

**ЎЗБЕК ТИЛИ НУТҚИЙ СЎЗЛАРИНИ ТАНИШНИНГ ЯШИРИН
МАРКОВ МОДЕЛИ***Берданов У.А.*

Ушбу мақолада Марков моделининг қўлланилиши натижасида ўзбек тили сўзларини аниқлаш модел структураси ишлаб чиқилган. Нутқий сигналларни қайта ишлаш босқичлари келтирилган. Нутқий сўзларни параметрларга ажратиш ва уларни гурухлаш масалалари ечилган. Сўзларни аниқлаш учун Марков моделнинг ишлаш принципи кенг ёритилган. Ўзбек тили нутқий сўзларни аниқлаш устида тажрибалар ўтказилган.

Таянч иборалар: нутқ сигнали, тезкор Фурье ўзгартириш, спектограмма, яширин Марков модели(ЯММ), ўзбек тили эталон сўзлари.

В этой статье, с применением Марковских моделей разработана модель распознавания узбекских слов. Приведены этапы обработки речевых сигналов, решаются задачи параметризации речевых сигналов и их группировки. Подробно описан алгоритм работы Марковских цепей при распознавании отдельных слов. Проведен ряд эксперименты по распознаванию узбекской речи.

Ключевые слова: речевой сигнал, быстрое Фурье-преобразование, спектограмма, скрытый Марковские модели, эталоны слов узбекского языка.

In this article, as a result of application of Markov models the recognition model Uzbek words is developed. Today, speech signals are widely used in recognition of hardware systems. Such speech recognition systems create the possibility recognizing speech signals in several languages. Analysing of their hardware's software and models is one of the most complex issues. Also, studying and analysing the principles of hardware and software processing of speech signals is being required. Besides, developping mathematical models of speech processing of the Uzbek language, creating of their algorithms and analyzing is one of the important tasks of the research. Principle of recognizing two and three-syllable words in Uzbek language by based on parameterization was performed. Problems of parametrization of speech signals and their group are solved. It is given stages of processing of speech signals. It is detailed described the principle of work of Markov chains when determining words. In the Markov model, the parameters of speech signals as incoming values were studied. Study for the two-syllable words found in 20 out of 17 and 3 of them are vaguely defined. Each found vaguely

defined word has an error. In this study, all the three-syllable words were completely clarified. Based on the hidden Markov model, the degree of detection in the processing of the words in the Uzbek language was 80-85%. Thus, the proposed model for processing of Uzbek language can be considered as effective.

Key words: speech signal, Fast Fourier transform, spectrogram, Hidden Markov models, etalon words Uzbek language.

I. КИРИШ

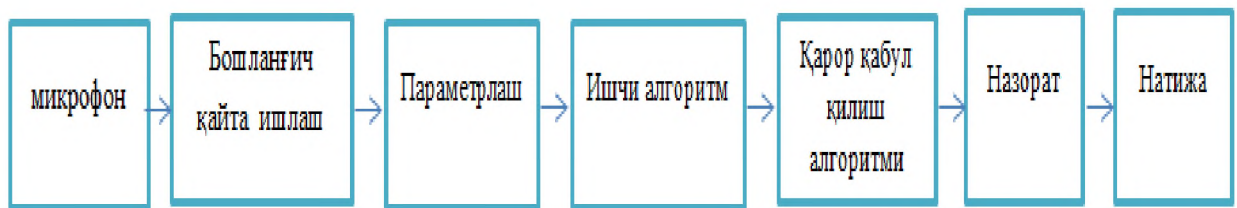
Бугунги кунда нутқий сигналларни таниш тизимлари кенг соҳаларда фойдаланилмоқда. Бундай нутқни таниш тизимлари бир нечта тилларда нутқий сигналларни таниш имкониятини яратиб беради. Нутқий сигналларнинг дастурий тамилотини яратиш ва моделларини таҳлил қилиш мураккаб масалалардан биридир. Шунингдек давр талаби нутқий сигналларни қайта ишловчи аппарат ва дастурий воситаларнинг ишлаш принципларини ўрганишни, таҳлил қилишни талаб этмоқда. Бундан ташқари ўзбек тили нутқий сигналларни қайта ишлаш жараёнининг математик моделларини ишлаб чиқиш, уларнинг алгоритмларини яратиш ва таҳлил қилишни амалга ошириш муҳим вазифалар саналади. Шунини алоҳида қайд этиш лозимки, ҳозирда инсон нутқи орқали бошқарилувчи аппарат воситаларининг ривожини куннинг замонавий талабига чиқмоқда. Бугунги кунда бундай тизимларни яратишда олимлар томонидан бир қанча алгоритмлар ва математик методлар яъни нутқни сигментларга ажратиш, филтрлаш усуллари, спектрал қийматлардан фойдаланиш, мел-частотаси, кепстр коэффицентлари, нейрон тармоқлари ва динамик дастурлаш алгоритмларидан унумли фойдаланилмоқда. Бундай алгоритмлардан фойдаланиш натижасида турли тиллардаги нутқий сўзларни аниқловчи дастурлар юзага келмоқда. Ҳозирги кунгача яратилган нутқни таниш тизимлари орасида ўзбек тилини танувчи дастурлар ва моделлар мавжуд эмас. Бундай дастурий воситалар ва моделларни яратилиш бугунги куннинг муҳим аҳамият касб этадиган муаммоли вазифаларидан биридир [1].

II. АСОСИЙ ҚИСМ

Инсон нутқини таниш тизимлари аппарат ва дастурий тамилотини таҳлил қилишда энг аввало инсон нутқининг шаклланиш жараёнини ўрганишни талаб қилади. Инсон нутқи сўзлаганида турли товушларни талаффуз қилади. Инсон талаффуз қиладиган товушлар нутқ товушлари дейилади. Инсон нутқ товушлари нутқ аъзолари иштирокида ҳосил бўлади. Буларга инсоннинг ўпка, кекирдак, товуш пайчалари, оғиз бўшлиғи, катта ва кичик тиллар, юмшоқ ва қаттиқ танглайлар, юқори ва пастки лаблар, юқори ва пастки тишлар, бурун бўшлиғи ва бўғиз бўшлиғи аъзоларини мисол сифатида келтириш мумкин. Ушбу инсон аъзолари нутқни тўлиқ талаффуз

килиш учун хизмат қилади. Нутқ товушлари сўз маъноларини фарқлаш учун хизмат қили, бу уларнинг асосий вазифасидир.

Инсон нутқининг ҳосил қилиниши учун бир қанча мураккаб жараён амалга оширилади. Нутқ товушларини ҳосил бўлишида ўпқадан чиқаётган ҳаво оқими нафас йўли, кекирдак орқали бўғизга, ундан оғиз бўшлиғи ёки бурун бўшлиғига ўтиб, ташқарига чиқади. Бундан келиб чиқиб, нутқий сигналларни ҳосил бўлиши дикторнинг характетистикасига боғлиқ деб тавсивлаш мумкин. Ўзбек тили нутқий сигналларини қайта ишлаш жараёнини масаланинг қўйилишига қараб бир неча кетма-кетликларга ажратиш мумкин. Ушбу кетма-кетликларда масаланинг қўйилишига қараб алгоритмлар ўзгариши мумкин. Ушбу кетма-кетликнинг имконияти инсон нутқини шахсини ва мазмунини аниқлаш масалалари учун фойдаланиш мумкин. Масаланинг қўйилишидан келиб чиқиб, бу кетма-кетлик ўз иш режимида бир неча алгоритмлар жамланмасини ўзида шакллантиради. Ўзбек тилини қайта ишловчи умумий алгоритмлар кетма-кетлигини қуйидаги тартибда тавсивдаш мумкин(1-расм).



1-расм. Ўзбек тилини қайта ишловчи умумий алгоритмлар кетма-кетлиги

Нутқий сигналларни бошланғич қайта ишлашдан олдин аналог кўринишдан рақамли кўришга ўтказиш ишлари амалга оширилади. Бунда кўпинча нутқий сигналлар .wav файл форматига келтирилиб, дискретлаш частотаси 22050 Hz ва ёзиш канали моно танланади.

Бошланғич қайта ишлаш алгоритми нутқий сигнални бошланғич филтрациясини амалга оширади. Бунда чекли импульс характеристикали филтр(КИХ) ва чексиз импульс характеристикали филтр(БИХ)нинг алгоритмларидан фойдаланиш мумкин. Филтрлаш жараёнидан сўнг нутқий сигналдаги жимлик ҳолатлари йўқотилади. Шундан сўнг фреймларга ажратиш жараёни амалга оширилади. Фреймларга ажратиш нутқий сигнални тенг кесимларга бўлиш ва ҳар бир ажратилган кесим Ҳемминг ойнасидан ўтказилади. Ажратилган кесимларнинг Ҳемминг ойнасидан ўтказилидан асосий мақсад нутқий сигналнинг спектрларини олишда спектрал кийматларнинг чексизликка интилишининг олдини олишдир[2].

Параметрлаш алгоритмида нутқий сигналнинг фреймлари акустик, спектрал, формант частота, кепстрал параметрлар аниқланади. Умумий ҳисобда маълум бир параметрнинг танланиши масаланинг қўйилишига

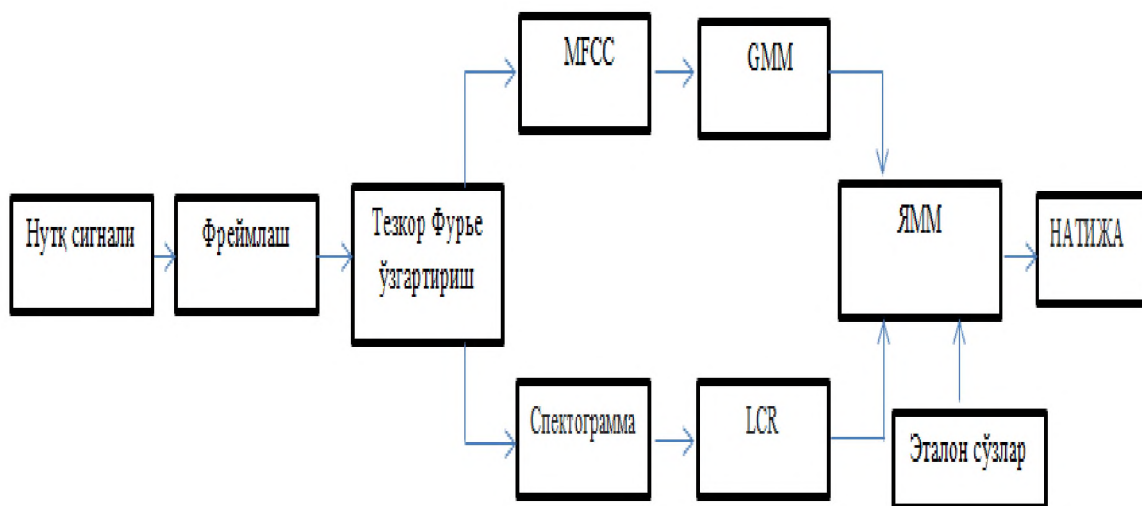
боғлиқ ҳолда танланади. Танланган параметрлаш алгоритмларидан олинган натижалар ишчи алгоритм қисмига юборилади. Таъкидлаш лозимки, ҳар бир параметрнинг ўз математик алгоритми мавжуд.

Ишчи алгоритмлар мавжуд аниқланган параметрларни устида иш олиб боровчи қисм ҳисобланади. Бунда параметрларнинг тўлиқлиги текширилади ва уларни гурухлаш ишлари амалга оширилади.

Қарор қабул қилиш алгоритми муҳим аҳамият қасб этадиган қисм ҳисобланади. Ушбу босқичда ишчи алгоритмлар асосида гурухланган параметрлар тўплами қарор қабул қилиш алгоритмлари асосида қайта ишланади. Ушбу босқич алгоритмлари гурухига корреляцион таҳлил, нейрон тармоқлари, марков занжирлари ва динамик дастурлаш алгоритмларини келтириш мумкин.

Назорат алгоритми қарор қабул қилиш алгоритмларидан олинган натижаларни умумлаштирилади. Яъни қайта ишланган фремерларнинг натижалари таҳлил қилинади ва якуний ҳулоса чиқарилади. Шундан сўнг натижаларни намойиш этиш вазифасини бажаради.

Ушбу алгоритмлар кетма-кетлиги асосида инсон нуқини қайта ишлашда қўлланилувчи бир қанча амалий масалаларни ечиш имконияга эга. Мисол сифатида, инсон нутқи асосида шахсини аниқлаш, инсон нутқининг мазмуни ва буйруқ орқали бошқарилувчи тизимларида фойдаланиш мумкин. Ушбу моделнинг имкониятидан фойдаланган ҳолда марков занжири асосида нутқий сўзларни аниқлашда ишлатилиши кўриб чиқамиз. Ушбу жараён куйидагича амалга оширилади (2-расм).



2-расм. Ўзбек тили сўзларини аниқловчи модел структураси

Ушбу структурадаги кетма-кетликни амалга ошириш учун нутқ сигнали ёзиб олинган. Бу модель сўзларнинг спектрал параметрлари ва спектограммасидан фойдаланиб яширин марков занжирлари асосида сўзларни аниқловчи алгоритмлар мажмуи ҳисобланади[3]. Ушбу моделда

кирувчи сигнал қийматлари КИХ фильтр асосида ташқи шумлардан тозаланади ва сегментлаш жараёни амалга оширилади. Сегментация жараёни – нутқ сигнали қийматларини маълум бир ўлчов асосида фреймларга ажратиш тушунилади. Ушбу моделда сегментлаш қиймати 256 га тенг. Яъни ҳар бир фрейм нутқий сигналининг 256 та қийматига эга бўлади ва қайта ишлаш жараёни тезкор Фурье ўзгартириш алгоритми асосида нутқий синалнинг спектрлари олинади.

Шундан сўнг олинган спектрал қийматлар икки қисмга ажратилган ҳолда қайта ишланади. Биринчи тур қайта ишлашда нутқ сигналининг олинган спектрлари MFCC(Mel-Frequency Cepstral Coefficients) алгоритмига йўналтирилади. Шундан сўнг ҳар бир фрейм учун қийматлар ҳисобланиб GMM(Gaussian Mixture Modeling) алгоритми асосида коэффицентлар олинади ва уларни гурухлаш ишлари амалга оширилади. Натижада чекли импульс характеристикали филтр яширин Марков модели (ЯММ) қайта ишлашга тайёр коэффицентлар гуруҳи шаклланади[4]. Иккинчи тур қайта ишлашда фреймларнинг спектрал қийматлари асосида спектограммаси қурилади. Шундан сўнг спектограмма LCR(left, center and right) методи ёрдамида бўлақларга ажратилади ва ҳар бир бўлақдан максимумлар, минимумлар ва чўққилар аниқланади. Аниқланган барча параметрлар гурухланади. Аниқланган гуруҳнинг барча параметрлар ЯММда қайта ишлашга тайёр ҳолатга келтирилади[5]. Шундан сўнг ушбу гуруҳланган параметрлар ЯММ асосида қайта ишланади. ЯММнинг умумий формуласи қуйидагича аниқланади. Берилган(гуруҳланган) параметрлар $Q = \{n_1, n_2, n_3 \dots n_m\}$ ҳолатлар учун иккита кетма-кет берилган s ва t ҳолатлар учун ўтиш эҳтимоллиги

$$p_{st} = P(x_i = t | x_{i-1} = s), \quad p_{in_1} + p_{in_2} + p_{in_3} + \dots + p_{in_m} = n_1$$

мавжуд бўлсин. Бунда барча ҳолатлар учун $i = n_1, n_2, n_3 \dots n_m$ ўринли бўлса ЯММ тўғри ташкил этилган бўлади(1). x_i ҳолатнинг аниқланиши x_{i-1} ҳолатга боғлиқ[6]:

$$P(x_i | x_0, x_1, \dots, x_{i-2}, x_{i-1}) = P(x_i | x_{i-1}) \quad (1)$$

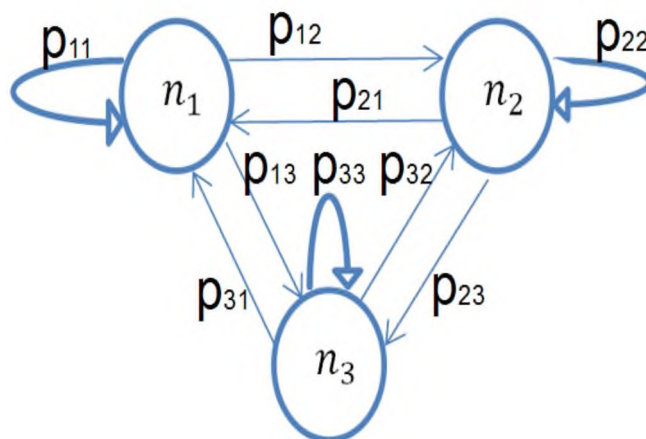
Бунда яқинлашиш эҳтимоли(2) х учун:

$$P(x) = P(x_1, x_2 \dots x_{L-1}, x_L) = P(x_2 | x_1) P(x_3 | x_2) \dots P(x_L | x_{L-1}) \quad (2)$$

қуйидаги шартли эҳтимоллик ўринли(3)

$$P(X, Y) = P(X|Y)P(Y) \quad (3)$$

Марков занжирининг амалий нуқтаи назардан қўлланилиши мисол сифатида учта ҳолат учун структураси қуйидаги 3-расмда келтирилган.



3-расм. Уч ҳолатли ЯММ структураси.

Бу ерда $Q = \{n_1, n_2, n_3\}$ яширин ҳолатлар, P_{in} – ўтиш эҳтимоллиги матрицаси. Ушбу моделдаги ўтиш эҳтимоллик матрицаси бир жинсли бўлганлиги учун қуйидагича бўлади(4):

$$P_{in} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} \quad (4)$$

Ўтиш эҳтимоллик матрицаси тартиби Марков занжирининг ҳолатлари сони m га тенг бўлади. Яширин Марков занжирида ўтиш эҳтимоллик матрицаси қуйидаги шартларни қаноатлантиради:

- ўтиш эҳтимоллик матрицасининг ҳар бир элементи $0 \leq p_{in} \leq 1$ оралиқда бўлади;
- ўтиш эҳтимоллик матрицасининг сатрлари йиғиндиси бирга тенг бўлиши керак;
- ўтиш эҳтимоллик матрицаси бирор устунининг барча элементлари нолга тенг бўла олмайди.

Нутқий сигналлар учун Марков занжирини ташкил этишда 29 та $Q = \{x_1, x_2, x_3 \dots x_{29}\}$ яширин ҳолат (ўзбек тилидаги ҳарфлар сони), ўтиш эҳтимоллиги ҳар бир яширин ҳолат учун 29 та (умумий сони $29 \cdot 29 = 841$) тугундан иборат ўтиш эҳтимоллик матрицаси тузилди. Оптимал йўлни аниқлаш учун Витерби алгоритмидан фойдаланилди[7]. Марков моделидаги олинган параметрларини маълумотлар базасидаги параметрларга мослиги аниқлаш вазифасини бажаради ва олинган натижаларни қайта ишлаш блокига юборади. Қайта ишлаш блоки қабул қилинган фреймларнинг натижа умумлаштиради ва саралаш вазифасини бажаради. Бунда мажуд аниқланган бир турдаги фреймлар кетма-кетлиги 15 тадан кам бўлса ташлаб юборилади. Агарда фреймлар сони 15 тадан юқори бўлса, ушбу фреймлар мажмуи натижа сифатида қайт этилади. Ушбу модель асосида ўзбек тили сўзларини аниқлаш учун икки ва уч бўғинли сўзлар

устида тадқиқот натижалар олинди. Ушбу сўзларни танлашда айтилиш нуқтаи назаридан энг содда сўзлар олинган ва ёзиб олиш жараёни шовқинсиз мухитда амалга оширилган.

Ўзбек тилидаги сўзларни таҳлил қилиш натижасида сўзнинг талаффузи, хар бир сўздаги фреймлар сони, харфлар, бўғинлар, унли ва ундош товушлар аниқланди. Сўзнинг аниқ топилганлиги 1 ёки 0 билан белгиланди. Тадқиқот учун танланган икки ва уч бўғинли сўзларнинг 20 тасидан 17 таси аниқ топилди ва 3 таси ноаниқ аниқланди. Ноаниқ аниқланган сўзларнинг хар бирида битта харфда хатолик мавжуд. Нутқий сигналларни қайта ишлаш натижасида олиб борилган таҳлил натижалари шуни кўрсатадики, ноаниқ топилган сўзлардаги харфлар сирғалувчи ундош товушлар синфига киради. Сирғалувчи ундош товушлар икки нутқ аъзоларининг ўзаро яқинлашиши ва хаво оқимининг ушбу нутқ аъзоларининг тор оралиғидан сирғалиб ўтиши натижасида ҳосил бўлади. Бундай товушлар сирасига қуйидаги товушлар v, f, j, z, y, s, x, sh, g', h киради. Энг кагта хатолик тил билан юқори милк орасидан ўтувчи j, z, s, sh ундош товушларда юзага келди.

Сўзларнинг аниқланиш даражаси топишда қуйидаги (5) формуладан фойдаланилади:

$$R = \left(\frac{n_s + n_i + n_d}{N} \right) \times 100 \quad (5)$$

Бу ерда,

R - тенглик тадқиқот натижасида қайта ишланган сўзларнинг аниқланиш фоизи;

N – аниқ топилган сўзларнинг умумий сони;

n_s – нотўғри аниқланган сўзларнинг умумий сони;

n_i – кўшилган сўзларнинг сони (кўшиш хатолиги: сўз маълумотлар базасига кўшилган, аниқланган сўз нутқда мавжуд эмас);

n_d – ўтказиб юборилган сўзлар сони (ўтказиш хатолиги: сўзлар маълумотлар базасида бўлган лекин тизим томонидан аниқланмаган).

Тажриба натижаларига асосан $N=20$, $n_s = 17$, $n_i, n_d=0$ га тенг. Юқоридаги формулани ҳисоблаш қуйидагича амалга оширилади(6).

$$R = \left(\frac{17+0+0}{20} \right) \times 100 = 85\% \quad (6)$$

Демак, (5) формула ёрдамида таклиф этилган моделнинг аниқлик даражасини 85% деб баҳолаш мумкин. Бу нисбатан юқори натижа деб ҳисобланади.

III. ХУЛОСА

Таҳлил натижаларидан шуни кўриш мумкинки, ўзбек тили нутқидаги сўзларни танишда ЯММдан фойдаланиш унимли ҳисобланади. Тажриба учун олинган икки ва уч бўғинли сўзларнинг аниқланишида танланган 20 та

сўздан 3 таси ноаниқ ва қолган 17 таси аниқ топилди. Ноаниқ топилган учта сўзнинг ҳар бирида биттадан товушда хатолик мавжуд. Ушбу товушлар сирғалувчи ундош товушлар эканлиги аниқланди. Агар сўзда сирғалувчи ундош товушлар мавжуд бўлмаса, сўзнинг аниқ топилиш эҳтимоли юқори эканлиги аниқланди. ЯММ асосида ўзбек тили сўзларини қайта ишлашда аниқланиш даражаси 85% аниқликни берди. Шундай қилиб, ўзбек тили сўзларини қайта ишлашда таклиф этилган моделни самарали деб баҳолаш мумкин.

АДАБИЁТЛАР

- [1] Мусаев М.М. Современные методы цифровой обработки речевых сигналов //” Вестник ТУИТ”, Ташкент, 2017, №2.с.2-13.
- [2] Берданов У.А. Ўзбек тили нутқининг қайта ишлашнинг корреляцион модели таҳлили. //” Вестник ТУИТ”, Ташкент, 2017, №3.с.10-18.
- [3] Bazzi, I, Glass, J. Modeling out of vocabulary words for robust speech recognition. Proc. ICASSP 2000, Beijing, China, Vol. 1, pp.401-404.
- [4] Daniel Ramage. Hidden Markov Models Fundamentals. CS229 Section Notes. 2007.
- [5] Yi Wang. The Variable-length Hidden Markov Model and Its Applications on Sequential Data Mining. Department of Computer Science, Tsinghua University, Beijing, China. 2005.
- [6] Elliott L., Ingham D., Kyne A., Mera N., Pourkashanian M., Whittaker S. «Efficient Clustering-Based Genetic Algorithms in Chemical Kinetic Modeling», GECCO 2004 Proceedings, Springer-Verlag Berlin Heidelberg, 2004, pp.932-944.