

ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМИЙ ДАРАЖАЛАР БЕРУВЧИ
DSc.13/30.12.2019.T.07.01 РАҚАМЛИ ИЛМИЙ КЕНГАШ

ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ

ШУКУРОВ КАМОЛИДДИН ЭЛЬБОБО ЎҒЛИ

**НУТҚ БУЙРУҚЛАРИНИ ТАНИШ АЛГОРИТМЛАРИ ВА АППАРАТ-
ДАСТУРИЙ ПЛАТФОРМАСИНИ ИШЛАБ ЧИҚИШ**

05.01.04 – Ҳисоблаш машиналари, мажмуалари ва компьютер
тармоқларининг математик ва дастурий таъминоти

ТЕХНИКА ФАНЛАР БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ

Тошкент – 2021

**Техника фанлари бўйича фалсафа доктори (PhD) диссертацияси
автореферати мундарижаси**

**Оглавление автореферата диссертации доктора философии (PhD)
по техническим наукам**

**Contents of dissertation abstract of doctor of philosophy (PhD)
on technical sciences**

Шукуров Камолиддин Эльбобо ўғли

Нутқ буйруқларини таниш алгоритмлари ва аппарат-дастурий
платформасини ишлаб чиқиш..... 3

Шукуров Камолиддин Эльбобо ўғли

Разработка алгоритмов и аппаратно-программной платформы
распознавания речевых команд..... 21

Shukurov Kamoliddin Elbobo o'g'li

Development of algorithms and hardware-software platform for speech
commands recognition..... 39

Эълон қилинган ишлар рўйхати

Список опубликованных работ
List of published works..... 43

ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ
ҲУЗУРИДАГИ ИЛМИЙ ДАРАЖАЛАР БЕРУВЧИ
DSc.13/30.12.2019.T.07.01 РАҚАМЛИ ИЛМИЙ КЕНГАШ

ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ УНИВЕРСИТЕТИ

ШУКУРОВ КАМОЛИДДИН ЭЛЬБОБО ЎҒЛИ

**НУТҚ БУЙРУҚЛАРИНИ ТАНИШ АЛГОРИТМЛАРИ ВА АППАРАТ-
ДАСТУРИЙ ПЛАТФОРМАСИНИ ИШЛАБ ЧИҚИШ**

05.01.04 – Ҳисоблаш машиналари, мажмуалари ва компьютер
тармоқларининг математик ва дастурий таъминоти

ТЕХНИКА ФАНЛАР БЎЙИЧА ФАЛСАФА ДОКТОРИ (PhD)
ДИССЕРТАЦИЯСИ АВТОРЕФЕРАТИ

Тошкент – 2021

Техника фанлари бўйича фалсафа доктори (PhD) диссертацияси мавзуси Ўзбекистон Республикаси Вазирлар Маҳкамаси ҳузуридаги Олий аттестация комиссиясида В2021.4.PhD/T902 рақам билан рўйхатга олинган.

Диссертация Тошкент ахборот технологиялари университетиде бажарилган.

Диссертация автореферати уч тилда (ўзбек, рус, инглиз (резюме)) Илмий кенгаш веб саҳифасида (www.tuit.uz) ва «Ziynet» ахборот-таълим порталида (www.ziynet.uz) жойлаштирилган.

Илмий раҳбар: Мусаев Муҳаммаджон Маҳмудович
техника фанлари доктори, профессор

Расмий оппонентлар: Назирова Элмира Шодмоновна
техника фанлари доктори, доцент

Авазов Юсуф Шодиевич
PhD, доцент

Ётакчи ташкилот: Тошкент давлат транспорт университети

Диссертация химояси Тошкент ахборот технологиялари университети ҳузуридаги DSc.13/30.12.2019.T.07.01 рақамли Илмий кенгашининг 2021 йил 28 «декабрь» да соат 16:00 даги мажлисида бўлиб ўтди. (Манзил: 100202, Тошкент шаҳри, Амир Темур кўчаси, 108-уй. Тел.: (99871) 238-64-43; факс: (99871) 238-65-52; e-mail: tuit@tuit.uz).

Диссертация билан Тошкент ахборот технологиялари университетининг Ахборот-ресурс марказида танишиш мумкин (Q237 рақам билан рўйхатга олинган). Манзил: 100202, Тошкент шаҳри, Амир Темур кўчаси, 108-уй. Тел.: (99871) 238-64-43).

Диссертация автореферати 2021 йил «16» декабрь кунини тарқатилди.
(2021 йил «14» декабрь даги 42 рақамли реестр баённомаси).



Р.Х.Хамдамов
Илмий даражалар берувчи илмий кенгаш раиси, техника фанлар доктори, профессор

Ф.М.Нуралшев
Илмий даражалар берувчи илмий кенгаш илмий котиби, техника фанлар доктори, доцент

М.А.Рахматуллаев
Илмий даражалар берувчи илмий кенгаш ҳузуридаги илмий семинар раиси, техника фанлар доктори, профессор

КИРИШ (фалсафа доктори (PhD) диссертацияси аннотацияси)

Диссертация мавзусининг долзарблиги ва зарурати. Жаҳонда инсон ва ускуна ўзаро алоқаларини ташкил этишда компьютерларнинг ўрни жуда муҳим ҳисобланади. Инсон ва ускунанинг ўзаро алоқасини ташкил этишнинг энг самарали воситаси бу нутқ ҳисобланади. Ҳозирги кунда инсон-машина интерфейсини амалга оширишни турли усул ва воситалари мавжуд бўлиб, энг самарали воситаларидан бири бу нутқ ҳисобланиб келмоқда. Нутқ буйруқлари орқали бошқарув турли соҳаларда, дастгоҳ ва ускуналарни овоз орқали бошқариш, кўнғироқ марказларида овозли ассистент, «ақлли уй» ва «ашёлар интернет»да (IoT), овоз орқали шахсини аниқлаш, имконияти чекланган инсонлар учун овозли бошқарув интерфейси ва бошқа соҳаларда кенг қўлланилиши мумкин. Бунда нутқ сигналларига ишлов бериш мураккаб таниб олиш алгоритмларини ўз ичига олади. Ушбу соҳада АКШ, Хитой, Япония, Жанубий Корея, Германия, Франция, Италия ва Россия Федерацияси каби дунёнинг етакчи мамлакатларида кенг қамровли илмий изланишлар олиб борилмоқда.

Жаҳонда бундай инсон-машина интерфейсларини амалга ошириш учун махсус аппарат-дастурий платформалар ишлаб чиқишни талаб этади. Лекин нутқга ишлов бериш алгоритмларининг мураккаблиги ва нутқ буйруқлари сонининг ошиши билан аппарат-дастурий платформанинг ҳисоблаш ресурси ва хотира хажмининг кескин ошишига олиб келади. Бундан ташқари аксарият инсон-машина интерфейслари реал вақт режимида ишлайди. Турли муҳитларда нутқ буйруқларини танишда нутқ сигналларига ташқи шовқин ва ҳалақитлар таъсир этади. Реал вақт режими эса ўз навбатида тезкорлик, ихчамлик ва фойдаланишга қулай бўлишликни талаб этади. Нутқ сигналларига ишлов бериш учун белгиланган ҳисоблаш ресурслари ва ишлов бериш тезлиги талаблари махсус алгоритмлар ва аппарат-дастурий ечимлар ёрдамида таъминланиши мумкин.

Республикамизда ахборот технологиялари ва сунъий интеллект соҳаларини ривожлантиришга қаратилаётган эътибор туфайли, нутқ сигналларига интеллектуал ишлов бериш, ўзбек тили нутқ сўзларини автоматик таниб олиш, овозли биометрия, овозли бошқарув тизимларини яратиш юзасидан илмий тадқиқотлар олиб борилмоқда. 2017-2021 йилларда Ўзбекистон Республикасини янада ривожлантиришнинг бешта устувор йўналиши бўйича Ҳаракатлар стратегиясида, жумладан, «...илғор ахборот-коммуникация технологияларини жорий этиш ва улардан фойдаланиш, иқтисодиёт, ижтимоий соҳа ва бошқарув тизимига ахборот-коммуникация технологияларини жорий этиш»¹ вазифалари белгилаб берилган. Мазкур масалаларни амалга ошириш, инсон нутқи орқали овозли биометрия ва нутқ

¹ Ўзбекистон Республикаси Президентининг 2017-йил 7-февралдаги ПФ-4947-сон «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида»ги Фармони

буйруқлари орқали турли техник агрегатларни бошқаришнинг математик алгоритмлари ва дастурий мажмуасини ишлаб чиқишда муҳим аҳамиятга эга.

Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги ПФ-4947-сон «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида»ги, 2018 йил 19 февралдаги ПФ-5349-сон «Ахборот технологиялари ва коммуникациялари соҳасини янада такомиллаштириш чора-тадбирлари тўғрисида»ги Фармони, 2021 йил 17 февралдаги ПҚ-4996-сон «Сунъий интеллект технологияларини жадал жорий этиш учун шарт-шароитлар яратиш чора-тадбирлари тўғрисида»ги, 2021 йил 26 августдаги ПҚ-5234-сон «Сунъий интеллект технологияларини қўллаш бўйича махсус режимни жорий қилиш чора-тадбирлари тўғрисида»ги қарорлари ҳамда мазкур фаолиятга тегишли меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишда ушбу диссертация тадқиқоти муайян даражада хизмат қилади.

Тадқиқотнинг республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги. Мазкур тадқиқот Республика фан ва технологиялари ривожланишининг IV. «Ахборотлаштириш ва ахборот-коммуникация технологияларини ривожлантириш» устувор йўналиши доирасида бажарилган.

Муаммонинг ўрганилганлик даражаси. Нутқ сигналларига рақамли ишлов бериш ва уларни автоматик таниб олиш усулларини ишлаб чиқиш ва такомиллаштириш масалалари ҳамда уларнинг амалий қўлланилиши бўйича бир қатор олимлар тадқиқотлар олиб борганлар, LeCun Y, Bengio Y, Abdel-Namid O, Grazina Korvel, Patel I, И.С. Кипяткова, Н.М. Марковников, А.А. Карпов, М.Н. Гусев, А.Л. Ронжин, О.Мамырбаев, С.Рустамов ва бошқа хорижий олимлар шулар жумласидандир.

Ўзбекистонда сигналларга ишлов бериш ва нутқни таниб олишнинг назарий асосларини ривожлантиришга М.М.Камилов, М.М.Арипов, Ш.Х.Фозилов, Н.А.Игнатъев, М.М.Мусаев, Х.Н.Зайнидинов, Ў.Р.Хамдамов, Н.С.Маматов, Э.Ш.Назирова ва бошқалар ўзларининг катта ҳиссаларини қўшганлар.

Нутқ сигналларига ишлов бериш мураккаб алгоритмик босқичлардан ташкил топганлиги сабабли, юқори ҳисоблаш ресурсларини талаб этади. Турли муҳитларда нутқ буйруқларини таниб олишни амалга ошириш, ишончли, ихчам ва юқори самарадорликга эга бўлган аппарат-дастурий таъминотини лойиҳалаштириш заруратини беради. Ушбу йўналиш бўйича тадқиқотлар шуни кўрсатдики, турли муҳит ва шароитларда нутқ буйруқларини таниб олиш аппарат-дастурий воситалари ва алгоритмлари етарлича тадқиқ этилмаган. Юқорида айтиб ўтилганлар ўзбек тили нутқини таниб олишга ҳам таалуклидир.

Диссертация тадқиқотининг бажарилган олий таълим муассасасининг илмий-тадқиқот ишлари режалари билан боғлиқлиги. Диссертация тадқиқоти Муҳаммад ал-Хоразмий номидаги Тошкент ахборот технологиялари университетининг илмий-тадқиқот ишлари режасининг

БЕОА-5-005 – «Инсон нутқини таниш тизимлари қурилмаларининг таҳлили ва нутқий сигналларни қайта ишловчи аппарат-дастурий воситалар яратиш» (2016-2018), ЁБВ-Атех-2018-134 – «Нутқий сигналларни таниш алгоритмларининг рақамли сигнал процессорларида таҳлили ва дастурий таъминотини яратиш» (2017-2019), Бирлашган миллатлар ташкилоти тараққиёт дастури ва Ўзбекистон Республикаси маҳалла ва оилани қўллаб-қувватлаш вазирилик қўмағида №32/20 «Разработка мобильного приложения автоматического распознавания Узбекской речи для слабослышащих людей» (2020-2021) мавзуларидаги лойиҳалари доирасида бажарилган.

Тадқиқотнинг мақсади ўзбек тили нутқ сигналларини танишнинг тезкор алгоритмлари, нутқ буйруқлари орқали овозли бошқарув интерфейсини ишлаб чиқиш ва аппарат-дастурий платформада амалга оширишдан иборат.

Тадқиқотнинг вазифалари:

нутқ буйруқлари орқали овозли бошқарув амалий дастурий интерфейсларини таҳлил қилиш ва ўрганиш;

кирувчи нутқ сигналларини иш фаолияти шароитларини ҳисобга олган ҳолда турли ҳалақит ва шовқинлардан тозалаш алгоритминини ишлаб чиқиш;

нутқ сигналларини реал вақт режимида сиқиш алгоритминини ишлаб чиқиш;

нутқ орқали овозли бошқарув тизимлари учун чекланган нутқ буйруқларини таниш алгоритминини ишлаб чиқиш;

реал вақт режимида нутқ буйруқларини таниб олиш аппарат-дастурий платформаларини ишлаб чиқиш.

Тадқиқотнинг объекти сифатида турли муҳитларда ёзиб олинган нутқ сигналлари ва улардан ҳосил қилинган спектограмма ва белгили хусусиятлар олинди.

Тадқиқотнинг предмети сифатида нутқ буйруқлари орқали усқуналарни овозли бошқарув аппарат-дастурий платформалари, нутқни таниб олиш усуллари ва алгоритмлари ташкил этади.

Тадқиқотнинг усуллари тадқиқот жараёнида сигналларга рақамли ишлов бериш, сонли усуллар, спектрал усуллар, мэл-кэпстрал усуллар ва қарор қабул қилишнинг усул ва алгоритмлари ташкил қилади.

Тадқиқотнинг илмий янгилиги қуйдагилардан иборат:

турли муҳит ва иш фаолияти шароитини ҳисобга олган ҳолда нутқ буйруқларини таниб олиш ва нутқ сигналларини шовқинлардан тозалашнинг адаптив филтрлаш алгоритми ишлаб чиқилган;

нутқ сигналларидан ахборот белгили хусусиятларини Сингуляр ёйиш орқали соддалаштириш асосида ишлов берувчи алгоритми яратилган;

векторли квантлаш ва яширин Марков модели асосида нутқ буйруқларини таниш алгоритмлари ишлаб чиқилган;

реал вақт режимида нутқ буйруқларини таниб олиш алгоритмлари асосида аппарат-дастурий платформанинг архитектураси ишлаб чиқилган ва амалга оширилган.

Тадқиқотнинг амалий натижалари қуйидагилардан иборат:

турли муҳитда нутқ буйруқларига ишлов бериш, суҳандонни идентификациялаш, нутқ буйруқ сўзларини таниб олиш ва матнга ўгириш алгоритмлари ва дастури ишлаб чиқилган;

нутқ буйруқлари орқали овозли бошқарув аппарат-дастурий платформаси функционал тузилмаси лойиҳаланган;

қулай фойдаланувчи интерфейси ва реал вақт режимида ишлайдиган аппарат-дастурий платформа ва дастурий мажмуаси яратилган;

ўзбек тили нутқ буйруқ сўзларини таниш ва улар орқали овозли бошқарув дастурлари саноат бошқарув тизимларида ва эшитиш нуқсонига эга болалар махсус мактаб-интернатларида тадбиқ қилинган.

Тадқиқот натижаларининг ишончлилиги Тадқиқот натижаларининг ишончлилиги диссертациядаги таклиф қилинган алгоритм ва усулларнинг тестлаш натижалари, алгоритм ва усулларнинг апробацияси босма нашрлар ва халқаро илмий конференциялар маърузаларида, тажрибавий тадқиқотлар ва яратилган аппарат-дастурий платформани амалиётга тадбиқ этишлиши билан тасдиқланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Мураккаб нутқ буйруқ сўзларини таниб олиш учун таклиф этилаётган нутқ сигналларини таҳлил қилиш ва таниб олишнинг интеллектуал алгоритмларини аппарат-дастурий платформаларида амалга оширилиши билан изоҳланади.

Тадқиқот натижаларининг амалий аҳамияти шундан иборатки, яратилган алгоритмлар турли муҳитларда ва ишлаб чиқариш корхоналарида нутқ орқали овозли бошқарув дастурий мажмуаси, суҳандонни идентификациялаш ва эшитиш нуқсонига эга инсонлар учун Ўзбек тили нутқ сўзларини таниб олиш дастурий мажмуаларини ишлаб чиқиш билан изоҳланади.

Тадқиқот натижаларининг жорий қилиниши. Диссертация доирасида олиб борилган тадқиқотлар натижасида ишлаб чиқилган адаптив фильтлаш, нутқ буйруқ сўзларини таниб олиш алгоритмлари ва яратилган аппарат-дастурий платформа асосида:

турли муҳитларда нутқ сигналларини адаптив фильтлаш, сўзловчи суҳандонни идентификациялаш ва нутқ буйруқлари орқали овозли бошқарув аппарат-дастурий платформаси «Ўзбекистон темир йуллари» АЖнинг ишлаб чиқариш корхоналарида мураккаб дастгоҳларни бошқаришда жорий қилинган (Ахборот технологиялари ва коммуникацияларни ривожлантириш вазирлигининг 2021 йил 9 ноябрдаги 33-8/7908-сон маълумотномаси). Натижада нутқ буйруқлари орқали мураккаб дастгоҳларни бошқаришда меҳнат унумдорлигини ошириш имконини берган;

эшитиш нуқсонига эга инсонлар учун ўзбек тили нутқ сўзларини матн, сурдо матн ва имо-ишорага ўгиришда қулай интерфейсга эга бўлган дастурий мажмуа 106-сон заиф эшитувчи болалар учун ихтисослаштирилган мактаб-интернатда жорий қилинган (Ахборот технологиялари ва коммуникацияларни ривожлантириш вазирлигининг 2021 йил 9 ноябрдаги 33-8/7908-сон маълумотномаси). Натижада мактаб тарбияланувчиларини

ўқитишда, янги Ўзбек тили сўзларини имо-ишора орқали ўрганиш вақти 15-20%га қисқаришига эришилган.

Тадқиқот натижаларининг апробацияси. Мазкур тадқиқот натижалари 15 та илмий-амалий анжуманларда, жумладан 5 та халқаро ва 10 та Республикада, муҳокамадан ўтказилган.

Тадқиқот натижаларининг эълон қилинганлиги. Диссертация мавзуси бўйича 23 та илмий иш жумладан, Ўзбекистон Республикаси Олий аттестацияси комиссиясининг диссертациялар асосий илмий натижаларини чоп этиш тавсия этилган илмий нашрларда 5 та мақола, жумладан 2 та чет эл ва 3 та Республика журналларида нашр этилган, шунингдек 4 та ЭҲМ учун дастурий маҳсулотларга гувоҳномалар олинган.

Диссертациянинг тузилиши ва ҳажми. Диссертация кириш қисми, тўртта боб, хулоса, фойдаланилган адабиётлар рўйхати, шартли белгилар ва атамалар рўйхати ҳамда иловалардан иборат. Диссертациянинг ҳажми 119 бетни ташкил этади.

ДИССЕРТАЦИЯНИНГ АСОСИЙ МАЗМУНИ

Кириш қисмида диссертация мавзусининг долзарблиги ва зарурийлик даражаси асослаб берилган, мақсад ва вазифалар шакиллантирилган, тадқиқот объекти ва предмети аниқланган, тадқиқотнинг Ўзбекистон Республикаси фан ва технологиялари тараққиётининг устувор йўналишига мослиги белгиланган, тадқиқотнинг илмий янгилиги ва амалий натижалари кўрсатиб ўтилган, олинган натижаларнинг ҳаққонийлиги асослаб берилган, олинган натижалар назарий ва амалий аҳамияти очиб берилган, тадқиқот натижаларининг амалиётга тадбиқ этилиши рўйхати, ишни синаш натижалари, нашр этилган ишлар ва диссертация тузилмаси тўғрисидаги маълумотлар келтирилган.

Диссертациянинг «**Нутқ сигналларига ишлов бериш алгоритмлари ва аппарат-дастурий воситалари таҳлили**» деб номланган биринчи бобида нутқ сигналларига рақамли ишлов беришнинг асосий амаллари, ахборот белгилари хусусиятларини ажратиш, замонавий аппарат-дастурий воситалар ва платформалар таҳлили, нутқ орқали овозли бошқарув дастурий интерфейслари кўриб ўтилган.

Нутқ сигнали – бу мураккаб динамик жараённинг модели бўлиб, уни таҳлил қилишда сигнал ва унинг бўлак қисмини тавсифловчи бир нечта кўрсаткич(хусусиятлар)ларига таянилади. Нутқ сигналларининг ушбу асосий хусусиятлари қуйидагилар: формант частотаси, асосий тон частотаси, спектрал ташкил этувчилардир.

Таҳлил қилиш учун нутқ сигналининг энг содда кўриниши бу ажратиб олинган сўзлардир. Одатда нутқ сўзлари ва уларнинг луғатлари чекланган бўлади. Замонавий мураккаб нутқни таниш тизимларида асосан ўзак сўзларни ажратиб таниб олиш амалга оширилади. Ҳозирги кунда нутқ сўзларини таниш қуйидаги соҳаларда кенг қўлланилмоқда: саноат дастгоҳларини ва транспорт воситаларини бошқаришда, «ақлли уй» тизимларида, «ашёлар интернет»

технологияларини амалга оширишда, имконияти чекланган инсонлар учун интерфейс, хизмат кўрсатиш соҳалари шулар жумласидандир.

Нутқ орқали бошқарув тизимини турли соҳада қўлланганда, амалга оширилаётган муҳитдан қўшиладиган шовқин ва ҳалақитлар частоталари турлича бўлади. Нутқ буйруқларини таниб олишда ушбу шовқин ва ўзбек тили товушларининг хусусиятларини инобатга олишни талаб этади

Кундан кунга реал вақтда нутқни таниб олиш ва таҳлил қилиш тизимларга талаб ортиб бормоқда. Бундай реал вақт тизимлари орқали катта маълумотлар базаси талаблари, нутқ сигналларининг шовқинга таъсирчанлиги ва юқори ҳисоблаш ресурслари талаблари бўлганлиги учун ягона аппарат-дастурий воситада амалга ошириш қийин

Нутқ сигналлари мураккаб сигнал бўлганлиги туфайли ушбу сигналларга ишлов бериш, хотирада сақлаш, алоқа каналлари орқали узатиш муаммолари мавжуд. Ушбу муаммоларни ечиш учун олимлар томонидан нутқ сигналларига ишлов беришда сигнални сиқиш, сигналдан хусусиятларни ажаратиб олиш, сигнал спектрал қийматлари билан ишлаш каби ёндашувлар таклиф этилган. Буларнинг барчаси сигналга ишлов бериш жараёнини такомиллаштиришга қаратилган.

Мураккаб нутқни таниш алгоритмларини реал вақт режимида амалга ошириш учун махсус аппарат-дастурий воситалар талаб қилинади. Нутқ буйруқларини таниш алгоритмларининг мураккаблик даражасига қараб, бир нечта ўрнатилган процессорлардан ягона аппарат-дастурий платформа сифатида фойдаланишни тақозо этади.

Аппарат-дастурий платформа(АДП) – бу белгилаган вазифани амалга оширишга мўлжалланган техник қурилма таъминоти ва дастурий таъминотлар мажмуидир. АДП модулли принципда қурилган бўлиб, махсус процессор жойлашган асосий плата блоки, хотира блоки, киритиш-чиқариш блоки, ташқи интерфейслар блокларидан ташкил топган.

Диссертациянинг **«Нутқ сигналларига параметрик ишлов бериш алгоритмларини амалга ошириш»** деб номланган иккинчи бобида нутқ сигналларига ишлов бериш босқичлари, турли муҳитларда нутқ сигналларини адаптив филтрлаш алгоритминини ишлаб чиқиш, нутқ сигналлари спектрал таҳлили алгоритмларини қўллашнинг аҳамияти ёритилган. Ҳисоблаш тизимига тушадиган юкламаларни камайтириш учун Сингуляр ёйиш орқали сигнал ўлчамларини қисқартириш ва процессорнинг ишлаш самарадорлигини ошириш, суҳандонга боғлиқ тизимлар учун суҳандонни идентификациялаш алгоритми, яширин Марков модели ёрдамида нутқ сўзларини таниб олишни амалга оширишга бағишланган.

Нутқни таниб олиш муаммоси - бу сигналларни қайта ишлаш соҳасида доимо муҳим мавзу бўлиб келган тимсолларни таниб олишдир. Тимсолларни таниб олишдан мақсад объектни сонли синфларга ажратишдир. Ҳар бир синф ёки намуна маълум бир сўзни ифодалайди. Барча нутқни танишда қўлланиладиган амаллар ва алгоритмлар маълум бир мураккаб сонли усуллар ва шунга мос ишлов бериш алгоритмларини талаб этади. Аксарият ҳолларда

нутқ сўзларини таниб олишда сигнални Хеннинг ойналаридан ўтказиш, тезкор Фурье ўзгартиришини қўллаш ва мел-кэпстрал қийматларини олишдан иборат.

Таклиф этилаётган реал вақт режимида суҳандонга боғлиқ ёки боғлиқ бўлмаган ҳолда нутқ буйруқларини таниб олиш алгоритмлари кетма-кетилги 1-расмда келтирилган.



1-расм. Таклиф этилаётган нутқ сигналларидан белгили хусусиятларни ажратиб олиш ва таниш алгоритмлари босқичлари

Нутқ буйруқларини таниб олиш қуйидаги алгоритмик босқичларда амалга оширилади.

1. *Кирувчи нутқ кўринишидаги аналог сигнал, 16 кГц частотали рақамли сигнал кўринишига ўгирилади.*

2. *Нутқ буйруқлари ташқи шовқин ва ҳалақитлардан тозалаш.* Реал вақт шароитида ва турли муҳитларда шовқин ва ҳалақитларнинг ҳам частота ва хусусиятлари турлича бўлади. Шуларни ҳисобга олиб нутқ буйруқларини танишда адаптив филтрлардан фойдаланиш мақсадга мувофиқ. Чунки ҳар бир муҳитнинг ўзига хос шовқин ва ҳалақитлари мавжуд булиб уларнинг хусусиятлари ҳам турлича булади.

Таклиф этилаётган алгоритм орқали филтрлашни амалга оширишда, аввал мавжуд бошқарув муҳитида микрофондан келадиган шовқинли

сигнални 5 секундлик вақт оралиғида ёзиб олинади ва ушбу сигналга мос равишда адаптив филтёр коэффициентлари NLMS (normalized least mean squares) алгоритми асосида шакллантирилади. Ушбу сошлаш ва мослашув жараёнидан сўнг микрофон орқали келаётган нутқ буйруқлари ушбу филтёрлаш коэффициентлари орқали филтёрлаш босқичидан ўтади.

3. *Нутқ сигналларидан жимлик соҳалари ўчириш.* Жимлик соҳаларини ўчириб ташлашдан аввал нутқ сигнали қийматлари $M=256$ га тенг бўлак сегментларга ажаратилиб, Q_i сегментга ўзлаштирилади. Ушбу сегментлардаги жимлик ҳолатини аниқлаш учун ҳар бир сегментнинг энергияси ҳисобланади:

$$E_n = \frac{1}{N} \sum_{i=0}^N Q_i^2$$

Бу ерда E - нутқнинг энергияси, Q_i - сигнал қийматлари.

4. *Фреймларга ажратиш.* Нутқ сигналдан жимлик ҳолатлари йўқотилгандан сўнг нутқ сигналини қайта сегментлаш натижасида фреймларга ажратиш зарур. Олиб борилган тадқиқот натижаларига кўра, $M=256$ бўлган фреймларга ишлов бериш максимал тезкорликни таъминлайди.

5. *Хеннинг ойнасидан ўтказиш.* Сигналларга рақамли ишлов беришда қуйидаги ойнали ўзгартиришлар қўлланилади: Бартлетт, Блэкман, Чебышев, Хеннинг, Кайзер ва Хэмминглар шулар жумласидандир. Шулардан нутқ сигналларига ишлов беришда Хемминг ва Хеннинг ойналари кенг қўлланилади. Бизнинг Хеннинг ойнасини танлашдан мақсад биз фойдаланадиган қисқа вақтли Фурье ўзгартиришида (STFT-short time Fourier transform) ойнадан ўтказишда сигналнинг боши ва охири тўлиқ ноль қийматига тенг бўлиши керак:

$$\omega(n) = 0.5 \left(1 - \cos\left(\frac{2\pi n}{N-1}\right) \right)$$

Бунда, N – ойналар кенглиги.

6. *Қисқа вақтли Фурье алмаштириш (STFT).* Нутқ сигналларидан белгили хусусиятларни тўғри ажратиш олиш, нутқ буйруқларини ва суҳандонни таниш ишончлигига таъсир қилади.

Инсон кулоқлари эшитиш тизими (улитка) механик вақт частотаси анализатори бўлгани учун, у ҳам доимо нутқ сигналнинг бироз вақт ўзгарган спектрларининг қисқа кетма-кетлигини қабул қилади, бу тахминан нутқ сигналларига STFTни қўллаганда ҳам содир булади:

$$STFT_x = STFT\{x[n]\} = X[k, n] = \sum_{m=-\infty}^{\infty} x[m]w[m-n]e^{-jkm}$$

бу ерда, j мавҳум қисми ва тўплам $\left\{e^{-\frac{2\pi j}{N}n}\right\}$ эса ортонормал базисдир.

Ҳосил қилинган спектограмма қийматлари турли хил интеллектуал тизимларда кирувчи параметр сифатида қабул қилинади.

7. *Ўлчамларни қисқартириш.* Сигнал ўлчамларини қисқартириш орқали ҳисоблаш юкмасини камайтириш ва процессор самардорлигини ошириш

учун сингуляр ёйиш усулидан фойдаланилди. Сингуляр ёйиш SVD (singular value decomposition) – бу хақиқий матрицани каноник шаклга келтириш мақсадида декомпозициядир. SVD матрицалар билан ишлашда жуда ҳам қулай усулдир, бундан ташқари матрицалар устида амаллар процессор самарадорлигини ҳам оширади.

SVDнинг моҳияти жуда ҳам оддий бўлиб, ҳар қандай матрица(хақиқий ёки мураккаб) учта матрицанинг ҳосиласи сифатида намоён этилади ва у қуйидаги формула орқали ифодаланади:

$$X = U_r D V_r^T = \sum_{i=1}^r \sigma_i u_i v_i^T$$

бу ерда,

$$U_r := [u_1, \dots, u_r] \in \mathbb{R}^{m \times r}, \quad V_r := [v_1, \dots, v_r] \in \mathbb{R}^{l \times r}$$

Сигнал ўлчамлари сонини камайтириш муаммосини ечишда айнан шу D матрицадан фойдаланилади. Шунинг учун сигнал ўлчамларини қисқартиришда, SVD алгоритми ҳисобга олинаётган компонентлар берадиган дисперсиялари йиғиндисига таянилади. Тажриба натижаларига кўра дисперсия компонентлари сони қанча юқори булса ўлчов қийматлари қисқартирилган сигнал шунча хақиқийсига яқинлашади.

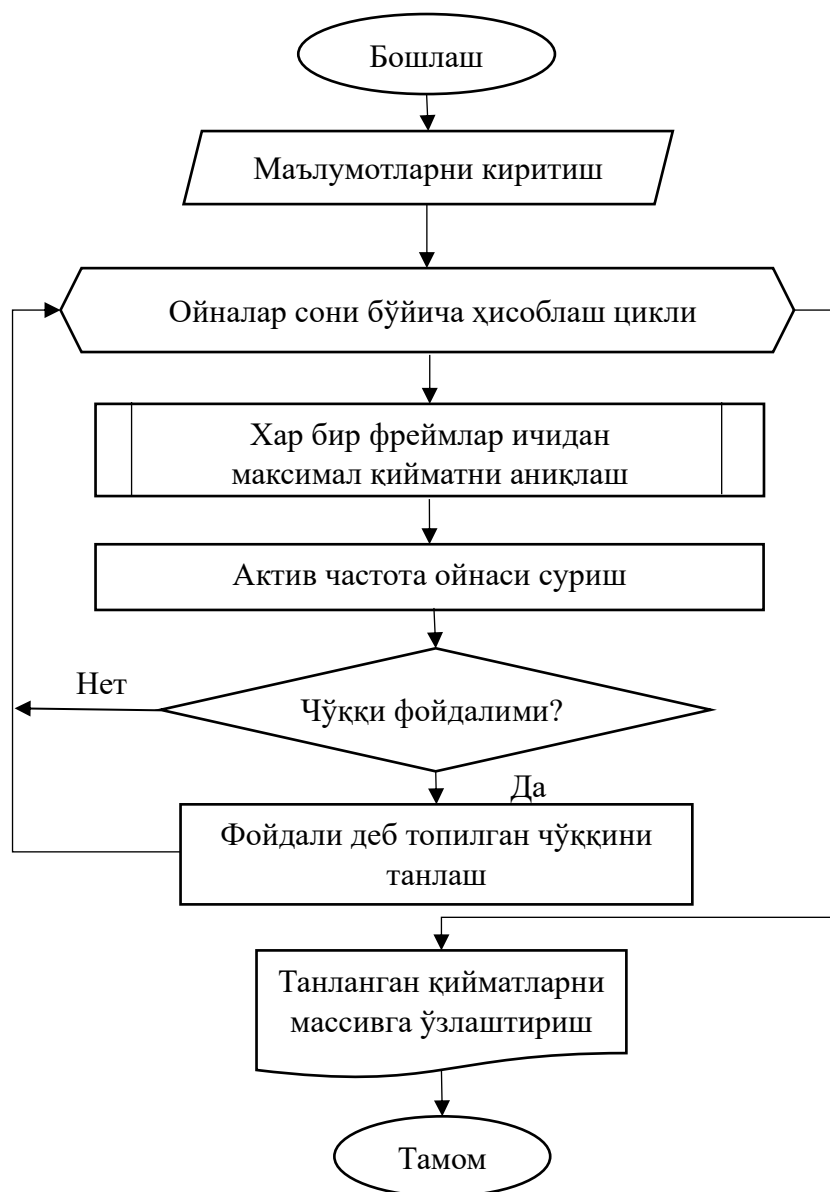
8. *Формант частоталарини аниқлаш.* Сухандонга боғлиқ бўлган нутқ буйруқларини танишда формант частоталарини қўшимча параметр сифатида олиш яхши аҳамият касб этади. Бу эса ўз навбатида таниб олиш аниқлик фойизини ошишига олиб келади.

Нутқ товушларини талаффуз қилишда оҳанг пульси ёки шовқин сигналлари овоз йўли орқали ўтади. Натижада, бир хил тонал ёки шовқин спектри бир қатор максимал ва минимал спектрга айланади. Спектрнинг максимумлари *форманта* деб аталади Нутқни талаффуз қилишда унинг спектри доимо ўзгариб туради ва формант ўтишлар ҳосил бўлади. Таклиф этилаётган асосий формант частоталарини аниқлаш алгоритми блок-схемаси қуйидагича 2-расмда келтирилган.

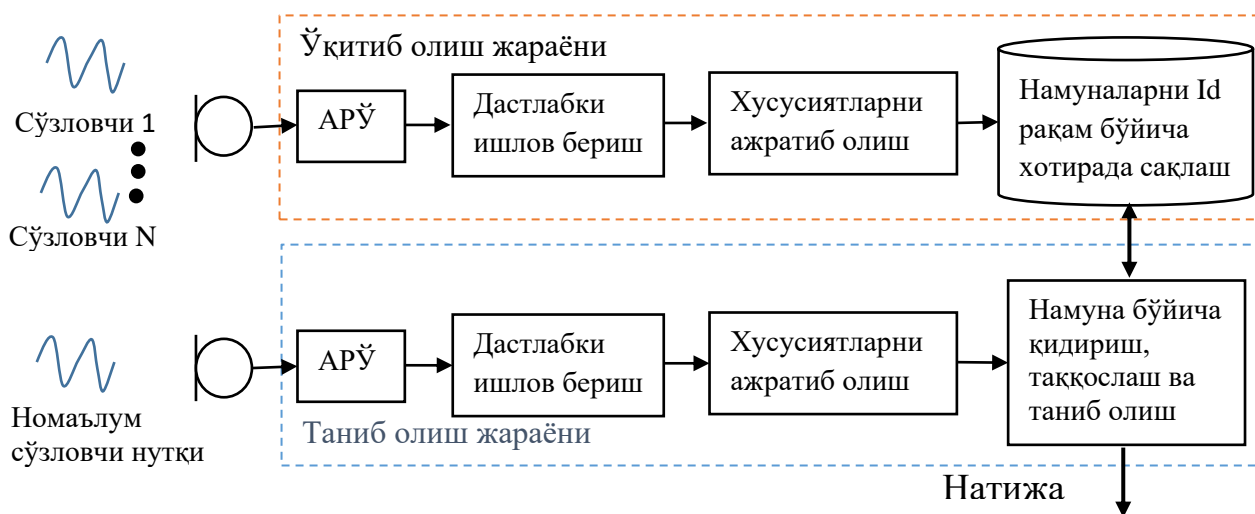
Аксарият овозли бошқарув ва маълум бир хавфсизлик тизимларида инсонни нутқи орқали шахсини идентификациялашда айнан аниқ буйруқ сўзи ёки айнан аниқ бир калит сўзлар ишлатилади.

9. *MFCC(Mel frequency cepstral coefficients) қийматларини олиш.* MFCC энг кенг тарқалган усуллардан бўлиб, инсон эшитиш тизимининг чизиқли бўлмаган таъсирини аппроксимациялайди. Бу эса ўз навбатида ушбу коэффициентлар инсон эшитиш тизимига жуда ҳам яқинлигини билдиради.

10. *Белгиларнинг мувофиқлигини текшириш ва таниб олиш.* Ягона белгилар векторида ажратиб олинган нутқ буйруқлари дастлаб мувофиқлиги текширилади ва мослиги аниқланса у таниб олинади.



2-расм. Асосий формант частоталарини аниқлашнинг тақомиллаштирилган алгоритми



3-расм. Сўзловчини шахсини идентификациялаш тизими

Ушбу босқичида нутқ буйруқларини турли синфлаш ва кластерлаш масалаларини ечиш орқали амалга оширилади. Бизнинг ҳолатда Яширин Марков моделига асосланган моделдан фойдаланилди.

Нутқ орқали сўзловчини шахсини аниқлаш тизимининг структурасини амалга ошириш босқичлари қуйида келтириб ўтилган(3-расм).

Сўзловчи ёки суҳандонни гапирётган нутқи орқали шахсини идентификация қилиш икки режимда, яъни ўқитиб олиш ва таниб олиш режимларида амалга оширилади.

Ўқитиб олиш жараёнида қуйида алгоритмик босқичлар амалга оширилади:

1-қадам. Микрофондан кирувчи аналог сигнал рақамли кўринишга ўгирилади.

2-қадам. Дастлабки ишлов бериш. Бунда кирувчи сигнал филтрланади, сегментланади, фреймларга ажратилиб ойналардан ўтказилади.

3-қадам. Ахборот белгили хусусиятлари ажратилаб олинади.

4-қадам. Сўзловчига идентификацион рақам (id) бириктирилади ва ушбу сўзловчининг овоз хусусиятлари хотирага сақлаб қуйилади.

Таниб олиш жараёнида қуйидаги алгоритмик босқичлар амалга оширилади:

1-қадам. Микрофондан ихиёрий гапирган сўзловчи ёки суҳандоннинг нутқи кирувчи аналог сигнал кўринишидан рақамли кўринишга ўгирилади.

2-қадам. Дастлабки ишлов бериш амаллари амалга оширилади.

3-қадам. Ахборот белгили хусусиятлари ажратилаб олинади.

4-қадам. Олинган хусусиятлар хотирасида сақланган сўзловчи суҳандонларнинг хусусиятлари билан векторли квантлаш (ВК) интеллектуал алгоритмлардан фойдаланган ҳолда таққосланади.

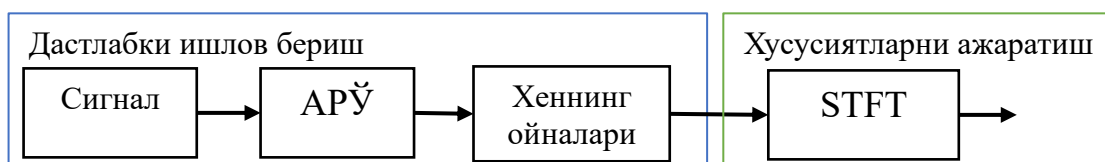
Тадқиқот натижалари шуни курсатдики суҳандонга боғлиқ бўлган нутқ буйруқ сўзларини реал вақт режимида таниб олишни аппарат-дастурий платформада амалга оширишда ВК алгоритми нисбатан ихчам ва амалга ошириш қулайлиги билан бошқа алгоритмларга нисбатан(яширин Марков модели, сунъий нейрон тармоқлари, DTW-dynamic time warping) яхши натижа берди.

5-қадам. Агар сўзловчининг нутқидан олинган хусусиятлар хотирадаги сақланган хусусиятлар билан тўғри келса натижада ушбу сўзловчининг шахси аниқлангани ва шунга мос хабар беради

Диссертациянинг «**Нутқ сигналига ишлов бериш алгоритмлари аппарат-дастурий платформасини яратиш**» деб номланган учинчи бобида нутқ буйруқлари орқали бошқарув тизимини амалга ошириш учун аппарат таъминотида қўйилган талаблар келтирилган. Турли дастурлаш муҳитларида ва аппарат-дастурий таъминотларида нутқ сигналларини таниб олишни амалга оширишни осонлаштириш учун аппарат таъминоти архитектурасига мос реал вақт операцион тизими амалга оширилган. Илмий ишда нутқ буйруқларга ишлов беришнинг аппарат-дастурий платформаси архитектураси таклиф этилган.

Аппарат-дастурий платформаларнинг турфа хиллиги, архитектура хусусиятларини ва уларнинг ҳисоблаш ресурслари чекланганлигини ҳисобга олган ҳолда, ўтказилган тажрибаларга асосланиб нутқ сигналларига ишлов беришнинг тўлиқ босқичларини алоҳида АДПда амалга ошириб бўлмайди. Шунинг учун турли АДПда нутқ сигналларига ишлов беришнинг қуйидагича вариантлари таклиф этилади.

1-вариант. Микроконтроллерга(МК) асосланган нутқни таниш тизимларида(4-расм.).



4-расм. МК тизимларида сигналларга ишлов бериш босқичлари

Таҳлиллар шуни кўрсатдики, *Микроконтроллерлар* нутқ сигналларига дастлабки ишлов бериш жараёнида жуда ҳам яхши натижалар кўрсатди. Сигналларни рақамлаштириш, филтрлаш, силликлаш, жимлик соҳаларини ажратиш алгоритмларини қийинчиликсиз амалга ошира олади.

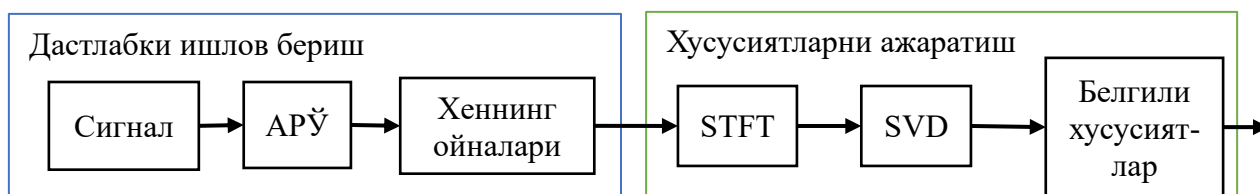
2-вариант. Махсус сигнал процессорлари ва қайта дастурланувчи мантиқий интеграл схемали(DSP, FPGA) нутқни таниш тизимларида(5-расм.).



5-расм. DSP, FPGA тизимларида сигналларга ишлов бериш босқичлари

Қайта дастурланувчи DSP ва FPGA қурилмаларининг арихтектуралари даражасида сигналларга рақамли ишлов бериш алгоритмларини амалга ошириш кўзда тутилган. Шунинг учун бу АДПда хусусиятларни ажаритиш, фреймлаш, турли ўзгартириш, сиқиш, ахборот белгиларни ажаратиш каби мураккаб амалларни осон амалга ошириш мумкин. Аммо, аналог кириши бўлмаганлиги учун албатта МКлар билан биргаликда ишлатилади.

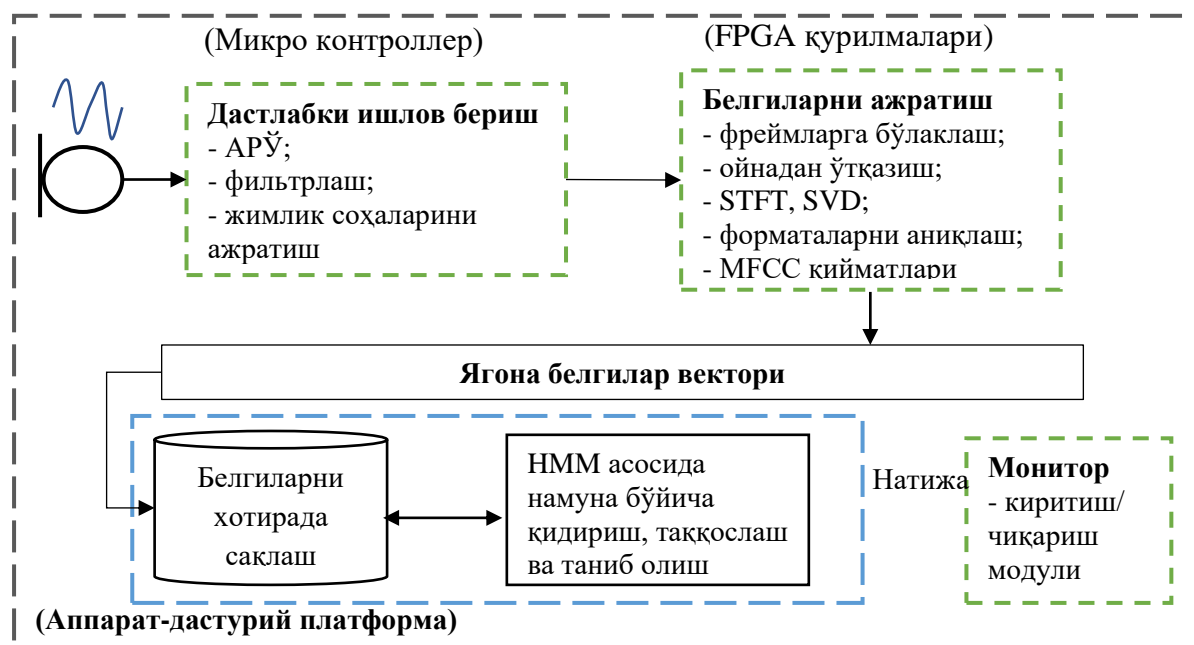
3-вариант. Махсус мобиль ўрнатилган тизим процессорларига асосланган нутқни таниш тизимларида(6-расм.).



6-расм. Мобиль процессорларга асосланган тизимларда сигналларга ишлов бериш босқичлари

Мобиль процессорларга асосланган аппарат-дастурий платформалар кўп ядроли бўлиб, оқимларни бошқариш ва кўп вазифалик амалларини бажара олади. Бундан ташақари ушбу тизимларда ихтиёрий юқори дастурлаш тилларининг компиляторларини ўрнатиш орқали мураккаб нутқ сигналларига

ишлов бериш алгоритмларини нисбатан осонроқ амалга ошириш имконияти пайдо бўлади. Кўп ядроли процессор мавжудлиги ва ўрнатилган реал вақт операцион тизимлари амалга ошириш имконияти мавжудлиги ва хотира хажми нисбатан катталигини ҳисобга олган ҳолда мураккаб нутқ сигналларига ишлов беришнинг тўлиқ циклини амалга ошириш имконияти мавжуд.

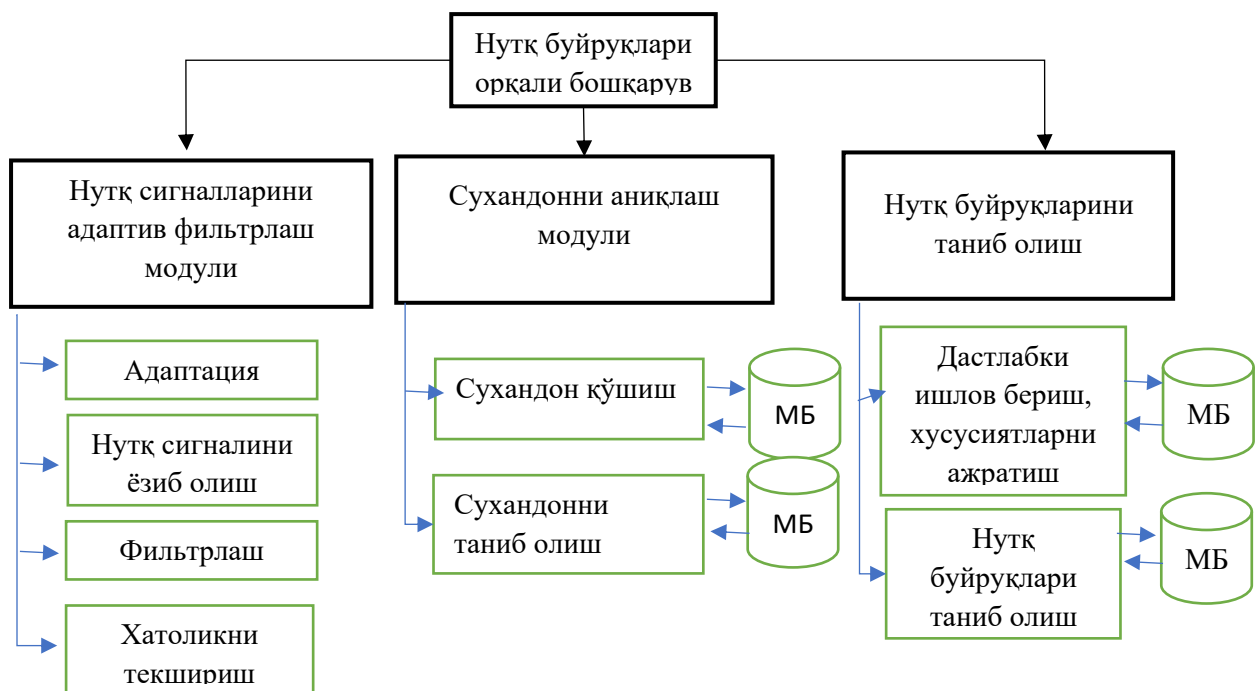


7-расм. Нутқ буйруқларини таниб олиш алгоритмларини АДПда амалга оширилиши

Натижада турли мураккаб нутқ сигналларига ишлов бериш алгоритмлари ва дастурларини ягона операцион тизим осида EM3288 аппарат-дастурий платформасида амалга оширилди. Ўтқазилган тажрибалар натижасига таяниб, нутқ буйруқларини таниб олишнинг мураккаб алгоритмлари таклиф этилаётган АДПда қуйидагича бирлаштирилди(7-расм).

Диссертациянинг «Ўзбек тили нутқ буйруқларини таниш платформаси интерфейси ва дастурий мажмуаси» деб номланган тўртинчи бобида чекланган нутқ буйруқларини таниб олиш ва овозли бошқариш мажмуаси модулларининг таркибий қисмлари, ҳамда нутқ буйруқлари орқали овозли бошқарув тизимларига мўлжалланган интерфейс алгоритми ва дастурлари келтирилган. Нутқ сигналлари орқали овозли бошқарув дастурий мажмуаси нутқ буйруқларига дастлабки ишлов бериш модули, адаптив филтрлаш модули, суҳандонни аниқлаш модули, нутқ буйруқларини таниб олиш модулларидан ташкил топган(8-расм).

Дастурий модуллар кросс-платформали Python дастурлаш тилида амалга оширилган бўлиб, Linux ва Window оиласига мансуб операцион тизимларда амалга ошириш имкониятига эга.



8-расм. АДПда амалга оширилган дастурий мажмуанинг тузулиши

Нутқ сигналларига ишлов бериш ва таниб олиш дастурий мажмуаси АДПда амалга ошириш учун Python дастурлаш тилида ёзилди. Бу ушбу дастурий воситани кросс-платформалилик хусусиятини таъминлайди. Нутқни таниш тизимини ишга тушуришдан аввал дастурий мажмуа соланади. Бунда дастлаб 5 секунд вақт оралиғида нутқ буйруқларини амалга оширилиши талаб этилган муҳитдан шовқин ёзиб олинади ва филтрлаш коэффициентлари танланади. АДП микрофони орқали келаётган нутқ сигнали ёзиб олиш бошланади. Ёзиб олинган нутқ сигналлари яширин Марков модели асосида таниб олиш ишлари амалга оширилади. Таниб олинган сўзлар экранга текст кўринишида чиқарилади.

Тадқиқот иши вазифасига кўра чекли лўғат сўзларини таниш ва овозли бошқарув тизими амалга оширилди. Бунда куйидаги 30 сўз буйруқлари танлаб олинди: “*talaba, malika, samalyot, lola, bola, ona, masala, mamlakat, olti, oltin, lochin, lobar, kitob, maktab, archa, chelak, gul, yulduz, sayyora, viloyat, tegirmon, dala, vagon, bekobod, angren, yangiyo`l, yunusobod, umid*”.

Ушбу тажриба натижасига асосан қисқа вақтли Фурье ўзгартириши ўрнатилган тизим процессорларида яхши тезкорлик берди(1-жадвал). Шунга кўра дастлабки ишлов беришдаги классик FFT+MFCC алгоритми ва таклиф этилаётган алгоритм тезкорлиги тажриба орқали тестланди.

1-жадвал.

Дастлабки ишлов беришдаги тезкорлик

Ишлов бериш усули	Тезкорлик (мсек.)
FFT+MFCC	60
STFT+SVD	15

Таниб олишда синфлаш алгоритмлари таҳлили(2-жадвал). Нутқ сигналларида ажратиб олинган ахборот белгили хусусиятлари орқали таниб олишда, сунъий нейрон тармоқлари, яширин Марков модели (ЯММ) ва ВК(векторли квантлаш) алгоритмлари қиёсий таҳлил қилинди.

2-жадвал.

Сухандонга боғлиқ буйруқларни таниш аниқлиги

Алгоритм	Нутқни таниш аниқлиги (%)		Бажарилиш вақти (сек.)
	Матнга боғлиқ	Матнга боғлиқ бўлмаган	
ВК	95	89	3
ЯММ	98	93	8

Унга кўра сўзловчига қатъий боғлиқ тизимларда қатъий матнни калит сўз сифатида ифодалаганда ВК ва ЯММ алгоритмлари мос равишда 96% ва 98% аниқликни кўрсатди.

3-жадвал.

Таклиф этилган алгоритмларни амалга ошириш

Яширин қатламлар сони	Формант частоталари сони	MFCC қийматлари сони	FFT + MFCC		STFT + SVD + MFCC	
			Натижа (%)	Тестлаш вақти (сек.)	Натижа (%)	Тестлаш вақти (сек.)
6	4	12	91	12	99	18
8			94	15	98	20
10			91	20	87	25

4-жадвал.

STFT + SVD + Spectrogram алгоритмларини амалга ошириш

Яширин қатламлар сони	Формант частоталари сони	Спектограмма қийматлари	Натижа (%)	Тестлаш вақт (сек.)
6	4	128x33	93	89
8			87	128
10			84	140

Ушбу 3-жадвал ва 4-жадвалдаги тажриба натижаларидан ҳулоса қилиб, 30-50та гача бўлган чекланган нутқ буйруқларида НММ моделининг яширин қатламлар сони 6-8ни ташкил этиб, таклиф этилаётган алгоритмлар кетма-кетлигида мос равишда 98-99% таниб олиш аниқлигини кўрсатди.

Тажрибалар шуни кўрсатдики, 8-расмда келтирилган дастурий мажмуа умумий структурасида саниб ўтилган алгоритмлар, аниқлилик ва ишлов бериш тезкорлиги бўйича энг самарали экан.

ХУЛОСА

Илмий тадқиқот ишининг бажарилиши мобайнида қуйидаги илмий натижаларга эришилди.

1. Нутқ буйруқларини таниб олишда, шовқин, ҳалақитларга бардошли ва турли ишчи муҳитга мослашувчи NLMS адаптив фильтри алгоритми ишлаб чиқилди. Ушбу филтрлаш алгоритмларини қўллаш нутқ буйруқларининг аниқлигини оширди.

2. Нутқ сигналларига дастлабки ишлов бериш босқичида фреймлаш, жимлик соҳаларини ажратиб олиш, ойналардан ўтқазиб, формант частоталарини аниқлашлар асосида сигнални параметрик кўринишга олиб келиш ва ягона ахборот хусусиятларини ҳосил қилиш алгоритмларининг умумий кетма-кетлиги яратилди.

3. Реал вақт режимида сигнал ўлчамларини қисқартириш мақсадида SVD алгоритмидан фойдаланиш таклиф этилди. Бу нутқ сигнали ахборот хусусиятларини сиқиш ҳисобига процессорга тушадиган юкломани 50%га камайтирди.

4. Нутқ буйруқларини таниб олишда интеллектуал ишлов беришнинг Векторли квантлаш ва яширин Марков моделига асосланган алгоритмлар яратилди. Векторли квантлаш суҳандонни идентификациялашда тезкор деб топилди ва 95% аниқликни кўрсатди. яширин Марков моделига асосланган алгоритм кам сонли кирувчи маълумот орқали ишлай олиш ва алгоритмик жиҳатдан соддароқ бўлганлиги сабабли, аппарат-дастурий платформада нутқ буйруқларини танишда нисбатан яхшироқ самара берди.

5. Имконияти чекланган инсонлар учун овозли бошқарув қулай дастурий интерфейсига эга дастурий мажмуа яратилди.

6. Реал вақт тизимларига қўйилган талаблар асосида нутқ буйруқларини таниб олиш аппарат-дастурий платформаси архитектураси ва функционал элементлари танлаб олинди.

7. Реал вақт режимида ишлайдиган нутқ буйруқларини таниб олишнинг кўп ядроли мобиль процессорига асосланган аппарат-дастурий платформаси яратилди.

8. Ишлаб чиқилган аппарат-дастурий платформада 30та ўзбек тили нутқ буйруқларини таниб олиш турли ишчи муҳитларда амалга оширилди. Аппарат-дастурий таъминотида хотира ҳажмини ошириш билан нутқ буйруқ сўзлари сонини ошириш имконияти мавжуд. Тизимнинг мослашувчанлиги ва мобиллиги аппарат-дастурий платформанинг Андроид ва Линукс операцион тизимларида ишлаш қобилияти билан таъминланади.

**НАУЧНЫЙ СОВЕТ DSc.13/30.12.2019.Т.07.01 ПО ПРИСУЖДЕНИЮ
УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТАШКЕНТСКОМ УНИВЕРСИТЕТЕ
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ**

**ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ**

ШУКУРОВ КАМОЛИДДИН ЭЛЬБОБО УГЛИ

**РАЗРАБОТКА АЛГОРИТМОВ И АППАРАТНО-ПРОГРАММНОЙ
ПЛАТФОРМЫ РАСПОЗНАВАНИЯ РЕЧЕВЫХ КОМАНД**

05.01.04 – Математическое и программное обеспечение вычислительных машин,
комплексов и компьютерных сетей

**АВТОРЕФЕРАТ ДИССЕРТАЦИИ
ДОКТОРА ФИЛОСОФИИ (PhD) ПО ТЕХНИЧЕСКИМ НАУКАМ**

Тошкент – 2021

Тема диссертации доктора философии (PhD) зарегистрирована в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан за B2021.4.PhD/T902

Диссертация выполнена в Ташкентском университете информационных технологий.

Автореферат диссертации на трех языках (узбекский, русский, английский (резюме)) размещен на веб-странице (www.tuit.uz) и на Информационно-образовательном портале «Ziynet» (www.ziynet.uz).

Научный руководитель: Мусаев Мухаммаджон Махмудович
доктор технических наук, профессор

Официальные оппоненты: Назирова Элмира Шодмоновна
доктор технических наук, доцент

Авазов Юсуф Шодиевич
PhD, доцент

Ведущая организация: Ташкентский государственный транспортный университет

Защита диссертации состоится 28 декабря 2021 г. в 16:00 часов на заседании научного совета DSc.13/30.12.2019.T.07.01 при Ташкентском университете информационных технологий. (Адрес: 100202, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-64-43; факс: (99871) 238-65-52; e-mail: tuit@tuit.uz).

С диссертацией можно ознакомиться в Информационно-ресурсном центре Ташкентского университета информационных технологий (регистрационный номер №0/237). (Адрес: 100202, г. Ташкент, ул. Амира Темура, 108. Тел.: (99871) 238-65-44).

Автореферат диссертации разослан 16 декабря 2021 года.
(протокол рассылки № 42 от 14 декабря 2021 г.).



Р.Х.Хамдамов
Председатель научного совета по
Присуждению учёных степеней,
достор технических наук, профессор

Ф.М.Нуралиев
Ученый секретарь научного совета по
присуждению учёных степеней,
достор технических наук, доцент

М.А.Рахматуллаев
Председатель научного семинара
при научном совете по присуждению
ученых степеней, достор технических наук,
профессор

ВВЕДЕНИЕ (аннотация диссертации доктора философии (PhD))

Актуальность и востребованность темы диссертации. В мире роль компьютеров в организации взаимодействия человека с оборудованием очень важна. Самым эффективным средством организации взаимодействия человека и техники является речь. Сегодня доступны различные методы и инструменты для реализации человеко-машинного интерфейса, и одним из самых эффективных инструментов является речь. Управление голосовыми командами широко используется в различных областях: голосовое управление устройствами и оборудованием, голосовой помощник в колл-центрах, в системах «умный дом» и «интернете вещей» (IoT), в голосовой идентификации, интерфейсах голосового управления для людей с ограниченными возможностями и в других областях. В этом случае для обработки речевых сигналов используются сложные алгоритмы распознавания. Научные исследования в этой области ведутся в ведущих странах мира, таких как США, Китай, Япония, Южная Корея, Германия, Франция, Италия и Российская Федерация.

В мире для реализации таких человеко-машинных интерфейсов требуются специальные программно-аппаратные платформы. Но с усложнением алгоритмов обработки речи и увеличением количества речевых команд аппаратно-программная платформа требует резкого увеличения вычислительных ресурсов и объема памяти. Кроме того, большинство человеко-машинных интерфейсов работают в режиме реального времени. При распознавании речевых команд в различных средах на речевые сигналы влияют внешние помехи, а режим реального времени, в свою очередь требует скорости, компактности и простоты использования. Указанные требования к вычислительным ресурсам и скорости обработки могут быть обеспечены применением специальных алгоритмов и аппаратно-программных решений при обработке речевых сигналов.

В Республике особое внимание уделяется развитию информационных технологий и искусственного интеллекта, проводятся научные исследования по созданию систем обработки речевых сигналов, автоматического распознавания речи на узбекском языке, биометрии голоса, систем управления голосом. В Стратегии по дальнейшему развитию Республики Узбекистан по пяти приоритетным направлениям в 2017-2021 годах определены такие задачи как «... внедрение информационно-коммуникационных технологий в экономику, социальную сферу, системы управления»². Реализация этих вопросов важна при разработке математических алгоритмов и программного обеспечения для управления различными техническими устройствами с помощью речевых команд и голосовой биометрии.

² Указ Президента Республики Узбекистан от 7 февраля 2017 года №УП-4947 «О Стратегии действий по дальнейшему развитию Республики Узбекистан»

Данное диссертационное исследование в определенной мере служит осуществлению задач, предусмотренных Указами Президента Республики Узбекистан от 7 февраля 2017 года №УП-4947 «О Стратегии действий по дальнейшему развитию Республики Узбекистан», от 19 февраля 2018 года №УП-5349 «О мерах по дальнейшему совершенствованию сферы информационных технологий и коммуникаций», Постановлением Президента Республики Узбекистан от 17 февраля 2021 года №ПП-4996 «О мерах по созданию условий для ускоренного внедрения технологий искусственного интеллекта», Постановлением Президента Республики Узбекистан от 17 февраля 2021 года №ПП-4996 «О мерах по внедрению Специального режима применения технологий искусственного интеллекта», а также другими нормативно-правовыми документами, касающимися данной сферы.

Соответствие исследования приоритетным направлениям развития науки и технологий Республики. Данное диссертационное исследование выполнено в соответствии с приоритетным направлением развития науки и технологии республики IV: «Информатизация и развитие информационно-коммуникационных технологий».

Степень изученности проблемы. Ряд ученых провели исследования по разработке и совершенствованию методов цифровой обработки речевых сигналов и их автоматического распознавания, а также по их практическому применению. В этом принимали активное участие известные зарубежные ученые LeCun Y, Bengio Y, Abdel-Hamid O, Grazina Korvel, Patel I, И.С. Кипяткова, Н.М. Марковников, А.А. Карпов, М.Н. Гусев, А.Л. Ронжин, О.Мамырбаев, С.Рустамов и другие.

Среди узбекских ученых значительный вклад в разработку алгоритмов, теоретических основ обработки сигналов и распознавания речи внесли М.М.Камилов, М.М.Арипов, Ш.Х.Фозилов, Н.А.Игнатъев, М.М.Мусаев, Х.Н.Зайнидинов, Ў.Р.Хамдамов, Н.С.Маматов, Э.Ш.Назирова и другие ученые.

Обработка речевых сигналов состоит из сложных алгоритмических этапов, поэтому она требует больших вычислительных ресурсов. Реализация распознавания речевых команд в различных средах требует разработки надежного, компактного и высокоэффективного аппаратно-программного обеспечения. Исследования в этой области показали, что аппаратно-программные средства и алгоритмы распознавания речевых команд в различных средах и условиях применения изучены в недостаточном мере. Все сказанное относится и к распознаванию узбекской речи.

Связь диссертационного исследования с планами научно-исследовательских работ высшего образовательного учреждения, где выполнена диссертация. Диссертационное исследование выполнено в соответствии с планами научно-исследовательских работ Ташкентского университета информационных технологий в рамках проектов: №БЕОА-5-005 «Анализ систем распознавания человеческой речи и создание аппаратно-программных средств обработки речевых сигналов» (2016-2018), ЁБВ-Атех-

2018-134 «Анализ алгоритмов распознавания речи на цифровых сигнальных процессорах и разработка программного обеспечения» (2017-2019), при поддержке Программы развития ООН и Министерства по поддержке махалли и семьи Республики Узбекистан №32/20 «Разработка мобильного приложения автоматического распознавания Узбекской речи для слабослышащих людей» (2020-2021).

Целью исследования является анализ быстрых алгоритмов распознавания речевых команд узбекского языка, создание интерфейса голосового управления с помощью речевых команд и реализация алгоритмов распознавания на аппаратно-программной платформе.

Задачи исследования:

анализ и изучение практических программных интерфейсов голосового управления с помощью речевых команд;

разработка алгоритма фильтрации входящих речевых сигналов от различных помех с учетом условий работы;

разработка алгоритма сжатия речевых сигналов в реальном времени;

разработка алгоритмов распознавания речевых команд для систем голосового управления;

разработка аппаратно-программной платформы для распознавания речевых команд в реальном масштабе времени.

Объектом исследования являются речевые сигналы, записанные в различных средах, спектрограммы и информативные признаки сигнала.

Предметом исследования являются программно-аппаратные платформы голосового управления оборудованием с помощью речевых команд, методы и алгоритмы распознавания речи.

Методы исследования включают методы и алгоритмы обработки сигналов, численные методы, методы спектрального анализа, мел-капстральные методы и алгоритмы принятия решений.

Научная новизна заключается в следующем:

разработан алгоритм адаптивной фильтрации для очистки речевых сигналов от шума в различных условиях работы при распознавании речевых команд;

предложен алгоритм выделения информативных признаков речевого сигнала на основе Сингулярного разложения для сокращения вычислительных операций;

предложены алгоритмы распознавания речевых команд на основе метода векторного квантования и скрытой Марковской модели;

разработана и реализована архитектура программно-аппаратной платформы в соответствии с алгоритмами распознавания речевых команд в реальном времени.

Практические результаты исследования заключаются в следующем:

разработаны алгоритмы и программы обработки речевых команд в различных средах, идентификации диктора, перевода речевых команд в текст;

спроектирована функциональная структура аппаратно-программной платформы голосового управления с помощью речевых команд узбекского языка;

разработана аппаратно-программная платформа с удобным интерфейсом пользователя и программный комплекс, работающий в реальном масштабе времени;

программы распознавания речевых команд узбекского языка и голосового управления применены в системах управления производством и специальных школах-интернатах для детей с нарушениями слуха.

Достоверность результатов исследования обоснована результатами апробации предложенных алгоритмов и методов в диссертации, апробацией алгоритмов и методов в публикациях и отчетах международных научных конференций, экспериментальными исследованиями и внедрением разработанной аппаратно-программной платформы.

Научная и практическая значимость результатов. Научная значимость результатов исследования обоснована анализом предложенных алгоритмов распознавания речевых сигналов для формирования речевых команд и реализацией интеллектуальных алгоритмов распознавания на аппаратно-программных платформах.

Практическая значимость результатов исследования подтверждается разработанными алгоритмами и программными комплексами голосового управления на производственных предприятиях, идентификации диктора и распознавания слов узбекской речи для людей с нарушениями слуха.

Внедрение результатов исследования. В результате проведенных исследований в рамках диссертации, на основе разработанных алгоритмов адаптивной фильтрации, распознавания речевых команд и аппаратно-программной платформы:

внедрена аппаратно-программная платформа голосового управления для управления сложным оборудованием на производственных предприятиях ОАО «Узбекистон темир йуллари»(Справка Министерства по информационным технологиям и развитию коммуникаций от 9 ноября 2021 года № 33-8/7908). В результате удалось повысить производительность труда при управлении сложными устройствами с помощью речевых команд;

в специализированной школе-интернате для слабослышащих детей №106 внедрен программный комплекс с удобным интерфейсом для перевода слов узбекской речи в текст, сурдо-текст и язык жестов для людей с нарушением слуха(Справка Министерства по информационным технологиям и развитию коммуникаций от 9 ноября 2021 года № 33-8/7908). В результате время, затрачиваемое на изучение новых узбекских слов с помощью жестов при обучении школьников сократилось на 15-20%.

Апробация результатов исследования. Результаты исследования обсуждались на 15 научно-практических конференциях, в том числе 5 международных и 10 национальных.

Опубликованность результатов исследования. По теме исследования опубликованы всего 23 научные работы, из них 5 статей в научных изданиях, рекомендованных Высшей аттестационной комиссией Республики Узбекистан для публикации основных научных результатов докторских диссертаций, из них 2 в зарубежных журналах и 3 в Республиканских журналах. Получены 4 свидетельства Агентства интеллектуальной собственности о регистрации программных продуктов для ЭВМ.

Структура и объем диссертации. Диссертация состоит из введения, четырех глав, заключения, списка использованной литературы, списка условных обозначений, терминов и приложений. Объем диссертации составляет 119 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обоснована актуальность и востребованность темы диссертации, сформированы цель и задачи, определены объект и предмет исследования, обосновано соответствие исследования приоритетным направлениям развития науки и технологий Республики Узбекистан, изложены научная новизна, практические результаты исследования, обоснована достоверность полученных результатов, раскрывается теоретическая и практическая значимость результатов исследования, приведены результаты внедрения, сведения об опубликованности результатов и структуре диссертации.

В первой главе диссертации «**Анализ алгоритмов обработки речевых сигналов и аппаратно-программного обеспечения**», рассматриваются основные операции цифровой обработки речевых сигналов, задачи извлечения информационных признаков, дан анализ современных аппаратно-программных средств и платформ, рассмотрены программные интерфейсы голосового управления.

Речевой сигнал - это модель сложного динамического процесса, анализ которого опирается на несколько показателей (признаков), описывающих сигнал и его части. Этими основными характеристиками речевых сигналов являются: формантная частота, частота основного тона, спектральные составляющие.

Наиболее распространенный способ анализа речевого сигнала – это анализ отдельных слов. В современных сложных системах распознавания речи распознавание осуществляется в основном по ключевым словам. В настоящее время распознавание речи широко используется в следующих сферах: в управлении промышленным оборудованием и транспортными средствами, в системах «умный дом», при реализации технологий «интернета вещей», создания интерфейсов для людей с ограниченными возможностями, в сферах обслуживания и сервиса.

При использовании систем управления речью в разных средах, частоты шума и помех, добавляемых из окружающей среды, в которой это происходит,

будут различаться. Распознавание речевых команд требует учета характеристик этого шума и особенности узбекских звуков.

Растет спрос на системы распознавания и анализа речи в реальном времени. Системы реального времени предъявляют высокие требования к скорости обработки, при этом объем данных и сложность алгоритмов обработки затрудняют решение задачи на аппаратно-программных средствах.

Поскольку речевые сигналы являются сложными сигналами, возникают проблемы с обработкой этих сигналов, хранением их в памяти и передачей по каналам связи. Для решения этих проблем ученые предложили такие подходы как сжатие сигнала при обработке речевых сигналов, извлечение признаков из сигнала, работа со спектральными значениями сигнала. Все это направлено на совершенствование процесса обработки сигнала.

Для реализации сложных алгоритмов распознавания речи в реальном времени требуются специальные аппаратные и программные средства. В зависимости от уровня сложности алгоритмов распознавания речевых команд требуется использование нескольких встроенных процессоров как единой программно-аппаратной платформы.

Аппаратно-программная платформа (АПП) - это комплекс аппаратного и программного обеспечения, предназначенный для выполнения определенной задачи. АПП построен по модульному принципу и состоит из блока основной платы со специальным процессором, блока памяти, блока ввода/вывода и блоков внешнего интерфейса.

Во второй главе диссертации **«Реализация алгоритмов параметрической обработки речевых сигналов»** описаны этапы обработки сигналов, предложен алгоритм адаптивной фильтрации речевых сигналов в различных средах, рассмотрена особенность применения алгоритмов спектрального анализа речевых сигналов. Описаны алгоритмы снижения нагрузки на вычислительную систему путём применения Сингулярного разложения и повышения эффективности обработки процессора, разработан алгоритм идентификации диктора для дикторозависимых систем распознавания и речевых слов с использованием скрытой Марковской модели

Проблема распознавания речи - это распознавание образов, что всегда было важной темой в области обработки сигналов. Цель распознавания образов - разделить объект на числовые классы. Каждый класс или образец представляет собой определенное слово. Вычислительные операции применяемые при распознавании речи, требуют определенного количества сложных методов и соответствующих алгоритмов обработки. В основном, распознавание речевых слов состоит из прохождения сигнала через окна Хеннинга, применения быстрого преобразования Фурье и получения значений мел-кэпстральных коэффициентов.

Предлагаемая последовательность алгоритмов распознавания речевых команд в режиме реального времени показана на рисунке 1.

Распознавание речевых команд выполняется в следующих алгоритмических шагах.

1. Аналоговый сигнал в виде входной речи преобразуется в цифровой сигнал с частотой 16 кГц.

2. Очистка речевых команд от внешнего шума и помех. В разных средах частота и характеристики шума и помех различаются. Поэтому целесообразно использовать адаптивные фильтры при распознавании речевых команд.



Поскольку каждая среда имеет свои особенности шума и помех, их характеристики также различаются.

Рис. 1. Этапы алгоритмов извлечения информативных признаков и распознавания речевых сигналов

При реализации фильтрации с использованием предложенного алгоритма зашумленный сигнал микрофона, записывается с 5-секундными интервалами и формируются коэффициенты адаптивного фильтра на основе алгоритма NLMS (normalized least mean squares). После процесса настройки речевые команды, проходят фазу фильтрации с помощью этих коэффициентов. При изменении среды коэффициенты перенастраиваются.

3. Удаление зон молчания от речевых сигналов. Перед удалением зон молчания значения речевого сигнала делятся на сегменты равные $M = 256$. Рассчитывается энергия каждого сегмента и выделяются участки с минимальной энергией.

$$E_n = \frac{1}{N} \sum_{i=0}^N Q_i^2$$

Здесь E - энергия речи, Q_i - значения отсчетов сигнала.

4. *Разделение на фреймы.* Необходимо разделить речевой сигнал на фреймы в результате повторной сегментации после удаления зон молчания из речевого сигнала. Согласно результатам исследования, обработка кадров с $M=256$ обеспечивает максимальную скорость.

5. *Применение окна Хеннинга.* Для цифровой обработки сигналов используются следующие оконные преобразования: Бартлетт, Блэкман, Чебышев, Хеннинг, Кайзер, Хэмминг. Из них окна Хемминга и Хеннинга широко используются при обработке речевых сигналов. Целью выбора окна Хеннинга является то, что в используемом нами кратковременном преобразовании Фурье (STFT-short time Furier transform) начало и конец сигнала при прохождении через окно должны быть равны нулю:

$$\omega(n) = 0.5 \left(1 - \cos\left(\frac{2\pi n}{N-1}\right) \right)$$

Здесь N - ширина окон.

6. *Кратковременное преобразование Фурье (STFT).* Правильное выделение информативных признаков речевых сигналов влияет на надёжность распознавания и идентификации диктора.

Поскольку человеческий слуховой аппарат (улитка) представляет собой механический частотно-временной анализатор, он также постоянно принимает короткую последовательность слегка изменяющихся во времени спектров речевого сигнала, что примерно происходит даже при применении STFT к речевым сигналам:

$$STFT_x = STFT\{x[n]\} = X[k, n] = \sum_{m=-\infty}^{\infty} x[m]w[m-n]e^{-jkm}$$

где j - мнимая часть, а множество $\left\{ e^{-\frac{2\pi j}{N}n} \right\}$ - ортонормированный базис.

Полученные значения спектограмм рассматриваются как входной параметр в различных интеллектуальных системах.

7. *Уменьшение размерности.* Метод Сингулярного разложения используется для уменьшения вычислительной нагрузки и повышения эффективности процессора за счет уменьшения размерности сигнала. Сингулярное разложение (SVD - singular value decomposition) - это декомпозиция для приведения реальной матрицы в каноническую форму. SVD очень удобный способ работы с матрицами, а операции с матрицами также повышают эффективность процессора.

Суть SVD проста, любая матрица (вещественная или сложная) представляется как произведение трех матриц и это выражается следующей

формулой:

$$X = U_r D V_r^T = \sum_{i=1}^r \sigma_i u_i v_i^T$$

Здесь,

$$U_r := [u_1, \dots, u_r] \in \mathbb{R}^{m \times r}, \quad V_r := [v_1, \dots, v_r] \in \mathbb{R}^{l \times r}$$

Именно эта D-матрица используется для решения проблемы уменьшения количества размеров сигнала. Следовательно, при уменьшении размерности сигнала алгоритм SVD полагается на сумму дисперсий, заданных рассматриваемыми компонентами. Согласно результатам экспериментов, чем больше количество составляющих дисперсии, тем ближе измеренные значения к реальному сигналу.

8. *Определение формантных частот.* При дикторозависимом распознавании речевых команд, важно использовать формантные частоты в качестве дополнительного параметра. Это в свою очередь приводит к увеличению процента точности распознавания.

При произношении звуков речи по звуковому тракту проходят тональные импульсы или шумовые сигналы. В результате один и тот же тональный спектр или спектр шума становится серией максимальных и минимальных спектров. Максимумы спектра называются формантой, при произнесении речи ее спектр постоянно меняется и образуются формантные переходы. Блок-схема предлагаемого алгоритма определения частот основных формант приведена ниже на рисунке 2.

Большинство систем голосового управления и некоторые системы безопасности используют определенное командное слово или определенное ключевое слово для идентификации человека с помощью речи.

9. *Получение значения MFCC (Mel frequency cepstral coefficients).* MFCC - один из наиболее распространенных методов, аппроксимирующих нелинейные эффекты слуховой системы человека. Это, в свою очередь, означает, что эти коэффициенты очень близки к человеческому слуху.

10. *Проверка соответствия признаков и распознавание.* Речевые команды, объединенные в одном векторе признаков, сначала проверяются на совместимость с сохраненными значениями, и если совместимость определена команда распознается. На данном этапе это реализуется путем решения задач различной классификации и кластеризации речевых команд. В нашем случае использовалась модель, основанная на скрытой Марковской цепи.

Шаги по реализации структуры системы идентификации диктора с помощью речи показаны ниже (рис. 3.).

Идентификация диктора или говорящего по речи осуществляется в двух режимах: в режиме обучения и режиме распознавания.



Рис. 2. Усовершенствованный алгоритм определения основных формантных частот

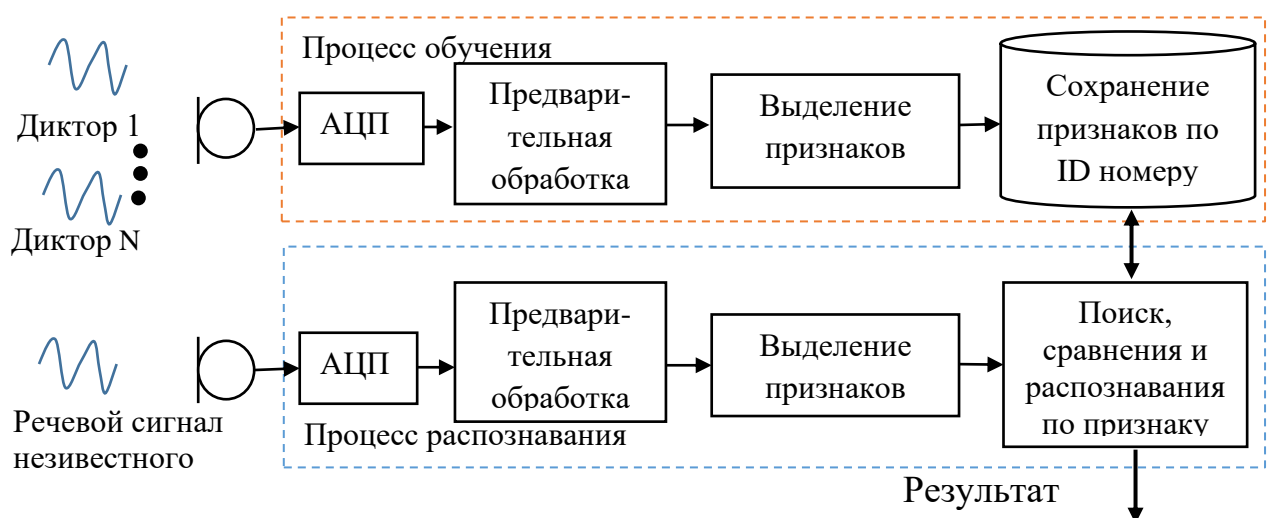


Рис. 3. Структура системы идентификации диктора

В процессе обучения выполняются следующие алгоритмические шаги.

Шаг 1. Входящий аналоговый сигнал с микрофона преобразуется в цифровой.

Шаг 2. Предварительная обработка. В этом случае входящий сигнал фильтруется, сегментируется, разбивается на фреймы и проходит через окна.

Шаг 3. Извлечение информативных признаков.

Шаг 4. Диктору присваиваются идентификационный номер (id) и его голосовые признаки сохраняются в памяти.

В процессе распознавания выполняются следующие алгоритмические шаги:

Шаг 1. Речь диктора, поступающая через микрофон, преобразуется из входящего аналогового сигнала в цифровой код.

Шаг 2. Выполнение предварительной обработки.

Шаг 3. Извлечение информативных признаков.

Шаг 4. Полученные информативные признаки диктора с использованием интеллектуального алгоритма векторного квантования VQ (vector quantization), сравниваются с заранее сохраненными информативными признаками в памяти.

Результаты исследования показали, что алгоритм VQ дал относительно хорошие результаты по сравнению с другими алгоритмами (скрытая Марковская модель, искусственные нейронные сети, DTW-dynamic time warping) при реализации дикторозависимого распознавания речевых команд в режиме реального времени на аппаратно-программной платформе.

Шаг 5. Если информативные признаки говорящего диктора совпадают с признаками, хранящимися в памяти, то в результате говорящий диктор идентифицируется и выдается соответствующий сигнал.

Третья глава диссертации «**Создание аппаратно-программной платформы алгоритмов обработки речевых сигналов**» содержит требования к аппаратным средствам для реализации систем управления с помощью речевых команд. Для облегчения реализации распознавания речевых сигналов в различных средах программирования и аппаратно-программном обеспечении, скомпилирована операционная система реального времени, совместимая с аппаратной архитектурой. В работе предлагается архитектура аппаратно-программной платформы обработки речевых команд.

Учитывая разнообразие аппаратно-программных платформ, архитектурных особенностей и их ограниченные вычислительные ресурсы, на основе проведенных экспериментов можно сказать, что невозможно реализовать полные этапы обработки речевого сигнала на отдельных аппаратных средствах. Поэтому рассматриваются следующие варианты обработки речевого сигнала в разных платформах.

Вариант 1. В системах распознавания речи на базе микроконтроллеров(МК)(рис. 4.).



Рис. 4. Этапы обработки сигналов в системах МК

Анализы показали, что микроконтроллеры дают очень хорошие результаты при предварительной обработке речевых сигналов. Можно легко реализовать алгоритмы оцифровки сигналов, фильтрации, сглаживания, удаление зон молчания.

Вариант 2. В системах распознавания речи на базе специальных сигнальных процессоров и перепрограммируемой логической интегральной схемы (DSP-digital signal processor, FPGA-field programmable gate array)(рис.5.).



Рис. 5. Этапы обработки сигналов в системах DSP, FPGA

На уровне архитектуры *DSP* и *FPGA*, предусмотрена реализация алгоритмов цифровой обработки сигналов. На этой платформе можно легко реализовать сложные операции, такие как выделение признаков, разделение на фреймы, различные преобразования, сжатие, разделение информационных признаков. Однако из-за отсутствия аналогового входа он обязательно используется совместно с МК.

Вариант 3. В системах распознавания речи на базе специальных мобильных процессоров(рис.6.).

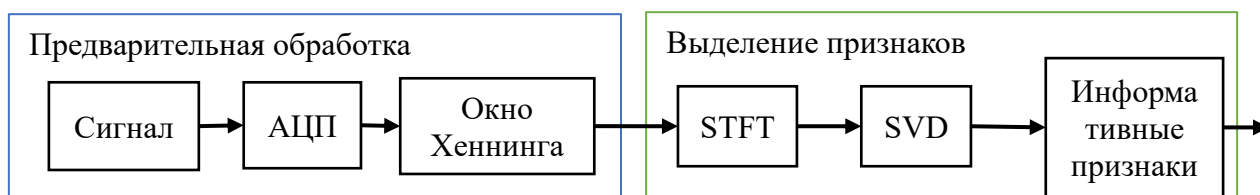


Рис. 6. Этапы обработки сигналов в системах на базе мобильных процессоров

Аппаратно-программные платформы на базе многоядерных мобильных процессоров способны управлять потоками и выполнять многозадачные операции. Кроме того, эти системы обеспечивают относительно простую реализацию сложных алгоритмов обработки речевых сигналов путем установки компиляторов произвольных языков программирования высокого уровня.

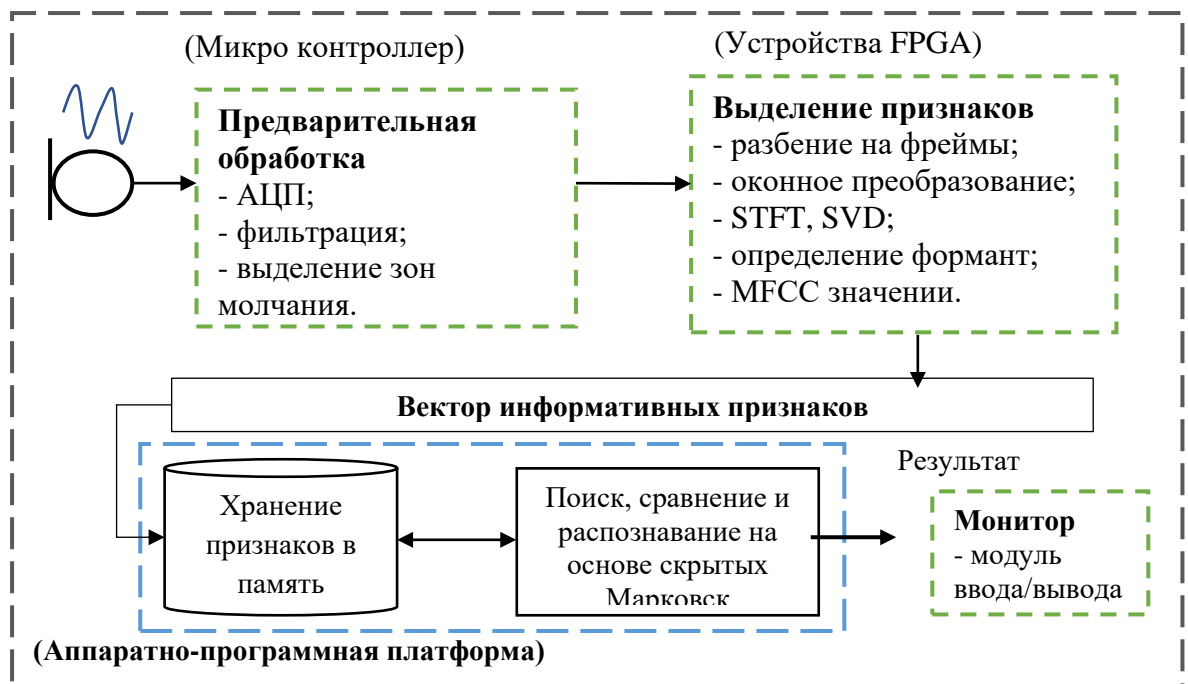


Рис. 7. Реализация алгоритмов распознавания речевых команд на АПП

Учитывая доступность многоядерных процессоров, возможность реализации встроенных операционных систем реального времени и относительно большой размер памяти появляется возможность реализовать полный цикл обработки сложных речевых сигналов.

В результате алгоритмы и программы обработки различных сложных речевых сигналов были реализованы на аппаратно-программной платформе EM3288 под единой операционной системой. По результатам проведенных экспериментов, алгоритмы распознавания речевых команд объединены в предлагаемой АПП(рис. 7.).

Четвертая глава диссертации «**Интерфейс и программный комплекс платформы распознавания узбекских речевых команд**» содержит компоненты модулей комплекса распознавания ограниченных речевых команд и голосового управления, а также алгоритмы и программы интерфейса для систем речевого управления. Программный комплекс речевого управления состоит из модуля предварительной обработки речевых команд, модуля адаптивной фильтрации, модуля идентификации диктора и модулей распознавания речевых команд(рис.8.).

Программные модули реализованы на кроссплатформенном языке программирования Python и могут быть реализованы в операционных системах, принадлежащих семейству Linux и Windows.

Программный комплекс обработки и распознавания речевых сигналов был написан на языке программирования Python для реализации в АПП. Это обеспечивает этому программному комплексу кроссплатформенности.

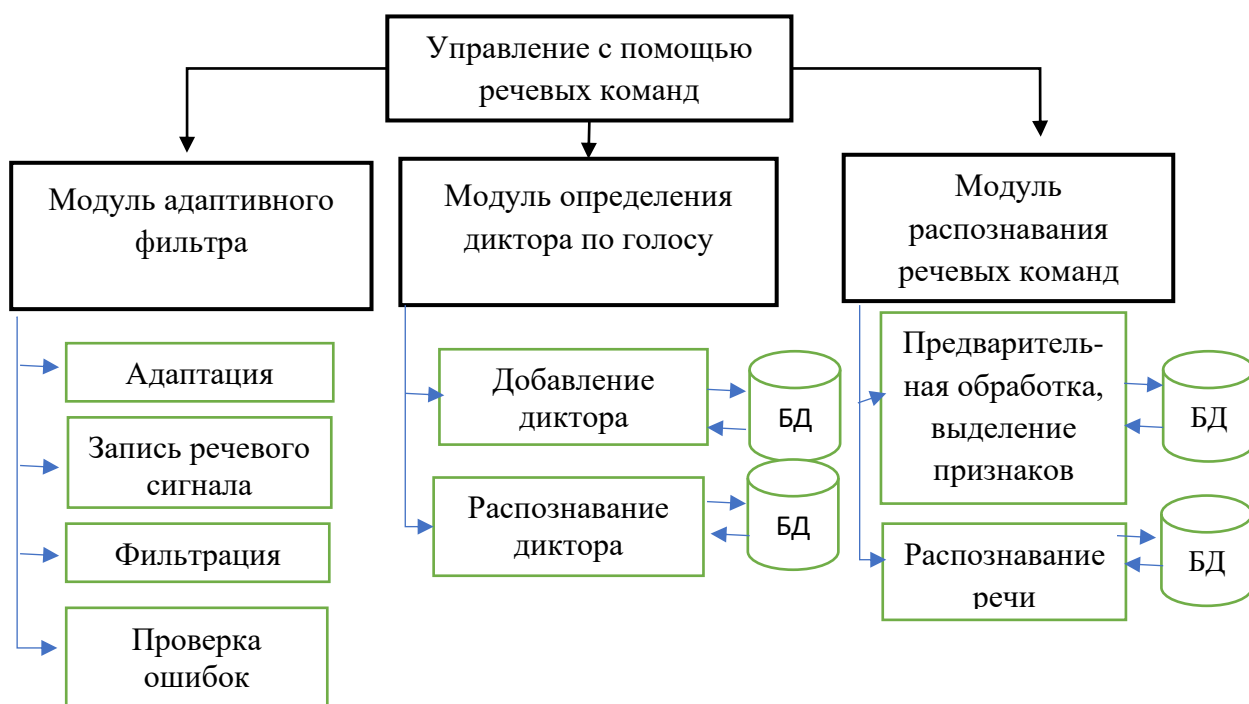


Рис. 8. Общая структура программного комплекса, реализованного в АПП

Программный комплекс настраивается перед запуском системы распознавания речи. В этом случае в течение 5 секунд записывается шум окружающей среды, который изначально требуется для выполнения речевых команд и выбираются коэффициенты фильтрации. Реализуется распознавание записанных речевых сигналов на основе языковых моделей НММ. Распознанные слова отображаются на экран в текстовой форме.

В соответствии с задачей исследовательской работы была реализована система распознавания ограниченного словаря и голосового управления. Были выбраны следующие 30 словесных команд: *“talaba, malika, samalyot, lola, bola, ona, masala, tamlakat, olti, oltin, lochin, lobar, kitob, maktab, archa, chelak, gul, yulduz, sayyora, viloyat, tegirmon, dala, vagon, bekobod, angren, yangiyo`l, yunusobod, umid”*.

По результатам эксперимента, кратковременное преобразование Фурье дало хорошую скорость обработки во встроенных процессорах (табл.1.). Соответственно, классический алгоритм FFT+MFCC предварительной обработки и быстродействие предлагаемого алгоритма были протестированы экспериментально.

Таблица 1.

Скорость предварительной обработки

Методы обработки	Время (мс)
FFT+MFCC	60
STFT+SVD	15

Анализ алгоритмов классификации при распознавании (табл.2). Проведен сравнительный анализ точности искусственных нейронных сетей, скрытой Марковской модели (СММ) и векторного квантования (ВК) в распознавании посредством выделенных информативных признаков речевых сигналов.

Таблица 2.

Точность распознавания дикторозависимого распознавания

Алгоритм	Точность распознавания речи(%)		Время (сек.)
	Текстозависимое	Текстонезависимое	
ВК	95	89	3
СММ	98	93	8

В дикторозависимых системах с произношением определенного текста, алгоритмы ВК и СММ показали точность распознавания 96% и 98% соответственно.

Таблица 3.

Реализация предлагаемых алгоритмов

Количество скрытых слоев	Количество формантных частот	Количество MFCC	FFT + MFCC		STFT + SVD + MFCC	
			Точность (%)	Время (сек.)	Точность (%)	Время (сек.)
6	4	12	91	12	99	18
8			94	15	98	20
10			91	20	87	25

Таблица 4.

Реализация алгоритмов STFT + SVD + Spectogram

Количество скрытых слоев	Количество формантных частот	Спектограмма	Точность (%)	Время (сек.)
6	4	128x33	93	89
8			87	128
10			84	140

Обобщая экспериментальные результаты в таблице 3 и таблице 4, можно видеть, что количество скрытых слоев Марковской модели в ограниченных речевых командах (30–50) составляет 6–8, предлагаемая последовательность алгоритмов показала точность распознавания 98–99% соответственно.

Проведенные эксперименты показали, что наиболее эффективными с точки зрения точности и скорости обработки являются алгоритмы, перечень которых приведен в общей структуре программного комплекса на рис.8.

ЗАКЛЮЧЕНИЕ

В ходе исследовательской работы были получены следующие научные результаты.

1. Для распознавания речевых команд был разработан алгоритм адаптивного фильтра NLMS, который устойчив к шумам, помехам и адаптируется к различным рабочим средам. Использование этого алгоритма фильтрации повысило точность распознавания речевых команд.

2. Для предварительного этапа обработки речевых сигналов было предложено приведение сигнала к параметрическому виду на основе последовательности алгоритмов разбиения на фреймы, выделения зон молчания, прохождения через окна, определение формантных частот и создания единого информационного признакового пространства.

3. В режиме реального времени для уменьшения размерности сигнала было предложено использовать алгоритм SVD. Это снизило нагрузку на процессор на 50% за счет сжатия информационных признаков речевого сигнала.

4. Разработаны интеллектуальные алгоритмы обработки, основанные на векторном квантовании и скрытой Марковской модели. Векторное квантование обеспечило быстрое действие при идентификации диктора и показало 95% точности распознавания. Алгоритм, основанный на скрытой Марковской модели, относительно эффективнее распознает речевые команды на аппаратно-программной платформе, поскольку он может работать с небольшим объемом входных данных и алгоритмически не является сложным.

5. Создан программный комплекс голосового управления, содержащий удобный интерфейс для людей с ограниченными возможностями.

6. На основе требований к системам реального времени были выбраны архитектура и функциональные элементы аппаратно-программной платформы для распознавания речевых команд.

7. Создана аппаратно-программная платформа распознавания речевых команд на многоядерных мобильных процессорах, работающая в реальном масштабе времени.

8. Реализовано распознавание 30 речевых команд узбекского языка в различных рабочих средах на аппаратно-программной платформе. Количество речевых команд может быть увеличено за счет увеличения объема памяти АПП. Гибкость и мобильность системы обеспечивается за счет возможности работы аппаратно-программной платформы на операционных системах Андроид и Линукс.

**SCIENTIFIC COUNCIL AWARDING SCIENTIFIC DEGREES
DSc.13/30.12.2019.T.07.01 AT TASHKENT UNIVERSITY OF
INFORMATION TECHNOLOGIES**

TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES

SHUKUROV KAMOLIDDIN ELBOBO O'G'LI

**DEVELOPMENT OF ALGORITHMS AND HARDWARE-SOFTWARE
PLATFORM FOR SPEECH COMMANDS RECOGNITION**

05.01.04 – Mathematical and software of computers, complexes and computer networks

**DISSERTATION ABSTRACT OF THE DOCTOR OF PHILOSOPHY (PhD)
ON TECHNICAL SCIENCES**

Tashkent – 2021

INTRODUCTION (abstract of PhD dissertation)

The aim of the research work is to analyze the speech signals of the Uzbek language on fast recognition algorithms, create an interface for voice control by means of speech commands, and implement them on a hardware and software platform.

The object of the research work is speech signals recorded in various environments, as well as the spectrogram and character traits formed on their basis.

The scientific novelty of the research work:

an adaptive filtering algorithm has been developed for cleaning speech signals from noise in various operating conditions when recognizing speech commands;

an algorithm for identifying informative features of a speech signal based on Singular value decomposition to reduce computational operations is proposed;

algorithms for recognition of speech commands based on the vector quantization method and the hidden Markov model are proposed;

the architecture of the hardware and software platform has been developed and implemented in accordance with algorithms for recognizing speech commands in real time.

Implementation of the research results. Based on adaptive filtering of speech command recognition algorithms and software and hardware platform, developed as a result of research carried out within the framework of the dissertation:

a hardware and software platform for voice control through adaptive filtering of speech signals in various environments, speaker identification and speech commands has been introduced into the management of complex equipment at the production enterprises of JSC "Uzbekiston Temir Yullari"(Certificate of the Ministry of Information Technologies and Development of Communications dated November 9, 2021 No. 33-8/7908). As a result, it was possible to increase labor productivity when controlling complex devices using speech commands;

in the specialized boarding school for hearing-impaired children No. 106, a software package with a convenient interface was introduced to translate Uzbek words into text, sign language and sign language for people with hearing impairment(Certificate of the Ministry of Information Technologies and Development of Communications dated November 9, 2021 No. 33-8 / 7908). As a result, the time spent on learning new Uzbek words with the help of gestures was reduced by 15-20% when teaching schoolchildren.

Publication of research results. The results of the dissertation were recommended by the Higher Attestation Commission of the Republic of Uzbekistan in 23 scientific articles, including the main scientific results of the dissertation in a total of 5, 2 foreign and 3 national journals. Also received four certificates for software products designed for computers.

Structure and volume of the dissertation. The dissertation consists of an introduction, four chapters, conclusion, a bibliography, references and applications. The volume of the dissertation is 119 pages.

ЭЪЛОН ҚИЛИНГАН ИШЛАР РЎЙХАТИ
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I бўлим (1 часть; part 1)

1. Shukurov K.E. Analysis of algorithms and implementation of real time speaker identification system. // Bulletin of TUIT: Management and Communication Technologies. Vol. 4, Article 2. April 2021. -P.9-18 (Олий аттестация комиссияси Раёсати қарори №283/7.1-сон, 30.07.2020 й).

2. M.M.Musaev, U.A.Berdanov, K.E.Shukurov. Hardware and software solution signal compression algorithms based on the Chebyshev polynomial. // International Journal of Information and Electronics Engineering, Vol. 4, №5, September 2014. -P.380-383(№ 12, Index Copernicus; IF=7.17).

3. Шукуров К.Э. Raspberry pi қурилмасида ўзбек тили нутқ буйруқларини таниб олиш тизимини амалга ошириш. // “ТАТУ ХАБАРЛАРИ” журнали. №2(54)/2020.- Б.45-61 (05.00.00; №31).

4. Шукуров К.Э., Хасанов У.К., Тўраев Б.Ш. Шахсни нутқи орқали идентификациялаш алгоритмларини амалга ошириш. // “ТАТУ ХАБАРЛАРИ” журнали. №1(57)/2021.- Б.32-43 (05.00.00; №31).

5. Shukurov K.E., Xasanov U.K., Allamuratova Z.J. Nutq signallariga ishlov berishda adaptiv filtrlarning ahamiyati. // БЕРДАҚ номидаги ҚОРАҚАЛПОҚ ДАВЛАТ УНИВЕРСИТЕТИНИНГ АХБОРОТНОМАСИ. №3 (52) 2021. -В.30-34 (05.00.00; №27).

II бўлим (2 часть; part 2)

6. Musaev M.M., Berdanov U.A., Rahimov M.F., Shukurov K.E., Parallel Algorithms for Acoustic Processing of Speech Signals. // Proceedings of 2016 IEEE International Conference on Signal and Image Processing. Part II. August 13-15, 2016 Beijing, China. -P. 421-425

7. Shukurov K.E., Kholdorov Sh.I., Khasanov U.K. Uzbek speech commands recognition and implementation based on HMM. // Proceedings of 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT). 07-09 October 2020, Tashkent

8. Berdanov U.A., Shukurov K.E., Nasimov R., Spectral analysis of speech signal on ADSP-BF561 signal processor. // Conference on Digital Policy of Korea “The Society of Digital Policy & Management” 2014. -P.298-300

9. Berdanov U.A., Rahimov M.F., Shukurov K.E., Analysis of performance with implementing algorithms of image processing. “PERSPECTIVES FOR THE DEVELOPMENT OF INFORMATION TECHNOLOGIES ИТРА-2014”. Tashkent-2014. -P.131-133

10. Шукуров К.Э. Нутқ нуқсонига эга инсонлар учун овозли бошқарув интерфейсини яратиш. // “Рақамли технологиялар: соҳаларда амалий жорий

этишнинг ечимлари ва муаммолари” Республика илмий-техник анжуманининг маърузалар тўплами. Тошкент-2021. -Б.51-53

11. Shukurov K.E., Ergashev S.B. Biometrik boshqaruv tizimlarida suxandonni aniqlash masalalariga bo‘lgan yondashuv. // “Иқтисодиётнинг тармоқларини инновацион ривожланишида ахборот-коммуникация технологияларининг аҳамияти” Республика илмий-техник анжуманининг маърузалар тўплами. 1-қисм. Тошкент-2019. -В.458-460

12. Shukurov K.E., Xasanov U.K. Murakkab nutq signallarini adaptiv filtrlash. // “Иқтисодиёт тармоқларининг инновацион ривожланишида ахборот-коммуникация технологияларининг аҳамияти” Республика илмий-техник анжуманининг маърузалар тўплами. Тошкент-2021. -В.300-302

13. Шукуров К.Э., Тўраев Б.Ш. Ўрнатилган рақамли сигнал процессорлари архитектураси. // “Ахборот коммуникация технологиялари ва дастурий таъминот яратишда инновацион ғоялар” Республика илмий-техник конференцияси маърузалар тўплами. Самарқанд-2020. -Б.123-125

14. Шукуров К.Э., Зоҳиров Қ.Р., Обзор современных аппаратных обеспечений обработки сигналов. // “Фан, таълим ва ишлаб чиқариш интеграциясида ахборот-коммуникация технологияларини қўллашнинг ҳозирги замон масалалари” Республика илмий-техник анжуманининг маърузалар тўплами. Нукус-2019. -С.305-307

15. Shukurov K.E., Xasanov U.K., Ob‘ektlarni boshqarishda o‘zbek tilida nutq buyruqlarini tanish va amalga oshirish. // “Ахборот-коммуникация технологиялари ва телекоммуникацияларнинг замонавий муаммолари ва ечимлари” онлайн Республика илмий-техник анжуманининг маърузалар тўплами. Фарғона-2020. -В.77-79

16. Shukurov K.E., Abdurazzoqov F.B., Signallarga spektral ishlov berish algoritmlari tahlili. // “Фан, таълим ва ишлаб чиқариш интеграциясида ахборот-коммуникация технологияларини қўллашнинг ҳозирги замон масалалари” Республика илмий-техник анжуманининг маърузалар тўплами. 1-қисм. Нукус-2015. -В.180-181

17. Шукуров К.Э., Жураев Д.Б., Аудио сигналларнинг дискретлаш частотасини спектрал усуллар асосида ўзгартириш. // “Фан, таълим ва ишлаб чиқариш интеграциясида ахборот-коммуникация технологияларини қўллашнинг ҳозирги замон масалалари” Республика илмий-техник анжуманининг маърузалар тўплами. 3-қисм. Нукус-2015. -Б.446-448

18. Шукуров К.Э., Рахматов Ф.А., Встроенные операционные системы. // “Ахборот-коммуникация технологияларини ривожлантириш шароитида инновациялар” мавзусидаги Республика илмий-амалий анжуман маърузалар тўплами. 1-қисм. Қарши-2019. -С.229-231

19. Shukurov K.E., To‘rayev V.Sh., Cortex-m mikrokontrollerlari yordamida raqamli signalni qayta ishlash. // “Иқтисодиёт тармоқларининг инновацион ривожланишида ахборот-коммуникация технологияларининг аҳамияти” Республика илмий-техник анжуманининг маърузалар тўплами. Тошкент-2021. -В.302-304

20. Мусаев М.М., Берданов У.А., Рахимов М.Ф., Шукуров К.Э. Nutqiy signallarning akustik parametrlarini aniqlovchi dastur. O‘zbekiston Respublikasi intellektual mulk agentligi. elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma № DGU 04020. 11.10.2016

21. Рахматов Ф.А., Берданов У.А., Шукуров К.Э., Рахимов М.Ф., Зохилов Қ.Р., Raqamli signal protsessorlarida nutqiy signallarni spektral tahlil qilish dasturi. O‘zbekiston Respublikasi intellektual mulk agentligi. elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma № DGU 06304. 07.05.2019

22. Шукуров К.Э., Мусаев М.М., Рахматов Ф.А., Хасанов У.К., Холдорев Ш.И., Тўраев Б.Ш., Nutq signallaridan ikki o‘lchovli belgilar kartasini shakllantirish dasturi. O‘zbekiston Respublikasi adliya vazirligi huzuridagi intellektual mulk agentligi. elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma № DGU 10366. 09.03.2021

23. Musaev M.M., Kaxharov A.A., Rahmatov F.A., Xasanov U.K., Xoldorov Sh.I., To‘raev B.SH., Suxandonni ovozi orqali tanib olish va shaxsini aniqlash dasturi. // O‘zbekiston Respublikasi adliya vazirligi huzuridagi intellektual mulk agentligi. Elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro‘yxatdan o‘tkazilganligi to‘g‘risidagi guvohnoma № DGU 11310. 07.06.2021

Автореферат «ТАТУ хабарлари» илмий журнали таҳририятида таҳрирдан ўтказилди ва ўзбек, рус ва инглиз тилларидаги матнларини мослиги текширилди.

