таъсир намерасонанд ё кам таъсир мерасонанд ва интиҳоб танҳо аз он басомадҳое, ки ба нутқи инсон хос аст, гузаронида мешавад. Чунин табдилдиҳӣ тавассути муайян намудани коэффисиентҳои кепстралии Мел-басомадҳо (Mel-frequency cepstral coefficients – MFCC) иҷро карда мешавад [4]. Мувофиқи маълумоти сарчашма [Википедиа] MFCC – ин як намуди тасвири энергияи спектри сигнал ба ҳисоб меравад. Бартариҳои истифодаи онро чунин шарҳ додан мумкин аст:

- спектри сигнал, яъне, паҳнкунии сигнал аз рӯи базиси функцияҳои ортогоналии косинусӣ (синусӣ) истифода мешавад, ки имконияти ба назар гирифтани хусусиятҳои мавҷии сигналро дар таҳлилҳои оянда фароҳам меорад;

- спектр ба mel-ҷадвали қиматҳои махсус проексия карда мешавад, ки имконияти ҷудокунии қиматҳои маънидори басомадҳо, ки барои қабули инсон мувофиқ аст;

- миқдори коэффисиентҳои ҳисобшавандаро метавон бо дилхоҳ қимат маҳдуд кард (масалан, дар тадқиқоти бурдашуда қимати 13 гирифта шудааст). Чунин интиҳоб имкон медиҳад, ки фреймҳо ба намуди ихчам табдил дода шаванд ва ба ин васила миқдори ҳисобкунии компютерӣ коҳиш меёбад.

Ҳисобкунии коэффисиентҳои MFCC-ро дар мисоли ягон фрейм дида мебароем. Бигзор фрейм дар намуди вектори  $\vec{x} = (x_1, x_2, \dots, x_N)$ , ки дар ин ҷо  $N$  – андозаи фрейм аст, тасвир карда шуда бошад.

Дар қадами аввал спектри сигнали ин фреймро тасвиркунанда бо ёрии табдилдиҳии дискретии Фурье табдил дода мешавад:

$$\tilde{x}_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i k n}{N}}, \quad 0 \leq k < N$$

Илова бар ин, барои киматҳои ҳисобкардашуда функсияи оинагии Хемминг

$$H_k = \frac{27}{50} - \frac{23}{50} \cos \frac{2\pi k}{N-1}$$

барои суфта кардани киматҳо дар канорҳои фрейм татбиқ карда мешавад. Дар натиҷа чунин векторро ба даст овардан мумкин аст:

$$\tilde{X}_k = \tilde{x}_k \cdot H_k, \quad 0 \leq k < N.$$

Бояд ба назар гирифт, ки пас иҷрои чунин табдилдиҳӣ тири **Ox** басомадҳоро ифода намуда, тири **Oy** магнитудаҳоро ифода мекунад.

Ҳисобкунии mel-коэффисиент воҳиди психо-физикии баландии овоз мебошад, ки ба дарккурии он аз тарафи инсон ба ҳисоби миёна асос карда шудааст. Он пеш аз ҳама аз басомади овоз (баландии овоз ва тембр) вобаста аст. Бинобар ин, тавассути он мо метавонем маълумотҳои дар фрейм мавҷударо аз рӯи маълумоти барои бозшиносӣ зарурӣ арзёбӣ кунем.

Барои табдилдиҳии басомад (**F**) ба мел (**M**) аз формулаи зерин истифода бурдан мумкин аст [5]:

$$M = 1127 \cdot \log \left( 1 + \frac{F}{700} \right) \quad (1)$$

Табдилдиҳии баръаксӣ бошад, намуди зеринро дорад:

$$F = 700 \cdot e^{\frac{M}{1127}} - 1 \quad (2)$$

Бигузур дар мисоли мавриди назар  $N = 256$  бошад, яъне фреймҳои чедошуда дорои андозаи **256** бошанд. Басомади маълумоти овозии коркардшаванда (онро аз сарлавҳаи файли овозӣ дастрас намудан мумкин аст) дар ҳамин фреймро **16000**hz қабул мекунем. Агар нутқи инсон дар порчаи **[300; 8000]** hz хобида бошад, онгоҳ миқдори mel-коэффисиентҳоро  $M = 10$  мегузорем.

Барои паҳнкунии спектри ба даст овардашуда аз рӯи mel-чадвал зарурати сохтани пайдапайии филтрҳо пеш меояд. Ҳар як mel-филтр (элементи mel-чадвал) ин худ функсияи оинагии секунҷавӣ мебошад, ки имконияти суммиронидани миқдори энергия дар порчаи муайяни басомадҳоро фароҳам меорад ва ба ин роҳ mel-коэффисиентҳоро ба даст овардан мумкин аст. Миқдори mel-коэффисиентҳо ва порчаи зарурии басомадҳои таҳлилшаванда ро пешакӣ муайян намуда, маҷмуи филтрҳоро ташкил намудан мумкин аст.

Бояд диққат дода шавад, ки ба mel-коэффисиентҳои рақами тартибиашон калон, асоси филтри мувофиқ ба рақами тартибӣ васеътар аст.

Ҳамин тариқ, дар ҳолати мавриди таҳлил басомадҳо пас аз табдилдиҳӣ ба mel-чадвал аз рӯи формулаи (3.1) порчаи **[300; 8000]** ба порчаи **[401,25; 2834,99]** табдил меёбад.

Барои ташкили **10** адад филтрҳои секунҷавӣ **12** адад нуқтаҳои тақиягоҳӣ заруранд. Порчаи киматҳои ба даст овардашударо ба 11 қисмҳои баробар аз рӯи формулаи зерин тақсим менамоем:

$$x_i = a + \frac{b-a}{n}$$

ки дар ин ҷо,  $x_i$  – киматҳои тақсимот;  $a$ ,  $b$  – мувофиқан кимати аввал (**401,25**) ва интиҳои (**2834,99**) порча;  $n$  – миқдори зарурии тақсимот ва  $i = \overline{1, n}$  мебошанд. Дар асоси тақсимот маҷмуи киматҳои

$$m_i = \{401,25; 622,5; 843,75; 1065; 1286,25; 1507,5; 1728,74; 1949,99; 2171,24; 2392,49; 2613,74; 2834,99\}$$

-ро ҳосил мекунем, ки порчаи додашударо ба **11** порчаи баробар тақсим мекунад. Бояд диққат дода шавад, ки дар mel-чадвал нуқтаҳои тақсимот аз ҳамдигар дар масофаи баробар меҳобанд.

Қиматҳои тақсимои ҳосилшударо бо истифода аз формулаи (3.2) баръакс ба басомадҳо табдил дода, ҳосил мекунем:

$$h_i = \{300; 517,33; 781,9; 1103,97; 1496,04; 1973,32; 2554,33; 3261,62; 4122,63; 5170,76; 6446,7; 8000\}$$

Қадвали тақсимои басомадҳо, чи хеле ки дида мешавад, баробар тақсим намешаванд ва монотонӣ васеъ мегарданд. Дар натиҷа, зиёдшавии қиматҳои барои басомадҳои қимати хурд дошта ва калондошта баробар мегардонанд.

Тақсимои ба даст овардари ба спектри фрейм мувофиқ гузошта, қиматҳои нуқтаҳои тақягоҳиро дар асоси таносуби ҳосилшаванда муайян кардан мумкин аст: дарозии спектр 256 элемент ва ба он 16000 hz дохил мешавад. Аз таносуби ҳосилшуда формулаи зеринро ҳосил мекунем:

$$f_i = \left[ (\text{Андозаи фрейм} + 1) \cdot \frac{h_i}{\text{басомад}} \right] = \left[ \frac{256 \cdot h_i}{16000} \right]$$

Дар мисоли мавриди баррасӣ чунин қиматҳои нуқтаҳои тақягоҳиро ба даст меорем:

$$f_i = \{4,8,12,17,23,31,40,52,66,82,103,128\}$$

Аз рӯи қиматҳои ҳисобкардашудаи нуқтаҳои тақягоҳӣ филтрҳои заруриро тартиб медиҳем:

$$H_m^k = \begin{cases} 0, & k < f_{m-1} \\ \frac{k - f_{m-1}}{f_m - f_{m-1}}, & f_{m-1} \leq k \leq f_m \\ \frac{f_{m+1} - k}{f_{m+1} - f_m}, & f_m \leq k \leq f_{m+1} \\ 0, & k > f_{m+1} \end{cases}$$

Барои истифодаи филтрҳои ба даст овардашуда, кифоя аст, ки онҳо чуфт-чуфт бо қиматҳои спектр зарб карда шаванд. Дар натиҷаи зарбзани mel-коэффисиентҳо ба даст овардан мумкин аст. Азбаски  $M$  адад филтр ҳисоб карда шуда аст, бинобар ин миқдори коэффисиентҳо низ  $M$  адад аст:

$$S_m = \log \left( \sum_{k=0}^{N-1} |X_k|^2 \cdot H_m^k \right), \quad 0 \leq m \leq M.$$

Вале, mel-филтрҳои ҳисобкарда на ба қиматҳои спектр, балки ба қимати энергияи онҳо зарб шаванд. Бинобар ин, пас аз суммиронӣ натиҷаҳо бояд логарифмиронӣ карда шаванд. Дар чунин ҳолат сатҳи таъсири ғашҳо ба коэффисиентҳо паст мегардад [4].

Дар интиҳо, mel-коэффисиентҳои ҳисобкарда тавассути табдилдиҳии дискретии косинусӣ мавриди коркард қарор гирифта, коэффисиентҳои кепстралӣ [6] ба даст оварда мешаванд. Табдилдиҳии дискретии косинусӣ (DCT, [5]) барои ҳамъ кардани натиҷаҳои ба даст овардашуда бо роҳи баланд намудани қимати коэффисиентҳои аввала ва паст кардани коэффисиентҳои охир истифода бурда мешавад. Дар тадиқотҳои гузаронидашуда табдилдиҳӣ бо ёрии DCT-II [7] иҷро карда шудааст (бидуни зарб кардан ба коэффисиенти андозавии  $\sqrt{2/N}$ ):

$$C_l = \sum_{m=0}^{M-1} S_m \cdot \cos \left( \frac{\pi}{M} \cdot \left( m + \frac{1}{2} \right) \cdot l \right), \quad 0 \leq l < M.$$

Ҳамин тариқ, ҳамаи коэффисиентҳои кепстралӣ ҳисоб карда мешаванд. Дидан мумкин аст, ки табдилдиҳии қайдшуда инъикоси вектори андозааш  $N$ -ро ба вектори андозааш  $M$  амалӣ мегардонад. Дар натиҷаи табдилдиҳӣ аз маълумотҳои сарчашмаи овозӣ танҳо он қисме, ки бевосита ба нутқи инсон тааллуқ доранд, гузошта шуда, таъсири маълумотҳои дигар ба назар гирифта, хориҷ карда мешаванд. Ин имконият медиҳад, ки муқоисаи мувофиқати нутқи рақамӣ бо қолабҳо дақиқ гардад ва раванди бозшиносӣ (яъне, мувофиқгузорӣ) нисбат ба вақт хурд карда шавад. Ба ҳамин васила, тезкории раванди шинохтагирӣ таъмин мегардад.

#### ПАЙНАВИШТ:

1. 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/>

2. Аграновский, А.В. Теоретические аспекты алгоритмов обработки и классификации речевых сигналов / Аграновский А.В., Леднов Д.А. - Москва, Радио и связь, 2014
3. Бабин, Д.Н. О перспективах создания системы автоматического распознавания слитной устной русской речи /Бабин Д.Н., Мазуренко И.Л., Холоденко.-Москва, 2018
4. Дадобаев, А. А. Компьютерные программы для обеспечения распознавания голоса и сегментации речи/А.А.Дадобаев//Сборник статей XXV Международной научно-практической конференции, состоявшейся 28 ноября 2022 г. в г. Петрозаводске, Российская Федерация МЦНП «Новая наука», 2022
5. Ли, У. А., Методы автоматического распознавания речи: В 2-х книгах. Пер. с англ. /Под ред. У. Ли/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. – М.: Мир, 2014 – -328 с.- 1 книга.
6. Ermilov, A. V. Speech Technologies in human computer interactions/A.V. Ermilov//International Journal of Modern Manufacturing Technologies. 2013. Vol. 4. P. 52-57.
7. Larochelle, H.An Empirical Evaluation of Deep Architectures on Problems with Many Factors of Variation/H.Larochelle, D.Erhan,D.Courville//International Conference on Machine Learning. 2017.

**REFERENCES:**

1. Analyzer of quality indicators of speech transmission, DSLA II - Digital Speech Level Analyzer [Electronic resource]/Unitary Enterprise "OPATOV"//Access mode: <http://www.Opatov.ru/content/view/125/64/lang,ru/>
2. Agranovsky A.B. Theoretical aspects of algorithms for processing and classification of speech signals / Agranovsky A.B., Lednov D.A. // Moscow, "Radio and Communications", 2014
3. 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 //Moscow, 2018.
4. Dadobaev A. A. Computer programs for voice recognition and speech segmentation: Collection of articles of the XXV International Scientific and Practical Conference, held on November 28, 2022 in Petrozavodsk, Russian Federation ICNP "New Science" 2022
5. Lee, W. A., 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. – М.: Mir, 2014 – Book. 1. 328 p.
6. Ermilov A. V. Speech Technologies in human computer interactions // International Journal of Modern Manufacturing Technologies. 2013. Vol. 4. P. 52-57.
7. Larochelle H., Erhan D., Courville D. An Empirical Evaluation of Deep Architectures on Problems with Many Factors of Variation // International Conference on Machine Learning. 2017.