

ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ
МИНИСТРЛІГІ

«Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті»
коммерциялық емес акционерлік қоғамы

Автоматика және ақпараттық технологиялар институты

Электроника, телекоммуникация және ғарыштық технологиялар
кафедрасы

Мұратбек Аяулым Сұлтанбекқызы

«Хопфилд нейрожелісін қолданып, дыбыс сигналын танудың
математикалық моделі»

ДИПЛОМДЫҚ ЖҰМЫС

6B06201 «Телекоммуникация» білім беру бағдарламасы

Алматы 2025 ж

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ
МИНИСТРЛІГІ

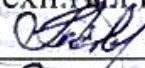
«Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті»
коммерциялық емес акционерлік қоғамы

Автоматика және телекоммуникациялық технологиялар институты
Электроника, телекоммуникация және ғарыштық технологиялар кафедрасы

ҚОРҒАУҒА ЖІБЕРІЛДІ

Кафедра меңгерушісі

техн. ғыл. кандидаты

 Е.Таштай

« 30 » 05 2025 ж.

ДИПЛОМДЫҚ ЖҰМЫС

Тақырыбы: «Хопфилд нейрожелісін қолданып, дыбыс сигналын
танудың математикалық моделі»

6B06201 – «Телекоммуникация» білім беру бағдарламасы

Орындаған: Мұратбек А.С.



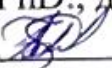
Рецензент

ҚазҰАЗУ, доктор PhD,
қауымд. профессоры

 Әлібек Н.Б.
« 15 » 05 2025 ж.

Ғылыми жетекші

ЭТЖҒТ каф. қауымд. профессор
PhD., докторы

 Тайсариева К.Н.
« 15 » 05 2025 ж.

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті

Автоматика және ақпараттық технологиялар институты

Электроника, телекоммуникация және ғарыштық технологиялар кафедрасы

6B06201 Телекоммуникация



Дипломдық жұмыс орындауға
ТАПСЫРМА

Білім алушы Мұратбек А.С.

Тақырыбы «Хопфилд нейрожелісін қолданып, дыбыс сигналын танудың математикалық моделі»

Университет ректорының «26» 01.2025 ж. бұйрығымен бекітілген.

Аяқталған жұмысты тапсыру мерзімі «30» сәуір 2025 ж.

Дипломдық жұмыстың бастапқы берілістері:

- 1) Спекторлық бөлу: жиіліктер жолағы 5-10 кГц;
- 2) Форматтық бөлу: 1-ші, 2-ші, 3-ші форматтағы диапазондар үшін, кГц:
а) қоңыраулар үшін (0,15-0,9; 0,55- 2,8; 1,5-3,4); б) саңырау дауыссыз (1-3,5; 2,5-5)

- 2) Дискретті Фурье түрлендіру алгоритмі: $Y = \frac{T}{N} WX$;

- 3) Хемминг функциясы үшін жақтау ұзақтығы 10 мс -тен 20 мс диапазонда;

- 4) Рекурентті нейрожелілер моделдері

$$f(X_z) = \begin{cases} 1, X_z \geq \Theta \\ 0, X_z < \Theta \end{cases}.$$

Дипломдық жұмыста қарастырылатын мәселелер тізімі:

- 1) Нейрожелілердің математикалық моделдерінің құрылымы
- 2) Хопфилд желісінің қабаттар арасындағы байланысы);
- 3) Нейрожелі моделдерін дыбысты тануға қолдануды өңдеу;
- 4) Matlab ортасында ұқсату модельдерін құру.

Сызбалық материалдар тізімі (міндетті сызбалар дәл көрсетілуі тиіс):

Ұсынылатын негізгі әдебиеттер:

1. Kim, J., & Park, D. (2022). Enhanced Speech Pattern Recognition Using Modified Hopfield Networks. IEEE Transactions on Neural Networks and Learning Systems, 33(4), 789-798

Дипломдық жұмысты (жобаны) дайындау
КЕСТЕСІ

Бөлімдер атауы, қарастырылатын мәселелер тізімі	Ғылыми жетекшіге және кеңесшілерге көрсету мерзімі	Ескерту
Дыбысты танудың сұлбасы және негізгі сипаттамалары	04.01.2025 - 01.02.2025	орындалды
Нейро желілердің жіктелуі Нейрожелілердің математикалық моделдерінің құрлымы	01.02.2025 - 01.03.2025	орындалды
Хопфилд нейро желілердің Matlab ортасында модельдеу	01.03.2025- 30.05.2025	орындалды.

Дипломдық жұмыс (жоба) бөлімдерінің кеңесшілері мен норма бақылаушының аяқталған жұмысқа(жобаға) қойған
ҚОЛТАҢБАЛАРЫ

Бөлімдер атауы	Кеңесшілер (аты, әкесінің аты, тегі, ғылыми дәрежесі, атағы)	Қол қойылған күні	Қолы
Диплом жұмысының тақырыбын талдау	Тайсариева К.Н., PhD., ЭТЖҒТ қауымдастырылған профессор.	13.03.25.	
Теориялық ақпарат	Тайсариева К.Н., PhD., ЭТЖҒТ қауымдастырылған профессор.	29.04.25	
Норма бақылау	Досбаев Ж.М. ЭТЖҒТ, PhD., аға оқытушысы	23.05.25	

Ғылыми жетекшісі



Тайсариева К.Н.

Тапсырманы орындауға алушы

алған білім



Мұратбек А.С.

Күні «2» ақпан. 2025 ж.

АНДАТПА

Бұл дипломдық жұмыс дыбыстық сигналдарды цифрлық форматқа түрлендіріп, оларды тану әдістерін зерттеуге арналған. Қазіргі таңда бұл сала мобильді және телекоммуникациялық технологиялардың дамуына байланысты қарқынды дамып келе жатқан ғылыми бағыттардың бірі болып отыр.

Жұмыс аясында дыбыстық сигналдарды тану әдістері мен олардың тарихи даму кезеңдері сараланып, нейрожелілердің құрылымдық ерекшеліктері мен қолдану бағыттарына жан-жақты шолу жасалды. Нақты түрде Хопфилд нейромоделі зерттеліп, MATLAB бағдарламасының `nntool` функциясы көмегімен дыбысты тануға арналған имитациялық модель жасалды. Тәжірибелік нәтижелер салыстырылып, әдістің тиімділігі бағаланды.

АННОТАЦИЯ

Данная дипломная работа посвящена исследованию методов преобразования звуковых сигналов в цифровой формат и их распознавания. В настоящее время эта область является одним из интенсивно развивающихся научных направлений благодаря активному росту мобильных и телекоммуникационных технологий.

В рамках работы были проанализированы методы распознавания звуковых сигналов и этапы их исторического развития, а также проведен подробный обзор структуры и сфер применения нейронных сетей. Особое внимание уделено модели Хопфилда, на основе которой с использованием функции `nntool` программной среды MATLAB была разработана имитационная модель для распознавания звука. Полученные экспериментальные результаты были сопоставлены, и дана оценка эффективности предложенного метода.

ANNOTATION

This thesis is devoted to the study of methods for converting audio signals into digital format and their recognition. Today, this field is one of the rapidly developing areas of science, driven by the active growth of mobile and telecommunication technologies..

The study analyzes audio signal recognition methods and their historical development stages, and provides a detailed overview of the structure and applications of neural networks. Special attention is given to the Hopfield model, based on which a simulation model for audio recognition was developed using the `nntool` function of the MATLAB software environment. The obtained experimental results were compared, and the efficiency of the proposed method was evaluated.

МАЗМҰНЫ

Кіріспе		
1	Дыбыстық сигналдарды танып білуге аналитикалық шолу	8
1.1	Дыбыстық сигналдарды танып білу тарихы	9
1.2	Дыбысты тану әдістеріне шолу	10
1.3	Дыбысты танудың сұлбасы және негізгі сипаттамалары	10
1.4	Дыбысты тану үшін пайдаланатын классикалық модельдер	23
1.5	Дыбыстық сигналды цифрлік сигналға түрлендіру	26
1.6	Вокодерлер түрінің сипаттамасы	26
2	Нейро желілер олардың моделдерін зерттеу	31
2.1	Нейро желілердің жіктелуі	33
2.2	Рекурентті нейрожелілер модельдері	36
2.3	Нейрожелілердің математикалық модельдерінің құрлымы	38
3	Нейрожелі моделдерін дыбысты тануға қолдануды өңдеу	41
3.1	Хопфилд нейрожелілі моделін қолданып дыбыс сигналын таанып білудің математитикалық моделі	43
3.2	Matlab ортасында модельдерін құру	45
Қорытынды		51
Пайдаланылған әдебиеттер тізімі		52
Қосымша А		53

КІРІСПЕ

Дыбыстық сигналдарды цифрлық форматқа түрлендіріп, оларды тану қазіргі таңда қарқынды дамып келе жатқан ғылыми бағыттардың бірі болып табылады. Бұл саланың негізгі ерекшеліктерінің бірі - күрделі және тиімді алгоритмдерді қолдану арқылы жоғары дәлдікпен нәтижелерге қол жеткізу. Мұндай жетістіктерге қол жеткізудің маңызды факторларының бірі - мобильді және телекоммуникациялық жүйелердің кеңінен дамып, осы технологияларға деген сұраныстың артуы.

Қазіргі уақытта дыбыстық сигналдарды тану мәселесіне үлкен назар аударылуда, бұл бағытта арнайы құрылғылар мен бағдарламалық жабдықтар жасалуда. Әсіресе, дыбысты автоматты түрде тану жүйелері кең көлемде қолданылуда және зерттелуде. Осы факторлар бұл мәселенің заманауи ғылым мен технология үшін өзектілігін нақтылай түседі.

Осы дипломдық жұмыстың негізгі мақсаты - дыбыстық сигналдарды тану әдістерін зерттеп, оларды нейрожелілік модельдер арқылы жүзеге асыру. Бұл мақсатты орындау үшін бірқатар міндеттер қойылып, төмендегідей мәселелер қарастырылды:

Дыбыстық сигналдарды тану әдістері мен олардың құрылымына талдау жасалып, тақырыптың тарихи даму кезеңдеріне шолу жүргізілді;

Жасанды нейрожелілердің құрылымы, классификациясы және қолдану салалары қарастырылып, олардың негізгі модельдері зерттелді;

Хопфилд нейрожелісі тереңірек зерттеліп, осы модель негізінде дыбыстық сигналдарды тануға арналған есептеулер жүргізілді;

Алынған математикалық модель MATLAB ортасының nntool функциясын пайдалану арқылы имитациялық модельге айналдырылып, Хопфилд нейромоделі негізінде дыбыстық сигналдарды тану тәжірибелік түрде жүзеге асырылды және нәтижелер салыстырылды;

1 Дыбыстық сигналдарды танып білуге аналитикалық шолу

1.1 Дыбыстық сигналдарды танып білу тарихы

Дыбыстық сигналдарды тану технологиясы алғаш рет 1950–1960 жылдары пайда болды. Ол кезде жүйелер тек цифрларды тануға қабілетті болатын және бұл бағыттағы зерттеулермен негізінен инженерлік мамандар айналысқан. Алғашқы жүйелердің бірі - Bell Laboratories компаниясы жасаған «Audrey» құрылғысы, ол бір адамның дауысын ғана танитын және тек айтылған сандарды түсіне алатын. Кейінірек, 1962 жылы IBM компаниясы ағылшын тілінде 16 сөзді тани алатын «Shoebox» құрылғысын ұсынды.

Сол кезеңде АҚШ, Жапония, Ұлыбритания және КСРО елдерінде жеке дыбыстарды тануға арналған құрылғылар жасала бастады. Бұл құрылғылар шектеулі функционалдылыққа ие болса да, компьютерлік технологиялардың дамуы үшін маңызды қадам болды.

1970 жылдары дыбысты тану саласы жаңа деңгейге көтерілді, бұл кезеңде АҚШ-тың Қорғаныс министрлігі тарапынан қолдау тапқан DARPA жобасы ерекше рөл атқарды. Carnegie Mellon университеті жасаған «Harpy» жүйесі үш жастағы баланың 1000-нан астам сөздік қорын тани алатын деңгейге жетті. Бұл жүйеде Beam іздеу алгоритмі қолданылып, сөз тізбектерінің ықтимал комбинацияларын тиімді іздеу ұсынылды.

Осы онжылдықта алғашқы коммерциялық жүйелер де пайда болды. Олар әртүрлі дауыс түрлерін тануға және нақты тапсырмаларды орындауға арналған болды. Ал 1980 жылдары жаңа статистикалық тәсіл – жасырын Марков модельдері (HMM) кеңінен қолданыла бастады. Бұл әдіс белгісіз дыбыстық үлгілерді ықтимал сөздермен сәйкестендіруге мүмкіндік берді және ұзақ уақыт бойы тану жүйелерінің негізі болды.

1990 жылдарда процессорлардың қуаты артты және алғашқы көпшілікке қолжетімді дыбыс тану бағдарламалары пайда болды. Dragon Systems компаниясы \$9000 бағасымен алғашқы өнімін ұсынды. 1997 жылы оның жетілдірілген нұсқасы - Dragon NaturallySpeaking - пайдаланушыға минутына 100 сөзге дейін тануға мүмкіндік берді, бірақ әлі де алдын ала оқыту кезеңін қажет етті.

1996 жылы BellSouth компаниясы VAL атты алғашқы интерактивті дауыс порталы жүйесін іске қосты. Ол пайдаланушының дауыстық командалары негізінде әрекет ететін алғашқы жүйе болды. 2000 жылдары сөйлеу тану дәлдігі 80%-ға дейін жетіп, жүйелердің сенімділігі артты.

2000 жылдан кейін бұл технология интернеттің дамуымен бірге кеңейе түсті. Лингвистикалық дерекқорлар мен статистикалық модельдер негізінде сөйлеу тану жүйелері неғұрлым тиімді жұмыс істей бастады. Windows Vista және Mac OS секілді операциялық жүйелерге кіріктірілген тану жүйелері пайдаланушыларға пернетақтасыз әрекет ету мүмкіндігін берді, дегенмен олар кең қолданысқа бірден ие бола қойған жоқ.

Көпшілікке танымал серпіліс Google Voice Search арқылы орын алды. Бұлттық технологияларды пайдалану арқылы Google пайдаланушылардың дауыстық үлгілерін үлкен дерекқормен салыстыра отырып, дәл тануға мүмкіндік жасады. 2010 жылы Android жүйесіне арналған Google Voice Search іске қосылып, смартфондарда сөйлеу арқылы басқару кеңінен тарады.

Осыдан кейін Apple компаниясы Siri-ді ұсынды. Бұл жүйе де бұлттық есептеулерге сүйеніп, пайдаланушы туралы мәліметтер негізінде жеке жауаптар ұсынуға қабілетті болды. Сөйлеу тану жүйелері осылайша күнделікті өмірде, бизнесте және түрлі салаларда кеңінен қолданыс тапты.

1.2 Дыбысты тану әдістеріне шолу жасау

Қазіргі заманғы дыбысты тану жүйелері лексикалық элементтерді сәйкестендіру және сөйлеуді тану мақсатында екі негізгі тәсілді қолданады. Бірінші тәсіл – алдын ала жазылған сөйлеу үлгілерін тану арқылы жұмыс істейтін жүйелер. Мұндай жүйелер, әдетте, белгілі бір дауыстағы, шектеулі командалармен шектелетін қарапайым міндеттерге арналған.

Екінші тәсіл – анағұрлым күрделі және икемді. Ол сөйлеу сигналын фонемалар мен аллофондарға, одан әрі морфемалар мен буындарға, кейін тұтас сөздерге бөліп талдайды. Бұл тәсіл «нақты» немесе толыққанды сөз тану жүйелерінде қолданылады.

Сөйлеу тану жүйелері шартты түрде екі топқа бөлінеді:

- Сөйлеушіге тәуелді жүйелер – алдын ала жеке қолданушының дауысын үйретуді қажет етеді, және басқа қолданушылармен жұмыс істеу үшін қайта оқытуды талап етеді.

- Сөйлеушіден тәуелсіз жүйелер – алдын ала жаттықтыруды қажет етпейді, бірақ дәлдігі төмен болуы мүмкін.

Қазіргі ұялы телефондардың көпшілігі дауыс арқылы басқарылатын функцияларға ие. Мысалы, телефон кітапшасындағы контактін дауыс арқылы табу немесе қоңырау шалу функциялары – бұл жүйелердің қарапайым түрлері. Кейбір құрылғыларда арнайы микрофондармен жабдықталған дауысқа жауап беретін пернетақталар да кездеседі.

Microsoft Windows секілді операциялық жүйелерде орнатылған сөйлеу тану функциялары арқылы бағдарламаларды дауыспен іске қосу, мәзір элементтерін таңдау және жалпы компьютермен өзара әрекет ету мүмкіндігі бар. Бұл жүйелер әсіресе мүмкіндігі шектеулі жандарға көмекші құрал ретінде қолданылады.

Егер құрылғыны (мысалы, телефон немесе пернетақта) нақты бір адамның даусына «үйретсе», ол динамикке тәуелді жүйе ретінде жұмыс істейді. Мұндай жүйелерде дауыс тегтері шектеулі болса да, олар қарапайым қолдану үшін жеткілікті тиімді жұмыс істейді. Сонымен қатар, қазіргі заманғы сөйлеуді тану

жүйелерінің басым бөлігі морфологиялық және лексикалық элементтерді тануға негізделген, неғұрлым жетілдірілген әдістерге сүйенеді.

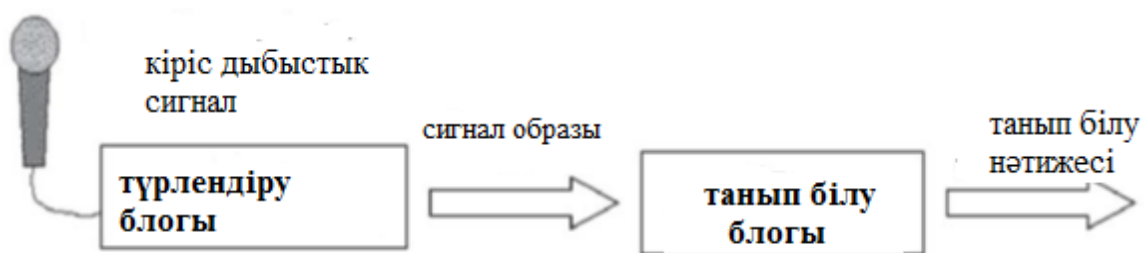
1.3 Дыбысты танудың сұлбасы және негізгі сипаттамалары

Дыбысты тану – бұл көпдеңгейлі үлгіге негізделген тану процесі. Ол дыбыстық сигналдарды фонемалар, буындар, сөздер және сөйлемдер деңгейінде құрылымдап, иерархиялық түрде талдау арқылы жүзеге асады.

Иерархияның әрбір деңгейінде сөйлеу құрылымының төменгі элементтері – мысалы, фонемалар немесе буындар – келесі, жоғары деңгейлі элементтерге біріктіріледі. Мұндай құрылым сөздерді тану дәлдігін арттыруға мүмкіндік береді және тану процесінде уақыт бойынша өзгертін сөйлеу бірліктерін нақты сипаттауға жағдай жасайды.

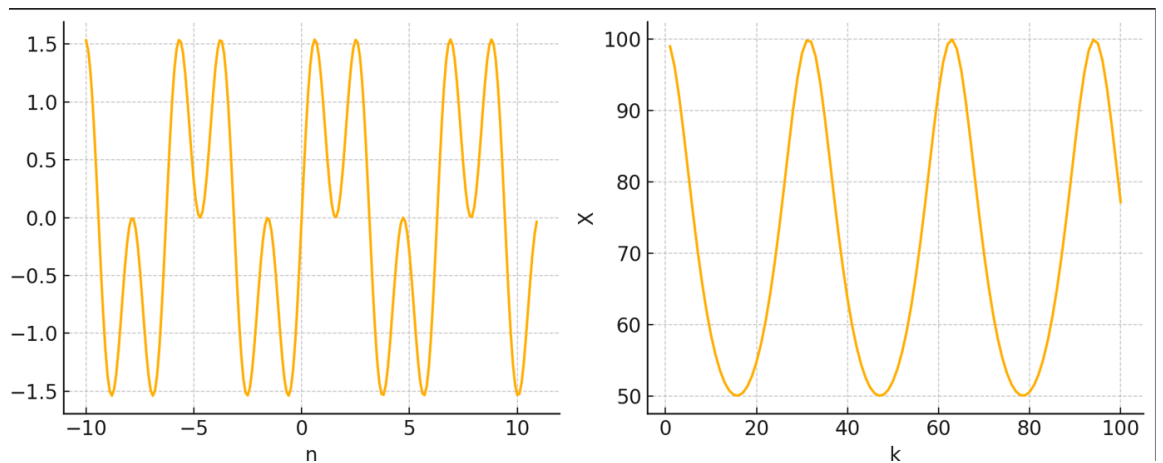
Егер кіріс сигнал туралы априорлық (алдын ала белгілі) ақпарат бар болса, жүйе тану сапасын айтарлықтай жақсарты алады. Бұл ақпарат дыбысты өңдеуде маңызды рөл атқарады.

Жалпылама түрде дыбысты тану жүйесінің құрылымы 1-суретте көрсетілген (әдебиет [2] негізінде).



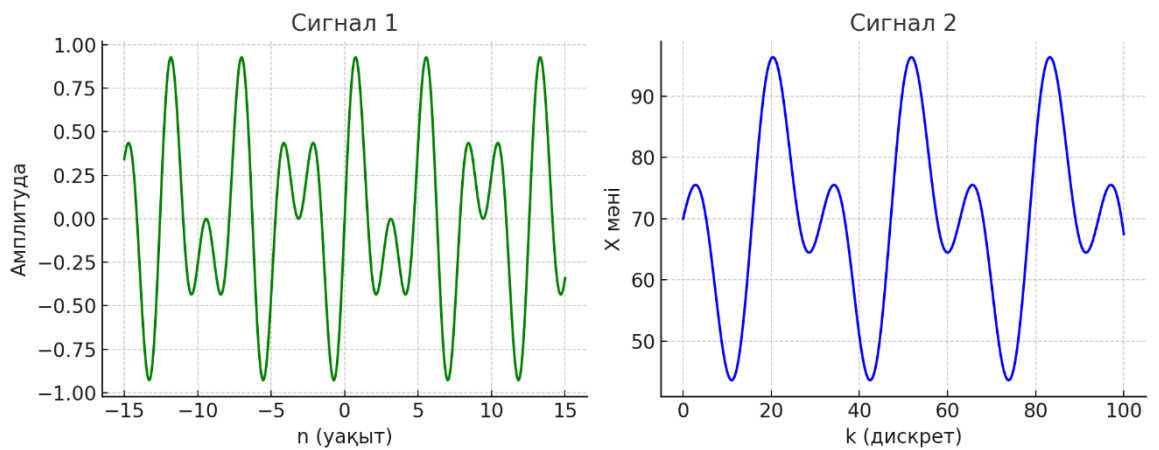
1.1-сурет – Дыбысты тану жүйесін жалпы сұлбасы

Әдетте, 1.2-суретте көрсетілгендей түрінде жазылған сөйлеу сигнал телефон желісі жазуды 8 кГц немесе микрофон 20 кГц жазу) жоғары сынамасын іріктеу [2].



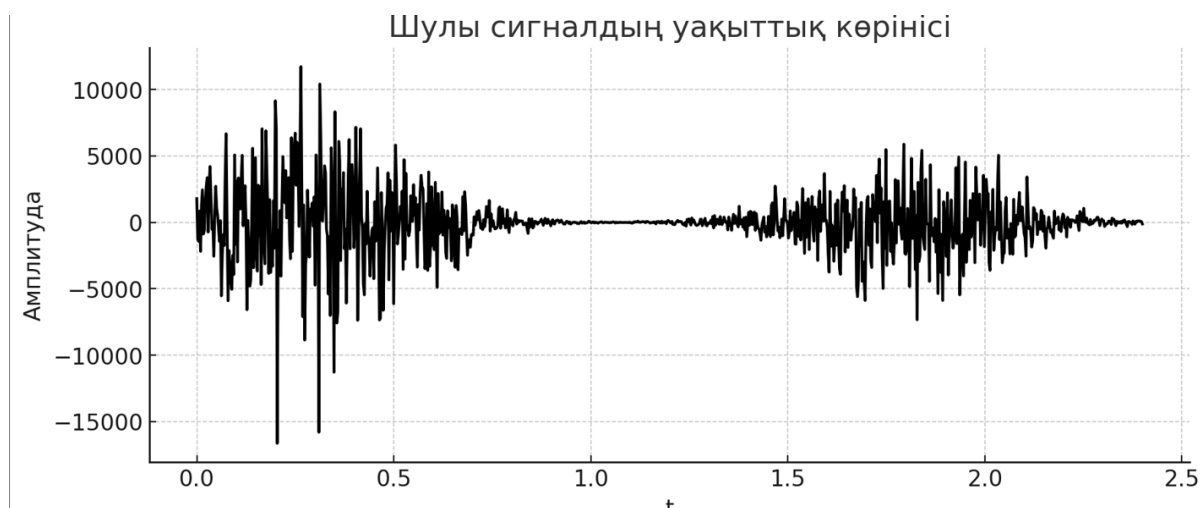
1.2-сурет – Дыбыстық сигнал және оның спектрі

1.3.1 Дыбыстық сигнал уақытша сипаттамалары



1.3-сурет – Ресейлік сөйлеуінің типтік толқынының пішіні

Сөйлеудің маңызды бағыттарын, қайталанатын және анық көрінетін элементтер ретсіз көрінеді. Шу мерзімді және квази - ретсіз болып саналады қайталанатын элементтерімен сөйлеу фрагменті 1.5-суретте көрсетілген.



1.4-сурет – Фраза фрагментінің формасы

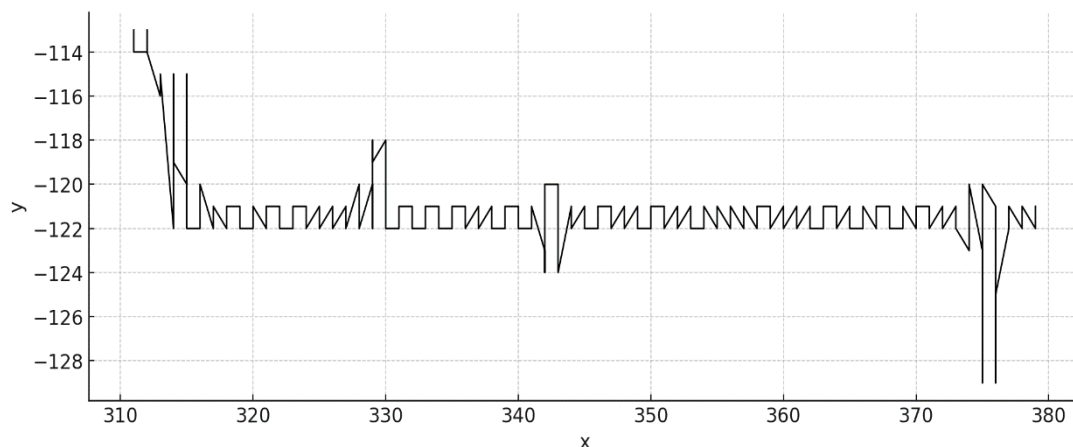
Сөйлеу сигналының құрылымын талдау барысында, дауыссыз дыбыстарда шу мен үзілістер байқалса, ал дауысты дыбыстар белгілі бір квазипериодты импульстік тізбектер түрінде көрініс табады. Бұл ерекшеліктер дауысты дыбыстарға тән фонемалардың акустикалық қасиеттерімен тығыз байланысты.

Фонологиялық зерттеулер көрсеткендей, дауысты дыбыстардың пайда болуы – вокалдық жолдардың, атап айтқанда дауыс сымдарының ретті түрде тербелуінен туындайды. Осы тербелістер ауаның дірілін тудырып, белгілі бір ырғақпен қайталанып отыратын импульстерге айналады.

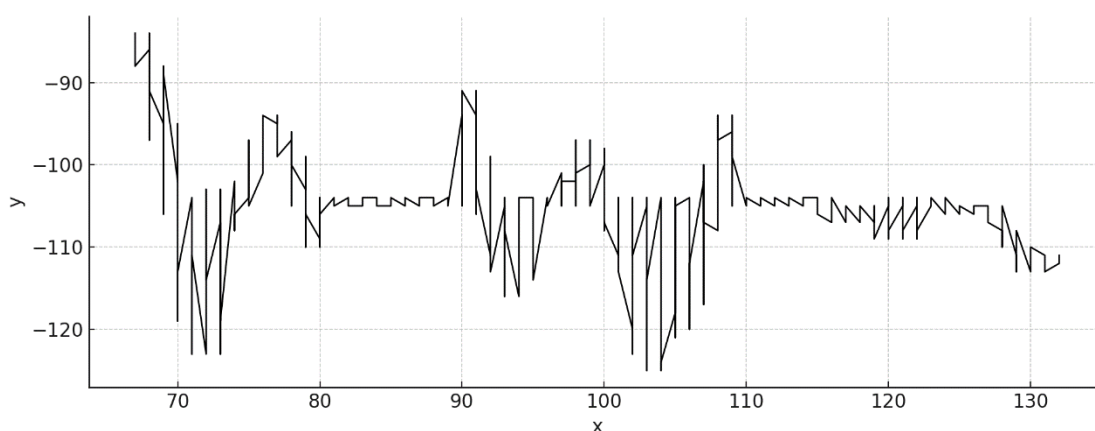
Дауыссыз дыбыстарда, керісінше, сөйлеу ағзаларындағы белгілі бір тарылу аймақтарында ауаның күшпен шығарылуы кезінде турбуленттілік (құйынды ағыс) пайда болады. Бұл кезде дауыс сымдары тербелмейді, оның орнына акустикалық шу пайда болады. Мұндай құбылыстар дауыссыз дыбыстардың спектрлік сипаттамаларын анықтайды.

Сөйлеу сигналдарының түрленуін (осциллограммасын) зерттей отырып, әртүрлі фонемаларды құрайтын вокалдық жол мүшелерінің қатысу деңгейін анықтауға болады. Мысалы, орыс тіліндегі дауысты дыбыстар («а», «о», «у» және т.б.) үшін осындай квазимезгілдік импульстар айқын көрінеді. Оларды 1.6-суреттен көруге болады.

Бұл импульстардың қайталану аралығы – период немесе pitch – фонема жайлы маңызды ақпаратты береді. Сонымен қатар, дыбыстың ұзақтығы және оның сөйлем ішіндегі орналасуы да сөйлеудің мағыналық құрылымына әсер етеді [3].



1.5-сурет – «А» дыбыс сигналдың мысалы

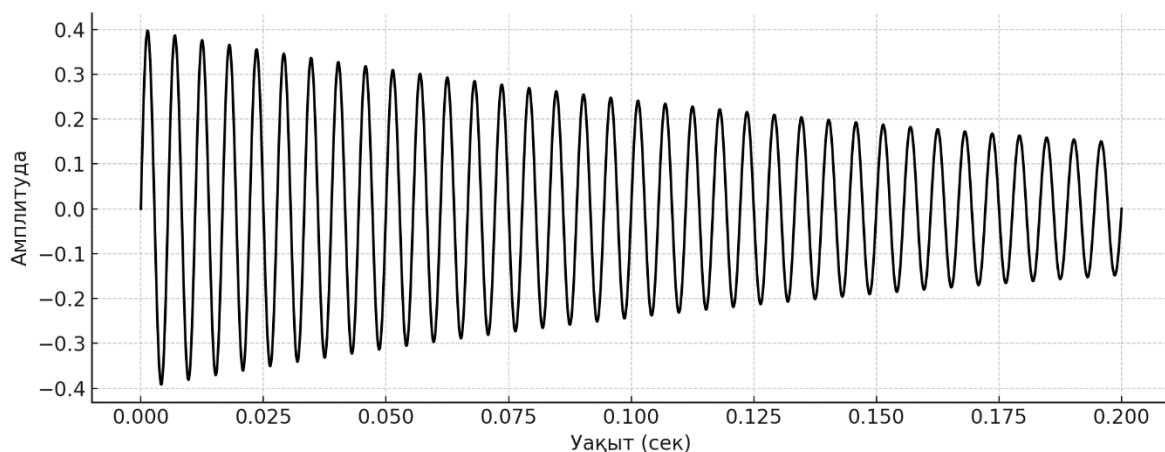


1.6-сурет – Мысал «Р» дыбыс сигналына

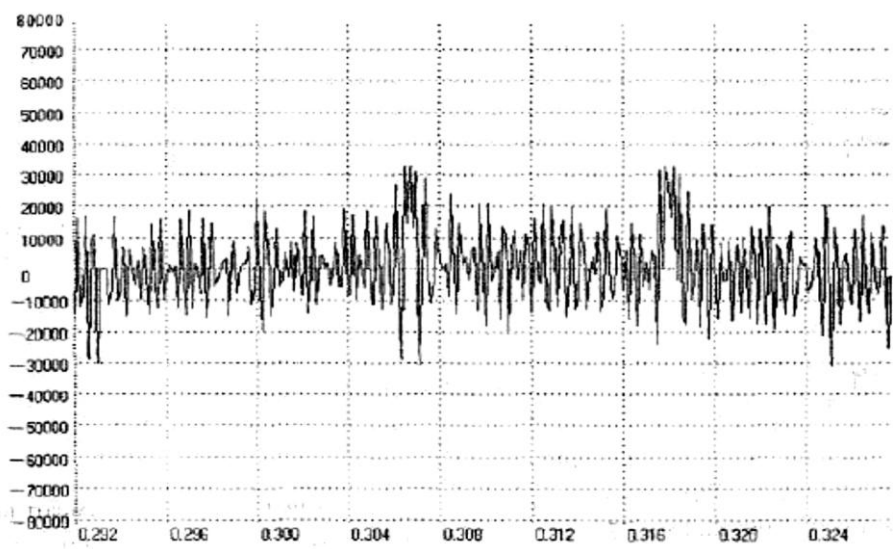
Дыбыстық сигналдардың уақыт доменіндегі құрылымы квазиреттілікпен қатар, ішінара шулы сипатқа ие болуы мүмкін. Мысалы, 1.7-суретте бейнеленген сигнал пішіні «б» сияқты сонант дыбысқа тән екендігін көруге болады. Мұндай фонеманы қалыптастыру кезінде сөйлеу ағзаларының қатысуымен ауа ағыны модуляцияланады және бір мезгілде қысыммен шығарылады.

Ал шулы, саңылаулы фонемалар үшін (мысалы, «ш» дыбысы) сигнал белгілі бір шуыл сегменттерінен тұрады. Бұл сипат 1.8-суретте көрсетілген. Мұндай дыбыстардың осциллограммасында окклюзионды форманттар – яғни, дыбыстың шығу кезінде пайда болатын уақытша үзіліс және энергияның күрт өзгеруі – айқын байқалады [4].

Мысалы, «һ» (немесе кейбір әдебиеттерде "Н") дыбысы кезінде сигналдың жылдам әлсіреуі мен ауытқуы осындай ерекшелікті білдіреді. Бұл құбылыстарды осциллограмма арқылы талдай отырып, дыбыстың пайда болу әдісін және сөйлеу аппаратындағы процестерді түсіндіруге болады.



1.7-сурет – «W» дыбысының сигналдың таза графигі



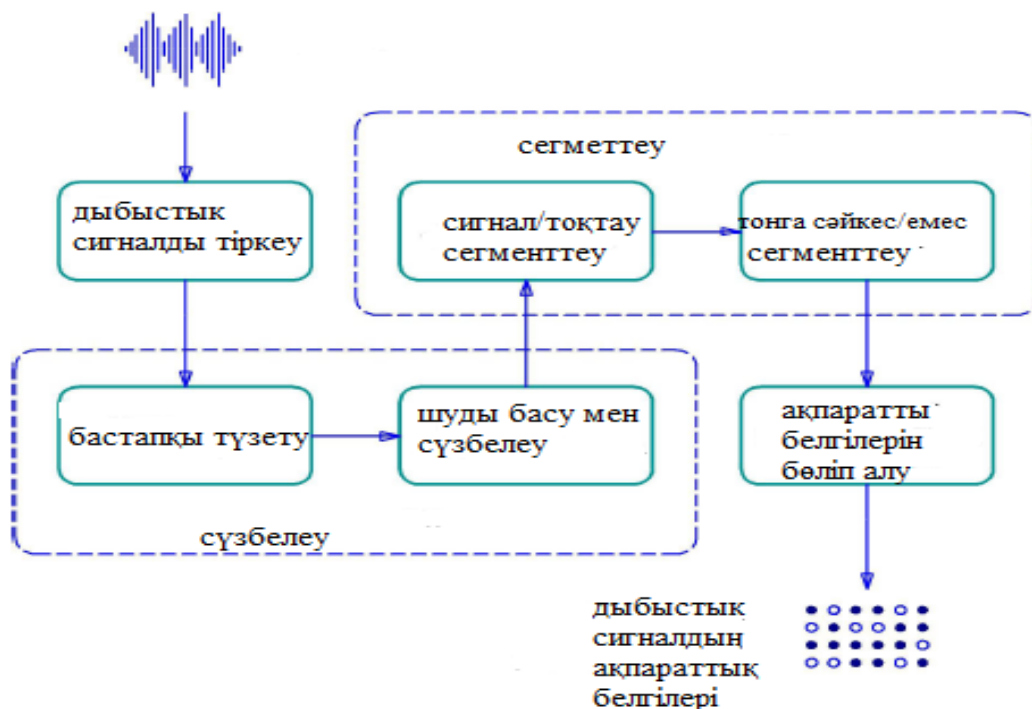
1.8-сурет – «Н » дыбысының сигнал пішінінің мысалы

Аффрикаттар – бұл акустикалық қасиеттері бойынша жарылғыш және саңылаулы дыбыстардың арасындағы аралық сипатқа ие фонемалар. Олар форманттардың ертерек пайда болуымен және ұзақ уақыт бойы сақталатын тегіс амплитудалы шуыл аймағында энергияның шоғырланған пішінімен ерекшеленеді.

Айтылған сөз тіркестері мен сөйлемдерге жүргізілген зерттеулер көрсеткендей, созылыңқы дауыссыз дыбыстар сөйлеу барысында жиі қолданылады. Бұл олардың ақпараттық жүктемесінің жоғары екенін білдіреді [6].

Сондай-ақ, белгілі бір дыбыстың жиілігі нақты бір фонеманың жүзеге асырылуына байланысты екені анықталды. Осы арқылы дыбыстың ақпараттық мазмұнын бағалауға болады. Зерттеулерге сәйкес, ең көп ақпаратты алып

жүретін – үнсіз (саңырау) дауыссыз дыбыстар болып саналады, өйткені оларда мағыналық ажырату мүмкіндігі өте жоғары.



1.9-сурет – Дыбыстық сигналды танып білу құрылымы

1.3.2 Дыбыстық сигнал спектрлік сипаттамалары

Сөйлеуді тануда ақпараттық сөздерді оқшаулау үшін олардың спектрлік көріністерін пайдалану маңызды рөл атқарады. Бұл бағытта қолданылатын негізгі спектрлік талдау әдістеріне Фурье түрлендіруі, сызықтық болжау және кепстральді (cepstral) түрлендірулер жатады.

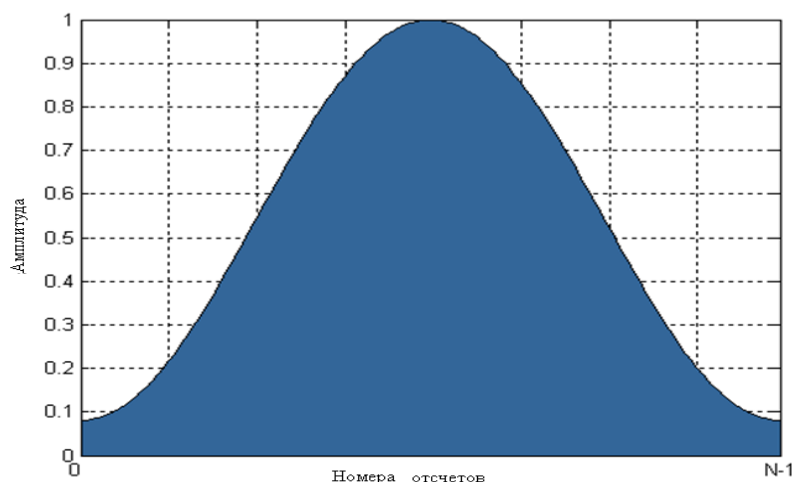
Сөйлеу сигналдарын өңдеуде қолданылатын бұл әдістер алдымен сигналды алу мен оны өңдеу үшін қажетті сандық сүзгілер мен алгоритмдердің жиынтығына сүйенеді. Адамның вокалдық жолдары тұрақты емес болғандықтан, сөйлеу сигналы да үнемі өзгеріп отырады, яғни ол тұрақсыз сипатта болады. Алайда, қазіргі зерттеулерге сәйкес, ұзақтығы 100 миллисекундқа дейінгі уақыт аралықтары квазистационарлық деп қарастырылады. Бұл кезеңдерде сөйлеу сигналын белгілі бір тұрақты күйде сипаттауға болады.

Басқаша айтқанда, сөйлеу сигналын талдауда кішкене уақыт терезелері (уақытша үзінділер) қолданылады. Осы терезелер ішінде сигналды сипаттайтын параметрлер – мысалы, Фурье спектрі, кепстраль коэффициенттері және вейвлет коэффициенттері – үздіксіз есептеліп отырады.

Бұл қысқа уақыт терезелер сигналдың бүкіл ұзақтығын қамтымайды, сондықтан сигналды шекара аймақтарында жоғалтпау үшін арнайы салмақтау

функциялары (weighting functions) пайдаланылады. Бұл функциялар «терезе» деп аталып, сигналды түрлендіру кезінде дәлдікті арттырады [7].

Сандық сигналдарды өңдеу саласында түрлі терезелер кеңінен қолданылады. Оларға тікбұрышты, Хемминг, Блэкман, Бартлетт және Кайзер терезелері жатады. Қазіргі кезде Хемминг терезесі сөйлеуді тану жүйелерінде ең жиі қолданылатын және тиімді нұсқалардың бірі болып саналады.



1.10-сурет – Хемминг функциясының спектрі

Сөйлеу сигналын тиімді талдау үшін уақытша терезе қолдану маңызды рөл атқарады. Бұл терезе сигналдың орталық бөлігінен үлгілер алу арқылы әртүрлі параметрлік шамаларды есептеуге мүмкіндік береді. Терезе түрі мен өлшемі спектрлік талдаудың дәлдігіне әсер етеді, әсіресе оның жиіліктік сипаттамалары мен негізгі бөлік ені маңызды мәнге ие.

Сигнал талдауы жақтау (кадр) негізінде жүргізіледі, мұндағы әрбір кадр белгілі бір уақыт аралығын қамтиды. Жақтау ұзақтығы – кадр ішіндегі сигнал ұзақтығын секундпен өлшейтін параметр. Ал жақтау аралығы – кадрлар арасындағы уақыттық қашықтықты білдіреді. Практикада сөйлеу тану жүйелері үшін бұл мәндер әдетте 10–20 миллисекунд аралығында таңдалады [8,9].

Бұл аралық сөйлеу сигналының жиілік спектріндегі өзгерістер мен жүйенің есептеу күрделілігі арасында теңгерімді қамтамасыз етеді. Таңдалатын кадр ұзақтығы көбінесе сөйлеу жылдамдығына және дауыс өзгерісінің қарқынына тәуелді болады.

Жақтаулардың негізгі сипаттарының бірі – терезе мөлшері (window size), яғни әрбір кадрда қанша үлгі есепке алынатыны. Терезе өлшемі уақыт бірлігімен (секунд немесе миллисекунд) көрсетіледі және есептеу кезінде қолданылатын орташа мәнді анықтауға мүмкіндік береді.

$$O = \frac{T_w - T_f}{T_w} \times 100\% \quad (1.2)$$

мұндағы T_w – секундтан;

T_f – кадрдың ұзақтығы.

$$T_w < T_f \quad (1.2)$$

Егер жақтаулар арасында қабаттасу болмаса, бұл жағдайда қабаттасу деңгейі нөлге тең болады. Мысалы, жақтау ұзақтығы 20 мс, ал терезе ұзақтығы 30 мс болса, бұл шамамен 33% қабаттасуға сәйкес келеді. Қазіргі заманғы сөйлеуді тану жүйелерінде 66%-ға дейін қабаттасу қолданылатын жағдайлар кездеседі.

Қабаттасу деңгейін арттырудың негізгі мақсаты – стационарлық емес шудың әсерін азайту және терезелерді шекараларында пайда болатын артефактілерді шектеу. Бұл тәсіл есептелетін параметрлердің нақтылығын арттырып, қосымша шудың ықпалын азайтуға септігін тигізеді.

Дегенмен, шамадан тыс тегістеу кейбір жағдайларда сөйлеу сигналының маңызды динамикалық өзгерістерін жасыруы мүмкін. Сондықтан жақтау мен терезе параметрлерін оңтайлы таңдау – жүйенің сенімділігі мен дәлдігін қамтамасыз етудің маңызды факторы.

Сөйлеу сигналдарын талдауда кең қолданылатын құралдардың бірі – сандық сүзгі банкісі. Бұл құрылым адам есту жүйесінің бастапқы деңгейдегі қабылдау механизмін имитациялайды. Сүзгі банкі әртүрлі жиіліктерді психоакустикалық шкалаға сәйкес бөліп, сөйлеу сигналының жиіліктік құрамын нақты сипаттауға мүмкіндік береді.

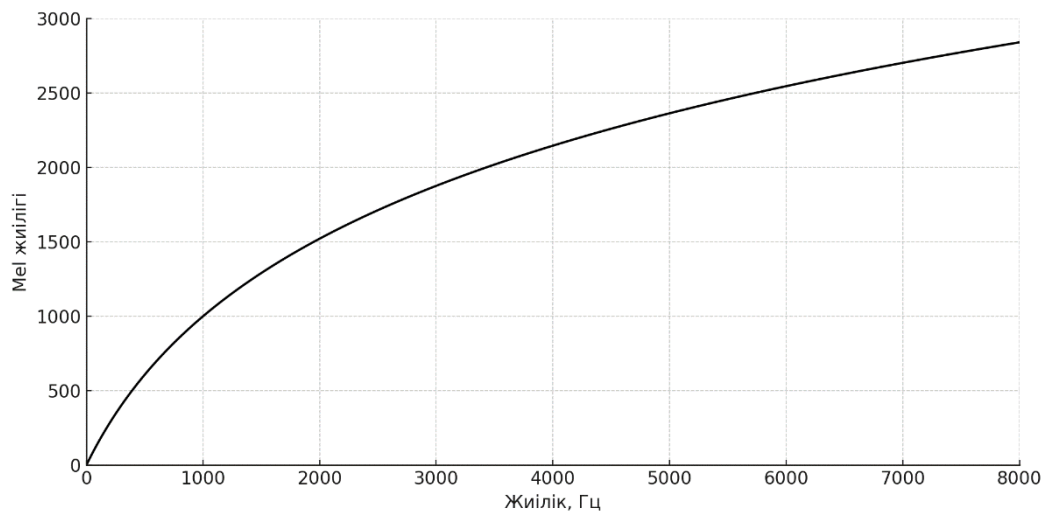
Бұл шкалада нақты жиілік мәндері адамның субъективті қабылдауына сәйкестендірілген, нәтижесінде сөйлеу тану сапасы айтарлықтай артады.

$$Bark = 13 \arctg\left(\frac{0,76f}{1000}\right) + 3,5 \arctg\left(\frac{f}{7500}\right)^2 \quad (1.3)$$

Дыбысты тану бөлу осы түріне көп танымал тәсіл Mel – шкаласы ретінде белгілі:

$$Mel = 2595 \lg\left(1 + \frac{f}{700}\right) \quad (1.4)$$

Mel ауқымы қабылданады жиілігі тоналдық шкаласын картаға тырысады. 0 жиілік диапазонында 1000 Гц сызықтық масштаб, 1000 Гц кейін - логарифмдік (1.11-сурет).



1.11-сурет – Mel ауқымды

Төмендегідей әрбір сүзгі жолақ енін білдіру болып табылады :

$$bw = 25 + 75[1 + 1/4(f/1000)^2]^{-0.69} \quad (1.5)$$

Сөйлеу сигналын талдауда жиілік шкаласын түрлендіру маңызды рөл атқарады. Бұл түрлендіру көбінесе Bark немесе Mel шкалалары негізінде жүзеге асырылады. Бұл шкалалар адамның есту қабылдауының психоакустикалық ерекшеліктерін ескере отырып, жиілікті логарифмдік түрде бейнелеуге мүмкіндік береді.

Bark және Mel шкалаларын кәдімгі сызықтық жиілік ауқымын адамның сезімтал қабылдау шкаласына түрлендіретін әдістер ретінде қарастыруға болады. Осы шкалалар бойынша жасалған фильтрлік банктер (filter banks) сөйлеу сигналының жиіліктік мазмұнын тиімді бөлшектеуге жағдай жасайды.

Бұл жағдайда сызықты фазаға ие жиілікке тәуелді фильтрлер (мысалы, ХҚФ – жиілікке қатысты фильтрлер) Bark немесе Mel шкаласының осіне тең қашықтықта орналастырылады. Әрбір фильтрдің өткізу қабілеттілігі арнайы формулаларға (мысалы, (1.3), (1.4), немесе (1.5)) сәйкес орталық жиілікке байланысты есептеледі.

Сөйлеу сигналын спектрлік түрлендірудің кең таралған әдісі – қысқа уақытты Фурье түрлендіруі (STFT). Бұл әдіс сөйлеу сигналын қысқа интервалдарға бөліп, әр интервал үшін дискретті Фурье түрлендіруін (DFT) қолдану арқылы жиіліктік мазмұнын анықтайды.

Көптеген сөйлеуді тану жүйелері спектрді есептеу алдында Хемминг терезесін қолданады. Бұл терезе сигналдың шеттерінде болатын кілттік мәндердің төмендеуін қамтамасыз етіп, спектрдің бұрмалануын азайтады. Нәтижесінде алынған спектр сигналдың уақыт бойынша өзгерісін дәлірек сипаттай алады [5,6].

$$4y_n = \sum_{k=0}^{N-1} x_k e^{-i\frac{2\pi}{N}kn} \quad (1.6)$$

Дыбыстық сигнал спектрлік талдауға дейін Хемминг терезесімен өңделіп, терезе ішіндегі N үлгі таңдалады.

мұндағы: N – терезе ішіндегі үлгілер саны,

K – жалпы сигнал үлгілерінің саны,

Y – спектрде есептелетін жиіліктер саны.

Дыбысты тану жүйелерінде алынған күрделі спектрдің амплитудасы ғана маңызды болып табылады, яғни сигналдың фаза компоненті жиі еленбейді. Бұл спектр Фурье түрлендіруінен кейін пайда болады, бірақ ол Фурье сүзгі банкінің тікелей нәтижесі ретінде қарастырылмайды.

Дискретті спектрді есептеу үшін жиі Кули–Тьюки алгоритмі негізінде құрылған жылдам Фурье түрлендіру (FFT) әдісі қолданылады. Бұл алгоритм есептеу уақытын қысқартуға және тиімділікті арттыруға мүмкіндік береді.

Алынған спектр формула (1.6) түрінде беріліп, оны матрицалық түрлендіру арқылы көрсетуге болады. Бұл әдіс дыбыстық сигналдың жиіліктік мазмұнын нақты және оңтайлы сипаттайды.

$$Y = \frac{T}{N} W X \quad (1.7)$$

онда Y , X - баған x және y мәндерін тиісті векторлары және W - мынадай нысанда матрицалық түрі:

$$W = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & q & q^2 & \dots & q^{N-1} \\ 1 & q^2 & q^4 & \dots & q^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & q^{N-1} & q^{2(N-1)} & \dots & q^{2(N-1)^2} \end{pmatrix} \quad (1.8)$$

мұндағы:

$$q = e^{-i\frac{2\pi}{N}} \quad (1.9)$$

Егер $N = r_1 r_2 \dots r_n$, қайда r_k – бүтін, матрицалық W әрбір матрицалық элементтер 0 екенін $N \times N$ осындай N матрицасы өнім ретінде ұсынылған. Трансформация (1.6) ретінде жазылған :

$$Y = \frac{T}{N} W_n W_{n-1} \dots W_2 W_1 X \quad (1.10)$$

Кули–Тьюки алгоритміне негізделген жылдам Фурье түрлендіру (FFT) күрделі матрицалық көбейтуді тиімді жүзеге асыру үшін нөлдермен көбейту операцияларын қысқартады және тек РКН (рекурсивті көбейту және қосу) операцияларын қолданады. Бұл тәсіл есептеу жүктемесін азайтуға және жылдамдықты арттыруға мүмкіндік береді.

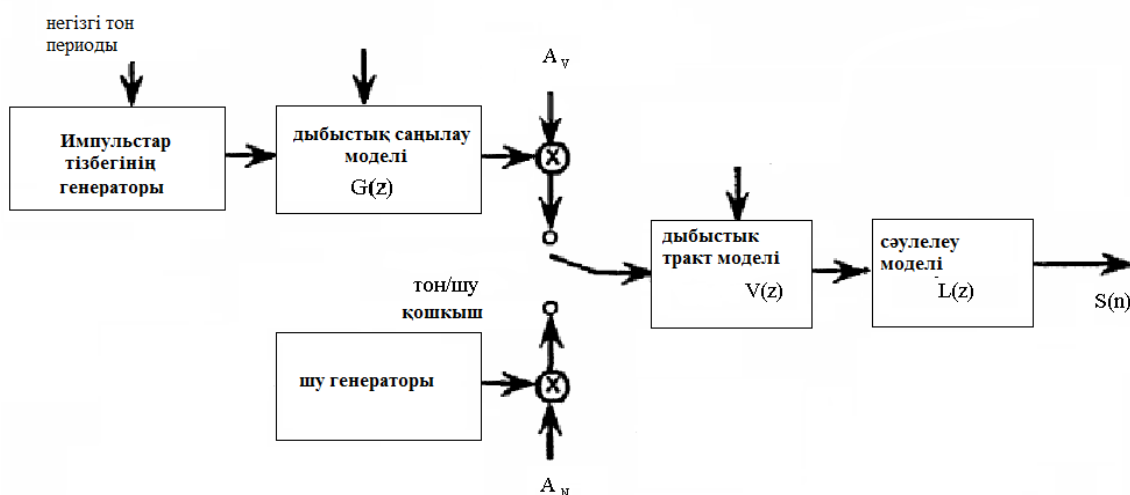
Сөйлеуді тану жүйелері сөйлеу сигналының адам вокалдық жолдарының үлгісі ретінде қалыптасатынына негізделеді. Яғни, адам дауысының пайда болу моделі – бұл сөйлеу аппаратындағы вокалдық-ішек жолдарының динамикасын сипаттайтын жүйе деп қарастырылады. Бұл жүйе арқылы алынған сигналды талдау арқылы тану жүзеге асырылады [6-9].

Сөйлеу сигналын сандық сипаттау үшін екі негізгі әдіс қолданылады:

- Сызықтық болжау әдісі (Linear Predictive Coding, LPC) – сөйлеу сигналының келесі мәндерін алдыңғы мәндерге негізделе отырып болжауға мүмкіндік береді.

- Кепстральдық түрлендіру (cepstral transformation) – сигналдың спектрлік құрылымын сипаттайтын тиімді параметрлерді анықтауға бағытталған әдіс.

Кепстральдық түрлендіру объективті модельді нақтылау үшін қолданылады, яғни сөйлеу сигналының жеке параметрлерін дәлірек анықтауға мүмкіндік береді. Бұл әдістер 1.12-суретте көрініс табатын сөйлеу өңдеу жүйесінің құрылымдық сұлбасы шеңберінде қолданылады.



1.12-сурет – Дыбысты сигналдың өңделуінің жалпылама құрлымы

Моделін Z - трансформация тұрғысынан [6] ретінде жазуға болады. Cepstral талдауда қолданылатын сөйлеу сигналының гомоморфты моделі сөйлеу сигналын сипаттау үшін жиі гомоморфты жүйелер теориясы қолданылады. Бұл теория суперпозиция принципіне бағынатын жүйелер класына негізделген және сөйлеу сигналдарын cepstral түрлендіру арқылы талдауға мүмкіндік береді.

Сөйлеу сигналы келесі түрдегі модельдік теңдеумен сипатталады:

$$S(z) = E(z)G(z)V(z)L(z) \quad (1.11)$$

мұндағы $E(Z)$ – жүйенің жалпы аударым (беру) функциясы, яғни сөйлеу сигналының жиіліктік сипаттамасын сипаттайды;

$G(Z)$ – қозу сигналының беру функциясы, яғни пек жиілігіне байланысты дыбысты және саңырау дыбыстардың импульстік немесе шулы дәйектілігі;

$V(Z)$ – вокалдық жолдардың жиіліктік сипаттамасы (адам дауысының резонанстық құрылымы);

$L(Z)$ – еріндердің радиациялық қасиеттерін сипаттайтын модельдік функция.

Егер $G(Z)$ және $V(Z)$ функцияларын біріктіріп бір $H(Z)$ функциясы ретінде қарастырсақ, онда сөйлеудің толық өндіріс моделі (1.12) сипатталады.

Бұл түрлендіру cepstrum талдауға негіз болады, өйткені логарифмдік доменде бұл функциялар қосындыға айналып, компоненттерді бөлек қарастыруға мүмкіндік береді. Осылайша, сөйлеу сигналының әрбір бөлігі (қозу, резонанс, радиация) жеке талданып, дәл сипаттама алуға болады.

$$S(z) = E(z)H(z) \quad (1.12)$$

Cepstral талдау гомоморфные жүйелерін түрінде трансформация (1.14) идеясына негізделген. Бұл жүйе трансформация D қатысты мынадай қасиеттерге ие болуы тиіс:

$$D[E(z)H(z)] = D[E(z)] + D[H(z)] \quad (1.13)$$

Мұндай мүлікті біз жазуға болады, ол үшін пайдалануға логарифм, жатады:

$$\log(S(z)) = \log(E(z)H(z)) = \log(E(z)) + \log(H(z)) \quad (1.14)$$

Вокалдық жолдарының аудару функциясының нерв функциясы: Осылайша, сөйлеу сигнал екі құрамдас бөлінген болуы мүмкін. (1.14) айырбастау сигнал компонентті шығарып алу үшін деп болжайды $z = e^{i\omega}$ Жиілік аудару функцияларын білдіреді. Содан кейін сигнал $S(N)$ жиілігі өкілдігі дискретті Фурье түрлендіру пайдалану және сигнал құрамдас формула сияқты (1.14) олардың Фурьесі болып көрінуі мүмкін. Осылайша, уақыт домендегі сигналдарды ажырату үшін Фурье Фурье логарифмін айналдыру кері есептеуді қажетті бастапқы сөйлеу сигналына түрлендіруге болады. Сондықтан [7] дыбысты сигнал кешенді cepstrum шықты. Нақты cepstrum, яғни берілген сөйлеу тану қосымшалар жеткілікті ақпарат алу үшін Фурье бастапқы сөйлеу сигнал Фурье модуль логарифмін кері айналдыру:

$$c_n = \frac{1}{N} \sum_{k=0}^{N-1} \lg|S(k)| e^{i\frac{2\pi}{N}kn} \quad (1.15)$$

мұндағы $S(k)$ – сөйлеу сигнал Фурье спектрі. Numbers - Cepstral коэффициенттер (MFCC) деп аталатын болды. Қазіргі уақытта, кең таралған тану жүйелер мынадай формула қайталану арқылы сөйлеу сызықтық болжау коэффициенттері тікелей алынған cepstral коэффициенттерді алды:

$$c_n = a_n + \frac{1}{n} \sum_{k=0}^{n-1} k c_k a_{n-k} \quad (1.16)$$

Cepstral коэффициенттері сөйлеу динамикалық сипаттамалары туралы кез келген ақпарат бермейді. Алайда, бұл ақпаратты, сондай-ақ сөйлеу тануы өте маңызды болуы мүмкін, сондықтан жиі уақыт cepstral коэффициенттеріне қатысты туындыны талдау үшін деректерді қамтиды. Коэффициенттері, екінші ретті туынды Delta - дельта коэффициенттері Бірінші ретті туынды Delta деп аталады. Зерттеулер бастауыш жуықтау түрі жақсы нәтиже бермейді, деп көрсетеді.

$$\Delta C_n = \frac{\sum_{k=-l}^l k \cdot c_{n+k}}{\sum_{k=-l}^l |k|} \quad (1.17)$$

$$\Delta\Delta C_n = \frac{\sum_{k=-l}^l k^2 \cdot c_{n+k}}{\sum_{k=-l}^l k^2} \quad (1.18)$$

мұнда келтірілген L- тұрақты әдетте 3. Стационарлық сигналдарды үздіксіз талдау аса маңызды құралдарының барлық әдістерін үздіксіз дүркін Фурье (STFT) түрлендіруі болып табылады.

1.3.3 Дыбыстық сигналдың статистикалық сипаттамалары

Сөйлеу өндірістік сөйлеу акустикалық теориясына сәйкес, вокалдық жолдарының резонанстық қуыстарын өзара нәтижесінде ұсынуға болады, дауысты және турбуленттік шу тықырды қалыптастыруға фермасы импульсы жатады.

Акустикалық теорияға сәйкес сөйлеу аузынан шығу нәтижесінде алынған сигнал $y(N)$ функциялары сигнал $x(n)$ және жинақталған және вокалдық жолдарының сағ(N) импульсті жауап ретінде ұсынылуы мүмкін:

$$y(n) = x(n) \cdot h(n) \quad (1.19)$$

Онда мұндай ағымдағы модуль спектр, форматтардың, кезеңі және т.б. ретінде параметрлерін, бірқатар сипатталады . негізгі түрлері және кестеде көрсетілген статистикалық сипаттамалары *reschi* өлшеу әдістері.

Кесте 1.1 – Дыбыстық сигналдың статистикалық сипаттамалары

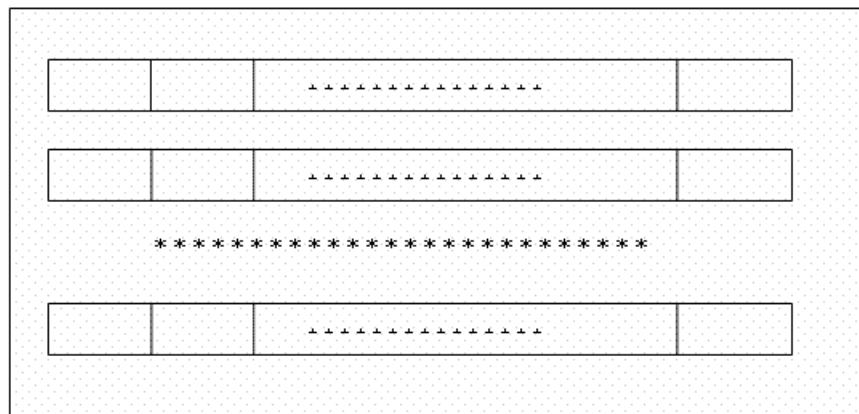
Мінездеме	Өлшеу параметрі	Басты нәтиже
Спекторлық бөлу	Дыбыстық қысым жиілік жолақтарында	Жиіліктер жолағы 5-10 кГц; спектр төмендеуіне қарай жоғарғы жиілік
Форманттық бөлу	Траекториясы формант бойынша видеограмм	1-ші, 2-ші, 3-ші форманты алады диапазоны, кГц: (0,15-0,9; 0,55- 2,8; 1,5-3,4); саңырау (1-3,5; 2,5-5) дауыссыз
Жиілігі тарату тоннасы	Сөйлеу сигнал ең биік шыңдарында арасындағы аралықтар	Жиілік диапазоны, Гц : ер дауыстары - (56-240) , әйелдер - (140-450)
Тарату аумақтарында және сөйлеу кідіртіледі	Ұзақтығы бөлімдер және тоқтата	Орташа ұзындығы: сөйлеу бағыттары - 0,3s , уақытша тоқтатады - 0.18 , жиілігі - 2 Pause / c

1.4 Дыбысты тану үшін пайдаланатын классикалық модельдер

Дыбысты тану үшін пайдаланатын классикалық модельдер ретінде Марков модельдерін қарастыруға болады. Жасырын Марков моделі (НММ) жүйелі параметр M құндылықтарының бірі болуы мүмкін, онда әрбір, N мемлекеттердің тұратын үлгісі болып табылады. i -ші жылғы J -ші мемлекет көшу ықтималдығы - ықтималдықтар $A = MOC$, MOC матрица айқындалған мемлекеттер арасындағы өтпелі ықтималдығы. J -ші мемлекетте параметрдің K -ші құнының ықтималдығы - N әрбір M параметр құндылықтарды әрбір ықтималдығы онда $VJ(K)$, векторлық $V = \{VJ(K)\}$ берген делінген. бастапқы уақытта жүйесі I -ші мемлекет болады ықтималдығы - бастапқы мемлекеттің ықтималдығы π векторының $\pi = \pi_i$, анықталады.

Осылайша, жасырын Марков моделі $= \{A, V, \pi\}$ үштік λ болып табылады. Екі жақындату негізінде дауысты тану үшін Жасырын Марков үлгілерді пайдалана отырып:

- Ол тұрақты болып саналады, әрбір сегментінің ішінде НММ сөйлеу параметрлерін мемлекеттерге тиісті үзінділерді бөлуге болады;
- Әрбір фрагменті ықтималдығы ғана жүйесінің ағымдағы жай-күйі туралы тәуелді және алдыңғы мемлекеттердің байланысты емес болады.



1.13-сурет – Құпия сөздік

Жасырын Марков модельдерін (НММ) қолдану кезінде, жүйені орнату кезеңінде алдын ала белгілі сөздік қор пайдаланылады. Бұл кезде Баум–Уэлч алгоритмі қолданылады. Ол бақылаулар тізбегіне сәйкес келетін ықтимал параметрлерді (яғни, A және B матрицаларын) оңтайлы түрде таңдауға мүмкіндік береді. Бұл матрицалар модельдің күйлік ауысу ықтималдықтарын (A) және бақылау ықтималдықтарын (B) сипаттайды [9].

Сөйлеуді тану кезінде дыбыс сигналы бірнеше сегменттерге бөлінеді. Әрбір сегментке белгілі бір модель сәйкестендіріліп, ықтималдықтар тура (forward) немесе кері (backward) алгоритмдер арқылы есептеледі. Егер белгілі бір сөзге сәйкес келу ықтималдығы белгіленген шекті мәннен асып кетсе, бұл сөз дұрыс танылған деп есептеледі.

Баум–Уэлч алгоритмі – бұл бақылау тізбегіне негізделген ықтималдықты ең жоғарғы деңгейге жеткізетін Жасырын Марков моделінің параметрлерін бағалауға арналған итерациялық әдіс. Алгоритм оқу деректері негізінде модельді тиімді жаттықтыруға мүмкіндік береді.

Кіріс ретінде акустикалық ерекшеліктер (AC – Acoustic Characteristics) пайдаланылады. Олар сөйлеу сигналының белгілі бір уақыт аралығындағы сипаттамаларын көрсетеді және НММ модельдері үшін бастапқы дерек болып саналады.

$$\xi t(i, j) = P(q_t = S_i, q_{t+1} = S_j | O, \lambda) \quad (1.20)$$

онда мемлекеттер S_i және S_j болады, тиісінше есе t мен $T + 1$, кем бақылау O жүйесінің берілген ретпен ықтималдығын көрсетеді.

Жиілік диапазоны 70-7000 Гц болып табылады. Сілтеме дыбыстық қысым P_0 ретінде дыбыс деңгейін бағалау үшін, яғни дыбыс естілетін болып 1 кГц, жиілігінде оның ең төменгі мәні таңдалады $P_0 = 2 \times 10^{-5} \text{ Н / м}^2$. дыбыс қысымының деңгейі беріледі

$$L = 20 \lg \frac{P}{P_0} \quad (Дб)$$

мұндағы P - дыбыстық қысымның мәні.

Динамикалық диапазон астында көп және ең аз сигнал деңгейлері арасындағы айырмашылықты түсіну керек. Сөздің динамикалық диапазоны 35-45 дБ. Сандық жазу жүйесі сандық түрге аналогтық дауыс сигналының берілуін талап етеді [7].

Нәтижесінде, аналогтық-цифрлық түрлендіру (ADC), үздіксіз сигнал дискретті үлгілерін сериясы айналады әр белгілі бір дәлдікпен сол сәтте аналогтық сигнал сипаттайтын бүтін сан болып табылады. Нақты алынған сандар ауқымын еніне байланысты өлшенеді, демек биттік АЦП туралы болады. Сигнал үлгілерін ыдырау процесі таңдама деп аталады. Секундына үлгілердің саны дискретизация жиілігі деп аталады. Түрлендірілген аналогтық сигнал кем дегенде екі есе барынша жиілігі болуы тиіс дәл осы теорема сәйкес таңдама жиілігін vD береді. Сигнал спектрі цифрлық нысанда үшін АЦТ фактісі мерзімді болып табылады. Сигнал спектрі - сигнал жиілігі компоненттерін қарқындылығын және таратуды сипаттайды оның дисплей түрінде болып табылады.

Аналогтық-сандық түрлендіруден (АЦТ) кейін алынған сигнал төмен жиілікті компоненттерге ие болады, олар жиілік спектрінің көрінетін бөлігін құрайды. Сонымен қатар, сигналда жоғары жиілікті компоненттер де болуы мүмкін, олар қайталаушы (еселік) жиіліктер түрінде, яғни негізгі жиіліктің (vD) еселенген мәндері маңында бүйірлік жолақтар ретінде байқалады.

Егер дискреттеу жиілігі жеткілікті түрде жоғары болмаса, aliasing деп аталатын құбылыс орын алады. Бұл кезде жоғары жиілікті құрамдас бөліктер төмен жиіліктерге «жабылып», спектрдің бұрмалануына әкеледі. Мұндай бұрмалауларды болдырмау үшін, АЦТ-ға дейін төмен жиілікті сүзгі (anti-aliasing фильтр) қолданылады. Іс жүзінде сүзгілердің өткізу жолағында тік шекара жасау мүмкін емес болғандықтан, дискреттеу жиілігі vD сөйлеу сигналының жоғарғы жиілігінен екі есе жоғары мәнде таңдалады, мысалы, $vD \approx 22,05$ кГц.

Кванттау кезінде сигналдың нақты мәні дөңгелектеніп, сандық мәнге айналады. Бұл процесс өлшеудің дәлдігіне әсер етеді. Егер сандық аудио жүйенің динамикалық диапазоны D белгілі болса, қажетті кванттау биттерінің саны мына формуламен анықталады: $D=6N+1.8D = 6N + 1.8D=6N+1.8$ мұндағы: D – динамикалық диапазон (децибелмен), N – кванттау биттерінің саны.

Бұл есептеуге сәйкес, дауысты сөйлеуді жазу үшін кемінде 8 бит қолданылуы қажет. Сөйлеу сигналының сандық көрінісі келесі түрде ұсынылады: Фрейм X (ұзындығы N) – сөйлеу сигналының $s_1, \dots, s_2, \dots, s_n$ түріндегі ақырғы тізбегі. Фразаның жүзеге асуы L фреймнен тұрады:

Мысалы, $N = 256$ үлгімен, дискреттеу жиілігі $\nu D = 22,05$ кГц болған жағдайда, әрбір фрейм шамамен 11,6 мс ұзақтыққа сәйкес келеді.

1.5 Дыбыстық сигналды цифрлік сигналға түрлендіру

Сөйлеу сигналын сандық түрде кодтау әдістері екі негізгі топқа бөлінеді: Сигнал пішінін кодтау әдістері (Waveform Coding): Бұл тәсілде аналогтық дауыс сигналы уақыт бойынша дискретті үлгілерге бөлініп, әр үлгінің амплитудасы ең жақын сан мәніне дейін дөңгелектеліп цифрланады. Осылайша, дауыс сигналының бастапқы пішіні барынша нақты сақталып, сандық форматқа түрлендіріледі. Бұл әдіс тікелей өлшеуге негізделеді және классикалық PCM, ADM, CVSD секілді тәсілдер осы топқа жатады [10].

Сөйлеу моделіне негізделген кодтау әдістері (Model-Based Coding): Бұл әдіс адамның сөйлеу аппараты мен дыбыс пайда болу процестерін модельдеуге негізделген. Мұнда сигналдың өзі емес, оның негізгі сипаттамалары (мысалы, дауыс пекі, формант жиіліктері) өлшеніп, сөйлеу сигналының параметрлік моделі құрылады. Алынған параметрлер қабылдаушы жаққа жіберіліп, ол жерде сөйлеу синтезі арқылы бастапқыға ұқсас сигнал қайтадан қалыптастырылады.

Бұл тәсілде нақты өлшеу нәтижесінде алынған параметрлер дауыстың мәнерін сақтай отырып, деректерді айтарлықтай қысуға мүмкіндік береді. Сондықтан мұндай әдістер жоғары тиімділікке ие және жиі вокодерлерде, телекоммуникацияда қолданылады.

1.6 Вокодерлер түрінің сипаттамасы

Қазіргі таңда дыбыстық сигналдарды цифрлық өңдеуде ең көп қолданылатын әдістердің бірі – PCM (Pulse Code Modulation), яғни импульстік кодтық модуляция. Бұл әдіс аналогтық сигналдарды сандық форматқа түрлендіруде, атап айтқанда сөйлеу, музыка, бейне және телеметриялық деректерді цифрлық байланыс арналары арқылы жеткізу кезінде кеңінен қолданылады. PCM әдісі өзінің қарапайымдылығы мен әмбебаптығына байланысты танымал болғанымен, кейбір жағдайда – әсіресе арнайы немесе қорғалған байланыс желілерінде – одан да тиімді әдістерді қолдану орындырақ.

Соңғы бірнеше онжылдық ішінде сөйлеу сигналдарын тиімдірек кодтауға арналған түрлі әдістер жасалды. Солардың бірі – Adaptive Delta Modulation (ADM), яғни PCM әдісінің бейімделген түрі. ADM амплитуда өзгерісін шектеулі интервалда бақылай отырып, сигналдың нақты мәні мен күтілетін мәнінің арасындағы айырмашылықты есептейді. Бұл айырмашылық төрт биттік кодтар арқылы жіберіліп, қабылдаушы жақта сигнал амплитудасы қайта құрылады.

Тағы бір жетілдірілген әдіс – CVSD (Continuously Variable Slope Delta Modulation), яғни айнымалы көлбеуі бар дельта модуляция. Бұл тәсіл бастапқыда әскери байланыс жүйелерінде қолданылған, себебі аналогтық сигналдарды шифрлау мен кодтауды жеңілдетеді. CVSD әдісінде әрбір өлшенген мән салыстырмалы деңгеймен бағаланады: егер сигнал тірек деңгейінен жоғары болса – «1», ал төмен болса – «0» ретінде кодталады. Алғашқы жүйелерде бұл процесс 8000 үлгі/сек жиілікте жүзеге асып, әр үлгіге 1 бит бөлінген. Нәтижесінде деректердің таралу жылдамдығы 8 кбит/с болды.

CML Microcircuits компаниясы ұсынған заманауи CVSD кодектері 128 кбит/с-қа дейінгі жоғары кодтау жылдамдығын қолдайды. Бұл кодектер қатарына FX619, MX629 және CMX639 микросхемалары жатады. Олар DECT, Eurocom секілді телекоммуникациялық стандарттарда және арнайы әскери байланыс жүйелерінде тиімді қолданылады [11].

CMX649 моделі – бұл PCM, ADM және CVSD форматтарында толық дуплексті кодтау мүмкіндігі бар көпфункционалды микросхема. Оның артықшылықтары – төмен құны, шағын өлшемі және портативті сымсыз байланыс құрылғыларына оңтайлы бейімделуі. Сондай-ақ, чип келесі қосымша функциялармен жабдықталған:

- Ұйқы/белсенді режим ауыстырғышы;
- Автоматты микрофон күшейткіші;
- Бағдарламаланатын кіріс-шығыс сүзгілері;
- Жылдамдықты бөлу схемалары.

Жаңа буындағы CMX7261 чипі – көп форматты транскодер, ол бірнеше кодтау және декодтау функцияларын бір құрылғыда біріктіреді:

- PCM, CVSD, G.729A форматтарында аналогтық сигналдарды декодтау;
- G.711 (μ -law, A-law) және G.729A форматтарында PCM-ды кодтау;
- Осы форматтар арасында транскодтау функциясын қамтамасыз ету.

CMX7261 өнімділікті төмендетпей, дауыстық байланыс сапасын арттыруға және жүйенің икемділігін кеңейтуге мүмкіндік береді. Бұл чип келесі салаларда кеңінен қолданылады:

- VoIP жүйелері;
- Қорғалған сымсыз PBX;
- Радиобайланыс құрылғылары;
- Дауыстық қауіпсіздік жүйелері.

Вокодер алғаш рет 1930 жылы әзірленіп, радио арқылы дауыстық хабарламаларды тарату кезінде жиілік ресурстарын үнемдеу мақсатында қолданылды. Бұл әдіс нақты сөйлеу сигналын жіберудің орнына, сөйлеу сигналының негізгі сипаттамаларын кодтап, оларды алушы жақтағы сөйлеу синтезаторына жеткізуге негізделген.

Сөйлеу синтезаторы үш негізгі элементтен тұрады:

- Үні генераторы – үндес дыбыстарды (мысалы, дауысты) қалпына келтіру үшін;

- Шу генераторы – дауыссыз дыбыстарды қалыптастыру үшін;
- Формант сүзгілер жүйесі – сөйлеушінің тембрі мен дыбыс ерекшеліктерін имитациялау үшін.

Нәтижесінде, синтезделген дауыс робот тәрізді естілсе де, ақпараттық мазмұны сақталады, бұл әсіресе шулы және тар жолақты арналарда тиімді.

RALCWI (Reliable Advanced Low Complexity Waveform Interpolation) – бұл ПДС Рух және CML Microcircuits бірлесіп әзірлеген төмен күрделілікті вокодер алгоритмі. Ол жоғары қысу дәрежесімен қатар, сапалы сөйлеуді қалпына келтіруді қамтамасыз етеді. Бұл алгоритм тар жолақты және шулы радиоарналарда қолдануға ерекше бейімделген.

RALCWI сигналды қысу мен кодтаудағы ерекше әдістерді пайдалана отырып, жоғары тиімділікті қамтамасыз етеді. Осылайша, бұл технология сандық радиобайланыс нарығында ең үздік баға/сапа арақатынасын ұсынады.

Тәуелсіз сарапшылардың бағалауы бойынша, RALCWI-вокодерлердің 2400 бит/сек жылдамдықта берілген дауыс сапасы, 4000 бит/сек жылдамдықпен жұмыс істейтін дәстүрлі вокодерлерден жоғарырақ деп танылады.

RALCWI вокодері, CML Microcircuits компаниясының жоғары интегралдық микросхемаларында іске асырылады және төрт негізгі тәуелсіз функцияны орындайды:

1. Дауыстық кодер (Speech Encoder);
2. Дауыстық декодер (Speech Decoder);
3. Кодер;
4. FEC декодері (Forward Error Correction – Алдыңғы қатені түзету декодері).

Бұл микросхемалар тар жиілік диапазонында, өте төмен жылдамдықпен жұмыс істей отырып, жоғары сапалы дауыс сигналын қамтамасыз етуге мүмкіндік береді. Сонымен қатар, олар функционалдық икемділігімен ерекшеленеді және жоғары деңгейлі интеграцияланған шешім ретінде сипатталады.

Тәуелсіз сарапшылардың пікірінше, RALCWI вокодерлері 2400 бит/с жылдамдықта дауысты стандартты 4000 бит/с вокодерлерге қарағанда сапалырақ жеткізе алады.

CMX608 және CMX618 микросхемалары – жартылай дуплексті вокодерлер, яғни олар екі бағыттағы байланысты кезекпен жүзеге асырады. Бұл чиптер 2050, 2400 немесе 2570 бит/с жылдамдықтағы цифрлық ағынды қолдайды (таңдау хост-контроллер арқылы жүзеге асады). Ал қателерді түзету режимінде (FEC) жұмыс істегенде, жалпы жылдамдық 3600 бит/с-ке дейін жетеді (2400 бит/с + 1200 бит/с FEC).

Аталған микросхемалар келесі қосымша мүмкіндіктермен жабдықталған:

- Voice Activity Detector (VAD) – сөйлеудің бар-жоғын анықтау;
- Comfort Noise Generator (CNG) – фондық шу генераторы;
- SND және DTMF тон генераторлары – бір және екі жиілікті сигналдарды генерациялау;

- Soft Decision Decoding – жұмсақ шешімді декодтау;
- Discontinuous Transmission Detection – үзілісті тарату режимін анықтау.

CMX618 моделі қосымша ретінде арналық сүзгілер және аудио кодек блогымен біріктірілген. Бұл чип толық дауысты аналогтық сигналға түрлендіру циклін қамтамасыз етеді, яғни кодталған төмен деңгейлі сигналдан тікелей дыбыстық шығару жүргізіледі.

CML Microcircuits компаниясы жақында жаңа радиопроцессор – CMX8341 чипін таныстырды. Бұл микросхема «жүйе-чипінде» (SoC) принципіне негізделіп жасалған және цифрлық пен аналогтық PMR-технологияларды бір уақытта қолдайды. Оның құрамында екі негізгі модуль бар:

- Аналогтық бөлік
- Сандық бөлік

Сандық модуль dPMR стандартындағы барлық негізгі функцияларды қамтиды, соның ішінде:

- 4FSK-модем,
- RALCWI вокодері,
- толық dPMR радиоинтерфейсі,
- физикалық және арналық деңгейлер,
- басқару деңгейі.

CMX8341 чипі FirmASIC технологиясы бойынша жасалған, бұл оның функционалдығын дайын бағдарламалық файлдар (Function Image) арқылы оңай бейімдеуге мүмкіндік береді. Бұл файлдар C-BUS интерфейсі арқылы хост-микроконтроллер немесе EEPROM жадынан жүктеледі және қажет болған жағдайда жаңартуға болады. Нәтижесінде бір ғана чип арқылы арзан әрі көп функционалды радиобайланыс құрылғыларын жасауға болады. Бұл құрылғылар PMR стандарттары бойынша жұмыс істейтін жүйелерге (мысалы, DuraTalk) өте ыңғайлы.

RALCWI технологиясы үшін бұл микросхема жоғары сапалы, бірақ төмен бағамен сөйлеу беру мүмкіндігін қамтамасыз етеді. Әсіресе бұл шешімдер dPMR/LMR (Digital Private Mobile Radio) нарығында сұранысқа ие.

RALCWI микросхемалары келесі салаларда кеңінен қолданылады:

- Транкингтік байланыс жүйелері;
- DMR TDMA және DMR FDMA технологиялары;
- Сандық шифрлау және дауыс арқылы байланыс;
- Сымсыз абоненттік жүйелер (WLL);
- Автохаттар мен дауыс хабарламаларды сақтау/ойнату жүйелері;
- Хабарламаларды тарату жүйелері;
- VoIP телефониясы;
- Дауыстық пейджерлер;
- Сандық регенеративті қайталағыштар.

CMX649 негізіндегі DE6491 демо-жинағы және RALCWI-вокодерлерге арналған құралдар бар.

DE6491 – бұл CMX649 микросхемасының жұмысын көрсетуге арналған демо-жинақ, ол төмен қуатты сымсыз гарнитуралар құрылымында қолданылады. Жүйе 863–865 МГц жиілік диапазонында жұмыс істейтін лицензияланбаған ISM радиодиапазонында жұмыс істейді.

Жинақ екі жеке тақшадан тұрады, олардың әрқайсысы:

- CMX649 негізіндегі аудио тізбектерді;
- Сигнал қабылдау және беру жүйесін;
- Басқару элементтерін қамтиды.

Бұл базалық схема басқа жиілік диапазондарына оңай бейімделе алады, сондықтан икемді әрі функционалды шешім болып табылады.

EV6180 және EV6380 – бұл CMX6x8 сериялы RALCWI вокодерлерінің жұмысын көрсету және жөндеу үшін арналған арнайы әзірленген баптау және тестілеу тақшалары. Бұл құралдар жартылай және толық дуплексті режимдерде вокодер өнімділігін талдауға және оңтайландыруға мүмкіндік береді [12].

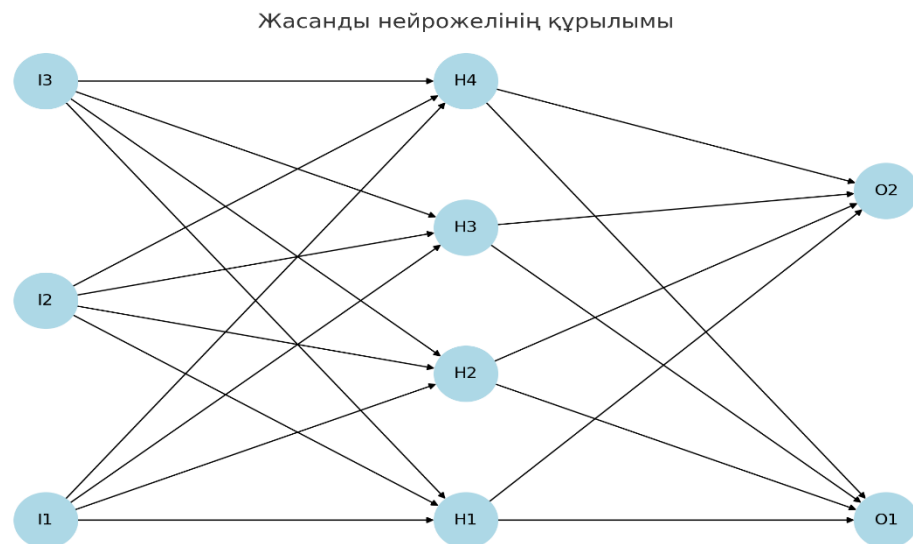
2 Нейро желілер олардың моделдерін зеттеу

Қазіргі таңда нейрондық желілер түрлі салалардағы күрделі мәселелерді шешуге арналған заманауи технологиялардың бірі болып табылады. Бұл технологиялар телекоммуникациялық жүйелерде күрделі есептерді тиімді шешуге кең мүмкіндіктер ұсынады.

Телекоммуникациялық жүйелерде нейрондық желілерді қолдану келесі артықшылықтарды қамтамасыз етеді:

- Ақпаратты параллельді түрде жоғары жылдамдықпен өңдеу мүмкіндігі;
- Күрделі әрі дәстүрлі әдістермен шешілуі қиын есептерді шешу мүмкіндігі;
- Сызықты емес процестер мен бөгеуілдерге ұшыраған деректерді
- бейімделген түрде модельдеу қабілеті.

Нейрондық желілердің құрылымы синапстардан, нейрон ұяшықтарынан және аксондардан тұрады. Бұл элементтер өзара байланысып, келесі негізгі құрылымды құрайды:



2.1-сурет – Жасанды нейрожелінің құрылымы

Перцептрон моделінде екі түрлі нейрон типі қарастырылады. Бірінші тип – рецепторлық нейрондар, олар кірісінде екілік сигналдарды қабылдап, оларды әрі қарай өңдеусіз жеткізеді. Екінші тип – ассоциативті нейрондар, олар алынған сигналдарды өңдеп, белгілі бір шешім қабылдауға қатысады.

Нейрондық желінің құрылымы өзара тәуелсіз екі қабаттан тұрады:

- Бірінші қабат рецепторлық нейрондардан құралған;
- Екінші қабат ассоциативті нейрондардан тұрады.

$$f(X_z) = \begin{cases} 1, X_z \geq \Theta \\ 0, X_z < \Theta \end{cases}$$

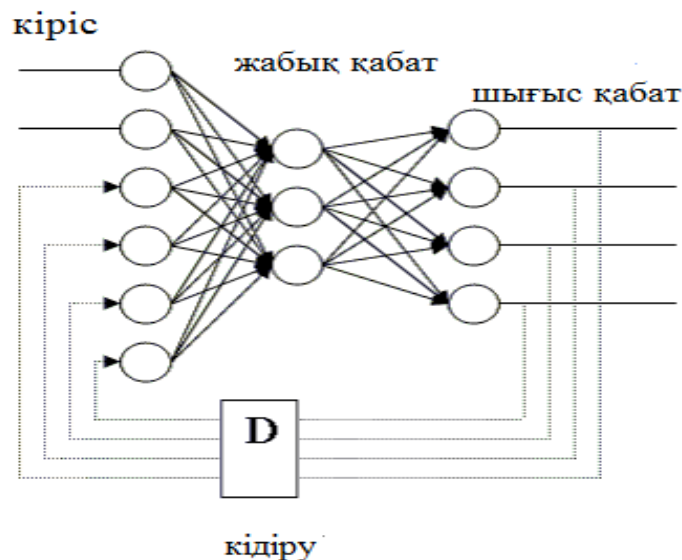
Бірінші қабаттағы нейрондар екінші қабаттағы нейрондармен тізбектей байланысады, яғни әрбір рецептор бір немесе бірнеше ассоциативті нейронға ақпарат береді [8].

Нейронның моделі персептронда екі түрлі нейрон болады.

Нейрондарың бірінші типі рецепторлар, олар кірісінде екілік сигналды қабылдап оны әрі қарай береді.

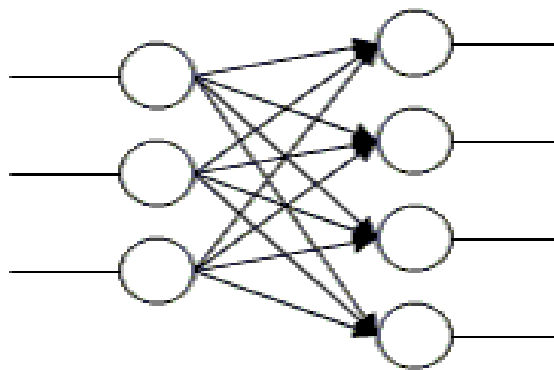
Нейронның екінші типі ассоциативті нейрондар, олар қарапайым ақырғы:

Нейрожелінің байланысы тәуелсіз екі қабаттан тұрады. Бірінші қабат рецепторлардан, ал екінші қабаты ассоциативті нейрондардан тұрады. Бірінші қабат нейрондары келесі қабат нейрондармен тізбектей байланысады. [8]



2.2-сурет – Нейрожелінің байланыс құрлымы

Бірінші қабат рецепторлардан тұрады ал екінші қабат нейрондардан тұрады. Мұндай желіні бір қабатты желі деп атайды.



2.3-сурет – Нейро желінің құрлымы

Айтып кетітін жағдайдың бірі персептронның кірісі мен шығысы ол екілік вектор, яғни персептрон екілік функция бейнесі болып табылады. Мұндағы нақты сан ретінде рецепттің мен ассоциативті қабатты байланыстыратын байланыс таразысы қарастыруға болады. Нейрон j -ші желісінің V_j векторы анықталған. Яғни персептрон $m \times n$ размерлі W матрицасымен анықталады. i -ші рецептордың j -шы ассоциативным нейронның арасында W_{ij} байланыс таразысы болып келеді.

2.1 Нейро желілердің жіктелуі

Инженерлік көзқарас тұрғысынан нейрондық желілер – бұл жоғары дәрежеде параллельденген динамикалық жүйелер, олар өз күйін кіріс сигналдарының әсерінен өзгерту арқылы ақпаратты өңдеуге қабілетті.

Нейрожелілік жүйелердің құрылымы әрбір нақты есепке байланысты жеке нейрондардан, синапстардан, дендриттерден және архитектуралық байланыстардан тұрады [13]. Осы желілерді құру барысында келесі қасиеттер ескерілуі қажет:

- Желінің мүмкіндіктері оның қабат саны мен қабаттар арасындағы байланыстар саны артқан сайын өседі;
- Кері байланыстың енгізілуі жүйенің динамикалық тұрақтылығын күшейтеді;
- Желінің алгоритмдік күрделілігі оның есептеу қуатын арттыруға мүмкіндік береді.

Нейрондық желілер құрылысы мен тағайындалуына қарай үш негізгі түрге бөлінеді:

- Иерархиялық желілер – мұнда ақпарат бір деңгейден келесі деңгейге тізбекті түрде беріледі. Бұл желілерде нейрондар коэнвергентті (ақпаратты біріктіретін) және дивергентті (ақпаратты тарататын) түрде әрекет етеді.

- Локальды желілер – нейрондар белгілі бір шектелген аймақта ғана өзара әрекеттеседі, яғни олардың байланыс ауқымы шектеулі.
- Дивергентті желілер – мұндай желілерде бір ғана кіріс сигналы болады, ал шығыстар бірнеше бағытта таралады.

Шығыс функциясының сипатына байланысты нейрондық желілер:

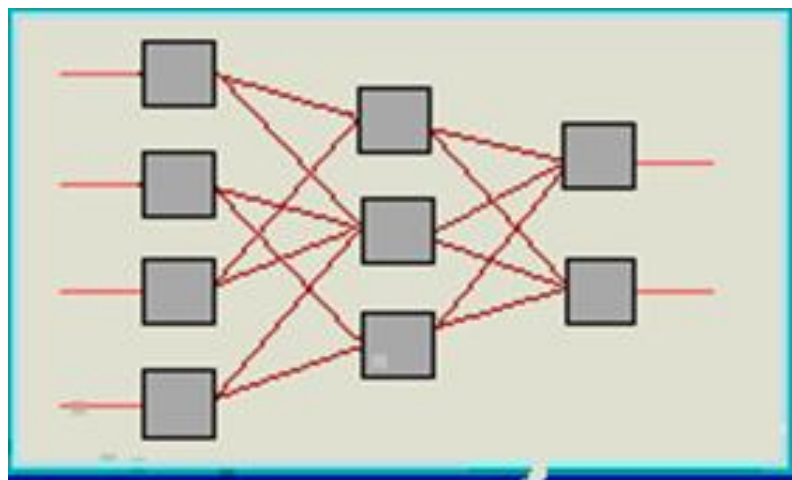
- Бинарлық – кіріс және шығыс мәндері тек екі күйде болады: логикалық 0 және логикалық 1;
- Аналогтық – шығыс мәндері үздіксіз интервалдағы нақты сандар ретінде беріледі, бірақ олар көбіне дискреттелген мәндермен алмастырылады.

Архитектуралық ерекшелігіне қарай нейрондық желілер екі негізгі топқа бөлінеді:

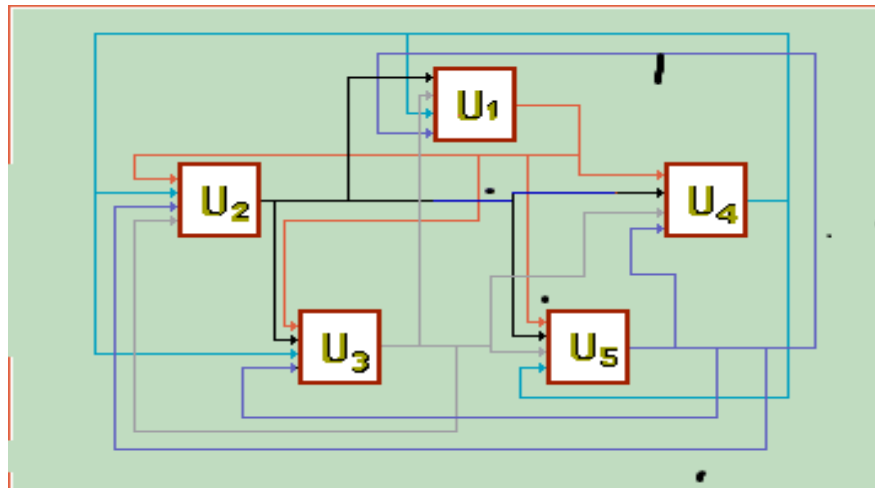
- Тура таралатын желілер (Feedforward Neural Networks) – ақпарат желі ішінде бір бағытта, артқа қайтпай таралады.
- Рекуррентті желілер (Recurrent Neural Networks) – бұл желілерде кері байланыс бар, яғни нейрондар арасындағы байланыстар уақыт бойынша кері бағытта әрекет етеді.

Тура тарататын желілерде байланыс түйідері болмайды. Тура таралатын желілер бір қабатты перцептрондар желісі және көп қабатты перцептрондар желісі болып бөлінеді. Төменде 2.5 -2.6 суретте келтірілген.

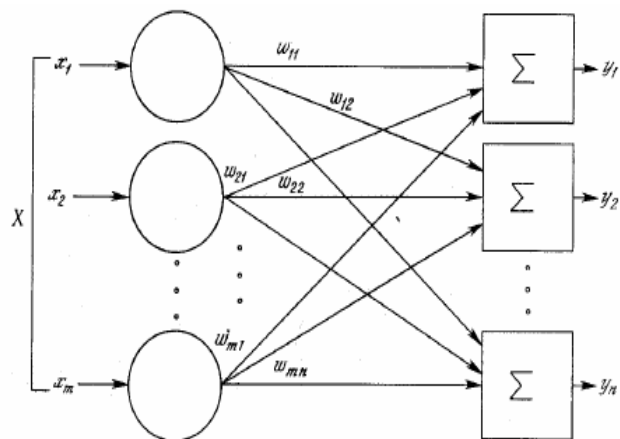
Рекуррентті желілер дегеніміз кері байланыстарды қолдану арқылы құралатын желер болып келеді [9].



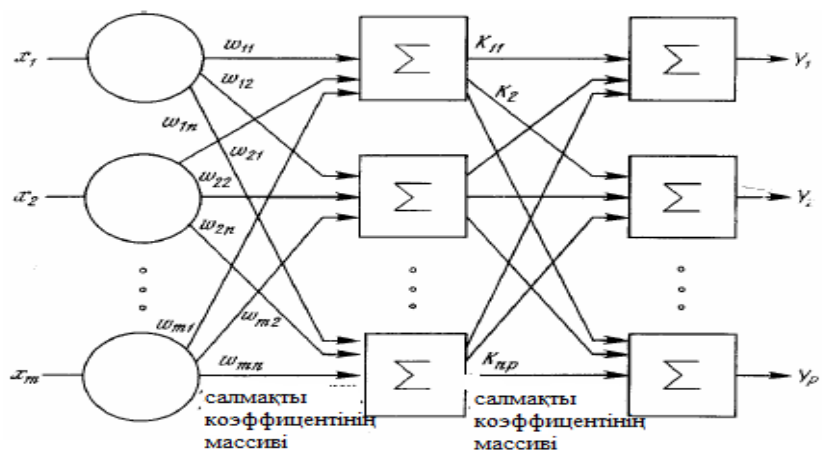
2.4-сурет – Тура таралу желілері



2.5-сурет – Реккурентті желілер



2.6-сурет – Бір қабатты перцептрондар нейро желі



2.7-сурет – Көп қабатты перцептрондар нейро желі

Көп қабатты перцептрондар нейро желілердің ерекшелігі кіріс және шығыс берілгендер арасында бірнеше жабық нейро қабаттардан тұрады, бұл модельге сызықты емес бірнеше байланыстарды қасады.

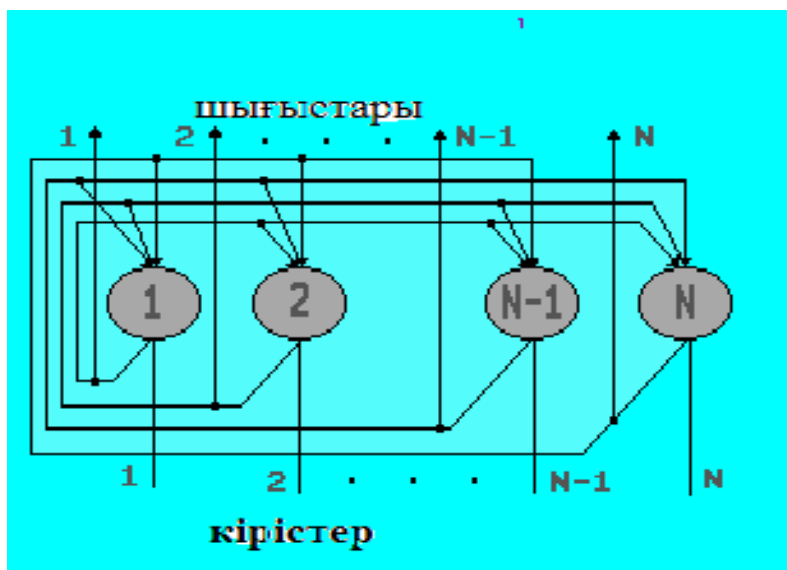
2.2 Рекуррентті нейрожелілер моделдері

Рекуррентті нейрожелілердің моделдері кеңінен таралған және олар құрлымы бойынша күрделі болып келеді.

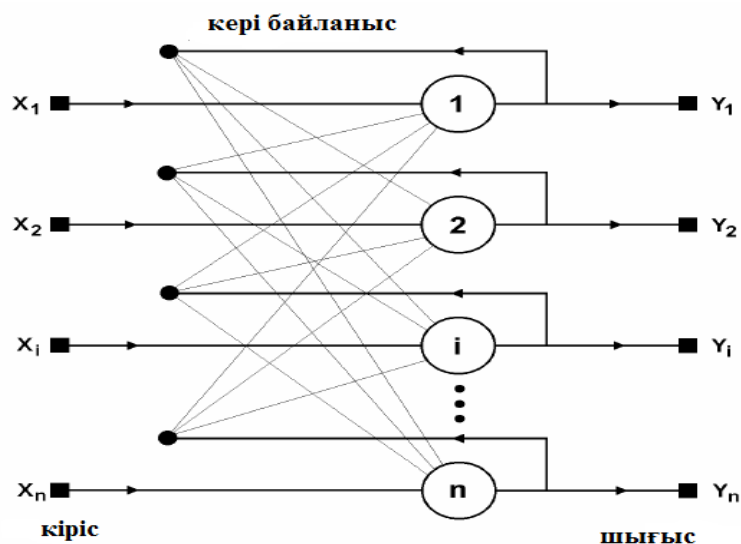
Рекуррентті желілердің жұмысы дифференциальды теңдеумен немесе айырымды теңдеулермен сипатталады. Бұл мұндай моделдердің кеңінен қолдануына мүмкіндік береді. Желінің жұмыс барысында әрбір нейрон кіріс ақпаратты басқа нейрондардан, өзінен және қоршаған ортадан алып отырады. Бұндай желілердің мәнісі жоғары болып келеді себебі олар сызықты емес динамикалық жүйелерді моделдеу мүмкіндігін береді.

Рекуррентті желілер арасынан Хопфилд және Кохонен желілерін атап өтуге болады.

Хопфилд желісінің көмегімен бейтарап, уақыт бойынша жинақталған немесе жазықтықтағы үлгілерді өңдеуге болады.



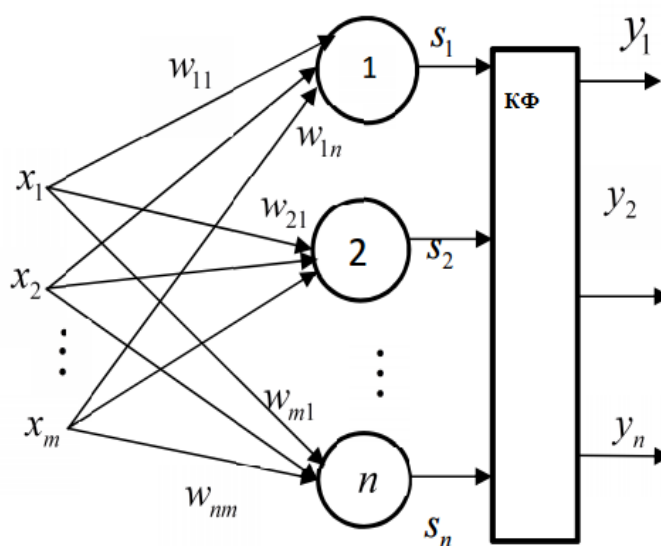
2.8-сурет – Хопфилд желісінің құрлымы



2

2.9-сурет – Хопфилд желісінің қабаттар арасындағы байланысы

Кохонен желісін «өзін өзі ұйымдастыратын қасиеттер картасы» деп атайды. Бұл желінің ерекшелігі ол өзін-өзі оқыту арқылы дұрыс шешім қабылдайды. Оқыту барысында желінің кірісіне түрлі үлгілер беріледі. Желі олардың арасындағы құрлымды анықтар кластерге бөледі, ал оқытылған желі келіп түскен мысалды белгілі кластерге тағайындайды. Желі бір кіріс және бір шығыс қабаттан тұрады. Шығыс қабаттағы элементтер саны желінің анықтайтын кластерлерінің санына байланысты болады. Әр байланыстың синаптикалық таразысы болады [10-13].

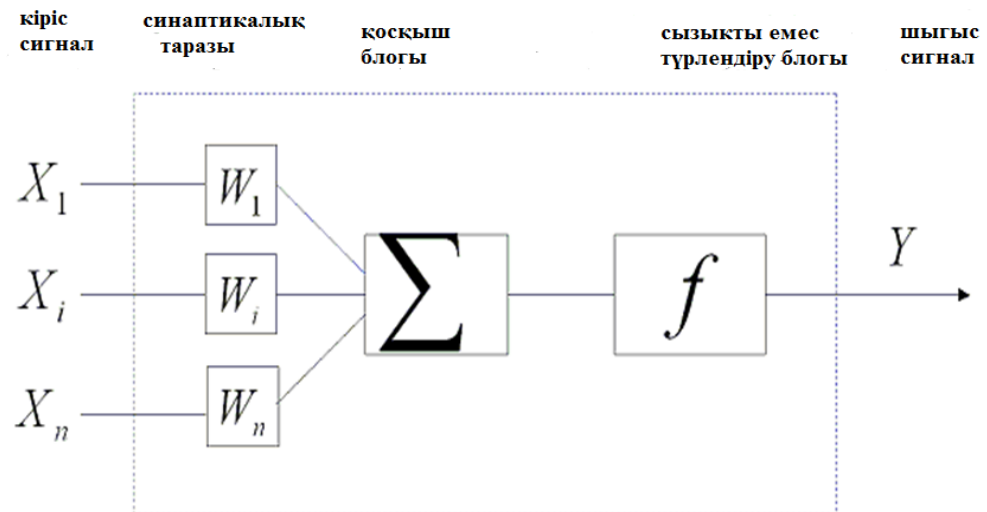


2.10-сурет – Кохонан желісінің құрлымы: КФ– конкуренция функциясы; X_m –кіріс; W_{m1} –Синаптикалық таразы; Y_n –шығыс.

2.3 Нейрожелілердің математикалық моделдерінің құрлымы

Нейро желінің математикалық моделінің құрлымы үш логикалық блоктардан тұрады:

- кіріс,
- түрлендірі функциясы,
- шығысы.



2.11-сурет – Нейро желінің құрлымы

Түрлендіру функциясын келесі бөліктерге бөлуге болады: синаптикалық таразы, қосу блогы, сызықты емес түрлендіру функциясы.

Түрлендіру функция блогының жұмыс алгоритмін сипаттайық:

- Шығыс сигналдар синаптикалық таразы коэффициентіне көбейтіліп отырылады. Әрбір кіріс үшін өз коэффициенті;
- Таразыға көбейтілген кірістер қосылады;
- Қосынды нәтижесі түрлендіру функциясының кірісіне беріледі;
- Функция нәтижесі шығысқа беріледі.

Кез-келген нейронды желіні сипаттау үшін оған келесі қалыптарды анықтау қажет:

- формалды нейрон моделі (түрлену функциясы және т.б.)
- нейрондық желілер байланысының құрылымы
- желіні оқу әдісі

Желіні оқыту үшін келесі шарттарды қарастырған жөн болады. Нейро желі керекті үлгілерді анықтау үшін, яғни әрбір кіріс берілгендер векторына керекті векторды шығысында құру үшін түптеу керек. Бұл процессті оқыту деп атайды. Оқытудың екі әдісі бар: оқытушымен және өзін-өзі ұйымдастыру арқылы.

Нейрожелі талап етілетін бейнені жүзеге асыру, яғни, әрбір кіріс мәліметтерінің векторына шығыста талап етілетін векторды жасау үшін оны

баптау қажет. Әдетте баптау үдерісі «оқу» деп аталады (бұдан әрі біз әрдайым оқу деп айтатын боламыз). Оқудың екі әдісі бар: тәлімгермен бірге және өзіндік ұйымдастыру негізінде [14].

Барлық желілерге ортақ тәлімгермен үйрену қағидасын қарастырамыз. Келесі түрдегі бірқатар жұптар жиынтығы бола берсін: кіріс векторы – желінің шығуы (яғни, бізге желі кейбір көп кіріс векторларында қалай нақты жұмыс істеуі керектігі белгілі). Мұндай жұптар жиынтығы «үйрететін іріктеме» деп аталады (бұдан әрі біз нақ осы терминді қолданатын боламыз). Үйрену үдерісі келесідей болады:

- Синаптикалық салмақтарға кейбір бастапқы мәндерді береміз;
- Үйрететін іріктемеден вектор желісін көрсетеміз;
- Желі шығысы жоспарлыдан ауытқуын есептейміз;
- Кейбір алгоритмдер бойынша (нақты бір архитектураға тәуелді болатын) алдыңғы қадамға алынған ауытқуды есепке ала отырып, синаптикалық салмақтарды құрамыз;
- Желі үйрететін іріктеме векторларындағы күтілетін шығысты бермегенше немесе ауытқу кейбір шектерден төмен болмағанша 2-4 қадамдарды қайталаймыз.

Басқаша айтқанда, желі салмақтарын желі белгілі мәліметтерде қанағаттанарлық дәлдікпен жұмыс істейтіндей етір құрамыз. Әдіс идеясы барлық ықтимал жиынтық кіріс векторларын азды-көпті біртекті қамтитын жеткілікті түрде репрезентативті үйрететін іріктемеде үйретілген нейрожелі оны оқытпаған векторларда дұрыс немесе дұрысқа жақын нәтиже береді деген болжамда болып табылады.

Мұндай болжам ережеге сәйкес екенін ескерейік, себебі, оқытылған желі үйрететін іріктемедегі векторлармен ұқсас түрде кіріс векторларын өңдейтін болады, тиісінше, ізделіп отырған түрлену жеткілікті дәрежеде үздіксіз және тегіс болатын болса, нәтиже де қанағаттанарлық дерлік болу керек.

Өзіндік ұйымдастыру көп емес нейрондық желілердің сипаты. Өзіндік ұйымдастыру желінің жұмыс үдерісінде өз бетімен қалаулы бейнелеуді жүзеге асыруға, яғни, үйрететін іріктемесіз жүргізуге бейімделу қабілетін білдіреді. Әдетте мұндай желілер кейбір жиынтықтарды кластерлеу есептерінде қолданылады. Оқудың бұл әдіс үшін бір жалпы абстрактілі алгоритмді тұжырымдау қиын, себебі, әдетте оқу әдісі нақты бір желінің архитектурасына қатты тәуелді болады.

Жасанды нейрон желісі – бұл математикалық моделі, сондай-ақ құрылғылар параллельді есептеу, білдіретін жүйесі америка құрама штаттары және өзара іс-қимылдағы өзара қарапайым процессорлар (жасанды нейрондық). Математикалық моделі ретінде жасанды нейрон желісі білдіреді жеке жағдайы әдістерін бейнелерді тану немесе дискриминантты талдау болып табылады.

3 Нейрожелі моделдерін дыбысты тануға қолдануды өңдеу

Адамның дыбыстық сөйлеуі өзінің жоғары дәрежелі өзгермелілігімен ерекшеленетіні белгілі. Бұл құбылыс бірнеше негізгі факторларға байланысты:

1. Индивидуалды вариациялар: Бір адам бір сөзді бірнеше рет айтқан жағдайда, әрбір айтылым акустикалық жағынан бір-бірінен айырмашылықтармен сипатталады. Бұл айырмашылықтар сөйлеудің спектралдық құрамында және айтылу ұзақтығында көрінеді. Мұндай ауытқулар адамның физиологиялық жағдайына, эмоционалдық күйіне, ортаға, және контекстке байланысты туындайды.

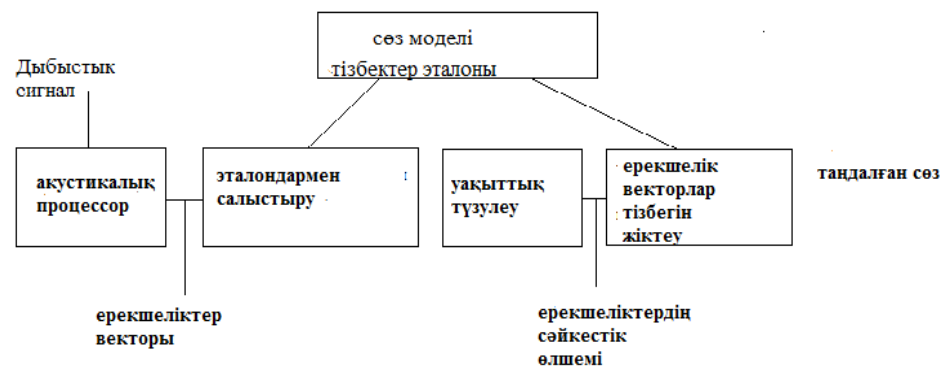
2. Коартикуляциялық әсерлер: Сөздің құрамындағы әрбір фонема өз контекстіне тәуелді түрде айтылады. Фонемалар бір-біріне әсер етіп, айтылымның акустикалық сипаттамаларын өзгертеді, бұл сөйлеуді тануда қосымша күрделілік тудырады.

3. Сыртқы кедергілер: Сөйлеу сигналына әртүрлі шу, фондық дыбыстар немесе техникалық бұрмаланулар әсер етуі мүмкін, бұл да тану процесіне кері әсер етеді.

Жоғары сапалы және нақты уақыт режимінде сөйлеуді тану үшін жоғары өнімді есептеу жүйелері қажет. Бұл талаптарды төмендетудің бір тиімді жолы – есептеу синхронизациясын қолдану. Бұл тәсіл жасанды нейрондық желілердің негізінде құрылған нейрокомпьютерлерді пайдалану арқылы табиғи түрде іске асады.

Қазіргі таңда қысқа сөйлеу сегменттері мен фонетикалық ұқсастығы бар оқшауланған сөздерді тану үшін жоғары тиімділікке ие нейрондық модельдер жасалып үлгерді. Бұл модельдердің бірқатары дәстүрлі сөйлеу тану әдістерімен біріктіріліп, оқшауланған сөздер мен тұтас сөйленісті тануға арналған кең сөздік негізіндегі тану жүйелерінде табысты қолданылып жүр.

Қарастырайық қарапайым сызбасын тану ұсынылған [6]



3.1-сурет – Сөз тізбегін танып білудің құрылымы

Сөйлеуді тану жүйесінің жұмысын үш негізгі кезеңге бөлуге болады:

1) Акустикалық алдын ала өңдеу (фичалар алу).

Бұл кезеңде акустикалық процессор кіріс сөйлеу сигналын белгілі уақыт аралықтарында алынатын векторлар тізбегіне түрлендіреді. Бұл векторлар спектрлік және кепстральдық коэффициенттерді қамтиды және сөйлеу сигналының қысқа фрагменттерінің акустикалық сипаттамаларын көрсетеді.

2) Салыстыру және жергілікті сәйкестік өлшемдерін есептеу.

Алынған векторлар эталондық үлгілермен салыстырылады. Бұл үлгілер – сөздердің алдын ала сақталған модельдері. Әр вектор үшін жергілікті метрикалар немесе сәйкестік көрсеткіштері есептеледі. Бұл метрикалар вектор мен модель арасындағы ұқсастық деңгейін сипаттайды [11-13].

3) Уақытша теңестіру және шешім қабылдау.

Осы кезеңде уақытша теңестіру (Dynamic Time Warping – DTW) әдісі қолданылады. Ол айтылу жылдамдығындағы ауытқуларды өтеуге көмектеседі. Векторлар мен эталондар тізбегі уақыт бойынша реттеліп, сәйкес келу дәрежесі есептеледі.

Барлық есептеулер аяқталғаннан кейін, тану жүйесі ең жоғары сәйкестік көрсеткіші бар сөзді таңдайды.

Егер сөйлеу тұтас сөйлемдер немесе фразалар түрінде берілсе (слитная речь), онда екінші кезеңде алынған жергілікті метрикалар жеке сөздер мен тіркестерді бөлу және сәйкестігін бағалау үшін пайдаланылады.

Жоғары сапалы тану нәтижесіне қол жеткізу үшін кейде қосымша лингвистикалық өңдеу кезеңі енгізіледі. Бұл кезеңде семантикалық, синтаксистік және прагматикалық шектеулер ескеріліп, ықтимал сәйкестік деңгейін арттыруға көмектеседі.

Сөйлеу тану схемасының екінші кезеңінде, яғни жергілікті метрикаларды есептеу барысында нейрондық желілер (НС) ерекше тиімділікпен қолданылады. Нейрожелілік модельдер сигналдар арасындағы ұқсастық функцияларын есептеуде жоғары дәлдік көрсетеді. Статистикалық тану жүйелерінде бұл көрсеткіштер монотонды функциялар ретінде қарастырылады және олар белгілер векторлары арасындағы ұқсастықты сипаттайды.

Сөйлеуді тану жүйелерінде ұқсастық функцияларын есептеу маңызды рөл атқарады. Мысалы, гаусстық таралу негізіндегі векторлар арасындағы ұқсастықты сипаттау үшін логарифмдік функциялар сияқты қарапайым функциялар қолданылады. Мұндай функцияларды есептеу бір қабатты нейрондық желілермен жүзеге аса алады, бұл желілер алдын ала оқыту арқылы бапталады.

Күрделірек метрикаларды есептеу қажет болған жағдайда, көп қабатты перцептрондар (MLP) қолданылады. Бұл желілер күрделілігі кез келген деңгейдегі функцияларды аппроксимациялауға қабілетті. Оларды оқыту барысында салмақ коэффициенттері реттеліп, желі *a posteriori* ықтималдықтарды (яғни, бір класқа тиістіліктің ықтималдығы) жақындатады. Бұл қасиет гибридті сөйлеу тану жүйелерінде, атап айтқанда Жасырын Марков модельдерімен

(НММ) бірге сәтті қолданылады. Мұнда көп қабатты нейрондық желі жағдайдың ықтималдықтарын есептеу блогы ретінде қолданылады.

Осындай тәсілдер векторлар арасындағы акустикалық контексті ескеруге мүмкіндік береді және таралу пішініне қатысты нақты болжамдарға тәуелділікті жояды.

Дискретті бақылау кезінде сөйлеу сигналының фича-векторлары алдымен векторлық кванттауға ұшырайды. Әрбір вектор кодтық кітаптағы белгілі бір символға сәйкестендіріледі. Одан кейін арнайы кестелер арқылы әр символ үшін эталондық үлгілермен сәйкестік ықтималдығы есептеліп, жергілікті метрикалар алынады [14].

Бұл есептеулер бір қабатты перцептрондар арқылы орындалуы мүмкін. Мұндай перцептрондар желісінің торап саны эталондар санына тең болады, ал әр кіріс нейрон ықтимал символдар санымен сәйкес келеді.

Векторлық кванттау және Хопфилд желісі

Векторлық кванттау процесі Хопфилд желісі арқылы жүзеге асырылуы мүмкін. Бұл желі – екі өлшемді тораптар массиві, мұнда әрбір торап кодтық кітаптағы бір ықтимал символды білдіреді. Әр торап кіріс векторы мен эталон арасындағы евклид қашықтығын есептейді, ал ең кіші қашықтыққа ие торап таңдалады.

Мұндай желінің салмақ коэффициенттері:

- Хопфилд алгоритмімен;
- оның модификациялары арқылы;
- Немесе дәстүрлі векторлық кванттау алгоритмдері көмегімен есептеледі, оларда евклидтік қашықтық негізгі метрика ретінде пайдаланылады.

3.1 Хопфилд нейрожелілі моделін қолданып дыбыс сигналын танып білудің математикалық моделі.

Есептеу шаралары, Хопфилд нейро желілерін қолдану арқылы жүргізіледі, яғни дыбыстық сигналдардың дискреттік мәндерінің үлгісін енгіземіз:

$$x_1 = [-11-111-1 \ 1 \ 1]$$

$$x_2 = [-11-111 \ 1 \ 1 \ -1]$$

$$x_3 = [-11-111-1-1 \ 1]$$

$$y = [-1 \ 1 \ -1 \ 1 \ -1 \ 1 \ 1 \ 1]$$

Матрицаны аламыз:

-1	1	-1	1	-1	1	-1	-1
1	-1	1	1	-1	1	1	1
-1	1	-1	-1	-1	1	-1	-1

$$\begin{array}{rcl}
1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 & = & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\
-1 & & & & & & & & & & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\
\\
-1 & & & & & & & & & & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 \\
-1 & & & & & & & & & & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & = & -1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 \\
-1 & & & & & & & & & & 1 & -1 & 1 & -1 & -1 & -1 & -1 & 1 \\
\\
-1 & & & & & & & & & & 1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 \\
-1 & & & & & & & & & & 1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 \\
1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 & = & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & -1 & -1 & 1 \\
-1 & & & & & & & & & & 1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 \\
-1 & & & & & & & & & & 1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 \\
1 & & & & & & & & & & -1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 & 1
\end{array}$$

Үлгілердің дискреттік мәндерінің матрицаларын жинақтаймыз және Хопфилд алгоритмінің барлық мәндерін аламыз.

$$\begin{array}{r}
3 & -3 & 3 & -3 & -3 & 1 & -1 & -1 \\
-3 & 3 & -3 & 3 & 3 & -1 & 1 & 1 \\
3 & -3 & 3 & -3 & -3 & 1 & -1 & -1 \\
-3 & 3 & -3 & 3 & 3 & -1 & 1 & 1 \\
-3 & 3 & -3 & 3 & 3 & -1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 & 3 & 1 & -3 \\
-1 & 1 & -1 & 1 & 1 & 1 & 3 & -1 \\
-1 & 1 & -1 & 1 & 1 & -3 & -1 & 3
\end{array}$$

Диагональды нөлге келтіреміз

$$\begin{array}{r}
0 & -3 & 3 & -3 & -3 & 1 & -1 & -1 \\
-3 & 0 & -3 & 3 & 3 & -1 & 1 & 1 \\
3 & -3 & 0 & -3 & -3 & 1 & -1 & -1 \\
-3 & 3 & -3 & 0 & 3 & -1 & 1 & 1 \\
-3 & 3 & -3 & 3 & 0 & -1 & 1 & 1
\end{array}$$

$$\begin{array}{ccccccc} 1 & -1 & 1 & -1 & -1 & 0 & 1 & -3 \\ -1 & 1 & -1 & 1 & 1 & 1 & 0 & -1 \\ -1 & 1 & -1 & 1 & 1 & -3 & -1 & 0 \end{array}$$

Ұқастылыққа тексереміз:

$$\begin{array}{cccccccc|c} 0 & -3 & 3 & -3 & -3 & 1 & -1 & -1 & -1 \\ -3 & 0 & -3 & 3 & 3 & -1 & 1 & 1 & 1 \\ 3 & -3 & 0 & -3 & -3 & 1 & -1 & -1 & -1 \\ -3 & 3 & -3 & 0 & 3 & -1 & 1 & 1 & 1 \\ -3 & 3 & -3 & 3 & 0 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 0 & 1 & -3 & 1 \\ -1 & 1 & -1 & 1 & 1 & 1 & 0 & -1 & 1 \\ -1 & 1 & -1 & 1 & 1 & -3 & -1 & 0 & 1 \end{array}$$

$$\begin{array}{cccccccc|c} 0 & + & 3 & - & 3 & + & 3 & - & 1 & + & 1 & + & 1 & = & 7 \\ -3 & + & 0 & - & 3 & + & 3 & + & 3 & - & 1 & + & 1 & + & 1 & = & 1 \\ -3 & + & 3 & + & 0 & + & 3 & + & 3 & - & 1 & + & 1 & + & 1 & = & 7 \\ -3 & + & 3 & - & 3 & + & 0 & + & 3 & - & 1 & + & 1 & + & 1 & = & 1 \\ 3 & - & 3 & + & 3 & - & 3 & - & 0 & + & 1 & - & 1 & - & 1 & = & -1 \\ 1 & - & 1 & + & 1 & - & 1 & - & 1 & + & 0 & + & 1 & - & 3 & = & -3 \\ -1 & + & 1 & - & 1 & + & 1 & + & 1 & + & 1 & + & 0 & - & 1 & = & 1 \\ -1 & + & 1 & - & 1 & + & 1 & + & 1 & - & 3 & - & 1 & + & 0 & = & -3 \end{array}$$

$$\begin{array}{cccccccc|c|c} 0 & - & 3 & + & 3 & - & 3 & - & 3 & + & 1 & - & 1 & - & 1 & = & -7 & -1 \\ 3 & - & 0 & + & 3 & - & 3 & - & 3 & + & 1 & - & 1 & - & 1 & = & 1 & 1 \\ 3 & - & 3 & - & 0 & - & 3 & - & 3 & + & 1 & - & 1 & - & 1 & = & -7 & -1 \\ 3 & - & 3 & + & 3 & - & 0 & - & 3 & + & 1 & - & 1 & - & 1 & = & 1 & 1 \\ -3 & + & 3 & - & 3 & + & 3 & + & 0 & - & 1 & + & 1 & + & 1 & = & 1 & 1 \\ -1 & + & 1 & - & 1 & + & 1 & + & 1 & - & 0 & - & 1 & + & 3 & = & -3 & -1 \\ 1 & - & 1 & + & 1 & - & 1 & - & 1 & - & 1 & - & 0 & + & 1 & = & 1 & 1 \\ 1 & - & 1 & + & 1 & - & 1 & - & 1 & + & 3 & + & 1 & - & 0 & = & 3 & 1 \end{array}$$

3.2 Matlab ортасында ұқсату модельдерін құру

Хопфилд нейро желісінің ұқсату моделін дыбысты сигналды тану мақсатында құру nntool графикалық интерфейсі көмегімен, Matlab біріктірілген ортасын қолданумен жүргізіледі.[15]

Жұмыстың орындалуы үшін Matlab-ты іске қосып, пәрменді терезеге өту керек. Тұтынушының графикалық интерфейсін қолдана отырып, операцияның орындалуына арналған нейронды желі мен сөздік сигналдың дискреттік мәндері үлгілерінің тапсырмаларын құрамыз.

$$x_1 = [-1 \ 1 \ -1 \ 1 \ 1 \ -1 \ 1 \ 1]$$

$$x_2 = [-1 \ 1 \ -1 \ 1 \ 1 \ 1 \ 1 \ -1]$$

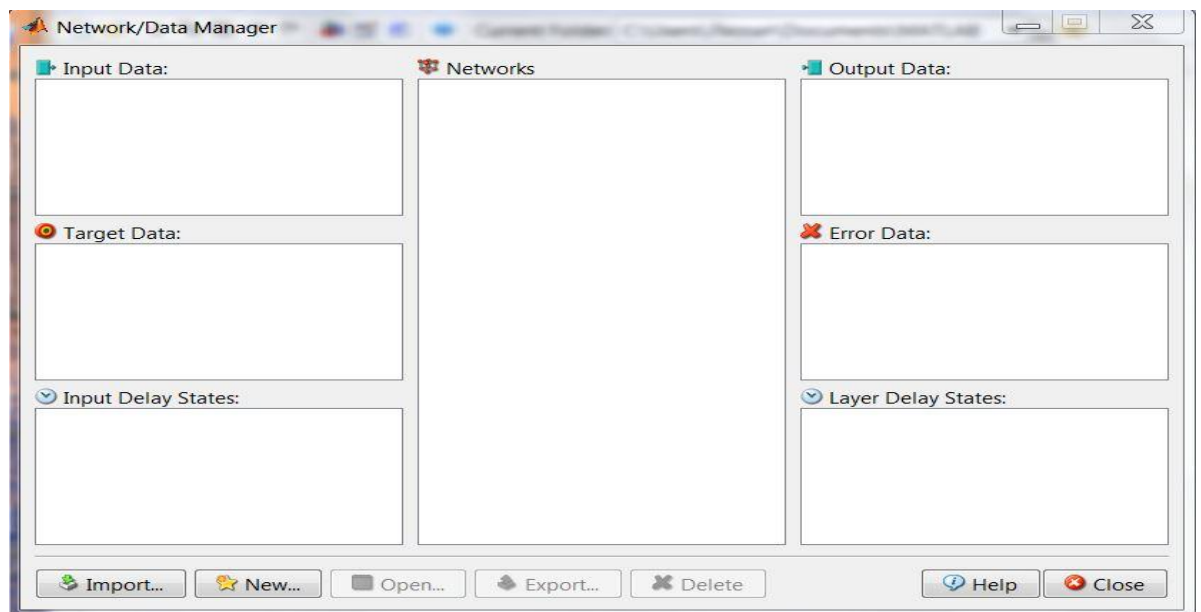
$$x_3 = [-1 \ 1 \ -1 \ 1 \ 1 \ -1 \ -1 \ 1]$$

Сондай-ақ мақсатты мәліметтердің векторы да беріледі

$$y = [-1 \ 1 \ -1 \ 1 \ -1 \ 1 \ 1 \ 1]$$

Интерфейстің негізгі терезесін ашу үшін MATLAB пәрменді терезесінде «>> nntool» пәрменін енгізу қажет

Пәрменді орындау Network/Data Manager нейронды желімен құрылған терезенің ашылуына алып келеді.



3.2-сурет – Нейронды желімен құрылған терезе

Create New Data терезесін қолдана отырып, GUI – интерфейсінің жұмыс аймағында кірістер мен мақсаттардың реттілігін құрамыз. Осы мақсатта New Data батырмасын басамыз, ары қарай Create New Data терезесінің Name алаңында алдымен x айнымалының атын енгіземіз, кейін Value мәндер алаңында мәндер векторын енгіземіз

$$[-1 \ 1 \ -1 \ 1 \ 1 \ -1 \ 1 \ 1]$$

және, Inputs (терезенің оң бөлігінде) радиобатырмасын қолдана отырып, айнымалылар типін көрсетеміз (Inputs - Кірістер).

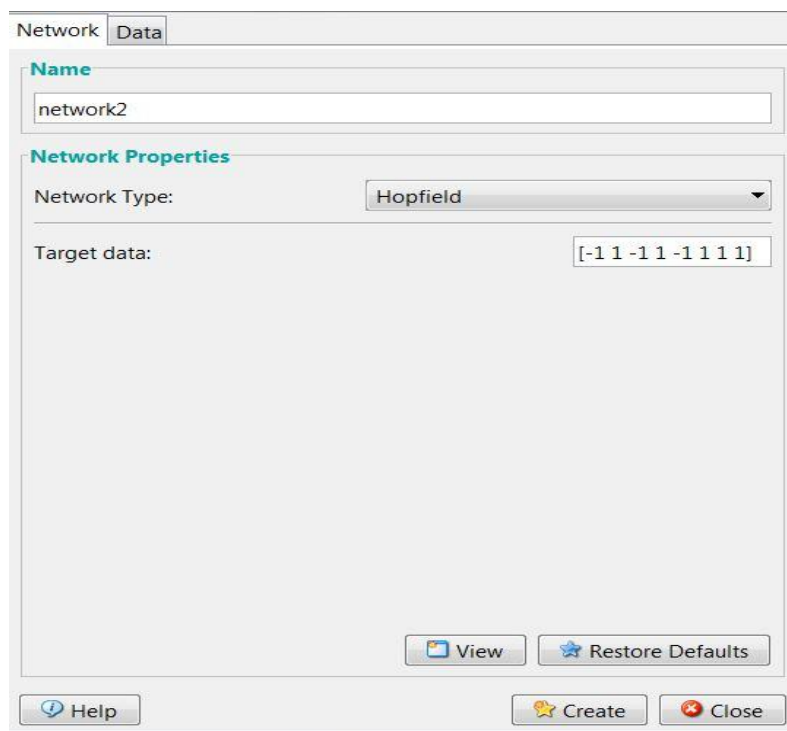
Енгізуді Create (Құру) батырмасын баса отырып аяқтаймыз.

Осыған ұқсас операцияны u векторына, бұл – мақсатты мәліметтер векторы, деп көрсете отырып (Targets радиобатырмасының көмегімен) жүргіземіз.

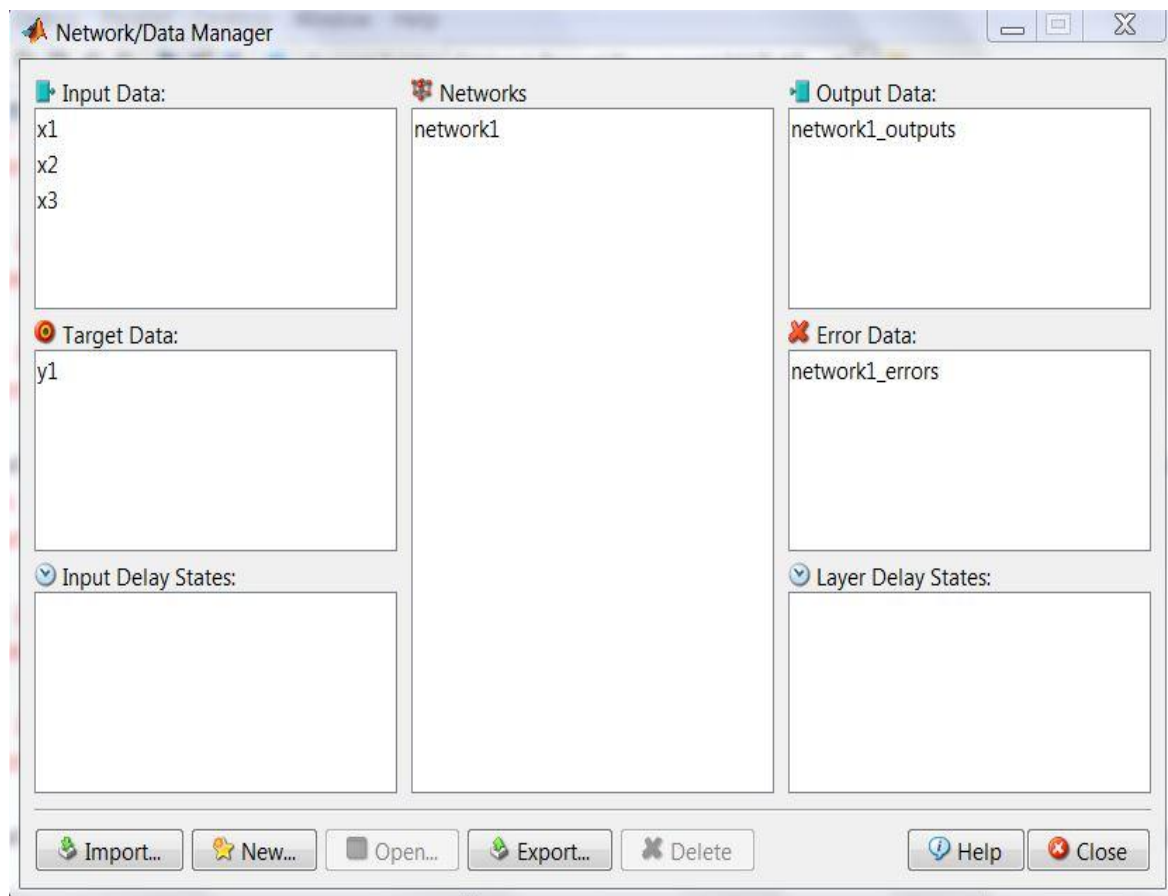
Жаңа нейронды желіні құрамыз.

Бұл үшін Network/Data Manager терезесінде New Network батырмасын басамыз. Ашылған Create New Network терезесінде сигналды тікелей жіберетін және қателіктің кері таралуына ие feed-forward backprop типіндегі нейронды желіні таңдаймыз. Желіні құру барысында әдеттегідей, Network1 бойынша берілетін атын сақтаймыз. Кірістер көлемін Get from input опциясының көмегімен анықтаймыз. Бірінші қабат (Layer 1) нейрондарының санын (Number of neurons) екіге тең деп орнатамыз. Желінің құрылу барысындағы қалған құрылымдарды әдепкі қалпы бойынша қалдырамыз. Желіні құруды Create батырмасын басумен аяқтаймыз.

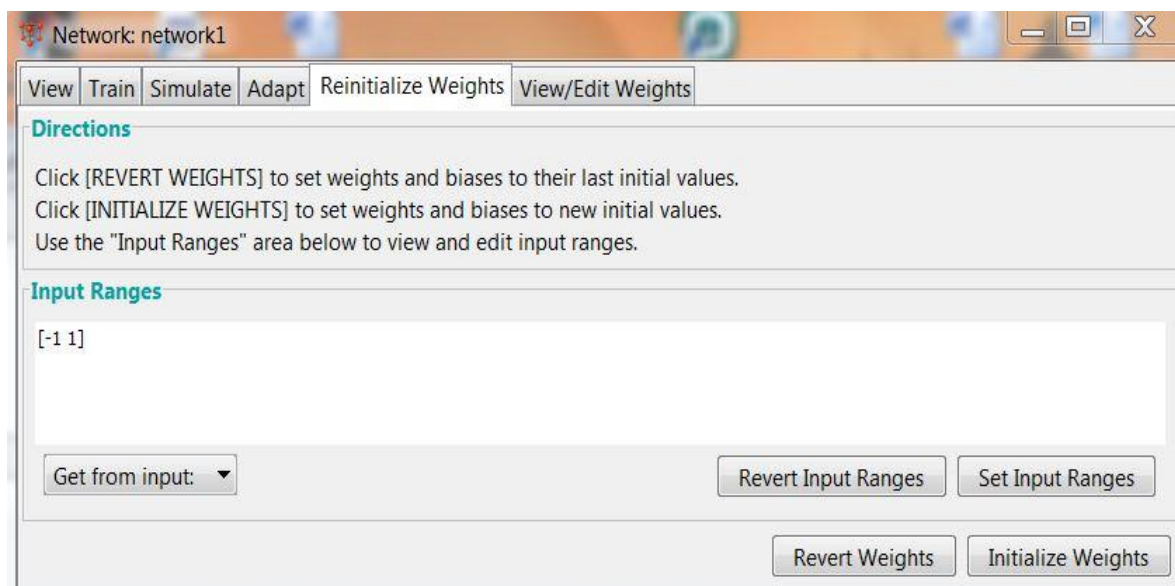
Осыдан кейін Network/Data Manager терезесіндегі Networks аймағында жаңадан құрылған – network 1 желісінің аты пайда болады. Осы атауды тышқанның көмегімен таңдаймыз, бұл көрсетілген терезедегі барлық батырмалардың активтенуіне алып келеді. Ары қарай Network: network1 диалогтық панелін ашамыз.



3.3-сурет – x және u кіріс векторларын енгізу

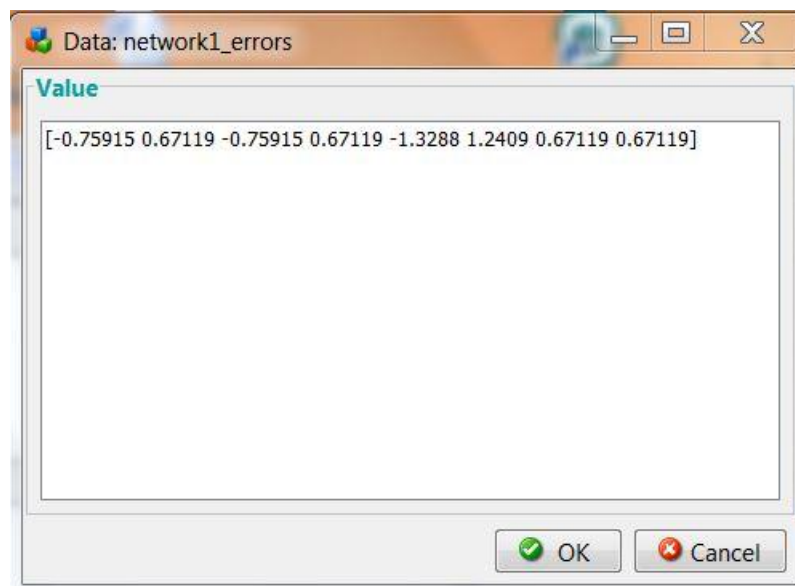


3.4-сурет – Network/Data Manager жалпы шолу



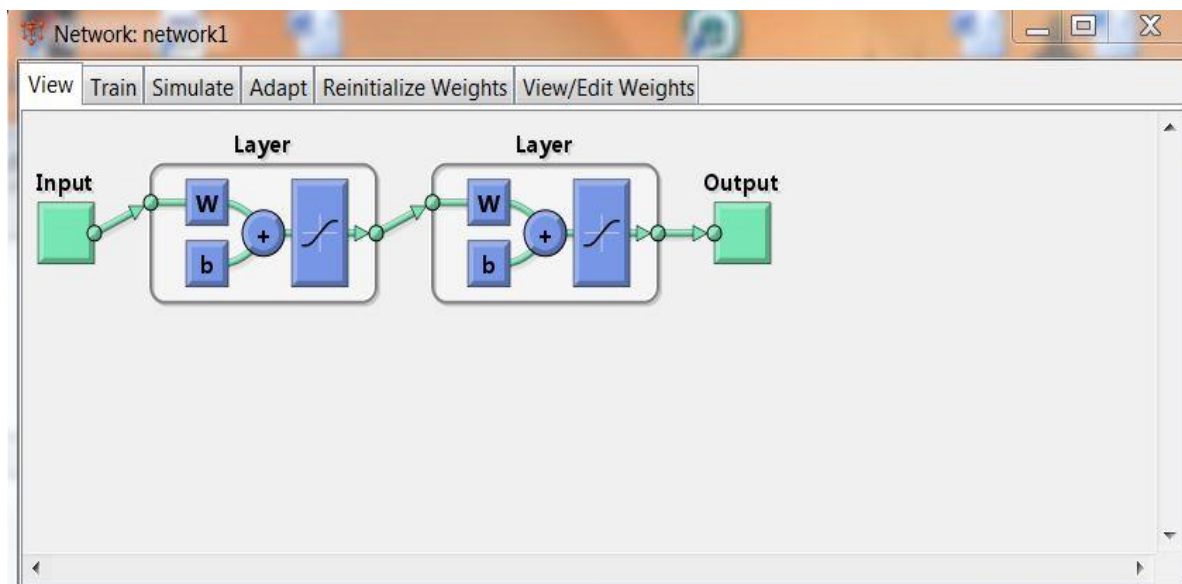
3.5-сурет – Орнатылған көлемдер

Үйрену нәтижелерін View батырмасын қолдана отырып, шығыстар немесе қателіктер реттілігінің атауларын активтеу арқылы Network/Data Manager терезесінде көруге болады.



3.6-сурет – Қате мәндерді шолу

Желінің құрылымдық сызбасын қарап шығу үшін network1 желісінің атауын тандап, View батырмасын қолдану қажет.



3.7-сурет – Нейронды желінің құрылымдық сызбасы

Жаңа дыбыстық сигнал x_{testx} желіге енгізіледі. Желінің күйлері асинхронды түрде жаңартылып, нәтижесінде ол алдын ала үйретілген үлгілердің біріне сәйкес келетін тұрақты күйге өтеді.

Бұл модель дыбыстарды тануды қамтамасыз етеді, егер енгізілген сигнал үйретілген үлгілерге ұқсас болса. Тіпті шу немесе сигналдың бұрмалануы жағдайында да желі дұрыс ассоциативті сәйкестендіру жүргізе алады.

ҚОРЫТЫНДЫ

Бұл дипломдық жұмыста дыбыстық сигналдарды танудың теориялық және практикалық аспектілері жан-жақты зерттелді. Жұмыстың негізгі мақсаты – нейрондық желілерді, атап айтқанда Хопфилд моделі мен әртүрлі жасанды нейрондық архитектураларды қолдану арқылы сөйлеу сигналдарын тану әдістерін талдап, модельдеу болды.

Зерттеу барысында дыбыстық сигналдардың физикалық және спектралдық сипаттамалары қарастырылып, сөйлеу процесіндегі коартикуляциялық, интонациялық және жеке ерекшеліктер сияқты күрделі факторларға талдау жасалды. Осы ерекшеліктерді ескере отырып, сөйлеуді тану процесінің негізгі үш кезеңі – алдын ала өңдеу, векторларды салыстыру және уақытша теңестіру мен шешім қабылдау процестері сипатталды.

Сонымен қатар, дыбыстық сигналдарды сандық өңдеу әдістері – PCM, ADM, CVSD және вокодерлер (RALCWI) – салыстырылып, олардың телекоммуникация мен әскери жүйелердегі қолданылу ерекшеліктері анықталды. CML Microcircuits компаниясының CMX649, CMX618 және CMX8341 микросхемаларының функционалды мүмкіндіктері сипатталып, олардың дыбысты тиімді өңдеудегі маңызы көрсетілді.

Жасанды нейрондық желілердің сөйлеу танудағы маңыздылығы дәлелденді: көп қабатты перцептрондар, бір қабатты желілер, векторлық кванттау әдістері мен Хопфилд желісіне негізделген классификация тәсілдері тәжірибеде қолдануға жарамды әрі тиімді екені көрсетілді. Бұл әдістер арқылы апостериорлық ықтималдықтарды бағалау, жергілікті ұқсастық метрикаларын есептеу және дауысты контексте тану сияқты міндеттер шешілді.

Жұмыс нәтижесінде дыбыстық сигналдарды тануға арналған нейрондық желі негізінде модель жасалып, MATLAB ортасында nntool құралымен оның имитациялық үлгісі құрылды және тиімділігі тексерілді. Алынған нәтижелер әзірленген әдістің нақты жүйелерде қолдануға жарамды екенін растады.

Жалпы алғанда, бұл зерттеу қазіргі заманғы дыбыстық тану жүйелерінің даму үрдістерін және нейрожелілік технологиялардың болашағын айқын көрсетті. Болашақта ұсынылған әдістер адамның дауысын тануға арналған автоматтандырылған және интеллектуалды жүйелерді дамытуда кеңінен қолданылады деп күтілуде.

ПАЙДАЛАНЫЛҒАН ӘДБИЕТТЕР ТІЗІМІ

- 1 Kim, J., & Park, D. (2022). Enhanced Speech Pattern Recognition Using Modified Hopfield Networks. *IEEE Transactions on Neural Networks and Learning Systems*, 33(4), 789–798.
- 2 Ahmed, S. (2021). *Advancements in Hopfield Neural Networks for Speech Signal Recognition*. PhD Dissertation, University of Technology.
- 3 Brown, R. L. (2021). *Advances in Hopfield Networks: Theory and Applications*. Wiley-IEEE Press.
- 4 Zhou, Y., & Zhang, L. (2020). Associative Memory and Pattern Completion with Hopfield Neural Networks. *Neural Computation*, 32(12), 2510–2531.
- 5 Nguyen, T. T. (2019). Hybrid Models Combining HMM and Hopfield Networks for Continuous Speech Recognition. *Pattern Recognition Letters*, 127, 41–48.
- 6 Wang, H., & Lee, J. (2018). Deep Learning Architectures in Hopfield Network-Based Audio Processing. *IEEE Access*, 6, 32421–32430.
- 7 Lopez, R. A. (2020). *Computational Models for Neuro-Inspired Audio Recognition*. Springer.
- 8 Patel, M., & Kumar, R. (2021). *Introduction to Hopfield Networks and Applications in Signal Processing*. CRC Press.
- 9 Chen, L. (2017). *Neural Network-Based Approaches to Speech Recognition*. Academic Press.
- 10 Ishikawa, M. (2016). *Advanced Neural Network Techniques for Pattern Recognition*. Elsevier Science.
- 11 Ghosh, D., & Roy, S. (2022). Hopfield Neural Networks for Time Series and Audio Signal Modeling. *Journal of Intelligent & Fuzzy Systems*, 42(2), 1701–1713.
- 12 Tanaka, K. (2018). *Associative Memory in Recurrent Neural Networks: Theory and Practice*. MIT Press.
- 13 Huang, Y. (2019). *Neural Network Applications in Speech and Audio Signal Processing*. IEEE Press/Wiley.
- 14 Santos, A., & Lima, C. (2020). *Neurocomputing Methods for Digital Speech Recognition*. Elsevier.

ҚОСЫМША А

Python бағдарламалық ортада жазылған Хопфилд нейрожелісінің коды

```
import numpy as np

class HopfieldNetwork:
    def __init__(self, n_units):
        self.n_units = n_units
        self.weights = np.zeros((n_units, n_units))

    def train(self, patterns):
        for p in patterns:
            p = np.array(p)
            self.weights += np.outer(p, p)
            np.fill_diagonal(self.weights, 0)
            self.weights /= len(patterns)

    def recall(self, pattern, steps=5):
        s = np.array(pattern)
        for _ in range(steps):
            for i in range(self.n_units):
                raw = np.dot(self.weights[i], s)
                s[i] = 1 if raw >= 0 else -1
        return s

# Үлгі шаблондары (мысалы, 3x3 тор түрінде)
patterns = [
    [1, -1, 1,
     -1, 1, -1,
     1, -1, 1],

    [-1, 1, -1,
     1, -1, 1,
     -1, 1, -1]
]
# Желі құру және үйрету
hopfield = HopfieldNetwork(n_units=9)
hopfield.train(patterns)

# Тануға ұсынылатын бұрмаланған үлгі
test_pattern = [1, -1, 1,
                 -1, -1, -1,
                 1, -1, 1]

# Желі тануы
result = hopfield.recall(test_pattern)
print("Input: ", test_pattern)
print("Recognized: ", result.reshape(3, 3))
```

СЫН ШҚІР
дипломдық жұмысқа

Мұратбек А.С.

6B06201 «Телекоммуникация» білім беру бағдарламасы

Тақырыбы: «Хопфилд нейрожелісін қолданып, дыбыс сигналын танудың
математикалық моделі»

Орындалды:

а) графикалық бөлім 3 парақ;
б) түсініктеме 57 бет.

ЖҰМЫСҚА ЕСКЕРТУ

Диплом жұмысы дыбыстық сигналдарды цифрлық өңдеу мен тану әдістерін зерттеуге арналған. Жұмыста нейрондық желілердің, атап айтқанда Хопфилд моделінің сөйлеуді танудағы мүмкіндіктері қарастырылып, MATLAB ортасында имитациялық модель жасалған. Сигнал өңдеудің PCM, ADM, CVSD, вокодерлер сияқты әдістері талданып, олардың қолдану салалары сипатталған.

Жасалған модель тәжірибе жүзінде тиімділігін көрсетіп, сөйлеуді тану процесінің негізгі кезеңдері — алып ала өңдеу, векторлық салыстыру, уақытша теңестіру — нақты сипатталған. CML Microcircuits микросхемаларының қолданылуы жұмыстың практикалық маңызын арттырады.

Жұмыстың ғылыми мазмұны жоғары деңгейде орындалған. Дегенмен, материалдың кейбір бөлімдерінде терминологиялық қайталауларды азайту және сөйлем құрылымын ықшамдау арқылы мәтіннің оқылымдылығын арттыруға болады. Сондай-ақ, тәжірибелік нәтижелердің графикалық түрде (гистограмма, диаграмма, спектрлік көрініс) ұсынылуы жұмыстың көрнекілігін арттырар еді.

Жалпы, диплом жұмысы ғылыми және қолданбалы тұрғыдан өте құнды. Автор зерттеу барысында өз бетімен жұмыс істеу, аналитикалық ойлау және техникалық модельдеу дағдыларын сәтті көрсете білген. Жұмыс қорғауға толық ұсынылады.

Жұмысты бағалау

Жалпы, дипломдық жұмысқа «өте жақсы» (95%) деген баға, ал студент Мұратбек Аяулымды 6B06201 – Телекоммуникация білім беру бағдарламасы бойынша «Ақпараттық коммуникациялық технологиялар бакалавры» дәрежесіне ұсынамын.

Рецензент

ҚазҰАЗУ, доктор PhD., ғылым
қауымд. профессоры

Әлібек П.Б.
«20» 05 2025 ж.

**Университеттің жүйе администраторы мен Академиялық мәселелер департаменті
директорының ұқсастық есебіне талдау хаттамасы**

Жүйе администраторы мен Академиялық мәселелер департаментінің директоры көрсетілген еңбекке қатысты дайындалған Плагиаттың алдын алу және анықтау жүйесінің толық ұқсастық есебімен танысқанын мәлімдейді:

Автор: Мұратбек Аяулым Сұлтанбек

Тақырыбы: Хопфилд нейрожелісін қолданып, дыбыс сигналын танудың математикалық моделі

Жетекшісі: Кырмызы Тайсариева

1-ұқсастық коэффициенті (30): 7

2-ұқсастық коэффициенті (5): 4.1

Дәйексөз (35): 1.3

Әріптерді ауыстыру: 8

Аралықтар: 0

Шағын кеңістіктер: 1

Ақ белгілер: 0

Ұқсастық есебін талдай отырып, Жүйе администраторы мен Академиялық мәселелер департаментінің директоры келесі шешімдерді мәлімдейді :

☒ Ғылыми еңбекте табылған ұқсастықтар плагиат болып есептелмейді. Осыған байланысты жұмыс өз бетінше жазылған болып санала отырып, қорғауға жіберіледі.

☐ Осы жұмыстағы ұқсастықтар плагиат болып есептелмейді, бірақ олардың шамадан тыс көптігі еңбектің құндылығына және автордың ғылыми жұмысты өзі жазғанына қатысты күмән тудырады. Осыған байланысты ұқсастықтарды шектеу мақсатында жұмыс қайта өңдеуге жіберілсін.

☐ Еңбекте анықталған ұқсастықтар жосықсыз және плагиаттың белгілері болып саналады немесе мәтіндері қасақана бұрмаланып плагиат белгілері жасырылған. Осыған байланысты жұмыс қорғауға жіберілмейді.

Негіздеме:

2025-04-29

Күні



Кафедра меңгерушісі



Протокол

о проверке на наличие неавторизованных заимствований (плагиата)

Автор: Мұратбек Аяулым Сұлтанбек

Соавтор (если имеется):

Тип работы: Дипломная работа

Название работы: Хопфилд нейрожелісін қолданып, дыбыс сигналын танудың математикалық моделі

Научный руководитель: Кырмызы Тайсариева

Коэффициент Подобия 1: 7

Коэффициент Подобия 2: 4.1

Микропробелы: 1

Знаки из других алфавитов: 8

Интервалы: 0

Белые Знаки: 0

После проверки Отчета Подобия было сделано следующее заключение:

☒ Заимствования, выявленные в работе, является законным и не является плагиатом. Уровень подобия не превышает допустимого предела. Таким образом работа независима и принимается.

☐ Заимствование не является плагиатом, но превышено пороговое значение уровня подобия. Таким образом работа возвращается на доработку.

☐ Выявлены заимствования и плагиат или преднамеренные текстовые искажения (манипуляции), как предполагаемые попытки укрытия плагиата, которые делают работу противоречащей требованиям приложения 5 приказа 595 МОН РК, закону об авторских и смежных правах РК, а также кодексу этики и процедурам. Таким образом работа не принимается.

☐ Обоснование:

2025-04-29

Дата



Заведующий кафедрой



Протокол

о проверке на наличие неавторизованных заимствований (плагиата)

Автор: Мұратбек Аяулым Сұлтанбек

Соавтор (если имеется):

Тип работы: Дипломная работа

Название работы: Хопфилд нейрожелісін қолданып, дыбыс сигналын танудың математикалық моделі

Научный руководитель: Кырмызы Тайсариева

Коэффициент Подобия 1: 7

Коэффициент Подобия 2: 4.1

Микропробелы: 1

Знаки из здругих алфавитов: 8

Интервалы: 0

Белые Знаки: 0

После проверки Отчета Подобия было сделано следующее заключение:

☒ Заимствования, выявленные в работе, является законным и не является плагиатом. Уровень подобия не превышает допустимого предела. Таким образом работа независима и принимается.

☐ Заимствование не является плагиатом, но превышено пороговое значение уровня подобия. Таким образом работа возвращается на доработку.

☐ Выявлены заимствования и плагиат или преднамеренные текстовые искажения (манипуляции), как предполагаемые попытки укрытия плагиата, которые делают работу противоречащей требованиям приложения 5 приказа 595 МОН РК, закону об авторских и смежных правах РК, а также кодексу этики и процедурам. Таким образом работа не принимается.

☐ Обоснование:

2025-04-29

Дата



Сұңғат Марксұлы

проверяющий эксперт