

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті

Ақпараттық және телекоммуникациялық технологиялар институты

Электроника, телекоммуникациялар және ғарыштық технологиялар кафедрасы

ҚОРҒАУҒА ЖІБЕРІЛДІ

Кафедра меңгерушісі

техн. ғыл. канд-ы

 Е. Т. Таштай

«29» 04 2019 ж.

ДИПЛОМДЫҚ ЖҰМЫС

Тақырыбы: «Сөйлеу сигналын кодтаудың оңтайлы әдістерінің анализі»

5B071900 – «Радиотехника, электроника және телекоммуникация» мамандығы

Орындаған:



Канатбекова Д.А.

Пікір беруші

ҚазҰАУ, ЭҮЖА каф. меңгерушісі

доктор PhD.,

қауымдастырылған профессор

 Ж. С. Шыныбай

«29» 04 2019 ж.

Ғылыми жетекші

ЭТЖҒТ кафедрасының ассоц.

профессоры

 Д.Б. Илипбаева

«26» 04 2019 ж.

Алматы 2019

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті

Ақпараттық және телекоммуникациялық технологиялар институты

Электроника, телекоммуникация және ғарыштық технологиялар кафедрасы

5B071900 – Радиотехника, электроника және телекоммуникация

БЕКІТЕМІН

Кафедра меңгерушісі,
техн.ғыл.канд.

 Е.Таштай

«20/» 01 2019 ж.

**Дипломдық жұмыс орындауға
ТАПСЫРМА**

Білім алушы Канатбекова Диана

Тақырыбы «Сөйлеу сигналын кодтаудың оңтайлы әдістерінің анализі»

Университет ректорының «16» қазан 2018 ж. № 1162-б бұйрығымен бекітілген.

Аяқталған жұмысты тапсыру мерзімі “25” сәуір 2019 ж.

Дипломдық жұмыстың бастапқы берілістері:

1) Сигналдың жиілігі $S = 4\text{кГц}$ болғанда, дискретизация жиілігі $F_s = 8\text{кГц}$;

2) Кванттаудың қадам өлшемі q ;

3) $P(x) = 1, 2, 3$ болғанда шектік мәні q мен кванттау шуының қуаты;

Дипломдық жұмыста қарастырылатын мәселелер тізімі:

а) Сөйлеу сигналын түрлендіру құрылымымен танысу;

ә) Сөйлеу сигналын кодтаудың негізгі әдістерін қарастыру;

б) Сөйлеу сигналының параметрлерін зерттеу;

в) Сөйлеу сигналын бастапқы кодтау әдістеріне есептеу жүргізу;

г) Сөйлеу сигналын Matlab бағдарламасында модельдеу.

Сызбалық материалдар тізімі (міндетті сызбалар дәл көрсетілуі тиіс)

Сөйлеу сигналын түрлендіруге қажетті аналогты сигналдың сандық сигналға түрлену схемасы. Сөйлеудің түрлену моделінің құрылымы көрсетілген. Ұсынылатын негізгі әдебиет 23 атау.

1. Dr.Shaila, D.Apte Speech and Audio Processing. New Delhi,2017

2. Солонина А.И Основы цифровой обработки сигналов, 2012

3. Ю.А.Ковалгин, Э.И. Вологдин Цифровое кодирование звуковых сигналов. Спб,2015

4. Н.К Верещагин Информация, кодирование и предсказание. Москва,2012

дипломдық жұмысты (жобаны) дайындау

КЕСТЕСІ

Бөлімдер атауы, қарастырылатын мәселелер тізімі	Ғылыми жетекшіге және кеңесшілерге көрсету мерзімі	Ескерту
Сөйлеу сигналын түрлендіру құрылымымен танысу	20.01.2019 - 01.03.2019	<i>орындалған</i>
Сөйлеу сигналын кодтаудың негізгі әдістерін қарастыру	02.03.2019-02.04.2019	<i>орындалған</i>
Сөйлеу сигналын Matlab бағдарламасында модельдеу	01.04.2019–15.04.2019	<i>орындалған</i>

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
Норма бақылау	PhD докторы, ЭТЖҒТ каф.сениор-лекторы Тайсариева К.Н.	29.04.19	<i>[Signature]</i>

Ғылыми жетекшісі *[Signature]* Л.Б. Илипбаева
(қолы)

Тапсырманы орындауға алған білім алушы *[Signature]* Д. Канатбекова

Күні “29” 04 2019 ж.

АҢДАТПА

Берілген дипломдық жұмыста сөйлеу сигналын кодтаудың оңтайлы әдістері қарастырылды. Сөйлеуді кодтау үшін аналогты сигналдың сандық сигналға түрлену схемасы және сөйлеудің түрлену моделінің құрылымы анықталды. Сөйлеу сигналын кодтаудың сызықты болжау әдісіне және оның түрлеріне аналитикалық шолу жасалынды. Сөйлеу сигналын кодтаудың негізгі параметрлері бойынша бастапқы кодтау әдістеріне есептеу жүргізілді. Сөйлеу сигналының уақыттық, спектральдық және кодтық сипаттамаларының моделі алынды. Қазақ тілінде «Даяшы» сөзінің сипаттамалары анықталып, эмоционалдық көрінісінің өзгеруі бақыланды.

АННОТАЦИЯ

В данной дипломной работе рассмотрены оптимальные методы кодирования речевого сигнала. Для кодирования речи была определена схема преобразования аналогового сигнала в цифровой сигнал и структура модели преобразования речи. Проведен аналитический обзор методов линейного предсказания кодирования речевого сигнала и его видов. По основным параметрам кодирования речевого сигнала произведен расчет методов первичного кодирования. Получены модели временных, спектральных и кодовых характеристик речевого сигнала. Были определены характеристики казахского слова «Даяшы» и наблюдались изменения эмоционального характера.

ABSTRACT

In this thesis we consider the best methods of coding the speech signal. To encode the speech, the scheme of converting the analog signal into a digital signal and the structure of the speech conversion model were determined. The analytical review of methods of linear prediction of speech signal coding and its types is carried out. The calculation of the primary coding methods was carried out according to the main parameters of the speech signal coding. Models of time, spectral and code characteristics of the speech signal are obtained. The characteristics of the Kazakh word were determined «Даяшы» and there were emotional changes.

МАЗМҰНЫ

Кіріспе	9
1. Сөйлеу сигналын түрлендірудің қысқаша сипаттамасы	10
1.2 Сөйлеу сигналын кодтаудың сызықты болжау әдісі	13
1.3 Сызықты болжауы бар вокодерлерді қолдану негізінде сөйлеуді талдау әдістері	16
1.4 Негізгі тонның жиілігінен қоздырылатын LPC кодтау әдісі	17
1.5 RELP немесе сөйлеуді гибридті кодтау әдісі	18
1.6 RELP негізінде ABS кодтау әдісі	20
1.7 Кодтық кітаптар негізінде қозуы бар кодтар немесе CELP кодтау әдісі	21
2. Импульстік-кодтық модуляция	25
2.1 Дифференциалды импульстік-кодтық модуляция	25
2.2 Сөйлеу сигналын дискреттеу	26
2.3 Сөйлеу сигналын кванттау	28
2.4 Біркелкі және біркелкі емес кванттаушы және кодер	30
2.5 Сөйлеу сигналының кванттау бойынша параметрлерін есептеу	33
2.6 DPCM бойынша параметрлерді есептеу	35
3. Matlab бағдарламасында сөйлеу сигналының параметрлерін зерттеу	38
3.1 MATLAB бағдарламасында «Даяшы» сөзіне зерттеу жүргізу	38
Қорытынды	44
Пайдаланылған әдебиеттер тізімі	45

КІРІСПЕ

Сөз - бұл адам өмірінің түрлі салаларында ақпарат алмасудың тәсілі. Қазіргі заманда сөйлеу сигналын әзірлеу, талдау, кодтау және сөйлеу синтезі саласындағы дәреженің айтарлықтай өсуі байқалады. Сөз алмасу технологиясын кеңінен қолданудың және дамытудың бірден бір жолы сөйлеу сигналын кодтаудың тиімді әдістерін падалану болып табылады. Бұл біріншіден кезекте заманауи желілер мен пакеттік коммутацияның желілік технологиялар үшін өзекті болып келеді.

Жұмыстың мақсаты – сөйлеу сигналын кодтаудың оңтайлы әдістерін қолданып, сөйлеу сигналының негізгі кодтау әдістерін сараптап талдау.

Сөйлеуді кодтаудың негізгі әдістерін қарастыру алдында жалпы сөйлеу сигналының түрлену құрылымы қарастырылады. Сөйлеуді кодтау үшін аналогты сигналдың сандық сигналға түрлену процесі және сөйлеуді түрлендіру моделінің құрылымы көрсетіледі. Дауысты және дауыссыз дыбыстырдың айырмашылықтары мен ерекшеліктері олардың фрагменттері мен спектралды көрінісі арқылы талданады.

Сөйлеу сигналын кодтаудың оңтайлы әдістерінің бірі – сызықты болжау әдісі. Бұл әдіс – негізгі тонның периоды, форманттар, спектр, сөйлеу трактісінің ауданының функциясы, сондай-ақ, төмен жылдамдықты беру және үнемді сақтау мақсатында сөйлеуді қысқартылған түрде көрсету кезінде және сөйлеу сигналының негізгі параметрлерін бағалау кезінде басымдылық көрсетеді. Дипломдық жұмыста сызықты болжау әдісінің негізгі түрлері қарастырылады, олар: негізгі тонның жиілігінен қоздырылатын LPC кодтау әдісі, RELP немесе сөйлеуді гибриді кодтау әдісі, көпимпульсті кодер негізіндегі кодтау әдісі, тұрақты импульстік қозуы бар кодектер негізіндегі кодтау әдісі, кодтық кітаптар негізінде қозуы бар CELP кодтау әдісі, PCM және дифференциалды DPCM кодтау әдістері[1].

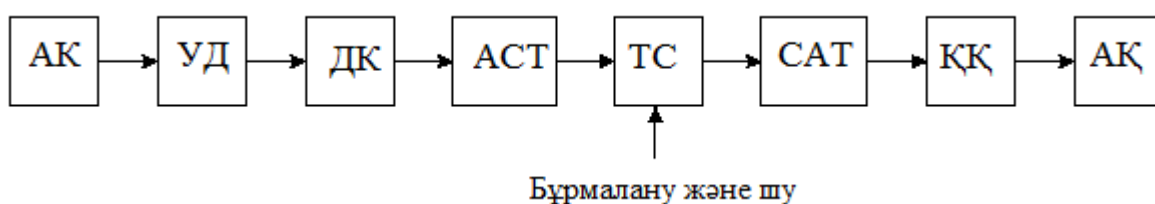
Сөйлеу сигналын кодтаудың негізгі параметрлеріне зерттеу жүргізіледі. Есептеулер сигналды дискреттеу мен кванттау параметрлерін есептеуге негізделген. Тарауда әртүрлі кванттау әдістері, біркелкі кванттау, дискретті кванттаудың біркелкі емес әдістері, дифференциалды кодтау әдістер негізінде келесі кванттаудың қадам өлшемі, кванттау шуы және кванттау шуының дисперсиясы, сигнал дисперсиясы, дифференциалды конфигурация үшін күшейту коэффициенті сияқты параметрлерге есептеу жүргізіледі[1].

MATLAB бағдарламасы көмегімен сөйлеу сигналын зерттеліп, сөйлеу сигналының уақыттық, спектральдық және кодтық сипаттамаларының моделі алынған. Қазақ тілінде алынған сөздің сипаттамалары анықталып, эмоционалдық көрінісі бойынша өзгерісі бақыланады.

1 Сөйлеу сигналын түрлендірудің қысқаша сипаттамасы

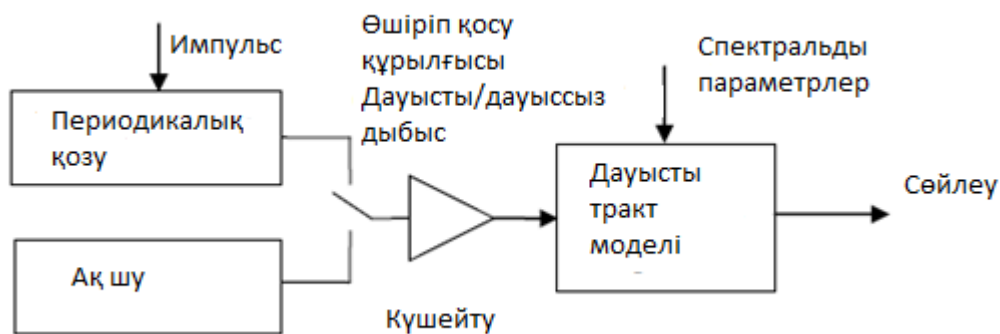
Соңғы жылдары сөйлеу сигналын әзірлеу, талдау, кодтау және сөйлеу синтезі саласындағы дәреженің айтарлықтай өсуі байқалады. Бұл негізінен адам қызметінің түрлі салаларында ақпараттық технологияларды кеңінен пайдаланумен, процестерді автоматтандырумен, ақпарат алмасумен және ақпараттық коммуникациялық технологияларды пайдаланушылар мен машиналық технологияларды кеңінен пайдаланумен байланысты.

Төменде сөйлеу сигналын түрлендіруге қажетті аналогты сигналдың сандық сигналға түрлену схемасы және жалпы сөйлеудің түрлену моделінің құрылымы көрсетілген (Сурет 1.1 және 1.2)[2].



Сурет 1.1 – Аналогты сигналдың сандық сигналға түрлену схемасы

Мұндағы, АК- үзіліссіз сигнал түріндегі ақпарат көзі, УД –уақыт бойынша дискретизатор, ДК- деңгей бойынша кванттаушы, АСТ-аналогты-сандық түрлендіргіш, ТС- тарату сызығы, САТ-сандық-аналогты түрлендіргіш, ҚҚ- дискретті аналогты сигналды үзіліссіз сигналға қалпына келтіру құрылғысы, АҚ- ақпарат қабылдағыш.

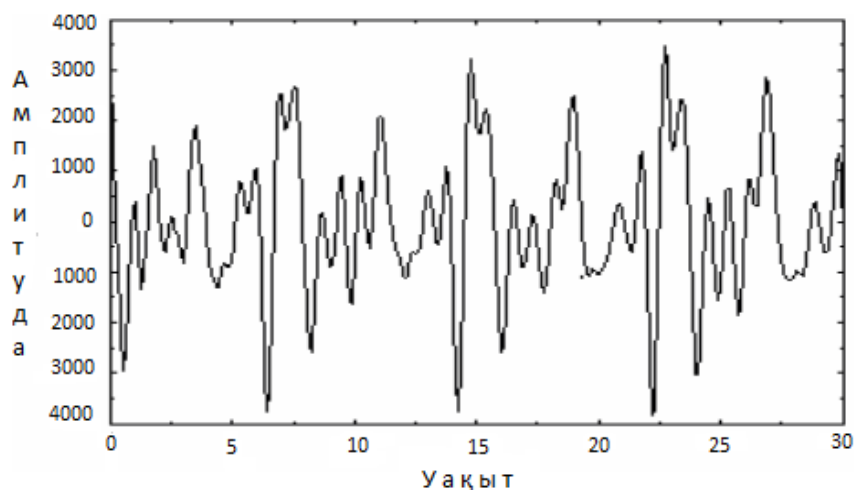


Сурет 1.2 – Сөйлеудің түрлену моделінің құрылымы

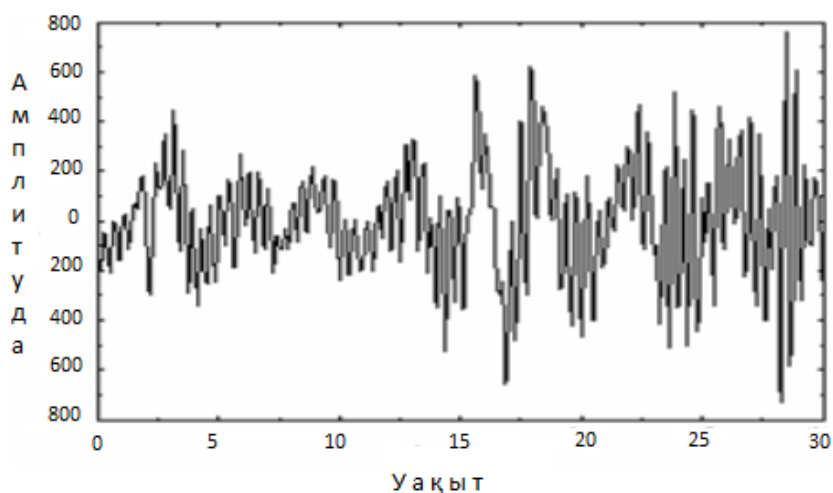
Сөз айтылатын сөздерге, дауыстың тембрына, интонацияға, сөйлеушінің жынысы мен жасына байланысты болатын тербелістің күрделі формасы болып табылады. Сөйлеу спектрі өте кең (шамамен 50-ден 10000 Гц-ке дейін), бірақ телефонияда сөйлеуді тарату үшін 0,3 - 3,4 кГц жиілік жолағынан тыс жатқан құрамдастардан бас тартты, себебі бұл жағдайда дыбыстардың қабылдануы

біршама нашарлады (мысалы, шуыл дыбыстардың энергиясының елеулі бөлігі сөйлеу спектрінің жоғарғы бөлігінде шоғырланған). Бірақ бұл жағдай сөз анықтылығына қатты әсер берген жоқ. Сонымен қатар, негізгі тонның төменгі жиілікті гармоникаларының жоғалуына байланысты төменгі жиілікті шектеу (300 Гц-ке дейін) қабылдауды біршама нашарлатады.

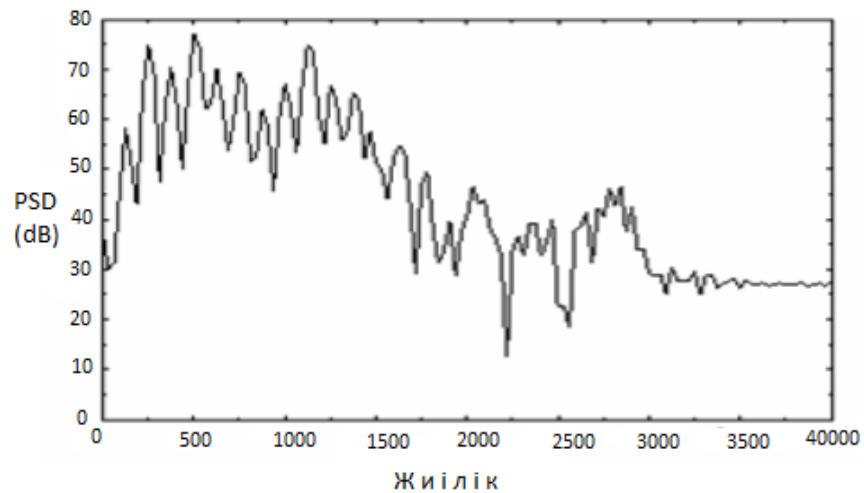
Төменде сөйлеу сигналының фрагменттері бейнеленген. Бұл суреттерде дауысты (Сурет 1.3) және дауыссыз (Сурет 1.4) дыбыстар, сондай-ақ, осы дыбыстардың спектрлері (Сурет 1.5 және 1.6) бейнеленген. Суреттерде сигналдардың сипаты жағынан айырмашылығы және бірінші және екінші жағдайларда сигнал спектрінің ені 3,5 кГц-тен аспайтыны анық байқалады. Сөйлеу сигналы спектрінің құрамындағы төменгі жиілікті (яғни, уақыт бойынша баяу) құраушылардың деңгейі жоғары жиілікті (жылдам) құраушылардың деңгейінен айтарлықтай жоғары екенін атап өтуге болады. Бұл сигнал спектрінің бірқалыпты болмауы және сигнал факторларының бірі болып табылады.



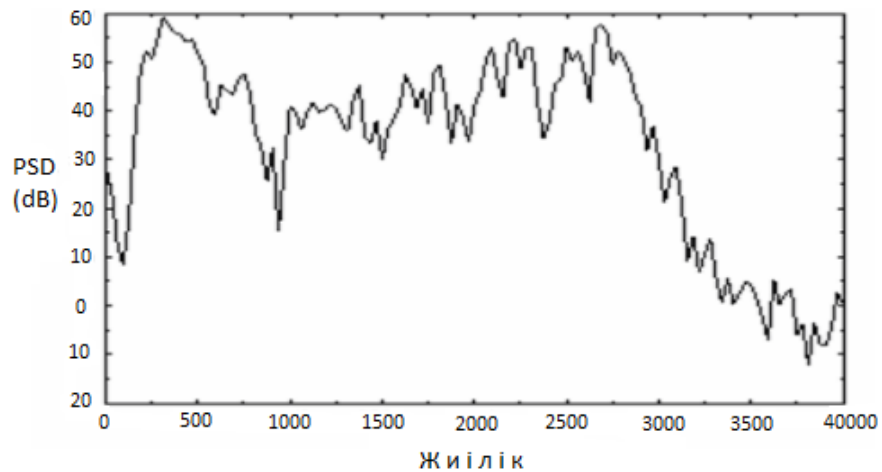
Сурет 1.3 – Дауысты дыбыстарды қамтитын сөйлеу сигналының фрагменті



Сурет 1.4 – Дауыссыз дыбыстарды қамтитын сөйлеу сигналының фрагменті



Сурет 1.5 – Дауысты дыбыстарды қамтитын сөйлеу сигналының спектрі



Сурет 1.6 – Дауыссыз дыбыстарды қамтитын сөйлеу сигналының спектрі

Сөйлеу сигналының біркелкі бөлінбеуінен, сигналдың үлкен деңгейлеріне қарағанда кіші деңгейлерінің орын алу ықтималдығы жоғары болуы мүмкін. Әсіресе, бұл сөйлеу белсенділігі жоғары емес ұзақ фрагменттерде байқалады[3].

Сөйлеу сигналы уақыт бойынша стационарлы болмауы мүмкін: әртүрлі бөліктерде сигналдың қасиеттері мен параметрлері айтарлықтай ерекшеленеді. Бұл ретте стационар аралығының көлемі шамамен бірнеше ондаған миллисекундты құрайды. Сөйлеу дауыстық байламдар мен дауыстық тракт арқылы жеңіл ауа ағынының өтуі кезінде қалыптасады.

Дауыстық тракт дауыстық байламдардан басталып ерінмен аяқталады және орташа ұзындығы шамамен 15-17 сантиметрді құрайды. Дауыстық тракт өзінің резонанстық қасиеттеріне байланысты қалыптасқан сигналға форманттар деп аталатын және дауысқа тембрлік бояу беретін әрбір адамға тән жиіліктік құрамдастардың жиынтығын енгізеді. Формант жиіліктері мен жолақтарын

басқаруды дауыстық тракт пішінін өзгерту арқылы жүзеге асыруға болады, мысалы, тіл жағдайын өзгерту арқылы дауыстық тракт пішіні өзгереді. Көптеген дауыстық кодер/декодерлердің маңызды бөлігі болып дауыстық трактты модельдеу болып табылады. Дауыстық тракт пішінінің салыстырмалы түрде баяу өзгеруін ескерсек (тілдің жағдайын секундына 20-30 реттен жиі өзгеруі мүмкін емес), онда мұндай сүзгінің параметрлері де салыстырмалы түрде сирек (әдетте әрбір 20 миллисекунд немесе тіпті сирек) жанартылуы тиіс[4].

Осылайша, дауыстық тракт оған дауыстық байламдар арқылы жіберілетін ауа ағынымен қозғалады. Қозу тәсіліне байланысты бұл кезде туындайтын дыбыстарды шартты түрде үш сыныпқа бөлуге болады:

1. Дауыстық байламдардың дірілі кезінде (вибрация) пайда болатын дауысты дыбыстар сол байламдардың ашылып қайта жабылуы кезінде өкпеден дауыстық трактқа ауа ағыны үзіліп жетуі кезінде пайда болады. Бұл ретте дауыс трактісінің қозуы квазипериодикалық импульстермен жүргізіледі. Байламдарды ашу және жабу жылдамдығы (жиілігі) туындайтын дыбыстың (тонның) биіктігін анықтайды. Ол дауыстық байламдардың пішіні мен кернеуінің өзгеруімен, сондай-ақ, ауа ағынының қысымының өзгеруімен басқарылуы мүмкін. Дауысты дыбыстар 2 - 20 мс кезеңімен негізгі тонның кезеңділігінің жоғары дәрежесіне ие.

2. Ауа ағыны арқылы ашық дауыстық байланыстырғыштар арқылы жоғары жылдамдықпен өтетін шу тәріздес турбулентті ағынмен қозғалғанда пайда болатын дауыссыз дыбыстар. Мұндай дыбыстарда дауыстық байламдардың дірілімен негізделген ұзақ мерзімді кезең іс жүзінде жоқ.

3. Ауаның артық қысымы әрекетінен, жабық дауыстық тракты кенеттен ашылғанда пайда болатын жарылыс сипатындағы дыбыстар.

Кейбір дыбыстар таза түрде жоғарыда сипатталған сыныптардың ешқайсысына сәйкес келмейді, бірақ олардың қоспасы ретінде қарастырылуы мүмкін.

Жоғарыда көрсетілген 1.3 және 1.4 суреттерден көріп тұрғандай, сөйлеу сигналы дауыстық байламдар тербелісінің мерзімділігіне және дауыстық трактының резонанстық қасиеттеріне байланысты қысқа мерзімді және ұзақ мерзімді болжамдылықтың жоғары дәрежесіне ие және бұл жүйе үшін ең қолайлы болып табылады. Кодер және декодерлердің көп бөлігінде осы болжамдылық қолданылады, сондай-ақ, код жылдамдығын азайту мақсататында сөз түзу жүйесі моделінің параметрлерінің аз жылдамдығы қолданылады.

1.2 Сөйлеу сигналын кодтаудың сызықты болжау әдісі

Сөйлеуді қысудың неғұрлым күрделі әдістері сигнал формасының кодтау элементтерімен бірге сөйлеуді сызықтық болжау әдісін қолдануға негізделген.

Бұл алгоритмдерде сигнал беру кезінде кодты оңтайландыру жүзеге асырылатын кері байланысты кодтау қолданылады. Сигналды кодтау арқылы процессор оның пішінін қалпына келтіруге тырысады және нәтижені бастапқы сигналмен салыстырады, содан кейін ең жақсы сәйкестікке қол жеткізе отырып, кодтау параметрлерін өзгертуді бастайды. Мұндай сәйкестікке жеткеннен кейін аппаратура алынған кодты байланыс желісі бойынша береді. Қарама-қарсы ұшында сөйлеу сигналын қалпына келтіру жүреді. Мұндай әдісті пайдалану үшін маңызды есептеу қуаттары қажет екені анық[5].

Бұл әдіс – негізгі тонның периоды, форманттар, спектр, сөйлеу трактысының ауданының функциясы, сондай-ақ, төмен жылдамдықты беру және үнемді сақтау мақсатында сөйлеуді қысқартылған түрде көрсету кезінде және сөйлеу сигналының негізгі параметрлерін бағалау кезінде басымдылық көрсетеді. Бұл әдіс алынған бағалардың жоғары дәлдігімен және есептеулердің салыстырмалы қарапайымдылығымен ерекшеленеді.

Сызықтық болжау – бұл сандық цифрлы сөйлеуді сүзуге негізделген талдау әдісі, бұл кезде сигналдың ағымдағы есебі шығыс реттіліктің өткен және қазіргі мәндерінің сызықты комбинациясына, содай-ақ, кіріс реттіліктің өткен мәндерін «алдын ала» болжануына мүмкіндік береді (мысалы, сөйлеудің автоматты синтезі кезінде).

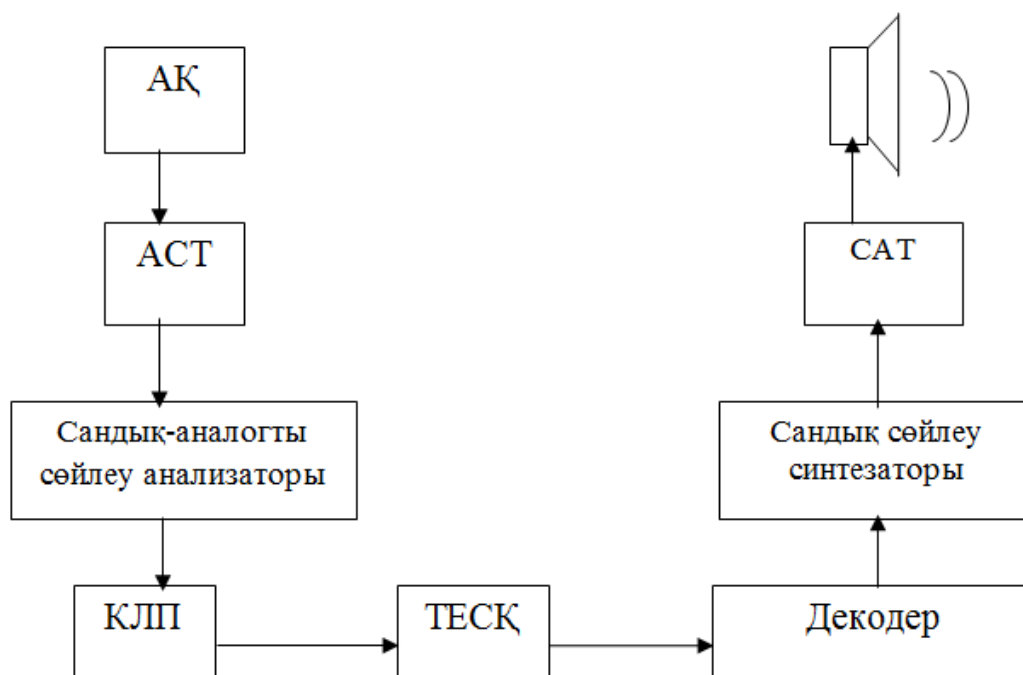
«Сызықтық комбинация» ұғымы белгісіз шығу есебін болжау (айқындау) үшін сызықтық болжамның тиісті коэффициенттеріне көбейтілген сигналдың белгілі дискретті есептеулері (кіріс және шығыс) туындыларының сомасын білдіреді. Сызықтық болжау кезінде сөйлеуді талдаудың негізгі міндеті – сигналды талдау бөлігінде болжаудың ең аз қатесін беретін осы сызықтық комбинацияның коэффициентін табу болып табылады.

Сөйлеуді сызықтық болжау коэффициентімен кодтау уақыт қатарының статистикалық талдау теориясына негізделеді. Уақытша қатар – бұл уақыт бойынша реттелген бақылау (есептеу) реттілігі.

Берілген әдістің мәнін түсіну үшін келесідей жағдай қарастырылады. Сөйлеу сигналын есептеу (таңдау) тізбегі бар болсын: x_1, x_2, \dots, x_t . Берілген таңдау үшін «өлшенген орташа» мән есептелінеді. 1020 мс аралығы үшін сөйлеу сигналының статистикалық қасиеттері өзгеріссіз деп саналады. Бұл аралықты болжаудың орташа квадраттық қатесін азайтатын a_s коэффициенттер жиынтығымен кодтайды, яғни бастапқы және тегістелген қатар арасындағы болжам қателігін азайтады. Коэффициенттерді есептеу өте қиын процесс болып табылады[6].

Сызықтық болжау коэффициенттерінің талдауы бойынша 1020 кГц жиілігімен келе жатқан есептерден құралған ИКМ сөйлеу көрінісі 50100 Гц жиілігімен келе жатқан параметрлер векторларының реттілігіне айналады. Бұл сөйлеу сапасы жақсы болған жағдайда сөйлеу сипаттамасын 50-100 рет қысады.

Төменде сызықтық болжау коэффициентін кодтаудың құрылымдық схемасы көрсетілген. Мұнда, АҚ-ақпарат көзі, КЛП - клиппирлеу, ТЕСҚ-тұрақты есте сақтау құрылғысы[7].



Сурет 1.7 – Сызықтық болжау коэффициентін кодтаудың құрылымдық схемасы

Сызықтық болжау коэффициент кадрның соңғы он элементі сөйлеу генерациясы үшін сандық көп буынды сүзгіште қолданылатын коэффициенттерге сәйкес келеді. Қарапайым процессорлардың тез әрекет етуі жеткіліксіз болғандықтан, практикалық іске асыру үшін арнайы есептеуіш немесе бағдарламаланатын сигналдар процессорларын қолданады.

Бұл әдіс жоғары жылдамдықты элементтік базаны талап ететіндіктен оны іске асыру қиынға соғады. Алайда, бұл әдіс ең перспективалы болып саналады, өйткені ол сөзжасау процестеріне негізделген. Дыбыстар мен сөздердің энергия деңгейі бойынша, қарқын бойынша, үндік сурет бойынша келісуді жүзеге асыру арқылы модель параметрлерін басқару мүмкіндігі пайда болады. Бұл ережелеге сүйене отырып, элементтер жиынтығынан күрделі сөздерді синтездеуге мүмкіндік береді.

Дыбыс болжауды тікелей пайдалану кезінде шығарылады, бірақ дыбыстың сапасы нашар болады. Сондықтан бұл әдістің сапасын жақсартудың бірнеше жолдары бар. Бұл әдістер қабылдау соңында генераторлардың қозу параметрлерін жақсарту үшін қолданылады. Сондықтан үш құраушы жүйенің ішінде-аппроксимация, болжау және сөйлеуді қалпына келтіру (генераторлардың қозуы) әдістері - барлық жетістіктер сызықтық болжау әдісінің соңғы құраушысына қатысты болады. Олар кейде гибриді кодтар деп де атайды, өйткені олар сигналдардың және кодтардың гибриді болып табылады.

Ақпарат шығындары бойынша сөйлеу синтезінің негізгі әдістерінің салыстырмалы мәліметтерін келесі кестеде[8] келтірілген (Кесте 1.1).

Кесте 1.1 – Ақпарат шығындары бойынша сөйлеу синтезінің негізгі әдістерінің салыстырмалы мәліметтері

Сөйлеуді кодтау әдісі	Ақпарат шығыны бит/с	Сигнал сапалылығы
ИКМ	$(40-100) \cdot 10^3$	-
Дельта-модуляция	$(20-50) \cdot 10^3$	-
АИКДМ	$(10-25) \cdot 10^3$	-
Клиппирлеу	$(3-60) \cdot 10^3$	Сөйлеудің төмен сапасы
Формантты талдау	$(2-4) \cdot 10^3$	-
Сызықтық болжау коэффициентінің талдауы	$(1,2-2) \cdot 10^3$	-
Фонемалық әдіс	50-100	-

Клиппирлеу - сөйлеуді өңдеу үшін ғана қолданылатын сигналдың шекті амплитудалық шектеуі. Сөйлеу сигналы ені әртүрлі тікбұрышты импульстердің тізбегіне айналады, яғни оң сигнал кезінде нөлмен, теріс сигнал кезінде бірмен кодталады.

Фонемалық әдіс сөйлеу синтезін басқаратын деректерді берудің ең аз жылдамдығына ие. Бұл әдістің тағы бір маңызды артықшылығы - бұл сөйлеу хабарламаларын ережелер бойынша қалыптастыру мүмкіндігі (синтезделетін сөздердің шексіз сөздігі).

Мультимедада сөйлеу синтезаторын кең қолданумен қатар, сөйлеу қарым-қатынасының әдістері мен құралдары телефондық автожауап бергіштерде, соқыр және сөйлейтін адамдарға арналған оқу құрылғыларында, сондай-ақ, қазіргі заманғы әскери техникаларда, мысалы ұшақтарда, ғарыш жүйелерінде және т. б. қолданылады.

Сөйлеуді талдаудың барлық әдістері уақыт бойынша сөйлеу сигналының қасиеттерінің баяу өзгеруін болжайды. Дауыстық трактының сипаттамаларын 10-20 мс аралығында өзгеріссіз деп санауға болады, яғни параметрлерін 1/20 мс = 50 Гц жиілікпен өлшеу керек[9].

1.3 Сызықты болжауы бар вокодерлерді қолдану негізінде сөйлеуді талдау әдістері

Вокодер – бұл қазіргі заманғы сандық телефон байланысында, оның ішінде – Internet-телефонияда кеңінен қолданылатын сөйлеуді параметрлік кодтау жүйесі. Оның себебі – ақпаратты сығудың жоғары дәрежесі, сондай-ақ, арналық кодтау және шифрлеу жүйелерімен байланыс жүйелерінің

кедергілерден және ақпараттың жылыстауынан жоғары қорғалуы салыстырмалы жеңіл қамтамасыз етіледі[10].

Дауыс түзуші тракт уақыт бойынша ауыспалы параметрлері бар желілік сүзгі ретінде ұсынылады. Ол ақ шудың (дауыссыз дыбыстарды қалыптастырған кезде) немесе негізгі тонның кезеңімен импульстер тізбектерімен қалыптасады.

Сөйлеу сигналын талдауының нәтижесінде вокодер алынған ақпаратты декодерге береді, бұл сөздік сүзгісінің параметрлері, көрсеткіш дауысты/дауыссыз дыбыс, қозу сигналының қуаты және кезең дауысты дыбыстар үшін негізгі тон. Бұл параметрлер әрбір 10 – 20 мс жаңартылуы керек, бұл сөйлеу сигналының стационарсыздығын бақылау үшін қажет. Вокодердің кодерден айырмашылығы - ол бастапқы сөйлеуге жақын келетін сөз формасын қалыптастыруға тырысады және оның формасының бастапқы сигнал формасынан айырмашылығына назар аудармайды.

1.4 Негізгі тонның жиілігінен қоздырылатын LPC кодтау әдісі

Негізгі тонның жиілігінен қоздырылатын кодер - бұл 2400 бит/с және одан төмен жылдамдықпен сөйлеу сигналының параметрлерін беру үшін қолданылатын LPC (Linear Predictive Coding) кодері.

Негізгі тонның жиілігінен қозуы бар кодер тіпті жоғары тарату жылдамдығы кезінде де синтезделген сөйлеудің қажетті сапасын қамтамасыз етпейді, яғни барлық дыбыстар үшін сөздің дәлме-дәл дауысты және дауыссыз болып бөлінуі мүмкін емес[11].

Дауыстық саңылауды түйістіру кезінде орын алатын негізгі тон жиілігінің негізгі қозуынан басқа, дауыстық саңылауды ажыратқанда ғана емес, сондай-ақ, түйістіру кезінде де екіншілік қозу бар екені белгілі. Сондықтан көпимпульсті қозғауда LPC қалдығының сигналы бірқалыпты бөлінбеген аралықтары және әртүрлі амплитудалары бар (10 мс үшін шамамен 8 импульс) импульстер тізбегі түрінде ұсынылады.

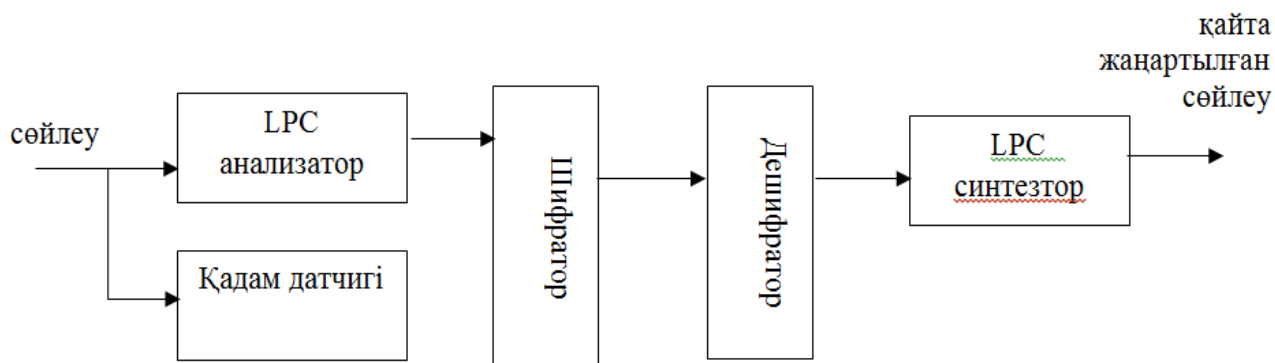
Әрбір кадрда LPC-параметрлермен бірге қозу импульстерінің ережелері мен амплитудалары туралы ақпарат кодермен қалыптастырылады.

Осы типті кодерлер дайын есептемелердің тұтас блоктарымен жұмыс істейтін болғандықтан, әрбір блок үшін LPC алгоритмі негізгі тонның жиілігін, оның амплитудасын, сигналдың сөйлеу немесе сөйлеу сигналына жатпайтын сигналдарды ажырату және басқа да параметрлерді есептеп береді. Содан кейін осы параметрлердің мәндеріне сүйене отырып, хабар беруге дайын сөйлеу кадры қалыптасады.

Кодты кодтауға осы тәсілмен бірінші кезекте есептеу қуатына қойылатын талаптар артады және екіншіден, беру кезінде кідіріс артады, себебі кодтау жеке мәндерге емес, түрлендіруді бастамас бұрын белгілі бір буферде жинақталатын олардың кейбір жиынтығына қолданылады.

Бұл әдіс LPC-вокодер сияқты адамның сөзінің нақты ойнатылуына тікелей байланысты болғандықтан, ол дауыстық жиілік жолағына түсетін дыбыссыз сигналдар үшін жарамсыз болып табылады[12].

Төменде қабылдағыштың бөлімі ретінде таратқыштан тұратын LPC-вокодерінің негізгі блок-схемасының суреті көрсетілген. Таратқыш LPC анализаторы, қадам датчигі, LPC және қадам үшін кодерден тұрады. Қабылдағыш LPC , қадам және LPC синтезторы үшін дешифратордан тұрады.



Сурет 1.8 – LPC-вокодерінің негізгі блок-схемасы

1.5 RELP немесе сөйлеуді гибриді кодтау әдісі

Сөйлеудің гибриді кодтау әдістері сигналдың шығу табиғатын мүлдем қарастырмайтын сигнал формасының кодтары мен сигналдың өзін емес, оның шығу көзі параметрлерін қарастыратын кодтар арасындағы алшақтықты толтырады.

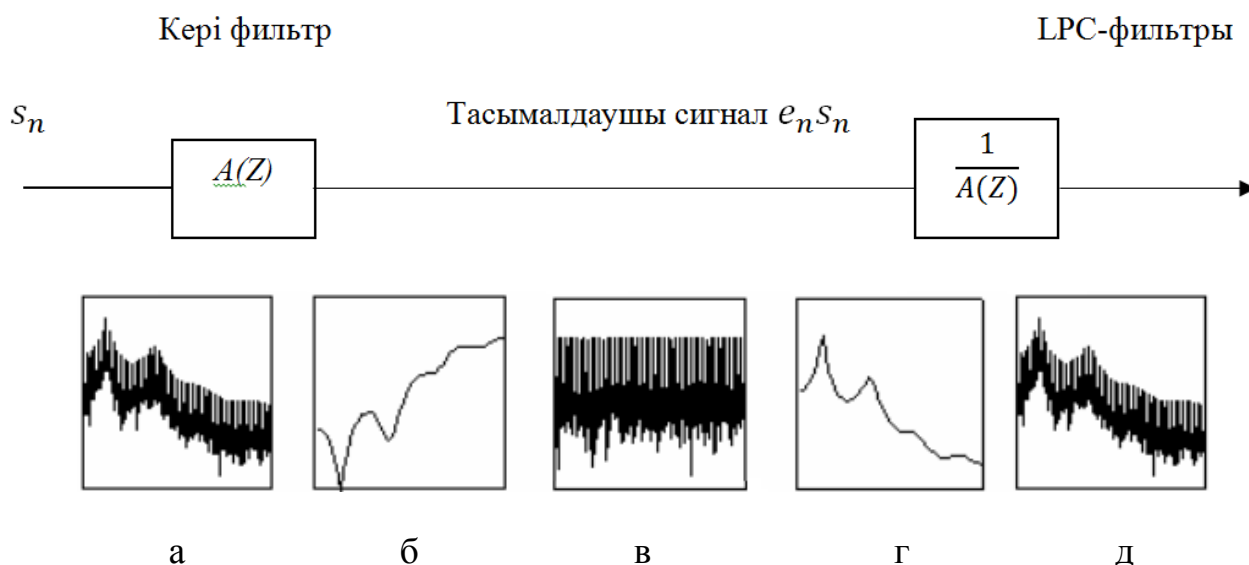
Сигнал формасының коды кодтау жылдамдығы 16 кбит/с жоғары болған кезде өте жақсы сөйлеу сапасын қамтамасыз етеді, бірақ бұдан төмен жылдамдықтарда мүлдем жұмыс жасамайды, ал вокодерлер 2,4 кбит/с және одан төмен кодтау жылдамдығы кезінде сөйлеудің анықтылығын қамтамасыз еткенімен, кез келген басқа жылдамдықтарда сөздің жақсы сапасын қамтамасыз ете алмайды.

Қазіргі таңда сигналдың синтезі бойынша талдауға негізделген (ABS-кодектер) уақыт аймағында (сигналдың спектрі немесе басқа да сызықтық түрлендірулермен емес, сигналдың өзімен) жұмыс істейтін гибриді кодтау әдістері кең қолданысқа ие[13].

ABS кодері алғаш рет бастапқы түрінде MPE-кодері (Multi - Pulse Excited-көпимпульсті қозу кодері) атауында ұсынылды. ABS тиімділігі мен сапасы неде екенін түсіну үшін, алдымен RELP-кодері (Residual Excited Linear Prediction) қарастырылады.

Егер сөйлеу сигналын (Сурет 1.9, а - да көрсетілген спектрі бар) сызықты болжағыш (Сурет 1.9,б - да көрсетілген жиіліктік сипаттамасымен) арқылы

өткізетін болса, шығу сигналын есептеу арасындағы корреляция (болжау қателері) айтарлықтай азаяды. Егер болжам жеткілікті түрде жақсы орындалса, онда болжаушының шығуы біркелкі спектрі бар ақ шу болып шығады (Сурет 1.9, в).



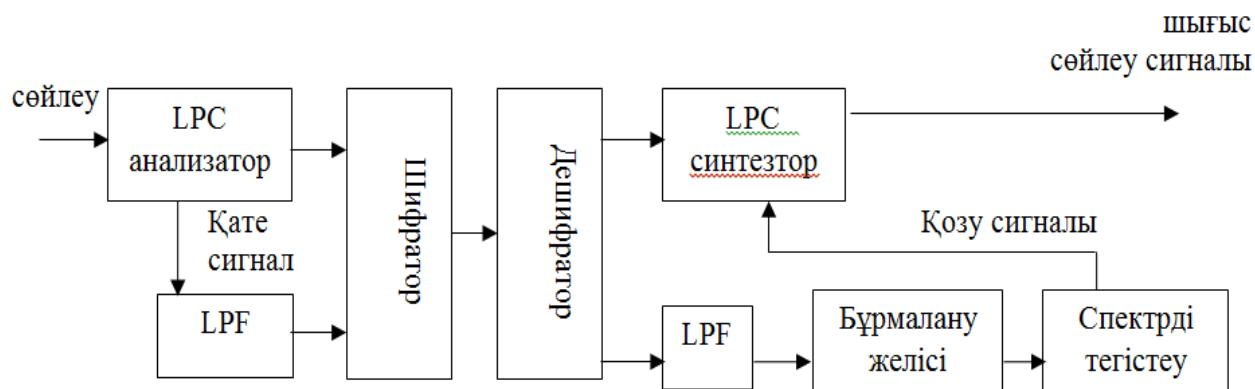
Сурет 1.9 – Сызықтық болжау коэффициенттері көмегімен сөйлеу сигналының анализі

Сонымен қатар, бұл ақ шу (болжау қатесі) барлық кодталған сөйлеу сигналы туралы ақпаратты сақтайды және егер оны қайтадан LPC-сүзгі арқылы (Сурет 1.9, г көрсетілген жиіліктік сипаттамасымен) өткізетін болса, онда бастапқы сөйлеу сигналын мүлдем дәл қалпына келтіруге болады. Өйткені бұл ақпарат болжау қателігінің спектрі бойынша біркелкі бөлінетіндіктен, болжам қателерінің спектрінің кейбір бөлігін ғана кодтау және беру идеясы пайда болды, ал қалғандары декодерде қалпына келтіріледі[14].

RELPC-кодерде сигналды болжау қателігі кесу жиілігі 1 кГц-ке жуық төмен жиілікті сүзгі арқылы өткізіледі. Шығу сигналы, мысалы ДИКМ-кодерын қолдану арқылы форма бойынша кодталады. Декодерде болжау қатесі оны төмен жиілікті сүзгілермен жойылған жиілік аймағына көшіру арқылы қалпына келтіріледі. Сызықтық болжау барысында ақ шу алу мүмкіндігі болатын болса, RELPC-кодерді бұл жағдайда мінсіз болатын еді. Алайда, сөйлеу сигналында квазипериодикалық форманттық құрамдастар бар болуынан, сызықтық болжағыш форманттың негізгі тон периоды бар ұзақ мерзімді корреляциясын жоя алмайды, сол себепті олар болжам қателіктерінің спектрінде анық байқалады. Егер болжау қатесін төменгі жиілік фильтры арқылы жіберетін болса, жоғары жиілікті формантты құрамдастар жоғалады және әрі қарай қалпына келу мүмкіндігінен айырылады[15].

RELPC кодеры 9.6 кбит/с шамасында код жылдамдығымен сигналдың жақсы сапасын алуға мүмкіндік береді, бірақ оларға тән бір кемшілік - қалпына

келтірілген сөйлеудің синтетикалық сипаты. Бұл сипаттама вокодерлерге де тән. Осыған байланысты олардың орнына ABS-кодерлер қолданыла бастады.

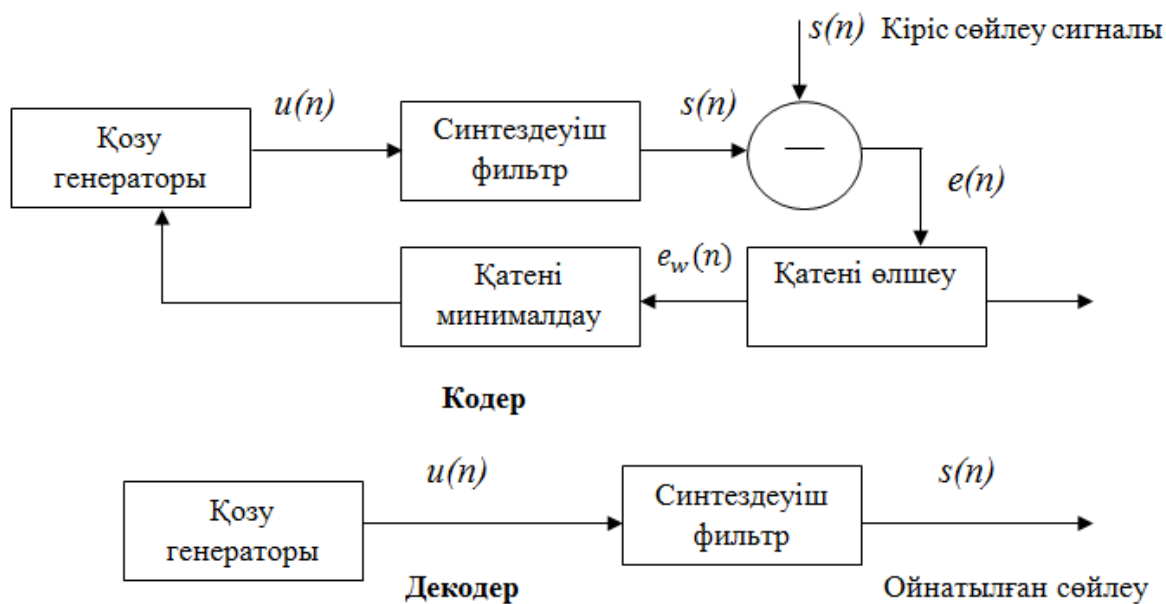


Сурет 1.10 – RELP-вокодерінің негізгі блок-схемасы

1.6 RELP негізінде ABS кодтау әдісі

ABS кодердің жұмыс істеу принципі келесідей жүреді. Кодталған кіріс сигналы (сандық нысанда, санау ағыны түрінде) ұзындығы 20 мс фрагменттерге бөлінеді, оның шегінде сигналдың қасиеттері елеусіз ғана өзгереді. Осы фрагменттердің әрқайсысы үшін синтездеуші сүзгінің (дауыс трактысының аналогы) ағымдағы параметрлері анықталады және одан әрі синтез сүзгісі арқылы өткізілген қозу сигналы таңдалып, кіріс және синтезделген сигнал арасындағы қатені азайтады.

Осылайша, ABS әдісінің мағынасы - кодер оған көпше жақындылықтар синтезі арқылы кіріс сөзін талдайды. Кейін кодер декодерге ағымдағы параметрлердің комбинациясын білдіретін синтездеу сүзгісі мен қозу сигналы туралы ақпаратты береді. Мүмкіндігінше, осы деректердің көлемі кішірек болғаны дұрыс. Осы параметрлер бойынша декодер кодталған сөзді қалпына келтіреді. Әртүрлі типті ABS-кодтар арасындағы айырмашылық олардың әрқайсысында синтездеуші сүзгінің қозу сигналы $u(n)$ таңдалады. Теория бойынша синтездеуші фильтр кірісіне әртүрлі сигналдардың шексіз көп санын беру керек, яғни бұл оның шығуында қандай сигнал болатынын көру үшін және оны кодталған сигналмен салыстыру үшін қажет. Бастапқы сөзбен синтезделген сөз арасындағы өлшенген қателердің минимумын беретін қозу сигналы кодтау нәтижесі ретінде таңдалады. Қозу сигналын анықтайтын бұл тұйық схема ABS-кодерлары үшін кодтаудың төмен жылдамдығы кезінде кодталатын сөздің жоғары сапасын қамтамасыз етеді (Сурет 1.11).



Сурет 1.11 – ABS-кодер/декодер құрылымдық схемасы

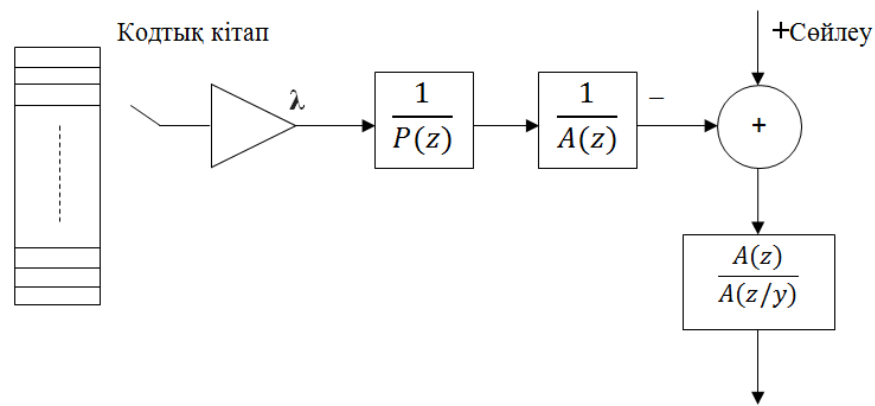
1.7 Кодтық кітаптар негізінде қозуы бар кодтар немесе CELP кодтау әдісі

Жоғарыда айтылып кеткен MPE кодтау әдісі кодтың жылдамдығы 10 кбит/с және одан да жоғары болғанда сөйлеудің жақсы сапасын қамтамасыз етеді, бірақ төмен жылдамдықтарда сигналды қатты бұрмалауға ұшыратады, себебі қозу сигналының қажетті параметрлері үшін импульстердің орналасуы мен амплитудасын қажетті дәлдікпен сипаттайтын биттердің саны жеткіліксіз болады.

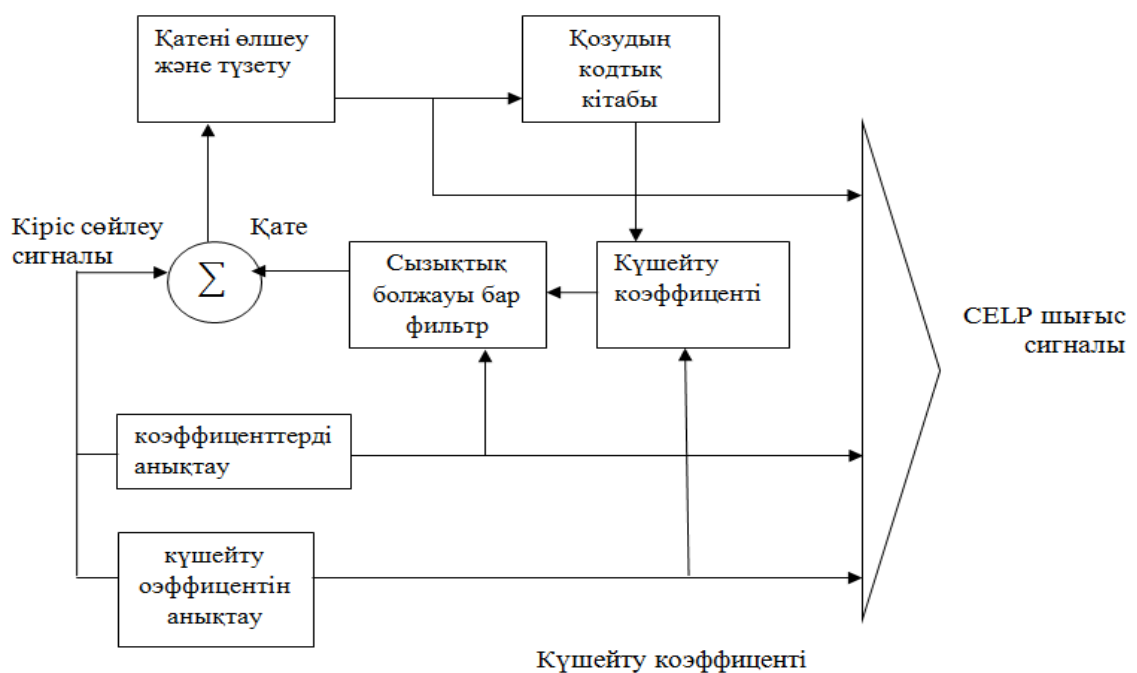
Осыған байланысты қозу сигналы ретінде өзіндік параметрлер жиынтығымен берілетін импустақ реттілік емес, әртүрлі формадағы қозулар сигналының есте сақтау құрылғысына арнайы дайындалған және жазылған кітапханана (кодтық кітаптар) – CELP (Code Excited Linear Predictive) қолданылады[16].

Кодтық кітаптардың көптеген түрлері бар:

1. векторлардың кодтарын іздеу принципі бойынша (толық аралықпен, екілік немесе дәйекті іздеу және т. б.);
2. оқыту тәсілдері бойынша (тіркелген немесе бейімделген кітап);
3. сақталатын ақпарат түрі бойынша (сөйлеуді таңдау немесе шуды іске асыру).



Сурет 1.12 – CELP-кодерінің қозу сигналының пайда болу құрылымдық схемасы



Сурет 1.13 – CELP- кодерінің құрылымдық схемасы

Бұл ретте кодтау нәтижесі болып қозу сигналының импульс параметрлері емес, кодтық кітаптың индексі (онда сақталатын нөмір), сондай-ақ оның амплитудасы болады. Мысалы, кодтық кітапта 1024 сигнал бар болса, ал сигнал амплитудасы 2 – 3% дәлдікпен кодталса, онда қажетті бит саны 10 (Индекс үшін) + 5 (амплитуда үшін) = 15 бит, яғни ұзақтығы 20 мс сигнал фрагментіне 15 бит (GSM RPE-кодектегі 47 битпен салыстырғанда).

CELP кодерінің құрылымдық схемасынан көрініп тұрғандай (Сурет 1.13), сигналдарды кодтау үшін "қозудың кодтық кітабы" қолданылады. Бұл жағдайда әрбір тасушы сигналға сигналдар жолы (үлгісі) салыстырылады. Бұл жол алынған қате мәнінің негізінде берілетін қалдық кезектілікке сәйкес келетін

санак жиынтығынан тұрады. Қабылдау соңында тасушы сигнал декодеры орнына «қозудың кодтық кітабы» қолданылады.

CELP алгоритмі сөйлеудің жоғары сапаын қамтамасыз етеді. Орташа сараптамалық баға: 3,5-3,5. Алайда, ол үлкен есептеу ресурстарын талап етеді, сондықтан оның негізінде көптеген кодтар құрастырылған. CELP-әдісі көпимпульсті әдіске қатысты бірдей жылдамдықпен сөйлеуді қалпына келтірудің жоғары көрсеткіштерін көрсетеді.

Сөйлеуді кодтаудың кейбір негізгі алгоритмдерінің сипаттамалары төменгі кестеде берілген:

Кесте 1.2 - Сөйлеуді кодтаудың кейбір негізгі алгоритмдерінің сипаттамалары

Алгоритм атауы	Ұсыну	Алгоритм жылдамдығы (кбит/с)	Сөйлеу кадрының мөлшері(октет)	Жинақталу кідірісі (мс)
CS-ACELP	ITU G.729	8	10	10
PCM	ITU G.711	64	40	5
		56	35	5
		48	30	5
ADPCM	ITU G.726	40	25	5
		32	20	5
		24	15	5
		16	10	5
LD-CELP	ITU G.728	16	10	5
MP-MLQ	ITU G.723.1	6.3	24	30
ACELP	ITU G.723.1	5.3	20	30

Кез келген вокодердің маңызды сипаттамасын оның дыбыстық сапасы көрсетеді. Бұл сапасын бағалау үшін орташа субъективті бағалау (MOS - mean opinion score) немесе адамның ойнайтын сөзге психологиялық реакциясы ұғымы енгізілді.

Адам ақпаратты алушы ретінде кез келген телекоммуникациялық жүйенің негізгі элементі болып табылатындықтан, сигнал сапасы оның сөйлеуді субъективті қабылдауы бойынша бағаланады. Сигнал сапасы MOS (mean opinion score - орташа субъективті бағалау) бес балдық шкала бойынша өлшенеді. MOS шкаласы бойынша тыңдаушылар тобы әртүрлі дауыс зорайтқыштармен ойнайтын бірнеше сөйлеу сигналдарына беретін бағаларды өңдеу жолымен анықталады. Әрбір тыңдаушы әр сигналдың бағасын шығарады: 1 - нашар, 2 - әлсіз, 3 - анық, 4 - Жақсы, 5 - өте жақсы. Содан кейін нәтижелер орташаланады[17].

Кесте 1.3- MOS негізгі сөйлеуді кодтау алгоритмдерінің көрсеткіштері

Алгоритм атауы	MOS
G.711 (PCM; 64 кбит/с)	4,1
G.726 (ADPCM; 32 кбит/с)	3,8
G.728 (LD-CELP; 16 кбит/с)	3,6
G.723.1 (ACELP; 5,3 кбит/с)	3,7
G.723.1 (MP-MLQ; 6,3 кбит/с)	3,9

Берілген кодтау әдістерінің арасында сөйлеу сапасы / ағын жылдамдығы арақатынасы тұрғысынан ең қолайлы G. 723 алгоритмі болып табылады.

Кодтау рәсімі өте үлкен есептеу шығындарын талап етеді, сондықтан CELP-кодтарын іске асыру тек соңғы уақытта ғана мүмкін болды. Арнайы сигналдық процессорларды пайдалану секундына 300 млн. және одан да көп операция орындап уақыты үнемдеуге мүмкіндік береді.

Celp алгоритмі қазіргі заманғы байланыс жүйелерінде 16-дан 4,8 кбит/с дейінгі жылдамдықпен көп қолданылады. Сонымен қатар CELP-кодерінің 16 кбит/с кодының жылдамдығы, 64 кбит/с ИКМ үшін, ал 4,8 кбит/с код жылдамдығы кезінде – 13 кбит /с GSM RPE үшін сигнал сапасын қамтамасыз етеді.

2 Импульстік-кодтық модуляция

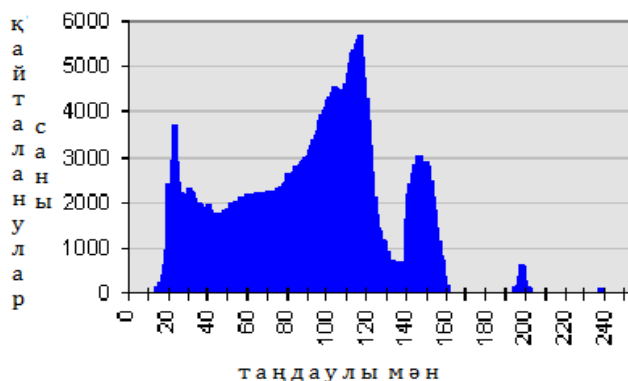
Үздіксіз бастапқы аналогтық сигналды сандық кодқа түрлендіру импульстік-кодтық модуляция (ИКМ) деп аталады. Телекоммуникацияда негізгі код ретінде ең аз аппараттық шығынмен жүзеге асырылатын екілік бірізділік таңдалады. ИКМ кезіндегі негізгі операциялар болып: уақыт бойынша дискреттеу, кванттау (уақыт бойынша дискреттеу деңгейі бойынша дискреттеу) және кодтау операциялары болып табылады[18].

Модуляциялық сигнал Заңы бойынша амплитудалық-импульстік модуляцияда (АИМ) импульс амплитудасы өзгереді, ал жүру ұзақтығы мен жиілігі тұрақты болып қалады. Сигналдарды аналогтыдан сандық түрге түрлендіру тарату кезінде олардың кедергіден қорғалуын айтарлықтай арттырады, өйткені қабылдағыш берілетін сигналдың екі күйін тіркеуі тиіс: берілетін сигналдың күйін немесе оның болуын (бірліктерді қабылдау) немесе оның болмауын (нөлді қабылдау).

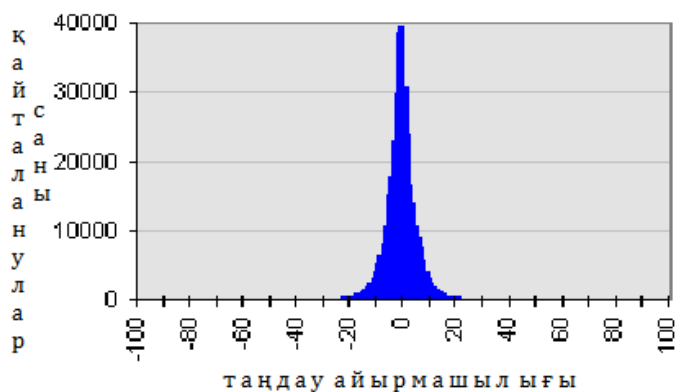
2.1 Дифференциалды импульстік-кодтық модуляция

DPCM әдісінде кодтық сөздер сигнал деңгейлерінің мәнін емес, олардың арасындағы айырмашылықты көрсетеді. Көршілес деңгейлерде аз айырмашылықтары бар сигналдардың көптеген түрлері бар. Егер оларға DPCM қолданса, онда деңгейлердегі айырмашылықты қамтитын кодтық сөздер басқалардан әлдеқайда жиі кездеседі, демек, сигналдың сығылуы жақсы болады. Бұл болжамды кодтаудың қарапайым мысалы, өйткені іс жүзінде сигналдың қазіргі ағымдағы мәні арқылы, оның келесі мәнін болжауға болады[19].

Қарапайым мысал ретінде төменде PCM және DPCM көмегімен кодталған бір суреттің 2 гистограммасы (суреттегі деңгей/сан) көрсетілген (Сурет 2.1).



а)



б)

Сурет 2.1 – PCM кодтаудың (а) және DPCM кодтаудың (б) дифференциалды гистограммасы

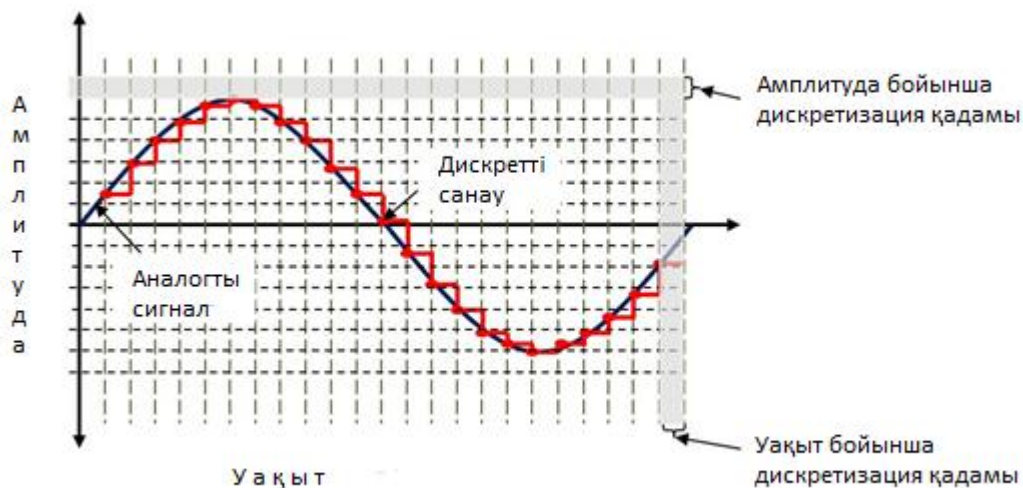
Кодтық сөздердің ауқымы айтарлықтай аз, осылайша тұрақты қысу коэффициентін алуға болады. DPCM практикасында әдетте JPEG немесе Adaptive (адаптивті) DPCM - дыбыс сигналдарын қысудың кең таралған әдісі қолданылады.

Бұл әдісте кванттау қадамының өлшемі қысылған сигналдың ағымдағы өзгеру жылдамдығына бейімделеді (яғни, бұл әдісте сөйлеуді беретін аналогтық сигналда қарқындылықтың күрт секіруі мүмкін емес, сондықтан сигналдың амплитудасын емес, оның ағымдағы мәнін алдыңғы мәнімен салыстырғандағы мәнін кодтайтын болсақ разрядтардың аз мәніне қол жеткізуге болады).

Бұл әдістің артықшылығы- процессордың ең аз жүктелуі, кемшілігі – кванттау шуы мен қысудың орташа дәрежесі 4:1.

2.2 Сөйлеу сигналын дискреттеу

АСТ (аналогтық-сандық түрлендіргіш) үшін бастапқы материал болып табылатын аналогтық сигнал уақыт бойынша дыбыс қысымының амплитудасы мен жиілігінің өзгеруін сипаттайтын үздіксіз сигнал болып табылады. Аналогтық сигналды сандық түрге айналдыру үшін оны дискреттеу қажет, яғни амплитуда бойынша да, уақыт бойынша да жеке (дискретті) элементтерге бөлу қажет (Сурет 2.1-де аналогтық сигналдың уақыт бойынша дискреттелуі көрсетілген)[20].



Сурет 2.2 – Уақыт бойынша дискреттелген сигнал

Сигналдарды өңдеу кезінде дискретизация – $S(t)$ аналогты үздіксіз сигналдың оның мәндерінің жиынтығымен берілуі, бұл жиынтықты t_0, t_1, \dots, t_{n-1} уақыт сәтінде алынған таңдау (выборка) $S(t_0), S(t_1), \dots, S(t_{n-1})$ деп атауға болады.

Жалпы жағдайда бір таңдаудан келесі таңдауға дейінгі уақыт кезеңі көрші таңдаулардың әрбір жұбы үшін әр түрлі болуы мүмкін, бірақ әдетте сигналды өңдеу кезінде таңдаулар белгіленген және тұрақты уақыт аралығы арқылы жүреді. Осы аралық бұл жағдайда дискретизация периоды немесе таңдау интервалы деп аталады және T әріпімен белгілененді. Дискретизация периодына кері шаманы $F_d = 1/T$ таңдау жиілігі немесе дискретизация жиілігі деп атайды.

Аналогтық сигналдардың мысалдары ретінде аудио немесе бейне сигналдар, әртүрлі өлшеу датчиктерінің сигналдары және т.б. бола алады. Келесі цифрлық өңдеу үшін аналогтық үздіксіз сигналдар міндетті түрде аналды-цифрлық түрлендіргіштердің (АЦТ) көмегімен деңгей бойынша дискретизациялаудан және квантталудан өтеді.

Қазіргі таңда аналогтық сигналдың бірқалыпты дискретизациясы қолданылатын цифрлық тарату жүйелері көп таралған (бұл сигналдың есебі бірдей уақыт аралықтары арқылы жүргізіледі). Бірқалыпты дискреттеу кезінде: Δt дискреттеу интервалы (дискретті сигналдың екі көршілес есептерінің арасындағы уақыт аралығы) және F_d дискреттеу жиілігі (дискреттеу интервалына кері шама) қолданылады. Импульстік-кодтық модуляцияланатын сигналды қалыптастыру принципінің негізіне Котельников теоремасы (Шеннон) қолданылады: кез келген аналогты (үздіксіз) сигнал дискретизацияланады және дискретизация жиілігі осы сигналдың жоғарғы жиілігінен екі есе артық болса, қарама-қарсы ұшында қалпына келтірілуі мүмкін.

$$F_d > 2F_{\max} \quad (2.1)$$

Мұндағы,

F_d – дискретизация жиілігі, Гц;

F_{\max} – жоғарғы екі еселенген жиілік, Гц;

Тоналдық жиілік арнасы (аналогтық телефон арнасының негізгі арнасы) 300...3400 Гц жолағын алуы тиіс. Сондықтан дискретизация жиілігі: $F_d = 2 \times 3400 = 6800$ Гц- тен кем болмауы керек.

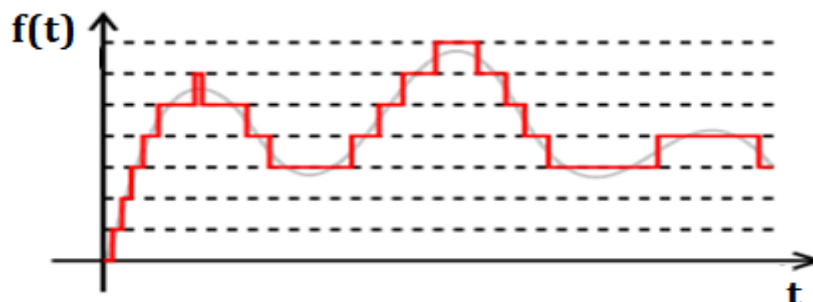
Тоналды жиілік арнасы бойынша берілетін сигнал үшін телефония және телеграфия жөніндегі Халықаралық консультативтік комитеттің (МККТТ) ұсыныстарына сәйкес $F_d = 8000$ Гц-ке тең дискреттеу жиілігі қабылданды. Мұндай жиілік ЦСП аппаратурасының сүзгілерін іске асыруды жеңілдетеді.

Дискреттелген импульстер осы сигналдың жылдам мәндерінің амплитудасына сәйкес келеді. Бұл түрлендірулер импульстік-амплитудалық модуляция деп аталады. Сигналдың жылдам мәндері АИМ сигналында болады. Әрбір кезеңде бір рет бос уақыт позициялары басқа арналардың таңдауларын берумен айналысады. Осылайша, дискреттер уақытша тығыздалған күйде бірінен соң бірі АИМ сигналдары түрінде беріледі.

Импульстік-кодтық модуляцияны берілетін ақпараттың жоғары кедергіден қорғалуын қамтамасыз ету үшін қолданады. ИКМ процесінде сигналдың әрбір дискреттік мәніне белгілі бір кодтық сөз беріледі[21].

2.3 Сөйлеу сигналын кванттау

Бұл ретте кванттауды дискреттеуден дұрыс ажырата білу керек (тиісінше, дискреттеу жиілігін кванттау қадамынан ажырату). Дискретизация кезінде уақыт бойынша өзгеретін шама (сигнал) берілген жиілікпен өлшенеді (дискретизация жиілігі), осылайша дискретизация сигналды уақыт бойынша бөледі. Ал квантау сигналды берілген мәндерге келтіреді, яғни сигналды оған жақын деңгейлерге дейін[22] дөңгелектейді (Сурет 2.3).



Сурет 2.3 – Квантованталған сигнал

АСТ-да дөңгелектеу ең жақын аз деңгейге дейін жүргізілуі мүмкін.

Кванттау сигналдарды өңдеу кезінде, соның ішінде дыбыс пен бейнелерді қысу кезінде жиі қолданылады.

Сигналды цифрлау кезінде кванттаудың бір деңгейін кодтайтын биттердің саны кванттау тереңдігі немесе разрядтылығы деп аталады. Кванттау тереңдігі және дискреттеу жиілігі көп болған сайын, сандық сигнал аналогтық сигналға дәлірек сәйкес келеді. Бірқалыпты кванттау кезінде кванттау тереңдігі децибелдерде өлшенетін динамикалық диапазонды анықтайды (6 дБ-да 1 бит).

Тарату жүйесі мен хабарламаларды өңдеудің кез-келген жүйесінің қорытынды рұқсат беру қабіліті бар, сондықтан үздіксіз сигналдардың барлық шексіз амплитудалық мәндерін берудің қажеті жоқ, оны соңғы жиынмен шектеуге болады. Бұл сигналдардың амплитудалық мәндері кванттау деңгейлері деп аталады, олардың санын таңдау арқылы электр сигналдарын беру сапасын анықтайды.

Дискреттеу кезінде алынған АИМ сигнал деңгей бойынша квантталады. Екі көршілес рұқсат етілген деңгейлер арасындағы айырмашылық кванттау қадамы - Δ деп аталады.

Сигналды есептеудің шынайы мәні мен оның квантталған мәні арасындағы айырмашылық қате немесе кванттау шуы деп аталады.

Қазіргі таңда кванттау шуын азайту үшін екі әдіс қолданылады. Бірінші әдісте тарату жүйесіндегі сигнал компандирлеуге ұшырайды. Компандирлеу – бұл екі өзара кері түрлендіруден тұратын процесс. Бұл жағдайда бірқалыпты квантталумен дискретті сигнал компрессияға ұшырайды, яғни дискретті сигнал әлсіз сигналдарда үлкен және үлкен сигналдарда аз болады. Сигналды қалпына келтіру кезінде қабылдау жағында кері түрлендіру - экспандирлеу жүреді және сигнал бастапқы түрге келтіріледі.

Сөйлеу сигналы табиғи сигнал болып табылады және кездейсоқ сипатқа ие. Бұл сондай-ақ, дәйекті үлгілер арасындағы өте жоғары корреляцияны көрсетеді. Бұл алдыңғы таңдаулардан келесі іріктеулерді болжауға мүмкіндік береді. Сызықтық болжамды талдау сөйлеу сигналдарын қысу және кодтау үшін жарамды болып келеді.

Сигнал кодталған кезде сигналдың ақпараттық мазмұндылығы жоғалмауы тиіс. Сөйлеу сигналдарын дауысты және дауыссыз сигнал деп кеңінен жіктеуге болады. Дауысты сөйлеу сигналы максималды энергияға ие, бірақ аз ақпараттан тұрады, ал дауыссыз сигнал керісінше төмен энергияға ие, бірақ максималды ақпараттан тұрады. Сол себептен, дауысты сегментті кодтауға қарағанда дауыссыз сегментті кодтауға назар көбірек аударылуы тиіс. Амплитуданың өте кішкентай өзгерістері адам құлағымен анықталмайды. Адам құлағының бұл қасиеті сөйлеу сигналдарын кванттауға мүмкіндік береді, себебі кванттау қателігі адам құлағымен анықталуы мүмкін емес[23].

Кванттаушы үшін екі үлкен санат бар: біркелкі кванттау және біркелкі емес кванттау.

2.4 Біркелкі және біркелкі емес кванттаушы және кодер

Біртекті квантаушы тұрақты қадам өлшемін пайдаланады. Қандай да бір сигнал компьютерлік жүйемен түйіскен кезде АСТ-ға кіру беріледі. Стандартты АСТ-ның кіру диапазоны -1 және $+1$ аралығында жатыр. Демек, компьютерлік жүйеге қосылуы тиіс сигнал әдетте тұрақтандырылады, сондықтан сигнал амплитудасы -1 және $+1$ арасында шектеледі. Сондықтан, тұрақтандырылған сөйлеу сигналының $2B$ динамикалық диапазоны бар болады. Деңгейлер саны кванттау үшін пайдаланылатын биттердің санымен анықталады. Ондық көріністің дәлдігін сақтау үшін, айталық, екі разрядты ондық сан үшін кем дегенде жеті битті екілік көрініс болу керек. Бұл кезде 8-bit кванттау және кодтау қолданылады. Егер B - кванттау үшін пайдаланылатын биттер саны болса, онда кванттау деңгейлері 2^B тең болады. Қадам өлшемін былай белгілеуге болады:

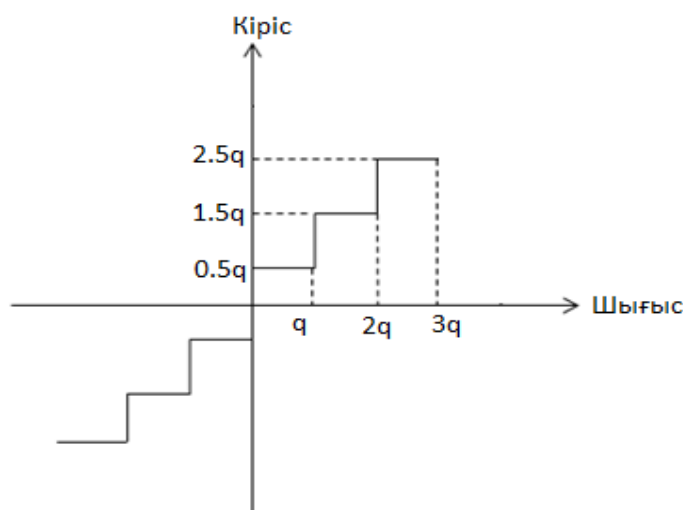
$$q = \frac{2}{2^B} \quad (2.2)$$

Мұндағы,

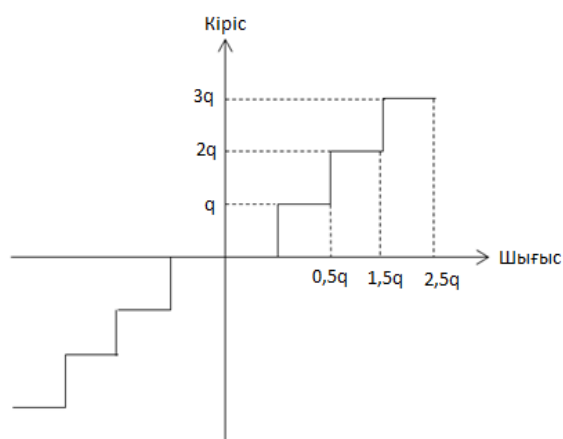
q - қадам өлшемі;

B - кванттау үшін пайдаланылатын биттер саны;

Дауыссыз сөйлеуде қадам өлшемі сигналдың амплитудасынан аз болуы керек. Егер қадам өлшемі сигналдан үлкенірек болса, онда фуксиленбеген (дауыссыз) сегменттердің барлығы нөлдермен кодталады. Біртұтас кванттаушының екі түрі бар: орташаланған көтергішпен (mid-riser) кванттаушы және орташа жіппен кванттаушы (mid-thread). Орташаланған көтергіші және орташа жіптері бар кванттаушылардың беру сипаттамалары тиісінше 1 және 2 суретте көрсетілген. Кейін сандық мән екілік B -код ретінде[24] ұсынылады.



Сурет 2.4 - Орташаланған көтергішпен кванттаушы

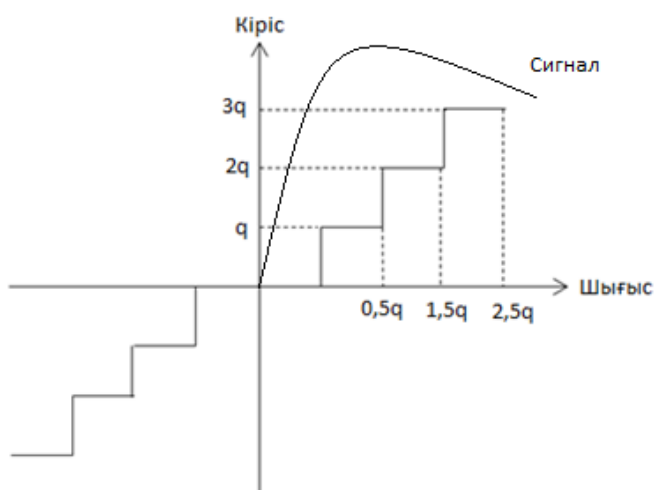


Сурет 2.5 - Орта жіптермен кванттаушы

Орташа көтергішпен кванттаушыны жағдайында нөлдік кіріс $0,54$ деп кодталады, ал орта жіппен кванттаушы жағдайында $0,5$ градустан төмен кез келген кіріс нөлге тең болады. Осындай бірыңғай кванттаушы сөйлеу сигналын кванттау үшін пайдаланылғанда екі қате болуы мүмкін - түйіршікті шу және кернеу шамасынан тыс шу.

Бірдей қадам өлшемімен дауыссыз сегментті кодттауда қадам өлшемі дауссыз бөліктің абсолютті амплитудасынан сәл аз болсын деп алынған. Сурет 2.5 көрсетілгендей кванттаушының шығысы $-q$, 0 және $+q$ арасында бірнеше рет ауысады. Шудың бұл түрі түйіршікті шу деп аталады.

Дауысты сөйлеуді кванттау қарастырылатын болса, дауысты сөйлеу сигналының дауыссыз сөйлеу сигналына қарағанда амплитудасы жоғары болады. Егер дауысты сегменттер үшін дауыссыз сегменттерде қолданылатын біркелкі қадам өлшемі қолданылса, сигнал бұрмалануға ұшырайды. Келесі суретте қадам өлшемінің кіші болуынан кванттаушы шығысы сигналды тез қуып жете алмайды, сол себептен кванттаушы шығысы сигналдан алыс қашықтыққа түседі.



Сурет 2.6 - Бұрыштық сигнал көрінісі

Кванттаушыға жүктеме көп түсіп сигналды анықтай алмағандықтан сигнал бұрыш жасап еңкеюі анық байқалады (Сурет 2.6).

Біркелкі қадам өлшемі сөйлеу сигналын кодтауға келмейді. Сөйлеу сигналын кодтау үшін ауыспалы қадам өлшемін қолданамыз. Дауыссыз сөйлеу үшін қадам өлшемі азайып, дауысты сөйлеу үшін үлкею керек.

Сөйлеу сигналы үшін ықтималдықты бөлу біркелкі емес Лаплас тығызды және Гамма тығыздымен анықталады[25].

Лаплас тығыздығын келесі формуламен анықтайды:

$$P(x) = \frac{1}{\sqrt{2}\sigma_x} e^{-\sqrt{2}|x|/\sigma_x} \quad (2.3)$$

Мұндағы,

$P(x)$ - Лаплас тығыздығы;

σ_x – сигнал дисперсиясы;

Гамма тығыздығының формуласы:

$$P(x) = \left(\frac{\sqrt{3}}{8\pi\sigma_x|x|}\right)^{\frac{1}{2}} e^{-\sqrt{3}|x|/2\sigma_x} \quad (2.4)$$

Мұндағы,

$P(x)$ - Гамма тығыздығы;

σ_x – сигнал дисперсиясы;

Екі шама шексіздікке ұмтылған. Лаплас тығыздығын ұсынған кезде сөйлеу таңдауының 0,35%-ы $-4\sigma_x \leq x(n) \leq 4\sigma_x$ аралығында жатады.

Ықтималықтарды бөлу және сигнал үшін дисперсияны назарға алып, кванттаушы үшін қадам өлшемін анықтау керек. Егер сигналдың шындық мәні болса, қадам өлшемі келесідей болады:

$$q = \frac{2X_{\max}}{2^B} \sigma_e^2 = \frac{q^2}{12} = \frac{X_{\max}^2}{3 \times 2^{2B}} \quad (2.5)$$

Мұндағы,

q – қадам өлшемі;

B – кванттаушының биттер саны;

σ_e^2 - кванттау шуының қуаты;

$$\text{SNR} = \frac{\sigma_x^2}{\sigma_e^2} = \frac{3 \times 2^{2B}}{\left(\frac{X_{\max}^2}{\sigma_e^2}\right)} \quad (2.6)$$

Мұндағы,

$$\text{SNR db} = 10 \log_{10} \left(\frac{\sigma_x^2}{\sigma_e^2}\right) = 6B + 4.77 - 2 \log_{10} \left(\frac{X_{\max}}{\sigma_x}\right) \quad (2.7)$$

Мұндағы,

SNR – өлшеусіз шама сигнал/шу қатынасы;

Квантау шуының дисперсиясы/кванттау шуының қуаты және σ_x сигнал дисперсиясы. Мұндағы $X_{\max}=4\sigma_x$, $SNR=6B-7.2$.

Біркелкі кванттаушы үшін сөйлеу сигналының жоғары сапасын қамтамасыз ету үшін 11 бит қажет. Қажетті биттер жылдамдығы $11 \times F_x$. Дискретизация жиілігі тұрақты. Битер жылдамдығын азайту үшін сәйкесінше биттер санын азайтатын кванттаушыны қолданады.

2.5 Сөйлеу сигналын кванттау бойынша параметрлерін есептеу

Келесі есептеуде сөйлеу сигналының параметрлері есептеледі.

Параметрлер кадам өлшеміне және кванттау шуының қуатына негізделген.

Ортаквадратты (RMS) амплитудасы 3В-қа тең сөйлеу сигналын алты битпен кодтау қажет деп алайық. Кванттау шуының қуатын σ_e^2 және кадам өлшемі q табу керек.

Сигналдың RMS мәні 3В-қа тең. Шыңдық мән табылады.

$$\text{Шыңдық мән} = 3 \times \sqrt{2} = 4.24264 \text{ В}$$

$$q = \frac{2 \times \text{Шыңдық мән}}{2^6} = \frac{2 \times 4.24264}{64} = 0.1325825$$

$$\sigma_e^2 = \frac{q^2}{12} = 0.00146484 \text{ тең болды.}$$

Келесі жағдайда сигнал үлгілері 0.2, 0.4, 0.55, 0.7, 0.75, 0.85, 0.9 және 1.0В мәндерге ие деп алынған. Барлық сегіз таңдау үшін төрт разрядты кодтау мәні мен шу қуатын табу керек.

Сигналдың динамикалық диапазоны 1В-қа тең деп алынған.

Қадам өлшемін табу қажет.

$$q = \frac{\text{Шыңдық мән}}{2^4} = \frac{1\text{В}}{16} = 0.0625$$

$$\sigma_e^2 = \frac{q^2}{12} = 0.00032552 \text{ тең болды.}$$

Енді кодталған шығуларды табу қажет. Әрбір код төрт разрядты код түрінде ұсынылады. Сигналдың нақты мәнінен аз қадам өлшеміне ие ең жақын

интегралды табу керек. Бірінші үлгі үшін 0.2 кодталған мән екілік нысанда 0011-ке тең, яғни үшінші деңгейдің мәні 0.1875 –ке тең үлгіге қарағанда аз болады.

Қадам өлшемінің міні 0.625-ке тең.

Біз үшін үшінші деңгей маңызды болып келеді, сол себептен $0.625 \times 3 = 0.1875$

Егер төртінші деңгейдің мәні қарастырылатын болса $0.625 = 0.2500 > 0.2$.

Төртінші деңгейдің мәні 0.25-ке тең және бұл мән үлгіге қарағанда жоғары болады. Осылайша, үшінші деңгей үшін сигналдың коды ретінде 0.2В алынады, яғни оның екілік нысаны- 0011. Оқырман кодталған нәрсені оңай тексере алады, 8 үлгілер үшін кодтық шығулар келесідей болады:

0011,0110,1000,1011,1100,1101,1110, және 1111.

Келесі жағдайда сөйлеу сигналының ортаквадратты амплитудасы 2В-қа тең деп алынған. Сигналды кодтау алты битпен жүреді. Сигналдың тығыздығы Лаплас бойынша алынады. Кванттау шуының қуатын σ_e^2 , қадам өлшемі q және SNR табу керек[26].

Сигналдың RMS мәні 2В-қа тең. Шыңдық мән есептеледі.

$$q = \frac{2 \times \text{Шыңдық мән}}{2^6} = \frac{2 \times 2 \times \sqrt{2} \text{ В}}{64} = 0.088388$$

$$\sigma_e^2 = \frac{q^2}{12} = 0.00065104$$

Лаплас тығыздығы үшін $\text{SNR} = 6\text{В} - 7.2 = 6 \times 6 - 7.2 = 28.8 \text{ dB}$ тең болды.

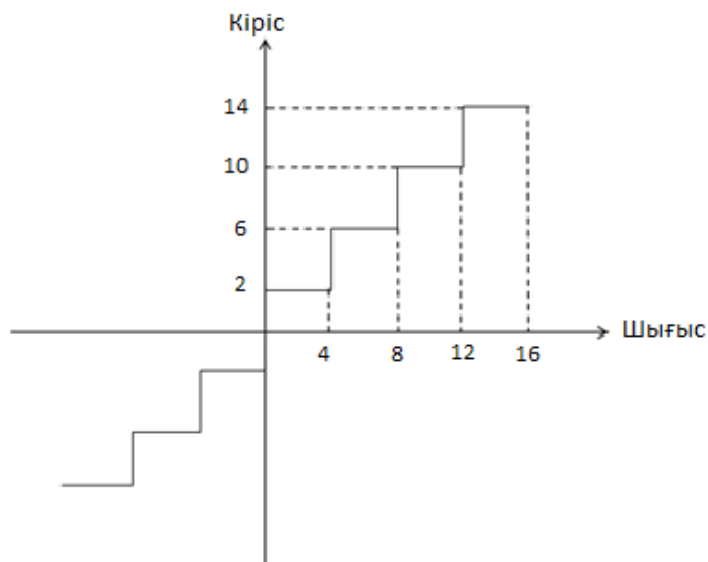
Орта мәні 0 және дисперсиясы 4-ке тең PDF форматындағы Лапласиан көзі үшін үш биттік кванттаушы көрінісін алу қажет.

Сигнал Лаплас тығыздығымен берілген. Осылайша біз сигнал үшін $X_{\max} = 4\sigma = 16$ деп болжанады, себебі $X_{\max} = 4\sigma_x$.

$$q = \frac{2 \times \text{Шыңдық мән}}{2^3} = \frac{2 \times 16 \text{ В}}{8} = 4\text{В}$$

$$\sigma_e^2 = \frac{q^2}{12} = 1.33$$

Орташа мәнді біркелкі кванттаушы келесі суретте көрсетілген (Сурет 2.7).



Сурет 2.7 – Орташа мәнді біркелкі кванттаушы

2.6 DPCM бойынша параметрлерді есептеу

Сигналды кванттау үшін оның динамикалық диапазонын азайту қажет. Сигналдың динамикалық диапазоны дифференциалды кодтау көмегімен азайтылуы мүмкін. Сигналды тікелей кодтаудың орнына сигнал кезекті таңдаулардың арасындағы айырмашылықтарды қарастырады. Сөйлеу сигналы кезекті таңдаулар арасында жоғары корреляцияға ие болғандықтан, сигналдар арасындағы айырмашылық үлкен болмайды.

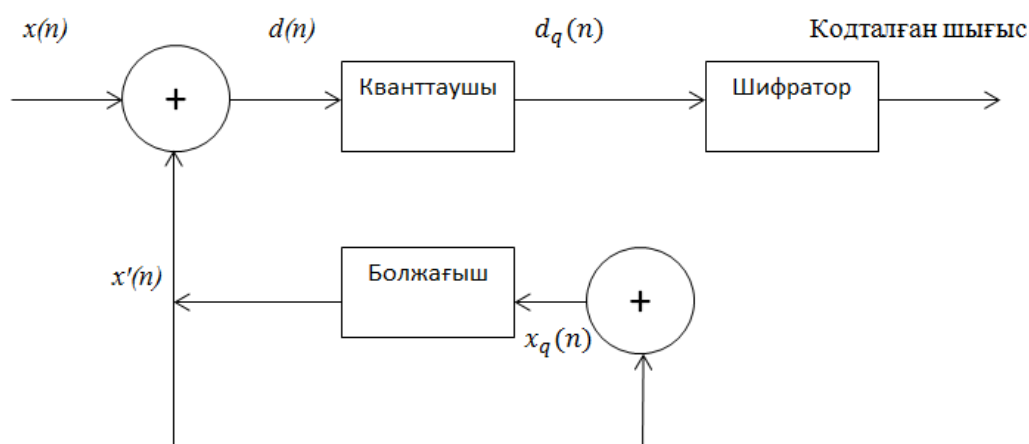
Бұл сигналдың динамикалық диапазонын азайтады.

Жүйенің өнімділігін қамтамасыз ету үшін келесі жағдайлар назарға аынуы тиіс:

1. Егер сигналдың динамикалық диапазоны аз болса, онда қадам өлшемі және кванттау шуы сәйкесінше аз болады. Бұл сигнал/кванттау шу қатынасына сәйкес келетін SNR өлшеусіз шаманы әлдеқайда үлкейтеді.

2. Егер SNR тұрақтысын сақтайтын болсақ, онда кванттау үшін қоланылға биттер санын азайту қажет болады. Бұл ақпараттық тарату жылдамдығын азайтуға және сөйлеу сигналын кодтаудың тиімділігін жоғарлатуға мүмкіндік береді.

Кванттау үшін сигналдың динамикалық диапазонын азайту үшін тұрақты және болжанатын сөйлеу таңдауларының арасындағы айырмашылықты кодтаймыз. Айырымды сигнал болжау қателігінің сигналы деп аталады. Болжау қателігі бір немесе одан да көп болуы мүмкін, бұл биттер санын азайтып, SNR-ді жақсартуға көмектеседі[27].



Сурет 2.7 – DPCM блок-схемасы

$$\begin{cases} d(n) = x(n) - x'(n) \\ d_q(n) = d(n) + e(n) \end{cases} \quad (2.8)$$

Мұндағы,

$x(n)$ – квантталған сигнал;

$x'(n)$ – болжанатын сигнал;

$e(n)$ -сигналды болжау қателігі;

$d_q(n)$ – квантталған айырымды сигнал;

$x_q(n)$ –квантталған сигнал.

$$\begin{cases} x_q(n) = x'(n) + d_q(n) \\ x_q(n) = x(n) + e(n) \end{cases} \quad (2.9)$$

Кездейсоқ сигналдың дисперсиясы сигнал дисперсиясынан аз болады, бұл болжағышқа байланысты, себебі кванттаушы қадам өлшемінің аз мәнімен спектрленуі мүмкін. Бұл сигнал/шу қатынасын әлдеқайда жақсартады.

Келесі есептеу формулаларында DPCM жүйесі үшін SNR есептеледі:

$$SNR = \frac{E[X^2(n)] \sigma_x^2}{E[e^2(n)] \sigma_e^2} \quad (2.10)$$

$$SNR = \frac{\sigma_x^2 \sigma_d^2}{\sigma_d^2 \sigma_e^2} = G_p SNR_Q \quad (2.11)$$

Мұндағы,

SNR - кванттаушының сигнал/шу қатынасы;

G_p -дифференциалды конфигурация үшін күшейту коэффициенті.

Келесі жағдайда сигнал Лаплас тығыздығына бағынышты болсын. Оның дисперсиясы 3В деп алынған. Айырымды сигнал дисперсиясы 1В.

Дифференциалды конфигурация үшін күшейту коэффициентін табу қажет.

Дифференциалды конфигурация үшін күшейту коэффициенті келесе формуламен есептеледі:

$$G = \frac{\sigma_x^2}{\sigma_d^2} = \frac{9}{1} = 9$$

Күшейту коэффициентінің 9-ға тең мәніне байланысты SNR 9 рет үлкееді.

Енді сигнал үлгілерін 1.8, 2.4, 3.5, 5.1, 6.2, 9.7, 11.2, 13.2, 12.6, 10.2, 7.3, 5.9, 8, 7.4, 5.2, 3.1 және 1.8 деп алатын болсақ, сигнал және айырымды сигнал үшін қадам өлшемін және кванттау шуының қуаты есетелу қажет. Бұл жағдайда төртразрядты кванттаушы қолданылады.

Сигналдың дианамикалық диапазоны 13.2-1.8В = 11.4В-ке тең.

$$q = \frac{11.4В}{2^4} = 0.7125В$$

$$\sigma_e^2 = \frac{q^2}{12} = 0.0423046$$

Айырымды сигналдың мәндері 1.8, 0.6, 1.2, 1.6, 1.2, 3.5, 3.5, 2, -0.6, -2.4, -2.9, 1.1, -0.6, -2.2, -2.1, -1.3.

Дианамикалық диапазон 5.5В-ке тең деп алынған.

$$q = \frac{5.5В}{2^4} = 0.34375 В$$

$$\sigma_e^2 = \frac{q^2}{12} = 0.009847$$

Қадам өлшемі динаикалық диапазон азайған сайын аз мәнге ие болады. Егер қадам өлшемін тұрақты етсек, таңдауда биттер саны келесідей болады:

$$\frac{5.5В}{2^4} = 0.7125 В$$

$$2^B = 7.7192$$

$$B = 3$$

3 Matlab бағдарламасында сөйлеу сигналының параметрлері зерттеу

MATLAB – бұл бағдарламалау, сандық есептеу және нәтижелерді визуализациялауға қажетті жоғарғы деңгейлі тіл және интерактивті орта болып келеді. MATLAB мәліметтерді талдауға, алгоритмдерді әзірлеуге және модельдер мен қосымшалары құруға мүмкіндік береді[28].

MATLAB келесі салаларда кең қолданысқа ие:

1. Сигналдарды өңдеу және байланыс;
2. Бейнелер мен мультимедияны өңдеу;
3. Басқару жүйелері;
4. Тестілеу және өлшемдерді автоматтандыру;
5. Қаржылық инженеринг және т.б.

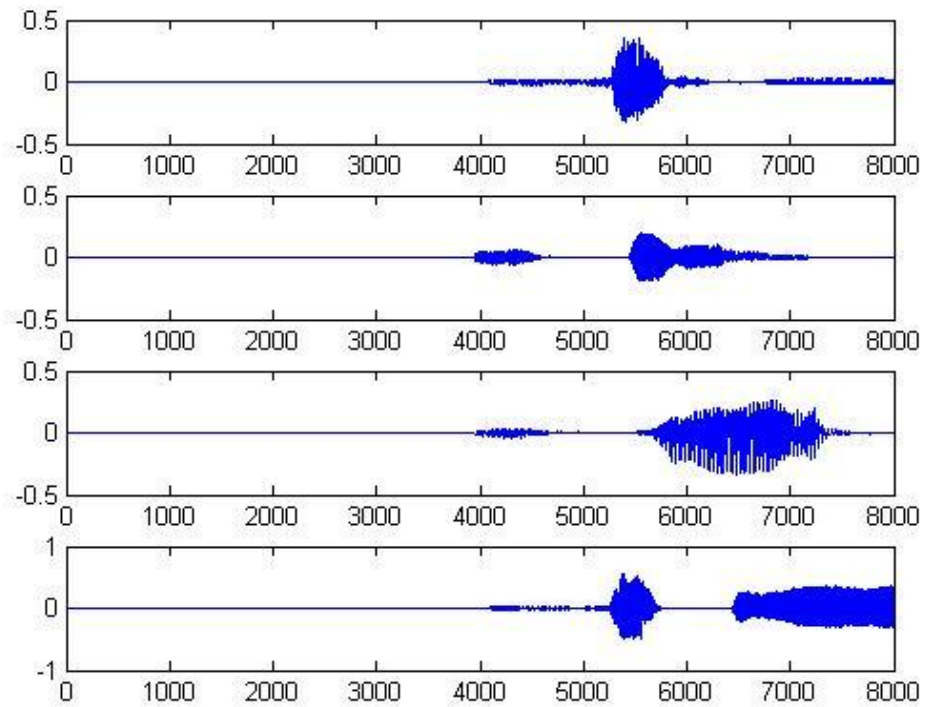
Бұл бөлімде MATLAB бағдарламасы көмегімен сөйлеу сигналын зерттеу, оның параметрлері мен көрінісінің бейнесін алу және сөйлеу сигналының кодтық моделін алу көзделеді[29].

3.1 MATLAB бағдарламасында «Даяшы» сөзіне зерттеу жүргізу

Қазақ тілінде алынған «Даяшы» сөзінің уақыттық, дискретті көрсеткіштерін және кодтық талдауын алып, оның эмоционалдық өзгерісін бақылау.

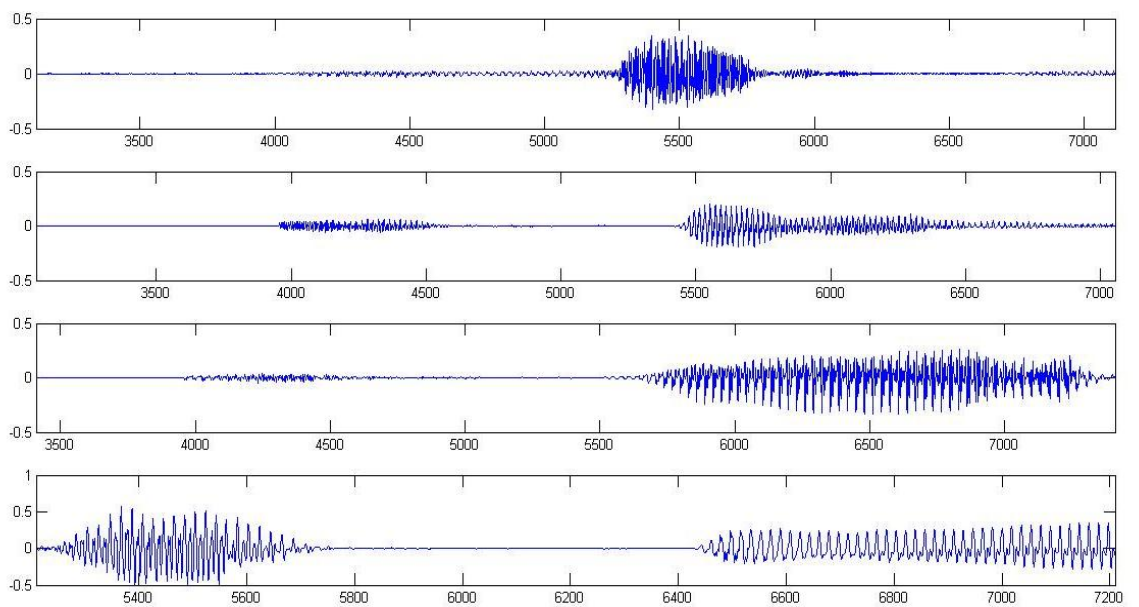
Төменгі суретте қыз дауысының әртүрлі интонацияда айтылған «Даяшы» сөзінің уақыттық сипаттамасы алынған (Сурет 3.1)

1. Ренжу эмоциясы
2. Көтеріңкі қуану эмоциясы
3. Көңіл толмаушылық ашулы эмоция
4. Көтеріңкі ойнақы эмоция



Сурет 3.1 – «Даяшы» сөзінің уақыттық көрінісі

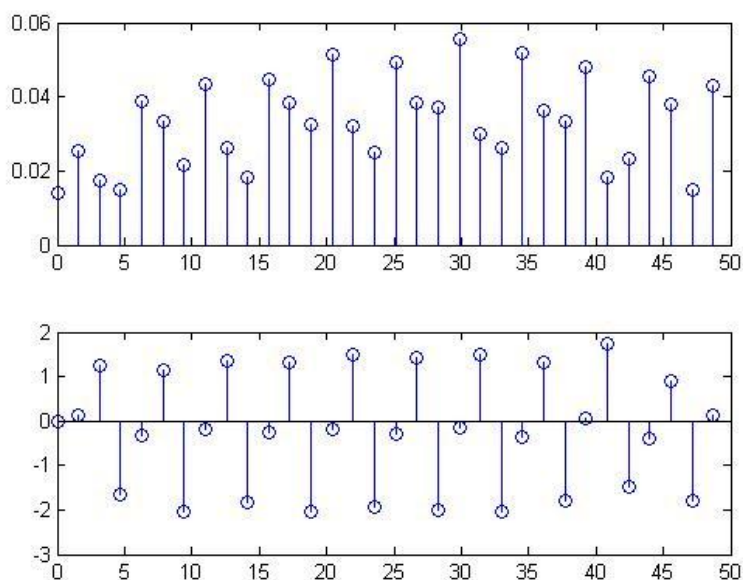
Абцисса осін созу арқылы дауыстың ерекшеліктерін қарауға болады (Сурет 3.2)



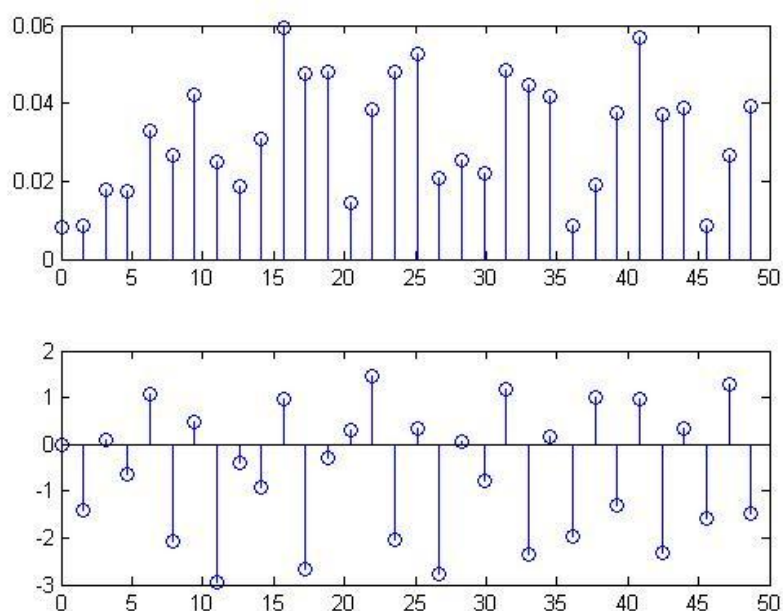
Сурет 3.2 - Абцисса осін созу арқылы дауыстың ерекшеліктерін бақылау

Келесі жағдайда «Даяшы» сөзінің амплитуда және жиілік бойынша дискретті сипаттамалар алынған (Сурет 3.1)

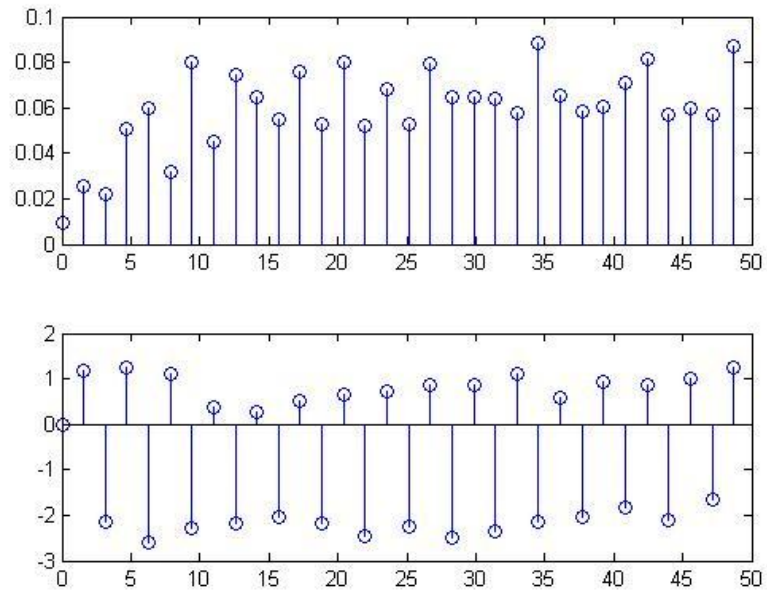
1. Ренжу эмоциясы (Сурет 3.2);
2. Көтеріңкі қуану эмоциясы (Сурет 3.3);
3. Көңіл толмаушылық ашулы эмоция (Сурет 3.4);
4. Көтеріңкі ойнақы эмоция (Сурет 3.5);



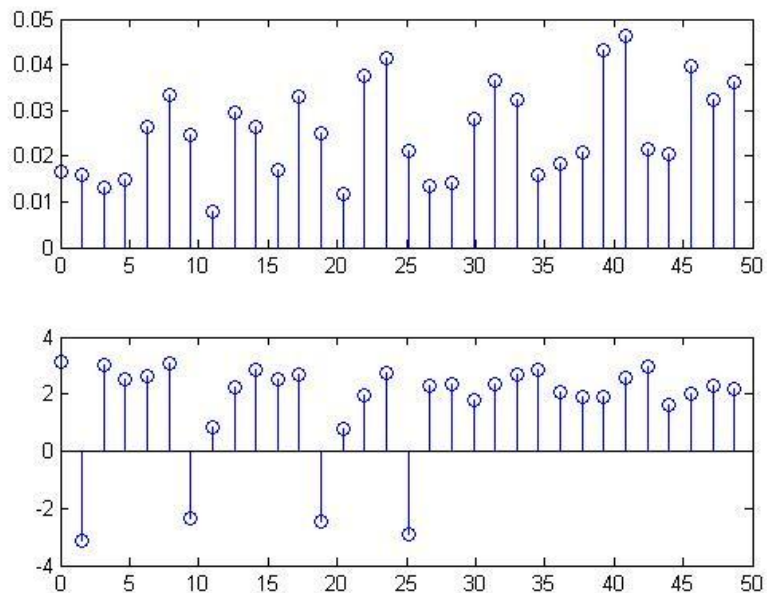
Сурет 3.2 - Ренжу эмоциясының амплитуда және жиілік бойынша дискретті сипаттамалар



Сурет 3.3 - Көтеріңкі қуану эмоциясының амплитуда және жиілік бойынша дискретті сипаттамалар



Сурет 3.4 - Көңіл толмаушылық ашулы эмоцияның амплитуда және жиілік бойынша дискретті сипаттамалар



Сурет 3.5 - Көтеріңкі ойнақы эмоцияның амплитуда және жиілік бойынша дискретті сипаттамалар

Келесі жағдайда дауыстың параметрлерін салыстыру кестесі ұсынылған. Сигналдарды бір-бірімен салыстыру үшін оның энергиясын немесе оның қуатын қолдану керек.

Сигналдың қуатын қолдану үшін сигнал периодты немесе статистикалық периодтығы болу керек. Адам дауысының периодтылығы болмайды, сондықтан

ты қолдану ыңғайсыз. Сонымен қатар адамның дауысы сөйлеу кезінде жиілікті өзгертіп тұрады.

Кесте 3.1-де дауыстың келесі параметрлері бойынша салыстыру анализы жүргізілді:

1) Сигналдың энергиясын табу үшін дыбыстың әр уақыт мезетіндегі мәнін, яғни амплитудасын квадраттап, оның толық суммасы табылды. MatLab-қа кодтық комбинация (код қосымшада тіркелген) терілді.

2) Амплитуда-жиіліктік сипаттама (АЖС) арқылы көңіл толмаушылық эмоциясында әр жиілікте амплитуда бірдей, қалған үш эмоцияда әр жиіліктің амплитудалары әртүрлі болады.

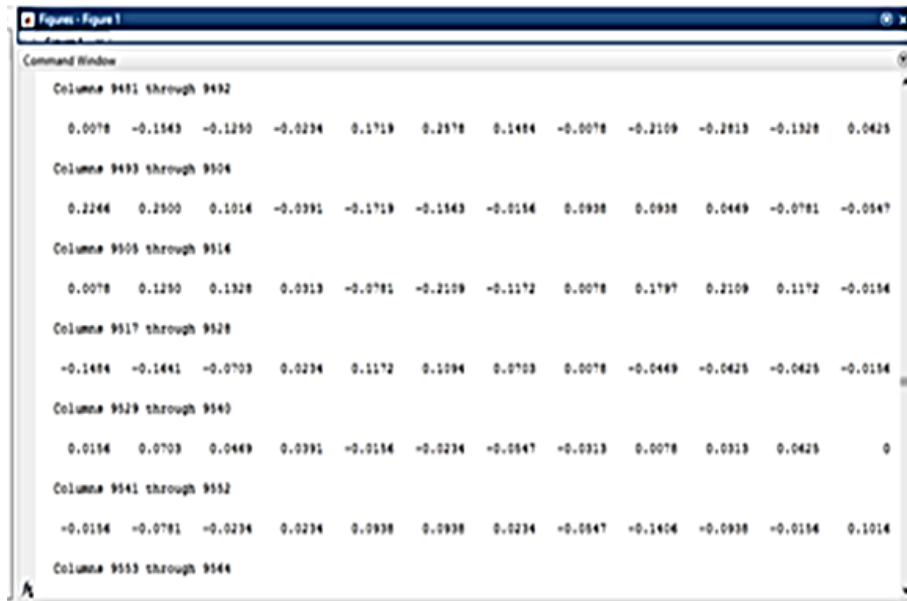
3) Фаза- жиіліктік сипаттама(ФЖС) арқылы көтеріңкі ойнақы эмоция кезінде толқындар фаза бойынша ығыспайды, ал қалған үш сигналдың фазалары өзгеріп тұрады.

Кесте 3.1- Дауыстың параметрлері бойынша салыстырмалы кесте

	Ренжу эмоциясы	Көтеріңкі қуану эмоциясы	Көңіл толмаушылық ашулы эмоция	Көтеріңкі ойнақы эмоция
Дауыс энергиясы	11,0810	4,7307	18,3014	57,1644
АЖС	0,03-0,05	0,02-0,05	0,07 (const)	0,02-0,04
ФЖС	2,5	1,25	1,25	0 (const)

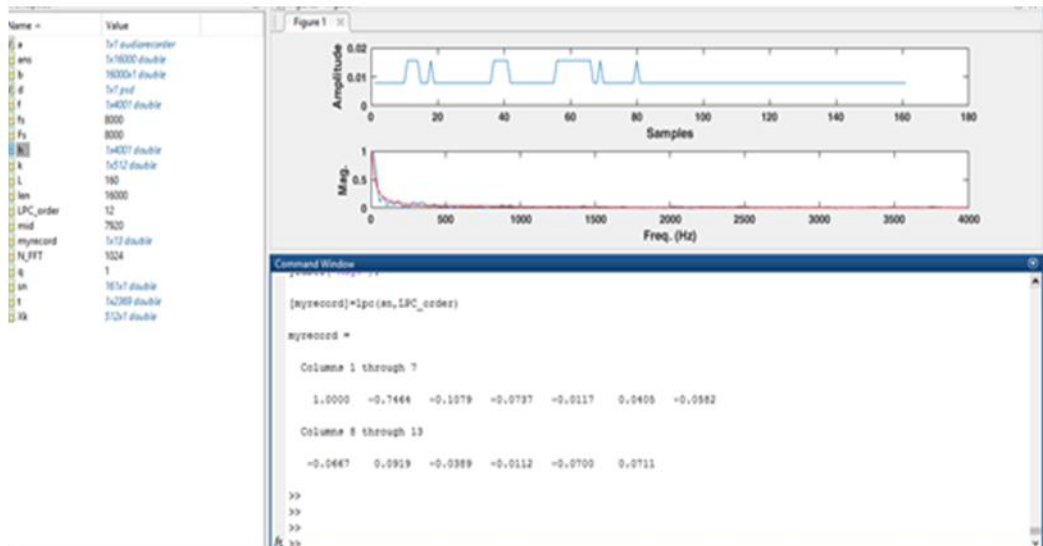
Келесі жағдайда «Даяшы» сөзінің бір фонемасын таңдалып, оған LPC кодтау әдісін қолдану арқылы кодтау анализы жүргізілді. Зерттелетін фонема ретінде «Даяшы» сөзіндегі «я» дауысты дыбысы алынған.

Әрбір фонеманың матрица түріндегі дискретті мәні болады. Келесі суретте «я» фонемасының дискретті мәні Matlab бағдарламасы арқылы алынған (Сурет 3.6).



Сурет 3.6 – Матрица түріндегі «я» фонемасының дискретті мәні

Келесі жағдайда фонеманың дискретті мәніне сүйене отырып, «я» фонемасының формантты жиіліктерінің мәндері алынды[30]. Бірінші фонемасының формантты жиілігі $F1 = 312,5$. Екінші формантты жиілігі $F2 = 1123$ және $F3 = 2201$ Гц и $F4 = 3443$ Гц.



Сурет 3.7 – «я» фонемасының сызықтық кодтық мәндері

Сурет 3.7-те «я» фонемасының сызықтық кодтық мәндері (LPC - коэффициенттері) алынды. Сызықтық болжаумен кодтау әдісі (LPC) - бұл негізінен дыбыстық сигналды өңдеу және сызықтық алдын ала болжайтын модель ақпаратын пайдалана отырып, сығылған формада сөйлеуді өңдеу үшін қолданылатын әдіс ретінде ұсынылған.

ҚОРЫТЫНДЫ

Дипломдық жұмыста сөйлеу сигналын кодтаудың оңтайлы әдістері қарастырылды. Сөйлеуді кодтау үшін аналогты сигналдың сандық сигналға түрлену процесі және сөйлеуді түрлендіру моделінің құрылымы көрсетілді. Сөйлеу сигналын кодтаудың оңтайлы әдісі – сызықты болдау әдісі және оның негізгі түрлері: негізгі тонның жиілігінен қоздырылатын LPC кодтау әдісі, RELP немесе сөйлеуді гибриді кодтау әдісі, ABS кодтау әдісі, кодтық кітаптар негізінде қозуы бар кодтар немесе CELP кодтау әдістеріне шолу жасалынды.

Сөйлеу сигналын кодтаудың негізгі параметрлерін зерттеу жүргізілді. Есептеулер сигналды дискреттеу мен кванттауға негізделген. Дискретизация периоды T және дискретизация жиілігі анықталып, сигнал жиілігі 4кГц болғанда дискретизация жиілігі $F_d=8000$ Гц-ке тең деп алынды. Сөйлеу сигналын кванттаудың бастапқы параметрлерін есептеуде: қадам өлшемі- q , кванттау шуының қуаты- σ_e^2 , сигналдың динамикалық диапазоны, кванттау шуының дисперсиясы, сигнал дисперсиясы, шындық мән, өлшеусіз шама SNR қолданылды. DPCM үшін дифференциалды конфигурацияның күшейту коэффициенті есептелінді. PCM және DPCM әдістерін есептеулер бойынша салыстыру жүргізілді.

MATLAB бағдарламасы көмегімен сөйлеу сигналына зерттеулер жүргізілді, оның параметрлері мен көрінісінің бейнесі және сөйлеу сигналының кодтық моделін алынды. Қазақ тілінде алынған «Даяшы» сөзінің уақыттық, дискретті көрсеткіштері және кодтық талдауы алынып, оның эмоционалдық өзгерісі бақыланды. Дауыс параметрлері бойынша зерттеу жүргізіліп, салыстырмалы кесте құрылды.

Бұл дипломдық жұмыс IP- телефония саласында теориялық және практикалық шешімдерде қосымша материал ретінде қолданыла алады.

ПАЙДАЛАНЫЛҒАН ӘДЕБИЕТТЕР ТІЗІМІ

- 1 Сергиенко А.Б. Цифровая обработка сигналов. 2 – изд. – СПб.: Питер, 2006. – 608 с.
- 2 Солонина А. И., Улахович Д. А., Арбузов С. М., Соловьева Е. Б. Основы цифровой обработки сигналов. 2-е изд. — СПб.: БХВ-Петербург, 2005.
- 3 Ст. преподаватель кафедры ПОСТ Киркоров С.И. на основе лекции Борискевича А.А. Курс: Цифровая обработка речи и изображения / Сжатие речевых и аудиосигналов на основе модели речеобразования. -Минск, 2010.
- 4 Солонина А. И., Улахович Д.А., Яковлев Д.А. Алгоритм и процессоры цифровой обработки сигналов — СПб.: БХВ-Петербург, 2001.
- 5 Сэломон Д. Сжаие данных. — М.: Техносфера, 2004.
- 6 Верещагин, Н. К. Информация, кодирование и предсказание / Н.К. Верещагин, Е.В. Щепин. - М.: ФМОП, МЦНМО, 2012. - 240 с.
- 7 <http://www.intuit.ru>. Интернет-Университет Информационных Технологий.
- 8 Доц. каф. 40 Большов О.А. Реферат: «Информационная безопасность и защита информации /Вокодеры с линейным предсказанием». — Москва, 2010.
- 9 Питер Кирн Цифровой звук. Реальный мир. — 2008. — С.65
10. www.rasip.fer
11. Ю.А. Ковалгин, Э.И. Вологдин. Цифровое кодирование звуковых сигналов. – М.: Корона-Принт, 2015. – 240 с.
12. William A. Pearlman, Amir Said. Digital Signal Compression: Principles and Practice. — Cambridge University Press, 2011. — P.83
13. Савчук В.Л. Интернет-ресурс <http://www.ie.tusur.ru>
14. <http://www.bnti.ru/>. Бюро научно-технической информации. По материалам 2-ой Всероссийской конференции "Теория и практика речевых исследований".
15. Запороженко Н.П., Карташевский В.Г., Клиентова Т.Г., Харченко Ю.Ю. Цифровая коммутационная система АХЕ-10: Учебное пособие для вузов.-М.: Радио и связь, 2000.
- 16 Dr.Shaila, D.Apte Speech and Audio Processing. New Delhi, 2017. — P.143
- 17 Method of Defining Multimodel Information Falsity for Smart Telecommunication Systems / O.O. Basov, A.L. Ronzhin, V.Yu. Budkov, I.A. Saitov // LNCS. – 2015. – P. 163–176.
- 18 Basov O.O. Optimization of Pitch Tracking and Quantization / O.O. Basov, An.L. Ronzhin, V.Yu. Budkov //LNAI. – 2015. – P. 317–324.
- 19 Вольф Д.А. Модель процесса сингулярного оценивания частоты основного тона речевого сигнала /Д.А. Вольф, Р.В. Мещеряков // Акустический журнал. – 2016. – Т. 62, № 2. – С. 216.
- 20 Ю.А. Ковалгин, Э.И. Володин, Л.Н. Кацнельсон. Стереофоническое радиовещание и звукозапись. Учебное пособие. – М.: Горячая Линия - Телеком, 2014. – 720 с.

- 21_Gerkmann T. Unbiased MMSE-Based Noise Power Estimation With Low Complexity and Low Tracking Delay / Gerkmann T., Hendriks R. // IEEE Trans Audio, Speech, Language Proc. – 2012. – No 20. – P. 1383–1393.
- 22 1. Лайонс, Р. Цифровая обработка сигналов / Р. Лайонс. - М.: Корона-Век, 2015. - 611 с.
- 23 Юкио, Сато. Цифровая обработка сигналов / Сато Юкио. - М.: Додэка-XXI, 2017. - 885 с.
- 24 Айфичер, Э. Цифровая обработка сигналов: практический подход / Э. Айфичер. - М.: Вильямс, 2016. - 992 с
- 25 Гадзиковский, В.И. Цифровая обработка сигналов / В.И. Гадзиковский. - М.: Солон-пресс, 2015. - 766 с.
- 26 Чан, Т.Т. Высокоскоростная цифровая обработка сигналов и проектирование аналоговых систем / Т.Т. Чан; Пер. с англ. К.В. Юдинцев. - М.: Техносфера, 2013. - 192 с.
- 27 Воробьев, С.Н. Цифровая обработка сигналов / С.Н. Воробьев. - М.: Academia, 2018. - 64 с.
- 28 Ian McLoughlin. Applied Speech and Audio Processing: With MATLAB Examples— Cambridge University Press, 2009
- 29 Luis F. Chaparro Signals and Systems Using MATLAB; Elsevier - Москва, 2011. - 768 с.
- 30 Солонина, А.И. Цифровая обработка сигналов и MATLAB: Учебное пособие /А.И. Солонина, Д.М. Клионский, Т.В. Меркучева. - СПб.: БХВ-Петербург, 2013. - 512 с.

ҚОСЫМША

```
script
clear
clc

%Дауысты жазу

Fs=8000;
z=zeros(Fs, 4);
for n = 1:4
    'Say'
y=wavrecord(1*Fs, Fs, 1, 'double');

l=410;
subplot(l+n), plot(y)
z(((n-1)*Fs)+1):n*Fs)=y;
end

%Матрицадан дауысты жазу

figure
subplot(411), plot((1:Fs), z(1:Fs));
subplot(412), plot((1:Fs), z((Fs+1):(2*Fs)));
subplot(413), plot((1:Fs), z((2*Fs+1):(3*Fs)));
subplot(414), plot((1:Fs), z((3*Fs+1):(4*Fs)));
figure

%АЖС және ФЖС 1 дауыс

g=z(1:Fs);
Ts=1/64; T0=4; N0=T0/Ts;
t=0:Ts:Ts*(N0-1); t=t';
g(1)=Ts*0.5;
G=fft(g);
[Gp, Gm]=cart2pol(real(G), imag(G));
k=0:N0-1; k=k';
w=2*pi*k/T0;
subplot(211), stem(w(1:32), Gm(1:32));
subplot(212), stem(w(1:32), Gp(1:32));
figure

%АЖС және ФЖС 2 дауыс

g=z((Fs+1):(2*Fs));
Ts=1/64; T0=4; N0=T0/Ts;
t=0:Ts:Ts*(N0-1); t=t';
g(1)=Ts*0.5;
G=fft(g);
[Gp, Gm]=cart2pol(real(G), imag(G));
k=0:N0-1; k=k';
w=2*pi*k/T0;
```

```
subplot(211), stem(w(1:32), Gm(1:32));
subplot(212), stem(w(1:32), Gp(1:32));
figure
```

```
%АЖС және ФЖС 3 дауыс
```

```
g=z((2*Fs+1):(3*Fs));
Ts=1/64; T0=4; N0=T0/Ts;
t=0:Ts:Ts*(N0-1); t=t';
g(1)=Ts*0.5;
G=fft(g);
[Gp, Gm]=cart2pol(real(G), imag(G));
k=0:N0-1; k=k';
w=2*pi*k/T0;
subplot(211), stem(w(1:32), Gm(1:32));
subplot(212), stem(w(1:32), Gp(1:32));
figure
```

```
%АЖС және ФЖС 4 дауыс
```

```
g=z((3*Fs+1):(4*Fs));
Ts=1/64; T0=4; N0=T0/Ts;
t=0:Ts:Ts*(N0-1); t=t';
g(1)=Ts*0.5;
G=fft(g);
[Gp, Gm]=cart2pol(real(G), imag(G));
k=0:N0-1; k=k';
w=2*pi*k/T0;
subplot(211), stem(w(1:32), Gm(1:32));
subplot(212), stem(w(1:32), Gp(1:32));
```

```
%Сигнал энергиясын табу коды
```

```
energy1=(z(1:Fs).*z(1:Fs));
energy2=(z((Fs+1):(2*Fs)).*z((Fs+1):(2*Fs)));
energy3=(z((2*Fs+1):(3*Fs)).*z((2*Fs+1):(3*Fs)));
energy4=(z((3*Fs+1):(4*Fs)).*z((3*Fs+1):(4*Fs)));
disp('Energy of 1-st voice: '), sum(energy1)
disp('Energy of 2-nd voice: '), sum(energy2)
disp('Energy of 3-rd voice: '), sum(energy3)
disp('Energy of 4-th voice: '), sum(energy4)
```

**ҒЫЛЫМИ ЖЕТЕКШІНІҢ
СЫН-ШІКІРІ**

дипломдық жұмысына

Канатбекова Д.А.

5B071900-Радиотехника, электроника және телекоммуникациялар

Тақырыбына: «Сөйлеу сигналын кодтаудың оңтайлы әдістерінің анализі»

Бұл дипломдық жұмыста сөйлеу сигналын кодтаудың оңтайлы әдістерін қолданып, сөйлеу сигналының негізгі кодтау әдістерін талдау мақсаты қойылған.

Қойылған мақсатқа сәйкес дипломдық жұмыс көлемінде келесі мәселелер қарастырылған:

- Сөйлеу сигналын түрлендіру құрылымымен танысу;
- Сөйлеу сигналын кодтаудың негізгі әдістерін қарастыру;
- Сөйлеу сигналының параметрлерін зерттеу;
- Сөйлеу сигналын бастапқы кодтау әдістеріне есептеу жүргізу;
- Сөйлеу сигналын Matlab бағдарламасында моделдеу;

Сөйлеуді кодтау үшін аналогты сигналдың сандық сигналға түрлену процесі және сөйлеуді түрлендіру моделінің құрылымы көрсетілген. Сөйлеу сигналын кодтаудың оңтайлы әдістерінің бірі – сызықты болжау әдісі және сызықты болжау әдісінің негізгі түрлері қарастырылған.

Сөйлеу сигналын кодтаудың негізгі параметрлері талданған. Кванттаудың бастапқы параметрлеріне және дифференциалды кодтық модуляция бойынша есептеулер жүргізілген.

MATLAB бағдарламасы көмегімен сөйлеу сигналы зерттеліп, сөйлеу сигналының уақыттық, спектральдық және кодтық сипаттамаларының моделі алынған. Қазақ тіліндегі сөздің сипаттамалары алынып, эмоционалдық көрінісі бойынша талдау жүргізілген. Дауыс параметрлері бойынша салыстырмалы кесте ұсынылған.

Бұл дипломдық жұмыс жоғарғы оқу орындарының талаптарына сай жеткілікті жоғары дәрежеде жазылған, нәтижелер ақпаратты бөгеуілден қорғау және тарату жүйелерінің ғылыми бағытына жауап береді.

Жалпы, дипломдық жұмысты "90" (А) өте жақсы деген бағаға, ал студент Канатбекова Д. А. 5B071900 – Радиотехника, электроника және телекоммуникация мамандығы бойынша техника және технологиялар «бакалавр» академиялық дәрежесіне ұсынылады.

Ғылыми жетекші

ЭТЖҒТ кафедрасының
қауымдастырылған профессоры

 Л.Б. Илипбаева

« 26 » _____ 2019 ж.

СЫН – ПІКІР

дипломдық жұмыс

Канатбекова Диана Аязбековна

**5B071900 - Радиотехника, электроника және телекоммуникациялар
мамандығы бойынша**

Тақырыбы: «Сөйлеу сигналын кодтаудың оңтайлы әдістерінің анализі»

Орындалды:

а) графикалық бөлімі _____ бет;

ә) түсіндірме жазбасы _____ бет.

ЖҰМЫСҚА ЕСКЕРТУ ЖАСАУ

Бұл дипломдық жұмыста Канатбекова Диана Аязбековна сөйлеу сигналын кодтаудың оңтайлы әдістерінің анализін сараптаған. Дипломдық жұмыс келесі мәселелерді қарастырады:

Біріншіден, сөйлеуді кодтау үшін аналогты сигналдың сандық сигналға түрлену процесі және сөйлеуді түрлендіру моделінің құрылымы көрсетілген. Дауысты және дауыссыз дыбыстырдың айырмашылықтары мен ерекшеліктері олардың фрагменттері мен спектралды көрінісі арқылы зерттелінген. Сөйлеу сигналын кодтаудың оңтайлы әдістерінің бірі – сызықты болжау әдісі және сызықты болжау әдісінің негізгі түрлері қарастырылған.

Екіншіден, сөйлеу сигналын кодтаудың негізгі параметрлеріне зерттеу жүргізілді. Есептеулер сигналды дискреттеу мен кванттау параметрлерін есептеуге негізделген. Тарауда әртүрлі кванттау әдістері, біркелкі кванттау, дискретті кванттаудың біркелкі емес әдістері, дифференциалды кодтау әдістер негізінде келесі кванттаудың қадам өлшемі, кванттау шуы және кванттау шуының дисперсиясы, сигнал дисперсиясы, дифференциалды конфигурация үшін күшейту коэффициенті сияқты параметрлерге есептеу жүргізілген.

Үшіншіден, MATLAB бағдарламасы көмегімен сөйлеу сигналын зерттеу, сөйлеу сигналының уақыттық, спектральдық және кодтық сипаттамаларының моделі алынған. Қазақ тілінде алынған сөздердің сипаттамалары анықталып, эмоционалдық көрінісі бойынша өзгерісі бақыланды.

Дипломдық жұмыста стилистикалық қателіктер орын алған, бірақ бұл дипломдық жұмыстың өзектілігін төмендетпейді.

Дипломдық жұмыс жоғарғы оқу орындарының талаптарына сай жеткілікті жоғары дәрежеде жазылған, алынған нәтижелер мамандығы бойынша ғылыми бағытқа жауап береді.

Жұмыс бағасы

Жалпы, дипломдық жұмыс "90/A/- өте жақсы" деген бағаға, ал студент Канатбекова Д. А. 5B071900 – Радиотехника, электроника және телекоммуникация мамандығы бойынша техника және технологиялар «бакалавр» академиялық дәрежесіне ұсынылады.

Пікір беруші

ҚазҰАУ, ЭУЖА каф. меңгерушісі,

доктор PhD,

қауымдастырылған профессор

Ж. С. Шыныбай

« 29 » 04 2019 ж

Ф. ҚазҰТЗУ-704-24. Пікір

