

**O‘ZBEKISTON RESPUBLIKASI OLIY TA’LIM, FAN VA
INNOVATSIYALAR VAZIRLIGI**
**“TOSHKENT IRRIGATSIYA VA QISHLOQ XO‘JALIGINI
MEXANIZATSIYALASH MUHANDISLARI INSTITUTI”
MILLIY TADQIQOT UNIVERSITETI**

**MAMATOV NARZULLO SOLIDJONOVICH,
DUSANOV XURSHID TOSHPULOTOVICH**

**SHAXSNI OVOZI ASOSIDA
IDENTIFIKATSIYALASH VA
AUTENTIFIKATSIYALASH ALGORITMLARI**

**Toshkent – 2024
“Fan ziyosi” nashriyoti**

UDK: 004.93

N.S.Mamatov, X.T.Dusanov. “Shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash algoritmlari”. Monografiya – T.: «Fan ziyosi» nashriyoti, 2024. 131 bet.

Mazkur monografiya shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash usul, algoritm va yondashuvlarini ishlab chiqishga bag‘ishlangan bo‘lib, unda signal amplituda qiymatlarini inobatga olgan holda nutq signalini diskretlash algoritmi taklif etilgan, shuningdek, nutqni tekis talaffuz qilish holatlarida o‘tkazib yuborish ehtimolligi yuqori bo‘lishini inobatga olgan holda kepsstral silliqlik o‘lchovini takomillashtirish orqali chegaralarni aniqlash algoritmi hamda ikki omilli autentifikatsiya shartlarini inobatga olgan holda takrorlanmaydigan so‘zlar matrisasini hosil qilish algoritmi taklif etilgan. Bu shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash xossa hamda xususiyatlarini inobatga olgan holda tanib olish masalasi uchun nutqli ma’lumotlar bazasini shakllantirish yondashuvi hamda signal-shovqin munosabatlari asosida shaxs ovoz modelini shakllantirishning chuqur neyron tarmoq modeli ishlab chiqilgan. tanib olish tizimi aniqligi va qobiliyatini oshirish imkonini ta’minlovchi mavjud va taklif etilgan algoritmlar asosida yaratilgan dasturiy modullarni qamrab oluvchi shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash dasturiy ta’minoti ishlab chiqilgan hamda shaxsni ovozi asosida identifikatsiyalash aniqligi oshirishga xizmat qiladi.

Monografiyadan 60610700-sun’iy intellekt, 60610200-axborot tizimlari va texnologiyalari, 60610600-dasturiy injiniring yo‘nalishlari bo‘yicha bakalavr va magistrilar hamda muhandis va doktorantlarni tayyorlashda foydalanish mumkin.

«Toshkent irrigatsiya va qishloq xo‘jaligini mexanizatsiyalash muhandislari instituti» milliy tadqiqot universiteti ilmiy kengashi tomonidan bosmaga tavsiya etilgan.

Taqrizchilar

S.S. Radjabov – «Toshkent irrigatsiya va qishloq xo‘jaligini mexanizatsiyalash muhandislari instituti» milliy tadqiqot universiteti professori, t.f.d., katta ilmiy xodim.

N. Mirzayev – Raqamli texnologiyalar va sin’iy intellektni rivojlantirish ilmiy tadqiqot insitituti bosh ilmiy xodim t.f.d., professor.

© N.S.Mamatov, X.T.Dusanov
© «FAN ZIYOSI» nashriyoti, 2024

Mundarija

Kirish	4
I BOB. OVOZLI BIOMETRIYA TEXNOLOGIYALARI.....	6
1.1-§. Ovozli biometriyaning asosiy tushuncha va ta'riflari.....	6
1.2-§. So'zlashuvchilarni tanib olish masalasi	13
1.3-§. Inson ovozing o'ziga xosligi	21
1.4-§. So'zlashuvchini tanib olish texnologiyalari va muommalari	25
II. BOB. NUTQ SIGNALLARIGA DASLABKI ISHLOV BERISH USUL VA ALGORITMLARI	40
2.1-§. Nutq signallarini normallashtirish algoritmlari	40
2.2-§. Segmentlash algoritmlarini ishlab chiqish	46
2.3-§. Xalaqitlardan tozalash algoritmlari.....	51
2.4-§. Diskretlash algoritmlari.....	57
2.5-§. Belgilarni shakllantirish algoritmlari	599
III BOB. SHAXSNI OVOZI ASOSIDA IDENTIFIKATSIYALASH VA AUTENTIFIKATSIYALASH ALGORITMLARINI ISHLAB CHIQISH	63
3.1-§. Nutq signallarini segmentlarga ajratish usullari	63
3.2-§. Tanib olish masalasi uchun nutqli ma'lumotlar bazasini shakllantirish....	72
3.3-§. Shaxsni ovozi asosida modelini qurish algoritmlari.....	79
3.4-§. So'zlovchini matnga bog'liq verifikatsiyalash.....	88
IV BOB. SHAXSNI OVOZI ASOSIDA IDENTIFIKATSIYALASH VA AUTENTIFIKATSIYALASH DASTURIY TA'MINOTI VA UNI AMALIY MASALALARDA QO'LLANISHI.....	95
4.1-§. Shaxsni ovozi asosida identifikatsiyalash masalasining ma'lumotlar bazasi modellashtirish	95
4.2-§. Shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash dasturiy ta'minot tuzilmasi va interfeysi	988
4.3-§. Tajribaviy tadqiqot natijalari	103
4.4-§. Dasturiy ta'minotning amaliy masalarni yechishga tadbiqu.....	108
XULOSA.....	113
FOYDALANILGAN ADABIYOTLAR RO'YXATI	11315

KIRISH

Hozirgi kunda biometrik namunalar bilan bog‘liq tizimlarni ishlab chiqishga katta e‘tibor qaratilmoqda. Ushbu yo‘nalishda biometrik belgilarni shakllantirish, qayta ishlash va tanib olishga mo‘ljallangan usul, algoritmlar va dasturiy ta‘minotlarni yaratish muhim masalalardan biri bo‘lib qolmoqda. So‘nggi yillarda ayniqsa, biometrik belgilar asosida identifikatsiyalash va autentifikatsiyalash tizimlari jadal rivojlanmoqda. Chunki biometrik namunalar har bir shaxsning o‘ziga xos bo‘lib, ular ishonchli va yuqori darajadagi xavfsizlikni ta‘minlaydi. Bundan tashqari, biometrik belgilarga asoslangan tizimlarga kirish-chiqishni nazorat qilish, qidiruv, boshqaruv, identifikatsiyalash, autentifikatsiyalash, verifikatsiyalash, xavfsizlik, elektron tijorat va bank kabi sohalarda talab yuqori hisoblanadi. Shaxsni ovozi asosida identifikatsiyalash, autentifikatsiyalash va verifikatsiyalash tizimlari boshqa biometrik belgilarga asoslangan tizimlaridan kam xarajatliligi bilan ustun hisoblanadi. Xorijiy mamlakatlar, jumladan Xitoy, AQSH, Rossiya Federatsiyasi, Yaponiya, Angliya, Hindiston, Germaniya, Janubiy Koreya kabi rivojlangan mamlakatlarda biometrik belgilarni shakllantirish, raqamli ishlov berish hamda ular asosida identifikatsiyalash, autentifikatsiyalash va verifikatsiyalash yo‘nalishlarining nazariy va amaliy masalalarini yechishga katta e‘tibor qaratilmoqda.

Ovozli ma‘lumotlarni qayta ishlash, belgilarini shakllantirish, ovoz asosida identifikatsiyalash, autentifikatsiyalash va verifikatsiyalash usul va algoritmlarini rivojlantirish, ishlab chiqish hamda hisoblash algoritmlarini yaratishga yo‘naltirilgan keng qamrovli ilmiy tadqiqot ishlari olib borilmoqda. Bu borada, jumladan, ovozli ma‘lumotlar sifatini oshirish, nutq signalidan shaxs ovozi joylashgan sohani ajratish va shaxsni xarakterlovchi belgilarini aniqlash, identifikatsiyalash, autentifikatsiyalash hamda verifikatsiyalash usul va algoritmlarini ishlab chiqish, takomillashtirish hamda shaxs ovoz modellari, avtomatlashgan va kontaktsiz xavfsizlik tizimlarini yaratish shu kunning muhim vazifalaridan biri hisoblanadi.

Ayni paytda ovozli ma‘lumotlarga intellektual ishlov berish va tahlil qilish

asosida shaxsni identifikatsiyalash, autentifikatsiyalash va verifikatsiyalashga mo'ljallangan avtomatlashgan xavfsizlik tizimlarini ishlab chiqish va amaliyotga joriy etishga alohida e'tibor qaratilmoqda. «Raqamli O'zbekiston – 2030» strategiyasida, jumladan, «robototexnika komplekslari va odamlar o'zaro ta'sirining algoritmlarini ishlab chiqish, ma'lumotlar uzatish tarmoqlari infratuzilmasini, o'rnatilgan sensorlar va sensor tarmoqlarni takomillashtirish, shuningdek, «bulutli» xizmatlarini taqdim etishning turli xil modellarini amalga oshirish uchun dasturiy ta'minot yaratish bo'yicha ilmiy ishlarni olib borish»¹ hamda 2021-yil 17-fevraldagi PQ-4996-son «Sun'iy intellekt texnologiyalarini jadal joriy etish uchun shart-sharoitlar yaratish chora-tadbirlari to'g'risida»gi Prezident qarorining 2021-2022 yillarda sun'iy intellekt texnologiyalarini o'rganish va joriy etish bo'yicha chora-tadbirlar dasturida «sun'iy intellektni mashina yordamida o'qitish uchun katta hajmda davlat tilidagi raqamli ma'lumotlarni shakllantirish, shuningdek, davlat tilidagi nutqni tahlil va sintez qilishni qo'llovchi dasturiy mahsulotlarni ishlab chiqish, ovozli biometriya va inson nutqini tanish texnologik yechimlarini joriy etish, nutqni tanish texnologik yechimini davlat tili nutqai nazaridan mahalliyashtirish» kabi ustovor vazifalar belgilab berilgan.

Yuqoridagi keltirib o'tilgan fikr va muloahazalar mazkur monografiya mavzusi ovozli biometriya asosida shaxsni identifikatsiyalash va autentifikatsiyalash, nutqni matnga o'tkazish hamda nutq asosida boshqarish tizimlarini ishlab chiqish axborot-kommunikatsiya texnologiyalarini rivojlantirishning dolzarb masalalaridan biri bo'lib, ovozli ma'lumotlarga raqamli ishlov berish model, usul va algoritmlarini, identifikatsiyalash, autentifikatsiyalash hamda verifikatsiyalash, ovozli biometriya uchun belgilar majmuasini shakllantirishni inobatga olgan holda rivojlantirish, shuningdek, ruxsatsiz kirish-chiqishlarni cheklash, nazorat qilish hamda boshqarish dasturiy ta'minotlarida qo'llash muhim ahamiyatga ega hisoblanadi. Mazkur monografiya mavzusi doirasida olib borilgan tajribaviy tadqiqotlar natijalari ichki ishlar tizimida xavfsizlikni oshirishga xizmat qiladi.

¹O'zbekiston Respublikasi Prezidentining 2020-yil 5-oktyabrdagi PF-6079-son «Raqamli O'zbekiston — 2030» strategiyasini tasdiqlash va uni samarali amalga oshirish chora-tadbirlari to'g'risidagi Farmoni

I BOB. OVOZLI BIOMETRIYA TEXNOLOGIYALARI

Mazkur bob ovozli biometriya texnologiyalarini o‘rganishga bag‘ishlangan bo‘lib, unda asosiy tushuncha va ta’riflar, so‘zlovchini tanib olish masalasi, inson nutqini yagonaligi, so‘zlovchini tanib olish usullarini rivojlanish bosqichlari bayon etilgan. Bundan tashqari, ovozli biometriya texnologiyalarining zamonaviy holati va istiqbollari yoritib berilgan. Shuningdek, shaxsni tanib olish muammolari o‘rganilib, ularni bartaraf etish yo‘llari va tadqiqot ishi masalasining qo‘yilishi keltirilgan.

1.1-§. Ovozli biometriyaning asosiy tushuncha va ta’riflari

“Biometriya” tushunchasi 19-asrning so‘nggi yillarida yuzaga kelgan bo‘lib, u insonni o‘ziga xos jismoniy yoki xulq-atvor xususiyatlariga asoslangan tanib olish tizimi hisoblanadi. Biometriya miqdoriy biologik belgilar bilan shug‘ullanadigan maxsus fan sohasi bo‘lib, unda har bir insonni o‘ziga xos va mos belgilari asosida tanib olish amalga oshiriladi. Insonni o‘ziga xos xususiyatlarga barmoq izlari, uch o‘lchamli fotosuratlar, ovoz namunasi, ko‘zining to‘r pardasi kabilar misol bo‘ladi. Biometrikada insonni o‘ziga xos belgi yoki xususiyatlariga ko‘ra turli avtomatlashtirilgan tanib olish tizimlarini yaratishda qo‘llaniladigan amaliy bilim sohasidan keng foydalaniladi [1; 8-24-b, 2; 16-18-b]. Ba’zi ilmiy adabiyotlarda “biometrics” inglizcha so‘z tarjimasida hosil bo‘lgan “biometriya” so‘zi o‘rniga “biometrika” so‘zidan ham keng foydalaniladi. Frensis Galton 1889 yilda nashr ettirgan tabiiy irsiyat muammolariga bag‘ishlangan ishida “biometry” tushunchasi kiritilgan [1; 25-36-b]. Biometriyaga biologik tajribalarda ma’lumotlarni o‘lchash va qayta ishlash, biologiyada matematik modellashtirish, tajribalarni loyihalashda biologik ilovalar kabilar kiradi. Mazkur ishda faqat so‘zlovchini identifikatsiyalash va autentifikatsiyalash masalalari qaralgani uchun “biometrika” tushunchasidan foydalanilgan.

Biometrik texnologiyalar deganda insonning biologik yoki xulq-atvor belgilari bo‘yicha uni tanib olishni avtomatik yoki avtomatlashtirilgan usullari tushuniladi. Har bir insonga individual bo‘lgan tug‘ma yoki hayoti davomida sekin

o'zgaruvchi xususiyati, ya'ni barmoq izi, yuz shakli, ko'z rangdor pardasi, ovoz, qo'l va kaft tuzilishi, tana hidi, qo'lyozmasi kabilar biologik belgilar sifatida qo'llaniladi. Bundan tashqari, yurishi, quloq chanog'i, jins belgilari kabilar ham belgilar sifatida qo'llanilishi mumkin [2; 16-18-b].

Ovoz esa maxsus bilim va ko'nikmalarni talab qilmaydi va u tabiiy vosita bo'lib, barchaga bir xil darajadagi imkoniyatlarni kafolatlaydi. Ovozli biometriya texnologiyalari esa hozirgi kunda eng tez rivojlanayotgan texnologiyalardan biridir.

Inson biometrik belgisi (IBB) deb uni barcha xossa va xususiyatlari yoki xulq-atvorini o'lchashga aytiladi [1; 56-63-b]. Bunday belgilar inson hayoti mobaynida odatiy fiziologik xususiyatlariga asoslangan statistik hamda xulq-atvoriga asoslangan dinamik turlarga ajratiladi. Statistik belgilar odatda ma'lum bir vaqt birligida o'lchanadigan fizik xususiyatlar bo'lib, u inson tug'ilgandan boshlab ega bo'lgan va o'ziga xos xususiyatlar hisoblanadi. Barmoq izi, ko'z qorachig'i, qon tomirlari joylashuvi kabilarni statistik belgilar sifatida keltirish mumkin. Dinamik belgilar esa ma'lum bir vaqt oralig'ida davom etadigan harakatlar bo'lib, bu belgilar insonni biror bir amallarni bajarishida ixtiyorsiz bajaradigan harakatlari natijasida shakllanadi. Dinamik belgilarga imzo, ovoz, yurish kabilar misol bo'la oladi.

Har qanday biometrik belgi universal, yagona, o'zgarmas, o'lchanuvchan, samarador, ijozatli kabi xususiyatlarga egadir [1; 77-81-b, 2; 16-18-b]. Universallik butun IBB mavjudligini, yagonalik bir xil IBBni bo'lmasligini, o'zgarmaslik vaqt bo'yicha IBB o'zgarmay qolishini, o'lchanuvchanlik har bir insondan biometrik belgini qulay va tez olishini, samaradorlik olingan belgi insonni tanib olishni ishonchli amalga oshirishni, ijozatlilik esa IBBni taqdim etishni qulay va tezkorligini ta'minlaydi.

Ma'lum bir biometrik texnologiyalarni yaratishda IBBlarni olish asosiy masala hisoblanadi. Mukammal IBBlar yuqorida keltirilgan xossalarga bajarishi shart. Ko'p hollarda IBB mukammal bo'lmaydi. Bu esa tanib olish va tasniflash masalalarni yechishda ayrim muammolarni keltirib chiqaradi. Shuning uchun olingan IBBdan foydalanishni zaruriylik sharti sifatida universallik va yagonalik olinadi. Odatda izlanuvchilar va biometrik tizimlar ishlab chiquvchilar yagona

biometrik tavsif bo'yicha ishlashni ma'qul ko'rishadi. Zamonaviy jamiyat esa multibiometriyadan foydalanishni talab qilmoqda. Biometrikaning asosiy tushunchalari verifikatsiya, identifikatsiya va autentifikatsiya hisoblanadi. Odatda bu tushunchalar ostida turli mazmunlar yashirilgan bo'ladi va u turli maxsus texnologiyalar bilan bog'liqdir.

Identifikatsiya va autentifikatsiya tushunchasini operasion tizimga kirishda foydalanuvchini autentifikatsiyalash orqali izohlaylik. Bunda foydalanuvchi dastlab avtorizatsiyadan o'tadi. Ro'yxatdan o'tgan foydalanuvchi esa yagona identifikatorga ega bo'ladi va bu jarayon verifikatsiya deb ataladi. Tizimga kirishda foydalanuvchi tizimga o'zini loginini taqdim etadi, tizim esa ushbu loginga ega bo'lgan foydalanuvchini mavjud yoki mavjud emasligini tekshiradi. Mazkur jarayon identifikatsiyalash jarayoni deb ataladi. So'ngra tizim mazkur foydalanuvchini haqiqatdan ham o'zini da'vo qilayotgan shaxs ekanligiga ishonch hosil qilishi shart. Shuning uchun tizim foydalanuvchidan o'z parolini kiritishni so'raydi. Ushbu jarayon autentifikatsiyalash jarayoni deb ataladi. Bunda identifikatsiyalash jarayonini IBBdan foydalanish kabi yagona identifikatorli holda amalga oshirishini autentifikatsiyalash sifatida qarash mumkin. Autentifikatsiyalashni identifikatsiyalashning umumiy tushunchasi bilan aralashtirmaslik kerak. Shaxsni identifikatsiyalash uni virtual ro'yxatini emas, balki bu real inson shaxs aniqlashdir. Biometrikada verifikatsiya, identifikatsiya va autentifikatsiya tushunchalari quyidagicha talqin etiladi.

Ro'yxatdan o'tgan barcha foydalanuvchilar orasida ma'lum bir foydalanuvchini mavjudligini aniqlash maqsadida o'tkaziladigan ko'p martalik taqqoslash (birga ko'p) identifikatsiyalash deb ataladi. Biror bir identifikasion raqam yordamida foydalanuvchi o'zini tizimga tanishtiruvchi identifikatsiyalash (birga bir) jarayoni verifikatsiya deb ataladi. Bunda ro'yxatdan o'tgan barcha foydalanuvchilar orasida ko'p martalik taqqoslash o'rniga yagona taqqoslash amalga oshiriladi va taqdim etilgan biometrik belgilar haqiqatdan ham foydalanuvchini ro'yxatda mavjudligi aniqlanadi. Umumiy ma'noda biometriyada autentifikatsiyalash deganda foydalanuvchini haqiqatdan ham da'vo qilgan shaxs ekanligini tekshirish imkonini

ta'minlovchi usul tushuniladi. Ovozli biometriyada esa so'zlovchini verifikatsiyalash va identifikatsiyalash tushunchalari quyidagicha talqin etiladi.

Tizimga ma'lum bo'lgan so'zlovchilar orasidan berilgan ovozli ma'lumot namunasiga mos yagona so'zlovchini aniqlash identifikatsiyalash deb ataladi. Namunaning so'zlovchilar to'plamida mavjud bo'lishiga bog'liq holda "noma'lum so'zlovchi" yechimiga mos identifikatsiyalash masalasi ochiq ("noma'lum so'zlovchi" so'zlovchilar to'plamida mavjud) va yopiq ("noma'lum so'zlovchi" so'zlovchilar to'plamida mavjud emas) turlarga ajratiladi. Mazkur ishda ochiq va yopiq identifikatsiyalash tushunchasi o'rniga umumiy bo'lgan "so'zlovchini identifikatsiyalash" tushunchasidan foydalaniladi.

Agar berilgan ovoz namunasi asosida uni shu so'zlovchiga tegishli degan xulosani berish mumkin bo'lsa, u holda bu verifikatsiyalash masalasida so'zlovchi ma'lum bo'ladi. Verifikatsiya masalasi ta'rifidan uni yagona so'zlovchiga o'qitilgan tizim bo'yicha ochiq identifikatsiyalash masalasi sifatida qarash mumkinligi kelib chiqadi. Verifikatsiyalash identifikatsiyalashga nisbatan qaror qabul qilish aniqligiga yuqori talab va amaliy muhimlik darajasi bo'yicha alohida sinflarga ajratiladi [3; 56-78-b].

Ovozli biometriyada autentifikatsiyalash deganda masala mohiyatini tushuntirish zarurati bo'lmaganda bir vaqtni o'zida ham verifikatsiyalash va identifikatsiyalash tushuniladi. Agar timsollarni tanib olish tushunchalari asosida autentifikatsiyalash ta'riflansa, u holda uni o'rniga "tanib olish" tushunchasini qo'llash maqsadga muvofiq bo'ladi.

Identifikatsiyalash tizimlari ma'lum bir belgi asosida ikki sinfga ajratiladi, ya'ni u foydalaniladigan ovozli ma'lumot grammatik ifodasidir: matnga bog'liq va matnga bog'liq bo'lmagan. Matnga bog'liq bo'lgan tizimlar grammatik jihatdan aynan bir ovozli ma'lumotdan foydalanadi, ya'ni o'qitish va autentifikatsiyalashda so'zlovchidan aynan bir jumlaning talaffuz qilishi talab qiladi. Matnga bog'liq bo'lmagan tizimlarda esa foydalaniladigan ovozli ma'lumot grammatik tuzilmasiga bog'liq bo'lmaydi, ya'ni o'qitish va autentifikatsiyalashda so'zlovchidan aynan bir jumlaning talaffuz qilishi talab etilmaydi, u ixtiyoriy jumlaning aytishi mumkin.

Kriminalistik identifikatsiyalash tizimlari soʻzlovchini ovozi orqali autentifikatsiyalash tizimlarining dastlabki koʻrinishi boʻlib, ular hozirgi kunda tezkor rivojlanmoqda [4; 1-22-b]. Suhbatlarni tahlil qilish, ovozli maʼlumotlardagi shaxslarni aniqlash, sud tizimi kabilarda bu tizimlardan keng va samarali foydalaniladi. Kriminalistikada ovozli maʼlumotlar asosida identifikatsiyalash texnologiyalarini joriy etishni ilmiy asoslari [5; 46-61-b] ishda batafsil yoritib berilgan. Ovoz maʼlumotlar asosida shaxsni tanib olishda oʻzgarishlar oʻta koʻp boʻlganligi uchun barmoq izi va gen orqali tanib olishdan tubdan farqlanadi. Bundan tashqari, aynan bir ovozli maʼlumotni maʼlum bir shaxsga tegishliligini aniqlovchi maxsus va ishonchli yondashuv hozirgi kungacha mavjud emas.

Soʻzlovchini tanib olish masalasi kriminalistikada faqat ehtimolli xususiyatga ega boʻladi, yaʼni bunda ikkita ovozli maʼlumotlarni aynan bir shaxsga tegishli boʻlish darajasini belgilash bilan cheklanadi. Mobil vositalar orqali soʻzlashuvlarda jinsni aniqlash esa koʻplab muammolarni yuzaga keltiradi. Ovozli maʼlumotlarni uncha katta boʻlmagan tanlanmasida ikkita ovozli yozuvini aynan bir soʻzlovchiga tegishli ekanligini baholash ishonchlilik oraligʻi esa katta boʻladi va bu bir qiymatli yechim olish imkonini bermaydi. Shaxsni ovozi asosida tanib olish masalasi oʻziga xos va u hal etilishi lozim boʻlgan vazifalarga bogʻliq boʻladi. Taqqoslanuvchi ovozli maʼlumotlarni fonetik tuzilmasi esa turlicha boʻlsada, hozirgi kunda kontekstga bogʻliq boʻlmagan holda soʻzlovchini tanib olishning aniqligi yuqori boʻlgan ishonchli tizimlariga boʻlgan talab yuqoridir. Ovoz boʻyicha shaxsni tanib olish ham oʻzining salbiy jihatlarga ega. Jumladan, shaxsiy hayotga aralashish va nazorat qilishda koʻplab muammolar mavjud, amaliy masalalarni yechishda shaxsni ovozi asosida tanib olish afzalliklari esa [6; 112-119-b] ishlarda batafsil yoritilgan. Shaxsni ovozi asosida tanib olishning asosiy afzalliklari quyidagilardan iborat:

- ovoz “tabiiy biometrik belgi”lardan biri boʻlib, undan insonlar ham bir-birini tanib olishda keng foydalanadi;
- ovoz namunalarini sezdirmay oson olish mumkin;
- ovozni insondan ajratib boʻlmaydi.
- ovoz qurilma bilan aloqa qilishni talab qilmaydi;

- tizimga nisbatan amalga oshiriladigan har bir harakatni muallifni aniqlash va u ekanligini isbotlash mumkin. Masalan, tajovuzkor biometrik ma'lumotlarni saqlash orqali;
- ovozni qayta ishlash qimmatbaho uskunalarni talab qilmaydi va mavjud telekommunikatsiya vositalaridan foydalanish mumkin;
- ovoz orqali ma'lum bir kichik vaqt oralig'ida shaxsni aniqlash mumkin;
- tanib olinishi lozim bo'lgan axborotni masofaviy tanlash murakkab;
- avtorizatsiyalash operasion muhit tili va belgilar kodirovkasiga bog'liq emas;
- ovoz bo'yicha tanib olish tilga bog'liq bo'lmaydi.

Har qanday shaxs moddiy yoki axborot manbalarini tassaruf etishi, ma'lumot olishi, hududga kirishi, seyf ochishi kabilarda o'z haq-huquqlarini tasdiqlashi talab etiladi. Odatiy holda tasdiqlash pasport, ID karta, ruxsatnoma, kalit, bank kartasi, kod, parol kabilar orqali amalga oshiriladi. Bunday verifikatsiyalash vositalari ko'p hollarda noqulay yoki yetarli daraja himoyalangan bo'lmaydi. Shaxsni bir omilli autentifikatsiyalash usuli, ya'ni SHIR-kod yoki belgi-raqamli parollar masofiy tizimlarga kirish uchun himoyalaniшни samarali vositasi hisoblanmaydi. Bunday an'anaviy himoyalaniş vositalari bilan birga IBBlarini birgalikda qo'llash xavfsizlik darajasini ortiradi. IBBlarni inson bilan doim birga yurishi, esdan chiqarmasligi, yo'qotmasligi, boshqa shaxsga berib bo'lmasligi, o'g'rilab bo'lmasligi kabi xossalarni mavjud emasligi biometriyaning asosiy afzalliklari hisoblanadi.

Biometriya texnologiyalarida foydalaniladigan belgilar o'zgaruvchan va o'zgarmas turlarga ajratiladi. Barmoq va kaft izi, ko'z qorachig'i, yuz tasviri kabilar har bir shaxs uchun o'ziga xos bo'lib, ular o'zgarmasdir. Biometrik usullarni asosiy kamchiligi undagi belgilarni o'zgarmasligi hisoblanadi va bu kamchilik shaxsni yuqori ishonchlilikda tanib olish talab etilganda ko'plab muammolarni keltirib chiqaradi. Chunki o'zgarmas belgi tanib olish ilovasiga oldindan tayyorlanib uzatilishi mumkin.

Ovoz asosida verifikatsiyalash o'zgarmas belgili biometriyadan farqli nisbatan uzun ovozli xabarlardan foydalanish orqali xatoliklarni kamaytirish uchun

katta imkoniyatga ega hisoblanadi. Ovozli ma'lumotlar asosida verifikatsiyalashdan qorong'u joylarda, turli masofalarda, aloqa kanallarida, tasvirga olish imkoniyati bo'lmagan joylarda samarali foydalanish mumkin.

O'ng miya ayrim kasalliklari ovozni tanib olish qobiliyatini pasayishiga, hattoki yo'qolishiga ham olib kelishi mumkin. Odatda bu kabi kasalliklar yopiq bo'ladi, ayrim hollarda esa ovozni avtomatik tanib olishni zaruratga aylantiradi. Biometriyani boshqa texnologiyalarida bo'lgani kabi ovozli ma'lumotlar bo'yicha so'zlovchini autentifikatsiyalash ham kamchiliklardan xoli emas [2; 16-18-b, 6; 124-130-b], ya'ni:

- berilgan matnni o'qish bo'yicha aniq bir shaxs ovozini taqlid qilishning sun'iy sintezlash imkoniyati mavjud;
- suxandonni ovozi bo'yicha tanib olish aniqligi signal sifatiga kuchli bog'liq bo'ladi. Identifikatsiyalash tizimi esa fon, kanal xalaqiti va mikrofonlar shovqinlariga sezgir hisoblanadi;
- jismoniy, ruhiy yoki karlik kasalliklari sababli ayrim insonlar gapira olishmaydi. Bundan tashqari, boshqa ayrim kasalliklar ovozni vaqtinchalik yo'qolishiga sababchi bo'lishi ham mumkin.
- yosh o'zgarishi, ayrim kasalliklar va hissiyot nutq xususiyatlarini o'zgartirishi ham mumkin.

Suhbatlardagi turli so'zlovchilarni segmentlash ovozli stenogramma, telekonferensiya, tele-radio eshittirish, so'zlashuvlar yozuvlarini so'zlovchilar bo'yicha yoyish kabilarda so'zlovchilarni belgilashda zarur hisoblanadi. So'zlovchi jinsi, suhbat mavzusi, nomlar kabilar avtomatik qidiruv va indekslash imkonini beradi. Kriminalistik ekspertizada bo'lgani kabi segmentlashda ham so'zlovchini hamkorlik qilishni xoxlamaydigan shaxs sifatida qarash kerak, chunki verifikatsiyadan farqli ravishda unda tanib olinish masalasi mavjud emas.

So'zlovchilar almashish vaqtini aniqlovchi va so'zlovchini tanib olishda qo'llaniladigan segmentlash yondashuvlari mavjud. So'zlovchini tanib olish EER xatoligi [5; 46-61-b] mualliflari ma'lumotlariga ko'ra 15.4%ni tashkil etadi.

Ma'lum bir shart-sharoitlarda segmentlash uchun so'zlovchi jinsini aniqlashni

o'zi yetarli bo'ladi. Agar ovozli ma'lumotni yetarli darajada uzun segmenti berilgan bo'lsa, u holda jinsni aniqlash deyarli bexato amalga oshiriladi. Unli tovushli qisqa segmentlarda so'zlovchi erkak ekanligini aniqlash o'rtacha 5.3% ni, ayol ekanligini aniqlash esa o'rtacha 3.1% ni tashkil etadi. Ko'plab ilovalar foydalanuvchidan login, parol va kodlarni eslab qolishini talab qiladi. Foydalanuvchilar esa ularni har doim ham eslab qola olmaydi yoki esidan chiqarib qo'yadi. Bu esa ularni tiklash yoki almashtirish kabi vazifalarni qayta-qayta bajarishni talab qiladi. Bundan tashqari, tiklash va almashtirish jarayonida beriladigan savollarga to'g'ri javoblarni berishga ham ayrim resurslar talab qilinadi. Bundan tashqari, foydalanuvchidan beriladigan savollarga undan to'g'ri javoblar berishi ham talab etiladi. Ayrim hollarda masofaviy, xususan telefon orqali kirishda ovozli verifikatsiya hal qiluvchi ahamiyatga ega bo'lishi mumkin.

Doimiy, kunlik ruxsat, shuningdek bank hisob raqamini boshqarish yoki o'tkazmalarni amalga oshirishda ovozli verifikatsiyadan foydalanish tez va qulay amalga oshirilishi mumkin. Ovozli muloqotlar orqali axborotlarga ruxsatni rasmiy verifikatsiyasiz ham amalga oshirish mumkin. Masalan, axborotlarni olish uchun ovozli so'rov faqat ma'lum bir shaxsga tegishli bo'lsa, u holda bunday so'rovni olishda bu faqat aynan shu shaxsga tegishli deb faraz qilinadi va firibgar hujumi ehtimolligi baholanadi hamda olingan baho asosida ruxsat berish yoki bermaslik bo'yicha qaror shakllantiriladi.

Foydalanuvchilarni avtomatik verifikatsiyalash tizim himoyalanganlik darajasini oshirish, vaqt va manbalarni tejash orqali ruxsat berish jarayonida inson omilini aralashishiga yo'l qo'ymaydi.

1.2-§. So'zlovchilarni tanib olish masalasi

Nutq texnologiyalarining asosiy va muhim masalalaridan biri bu so'zlovchilarni tanib olish masalasi hisoblanadi va unda so'zlovchilarni identifikatsiyalash, autentifikatsiyalash va verifikatsiyalash masalalari mujassamlashgandir. Mavjud ma'lumotlar bazasida saqlangan ovoz ma'lumotlari bilan berilgan ovoz ma'lumotlarini taqqoslash orqali shaxsni aniqlash so'zlovchini

identifikatsiyalash masalasi deb nomlanadi. Identifikatsiyalash natijasi nomzodlar ro'yxati hisoblanadi va bu ro'yxat sanoqli bo'lib, u ma'lum bir bo'sag'a asosida shakllantiriladi. Agar identifikatsiyalash jarayonida ro'yxatdan o'tmagan foydalanuvchini ishtirok etishi nazarda tutilgan bo'lsa, u holda bu ochiq to'plamda, aks holda yopiq to'plamda identifikatsiyalash deb ataladi. Umuman olganda identifikatsiyalash natijasi bo'sh to'plam bo'lishi ham mumkin.

Berilgan ovoz namunasini mavjud ma'lumotlar bazasidagi namunalar bilan aynan o'xshashligini tekshirish jarayoniga suxandonni verifikatsiyalash deb ataladi. Verifikatsiyalash natijasi ijobiy yoki salbiy bo'ladi [7; 1-30-b]. Yuqorida so'zlovchi bilan bog'liq bo'lgan ikki asosiy masala, ya'ni verifikatsiyalash va identifikatsiyalash masalalari bayon etildi. Kirish-chiqishni nazorat qilish tizimlarida so'zlovchini identifikatsiyalash foydalanish shartlariga muvofiq faqat yagona so'zlovchini shaxsni tekshirishni nazarda tutadi. Bunday holda shaxsni aniqlanayotgan so'zlovchi identifikasion raqamini ko'rsatishi shart bo'lmaydi va yakuniy qaror qabul qilish yagona foydalanuvchini kiruvchi ovoz belgilari asosida amalga oshiriladi. Ayrim hollarda alohida so'zlovchini emas, balki ma'lum bir guruhdagi so'zlovchi identifikatsiyalash ham talab etiladi. Bunda so'zlovchi ovozini guruhga tegishliligini aniqlashda ham identifikasion raqamini ko'rsatish talab etilmaydi. So'zlovchilar guruhi ayollar va erkaklardan iborat bo'lganda esa tanlovni qisqartirish maqsadida dastlab jinsni aniqlash amalga oshiriladi. Ovoz namunasini davomiyligi va tanib olish usuliga bog'liq holda jinsni to'g'ri aniqlash xatoligi o'ta kichik, ya'ni nolga yaqin bo'ladi. Foydalanuvchilar soni unchalik katta bo'lmaganda aniq bir so'zlovchini identifikatsiyalash uni identifikasion raqamidan foydalanmasdan ham amalga oshirilishi mumkin. Bunday holda tekshiruv tartib bilan amalga oshiriladi. Identifikatsiyalash o'xshashlikni eng katta qiymatiga ega bo'lgan so'zlovchini tanlash orqali amalga oshiriladi. Bu holda so'zlovchini verifikatsiyalash masalasi dixotomiya masalasi sifatida qaraladi. Biroq, aniqlanayotgan so'zlovchini bazadagi so'zlovchilar belgilari bilan taqqoslashda turli algoritmlar qo'llanilishi mumkin. Odatda belgilarni taqqoslash so'zlovchilarni bazaga birlashtirilgan belgilari asosida ketma-ket yoki sinflashtirish orqali aniqlangan

sinfga xos soʻzlovchilar belgilari asosida amalga oshiriladi. Odatda firibgarlar oldindan maʼlum yoki nomaʼlum boʻladi va ularni guruhini tanib olish maʼlum bir guruhda amalga oshirilganida shakllantiriladi.

Soʻzlovchini identifikasion raqamni verifikatsiyalash tizimlarida kiritish turlicha amalga oshirilib, unda agar foydalanuvchi verifikatsiyalash tizimiga kompyuter orqali kirsam, u holda nisbatan sodda va ishonchli boʻlgan raqam va harfdan iborat boʻlgan kod yoki biror bir identifikatorni tanlash maqsadga muvofiq hisoblanadi. Bunday yondashuvdan foydalanish hujumlarga nisbatan verifikatsiya tizimi turgʻunligini pasaytirmaydi, chunki identifikator faqat kiruvchi ovozli maʼlumot bilan taqqoslanadigan ovoz parametrlari qiymatlari sohasini koʻrsatadi. Ayrim hollarda insonni boshqa biometrik belgilaridan, yaʼni barmoq izi, koʻz qorachigʻi tasviri kabilaridan identifikator sifatida foydalaniladi. Bunday identifikatorlarni kod sifatida ifodalash mumkin boʻlganligi uchun ulardan indentifikasion raqam sifatida foydalanish ham mumkin. Odatda bunday tizimlarni tanib olish xatoligi 100%ni tashkil etmaydi. Shuning uchun qaror qabul qilish boʻsagʻa qiymatini tanib olishni inkor etish ehtimolligini eng kam taʼminlaydigan etib olinishi shart. Biroq, bu kabi yondashuvlardan foydalanish verifikatsiyalashni murakkablashtirib, turli muammolarni yuzaga keltiradi. Masalan, identifikasion raqam sifatida soʻrov yuborilayotgan telefon raqamidan foydalanish mumkin. Foydalanuvchi faqat ovozli xabarlarni yuborishi mumkin boʻlgan hollarda esa tanib olish tizimini oʻzi yoki biror bir maxsus qurilma identifikasion raqamni aniqlashi mumkin. Bunda identifikasion raqam yagona kod koʻrinishida shakllantirilishi shart.

Soʻzlovchini tanib olish ovozli mulohazaga koʻra ikki turga ajratiladi, yaʼni matnga bogʻliq va bogʻliq boʻlmagan turlarga. Umuman olganda tanib olishda foydalaniladigan nutq belgilari matnga ham tilga ham bogʻliq boʻlmaydi. Kriminalistikada ovozli maʼlumotlar ixtiyoriy boʻlib, unda matnga bogʻliq boʻlmagan usullardan keng foydalaniladi. Biroq, mazkur yondashuvni kirishni nazorat qilish tizimlarida qoʻllash istiqbolli hisoblanmaydi. Bunda soʻzlovchi uchun ixtiyoriy jumlaning talaffuz qilish qulay va boshqalari bilan bir xil deb faraz qilinadi. Biroq, bu soʻzlovchidan har safar ongli ravishda yangi matn yaratish va uni qulay

hamda tabiiy talaffuz qilishini talab qiladi. Bu esa soʻzlovchiga maʼlum bir aqliy qoʻshimcha vazifalarni yuklaydi. Shuning uchun yodlab olingan aynan bir xil jumladan foydalanish tavsiya etiladi. Biroq, bu firibgarlar uchun kirishda muayyan parolda boʻlgani kabi eng kam qarshilikka ega boʻlgan yondashuv hisoblanadi. Bu kontekstga bogʻliq boʻlmagan tanib olish tizimini eshitilgan va yozib olingan parolni takrorlash orqali firibgarlarni kirishiga eng kam qarshilik koʻrsatuvchi fiksirlangan parolga ega tizimga aylantiradi. Bundan tashqari, muayyan parol har bir foydalanuvchi uchun ixtiyoriy, biroq soʻzlovchilar ovozli maʼlumotlari bazasini yaratishda esa koʻpincha boshqa jumladan foydalaniladi. Bu esa tanib olish aniqligini pasayishiga olib kelishi mumkin.

Foydalanuvchi va tanib olish tizimi uchun qulay hamda maqbul yechim soʻzlovchiga maʼlum boʻlgan imkon qadar kam soʻzlardan iborat boʻlgan muayyan lugʻatdan foydalanish hisoblanadi. Bunda har bir soʻzlovchiga ajratiluvchi soʻzlar oʻzaro kesishmasligi ular har safar taqdim etilganda tartibi tasodifiy oʻzgartirilishi shart. Bu bilan tizim ishonchliligini yanada oshirish mumkin boʻladi. Maxsus tanlangan jumla raqamlardan tashkil topgan jumlalarga nisbatan xatolikni deyarli ikki barobarga kamaytirish imkonini beradi. Odatda tanib olish tizimlaridan individual yoki jamoaviy koʻrinishda foydalaniladi. Biror bir dastur yoki maʼlumotlarga ruxsat berishda soʻzlovchini tanib olish bevosita shu qurilma orqali amalga oshiriladi. Masofaviy ruxsat berishda esa tanib olish individual yoki jamoaviy amalga oshirilishi mumkin. Bunda asosiy muammo soʻzlovchini ovozini taqlid qilish boʻlib, mohir taqlidchilar bilan tajribalarida ovozni soxtalashtirish darajasi unchalik yuqori emasligi aniqlangan.

Taqlidchilar intonasion konturlarni aʼlo darajada oʻxshatishadi va nutq aniq xususiyatlarini ifodalashadi, biroq lab bilan bogʻliq boʻlgan individual xususiyatlarni yuqori aniqlikda oʻxshata olishmaydi. Bir xil jinsga mansub boʻlgan yaqin qarindoshlar ovozlari maʼlum bir oʻxshashlikka ega boʻladi. Hattoki ovoz yoʻlini oʻxshash anatomiyasiga ega boʻlgan egizaklar ovozlari ham boshqarish dinamikasi boʻyicha bir-biridan farqlanadi.

Soʻnggi yillarda bir shaxs ovozini boshqa shaxs ovoziga oʻxshatish usullarini

ishlab chiqilmoqda. Bu shaxsni ovozi asosida tanib olish tizimlariga jiddiy xavf tug'diradi [7; 1-30-b]. Foydalanuvchi ovoz belgilari asosida soxta ovoz namunalarini yaratish verifikatsiyalash xatoligini keskin oshirib yuborishi mumkin. Ayrim hollarda bu 50%dan ortib ketadi. Bu esa nutqni tahlil qilish usullarini rivojlantirishni, shuningdek soxta ovozlarni hosil qilishda muammolarni yuzaga keltiradigan belgilarni shakllantirishni talab qiladi.

Muayyan paroldan foydalanuvchi verifikatsiyalash tizimlari uchun so'zlovchi nutqini yashirin yoki aloqa kanallari orqali olinishi jiddiy xavf tug'diradi. Bunday tizimlarga yozib olingan nutq asosida uyushtiriladigan hujumlarni aniqlashni ikkita bir xil so'z yoki jumlani taqqoslash yoki so'zlovchini talaffuzini boshqarish tizimi orqali amalga oshirish mumkin. Agar talaffuz to'liq ustma-ust tushsa, u holda buni hujum sifatida baholanadi. Hujumni aniqlashni bunday usuli firibgarga ma'lum bo'lsa, so'zlovchi tomonidan unga berilgan so'zlar talaffuzini bir necha marta yozib olish orqali hujumni muvafaqqiyatli amalga oshirishi mumkin.

Mikrofon orqali verifikatsiyalash tizimlarida ovozli ma'lumotlarni eshitish orqali hujumlar xavfini kamaytirishni fizik usuli mavjud. Bunda asosiy omil sifatida verifikatsiyalash tizimiga amalga oshiriladigan so'zlovchiga nisbatan mikrofonni joylashuvi va mikrofon xususiyatlaridagi farqlar, boshqa bir omil sifatida esa eshittiruvchi tizim tomonidan akustik xossalari asosida hosil qilinadigan signal amplituda-chastotali buzilishlari olinadi. Odatda ovoz qabul qiluvchi va aloqa kanallari orqali hosil bo'ladigan ovozli ma'lumotlardagi buzilishlar ziyon keltiruvchi va verifikatsiyalash tizimi sifatini pasaytiruvchi omil deb qaraladi. Biroq, haqiqiy ovozni soxtasiga almashtirishga urinishlar amalga oshirilganda mazkur omillar foydali bo'lishi ham mumkin. Bu esa o'qitish jarayonida foydalanilgan so'zlovchi ovozi belgilari bilan hujum amalga oshirilayotgan nutq signali orasidagi farqni kattalashishiga olib keladi.

Agar ovozli ma'lumot yoki uning belgilari aloqa kanali orqali firibgar tomonidan olinayotgan va undan ushbu kanal orqali hujum qilishda foydalanilayotgan bo'lsa, u holda bunday hujumga qarshi chora ovozli ma'lumotlarni maxsus kodni qo'shish orqali undan foydalanish hisoblanadi. Bunda

kodni har bir signal uzatilishida almashtirish zarur. Hujumlarni aniqlashda signal uzatiladigan va qabul qilinadigan vaqt momentidan nishon sifatida foydalanish yoki signaldan biror bir belgini o‘chirish va uni oldingi yozuv bilan taqqoslashdan ham foydalanish mumkin. Agar signalni kodlash yoki dekodlash usuli oldindan ma’lum bo‘lsa, u holda uni oson yasash mumkin. Bu holda signalni yanada yaxshiroq maxfiylashtirilishi lozim. Maxfiylashtirishda qo‘llaniladigan nishon tanib olishga xalaqit bermasligi talab etiladi. Shuni esdan chiqarmaslik kerakki, hujumga ko‘riladigan har qanday chora hujumdan to‘liq himoyalaniшни kafolatlamaydi, u shunchaki buzishni murakkablashtiradi.

So‘zlovchilarni avtomatik tanib olish tizimlarini yaratishda xatolikni minimal ehtimolligini bilish hamda undan ma’lum bir tizim samaradorligini baholash ko‘rsatkichi sifatida foydalanish maqsadga muvofiq hisoblanadi [7; 1-30-b]. Shaxsni ovozi asosida tanib olishda bu ehtimollik xatolik ehtimolligi sifatida olinishi mumkin.

Mavjud ishlarni katta qismida tanib olish uchun signal davomiyligi muhimligi ta’kidlab o‘tilgan. Tadqiqotlardan tadqiqotlarga o‘tgan sari aniq baholash o‘zgarib boradi. Biroq, ovoz namunasi qanchalik davomiy, ya’ni uzun bo‘lsa, u holda so‘zlovchini tanib olish aniqligi shunchalik yaxshi bo‘ladi. Dastlabki tajribalarda davomiyligi katta bo‘lgan ovozli ma’lumotlar segmentlarida tanib olish ehtimolligi deyarli ikki barobarga yaxshilangan. Ayrim tajribalarda esa 30 sekundgacha bo‘lgan signaldan foydalanilganda aniqlik ancha past bo‘lgan. Bu kabi natijalar ko‘plab tajribalarda takrorlangan bo‘lib, tanib olish aniqligi kutilgan darajada bo‘lmagan. Ikkita jumlaning aynan bir suxandonga tegishli ekanligini aniqlash masalasi [7; 1-30-b] ish mualliflari tomonidan o‘tkazilgan tajribalarda ham qarab chiqilgan bo‘lib, o‘rtacha 78% aniqlikka erishilgan.

Odatda tanib olish aniqligi tajriba o‘tkaziladigan shart-sharoitga bog‘liq bo‘ladi. Bir xil sharoitlarda yozib olingan nutq signallarini taqqoslashda aniqlik birmuncha yuqori bo‘ladi. Bunda nutqni qayd etish kanallarini ham esdan chiqarmaslik lozim. Kanallar turlicha bo‘lganda esa tanib olish xatoligi deyarli ikki barobarga ortib ketishi mumkin. Ovozni o‘ziga xosligini ovoz yo‘li anatomiyasi,

qo'zg'alish manbasi xususiyatlari, talaffuzni boshqarish tizimi bilan belgilanadi. Shuning uchun turli kontekstlarda suxandonni tanib olish aniqligida farqlar yuzaga keladi.

Nutqni tanib olish aniqligi shaxsni o'z tilida gapirishiga ko'p jihatdan bog'liq bo'ladi. Chunki shaxs o'z fikrini o'z tilida ravon va tushunarli bayon etadi. Bundan tashqari, o'z tilida so'zlaganda nutqni hosil qilish organlari optimal holatda bo'ladi. Talaffuzni boshqarish tizimi esa ovoz yo'li xususiyatlarini inobatga olib, talaffuzni nutq signali akustik parametrlarini gapirilayotgan til diapazoniga moslashtiradi. Shuning uchun ovoz yo'li qanchalik kuchli deformatsiyalansa, ovoz yo'li anatomiyasi akustik xususiyatlarga kamroq ta'sir ko'rsatadi. Unli va unsiz tovushlar bo'yicha shaxsni tanib olish bo'yicha tajribalar [7; 1-30-b] ish mualliflari tomonidan o'tkazilgan bo'lib, unda nutq signalini semantik tarkibini tanib olishga ta'siri ham o'rganilgan.

So'zlovchi tomonidan gapirilgan aynan bir xil jumla nutq signali parametrlari tashqi muhit va so'zlovchi holati, ya'ni ichki muhit bo'yicha o'zgaradi. So'zlashuv uslubi, nutq tezligi va balandligi, shovqinli ovoz kabilar ichki omillar hisoblanadi. Aloqa va akustik kanallardagi shovqin darajalari, qabul qilishdagi buzilishlar hamda muhitdagi aks-sado kabilar tashqi omillar bo'lib, ular tanib olish aniqligiga kuchli ta'sir ko'rsatadi. Bundan tashqari, sheva bilan bog'liq bo'lgan xususiyatlar ham ovozli ma'lumotlar bazasini shakllantirish va tanib olishda turli muammolarni keltirib chiqaradi [7; 1-30-b].

Tabiiy, ya'ni odatiy va so'zlashuv uslubidan tashqari, kuchaytirilgan yoki uncha tushunarli bo'lmagan talaffuz turlari ham nutqni tanib olishda keng qo'llaniladi. Nutq signali akustik xususiyatlari va tezligida nutq uslublari yuzaga keladi. Odatda nutq tezligi jumla uzunligi, muhokama qilinayotgan obyekt murakkabligi, so'zlashuvchini kayfiyati va hissiy holatiga bog'liq bo'ladi. Har bir inson o'ziga xos nutq tezligiga ega bo'lib, u yosh o'zgarishi bilan o'zgarib boradi. Bundan tashqari, u geografik kelib chiqish va jinsga ham bog'liqdir. Mazkur omillar ovozli bazani yaratish va tanib olishda albatta inobatga olinishi zarur.

So'zlovchini ovozi balandligini o'zgarishi nutqni amplituda-chastotali

xususiyatlarini o'zgarishiga sababchi bo'ladi. Xususan, xalaqitli sharoitda ovozni ixtiyoriy bo'lmagan ko'tarilishi nutq signalini yuqori chastotali komponentlari darajasini oshishiga olib keladi va bu Lombard ta'siri deb ataladi. Vaqt o'tishi bilan inson ovozi xossalari uzluksiz o'zgarib turadi. Shuning uchun aynan bir so'zlovchini o'quv tanlanmadagi ovozli ma'lumoti va joriy vaqtdagi ovozli ma'lumoti orasidagi farq verifikatsiyalash tizimi samaradorligiga kuchli ta'sir ko'rsatadi. Bundan tashqari, tanib olish aniqligiga inson salomatligi bilan bog'liq bo'lgan omillar ham kuchli ta'sir ko'rsatishi mumkin. Shuningdek, mikrofon turi va yo'nalishi, so'zlovchidan mikrofondacha bo'lgan masofa nutq signalini amplituda-chastotali xususiyatlarini o'zgarishiga sababchi bo'ladi. Masalan, so'zlovchi va mikrofonni bir-biriga yaqin joylashuvi muhit akustik xalaqiti va nutq signali munosabatini nisbatan yaxshilaydi. Biroq, quyi chastotali sohada mikrofondacha bo'lgan masofaga kuchli bog'liq bo'lgan nutq signali amplituda-chastotali xususiyatlari yaqin akustik maydon effekti yuzaga keladi. Mikrofonli quloqchinlardan foydalanishni esa ko'plab foydalanuvchilar yoqtirishmaydi. Mobil telefonlar turlicha holatida so'zlashuv esa chastotalarni siljishi yoki buzilishiga olib keladi.

Binolarda ovozni asta sekin so'nishi esa uning amplituda-chastotali xususiyatlarni buzilishiga hamda qoldiq signalda yolg'on cho'qqilarni hosil bo'lishiga olib keladi. Shuningdek, ovozni ulanish sohalari formantli chastotalarda tebranishlarni uzoq so'nishiga olib keladi. Shaxsni ovozi asosida tanib olish samaradorligi ovoz belgilarini tanlashga emas, balki kanal xususiyatlarini to'ldirish usuliga ko'proq bog'liq bo'ladi. Hozirgi kunda kanal xususiyatlarini to'ldirishning modelga (model-based), ballga (score-based) va belgiga (feature-based) asoslangan usullari ishlab chiqilgan. Modelga asoslangan usullarda foydalaniladigan kanal modelidan foydalanadigan modelga asoslangan usul ko'p turdagi mikrofonlar uchun stereo yozishni, keyinchalik ular orasidagi almashtirishni xisoblash yoki so'zlovchiga bog'liq bo'lmagan tanib olish rejimida mikrofon turini tanib olishdan foydalanadi. O'rtacha kepstrni ayirish ham qo'llaniladi. Spektr yoki kepstrni logarifmlash kanal ta'sirini multiplikativ xalaqitdan additivga o'zgartiradi, bu esa spektral yoki kepstral ayirish usullaridan foydalanishga imkon beradi. RASTA-PLP

bo'yicha kanal kompensatsiyasi 1 dan 16 Gsgacha bo'lgan modulyasiya bilan signallarni filtrlash imkonini beruvchi kanal xarakteristikalarining stasionarligi faraziga asoslanadi.

Ballga asoslangan usul o'qitishda ham, tanib olishda ham mikrofon turini aniqlaydi, bunda turli mikrofonlar orasidagi so'zlovchi modeli farqi uchun bitta normal taqsimot funksiyasi hisoblanadi va to'ldirish mikrofonga bog'liq bo'lgan siljishni hamda o'rtacha hamda dispersiya bo'yicha masshtablashni ayirish orqali amalga oshiriladi. Agar o'qitish turli mikrofonlar uchun amalga oshirilsa, normallashtirish uning parametrlari har bir kanal uchun alohida hisoblanadi va verifikatsiyada kanal turi xatolar bilan bog'liq bo'lsada, eng yuqori ehtimollikni tanlash orqali aniqlanadi.

Belgiga asoslangan usul o'qitish yoki tanib olishda mikrofon turlari uchun stereo yozish va qo'lda yorliqlashni talab qilmaydi, lekin noxiziqli almashtirish log-spektr yoki kepstr kabi xususiyatlarga qo'llaniladi. Mikrofonlardan bo'lgan masofadagi farq bilan bog'liq bo'lgan kanal xususiyatlarini xususan, kepstrning o'rtacha qiymati to'ldirish signal kelishi vaqtiga T-shaklidagi tekislikda joylashgan 4 tagacha mikrofon masofasini baholash orqali amalga oshirildi. Agar mikrofon masofasi dinamik bo'lsa, u holda real vaqtda moslashish muammosi yuzaga keladi. Verifikatsiya uchun yaroqli, ammo identifikatsiyalash uchun yaroqli bo'lmagan yanada samaraliroq normallashtirish usuli [8; 230-244-b] ishda tavsiflangan, bunda belgi taqsimoti normal taqsimotga keltiriladi. Kanallarni normallashtirishning samarali usuli qator ishlarda taklif qilingan [7; 1-30-b]. Kanallar orasidagi farq qo'shma omilli tahlil (joint factor analysis (JFA)) yordamida aniq ko'rinishda modellashtiriladi. Har bir kanal parametrlari katta ma'lumotlar bazasida joylashgan "xususiy kanallar" (eigenchannels) vektorlari bilan ifodalanadi.

1.3-§. Inson ovozi o'ziga xosligi

Ovoz akustik xususiyatlarini individualligi ovoz burmalari tebranishlari mexanikasi, ovoz yo'li anatomiyasi va artikulyasiyani boshqarish tizimi bilan belgilanadi, shuningdek o'z-o'zidan yuzaga keladigan ovozda lug'atdan

foydalanishni individual xususiyatlari va ovoz aylanishlari ham namoyon bo‘ladi.

Burmalarini o‘z-o‘zidan tebranish jarayonlari asosida tovush burmalar o‘lchamlari, massasi, qattiqligi va yopishqoqlik xususiyatlari, o‘pkadagi bosim yotadi. Burmalarni tebranish chastotasi va tovush chiqaruvchi oraliq orqali o‘tuvchi oqim tezligi impulslari shakli ovoz signalini egiluvchan spektrining shakliga va uning vaqt parametrlariga ta’sir qiladi. Ovoz yo‘lining turli bo‘limlarini geometrik o‘lchamlari va yon bo‘shliqlari, ya’ni halqumdagi nok shaklidagi bo‘shliqlar, ikkita burun bo‘shlig‘i, yuqori jag‘ bo‘shliqlari, shuningdek, ovoz yo‘llari to‘qimalarining mexanik xususiyatlari, uning rezonans chastotalarini va rezonans chastotalarda tebranishlarning yemirilish tezligini aniqlaydi. Nutq signali spektrida bu uning cho‘qqilari chastotasi va kengligi sifatida namoyon bo‘ladi.

Artikulyasiyani boshqarish tizimi prosodik xususiyatlarni, ya’ni asosiy ohang chastotasi dinamikasi, fonetik segmentlar davomiyligi, artikulyatorlar harakat tezligi, shuningdek, turli insonlarda o‘zini turlicha namoyon bo‘ladigan koartikulyasiya effektlarini hosil qiladi. Xususan, individuallikni eng katta ta’siri /u/ uchun aniqlangan. Burunli tovushlar uchun koartikulyasiya so‘zlovchi haqida ma’lumotni o‘z ichiga oladi.

Ovoz uslubining individualligi ancha uzun so‘zlarda namoyon bo‘ladi, masalan, bir nechta so‘zlovchilar ovozini ichiga olgan nutq signallari oqimida so‘zlovchini segmentlash masalalarida foydali bo‘lishi mumkin. Akustik jihatdan uslub ohang chastotasi konturi, so‘zlar va uning segmentlari davomiyligi, zarbdor segmentlar ritmi, to‘xtashlar davomiyligi va ovoz balandligi ko‘rinishida amalga oshiriladi.

So‘zlovchi shaxsi to‘g‘risida qaror qabul qilinadigan belgilar fazosi ovoz hosil bo‘lish jarayonini ovoz manbai, ovoz yo‘li rezonans chastotalari va ularni zaiflashishi, shuningdek artikulyasiyani boshqarish dinamikasi kabi barcha omillarini hisobga olgan holda shakllantirilishi shart. Bunda ovoz manbaining asosiy ohangi o‘rtacha chastotasi, chastota konturi, chastota flyuktuatsiyasi va qo‘zg‘alish zarbasi shakli kabi parametrlari ko‘rib chiqiladi. Ovoz yo‘li spektral xarakteristikalarini spektr egilmasi va uni o‘rtacha qiyaligi, formant chastotalari va

ularni chiziqlari, uzoq muddatli spektr yoki kepsr bilan tavsiflanadi. Bundan tashqari, soʻzlarni davomiyligi, ritm, signal darajasi, chastotasi va toʻxtalishlar davomiyligini ham inobatga olish zarur. Belgilar sifatida asosiy ohang chastotasidan, unli tovushlarni oʻtish va stasionar qismlarida uchta formatli chastotadan, frikativlarni egilma spektri parametrlaridan, shuningdek, soʻzni umumiy davomiyligi va nutq segmentlari nisbiy davomiyligidan foydalanish mumkin.

Prosodik xususiyatlar va yuqori darajadagi xususiyatlar nisbatan barqaror biroq, nisbatan past farqlovchi qobiliyatga ega deb hisoblanadi. Bundan tashqari, ularni taqlid qilish juda oson. Tanlangan parametrlar vaqtli qatorlar yoki uzoq muddatli baholar koʻrinishida qoʻllanilishi mumkin. Bu kontekstga bogʻliq boʻlmagan verifikatsiya tizimlarida faol qoʻllaniladi. Bunday holda, ovoz signali yetarlicha uzun segmentlarini tahlil qilish talab etiladi. 18-20 sekundgacha davom etadigan uzoq muddatli spektr kontekstga bogʻliq va maqbul natijalarga erishish uchun oʻqitish bosqichini odatiy davomiyligi 5 minutdan kam boʻlmasligi talab etiladi, garchi bu ayrim ilovalarda 10 soniyagacha boʻlishi mumkin.

Bola individual taqlid tovushlarini hayotni birinchi yilida talaffuz qilishni oʻrganishi mumkin, keyin bolalar tovush va soʻzlarni qoʻshishni oʻrganadilar, faqat 3-5 yoshga kelib, bolalarcha gapirishni oʻrganadi. Uzoq vaqt davomida inson gapirishni oʻrganishni davom ettirishi shart va bu vaqt davomida ovozini barqarorligi va oʻziga xosligi shakllanadi. Ovoz barqarorligi va oʻziga xosligini asoslash soʻzlovchini tanib olish masalasi uchun juda muhim hisoblanadi. Inson ovozini individualligini namoyon boʻlishini ikkita asosiy belgilar guruhida izlash zarurligini amaliyot koʻrsatgan [9; 20-34-b]. Birinchi guruh ovoz trakti anatomik xususiyatlari bilan bogʻliq boʻlsa, ikkinchisi - markaziy asab tizimi faoliyati bilan bogʻliq boʻlgan uni harakatga keltirishni oʻziga xos mexanizmiga bogʻliqdir. Ushbu guruh inson nutq apparati statik xususiyatlarini aks ettiradi, ikkinchi guruh esa vaqt boʻyicha uni hatti-harakatlarini tavsiflaydi. Individuallik ogʻzaki tovushlarga bogʻliq ekanligini alohida taʼkidlab oʻtish joiz. Ayrim soʻzlovchilar bitta soʻzni talaffuz qiladilar biroq, ayni paytda buning uchun turli xil tovush va allofonlardan

foydalanadilar. Bundan tashqari, soʻzlovchini shevasi va iboralarni talaffuz qilishda uni individualligiga bogʻliqlik ham mavjud ekanligini esdan chiqarmaslik lozim.

Inson anatomik xususiyatlari uning ovoz apparati bilan bogʻliq, jumladan, asosiy ohang chastotasi toʻgʻridan-toʻgʻri ovoz qatlamlari tebranishlariga bogʻliq boʻlib, ular oʻz navbatida navbatdaxisini uzunligi, qalinligi va kuchlanishiga bogʻliq boʻladi, yaʼni turli anatomik xususiyatlar individual boʻladi. Biroq, anatomik asoslash alohida emas, balki artikulyar bilan chambarchas bogʻliqdir. Quyida ovoz unikalligini artikulyar asoslash haqida batafsilroq maʼlumotlar keltirilgan. Bunda nutq yaratilishini kundalik hayotdagi dinamik masala sifatida qarash yetarli. Bunda dastlab mushaklarni boshqarish masalasini hal etish kerak. Odatda nutq hosil qilishda 120 ga yaqin mushak ishtirok etadi. Bular koʻkrakni 44 ta, qorin va qorin boʻshligʻini 9 ta, yuz va jagʻni 28 ta, 12 ta til, 9 ta halqum, 6 ta yumshoq tanglay va halqum mushaklaridir. Shu bilan birga, mushaklarni boshqarish yetarli darajada yuqori, yaʼni 200-3400 Gs oraligʻidagi chastotali jarayonni hosil qiladi.

Ovozni shakllantirish jarayonida harakat imkoniyatlarni haddan tashqari ortiqcha boʻlishi miya va orqa miya kabi hisoblash resurslarini cheklanganligi muammosini keltirib chiqradi. Fiziologlarni miya haqidagi zamonaviy qarashlariga asoslanib neyronlardan hosil boʻlgan miya uch oʻlchamli maʼlumotlarni bir siklda qayta ishlay olmaydi deb aytish mumkin [10; 109-114-b]. Yuqorida taʼkidlab oʻtilganidek, ovoz hosil boʻlishida birinchi oʻrinda 120 ga yaqin mushaklarni boshqarish masalasini hal qilish kerak va bu masalani har bir mushakni boshqarish nohiziqiligi bilan murakkablashishini inobatga olish shart [8; 24-33-b].

Inson kundalik hayotida hal qiladigan boshqaruv masalasi hisoblagich oʻlchamidan ancha katta hajmga ega boʻlishi nafaqat butunlay yangi boshqaruv masalalarini tez va aniq bajara olmasligini, balki qayta oʻqitishga ham qodir emasligidan dalolat beradi. Ushbu masalalarni hal qilish uchun uzoq vaqt davomida oʻz yechimlarini tanlab olish va ularni uzoq vaqt davomida oʻqitish kerak. Ovoz unikalligi miyani ovoz hosil boʻlishi oʻqitish prosedurasini tasodifiyligi va betakrorligi sababli yuzaga keladi, har bir inson ovoz hosil qilishda tez ongsiz harakatlar dinamikasini oʻziga xos tarzda rivojlantiradi. Butun umri davomida inson

ovoz hosil bo'lishini ishlab chiqilgan sxemasidan foydalanadi. Bunday holda, mushaklarni xotirasi deb ataluvchi, mushaklarni barcha boshqaruvi ongsiz darajada sodir bo'lganda paydo bo'ladi. Ovoz hosil bo'lishini individual sxemasini tubdan qayta o'qitishni murakkabligi va ovozli ma'lumotlarni qayta tiklashni ongsizligi darajasi tufayli mushaklar xotirasi vaqt o'tishi bilan ancha barqaror bo'ladi. Inson ovozi barqaror va unikal bo'lib, garchi bu masala qisman hal etilgan bo'lsada, bu so'zlovchini ovozi bo'yicha muvaffaqiyatli tanib olish imkoniyati mavjudligini tasdiqlaydi.

1.4-§. So'zlovchini tanib olish texnologiyalari va muammolari

Ovoz - bu insonni etnik kelib chiqishi, yoshi, jinsi va his-tuyg'ulari kabi inson xususiyatlariga oid ma'lumotlarni uzatuvchi xulq-atvor biometrik ko'rsatkichdir. Shaxsni ovozi asosida tanib olish uni ovoziga ko'ra amalga oshiriladi. Bugungi kunda buyumlar interneti, aqlli qurilmalar, ovozli yordamchilar, aqlli uylar va gumanoidlar kabi texnologiyalarda ovozli ma'lumotlardan foydalanish ommalashdi. Shaxsni identifikatsiyalash va autentifikatsiyalashning avtomatlashtirilgan tizimlari esa xavfsizlik, saylov jarayonlari, kiberjinoyatlar va kirish-chiqishni boshqarishga oid ko'plab ilovalarda keng qo'llanilmoqda. Identifikatsiyalashda biometrik belgilardan foydalanish ismlar, shaxsiy identifikatsiya raqamlari kodlari, parol kabi ananaviy usullarni tezda almashtirmoqda, chunki tabiat insonlarga alohida unikal belgilar bergan. Shaxsni identifikatsiyalashda yuz, barmoq va kaft izi, quloq chanog'i bo'yicha tanib olishga qadar turli choralar ko'rildi. Ushbu paragrafda so'nggi besh o'n yillikda ovozga asoslangan shaxsni identifikatsiyalash tizimlari taklif etilgan asosiy yondashuvlari, ovozga asoslangan mashhur tanib olish usullari, turli biometrik tizimlarni afzalliklari va muammolari yoritib berilgan.

Shaxsni identifikatsiyalash bu shaxsni o'ziga xos xususiyatlar asosida uni tanib olish jarayoni hisoblanadi. Eng keng tarqalgan turi shaxsni ism bo'yicha tanib olish usulidir. Shaxsni dastlabki identifikatsiyalash texnologiyalari maxfiy bilimlar asosida ishlab chiqilgan. Hozirgi kunda shaxsni identifikatsiyalashning ikki asosiy usuli mavjud, ya'ni biometrik va biometrik bo'lmagan identifikatsiyalash. Biometrik

bo'lmagan usullar umumiy maxfiy bilimlar va jismoniy tokenlardan foydalanadi. Yashirin bilim PIN-kod, parol yoki biror bir maxfiy savolga javob shaklida bo'lishi mumkin. Jismoniy tokenlarga esa kalitlar, ID kartalar, xavfsizlik breloklari, haydovchilik guvohnomalari va pasportlar misol bo'ladi.

Biometrik identifikatsiyalash ma'lum bir shaxsni o'ziga xos xususiyatlarini o'lchashga asoslanadi. Bunday xususiyatlarga sifatida inson barmoq izi, ko'z rangdor pardasi, DNKsi, xulq-atvori kabilar kiradi. Bundan tashqari, inson ovozi uni yurishi, imo-ishoralari, qo'lyozmasi kabilar ham insonni o'ziga xos xususiyatlari hisoblanadi [11; 1-3-b.]. Biometrik ma'lumotlardan foydalanish shaxsni tanib olishning eng yaxshi usuli sifatida yuzaga keldi. Chunki ikkita shaxs aynan bir xil biometrik belgilarga ega bo'lmaydi [12; 1-5-b.]. Jismoniy xususiyatlarni o'lchashga asoslangan biometrik belgilar umumiy bo'lib, barcha toifadagi shaxslar uchun qulay shakllantiriladigan xususiyatlardan foydalanadi. Bu xususiyatlar o'zaro ajralib turadi, oson yig'iladi va ma'lumotlar sinfini takrorlash uchun yuqori o'zgaruvchanlikka ega bo'ladi.

Ovozga asoslangan tanib olish tizimida shaxs ovozi xususiyatlari ularni ovoz yo'llari, burun bo'shliqlari va tovush hosil qilishda foydalaniladigan artikulyatorlar xususiyatlariga asoslanadi. Bu xususiyatlar shaxs uchun o'zgarmas, biroq xulq-atvor xususiyatlari vaqt o'tishi bilan yoshi, joylashuvi, tibbiy sharoitlar yoki hissiy holatga ko'ra o'zgarishi mumkin. Ovozga asoslangan tanib olish usullari shaxsni avtomatik verifikatsiyalash (SHAV) va shaxsni avtomatik identifikatsiyalash (SHAI)ga ajratiladi [13; 1-6-b.]. Odatda shaxsni identifikatsiyalash tizimi cheklovlari murakkab, shuning uchun eng zamonaviy biometrik texnologiyalar va yangi biometrik xususiyatlarni o'rganish zarur hisoblanadi. Biometrik autentifikatsiyalash tizimini identifikatsiyalash yoki verifikatsiyalash tizimi sifatida ham qarash mumkin. Identifikatsiya tizimi biometrik ma'lumotlar bazasida saqlangan bir nechta shaxs ma'lumotlarini taqqoslash orqali shaxsini tanib olishda foydalaniladi. Bunda biometrik ma'lumotlar kirishni boshqarishda shaxsini tasdiqlash uchun qo'llaniladi [14; 55-61-b, 15; 113-124-b].

J. Kaur va boshqalar davomat tizimida so'zlovchilarni verifikatsiyalash

ishlarini amalga oshirishgan bo'lib, ular talabalarni avtomatik davomati uchun qo'llaniladigan turli usullarni ko'rib chiqib, ovozdan foydalanishni taklif qilishdi. Bunda Mel filtri banki o'rniga Gammatone filtri bankidan foydalanib, so'ng alohida ovoz signallari uchun diskret kosinus almashtirishi qo'llanilgan. Gauss aralashmali modeli va sun'iy neyron tarmoqlari bilan Gammatone chastotasi kepral koeffitsiyentidan (GFCC) foydalanish o'qitish va moslashtirishda foydalanilgan [16; 8-15-b].

D. Tran va boshqalar ovozga asoslangan tekshirish tizimi samaradorligini oshirish maqsadida noravshan to'plamlar nazariyasiga tegishli bo'lgan normallashtirish texnikasini joriy qildi. Bunda da'vo qilingan shaxs haqiqiylikini tekshirish uchun unga ruxsat berish yoki rad etish uchun ehtimollik chegara qiymati bilan baholanadi. Noto'g'ri qabul qilish darajasiga ta'sir qiluvchi nisbat tipidagi ballar muammosini bartaraf etish shovqinli klasterizatsiya, shuningdek k-yaqin qo'shnilar tegishlilik funksiyasidan foydalanish joriy etilgan. Natijada, noto'g'ri qabul qilish va noto'g'ri rad etish darajasi sezilarli darajada pasaygan [17; 1958–1963-b].

J. Delak, S. Grgich, L. Dyugelay va boshqalar yakka va ko'p biometrik tizimlardan iborat bo'lgan turli biometrik usullarni qisqacha bayon etishgan [18; 4060-4063-b]. Ovozga asoslangan biometrik tizim ma'lum bir shaxs uchun o'zgarmas bo'lgan inson nutqini ayrim xossalardan foydalanadi. Aynan bir shaxs xulq-atvor va nutq xususiyatlari yoshga, tibbiy, hissiy va atrof-muhit sharoitlariga ko'ra vaqt o'tishi bilan o'zgarib turadi. SHAV yakka tartibda verifikatsiyalash tartibida tasdiqlash xususiyati sifatida ovozdan foydalanadi. SHAI esa insonni kimligini aniqlash uchun ovozdan foydalanadi. Shaxsning ma'lum bir ovoz xususiyati ma'lumotlar bazasida saqlangan xususiyatlarga mos keladi. Odatda ovoz xususiyati har bir shaxsni vokal traktiga xos bo'lgan formatlar yoki boshqa har qanday tovush xususiyatlari bo'lishi ham mumkin.

P. Belin va boshqalar neyrokognitiv nutqai nazardan ovozni hosil bo'lishini neyron assosiativiyasini tahlil qilishgan. Mazkur tadqiqotni rivojlanishiga shaxs ovozini tinglash orqali uni jinsi va yoshiga baho berish qobiliyatini aniqlash turtki

bo'ldi. So'ngra ular Bryus va Yangning yuzni tanib olish modelidan ovozni tanib olishda mavjud bo'lgan kognitiv jarayonlarni tushunish uchun tuzilma sifatida foydalanishni taklif etishgan[19; 129-135-b].

Aqlli telefon va mobil internet-bank tizimini jadal rivojlanishini o'ta kuchli autentifikatsiya tizimiga ehtiyojni yuzaga keltirdi [20; 1-9-b]. Bunda telefonlarni qulfdan chiqarish uchun parol yoki yuzdan foydalanish yetarli bo'lmadi. Ovozli biometrik tizimdagi joriy ishlanmalar ovoz yordamida mobil smartfonni ishonchli autentifikatsiyalash uchun katta imkoniyatlarni taqdim etadi. Bu esa moliyaviy va bank sohasi uchun o'ta muhim hisoblanadi. Bunda moliyaviy tashkilotlar moslashuvchan mobil mijozlarga xizmat ko'rsatish va oson autentifikatsiyalashni joriy etish yo'llarini o'ylashmoqda. Shu bilan birga xavfsizlikni ta'minlash va firibgarliklarni kamaytirish yondashuvlari ham tadqiq etilmoqda. Aronowitz va boshqlar joriy texnologiyadan foydalanishni tavsiya etdilar va unda matnga bog'liq bo'lgan so'zlovchini autentifikatsiyalash ma'lum raqamli satrlar, matndan mustaqil tekshirish uchun qo'shma omillar tahlili (Joint Factor Analysis (JFA)) asosida baholanadi. Tizim xavfsizligi va ishonchliligini oshirish uchun noqulay atribut proyeksiyasi (Gaussian mixture models with nuisance attribute projection (GMM-NAP) va NAP bilan Yashirin Markov modeli (HMM) hamda Gauss aralashmasi modellari qo'llanilgan.

R. Korshunov va S. Marsel biometrik texnologiyalarga asoslangan ko'plab tizimlarni firibgarlikka qarshi himoyasi bo'yicha tadqiqotlar o'tkazishgan. Natijada firibgarlikka qarshi aniqlash usullarini ishlab chiqish zarurligini ko'rsatdilar va ASV-PAD gibrid tizimiga asoslangan i-vektorli mahalliy ikkilik namunalari histogrammalardan foydalangan holda PAD va avtomatik dinamik verifikatsiyalash (SHAV) tizimlari integratsiyasini taklif etishdi[21; 5-15-b].

S. Krawczyk va A. Jain qog'ozga asoslangan tibbiy yozuvlardan elektron tibbiy yozuvlarga o'tish jarayonini tahlili asosida shaxsiy va maxfiy ma'lumotlar xavfsizligini kafolatlab bo'lmaydi degan xulosaga kelishgan [22; 1110–1119-b]. Chunki sog'liqni saqlash xodimidan faqat planshet, shaxsiy kompyuterlar yoki smartfonlarda bemor yozuvini tahrirlashi va yangilashi talab etilishi shart. Bunday

yozuvlarga kirish uchun xavfsiz autentifikatsiya tizimi talab etiladi. Bu esa biometrik ma'lumotlarga asoslangan bo'lishi zarurligini bildiradi. Ovozli rejimlar bilan birlashtirilgan onlayn imzo bunday tekshirish tizimidagi foydalanuvchilar uchun eng qulay vosita hisoblanadi.

L. Mazaira-Fernandez va boshqalar atrof-muhit nutqai nazaridan ijtimoiy tarmoqlardan o'z xabarlarini uzatish uchun foydalanishlarini o'rganishgan [23; 1-14-b]. Bunda yuz yoki barmoq izlari kabi biometrik tanib olish o'rniga ovoz kabi boshqa biometrik belgilardan foydalanish maqsadga muvofiq hisoblanadi. Chunki ularni qulay va oson olish mumkin. Shuningdek, mazkur ishda jinsga bog'liq kengaytirilgan biometrik omillarni (JBKO) taklif etishgan. JBKO akustik-fonetik, fiziologik farqlarni inobatga olgan holda, ovoz manbasi va trakt omillaridan olingan xususiyatlar va format ma'lumotlari kabi xususiyatlarni tasniflaydi. Bunda asosiy g'oya kam sonli belgilar yordamida so'zlovchilarni tezkor va yuqori aniqlikda tanib olishni amalga oshirish hisoblanadi.

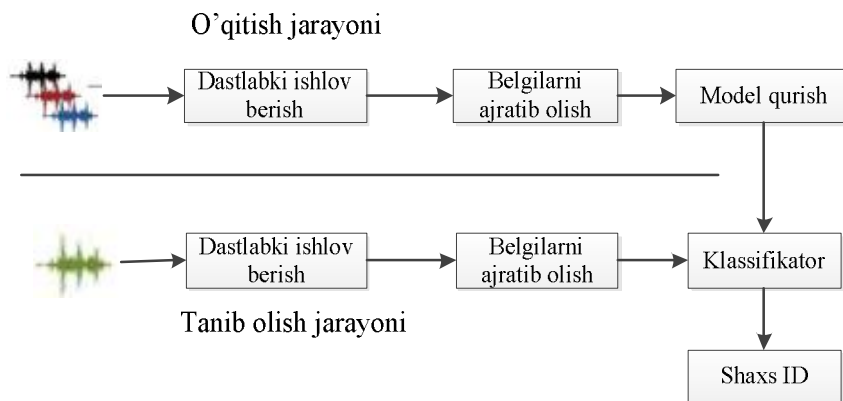
N. Scheffer va boshqalar ovozli biometriyaning ikki muammosi ustida tadqiqotlar o'tkazishgan. Birinchi muammo ideal bo'lmagan sharoitlarda yozib olingan ovozli ma'lumot asosida shaxsni tanib olish bo'lsa, ikkinchisi IARPA, DARPA, RATS kabi loyihalarda yuzaga keladigan natijalarni SRIga moslashtirishdir [24; 447–452-b]. SRI yondashuvida ovozdin ajratib olingan turli belgilar qo'llaniladi va bu belgilar i-vektorlar orqali shakllantiriladi.

K. Bhogal va boshqalar ovozni nisbatan tabiiy muloqot vositasi ekanligini ko'rsatib, uni boshqa aloqa vositalaridan tezkor ekanligini isbotlashgan. Ularning asosiy tadqiqotlari ro'yxatdan o'tgan foydalanuvchini ovozi orqali tasdiqlashdan iborat bo'lib, unda biometrik ma'lumotlarni internetdagi imkoniyatlari o'rganilgan [25; 2-9-b].

N. Zhang va boshqalar identifikatsiyalash tizimini shaxsni frontal bo'lmagan holda tanib olmasligi ustida tadqiqotlar olib borishgan va shaxsni yuz tasviri asosida holatiga invariant holda tanib olish uchun algoritmi taklif etishgan. Olingan natijalar taklif etilgan algoritmi hozirgi kunda aniqligi eng yuqori bo'lgan algoritmlardan ustun ekanligini ko'rsatdi [26; 4804–4813-b].

Nutq texnologiyalari markazi xodimi M. Xitrov esa ovozli biometriyani ma'lumotlarga kirish va xavfsizligida qo'llash bo'yicha katta ishlarni amalga oshirgan bo'lib, u ovozni boshqa biometrik belgilar bilan birgalikda qo'llash masalasini o'rgangan [27; 9-11-b]. M. Vatsa va boshqalar turli xil biometrik texnologiyalarni tadqiq natijasida ko'plab jiddiy muammolarni aniqladilar va ularni aniqlik, hisoblash tezligi, xavfsizligi, sarf-harajati, real vaqt rejimidagi hujumlarga bardoshlilik kabi turlarga ajratdilar. Shuningdek ular biometrik texnologiyalarga qilinadigan hujum turlarini ham tasniflashgan. Masalan, ovozli biometriyada taqlid qilish, majburlash, takroriy hujumlar, shuningdek, xususiyat ekstraktoriga, shablonlar bazasiga, moslashtiruvchi va mos keladigan natijalarga hujumlarni tahlil qilishgan. Biometrik texnologiyalarni rivojlantirish uchun ular biometrik ma'lumotni bu kabi hujumlardan himoyalashni ikkita asosiy usulini taklif etishgan, ya'ni shifrlash va suvli belgilar. Bunda aniqlik 90% dan ortiqni tashkil etgan [28; 293– 304-b].

Biometrik tizim arxitekturasi. Identifikatsiyalashning biometrik tizimi bir nechta bosqichlardan iborat bo'lib, u ovozli ma'lumotlarni kiritish, dastlabki ishlov berish, belgi shakllantirish, so'zlovchi modelini qurish va tanib olish kabi bosqichlarda amalga oshiriladi.



1.1-rasm. Shaxsni identifikatsiyalash tizimining fryeymvorki

Tanib olishda tizim noma'lum shaxsni biometrik ma'lumotlar bazasidagi shaxslar bilan birga ko'p taqqoslashni amalga oshiradi va no'malum shaxsni kim ekanligini aniqlaydi. Agar biometrik namunani ma'lumotlar bazasidagi shablon bilan taqqoslash oldindan berilgan bo'sag'a qiymatidan oshib ketmasa, u holda tizim

shaxsni muvaffaqiyatli aniqlangan deb hisoblanadi. Bunda shaxsdan biometrik tizimdan ro'yxatdan o'tishi talab qilinadi va bu bir marta amalga oshiriladi. Ushbu jarayon davomida shaxsni biometrik ma'lumotlari olinadi va u yoki bu ko'rinishda saqlanadi. Biometrik tizim mukammal bo'lishi uchun u qat'iy himoyalangan bo'lishi shart. Ma'lumotlarni yig'ishda zarur bo'lgan ma'lumotlarni olish uchun bir nechta datchiklar qo'llanilishi mumkin, ya'ni ular inson va tizim orasidagi bog'lovchi vazifasini bajaradi. Dastlabki ishlov berishdan maqsad xalaqitlar va ortiqcha ma'lumotlarni kamaytirish yoki yo'qotish hisoblanadi. Belgilarni shakllantirish bosqichida signalni imkon qadar yaxshi xarakterlovchi, bir-biridan tubdan farq qiluvchi va tanib olish yuqori aniqligini ta'minlovchi belgilar shakllantiriladi. Mazkur bosqich har qanday biometrik tizimning asosiy va muhim bosqichi hisoblanadi. Shablonlarni shakllantirishda esa shablonlarni shakllantirish uchun vektor yoki tasvirlardan foydalaniladi. Shablon bu ko'plab tavsiflovchi funksiyalarni ifodasi bo'lib, u oldindan shakllantiriladi va keyinchalik foydalanish uchun saqlanadi. Ro'yxatdan o'tishda shablon qulay saqlash vositalarida saqlanadi. Taqqoslashda esa shakllantirilgan belgilar boshqa shablonlar bilan taqqoslash uchun uzatiladi va biror bir metrika asosida taqqoslash amalga oshiriladi. So'ngra masofaga ko'ra qaror qabul qilinadi.

Biometriyadan verifikatsiyalashda foydalanish odatda uch bosqichda amalga oshiriladi. Birinchi bosqichda biometrik belgilar namunalari yig'iladi va navbatdagi bosqichda foydalanish uchun etalon namuna sifatida bazaga saqlanadi. Uchinchi bosqichda esa etalon modellar bilan namunalar taqqoslanadi va ruxsat uchun qaror qabul qilinadi. Bundan tashqari, pozitiv tanib olish tushunchasi ham mavjud bo'lib, bu tekshirishning odatiy rejimi hisoblanadi va unda identifikasion ma'lumotlari o'xshash bo'lgan ko'plab shaxslarni tizimdan ro'yxatdan o'tkazish amalga oshiriladi [29; 3122-3126-b]. Biometrik texnologiyada keng qo'llaniladigan algoritmlar batafsil bayoni [30-33] keltirib o'tilgan.

Xavfsizlikka bo'lgan talabning ortishi biometriyani yanada rivojlanishiga olib keldi. Bu esa biometrik tizimlarga qo'yiladigan talablarni kuchaytirmoqda. Ayrim hollarda insonni yuz tasviri asosida tanib olishdan ko'ra, yanada murakkab bo'lgan

o‘ziga xos belgilari orqali ham tanib olish mumkin. Masalan, ko‘z qorachig‘i, ovoz, barmoq yoki kaft izi kabilar orqali. Odatda biometriya ikki katta turga ajratiladi, ya‘ni fiziologik va xulq-atvor turlariga ajratiladi [34; 1–22-b, 35; 153–160-b]. Birinchi turga yuz, barmoq izi, ko‘z qorachig‘i kabilar misol bo‘ladi va bu yo‘nalishda ko‘plab tizimlar ishlab chiqilgan. Ikkinchi turga esa yurish, ovoz, imzo kabilar misol bo‘ladi. Bu yo‘nalishlarda katta ishlar amalga oshirilgan bo‘lsada, bajarilishi lozim bo‘lgan ishlar ham yetarlicha. 1.1-jadvalda biometrik texnologiyalar xususiyatlari va sarf-xarajatlari bo‘yicha solishtirish natijalari keltirilgan [36; 1-17-b, 37; 23-33-b].

1.1-jadval

Turli xil biometrik texnologiyalar xususiyatlari

Biometrik belgi	Biometrik texnologiyalar xususiyatlari				
	Aniqlik	Foydalanish qulayligi	Foydalanuvchini qabul qilish	Amalga oshirish qulayligi	Narxi
Ovoz	O‘rta	Yuqori	Yuqori	Yuqori	Past
Yuz	Past	Past	Yuqori	O‘rta	Past
Ko‘z	O‘rta	O‘rta	O‘rta	O‘rta	Yuqori
Barmoq izi	Yuqori	O‘rta	Past	Yuqori	O‘rta
Yurish	O‘rta	Yuqori	O‘rta	O‘rta	Yuqori
Imzo	O‘rta	O‘rta	Yuqori	Past	O‘rta

Ma‘lumotlar asosida ovoqli biometriya foydalanish qulay bo‘lgan arzon va keng tadbiiq etish mumkin bo‘lgan texnologiya ekanligini aytish mumkin. Ovozni foydali va yuqori xavfsizlik darajasini ta‘minlovchi biometrik ko‘rsatkich ekanligi [38; 1-8-b] ish mualliflari tomonidan ta‘kidlab o‘tilgan. Insonlarni bir-biridan farqlashda ovozdan foydalanish mumkinligi va inson ovozini o‘ziga xosligi [39; 9-18-b] ish ishda asoslab berilgan. Umumiy holda axborot uzatuvchi har qanday tovush ovoz deb ataladi [40; 6-17-b]. Ovoz insonni yoshi, irqi, jinsi va sezgisi kabilarni o‘zida saqlaydi. So‘zlovchini tanib olish bo‘yicha ilmiy izlanishlar bir necha o‘n yillardan buyon amalga oshirilmoqda. Apparat-dasturiy ta‘minotlar, algoritmlar va signallarni qayta ishlashdagi yutuqlar sabab ularni soni so‘nggi yillarda keskin ortdi [41; 1-5-b]. So‘zlovchini avtomatik tanib olish (SATO)

signallarga raqamli ishlov berishning asosiy yo‘nalishlaridan biri bo‘lib, u inson ovozini tanib olish bilan uzviy bog‘liq. Nutq organlari xususiyatlariga ko‘ra har bir shaxs o‘ziga xos ovoz belgilariga ega va ular yagonadir.

So‘zlovchini tanib olish. Bu so‘zlovchi va nutqni tanib olishda ikki xil masala bo‘lib, u tadqiqotlarda aytilgan jumlaning tahlil qilishda qo‘llaniladi [42; 2-7-b]. Bunda nutqni tanib olish aytilgan jumlaning matnga o‘kazish hisoblanadi. So‘zlovchini tanib olish esa jumla aytgan shaxsni aniqlash masalasidir [43; 1-4-b]. Nutqni tanib olishning o‘ziga xos tadbiriq etish yo‘nalishlari mavjud. So‘zlovchini tanib olish esa identifikatsiya va verifikatsiya masalalarini yechishda keng qo‘llaniladi [44; 81-88-b]. 1.2- jadvalda so‘zlovchini tanib olish va nutqni tanib olish texnologiyalari farqlari keltirib o‘tilgan [45; 1-5-b].

1.2- jadval

Shaxsni va nutqni tanib olish texnologiyalari farqlari

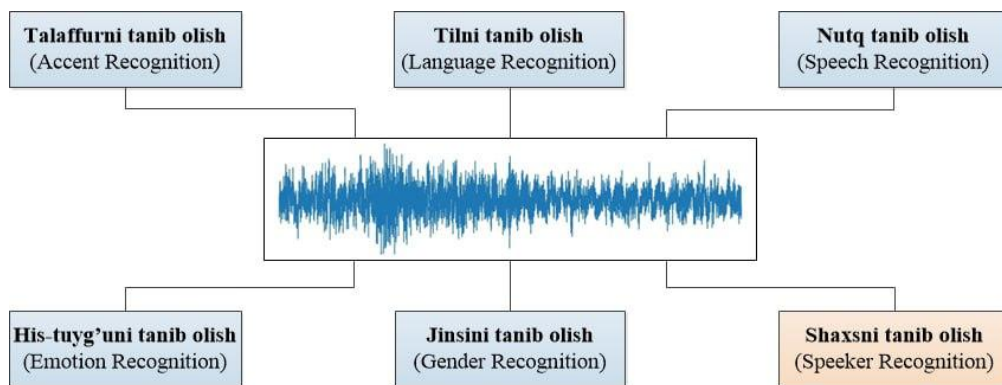
Xususiyat	Shaxsni tanib olish	Nutqni tanib olish
Tanib olish	Ovoz shakli, nutq uslubi va boshqa og‘zaki xususiyatlarni o‘lchash orqali kim gapirayotganini tan oladi	Aytilgan gaplarni tanib oladi va matnga o‘tkazadi
Maqsad	Shaxsni tanib olish uchun	Shaxs aytganlarini aniqlash va raqamli yozib olish.
Focus	Shaxsni tanib olish uchun uning balandligi, intensivligi va boshqalar kabi biometrik jihatlari	Shaxs tomonidan aytilgan so‘zlarning lug‘ati va so‘zlarni raqamli matnlarga aylantiradi
Ilova	Ovozli biometrika	Nutqdan matnga

Turli sohalarni rivojlanishi so‘zlovchini tanib olish tizimlari ahamiyatini yanada oshirdi. Bu ayniqsa, shaxsni tanib olishda yaqqol namoyon bo‘lmoqda. So‘zlovchini tanib olish so‘zlovchini shaxsini aniqlash va mavjud ovozi ma’lumotlar bazasidagi ovozlarni bilan solishtirish jarayonini tashkil etadi va bu birga ko‘p solishtirish orqali amalga oshiriladi [46; 1580–1588-b].

Hozirgi kunda matnga bog‘liq bo‘lmagan nutqni tanib olishning tijoriy tizimlari ko‘plab tillar uchun ishlab chiqilgan bo‘lib, ular kichik tanib olish xatoligiga ega hisoblanadi va bu ularni ko‘plab ilovalarda foydalanish imkoniyatini oshiradi. 2010 yilda asosiy shartlarda eng yaxshi tizim xatoligi 2%dan kam bo‘lgan

[48; 1-10-b]. So‘nggi yillarda timsollarni tanib olish usullarni kengayishi nutqni tanib olishni rivojlanishiga olib keldi. Jumladan, birgalikdagi omilli tahlil samarali tizimlarni ishlab chiqishda asosiy o‘rinni egalamoqda [49; 4410-4413-b, 50; 4205-4208-b]. Shuningdek, bosh komponentalar tahlili, chiziqli diskriminat tahlil, latent omillar tahlili va tasniflashning boshqa ko‘plab usullari ovozli ma’lumotlarni tanib olishda keng qo‘llanilmoqda [51; 308-311-b, 52; 1-8-b]. Quyi va yuqori darajadagi nutq funksiyalarini baholashni umumlashtirish esa xatolik darajasini pasayishiga olib keldi. Biroq, bunda hisoblash sarf-xarajatlari keskin oshdi. Yuqori darajadagi nutq funksiyalariga ega bo‘lgan tizimlarni birlashtirish real vaqt rejimi talablariga javob bermasligi [48; 1-10-b] ishda ta’kidlab o‘tilgan.

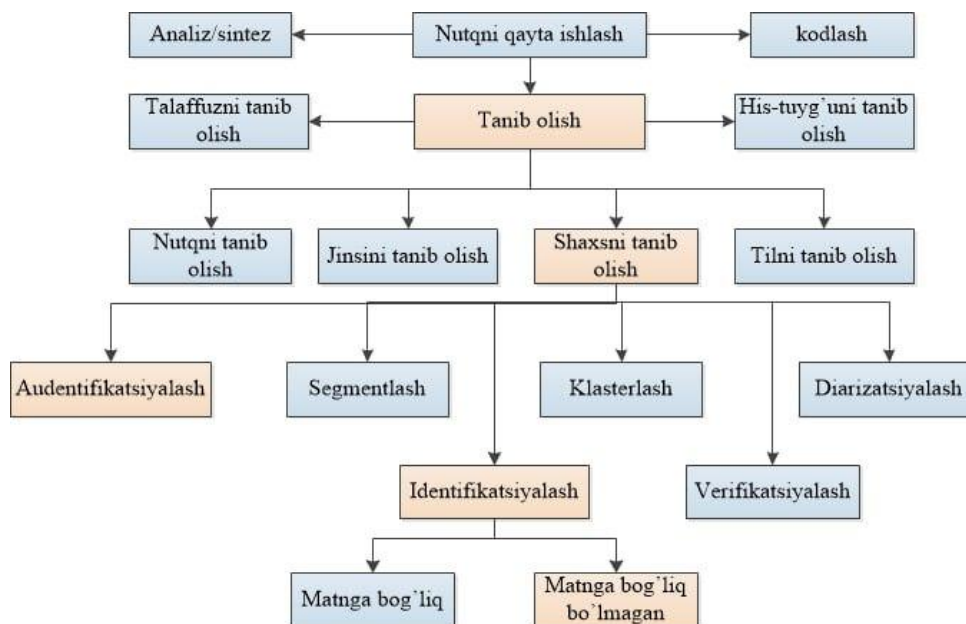
Ayni paytdagi tadqiqotlar ovozni aynan qanday xususiyatlari so‘zlovchilar uchun umumiy degan savolga javob topishga qaratilgan bo‘lib, ovozli xususiyatlar yosh o‘tishi, irqi, ona tili, emosional holati va shevalarga bog‘liq ekanligi aniqlangan [53-54]. Shunday bo‘lsada, mazkur yo‘nalishda asosiy tadqiqotlar atrof-muhitni o‘zgarishi bilan bog‘liq bo‘lgan muammolarni hal etishga qaratilgan [55, 56; 1-9-b]. Bundan tashqari, ekologik muammolar ham asosiy muammolardan biri hisoblanadi. Shuning uchun kelgusi tadqiqotlarni aynan ushbu muammo bilan bog‘liq bo‘lgan masalalarni yechishga yo‘naltirish zarur.



1.2-rasm. Og‘zaki tilda mavjud bo‘lgan ba’zi ma’lumotlar

Nutq qayta ishlash texnologiyalari mashhur aloqa texnologiyasiga aylandi, chunki ko‘plab ilovalari esa inson hayot faoliyatini yaxshilashga yo‘naltirilgan. Signallarga raqamli ishlov berish asosiy yo‘nalishlaridan biri bu ovozi asosida shaxsni tanib olish hisoblanadi [57, 58; 34–40-b]. Bunda inson ovozi muhim

ahamiyatga ega bo‘lib, u insonni muhim hayotiy xususiyatlarini o‘zida jamlaydi [59; 1-9-b]. Sheva, aksent, nutq, jins kabilar ovozda saqlanadigan asosiy ma’lumotlar hisoblanadi [60; 138–43] (1.2-rasm). So‘zlovchini tanib olish yo‘nalishi so‘nggi yillarda intellektual qurilma, ovozli yordamchilar, aqlli uy, buyumlar interneti kabilarda keng qo‘llanilganligi uchun mashhurlikka erishdi. Quyidagi rasmda nutqni qayta ishlashning batafsil tasnifi keltirilgan.



1.3-rasm. Nutq qayta ishlash tasnifi [61,62].

So‘zlovchini tanib olish turlari. So‘zlovchini tanib olish identifikatsiyalash, verifikatsiyalash, aniqlash, segmentlash, klasterlash va diarizatsiyalash kabi turlarga ajratiladi [63; 1085–95] (1.4-rasm) va ularni qisqacha bayoni 1.3-jadvalda keltirilgan.



1.4-rasm. Shaxsni ovozi asosida tanib olish

1.3-jadval

Shaxsni tanib olish turlari

Shaxsni tanib olish	Tavsif
Shaxsni identifikatsiyalash	Shaxsni identifikatsiya qilish vazifasi alohida aytilgan noma'lum ovozni N ta mos yozuvlar shaxslaridan biriga tegishli deb tasniflashdan iborat. [64].
Shaxsni verifikatsiyalash	Shaxsni verifikatsiyalash vazifasi noma'lum ovozning ikkita mumkin bo'lgan ma'lum bir ma'lumotli shaxsga tegishli yoki yo'qligini aniqlashdir: mos yozuvlar shaxsni qabul qilish yoki mos tushmaganini rad etish. [64].
Shaxsni aniqlash	Shaxsni aniqlash vazifasi - maqsadli suxondon tizimga sinov nutqlari bilan birga taqdim etilganda, maqsadli suxondonning nutqlarini to'g'ri belgilashdir. [63].
Shaxsni segmentatsiyasi	Shaxsni segmentatsiyasining vazifasi tizimga audio oqim taqdim etilganda shaxsning o'zgaruvchan nutqalarni topishdir
Shaxslarni klasterlash	Shaxslarni klasterlash vazifasi tizimga taqdim etilgan ko'plab so'zlarni to'g'ri klasterlashdir va odatda vazifa onlayn tarzda amalga oshiriladi. [65].
Shaxsni diarizatsiyalash	Shaxsni diarizatsiyalash vazifasi audioni avtomatik ravishda shaxs ovoz segmentlariga bo'lish va qaysi segmentlar bir xil shaxs tomonidan aytilishini aniqlashdir. [65].

So'zlovchini tanib olish tizimlari matnga bog'liq, matnga bog'liq bo'lmagan, ochiq va yopiq to'plamli turlarga ajratiladi [57, 66; 1431–4] (1.5-rasm). So'zlovchini tanib olish masalasi o'ta murakkab masala bo'lib, tanib olish modellarini o'qitish va testlashda sog'lig'idagi muammolar, gapirish tezligi, muhit kabi ko'plab salbiy omillar ta'sir ko'rsatadi [67; 1278–83, 68].



1.5-rasm. So'zlovchini tanib olish tizimlari turlari

So'zlovchini tanib olish muammolari. Nutqni tanib olish inson uchun sodda masala bo'lib tuyulsada, apparat-dasturiy vositalar orqali uni tanib olish o'ta murakkab masala hisoblanadi. Masalan, telefon orqali suhbatlar matnini

shakllantirish, soʻzlovchini tanib olish kabi masalalarni turli omillar sabab toʻliq hal etishni imkoni yoʻq. Umuman olganda soʻzlovchini tanib olishning koʻplab muammolari mavjud. Ana shunday murakkab muammolardan biri bu maʼlumotlar yetishmasligi va oʻzgarishi hisoblanadi [69; 14-17-b]. Ushbu ikki muammo safiga [46] ishda tanib olish sifatiga jiddiy salbiy taʼsir koʻrsatuvchi fon xalaqitlari ham qoʻshilgan. Bunda muammolar soʻzlovchi va texnik xatoliklar bilan bogʻliq boʻlishi ham mumkin [44; 23-34-b]. Quyida mazkur holatlarni batafsil bayoni keltirilgan.

Oʻzgaruvchanlik muammosi. Ovozni oʻzgaruvchanligi soʻzlovchini tanib olishning har qanday tizimi uchun muammo hisoblanadi. Oʻzgaruvchanlik ichki va oʻzaro turlariga ajaratiladi [57]. Ichki oʻzgaruvchanlik his-hayajon, jumla tezligi, kasallik, hatti-harakat, urgʻu, kayfiyat kabi turli omillar asosida yuzaga keladi [70; 1-10-b]. Oʻzaro oʻzgaruvchanlik esa nutq organlarini anatomik jihatdan oʻzaro farqlari natijasida yuzaga keladi [57, 70; 1-10-b]. Nisbatan keskin oʻzgarishlar oʻqitish va testlash jarayonida yuzaga keladi, yaʼni oʻqitishda soʻzlovchi toza muhitda gapiradi [57]. Biroq, testlash jarayonida soʻzlovchi xalaqitli muhitda gapiradi. Shuning uchun [70; 12-18-b] ish mualliflari soʻzlovchi ovozidagi oʻzgarishlarni oldini olish maqsadida ovozli maʼlumotlarni maʼlum bir vaqt intervallarida yozishni taklif etishgan. Bundan tashqari, ovozli maʼlumotlar uzatishni turli kanallarida ham oʻzgarishi mumkin. Masalan, turli xildagi mikrofonlardan foydalanishda ovozli maʼlumotlar oʻzgaradi. Bu esa modelni xosmas holatda ham ishlashi zarurligini bildiradi.

Maʼlumotlarni yetishmaslik muammosi. Aniq yechimlarni olish uchun modellarni oʻqitish uchun maʼlumotlarni yetarli boʻlmasligi maʼlumotlarni yetishmasligi hisoblanadi. Bu esa jiddiy va keng tarqalgan muammolarni yuzaga keltiradi. Koʻplab ilovalar kam sonli roʻyxatdan oʻtish orqali kichik hajmli oʻquv tanlanmalari asosida ishlashni talab qiladi [68; 1-9-b]. Test maʼlumotlari taqsimoti soʻzlovchi modeliga mutanosib boʻlishi zarurligi [70; 12-18-b] ish mualliflari tomonidan taʼkidlab oʻtilgan. Soʻzlovchi modeli yetarli darajada ovozli maʼlumotlar bilan toʻliq oʻqitilsa va test namunalari soni yetarli boʻlsa tanib olish tizimi kutilgan natijalarni taʼminlaydi. Singh tizim samaradorligi oʻqitiladigan tanlanma hajmiga

bog‘liq ekanligini alohida ta’kidlab o‘tgan [64; 1651-1665-b]. Biroq, o‘quv tanlama hajmini kattaligi kutilgan natijani ta’minlaydi degan mulohazani isboti mavjud emas [57]. O‘quv tanlanma hajmini yetarli bo‘lishi zarur, biroq o‘qitish uchun uzun jummalarni yig‘ish va so‘zlovchini uzoq vaqt davomida gapirtirish ham mumkin emas [70; 1-10-b]. Bu esa so‘zlovchini tanib olish yo‘nalishi bo‘yicha tadqiqotlarni davom ettirish zarurligini ko‘rsatadi.

Fon xalaqiti muammosi. So‘zlovchini tanib olishdagi yana bir asosiy muammo fon xalaqiti hisoblanadi [57]. Odatda o‘qitishda xalaqitsiz ma’lumotlardan keng foydalanilsa, testlashda esa xalaqitli ma’lumotlar qo‘llaniladi. Fon xalaqiti tanib olish aniqligiga kuchli ta’sir ko‘rsatuvchi omil bo‘lib, toza ovozli ma’lumot uchun aniqlik yuqori, biroq xalaqitli uchun aniqlik past bo‘ladi [57]. Xalaqitlarga ega bo‘lgan ovozli ma’lumotlar so‘zlovchini modelini yaratishda muammo hisoblanadi va u tanib olish aniqligiga ham kuchli ta’sir ko‘rsatadi.

Umuman olganda so‘zlovchini tanib olishdagi yana bir jiddiy muammo hujumlar hisoblanadi. Hujumlar ikki turda, ya’ni maqsadli va maqsadsiz amalga oshirilishi mumkin. Ovozli ma’lumotlarga qilinadigan hujumlar bayoni [71, 72; 49–64-b] ishlarda batafsil keltirib o‘tilgan. So‘zlovchini tanib olish tizimlarini rivojlanishiga qaramay ayrim muammolar ularni amaliyotga joriy etilishiga kuchli ta’sir ko‘rsatmoqda [73; 1-12-b]. Biometriya texnologiyalari nuqtai nazaridan ba’zi bir jiddiy muammolar 1.4-jadvalda keltirib o‘tilgan.

1.4- jadval

Biometriya texnologiyalari nuqtai nazaridan ba’zi bir jiddiy muammolar

Biometrik tizimlar	Afzalliklari	Kamchiliklar
Qo‘lga asoslangan: Barmoq izi yoki barmoq skanerlash Qo‘l geometriyasi	Foydalanish oson va narxi arzonligi. Yuqori aniqlik va uzoq muddatli barqarorlik. Turli barmoqlarni ro‘yxatga olish qobiliyati.	Kimyoviy ishlab chiqarish kompaniyalarida barmoq izlari tez-tez ta’sirlanadi. O‘xshash holatga ta’sir qiladi. Qisqa muddatli yoki doimiy ravishda bir nechta kesilgan, pufakchalar yoki jarohatlardan ta’sirlangan.
Face-Eye	Yashirin faoliyat yuritish	Egizaklarning yuz

asoslangan yuz asosida tanib olish ko'z qorachig'ini skanerlash Yuz termografi	qobiliyati. Intruziv bo'lmagan. Atrof-muhit holatiga ta'sir qilmaydi.	xususiyatlari o'xshash bo'lishi mumkin. Qandli diabet bilan og'rigan odamlarning ko'zlari ta'sir qilishi mumkin, shuning uchun farqlar inson yoshi ulg'aygan sari o'zgarishi mumkin.
Xulq-atvor xususiyatlari Ovozni aniqlash Imzoni tanib olish Klaviatura zarbasini skanerlash Shaxsni tanib olish, Yurishni aniqlash	U joriy telefoniya sozlamalaridan foydalanadi. Foydalanish oson va narx arzonligi. Ovozli biometrika quyidagi afzalliklarga ega: Qo'lga olish uchun u odam bilan aloqada bo'lishi shart emas, shuning uchun uni masofadan turib olish mumkin. Uni haydash paytida, boshqa xonadan, mobil qurilma orqali, odam boshqa faoliyat bilan band bo'lganda va hokazolarda ishlatish mumkin. Bu oddiy va ishlatish uchun qulay Yozib olingan ovozlardan ushbu tizimni aldash uchun ishlatib bo'lmaydi.	Imzoga nomuvofiqlik ta'sir qiladi, bu esa foydalanishni qiyinlashtiradi Ovoz sovuqdan ta'sirlanadi Atrof-muhit shovqini. Yoshi ulg'aygan sari u o'zgarishi mumkin. Xulosa va buzilish chastotasi javobi, kanal kodlash kabi kanal effektlari Taqdimot harakatlari nutq namunasining davomiyligi, apeketning psixofiziologik holati (kasallik, his-tuyg'ular kabi), vokal zo'riqishining ta'siri va boshqalar bilan bog'liq.

So'zlovchini tanib olishdagi mavjud muammolarni bo'lishiga qaramay, unga asoslangan tizimlar ko'plab amaliy vazifalarni bajarishda qo'llanilmoqda va ular bir necha o'n yillardan buyon mavjud. Mavjud tizimlar samaradorligini yanada oshirish uchun yuqorida keltirib o'tilgan muammolarni bartaraf etish zarur.

II. BOB. NUTQ SIGNALLARIGA DASTLABKI ISHLOV BERISH USUL VA ALGORITMLARI

Mazkur bob nutq signallariga dastlabki ishlov berish usul va algoritmlarini ishlab chiqishga bag'ishlangan bo'lib, unda nutq signallarini normallashtirish, nutqli qismlarni aniqlash, nutq signalining foydali va to'xtalish qismini ajratish, xalaqitlardan tozalash, diskretlash hamda tanib olishda qo'llaniladigan belgilarni shakllantirish usul va algoritmlari keltirib o'tilgan.

2.1-§. Nutq signallarini normallashtirish algoritmlari

Inson ovozi signali amplitudasiga ovozining balandligi, mikrofondacha bo'lgan masofa kabi ko'plab tashqi va ichki omillar ta'sir ko'rsatadi. Buning natijasida signalda turli o'zgarishlar kelib chiqadi. Bu kabi omillar nutq signali ovoz holatiga ta'sir ko'rsatishi va bu ta'sir ovozni pastlashishi yoki balandlashishida o'z ta'sirini namoyon qiladi. Bu holat asosan ovoz yozish apparatlari bir jinsli bo'lmaganda kuchli seziladi. Ovoz baland bo'lganda o'zgarishlarni kamaytirish amplituda bo'yicha normallashtirish asosida amalga oshiriladi, ya'ni signal amplitudasi $\frac{d}{k}$, $\frac{d}{k}$ oraliqqa keltiriladi (odatda $k=2$ deb olinadi) va $\bar{s}[n]$ signal quyidagi formula asosida qayta shakllantiriladi:

$$\bar{s}[n] = \frac{d \times s[n]}{\max_i |s[i]| + 1} \quad (2.1)$$

bu yerda d - normallashtirish uzunligi va u absissa o'qiga nisbatan simmetrik.

Jumla yoki gap talaffuzi ovoz balandligi o'zgarishini baholashda K ta to'plamdan foydalaniladi. Bunda j - chi namuna uchun $M(j)$, balandlikni o'rtacha qiymati N namunalar hajmi va K namuna uchun $M(j)$ o'rtacha qiymat quyidagi formulalar asosida hisoblanadi:

$$M(j) = \sqrt{\frac{1}{N} \sum_{n=1}^N s^2[n]}, \quad j = 1, \dots, K; \quad (2.2)$$

$$M_K = \frac{1}{K+1} \sum_{j=1}^K M(j). \quad (2.3)$$

Namunalar balandligi va o'rtacha qiymat nisbiy farqi (2.4)-formula orqali hisoblanadi:

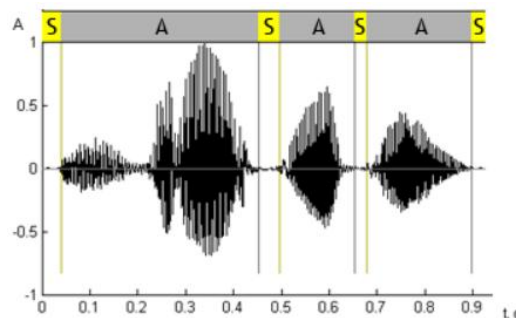
$$D(j) = \frac{|M_K - M(j)|}{M_K}. \quad (2.4)$$

Natijaga namunaning absolyut qiymati va ularning umumiy soni ham ta'sir qiladi. Bir sinfga tegishli bo'lgan namunalar to'plami balandligini o'zgarishi, uni uzunligi taqriban mos tushishini hamda bazadagi ovoz balandligi o'zgarishini baholanishi talab etiladi.

Nutqli sohalarni ajratish. Aksariyat holatlarda nutq bilan bog'liq tizim uchun kiruvchi signallar tabiiy xalaqitlarga boy bo'ladi, bunga mikrofon xalaqiti, nafas takti va boshqa omillar misol bo'ladi. Ularni bartaraf etish, tanib olish sifati va tezkorligini oshirishda dastlab nutq signalidan nutqli sohalar ajratib olinadi.

Yozuvda ovozli ma'lumot joylashgan oraliq nutqning faol sohalari deb ataladi va barcha boshqa sohalari esa fon xalaqitini bo'lishidan qat'iy nazar to'xtash qismlari sifatida e'tirof etiladi.

Adabiyotlarda nutq faolligi vaqt chegaralarini aniqlash algoritmi VAD (Voice Activity Detection) deb ataladi. VAD algoritmi signalni ikkita faollik A (activity) va S (silence) to'xtash segmentga ajratish imkonini beradi (2.1-rasm).



2.1-rasm. VAD segmentatsiyasi

Qisqa muddatli energiya va noldan o'tish chastotasi. $s(n)$ signaldan nutqni ajratishning mavjud usullari vaqtli sohada ishlaydi [74; 62-63-b]. Bularga qisqa muddatli energiya $E_s(m)$, qisqa muddatli quvvat $P_s(m)$ va noldan o'tish chastotasi

$Z_s(m)$ kiradi:

$$E_s(m) = \overset{m}{\underset{n=m-L+1}{\overset{\circ}{\mathbf{a}}}} s^2(n) \quad (2.5)$$

$$P_s(m) = \frac{1}{L} \overset{m}{\underset{n=m-L+1}{\overset{\circ}{\mathbf{a}}}} s^2(n) \quad (2.6)$$

$$y(x, y) = \begin{cases} 1, & x \neq y \neq 0; \\ 0, & \text{aks holda} \end{cases}$$

$$Z_s(m) = \frac{1}{L} \overset{m}{\underset{n=m-L+1}{\overset{\circ}{\mathbf{a}}}} y(s(n), s(n-1)) \quad (2.7)$$

Mazkur xarakteristikalar $s(n)$ signal L uzunligi m - qismida hisoblanadi. Quvvat olingan qism uzunligiga bog‘liq bo‘lgan signal energiyasiga ekvivalent. Hal qiluvchi qoida VAD (Voice Activation Detection) bo‘lib, unda yuqoridagi barcha xususiyatlar inobatga olinadi [74; 62-63-b]:

$$VAD(m) = \begin{cases} 1, & W_s(m) \geq t_w \\ 0, & W_s(m) < t_w \end{cases} \quad (2.8)$$

$$W_s(m) = P_s(m) \times (1 - Z_s(m)) \times S_c \quad (2.9)$$

$$t_w = m_w + a d_w \quad (2.10)$$

$$a = 0,25 \times d_w^{0.8} \quad (2.11)$$

$Vad(m)$ funksiyasida bo‘lag‘a qiymati t_w bo‘lib, u m_w xalaqit namunalari o‘rtacha qiymati va d_w - xalaqit namunalari qiymatlarini o‘rtacha qiymatidan og‘ishidan tashkil topgan bo‘lib, unda masshtablash koeffitsiyentlari tajribalar asosida aniqlanadi masalan, $S_w = 100$.

VAD algoritmini qo‘llashga doir misollaridan biri bu GSM standartidagi uyali radioaloqa tizimidir (Global System for Mobile Communications, global mobil aloqa tizimi) [75; 45-61-b]. Nutqga ishlov berish DTX (Discontinuous Transmission) uzlukli uzatish tamoyiliga muvofiq foydalanuvchi faol nutq davomiyligida uzatgichni yoqish va uni to‘xtash vaqtida yoki suhbat so‘ngida o‘chirish imkonini beradi. DTX VAD detektori tomonidan boshqariladi, bu hatto

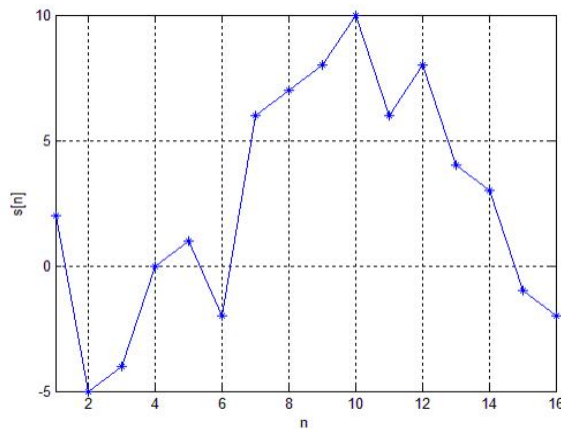
SHM 0dB dan past bo'lsa ham, nutqli va nutqsiz oraliqlarini aniqlash va ajratishda undan foydalaniladi. GSM tizimlarida VAD detektor iste'molni qisqartirishda hal qiluvchi rol o'ynaydi, chunki monologda so'zlovchining o'rtacha nutq faolligi 50% dan past va dialogda ishtirokchining faolligi suhbat vaqtining 30% igacha pasayishi mumkin [76; 53-58-b].

Lokal ekstremumlar tarqalishi asosida nutqli sohalarni ajratish usuli. Nutq signalining 2.2-rasmda keltirilgan qismida 2, 5, 6, 10, 11, 12 raqamli ekstremumlar mavjud. Nutq signali qismdagi ekstremumlarni umumiy soni M bilan, ekstremumni maksimal qiymati esa e bilan va mos namunalar razryadi k orqali belgilanadi:

$$e \hat{=} \left[\frac{2^{k-1}, 2^{k-1} - 1}{2^k} \right] \quad (2.12)$$

Masalan, nutq $k = 16, 2^{k-1} = 32768$ razryad bilan raqamli ko'rinishga o'tkazilsa, u holda E_j orqali ekstremumlarni j -oralig'ini belgilab olinadi va bunda oraliqlar uzunligi $L = e_{jmax} - e_{jmin}$ ga teng bo'ladi.

$$E_j = \left[e_{jmin}, e_{jmax} \right], \quad j = 0, \frac{1}{4}, \frac{2^k}{L} \quad (2.13)$$



2.2-rasm. Diskret nutq signali olti ekstremumi

E_j oraliqda ekstremumlar soni D_j ga teng va u quyidagi qoidalar yordamida aniqlanadi. Bunda dastlab $i = 0$ deb olinadi.

Agar $(s(n) < s(n-1) \text{ va } s(n) < s(n+1))$ yoki $(s(n) > s(n-1) \text{ va } s(n) > s(n+1))$ bo'lsa, u holda $i = i + 1$ va $g_i = 1$ aks holda $i = i + 1$ va $g_i = 0$.

$$D_j = \underset{i=1}{\overset{M}{\mathbf{a}}} g_i, j = 0, \frac{1}{4}, \frac{2^k}{L} \quad (2.14)$$

Nutq signalining berilgan har bir qismi uchun $P = (P_0, \dots, P_j, \dots, P_{\frac{2^k}{L}})$ ekstremumlar taqsimoti vektori shakllantiriladi va unda P_j ekstremum qiymatini E_j oraliqga tushish ehtimolligi va $\underset{j}{\mathbf{a}} P_j = 1$ bo'ladi. P vektor tanib olish tizimiga kiritilishi mumkin bo'lgan belgilar vektoridir. U sof ton uchun ekstremumlar taqsimotida aniq simmetrik cho'qqilli ko'rinishida va oq xalaqitli signalda esa taqsimot asimmetrik va cho'qqilarsiz tekis bo'ladi. Nutq signaliga nisbatan yuqori chastotali obertonlar qo'shilsa, taqsimot oraliqlari 0ga teng bo'lmagan qiymatlar orqali to'ldiriladi, bunda $\frac{2^k}{L} / 2$ ga nisbatan simmetriya saqlanib qoladi. Mazkur ta'sirni yo'qotish uchun past amplitudali tebranishlarni silliqilashtiruvchi va obertonlarni betaraf etadigan signalni to'liq qamrab oluvchi egri chiziq quriladi [77; 24-30-b, 78], ya'ni e^{max} va e^{min} signalni qamrab oladigan ikki egri chiziq quriladi va ular faqat maksimal va minimal qiymatni o'z ichiga oladi. Chiziq uzunligi asl signal uzunligiga teng bo'ladi va u N bilan belgilanadi va bunda $M < N$ bo'lib, ekstremumlar orasidagi namunalar qiymatini chiziqli interpolyasiya orqali aniqlash mumkin. So'ng o'rtacha signal s hosil qilinib, har bir namuna signal s_i boshlang'ich signal maksimal va minimal qiymati o'rtasiga to'g'ri keladi:

$$s_i = \frac{(e_i^{max} + e_i^{min})}{2}, i = 0, \dots, N. \quad (2.15)$$

Ekstremumlar taqsimotini yuqorida keltirilgan xossaligidan foydalanib dastlabki signalidan nutq signalini ajratib olish algoritmi taklif etildi va u quyidagi qadamlarda ishlaydi:

1-qadam. Normallashtirish amalga oshiriladi. Bunda mikrofon orqali kiritilgan signal amplitudasi 10-25 ms kenglikdagi sirpanuvchi oyna orqali normallashtiriladi. Ushbu jarayon ovoz va qurilmaning individual xususiyatlariga ko'ra signal balandligi o'zgaruvchanligini bartaraf etish uchun zarur bo'lib, unda

barcha qismlar bir xilda ovoz balandligi qiymatiga ko'tarilishi uchun oyna ichida normallashtirish amalga oshiriladi.

2-qadam. Signalni qamrab oluvchi egri chiziq quriladi. Bunda past amplitudali tebranishlar tekislanadi.

3-qadam. Normallashtirish amalga oshiriladi. Bunda qamrab oluvchi egri chiziq qurilishi sababli signalni takroriy normallashtirish amalga oshiriladi.

4-qadam. signalning m -qism uchun ekstremumlar taqsimotini aniqlash amalga oshiriladi.

5-qadam. Hal qiluvchi mantiqiy qoidalar qo'llaniladi. Bunda signal m - qismi nutq yoki xalaqitga ega ekanligi bo'yicha qaror qabul qilinadi. Hal qiluvchi mantiq sifatida 160-240 ta kiruvchisi bo'lgan uch qatlamli to'liq bog'langan neyron tarmog'i (NT) tanlandi.

Ko'p qatlamli neyron tarmog'idan foydalanish shartlarini quyidagilardan iborat:

1) Bir o'lchovli $W_s(m)$ kattalikni P vektor bilan almashtirish bo'sag'a funksiyasidan qutilish va tanlangan belgilar fazosida tasniflagichni qo'llashga majbur qiladi. NT - bu tasniflash muammolarini hal qilish uchun ishlab chiqilgan juda yaxshi vositadir.

2) Timsollarni ishonchli ajratish chiziqli bo'lmagan ajratuvchi funksiyalarni qurish qobiliyati talab qiladi. Ko'p qatlamli NT ushbu masalani to'liq hal eta oladi [78].

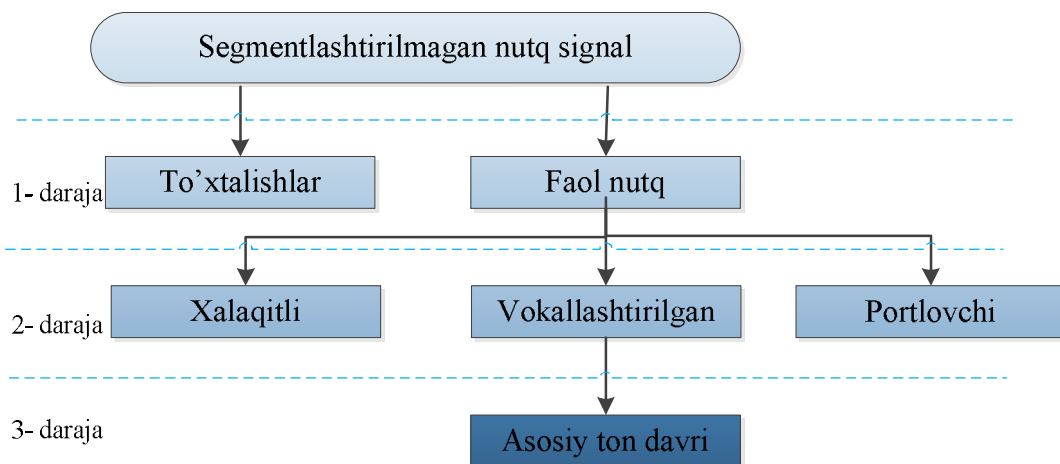
3) NT ham dasturiy ham apparat vositalarida mavjud. Mazkur ishda dasturiy vosita ishlab chiqilgan bo'lib, uni samarali apparat vositasini ishlab chiqishda ham foydalanish mumkin. Algoritmni o'qitish uchun barcha frikativ undosh tovushlarni inobatga olmaganda, tasniflash jarayonida oq xalaqitdan ajratish juda qiyin bo'lganligi uchun tizim duch keladigan nutq tovushlaridan foydalanish tavsiya etildi.

Tajriba-sinovlar turli xalaqitlarni nutqdan ajratishga o'qitish taklif qilingan tizim xalaqitlar mavjud bo'lganda yaxshi ishlashini ko'rsatdi. Biroq, ikkala yondashuv ham frikativ tovushlarni o'tkazib yuboradi.

2.2-§. Segmentlash algoritmlarini ishlab chiqish

Foydali nutq chegaralarini aniqlash masalasini yechish uchun VAD algoritmi foydali nutq signal uchun xarakterli bo'lgan belgilar mavjudligini aniqlashi shart. Bunday oddiy belgilardan biri signal energiyasi hisoblanadi va ideal holat uchun to'xtalishlar bo'limida signal kuchi nolga teng bo'ladi. Biroq xalaqit fonining mavjudligi VAD segmentlashi uchun signal energiyasidan parametr sifatida foydalanishni murakkablashtiradi va yanada murakkabroq algoritmlardan foydalanishni talab qiladi.

Nutq signal tuzilishining ma'lumotlariga asoslanib, o'zbek tilidagi tovushlarni amalga oshirilishini o'ziga xos xususiyatlaridan kelib chiqqan holda fonogrammalarni vaqt bo'yicha segmentlash masalasi ketma-ket amalga oshiriladigan bir nechta bosqichga bo'linishi mumkin. Ularning har biriga tegishli segmentlash darajasi topiladi (2.3-rasm).

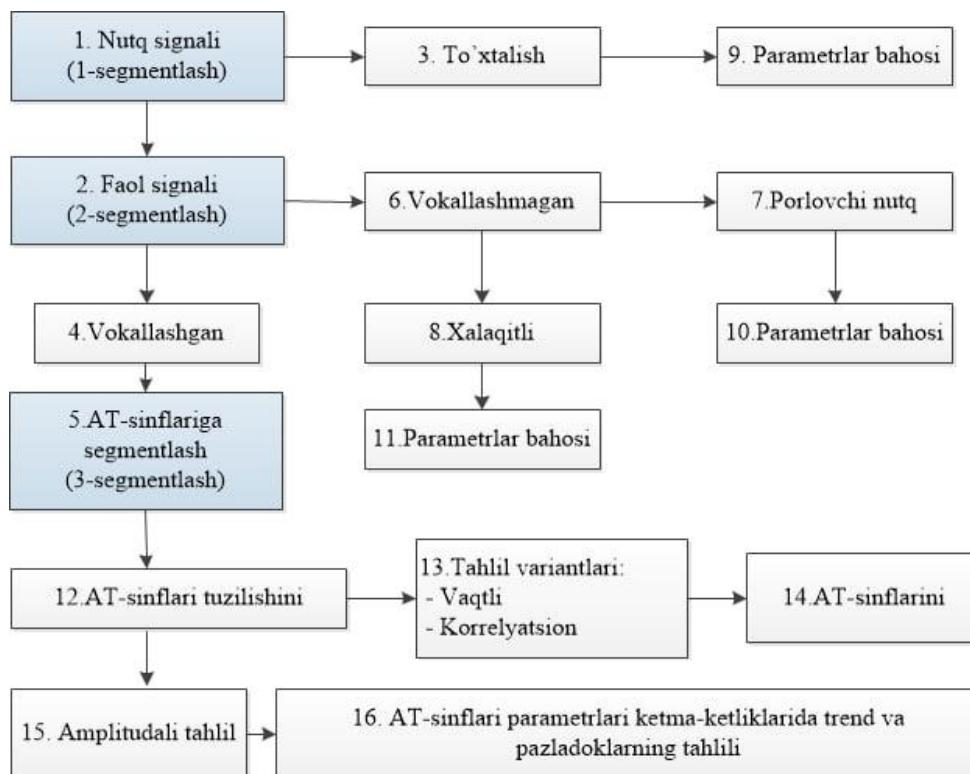


2.3-rasm. Vaqt bo'yicha nutq signalini segmentlashni darajalari

Segmentlashning birinchi darajasida nutq signalini foydali va to'xtalishlar qismiga ajratish amalga oshiriladi (VAD algoritmi qo'llaniladi). Bunda to'xtalishlar sifatida birlashmalarni hisobga olish mumkin. Segmentlashning ikkinchi darajasida nutqning foydali qismi uchta asosiy tovushga mos keladigan qismlarga ajratiladi: vokallashtirilgan, xalaqitli va portlovchi. Uchinchi darajali segmentatsiya faqat vokallashtirilgan qismlarga tegishli bo'lib, ularni quyi segmentlash davrlarga

ajratishdan iborat. Davrlar bo‘yicha bo‘linish segmentlashning yakuniy bosqichi emasligini alohida ta’kidlash joiz: keyinchalik vokallashtirilgan fragmentlarni bo‘linishi asosida bir xil tipli allofonlar guruhidan turlicha allofonlarga ajratilishi mumkin. Mazkur holat turli xil fonemalar uchun neyron tarmoqlar, Markov modellari va etalonlarni shakllantirib olish imkonini beradi.

Umumlashtirgan segmentatsiya algoritmining tarkibi. 2.4-rasmda segmentlash hamda nutq signalini parametrlashtiruvchi tizimni qurish yondashuvi keltirilgan. Birinchi bosqich ishlov berishda (1-3-modullari) VAD algoritmidan foydalaniladi. To‘xtalishlar segmentlarining asosiy informativ parametrlari (9-modul) ularni davomiyligi hisoblanadi.



2.4-rasm. Tizimli yondashuv funksional sxemasi

Foydali yoki faol nutq qismlari ustida segmentlashning ikkinchi darajasi qo‘llanilib, unda tovush turlariga mos keladigan segmentlar ajratiladi: vokal (4-bo‘lim), xalaqitli (8-modul), portlovchi (7-modul). Turli xil tipdagi nutq segmentlarini parametrlashda har xil parametrlar to‘plamidan foydalanish mumkin. Nutq signal fragmentlarini yagona almashtirishi bir xil turdagi parametrlarini hosil qilishda qo‘llanilishi mumkin. Masalan, nutq signalini spektral almashtirishda,

korrelasion funksiyalarni shakllantirishda, uni parametrlash va boshqalarda.

Quyidagi parametrlash, trenlar va buzilishlarni tahlili orqali vokallashtirilgan fragmentlar asosiy ton davrlari uchun qo‘shimcha segmentlashdan o‘tkaziladi (5, 12, 14 ... 16- modullar).

Bir tipli segmentlash algoritmlarini taqqoslash usuli. Segmentlash algoritmlarini ishlab chiqishdagi asosiy savol bu- ishlab chiqilgan algortm mavjud bo‘lgan algoritmlarga qaraganda samaraliroqmi? Ushbu savolga javob, ishlab chiqilgan algoritmni nutq tizimlarida qo‘llash maqsadga muvofiqligini aniqlaydi. Quyida bir tipli nutq signalini segmentlash algoritmlari aniqligini taqqoslash uchun tizimli yondashuv keltirilgan. Nutq signalining har bir kichik fragmenti uchun parametrlashtirish natijalariga ko‘ra, ma’lum bir parametrlning miqdoriy bahosi hisoblanadi. Bunda parametrlashtirishning asosiy usullaridan quyidagilarni keltirish mumkin: chiziqli bashoratlash, keprstral tahlil, veyvlet-almashtirish, modulyasiya spektrini tahlil qilish [79; 3023–3035-b]. Shunday qilib, parametrlar vektori, signalning vaqt fragmentiga mos keladigan har bir muhim bahoni sonli to‘plamlar bahosi ko‘rsatadi.

Segmentlash masalasi ma’lum umumiy xususiyatlarga ega bo‘lgan signal fragmentlari o‘rtasida vaqt chegaralarini taqsimlash hisoblanadi. Segmentlash natijasida ajratish chegaralarini vaqtli belgilarini tavsifisiz olish mumkin yoki cheklangan segment tipini ko‘rsatuvchi vaqtli belgilarini, ya’ni sifatini tavsifiya qilish orqali ham olish mumkin (masalan, VAD algoritmi). Parametrlashtirish va segmentlash jarayoni bir–biri bilan uzviy bog‘liqdir. Signalni segmentlash baholari oldindan ma’lum bo‘lgan parametrlar asosida amalga oshiriladi va ushbu parametrlar uchun bo‘lag‘a qiymatlari beriladi va bo‘lag‘adan o‘tuvchi parametrlarni vaqtga bog‘liq funksiyasidan nutq signalini ajratishda foydalaniladi. Masalan, VAD algoritmidagi kichik vaqt oralig‘ida o‘rtacha signal quvvati parametrini olish mumkin [80; 58-70-b]. Bu esa, keyinchalik segmentlash natijalariga asoslanib har xil turdagi segmentlarga mos parametrlar turli guruhlarini hisoblanadi.

Vokallashtirilgan segmentda suxandonni AT davriyligi parametrini olish maqsadga muvofiq. Bunda xalaqitli tovush va to‘xtalishlarda mazkur parametr ma’noga ega emas. Segmentlash va parametrlash jarayonini umumlashtirish maqsadida segment turini sifat jihatdan ifodalash va uning “sifatli parametrlar orqali segmentlash” orqali ko‘rish talab qilinadi. Bu holda parametr qiymati raqam bo‘lmaydi, balki segment turlari alifbosidagi ayrim ramziy kod bo‘ladi. Segmentlash aniqligi va ishonchliligini baholashda universal bo‘lgan yagona yondashuv mavjud emas, bundan tashqari, bu kabi yondashuvlarni yaratish nutq signalini segmentlash masalasini murakkablashishiga olib keladi.

VAD-algoritmi samaradorligi va boshqa algoritmlarni baholash maqsadida quyidagi parametrlar taklif etilgan: Front End Clipping (FEC) – to‘xtalishdan nutqga o‘tish chegarasini aniqlash xatoligi, Mid Speech Clipping (MSC) – nutqni to‘xtalish sifatida noto‘g‘ri aniqlash, OVER – nutqdan to‘xtalishga o‘tish chegarasini aniqlash xatoligi va NDS – to‘xtalishni nutq deb noto‘g‘ri olishini aniqlash. Shuni alohida ta’qidlash kerakki, mazkur tahlilda samaradorlik segment chegaralarini aniqlash xatoligi (FEC va OVER) va segment turini aniqlash xatoligi (MSC va NDS) orasidagi farqlar yordamida aniqlanadi. [81; 145-160-b] ishda VAD algoritmi aniqligini baholash uchun boshqa ko‘rlab parametrlar taklif etilgan, ya’ni bular HR0 va HR1 muvaffaqiyat koeffitsiyentlaridir:

$$HR0 = \frac{N_{0,0}}{N_0^{ref}}, HR1 = \frac{N_{1,1}}{N_1^{ref}}, \quad (2.16)$$

bu yerda $N_{0,0}$ va $N_{1,1}$ – to‘g‘ri aniqlangan to‘xtalish va nutq oralig‘i soni (bunda signal qisqa muddatli oyna orqali intervallarga ajratiladi). N_0^{ref} va N_1^{ref} – esa nutq etalonini segmentlashiga nisbatan to‘xtalish va nutq intervali soni.

Yuqorida aytib o‘tilgandek, segmentlash samaradorligini baholash masalasi nafaqat VAD-algoritmlar uchun, balki segmentlashning boshqacha ko‘rinishlarini amalga oshiruvchi algoritmlar uchun ham muhim hisoblanadi. Bunda segmentlash xatoligi, segmentlashning turlari va ularni nisbiy holatlariga ko‘ra turli xil baholanishi mumkin. Ayni paytda mazkur noqulayliklarni inobatga oluvchi

segmentlash aniqligi parametrlarini qurishga mo'ljallangan aniq bir yagona yondashuv mavjud emas.

Nutq signal so'navchi qismini aniqlash algoritmlari. Fonogrammalarga o'xshash alohida tovushlarni vaqt tarkibiy qismlaridan biri bu so'navchi signal yoki tovush xususiyatlari amplitudasini modullash funksiyasini (MF) inobatga olishi hisoblanadi.

Segmentlash natijalaridan foydalanib nutqga ishlov berish ko'plab amaliy masalalarida bir dona fundamental ton o'zgarishiga qadar segmentlash aniqligi muhimdir. Shuning uchun signalni filtrlash va tanib olishiga yo'naltirilgan so'navchi qismini ajratish algoritmlari, natijalarni yetarli emasligi ularni segmentlash masalalariga har doim ham qo'llash imkonini bermaydi. Shu sababli signalni so'navchi qismi qiymatlarini olish imkonini ta'minlaydigan algoritm ishlab chiqilgan bo'lib, uni barcha qiymatlari dastlabki signal qiymatlari o'zi bo'ladi. Mazkur yondashuv olingan so'navchi qismni qo'llash uchun ko'proq imkoniyatlarni beradi. Jumladan, so'zlovchini asosiy ton davrini maksimal darajada ko'rish orqali baholashga qodir.

Algoritm ma'lum bir sanoqlarda nutq signalini dastlabki sanoqlaridan foydalanadi va bu MF amplitudasini tashkil qiladi. Signal so'navchi qismini aniqlash algoritmi hisoblash murakkabligini kamaytirish maqsadida dastlabki qadamda nutq signali lokal maksimumlari to'plami M quyidagi qonun-qoidalar asosida shakllantiriladi, dastlab $M = \overline{AE}$ deb olish talab etiladi.

Agar $(s(i) > s(i-1))$ va $(s(i) > s(i+1))$ shart o'rinli bo'lsa, u holda $M = M \dot{\bar{E}} i, i = \overline{1, N}$, bu yerda i – nutq signali dastlabki sanoqlari soni, $s(i)$ – nutq signali son qiymatlari. Keyinchalik so'navchi qismni sanoq raqamlari iterativ ravishda M to'plam elementlaridan orasidan olinadi. Agar $s(n)$ da so'navchi qismlarning joriy sanoq sonlari sifati qabul qilinsa, u holda navbatdagi $s(n+1)$ so'navchi qism sanoqlari quyidagi formula orqali aniqlanadi:

$$\begin{aligned} \hat{K} &= \{m \mid M \mid n < m < n + L\} \\ \hat{s}(n+1) &\in \hat{K} \\ \hat{k} &\in \hat{K} \mid s(n+1) \neq s(k) \\ \hat{K} &= \{k \mid K \mid (s(n+1) = s(k) \wedge \dot{U}(n+1) < k)\} = \mathcal{A} \end{aligned} \quad (2.17)$$

Taklif etilayotgan nutq signali soʻnuvchi qismlarini chiqarib yuborish algoritmi quyidagi qadamlarda amalga oshiriladi.

1-qadam. Nutq fayli asosida f vektor shakllantiriladi;

2-qadam. ϵ ga qiymat oʻzlashtiriladi;

3-qadam. y ning qiymati hisoblanadi:

$$y = \begin{cases} 0, & f(i) < \epsilon, \\ 1, & f(i) \geq \epsilon. \end{cases}$$

4-qadam. f vektor oʻlchamini kamaytirish amalga oshiriladi, yaʼni $y=0$ boʻlsa, u holda unga mos element f vektordan chiqarib yuboriladi;

5-qadam. Tamom.

2.3-§. Xalaqitlardan tozalash algoritmlari

Suxandonni identifikatsiyalash toza akustik sharoitda ishonchli amalga oshirilishi mumkin, biroq fon xalaqitini mavjudligi identifikatsiyalash sifatiga salbiy taʼsir qiladi. Suxandonni identifikatsiyalash tizimlarini xalaqitga nisbatan barqarorligini oshirishda eng samarali usullaridan biri bu xalaqitni kamaytirish algoritmlari yordamida kirish signallariga dastlabki ishlov berish hisoblanadi [82; 137-144-b]. $y(t)$ xalaqitli nutq signal biror-bir F_q diskretizatsiya chastotasi bilan raqamlashtirishi mumkin: $y(t) \approx \sum_{n=0}^{N-1} y_n \delta(t - nT)$, bu yerda n -sanoqlar raqami. Chastotali sohada xalaqitni kamaytirish quyida keltirilgan qadamlar orqali amalga oshiriladi.

1-qadam. Nutq signali kiritiladi.

2-qadam. Har bir spektral namunaga mos xalaqit spektral zichligi \hat{f}_k^D baholanadi.

3-qadam. y_n ga belgilar sonini kamaytirish algoritmi qoʻllaniladi va

$M = T \times F_{\eta}$ ta namuna uzunligidagi bir-birini qisman qoplovchi oynalar yordamida ajratish amalga oshiriladi, T - stasionar nutq davri.

$$y_{m,w}^w = F_m^M y_{(w-1)O+m}, \quad m=1 \dots M, \quad w=1 \dots W, \quad W=1 + \text{fix} \frac{\text{æV ö}}{\text{ëO ÷}}$$

bu yerda w - oyna, m - esa oynadagi namuna raqami; N - signal namunalari soni; W - signal oynalari soni; $\text{fix}(x)$ - sonni butun qismi; O - oynani surilishi bo'lib u qoplanish darajasini aniqlaydi ($O = M$ bo'lganda qoplanish bo'lmaydi).

4-qadam. Quyidagi formula asosida DFA hisoblanadi.

$$Y_{k,w} = \mathring{\mathbf{a}}_{m=1}^M y_{m,w}^w \times \exp \frac{\text{æ}}{\text{ë}} j \frac{2p}{M} (k-1)(m-1) \frac{\text{ö}}{\text{÷}} \quad k\text{- spektral namuna raqami.}$$

5-qadam. Signal amplitudali spektrini quriladi, ya'ni: $R_{k,w} = |Y_{k,w}|$.

6-qadam. Har bir oyna va spektral namunasining aposterior signal xalaqit munosabat $\hat{g}_{k,w}$ lari baholanadi.

7-qadam. Har bir oyna va spektral namunasi aprior $\hat{x}_{k,w}$ signal xalaqit munosabatlari baholanadi.

8-qadam. Spektrni korrektsiyalash funksiyasini tanlash amalga oshiriladi hamda har bir oyna va namuna uchun uni qiymatlari hisoblanadi. $G(x_k, g_k)$.

9-qadam. $G(x_k, g_k)$ va $R_{k,w}$ elementlar bo'yicha ko'paytirish amalga oshirilib, toza signal spektri amplitudalari baholanadi:

$$\hat{A}_{k,w} = G(x_k, g_k) \times R_{k,w}.$$

10-qadam. Toza signal kompleks spektri hisoblanib baholanadi:

$$\hat{X}_{k,w} = \hat{A}_{k,w} \exp(j \times \arg(Y_{k,w})).$$

11-qadam. Teskari DFA amalga oshiriladi:

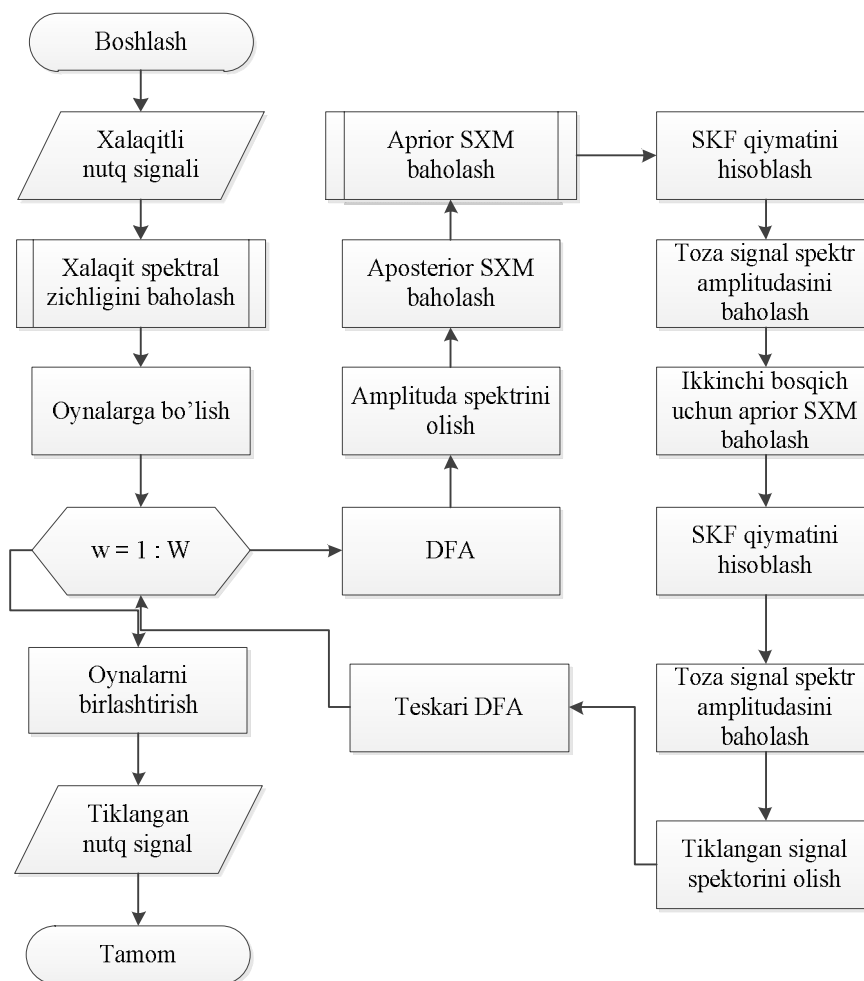
$$\hat{x}_{m,w}^w = \frac{1}{M} \mathring{\mathbf{a}}_{k=1}^M \hat{X}_{k,w} \times \exp \frac{\text{æ}}{\text{ë}} j \frac{2p}{M} (k-1)(m-1) \frac{\text{ö}}{\text{÷}}$$

12-qadam. Qoplanish hisobiga oynalarni birlashtirish amalga oshiriladi:

$$\hat{x}_{m,w}^w = \frac{3}{4} \frac{\text{æ}}{\text{ë}} \frac{\text{æ}}{\text{ë}} \frac{\text{æ}}{\text{ë}} \frac{\text{æ}}{\text{ë}} \frac{\text{æ}}{\text{ë}} \frac{\text{æ}}{\text{ë}} \frac{\text{æ}}{\text{ë}} \frac{\text{æ}}{\text{ë}} \frac{\text{æ}}{\text{ë}} \hat{x}_n.$$

13-qadam. Tamom.

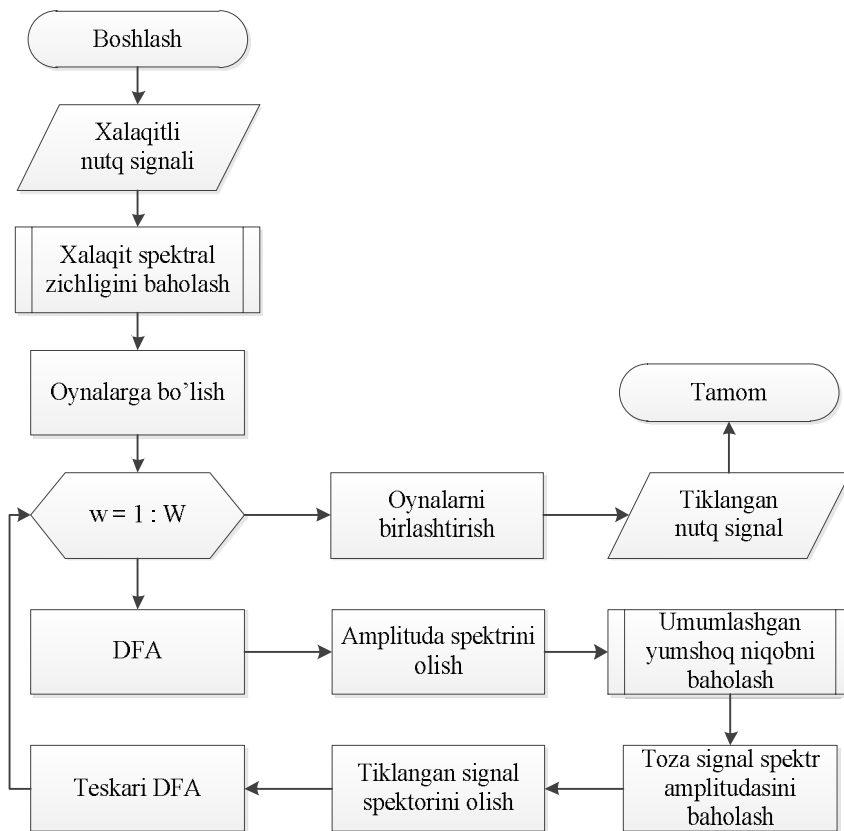
Mazkur ishda ikki bosqichli xalaqitni kamaytirish algoritmi [83; 39-44-b, 84; 153-156-b] ishlardagi kabi quyidagicha amalga oshirilgan (2.7-rasm).



2.7-rasm. Xalaqitni kamaytirish uchun ikki bosqichli algoritmi blok sxemasi

Binar niqoblar. Xalaqitlarni kamaytirishga yo'naltirilgan yana bir usul vaqt chastotasining binar niqoblariga asoslangan bo'lib, ularni qurishda additiv xalaqit signalini vaqt chastotasi individual qiymatlarini kuchli buzadi, boshqalari esa kamroq buzadi degan g'oyaga asoslanib qurilgan. Mazkur g'oya signallarni "xalaqitli" va "uncha xalaqitli bo'lmagan" turlarga ajratish imkonini beradi. Binar niqoblardan eshitish qiymatlarini tahlil qilish va hisoblashda keng foydalaniladi [85; 92-102-b, 86; 3501-3504-b].

Umumlashgan yumshoq niqob asosida xalaqitlarni kamaytirish algoritmi quyida blok sxema asosida amalga oshiriladi (2.8-rasm). Mavjud yondashuv asosida umumlashgan yumshoq niqobni baholash blok-sxemasi 2.9-rasmda keltirilgan.



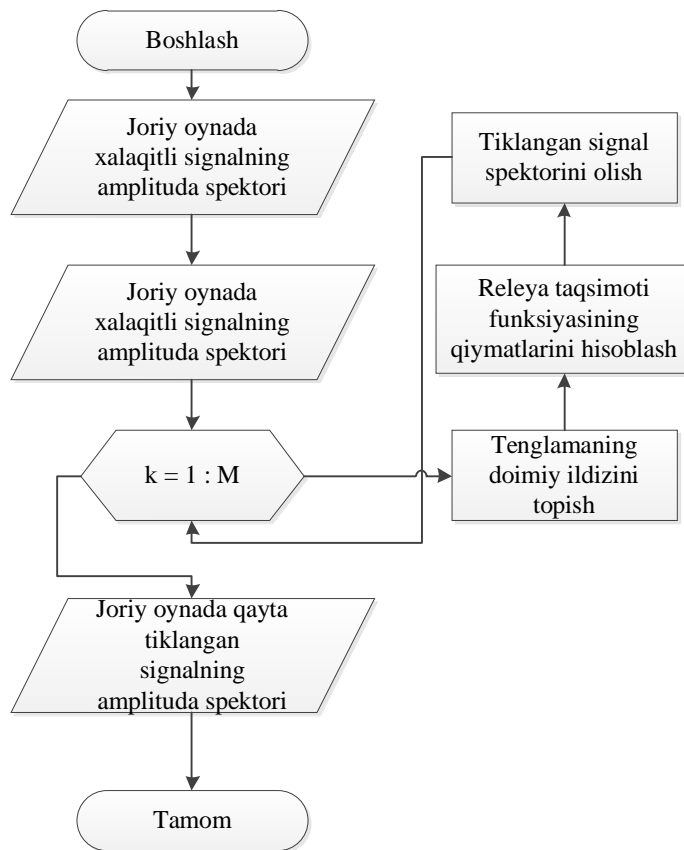
2.8-rasm. Umumlashgan yumshoq niqob asosida xalaqitlarni kamaytirish algoritmi blok-sxemasi

Yumshoq niqobga asoslangan ikki bosqichli algoritm. Yumshoq niqobli xalaqitni kamaytirishning takomillashtirilgan algoritmi ikki bosqichli xalaqitni kamaytirish algoritmini dastlabki qadamida amalga oshiriladi va uning blok-sxemasi 2.10-rasmda keltirilgan.

Ikkinchi bosqichda esa spektrni sozlash funksiyasi sifatida qisqa muddatli spektr amplitudasi minimal o‘rtacha kvadratik xatoligi spektrni korreksiyalash funksiyasini tanlash tavsiya qilinadi. bunda boshqa spektrni korreksiyalash funksiyasidan foydalanib yuqori natijalar olindi, bunda Viner va spektr qisqa muddatli amplitudasi logarifmi minimal o‘rtacha kvadratik xatoligidan foydalanildi.

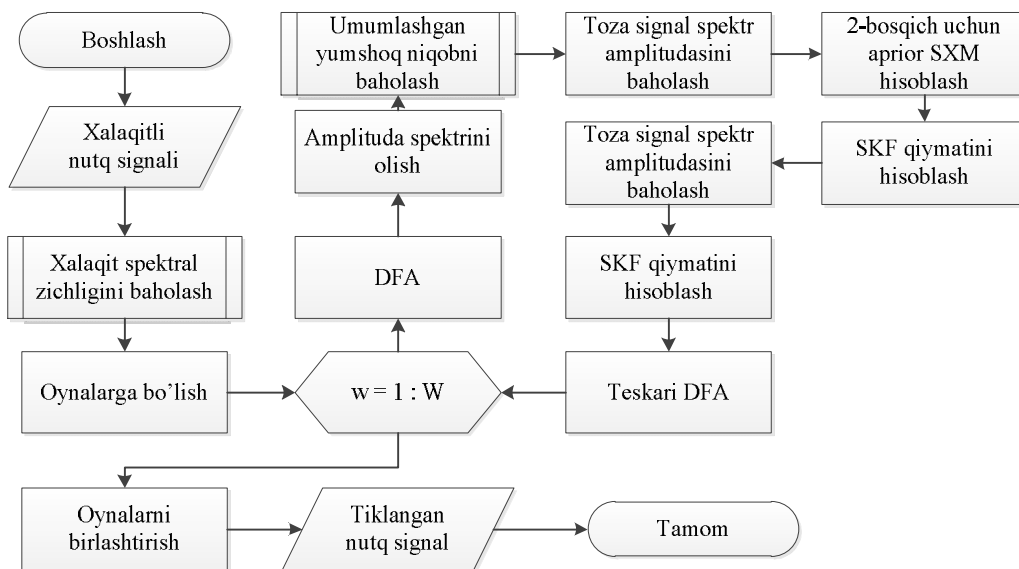
Nutq signali sifatini baholash. Nutqni eshitayotgan shaxs, uni sifat darajasini intuitiv baholashi mumkin. Shuning uchun mazkur ko‘rsatkich miqdor orqali baholashni talab qiladi va u nutqni tanib olishda dolzarb muammodir [87; 39-44-b]. Nutq signali sifati subyektiv va obyektiv usullar asosida baholanadi. Subyektiv ko‘rsatkichlar eshituvchilar tomonidan dastlabki nutq signali va qayta ishlangan

nutqni eshitish orqali solishtirishga asoslanadi. Har bir eshituvchi shaxsning bahosi albatta bir-biridan farq qiladi, shuning uchun nutq signali sifatini yetarlicha baholashda guruh o'rtacha bahosini olish kerak [88; 123-132-b].



2.9-rasm. Mavjud yondashuv asosida umumlashgan yumshoq niqobni

baholash blok-sxemasi



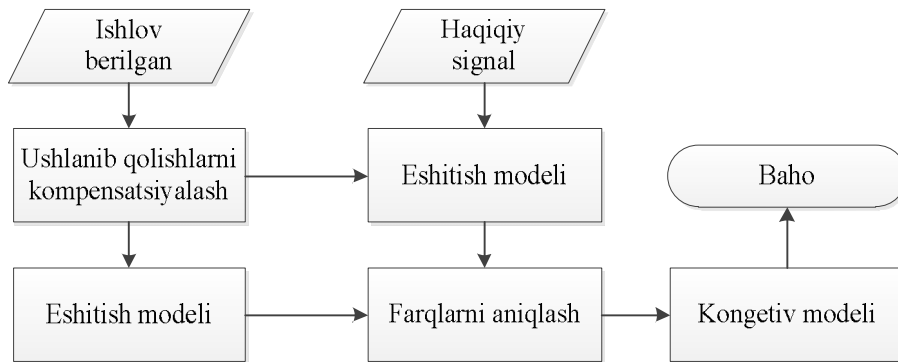
2.10-rasm. Yumshoq niqobli ikki bosqichli algoritim blok sxemasi

Obyektiv ko‘rsatkich boshlang‘ich va qayta ishlangan signal orasidagi qiymatlar asosida hisoblanadi. Bunda eng ko‘p foydalaniladigan sifat ko‘rsatkichi, ya‘ni signal/xalaqit munosabatidan foydalaniladi va u quyidagi foymula orqali hisoblanadi:

$$Q^{SSHM} = 10 \log_{10} \frac{\sum_{n=1}^N x_n^2}{\sum_{n=1}^N (y_n - x_n)^2} \quad (2.18)$$

bu yerda x_n - signal xalaqitsiz sanoqlari, y_n - signal xalaqitli sanoqlari.

Nutqni tinglash sifatini baholash. Nutqni tinglash sifatini baholash (PESQ, perceptual evaluation of speech quality) [89; 1123–1137-b] – bu MSE-T P.800 standartiga mos ekspert eshituvchilar tomonidan approksimatsiyalanuvchi sub’yektiv baholash MOS (mean opinion score) sifat ko‘rsatkichi hisoblanadi. PESQni yuzaga kelishi bilan, MOS [88] baholashda murakkab va aniq obyektiv usul sifatida ko‘riladi. Ayni paytda ham u nutq sifatini baholashning ko‘p foydalaniladigan obyektiv usullardan biridir, uni amalga oshirish blok sxemasi 2.11-rasmda keltirilgan.



2.11-rasm. PESQ algoritmi blok sxemasi

Nutq sifatini baholovchi PESQ dastlabki va qayta ishlangan nutq signalini almashtirishda eshitish modelidan foydalanishga asoslangan. Bunda qayta ishlashga ketadigan vaqtni qoplash maqsadida qayta ishlangan nutq signal berilgan nusxasiga moslashtirish amalga oshiriladi. Uning ichki ko‘rinishi bo‘yicha 2 signal farqi MOS orqali baholangan kognitiv modelga asoslanadi.

2.4-§. Diskretlash algoritmlari

Shaxsni ovozi bo'yicha avtomatik identifikatsiyalash tizimini qurishda kiruvchi nutq signali belgilar to'plamini shakllantirish va qayta ishlash, model parametrlarini shakllantirish hamda baholash algoritmini ishlab chiqish, komponentalar sonini aniqlash masalalarini yechish talab etiladi. Bunda dastlab nutq signalda analog-raqamli almashtirish bajariladi. Diskretlashda esa signal biror bir vaqt oralig'i bo'yicha kvant amplitudaning alohida qiymatlariga ajratiladi.

Shaxsni ovozi asosida identifikatsiyalashda nutq signali hajmi va unga sarflanadigan vaqtni kamaytirish muhim ahamiyat kasb etadi. Shu sababli, ushbu ishida nutq signali belgilari sonini kamaytirish hisobiga qayta ishlanadigan nutq signali hajmini va tizim ishlashi uchun sarflanadigan vaqtni kamaytirishga alohida e'tibor qaratilgan. Mazkur ishda nutq signallarini «Delta diskretlash» va «Maksimum-minimum diskretlash» algoritmlarini taklif etilgan bo'lib, ular qadamlar ko'rinishda quyida keltirilgan.

“Delta diskretlash” algoritmi A1 orqali belgilangan bo'lib, u quyidagi qadamlarda amalga oshiriladi:

1-qadam. Nutq signali qabul qilinadi.

2-qadam. Nutq signali kerakli formatga o'tkaziladi.

3-qadam. Nutq signalini normallashtirish amalga oshiriladi.

4-qadam. Yuqori va quyi nuqtalar to'plami shakllantiriladi, ya'ni X_{\max} va X_{\min} .

5-qadam. $Dx = \frac{\overline{X}_{\max} + \overline{X}_{\min}}{2}$ hisoblanadi.

6-qadam. Noldan o'tishlarni hosil qilish maqsadida $x_i = x_i - Dx, i = \overline{1, N}$ amalga oshiriladi.

7-qadam. Nutq signalidan quyidagi qonuniyat asosida belgilar vektori shakllantiriladi:

$$\hat{x}_i = \frac{x_{2i+1} - x_{2i-1}}{2}, x_i \hat{X}_{\max}$$

$$\hat{x}_i = \frac{x_{2i} - x_{2i+2}}{2}, x_i \hat{X}_{\min}$$

bu yerda, x - berilgan nutq signali sanoqlaridan iborat vektor, $i = 1, 2, \dots, N$.

8-qadam. Nutq signalidan belgilar to‘plami shakllantiriladi.

9-qadam. Shaxs modelini qurish va uning asosida shaxsni identifikatsiyalash amalga oshiriladi.

10-qadam. Tamom.

Maksimum-minimum diskretlash algoritmi quyidagi qadamlarda amalga oshiriladi:

1-qadam. A1 algoritmnining 1-6-qadamlari amalga oshiriladi.

2-qadam. Nutq signali L uzunlikdagi freymlarga ajratiladi.

3-qadam. Har bir freym uchun k ta maksimum va minimum elementlar tanlab olinadi.

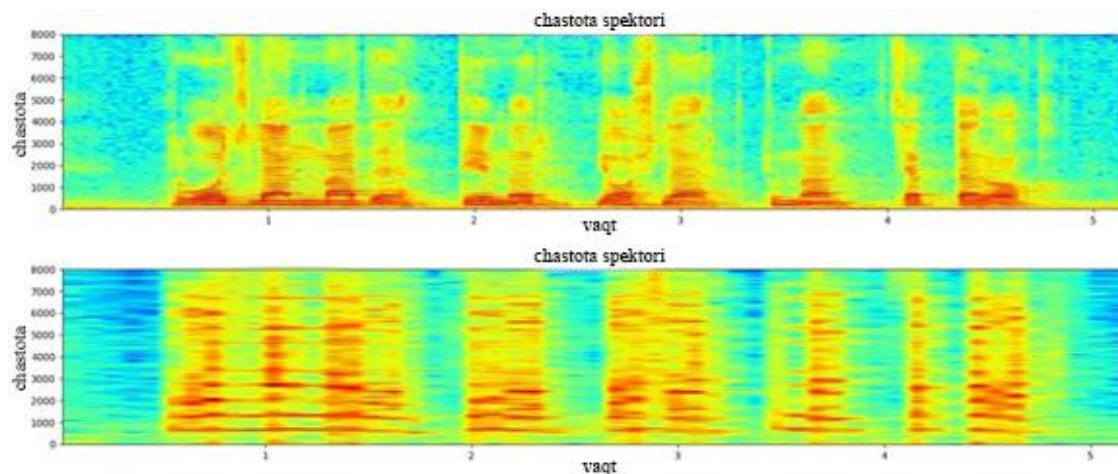
4-qadam. Freymalar bitta vektorga birlashtiriladi.

5-qadam. Nutq signalidan belgilar to‘plami shakllantiriladi.

6-qadam. Shaxs modelini qurish va uning asosida shaxsni identifikatsiyalash amalga oshiriladi.

7-qadam. Tamom.

Nutq signalining spektrogrammasi va ungi Maksimum-minimum usuli orqali diskretlashdan keyingi spektrogrammalari 2.5-rasmda keltirilgan.



2.5-rasm. Nutq signali va unga Maksimum-minimum diskretlash algoritmi natijasi

Natijalarni baholash. Yuqorida taklif etilgan usullarni samaradorligini baholash uchun 50 ta shaxsdan iborat bo'lgan nutq bazasida shaxsni ovozi asosida identifikatsiyalash masalasi yechildi. Xar bir shaxs uchun 10 ta nutq fayli mavjud bo'lib, shundan 3 tasi o'qitish va qolgan 7 tasidan testlashda foydalanildi. Har bir fayl davomiyligi 3-5 sekundni tashkil qiladi. Quyidagi jadvalda yuqorida keltirilgan algoritmlar asosida shakllantirilgan belgilardan foydalanib identifikatsiyalash natijalari keltirilgan.

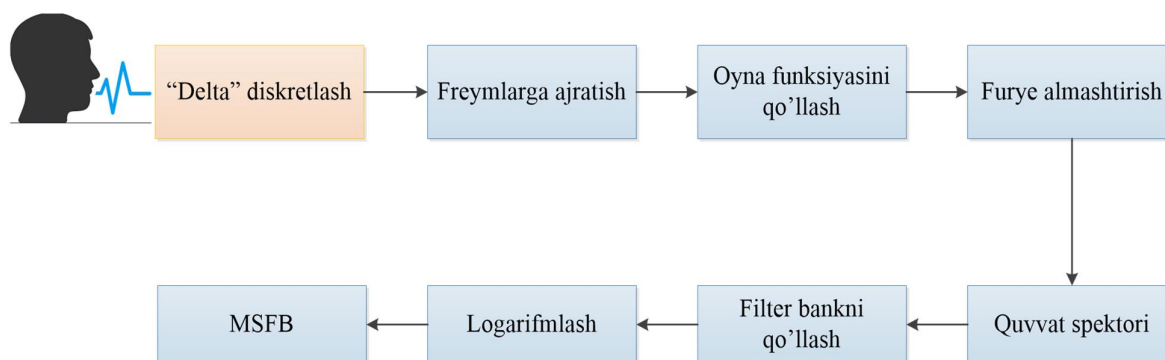
2.1-jadval

Asosiy va taklif etilayotgan usullar natijalarini baholash

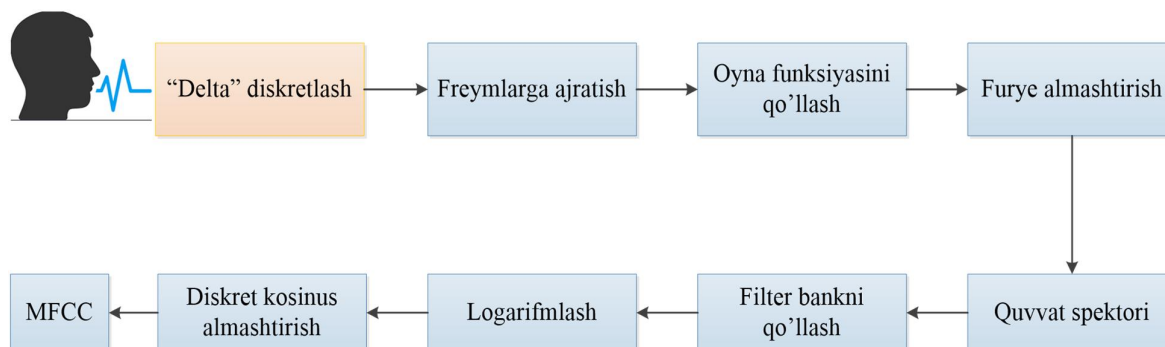
Usul	Ma'lumotlar hajmi (MB)	Aniqlik (%)	O'qitish uchun sarflangan vaqt (sek.)	Identifikatsiyalashga sarflangan vaqt (sek.)
Asosiy	135,3	99,8	27,2	37,4
Delta diskretlash	15,8	97,6	3,5	7,6
Maksimum-minimum	15,8	98,6	3,4	7,2

2.5-§. Belgilarni shakllantirish algoritmlari

Shaxsni uning ovozi asosida identifikatsiyalash masalasini yechish uchun dastlab shaxs nutq signalidan uni xususiyatlarini o'z ichiga olgan belgilar to'plamini aniqlash zarur. Nutq signallaridan belgilar to'plamini shakllantirish uchun mel-masshtablangan filtr banklar va mel-chastotali kepsstral koeffitsiyentlar algoritmlari blok sxemasi 2.6 va 2.7-rasmlarda keltirilgan.



2.6-rasm. Mel-mashtablangan filtr banklar algoritmi blok-sxemasi



2.7-rasm. Mel-chastotali kepsstral koeffisiyentlar algoritmi blok-sxemasi

Mel-mashtablangan filtr banklar algoritmi. Odatda insonlar past chastotalarda tovush o'zgarishlarini osonroq qabul qiladilar. Uning eshitish qobiliyati taxminan 1000 Gs gacha deyarli chiziqli, keyin u yuqori chastotalar uchun ko'proq logarifmik bo'ladi, shuning uchun spektrogramma logarifmi olinadi. Mel shkala - bu spektrdagi yuqori chastotalarni ajratib turadigan mel oralig'idagi filtrlarni joriy qiluvchi shkala bo'lib, unga o'tish quvvat spektral zichligi va uchburchak filtrlar soni skalyar ko'paytmasini hisoblash yo'li bilan amalga oshiriladi. Nutq signallaridan ushbu spektral xususiyatlari asosida belgi ajratish mel-mashtablangan filtr banklar (MSFB) belgilar to'plami sifatida tanilgan.

Nutq signalidan mel-mashtablangan filtr banklar belgilar to'plamini shakllantirish algoritmi quyida keltirilgan.

1-qadam. A1 algoritmini 1-7-qadamlari amalga oshiriladi.

2-qadam. Paydo bo'lgan vektor 32 ms li 10 ms qadam bilan freymlarga ajratiladi.

3-qadam. Har bir freym uchun quyida berilgan Xemming oynasi qo'llaniladi

$$w(t) = 0,54 - 0,46 \cos \frac{2\pi t}{N-1}$$

4-qadam. Har bir freym uchun spektral tahlillar olish maqsadida quyida keltirilgan Furye almashtirishi qo'llaniladi

$$H(n, k) = \sum_{n=1}^N x(n) w(n) e^{-\frac{2\pi i k n}{N}}$$

5-qadam. Har bir freym uchun quyida keltirilgan ifoda orqali quvvat spektri hisoblanadi.

$$S(n, k) = |H(n, k)|^2$$

6-qadam. Mel-filtr banki qo'llaniladi

$$f_{mel} = 2595 \log_{10} \left(\frac{f}{700} + 1 \right)$$

7-qadam. Logarifmlash orqali mel-masshtablangan filtr banklari olinadi

$$MSFB = 20 \log_{10} (S(n, k) \cdot F_m(k))$$

8-qadam. Tamom.

Mel-chastotali kepsral koefitsiyentlar algoritmi. Mel-chastotali kepsral koefitsiyentlari (MFCC) nutq signalini tavsiflash uchun qo'llaniladigan taniqli belgilar to'plami hisoblanadi. Ular nutq signali past chastotali komponentlari tomonidan olib boriladigan ma'lumotlar yuqori chastotali komponentlarga qaraganda fonetik jihatdan inson uchun muhimroq ekanligi haqidagi ma'lum dalillarga asoslanadi. Mel-chastotali kepsral koefitsiyentlarini hisoblash texnikasi qisqa muddatli tahlilga asoslanadi va shuning uchun har bir freymdan MFCC vektori hisoblanadi.

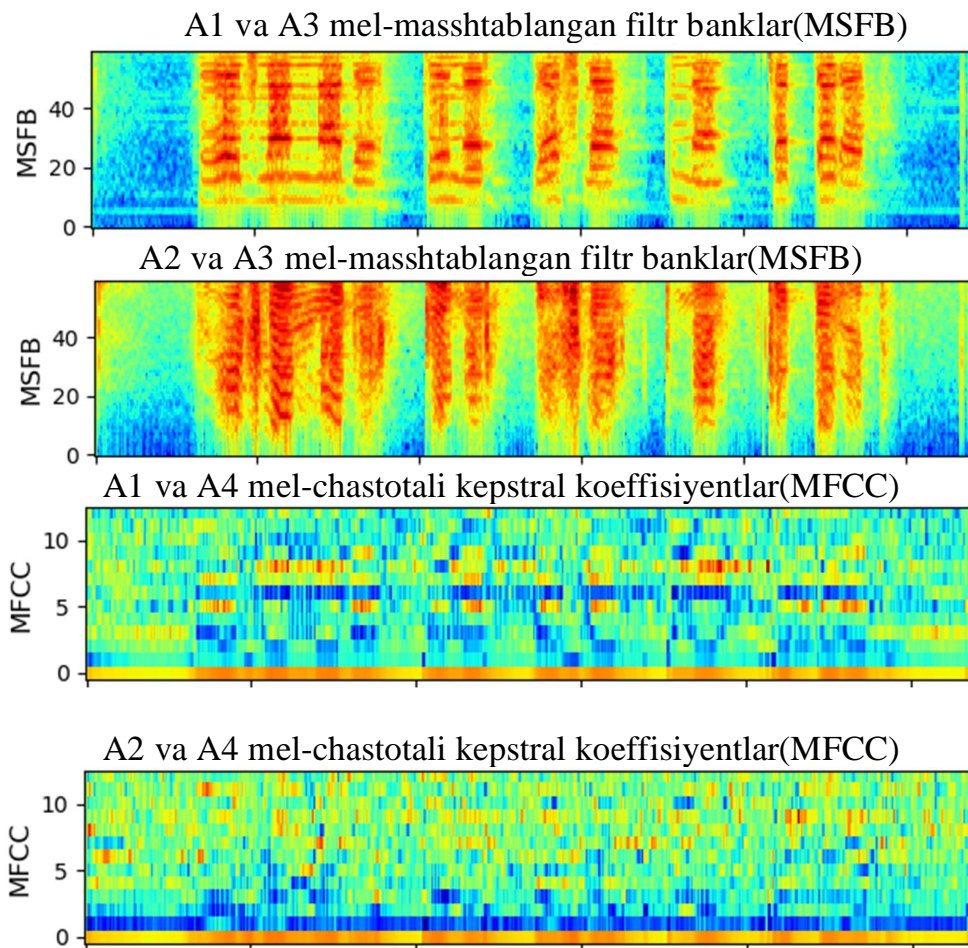
Nutq signalidan mel-chastotali kepsral koefitsiyentlar belgilar to'plamini shakllantirish algoritmi quyidagi qadamlarda amalga oshiriladi.

1-bosqich. A2 algoritmi qo'llaniladi.

2-bosqich. Diskret kosinus almashtirishi amalga oshiriladi va natijada mel-chastotali kepsral koefitsiyentlari shakllantiriladi.

$$C_k = a_k \sum_{i=1}^M X_i \cos \frac{\rho k (2i+1)}{2M}, k = 0, 1, 2, \dots, K$$

Nutq signallaridan mel-masshtablangan filtr banklar va mel-chastotali kepstral ko'effitsiyentlar belgilar to'plamini shakllantirishga spektr ko'rinishida misollar 2.8-rasmda ko'rsatilgan.



2.8-rasm. Shakllantirilgan nutq spektrogrammasi

III BOB. SHAXSNI OVOZI ASOSIDA IDENTIFIKATSIYALASH VA AUTENTIFIKATSIYALASH ALGORITMLARINI ISHLAB CHIQUISH

Mazkur bob shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash algoritmlarini ishlab chiqishga bag'ishlangan bo'lib, unda nutq signallarini segmentlarga ajratish, tanib olish masalasi uchun nutqli ma'lumotlar bazasini shakllantirish, chuqur neyron tarmoq asosida shaxsni ovozi asosida identifikatsiyalash modelini amalga oshirilgan. Bundan tashqari, mazkur bobda so'zlovchini matnga bog'liq hamda bog'liq bo'lmagan holda verifikatsiyalash masalasi bayon etilib, ikki omilli autentifikatsiyalash uchun takrorlanmaydigan so'zlar matrisasini hosil qilish algoritmi taklif etilgan.

3.1-§. Nutq signallarini segmentlarga ajratish usullari

Nutqni spektral tahlil qilish. Nutq spektral tahlilida, umumiy holda, signal diskret Furye almashtirishiga uchraydi, natijada paydo bo'lgan spektr uchun amplituda va chastotalar fazosida logarifmik masshtab o'zgartirilishi amalga oshiriladi (mel-chastota almashtirishi), spektr tekislanadi va uni egri chizig'i shakllantiriladi. [90; 323-326-b] ishda nutqni fonemalarga mos bo'laklarga ajratish lateral tormozlanish mexanizmi bo'yicha o'zgartirilgan sonogrammani tahlil qilish va inosonni nutqni idrok etish matematik modelini (chastotali sohada lokal vaznlashtirilgan integratsiyalash mel - chastotasi almashtirishi) inobatga olgan holda amalga oshirilgan. Segmentlar chegaralari signal spektral tuzilishidagi sezilarli o'zgarish omili bilan belgilanadi. Joriy segment boshlanish vaqti ma'lum deb faraz qilinsa, segmentni oxiri spektr o'rtachasi va har bir profil yaqinlik o'lchovi muayyan bo'sag'adan oshib ketgan vaqt momenti hisoblanadi.

Joriy segment har bir vaqt momentidagi o'rtacha spektr rekursiv hisoblanadi [90; 323-326-b], ya'ni:

$$\bar{S}_i(\omega) = \frac{\bar{S}_{i-1}(\omega) \times (i - 1) + S(\omega, t)}{i}, \quad \bar{S}_0(\omega) = S(\omega, t_0), \quad (3.1)$$

bu yerda $\bar{S}_i(\omega) - [t_0, t_i]$ oraliqdagi o'rtacha amplitudaviy spektr, $S(\omega, t_i) - t_i$ vaqt

momentidagi sonogramma profili.

[91; 323-326-b] da alveolyar frikativlarni va palatal ajratish uchun tovush spektrlari ko‘rib chiqiladi: palatal tovushlar kompakt spektr bilan xarakterlanadi, uning asosiy cho‘qqisi nisbatan past chastotada joylashgan, alveolyar tovushlar cho‘qqisi yuqori chastotada bo‘ladi.

Nutq segmentlarini unli va undosh tovushlar guruhlariga avtomatik klassifikatsiyalash uchun [92; 161-171-b] ishda nutq signalini vokallashgan (unlilashgan) qismlarini korrelyasion funksiyasidan foydalanish bilan turli xil tovushlarni chastota spektri bo‘yicha energiya taqsimoti ko‘rib chiqilgan. Nutq signali segmentining past chastotali va yuqori chastotali pastki diapazonlari orasidagi energiya taqsimotini tasniflash uchun nisbiy polosali energetik mezondan foydalaniladi. Bunday holda, segmentatsiya mezoni sifatida ikkala diapazon osti uchun ham nutq signali spektri namunalari ketma-ketligi amplitudalari kvadratlari yig‘indilari nisbati hisoblanadi [92; 161-171-b].

Nutq kepral tahlili. Kepral tahlil nutqni gomomorf ishlov berishga asoslangan va nutq signalida generator va filtr qismlarni ajratishga imkon beradi [93; 2431–2434-b]:

$$C_s(q) = \frac{1}{2\rho} \int_{-\infty}^{\infty} \ln(S(w))^2 e^{iwq} dw, \quad (3.2)$$

bu yerda $s(w)$ nutq signalining amplitudaviy spektri.

[94] ishda keng fonetik klassifikatsiyada nutq segmentatsiyasini amalga oshirish uchun mel-chastotali kepral koefitsiyentlari (MFCC) koefitsiyentlari qo‘llaniladi. MFCC koefitsiyentlar signalning o‘zgarishini o‘zida aks ettirmaydi, dinamik o‘ziga xosliklarni aniqlash uchun kepral hosilasidan delta-kepral koefitsiyentlar DCC (Delta-Cepstral Coefficients) va delta-delta kepral koefitsiyentlar DDCC (Delta-Delta Cepstral Coefficients “tezlanish koefitsiyentlari”)dan foydalaniladi [95; 1-8-b].

Nutqni fonema bo‘yicha segmentatsiyasi masalasida kepral tahlildan foydalanish misoli [95; 1-8-b] ishda keltirilgan. Bayes informativ mezoni to‘g‘ridan-to‘g‘ri segmentatsiya uchun qo‘llaniladi. 40 ta uchburchak filtdan iborat Grebenka

ham xuddi shu tarzda [96; 450-458-b] ishda 12 ta MFCC koefitsiyentini hisoblashda qo'llanilgan. Bunda har bir tahlil vaqt oralig'i 20 ms davomiylikka ega va 10 ms bilan bir-birini qoplash mavjud. [97; 135-136-b] ishda spektral o'tish o'lchoviga asoslangan fonemalarni segmentatsiyalash usuli (Spectral Transition Measure, STM) taklif qilingan:

$$STM(m) = \frac{\mathring{\mathbf{a}}^D a_i^2(m)}{D}, \quad a_i(m) = \frac{\mathring{\mathbf{a}}^I c_i(n+m) \times \mathcal{N}}{\mathring{\mathbf{a}}^I n^2}, \quad (3.3)$$

bu yerda m - joriy oyna tartib raqami; D - belgilar vektori o'lchami (MFCC koefitsiyentlari soni); $c_i(m)$ - m oyna uchun hisoblangan i -chi MFCC koefitsiyenti; $a_i(m)$ - regressiya koefitsiyentlari (MFCC koefitsiyentlari vektorining o'zgarish tezligi); I - regressiya koefitsiyentlarini $a_i(m)$ hisoblashda foydalanilgan oynalar soni (joriy oyna chap va o'ng tomonidagi oynalar uchun alohida sanaladi).

Ushbu algoritmni kamchiligi unli va unililashgan undosh tovushlar orasidagi chegaralarni koartikulyasiya natijasida yoki nutqni tekis talaffuz qilinishi holatlarida o'tkazib yuborish ehtimoli katta ekanligi hisoblanadi. Shuning uchun [98; 306-310-b] ishda keltirilgan algoritmni quyidagicha o'zgartirish taklif etiladi, ya'ni eksperimental ravishda aniqlangan kamchiliklar holatlarida kepstral tasvir chiziqli bog'liqlikka yaqin shaklga ega bo'ladi. Shuning uchun kepstral silliqlik o'lchovi (CSM) quyidagicha kiritiladi.

$$CSM(m) = \frac{\frac{c_{\max}(m) - c_{\min}(m)}{\sqrt{\mathring{\mathbf{a}}^{D-1} (|c_{i+1}(m) - c_i(m)|)^2}}, c_{\max}(m) \wedge c_{\min}(m)}{1}, \quad c_{\max}(m) = c_{\min}(m) \quad (3.4)$$

bu yerda m - joriy oyna raqami; $c_i(m)$ - m oyna uchun hisoblangan i -chi MFCC koefitsiyenti; $c_{\max}(m)$ - m oyna uchun maksimal MFCC koefitsiyenti; $c_{\min}(m)$ - m oyna minimal MFCC koefitsiyenti; D - bu kepstrum o'lchami.

Ushbu o'zgartirilgan STM + CSM algoritmining dastlabki STM algoritmi bilan taqqoslash natijalari 3.1-jadvalda keltirilgan.

STM va STM + CSM algoritmlari natijalari

	To'g'ri aniqlangan chegaralar	O'tkazib yuborilgan chegaralar	Noto'g'ri topilgan chegaralar
STM	82.9%	17.1%	20.2%
STM + CSM	90.9%	9.1%	22.3%

3.1-jadvaldan takomillashtirilgan segmentatsiya algoritmi katta miqdordagi kamchiliklar va noto'g'ri chegaralarni aniqlashga imkon berishini ko'rish mumkin, chunki ba'zi hollarda fonemadan fonemaga o'tishda ham spektral o'tish yoki kepstsal silliqlash bilan biron bir muhim o'zgarish bo'lmaydi.

Nutq signallariga ishlov berishda veyvlet-almashtirishlaridan foydalanish. Nostasionar signallarni tahlil qilish uchun Furrye almashtirishi va chiziqli bashoratlash koeffitsiyentlari bilan parametrizatsiyasi moslashtirilmagan, shuning uchun bu holda signalni vaqtinchalik o'ziga xos xususiyatlari haqida ma'lumot yo'qoladi [99; 138-146-b].

Nutq signalida fonemali bo'laklar spektral tasvirni nisbatan sekin o'zgarishi, shuningdek nutq apparatini tezkor qayta qurilgan bo'laklari (fonemalararo o'tish, portlovchi fonemalar), ya'ni signal spektrini tez o'zgaradigan qismlari mavjud bo'ladi. Bunday nostasionar bo'laklar nutq signali xususiyatlarini o'rganish uchun veyvlet - tahlildan foydalanish asosliroq bo'ladi. Veyvlet - almashtirishlari natijasida olingan veyvlet spektrogrammalari shakllanadigan chastotalar va boshlang'ich nutq signali garmonik tuzilishi haqida ma'lumotni o'z ichiga oladi [96; 450-458-b], chunki veyvlet-yoyilish bazis funksiyalari tahlil qilingan signalda ham chastota, ham vaqt xususiyatlarini aniqlash qobiliyatiga ega. Bu esa nutq signalini vaqtga bog'liq xususiyatlarini ajratish va lokallashtirishga imkon beradi [99; 138-146-b]. Bir fonemadan boshqasiga o'tish nutq apparati konfiguratsiyasi o'zgarishi bilan izohlanadi, bu nutq signalini bir yoki bir nechta yoyilishdagi veyvlet koeffitsiyentlari keskin o'zgarishida aks etadi [100; 43-48-b]. Segmentatsiya algoritmlarida bir vaqtning o'zida bir nechta bazislardan foydalanishga ruxsat beriladi, so'ngra ularni har biri uchun olingan natijalar birlashtiriladi [101; 169-174-b]. Gauss oynalari (Morlet veyvleti [102; 83-89-b], Xanning va Xanning [103; 209-211-b] nutq

signallarini tahlil qilishda modulyasiya qiluvchi veyvlet sifatida foydalaniladigan oyna funksiyalari keng qo'llaniladi. MHat veyvleti (Meksika shlyapasi) timsollarni tahlil qilish masalalarida taniqli hisoblanadi biroq, formanta chastotalarini past aniqligi sababli nutqni tahlil qilishda qo'llashga yaroqsiz. [104] ishda diskret veyvlet - almashtirishiga asoslangan nutq signali segmentatsiyasi fonema turlaridan qat'iy nazar 85% ishonchlikni ko'rsatgan. Avtomatik segmentatsiyada segmentatsiya momentlari bo'yicha qaror xatolik funksiyasini o'zgarish tezligi bo'yicha qabul qilinadi. Buning uchun optimal kvadratik interpolyasiya bilan xatolik o'zgarish tezligi funksiyasi hisoblab chiqiladi va natijada paydo bo'lgan funksiya maksimumlari qidiriladi. Xuddi shunday, [105; 273-302-b] ishda signal bir-birini ustiga tushadigan qismlarga ajratilib, ularni har biri uchun diskret veyvlet-almashtirishi qo'llanilgan. [105; 273-302-b] ishda o'tkazilgan tajribalar natijalari Mayer, Dobeshi 16, Dobeshi 8 va Simlet 6 veyvletlaridan foydalanish samaradorligidagi ozgina farqni ko'rsatgan. Bu veyvletlarni qatorga yoyish bazisi sifatida segment chegaralarini tanib olish darajasini oshirish uchun kelajakda natijalarni birlashtirish orqali qo'llash mumkin degan xulosaga beradi. Veyvlet-segmentatsiya usullari ajratilgan segmentlarni tiplarga ajratish, ya'ni ularni asosiy tovush turlari bilan o'zaro bog'liqligini aniqlash imkoniyatini ta'minlamaydi.

Nutq korrelyasion tahlili. Signallarni korrelyasion tahlili nutq signali segmentatsiyasi masalalarida keng qo'llaniladi, chunki bu tahlil qilinayotgan fragment energiyasi kattaligini, vokalizatsiya mavjudligini va uni chastotaviy lokalizatsiyasini baholashga imkon beradi [106; 9-21-b]. Avtokorrelyasiya funksiyasi (AKF) bosh maksimumi nutq signalini tahlil qilingan fragmenti energiyasiga mos keladi, uni pasayishi esa nutq signalida xalaqitli tarkibiy qismini mavjudligini anglatadi. Unlilashgan fragmentlar uchun avtokorrelyasiya funksiya ko'rinishi to'g'riburchakli radio impulsning avtokorrelyasiya funksiyaga o'xshash bo'ladi. AKFni birinchi lokal maksimumiga qarab, ko'rib chiqilgan nutq fragmentida unlilashish mavjudligini va asosiy ton davri davomiyligini o'rtacha qiymatini aniqlash mumkin [107; 49-57-b]. AKF xalaqitlarga nisbatan barqaror, biroq unlilashgan fragment davriyligini baholash uchun yetarli darajada katta tahlil

oralig‘i talab etiladi. Natijada, unlilashgan segmentlarni o‘tish qismlarini ko‘rib chiqishda, ovoz bog‘lamlari tuzilishi va tebranish chastotasini keskin o‘zgarishi bilan birga, lokal maksimum sezilarli bo‘lmay qoladi. Algoritmarda [108; 1547–1557-b] ushbu kamchilikni bartaraf etish uchun asosiy ton davrining mumkin bo‘lgan qiymatlari ro‘yxatini aniqlash uchun normallashtirilgan kross-korrelyasion funksiya (NKKF) qo‘llaniladi. Agar nutq signalidan namuna olish oralig‘i $T = 1 / F_s$, F_s bo‘lsa va tahlil oralig‘i davomiyligi t bo‘lsa, $z = t / T$ o‘zgaruvchisi kiritiladi, so‘ngra NKKF quyidagicha hisoblanadi:

$$f_{i,k} = \frac{\overset{m+n-1}{\underset{\circ}{\mathbf{a}}} s_j s_{j+k}}{\sqrt{e_m e_{m+k}}}, \quad k = \overline{0, K-1}, m = iz, i = \overline{0, M-1} \quad (3.5)$$

$$e_j = \overset{j+n-1}{\underset{\circ}{\mathbf{a}}} \underset{l=j}{s_l^2},$$

bu yerda i - tahlil oralig‘i raqami, k - kechikish, n - tahlil oralig‘idagi nutq signali namunalari soni. Tahlil oralig‘i davomiyligi asosiy ton davrini kutilgan qiymatiga teng tanlanadi.

Davriy va tasodifiy jarayonlar uchun avtokorrelyasiya funksiyasini umumiy shaklidagi aniq farqlari korrelyasion tahlil nutq faolligini “ton/ton emas” (unlilashish mavjudligi yoki yo‘qligi)ga tasniflashda ham qo‘llaniladi. Korrelyasion tahlil usullari kvazigarmonik strukturaning stasionar unlilashgan fragmentlarini ko‘rib chiqishda yaxshi natijalarni beradi. Shu bilan birga, algoritmarni ishlashi asosiy ton tebranishlarini murakkab tuzilishi bo‘lgan tovushlarni mavjudligida sezilarli darajada murakkablashadi. Tahlilning yetarli davomiylikdagi oralig‘iga bo‘lgan ehtiyoj, artikulyasiya apparatini tez qayta qurilish fragmentlarida algoritmarni ishlashida xatolarga olib keladi. Bundan tashqari, korrelyasion tahlil usullari katta hisoblashlarini talab qiladi [109; 1-9-b], shuning uchun tadqiqotchilar sa‘y-harakatlari nutq signaliga ishlov berish uchun korrelyasion algoritmarni tezligini oshirishga qaratishadi [110; 783–789-b].

VAD (Voice Activity Detection) GSM algoritmi mashina vaqti nutqai nazaridan ko‘p manba sarf qiladi, Rabiner-Sambur algoritmi [107; 49-57-b] unga

nisbatan tejamkor bo‘lib, bunda bo‘lag‘ani shakllantirish uchun (3.6) va (3.7) formulalarda berilgan noldan kesib o‘tishlar funksiyasi Z_n va 10 ms davomiylikli oynada olingan nutq signali $x(m)$ ning o‘rtacha qiymati M_n funksiyalaridan foydalaniladi. Fon xalaqiti statistik parametrlarini baholash uchun algoritm birinchi 100 ms nutq signalini ko‘rib chiqadi (bu vaqt oralig‘ida nutq faolligi mavjud emas deb hisoblanadi).

$$Z_n = \sum_{m=-N}^N \left| \operatorname{sgn} x(m) - \operatorname{sgn} x(m-1) \right| w(n-m)$$

$$\operatorname{sgn} x(n) = \begin{cases} 1, & x(n) \geq 0 \\ -1, & x(n) < 0 \end{cases} \quad (3.6)$$

$$w(n) = \begin{cases} 1/2N, & 0 \leq n \leq N-1 \\ 0, & \text{bo‘shqa holatlar} \end{cases}$$

$$M_n = \sum_{m=-N}^N |x(m)| w(n-m) \quad (3.7)$$

Ushbu xususiyatlarni inobatga olgan holda, nutq faolligini aniqlashni mos bo‘lag‘alari hisoblanadi. M_n qiymati ITUning eng yuqori bo‘lag‘asidan oshib ketadigan nutq signali segmenti ajralib turadi. Ushbu fragment tashqarisida so‘zning boshi va oxiri yotadi, deb taxmin qilinadi. So‘ngra, M_n birinchi marta ITU bo‘lag‘a qiymatidan oshib ketgan undan vaqt o‘qi bo‘ylab harakatlanib, M_n birinchi marta ITL pastki bo‘lag‘asidan pastroq bo‘lgan moment aniqlanadi. Ushbu moment nutq faolligi taxmin qilingan boshlang‘ich moment sifatida olinadi. Nutq faolligini taxmin qilingan yakunlanish momenti ham shu kabi aniqlanadi.

Ikki bo‘lag‘ali algoritm nutq signali o‘rta qiymati dinamikasidagi qisqa muddatli tushishlar so‘zni boshlanish va oxirgi momentlarini noto‘g‘ri tanlashga olib kelmasligini ta‘minlaydi. Algoritmning afzalliklari soddaligi va hatto past energiyali tovushlarda ham nutq faolligi chegaralarini aniqlashda yetarlicha aniqligi hisoblanadi.

Faol va faol bo‘lmagan nutqni farqlash uchun alohida mezonlarga ega bo‘lgan VAD algoritmi [111; 61-65-b] ishda taklif etilgan. VAD-algoritmilarining ishlashiga

fon xalaqtlari sezilarli to'sqinlik qiladi, shuning uchun bunday vaziyatlarda yanada murakkab yondashuvlar qo'llaniladi. Moslashuvchan bo'sag'a qiymatiga ega VAD algoritmi [112; 880-883-b] ishda taklif etilgan. Ushbu algoritm hatto vaqt bo'yicha o'zgaruvchi signal-xalaqit nisbatida ham yaxshi natijalarni ko'rsatadi. Nutq faol qismini ajratib olish algoritmlarini ko'plab zamonaviy ishlanmalari tijoriy bo'lib, ular ochiq emas. Mavjud VAD-algoritmlarni asosiy muammolari fon xalaqitida ishonchsiz ishlashi va katta hisoblash resurslari talab qilishi hisoblanadi. Ushbu usulni yana bir kamchiligi, nutqni xalaqitga qaraganda balandroq tovush sifatida qarashi hisoblanadi. Yana bir kamchilik sifatida masshtablash koeffitsiyentlarini sinchkovlik bilan eksperimental tanlash zaruratini keltirish mumkin. Ko'rsatilgan kamchiliklarni bartaraf etish maqsadida mazkur ishda lokal ekstremumlar tarqalishini tahlil qilishga asoslangan yondashuv taklif etilgan.

Nutq faolligini asosiy turlarini ajratish. Nutq faolligini unlilashgan / undosh / portlovchi / sirg'aluvchi tovushlarga ajratish zarurati nutq signalilarning bu qismlarining signal xususiyatlarini tubdan farq qilishi va shunga mos ravishda funksional nutq algoritmlarida turli xil ishlov berish usullaridan foydalanish bilan bog'liq. Ko'rib chiqilgan usullardan tashqari, tegishli segmentlarni o'ziga xos xususiyatlaridan ayrim tovush turlarini ajratish uchun ham foydalanish mumkin. Bunday belgilar signal tebranish tezligi, o'rtacha quvvat va davomiylik hisoblanadi.

Tasniflash parametrlaridan biri bu nolni kesib o'tish chastotasi (ZCR, Zero Crossing Rate): nutq signalini unlilashmagan qismlari tezroq tebranadi va shuning uchun ZCR qiymati ancha yuqori, bu esa ushbu parametрни "ton / ton emas" segmentatsiya uchun foydalanishga imkon beradi [113; 90-95-b]. Energetik nutqai nazardan, unli tovushlarni farqlash uchun oddiy biroq, samarali parametr o'rtacha quvvat: unlilashgan tovushlar uchun odatda nisbatan katta qiymatlarni oladi [113; 90-95-b]. Segmentatsiya algoritmining samaraliroq ishlashi uchun segmentlar o'rtacha quvvatini ma'lum chastota diapazonida ko'rib chiqish mumkin [114; 23-29-b]. Nihoyat, talaffuz nutq apparatida to'siqni keskin ochilishi bilan bog'liq bo'lgan portlovchi tovushlar sirg'aluvchi va unli tovushlarga nisbatan qisqa davomiylik bilan ajralib turadi.

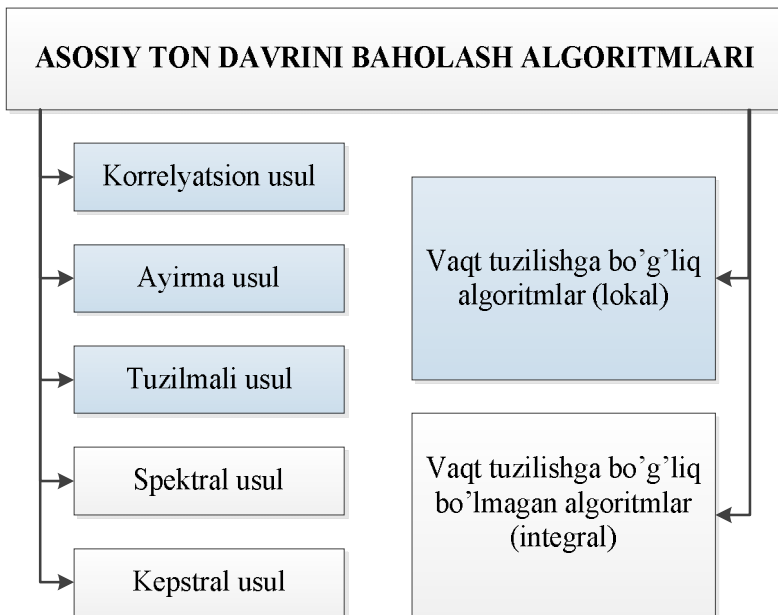
Asosiy ton davrlarini ajratib olish. Nutq signalini asosiy ton davrlariga segmentlash bu ovoz bog‘lamlarini har bir individual tebranishi vaqtli chegaralarini aniqlash bilan bog‘liq. Asosiy ton chastotasi nutq talaffuzi jarayonida uzluksiz o‘zgarib turadi. So‘zlovchi uchun o‘rtacha qiymatdan asosiy ton chastotasi og‘ishini uchta birlamchi sabablarini ajratish mumkin [115; 124-129-b]:

- vokal sikllar davomiyligidagi tasodifiy og‘ishlar (djitter-effektlar), chastotalar og‘ishi katta hamda kichik tomonga bo‘lishi mumkin;
- asosiy ton chastotasini davriy o‘zgarishi; tovush chastota modulyasiyasiga asoslangan vokal nutq xarakteristikasi, modulyasiya chastotasi 5-8Gs oraliqda yotadi;
- intonasion rang berish vositasi sifatida nutq jarayonida asosiy ton chastotasini bir tekisda o‘sishi va kamayishi; ovoz bog‘lamlari ostidagi bosim ortishi nutq balandligi ortishiga ham olib keladi.

Asosiy ton (AT) davrlarini fonogrammadan ajratish uchun turli matematik yondashuvlar asosida ishlovchi turli xil algoritmlar mavjud [117]. Asosiy ton davrini baholash uchun algoritmlarni ikkita asosiy kichik guruhlar mavjud [117; 19-32-b], ya’ni vaqt strukturasi bog‘liq va bog‘liq bo‘lmagan algoritmlar. Nutq signali asosiy toni davrini baholash algoritmlari shartli tasnifi 3.1-rasmda keltirilgan. Kulrang fonli bloklarda keltirilgan usullar, ma’lum bir tahlil qilish oraliqda asosiy ton davrini o‘rtacha (“integral”) bahosidan tashqari, nutq signalidagi muayyan namunalardan bilan bog‘langan asosiy ton davri joriy qiymatlarini ham olish imkonini beradi.

“Lokal” algoritmlar kichik guruhini mohiyati dastlabki o‘zgartirilmagan fonogrammada asosiy ton davri har bir qiymatini o‘lchashda hisoblanadi. Asosiy ton davrini baholashning strukturaviy usullari, nutq signali tuzilmasining ba’zi xarakterli davriy takrorlanuvchi xususiyatlarini asosiy ton davri oralig‘ida mavjud bo‘lgan davrlarga “bog‘laydi”. Bunday algoritmnin ishlashi natijasida nutq signalini asosiy ton davrlari bilan belgilash, shuningdek, fonogrammada asosiy ton davrini joriy bahosi shakllanishi mumkin. Ushbu kichik guruh usullari nutq signali asosiy ton davrlarida belgilar qo‘yishni amalga oshirishda algoritmlarda vaqt

segmentatsiyasidan foydalaniladigan samarali vosita hisoblanadi.



3.1-rasm. Asosiy ton davrini baholash algoritmlari tasnifi

ATning joriy chastotasini ajratish uchun uni garmonikasini birinchi garmonikasi chastotasiga nisbatan karrali bo'lish xususiyatidan foydalanish mumkin [118; 1-6-b]. Algoritm kirishiga nutq signali fragmenti berilishi kerak, uni davomiyligi AT davridan kattaroq bo'lishi ta'minlanadi. Algoritm ishlashi davomida spektrni butun son qiymatida bir necha marta siqish bajariladi. Bunda har bir siqishda mos yuqori garmonikali AT chastotasi birinchi garmonika chastotasiga to'g'ri keladi.

3.2-§. Tanib olish masalasi uchun nutqli ma'lumotlar bazasini shakllantirish

Shaxsni ovozi asosida identifikatsiyalash – uni oldindan, turli holat va muhitlarda yozib olingan nutq bo'laklari asosida yangi kiritilayotgan nutq bo'lagini tahlil qilish hamda tasniflash natijalariga ko'ra nutq egasi shaxsini aniqlash masalasi bo'lib, bunda tayanch nutq ma'lumotlar to'plami vazifasini ovozli ma'lumotlar bazasi bajaradi.

Ovozli ma'lumotlar bazasi - bu so'zlar, iboralar, jummalarni turli xil suxandonlar tomonidan yozilgan audioyozuvlar jamlanmasidir. Nutq signalini qayta

ishlash orqali ovozli ma'lumotlar bazasiga tayangan holda kirish nutq signalini tahlil qilish va tasniflash hamda uni shaxsini aniqlash masalasi uchun neyron tarmoq modellarini ishlab chiqish – eng maqbul yo'l hisoblanadi. Uzoq yillar davomida ushbu yo'nalishga oid masalalarni hal etishdagi asosiy to'siqlardan biri – bu o'zbek tilidagi nutq ma'lumotlar to'plamini yetarli emasligi bo'lib kelgan. So'nggi yillarda bunday ma'lumotlar bazalarini yaratish bo'yicha mamlakatimizda ham keng qamrovli ishlar amalga oshirilmoqda, biroq ularni umumiy hajmi hali ham yetarli darajada emas. Bundan tashqari, ushbu ishlanmalar uchun yetarlicha moliyalashtirish amalga oshirilmaydigan bo'lsa, nutqni va shaxsni tanib olish mukammal tizimini yaratish imkonsiz hisoblanadi[118; 1-6-b].

Nutq ma'lumotlar bazasini yaratishda bir nechta omillar inobatga olinishi shart. Bunda to'planuvchi so'zlar yoki nutqni o'zaro ketma-ket keluvchi ba'zi bo'laklari birgalikda ham talaffuz etilgan bo'lishi mumkin. Bundan tashqari, yozib olish parametrlari ham boshqacha bo'lishi ya'ni, tor polosali telefon yozuvlaridan (mono, namuna olish tezligi 8 kGs, har bir namunaga 8 bit), keng polosali mikrofondan (mono, 16 kGs, har bir namunaga 16 bit) va sinxron ko'p kanalli yozuvlardan (telefon va mikrofon, bir nechta mikrofon va h.k.) olinishi mumkin. Odatda telefon so'zlashuvidan olingan nutqni tanish tizimini yaratishda tor polosali ma'lumotlar bazalaridan foydalaniladi.

Nutq ma'lumotlar to'plamini yozib olishda bir qator omillar va vositalar muhim rol o'ynaydi. Masalan, texnik vositalar sifatida qo'llaniluvchi ba'zi asbob-uskunalar, yozib olish muhiti va yozib olish texnikasini shular toifasiga kiradi. Quyida ushbu omillar va texnik vositalarga ba'zi misollar keltirilgan[118; 1-6-b].

Texnik vositalar. Bunda asosan ovoz yozish vositalari nazarda tutiladi. Nutq ma'lumotlar to'plamini shakllantirishda ovoz yozish uchun mikrofon, telefon, gadjet kabi qurilmalardan foydalanilishi mumkin. Amaldagi yozish uskunasi sifati olingan nutq namunalari sifatiga sezilarli ta'sir ko'rsatadi. Aniqligi yuqori bo'lgan yozuvlarni olish uchun yuqori sifatga ega bo'lgan mikrofon va yozib olish dasturi talab etiladi.

Audio ma'lumotlarni yozish muhiti. Nutqni yozib olish muhitini yaratilayotgan tizim maqsadini inobatga olgan holda tanlash zarur. Agar ma'lumotlar to'plamidagi audiolar turli xil muhitda bo'lishi talab etilsa, u holda audiolarni hosil qilish xuddi shunday tamoyil asosida amalga oshiriladi. Shaxsni ovozi asosida tanib olish tizimi shovqinga bardoshli bo'lishi uchun turli xil muhitda yozib olingan audiolar zarur bo'ladi. Shovqinlarni ko'pligi tanib olish aniqligini pasaytiradi. Shuning uchun yozib olish muhiti atrofdagi shovqin va xalaqitlarni minimallashtirish uchun ehtiyotkorlik bilan tanlanishi shart. Yozib olish muhiti, shuningdek, nutq namunalari sifatiga ta'sir qilishi mumkin bo'lgan aks-sadolar yoki boshqa shaxs ovozlaridan holi bo'lishi maqsadga muvofiq hisoblanadi.

Yozib olish formati. Yozilgan nutq namunalari fayl formati nutq fayli parametrlarini o'qish samaradorligini oshishi yoki pasayishiga bevosita ta'sir qiladi. Nutq ma'lumotlar to'plamlari uchun eng ko'p qo'llaniluvchi ovoqli fayl formatlariga Wav, MP3 va FLAC kabilarni misol sifatida keltirish mumkin. Ushbu fayl formatlari ovoz sifati va fayl hajmi bilan bog'liq bir qator variantlarni taklif qiladi. Wav – fayl formati siqilmagan holda taqdim qilinuvchi va undan ma'lumotlarni o'qish hamda unga ma'lumot yozishda eng ko'p samaradorlik taklif qiluvchi format hisoblanadi. Unda ma'lumotlarni ortiqcha shifrlanishsiz (siqilmasdan) saqlanishi fayl hajmini keskin oshishiga olib keladi. Hozirgi kunga kelib mazkur fayl formati talab qiladigan darajada xotira hajmi bilan bog'liq muammolar bartaraf qilingach, nutq fayllari bilan ishlashda Wav – fayl formati eng qo'llanilayotgan fayl formati hisoblanadi.

Shevalar. Nutq ma'lumotlar to'plamini shakllantirishda muayyan tilda foydalaniladigan turli sheva va urg'ular ham inobatga olinishi shart. Nutq ma'lumotlar to'plami turli shevalarni o'z ichiga olishini ta'minlash uchun turli mintaqalardan kelgan so'zlovchilarni ushbu so'zlashuv tilida so'zlagan nutq namunalarini ham yozib olish talab etiladi.

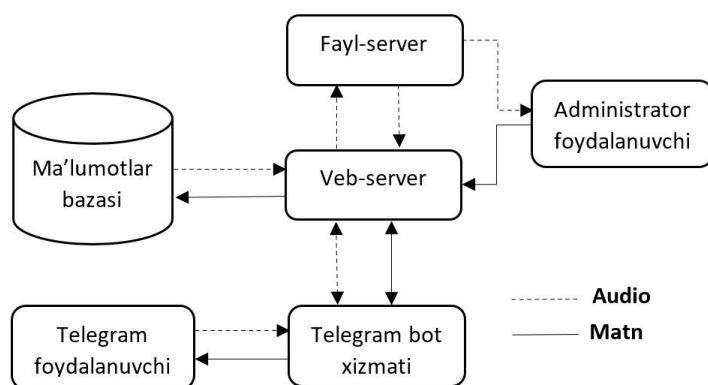
His-tuyg'ular. Nutq ma'lumotlar to'plamini shakllantirishda nutq namunalarini turli his-tuyg'ular ostida shakllanganligini ham inobatga olinishi zarur. Nutq ovoz fayliga yozilgan belgilarni shakllanishi uni ijro etuvchini ichki kechinmalariga va holatiga ham bog'liq bo'ladi. Nutq ma'lumotlar to'plami

quvonch, g‘azab, qayg‘u va qo‘rquvni o‘z ichiga olgan turli xil his-tuyg‘ularni aks ettiruvchi nutq namunalarini ham ichiga olishi mumkin. Bu orqali hosil qilingan model asosida shaxsni uni ichki kechinmalari va joriy holatini ham inobatga olgan holda tanib olish imkonini taqdim etadi.

Nutq ma’lumotlar to‘plami hajmi. Nutq ma’lumotlar to‘plamini shakllantirish nutqiy bo‘laklarni to‘plash orqali amalga oshiriladi. Nutq ma’lumotlar to‘plami hajmi deganda, mazkur nutqiy bo‘laklar hajm ko‘rsatkichlarini umumlashtirilgan miqdorini nazarda tutiladi. Audio ma’lumotlar umumiy davomiyligi vazifalariga ko‘ra tanlanadi. Tanib olish tizimining har bir audio fayli uchun 1 ta jumla(gap)da kamida 5 ta so‘z davomiyligi 5-10 s bo‘lishi yetarli hisoblanadi. Odatda kichik baza uchun - 40 soat, katta baza uchun 1000 soatdan kam bo‘lmagan audio ma’lumotlar to‘plami talab etiladi. Nutqli ma’lumotlar to‘plami hajmi (davomiyligi) uni aniqlik darajasini belgilab beruvchi muhim omil hisoblanadi. Hozirgi kunda nutqli ma’lumotlar to‘plamini shakllantirishni bir nechta yondoshuvlari taklif etilgan va ular qo‘llanilish maqsadlariga ko‘ra o‘zaro farqlarga ega. Mazkur ishda qo‘yilgan maqsadlarga mos keluvchi yondoshuvlardan quyidagi ikki turi qaraladi[119; 140-143-b]:

1. Birinchi yondashuvda audio kitoblardan tanlangan audiolarni saralash orqali turli shaxslar ovozlari yig‘ib olinadi va shaxsiy kompyuterda maxsus audio fayllar bilan ishlaydigan dasturlari yordamida audio fayllar 5-15s oraliqqa bo‘laklanib, mos foydalanuvchi nomi bilan saqlangan papkaga jamlanadi.

Ikkinchi yondashuv ijtimoiy tarmoq orqali foydalanuvchilar tomonidan yozilgan nutq bo‘laklarini to‘plashga asoslanadi. Bunda foydalanuvchiga oldindan yaratib qo‘yilgan matnlar taqdim etiladi, foydalanuvchidan unga taqdim etilgan matnlar ovoz chiqargan holda o‘qib undan yozib olingan nutq faylini yuborishi talab etiladi. Bunda telegram botdan foydalanish maqsadga muvofiq hisoblanadi. Yuqorida keltirilgan bosqichlarni avtomatlashtirish maqsadida maxsus tizim ishlab chiqish zarurati tug‘iladi. Taklif qilingan nutqli ma’lumotlar to‘plamini shakllantirishga mo‘ljalgan tizim funksional modullari quyidagi rasmda keltirilgan.



3.2-rasm. Nutqli ma'lumotlar to'plamini hosil qilish tizimi funksional modullari

Ikkinchi yondashuv haqida dastlabki tushunchalar berilganda ijtimoiy tarmoq ilovasi funksiyalari haqida umumiy fikrlar bildirib o'tildi. Alohida misol sifatida telegram ilovasini ko'rib chiqadigan bo'lsak, unda qo'llaniluvchi telegram bot funksiyalarini hamda ushbu jarayon bosqichlarini kengroq yoritib berish mumkin. Bunda maxsus telegram bot ishlab chiqiladi va u orqali veb-server bilan ijtimoiy tarmoq foydalanuvchilarini bog'lash amalga oshiriladi. Tizimni ishlashini nazorat qilib turish uchun maxsus administrator mavjud bo'lib, u bir nechta vazifalarni bajaradi, ya'ni foydalanuvchilarga tizimdan foydalanish huquqlarini belgilash, o'qilishi zarur bo'lgan matnlarni tahrirlash (yangi kiritish, o'zgartirish, o'chirish); fayl-server ishini nazorat qilish kabi vazifalarni bajaradi.

Mazkur monografiyada shaxsni ovozi asosida identifikatsiyalash masalasini yechish uchun ishlab chiqilgan usullar va algoritmlarini tajriba-sinovdan o'tkazish va tadqiq etish orqali samaradorligi aniqlashda VCTK, TIMIT, SITW, VoxCeleb1, VoxCeleb2 kabi xalqaro ochiq turdagi nutq ma'lumotlar to'plamlaridan foydalanilgan. Shu bilan birga o'zbek tilidagi nutqlardan iborat, oddiy sharoitda yozilgan nutqli ma'lumotlar to'plami–NutqUzData shakllantirildi. Ushbu nutq ma'lumotlar to'plamlari bo'yicha ma'lumotlar 3.2-jadvalda keltirilgan.

3.2-jadval Nutq ma'lumotlar to'plamlari bo'yicha ma'lumot

Nutq ma'lumotlar to'plami	Shaxslar soni	Audio fayllar soni
VCTK	110	44 000
TIMIT	630	6 300

VoxCeleb1	1 211	153 516
SITW	299	40 000
VoxCeleb2	5994	1 128 246
NutqUzData	6324	75888

VCTK - nutq texnologiyasi tadqiqot markazi (CSTR VCTK Corpus - Centre for Speech Technology Voice Cloning Toolkit) nutqni aniqlash tizimlarini ishlab chiqish va baholash uchun nutq ma'lumotlarini taqdim etish uchun mo'ljallangan. Nutqli to'plami 110 ta shaxsni ingliz tilida turli xil urg'u bilan aytilgan ovozlarni o'z ichiga oladi. Har bir shaxs tomonidan gazeta, jurnal parchasidan tanlangan 400 ga yaqin jumlar o'qilgan. Barcha nutq ma'lumotlari bir xil yozish sozlamalari yordamida yozilgan[120; 138-143-b].

TIMIT(Acoustic-Phonetic Continuous Speech Corpus) o'qish nutqi korpusi akustik-fonetik tadqiqotlar olib borish, nutqni avtomatik tanib olish tizimlarini ishlab chiqish va baholash uchun nutq ma'lumotlarini taqdim etishga mo'ljallangan. Nutq ma'lumotlar to'plamida 630 ta shaxsni keng polosali ovozlarni jamlangan bo'lib, ularni har biri 10 ta fonetik jumlarini o'qigan, har bir so'z uchun 16 bitli, 16 kGs chastotali nutq to'liq shakli fayl shakllantirilgan.

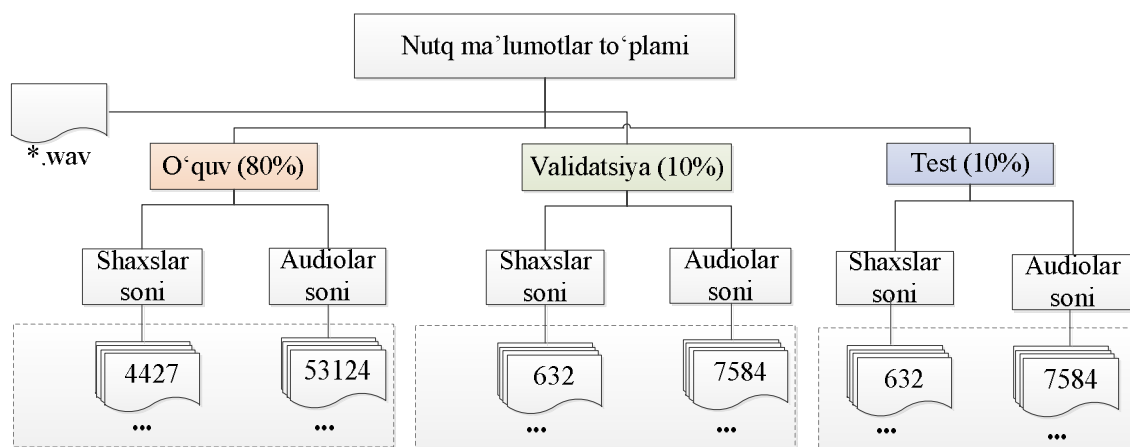
SITW (Speakers in the Wild) shaxsni tanib olish ma'lumotlar to'plamida cheklanmagan yoki turli muhit va sharoitlarda olingan nutq namunalarini qamrab olingan. Ma'lumotlar to'plami 299 ta shaxsni ovozidan iborat bo'lib, har bir shaxs tomonidan o'rtacha sakkiz xil jumlar o'qilgan. Shaxsni ovozi asosida tanib olishga mo'ljallangan mavjud ma'lumotlar to'plamlaridan farqli ravishda, ushbu ma'lumotlar to'plamida muayyan talablarga mos keluvchi muhitlardan nutq bo'laklari to'planmagan va shuning uchun u tabiiy shovqinlarni tarkibiga qo'shib olgan.

VoxCeleb - bu ochiq manbali mediadan avtomatik ravishda olinadigan keng ko'lamlı shaxsni tanib olish ma'lumotlar to'plamidir. Bu ikki nutqli ma'lumotlar to'plamida YouTube videolaridan olingan jumlar mavjud. VoxCeleb1 - nutq to'plamida 1211ta shaxsni 153 516 nutq segmentlaridan VoxCeleb2 - 5994 ta shaxsni 1 128 246 ta nutq segmentlaridan iborat. Ma'lumotlar to'plami turli xil

muhit va sharoitda to‘planganligi uchun undagi nutq bo‘laklari kulgi, o‘zaro suhbatlar, kanal effektlari, musiqa va boshqa tovushlar kabi tabiiy shovqin ta’sirida buzilishlarga uchragan. Shuningdek, ma’lumotlar to‘plami ko‘p tilli bo‘lib, 145 xil millat vakillari nutqlari turli xil aksyentlarda, yosh oraliqlarida, etnik mansubliklar va so‘zlashuv tillarida yozib olingan.

NutqUzData – nutq ma’lumotlar to‘plami shaxsni ovozi asosida tanib olishda neyron tarmoq modelini chuqur o‘qitishda foydalanilgan bo‘lib, unda jami 6324 ta shaxs va 75888 ta jumla mavjud. Neyron tarmoqni qo‘llagan holda shaxsni ovozi asosida tanib olish modeli tashqi shovqinlarga bardoshlilikini oshirish maqsadida, nutqni qayta ishlash uchun turli muhitlarda yozilgan ma’lumotlardan foydalanildi (3.3-rasm). Taklif etilayotgan neyron tarmoq modelini o‘qitish uchun davomiyligi 312 soat bo‘lgan audiokitoblar va telegram botdan yig‘ilgan audioyozuvlaridan foydalanildi. Ma’lumotlar to‘plamini yaratishda 6324 ta shaxs ishtirok etdi va neyron tarmoq modelini o‘qitish uchun dastlab qisqa muddatli (3-10 soniya) audio fayllar ketma-ketligidan modelni yaqinlashishi ortib borishi bilan o‘qitildi. Audio fayllar 16 kGs namuna olish chastotasiga ega va *.wav formatida saqlanadi. Nutq ma’lumotlar to‘plami umumiy hajmi 46 Gb ni tashkil qiladi. Nutq ma’lumotlar to‘plamini 90 foizidan neyron tarmoq modelini o‘qitish uchun foydalanilgan, qolgan 10 foizi esa modelni sinab ko‘rishda qo‘llanilgan[121; 25-32 b].

Shaxsni ovozi asosida identifikatsiyalashda qo‘llaniluvchi nutq ma’lumotlar to‘plami o‘ziga qo‘yilgan bir qator talablarga mos ravishda ishlab chiqilishi lozim. Bunday talablarga bazaviy nutq bo‘laklari olinuvchi shaxslarni turli yoshlarda, holatlarda va muhitlarda bo‘lishi, shu bilan birga ularni turlicha ichki kechinmalarida bo‘lishlarini inobatga olinishini kiritish mumkin.



Yozib olishga quyilgan talablar:

1. formati - *.wav
2. Kanallar soni – mono
3. Namuna olish tezligi – 16kGs
4. har bir namuna uchun joy – 16 bit

Yozib olish sharoiti – oddiy (uy, ko'cha, ertalab, abet, kechqurun) sharoitlarda yozib olingan



Nutq ma'lumotlar to'plami:

O'quv-5060 ta shaxslar soni, 60710 ta audiolar soni.

Validatsiya-632 ta shaxslar soni, 7584 ta audiolar soni.

Test-632 ta shaxslar soni, 7584 ta audiolar soni.

Jami: 6324 ta shaxslar soni, 75888 ta audiolar soni.

3.3-rasm. NutqUzData – nutq ma'lumotlar to'plamining umumiy ko'rinishi

3.3-§. Shaxsni ovozi asosida modelini qurish algoritmlari

Nutq signalidan belgilar to'plamini shakllantirib olgandan so'ng, shaxsni ovozi asosida identifikatsiyalash uchun ularni ma'lumotlar bazasida saqlanadigan mos etalon belgilar to'plami bilan solishtirish zarur.

Har bir shaxs ovozi uchun foydalaniladigan belgilar to'plami soniga va olingan belgilar to'plamini qanday modellashtirishni xoxlayotganligiga ko'ra, belgilarni moslashtirish va modellashtirishni bir nechta turlaridan foydalanish mumkin. Quyida belgilar to'plamini modellashtirishning eng keng tarqalgan usullari keltirib o'tilgan:

- shablonli modellashtirish (masalan, vektorli kvantlash - VQ);
- stoxastik modellashtirish (masalan, Gauss aralashma modeli - GMM);
- neyron tarmoqli modellashtirish.

Vektorli kvantlash (VQ) [71, 85-86] ma'lumotlarni siqishni samarali usuli bo'lib, turli ilovalarda, jumladan vektor kvantlash asosida kodlash va

identifikatsiyalashda muvaffaqiyatli qo‘llanilgan.

LBG algoritmi kod kitoblarini generatsiyalashda qo‘llaniladi [122; 140-143 b.]. LBG algoritmi quyidagi bosqichlarda amalga oshiriladi:

1. Vektorli kod kitobi yaratiladi. Bu o‘qitish vektorlari butun to‘plamining markaziy qismidir.

2. Har bir joriy kod kitobi y_n qoida bo‘yicha ajratish orqali kod kitobi o‘lchami ikki baravar oshiriladi:

$$\begin{aligned} y_n^+ &= y_n(1 + \epsilon) \\ y_n^- &= y_n(1 - \epsilon) \end{aligned} \quad (3.8)$$

bu yerda $n - 1$ dan joriy kod kitobi o‘lchami oralig‘ida va ϵ – ajratish parametri.

3. Ajratilgan kod kitobi uchun markazlar aniqlanadi, ya’ni bunda kod kitobi ikki baravar katta bo‘ladi.

4. M o‘lchamli kod kitobi hosil bo‘lmaguncha 2- va 3-qadamlar takrorlanadi.

Evklid masofasi nutq signalidan ajratib olingan belgilar to‘plami kod kitobiga kvantlangandan so‘ng hisoblanadi va u belgilar to‘plami orasidagi o‘xshashlik yoki farqni o‘lchashda foydalaniladi. Noma’lum shaxsni aniqlash noma’lum belgilar to‘plami va ma’lumotlar bazasidagi ma’lum so‘zlar modeli (kod kitobi) orasidagi Evklid masofasini o‘lchash orqali amalga oshiriladi. Bunda eng kichik o‘rtacha minimal masofaga ega bo‘lgan belgilar to‘plami tanlanadi. Quyida Evklid masofasini hisoblash formulasi keltirilgan.

$$d(x, y) = \sqrt{\sum_{i=1}^M (x_i - y_i)^2} \quad (3.9)$$

bu yerda x_i – kirish belgilari i -vektori, y_i – kod kitobidagi i -vektor, $d - x_i$ va y_i orasidagi masofa.

Gauss aralashmasi modellari (GMM) [123; 19-21-b.] shaxs ovozi ehtimolli modelini ta’minlovchi generativ shaxsni ovozi asosida identifikatsiyalash yondashuvi hisoblanadi. GMM shaxsni ovozi asosida identifikatsiyalashning eng samarali usullaridan biri bo‘lib, u ushbu sohada asosiy modellashtirish usuli hisoblanadi. Aralashmaning Gauss taqsimoti M komponentlari zichligi vazndagi

yig'indisidir: $p(\mathbf{x} | I) = \prod_{i=1}^M p_i b_i(\mathbf{x})$, bu yerda \mathbf{x} - D o'lchovli vektor, $b_i(\mathbf{x})$ - i -

komponent zichligi va p_i - i -komponent vazni. Bunda aralashma vazni quyidagi

ifodani qanoatlantiradi: $\sum_{i=1}^M p_i = 1$

Aralashmaning har bir komponenti D o'zgaruvchisi bo'lgan Gauss zichligi funksiyasidir:

$$b_i(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^{D/2} |\mathbf{S}_i|^{1/2}}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m}_i)^T \mathbf{S}_i^{-1} (\mathbf{x} - \mathbf{m}_i) \right\} \quad (3.10)$$

\mathbf{m}_i - matematik kutilma va \mathbf{S}_i - kovariatsiya matrisasi.

GMM vektori matematik kutilmasi, kovariatsiya matrisasi va barcha komponentlari vaznlari bilan tavsiflanadi. Shunday qilib, uni quyidagicha ixcham ko'rinishda ifodalash mumkin: $I = (p_i, \mathbf{m}_i, \mathbf{S}_i) \quad i = 1, 2, \dots, M$

Shaxsni ovozi asosida identifikatsiyalashda har bir shaxs o'zining λ parametrlari bilan GMM orqali tavsiflanadi. Kovariatsiya matrisalarini ko'plab turli xil variantlari mavjud [124; 138-143-b], masalan, model har bir komponent uchun bitta kovariatsiya matrisasidan, barcha komponentlar uchun bitta kovariatsiya matrisasidan yoki shaxs modelidagi komponentlar uchun bitta kovariatsiya matrisasidan foydalanish mumkin. Kovariatsiya matrisalari to'liq yoki diagonal ko'rinishga ega bo'ladi. X o'qitish tanlanmalari to'plamini hisobga olgan holda, GMM ni o'qitishning eng mashhur usuli, maksimal ehtimollikni baholashdir (ML-maximum likelihood). GMM ehtimolliqi quyidagicha hisoblanadi:

$$p(X | I) = \prod_{t=1}^T p(\mathbf{x}_t | I) \quad (3.11)$$

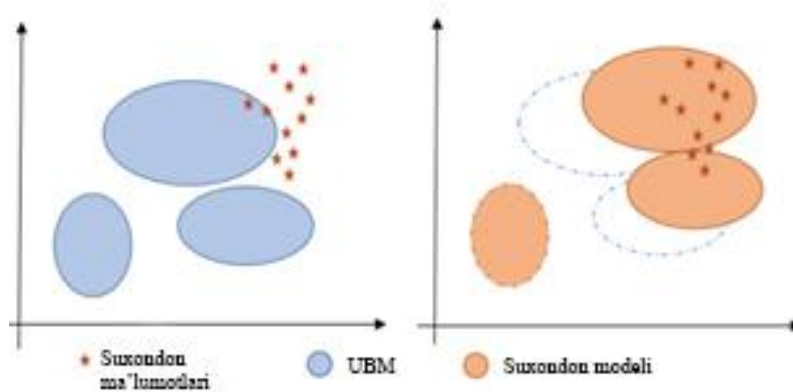
ML parametrlari odatda kutilma ehtimollikni maksimallashtirish (EM - expectation-maximization) algoritmi [125] yordamida baholanadi.

I_1, I_2, \dots, I_n parametrlar bilan tavsiflangan shaxslar to'plami orasida GMM tizimi ma'lum X so'zining aposterior ehtimolini maksimal darajada oshiradigan shaxsni qaytarish orqali o'z bashoratini amalga oshiradi:

$$\hat{s} = \arg \max_{I_k \in \mathcal{I}_n} P(X | I_k) = \frac{P(X | I_k) P(I_k)}{P(X)} \quad (3.12)$$

Agar barcha shaxslarni aprior ehtimolligi teng bo'lsa, masalan, $P(I_k) = 1/n$ k , ya'ni $P(X)$ barcha shaxslar uchun bir xil va logarifm monoton bo'lgani uchun (3.13)ni quyidagicha ifodalash mumkin:

$$\hat{s} = \arg \max_{I_k \in \mathcal{I}_n} \log P(X | I_k) = \arg \max_{I_k \in \mathcal{I}_n} \sum_{t=1}^T \log p(x_t | I_k) \quad (3.13)$$



3.4-rasm. Shaxs modelini yaratish uchun GMM ni moslashtirish

Yutuqlariga qaramay GMM hali ham ayrim kamchiliklarga ega [126; 1-9-b]. Birinchidan, GMMdan foydalanganda ko'p sonli parametrlarni o'qitish talab etadi. Ushbu fakt nafaqat qimmat hisob-kitoblarga olib keladi, balki yetarli miqdordagi o'qitish ma'lumotlarini ham talab qiladi. Shunday qilib, kichik ma'lumotlar to'plamida o'qitilganda GMM ishlashi ishonchsizdir. Ikkinchidan, generativ model sifatida GMM model o'qitilmagan ma'lumotlar bilan yaxshi ishlamaydi, bu esa past ehtimollik baholarini beradi. Ushbu ikki muammoni shaxsni moslashtirish orqali hal qilish mumkin. Shaxsni moslashtirishni asosiy g'oyasi shaxsning barcha ma'lumotlaridan foydalangan holda shaxsdan mustaqil tizimni moslashtirish orqali shaxsga bog'liq tizimlarni qurishdir. GMM-UBM bu barcha shaxslar ovozlari uchun o'qitilgan universal fon modeli hisoblanadi. GMM-UBM maksimal aposterior moslashuv yordamida shaxs modeliga modifikatsiyalanishi mumkin [127; 13-20-b].

Tizim samaradorligini baholash. Shaxsni ovozi asosida identifikatsiyalash tizimi samaradorligini baholash uchun bir nechta usullar mavjud bo'lib, ular turli toifalar uchun bir oz farq qiladi. Shaxsni verifikatsiyalash yoki ochiq-to'plamli

shaxsni identifikatsiyalash masalasi uchun Detection Error Trade-offs (DET) [128; 226-231-b] egri chizig'i va Equal Error Rate (EER) baholash ko'rsatkichlaridan foydalaniladi.

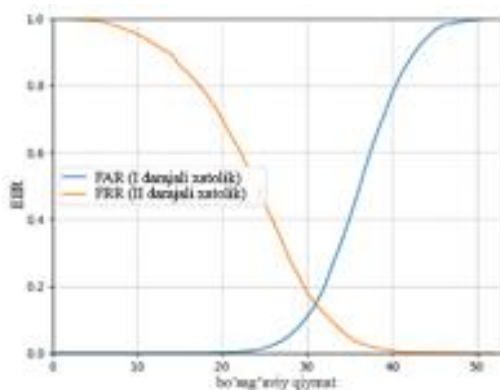
Shaxsni verifikatsiyalash yoki ochiq-to'plamli shaxsni identifikatsiyalash tizimlari samaradorligini baholashda esa ikkita muhim ko'rsatkich mavjud, ya'ni noto'g'ri qabul qilish darajasi (False Acceptance Rate (FRR)) va noto'g'ri rad etish darajasi (False Rejection Rate FAR)). FAR - bu tizim firibgarni identifikatsiya so'rovini noto'g'ri qabul qilish va FRR – tizim tomonidan haqiqiy shaxsni identifikatsiya so'rovini noto'g'ri rad etishi ehtimoli o'lchovidir. FAR va FRR orasidagi bog'liklik grafigi 3.5-rasmda keltirilgan va undagi kesishish nuqtasi teng xatolik darajasi EERni beradi. FAR va FRR mos ravishda quyidagi formulalar orqali hisoblanadi:

$$FAR = \frac{\text{noto'g'ri qabul qilinganlar soni}}{\text{barcha yolg'ondan urinishlar soni}} \cdot 100\%$$

$$FRR = \frac{\text{noto'g'ri rad qilinganlar soni}}{\text{barcha to'g'ri urinishlar soni}} \cdot 100\%$$

FAR ham FRR ham qaror qabul qilish jarayonida qo'llaniladigan q bo'sag'a qiymatiga bog'liq. Chunki past bo'sag'a qiymatida tizim har bir identifikatsiya so'rovini osongina qabul qilishga intiladi, shuning uchun kam noto'g'ri rad etishlarni, shu bilan birga ko'p noto'g'ri qabul qilishlarni amalga oshiradi. Aksincha, agar bo'sag'a qiymati yuqori bo'lsa, tizim barcha so'rovlarni osongina rad etadi va juda kam noto'g'ri qabul qilishlarni amalga oshiradi, lekin ko'plab yolg'ondan rad etishlar sodir bo'ladi.

Ikkita xato darajasi q bo'sag'a qiymatini qaror qabul qilish funksiyalari hisoblanadi. Demak, FAR ni FRR funksiyasi sifatida ifodalab, tizim samaradorligini baholash mumkin. Shaxsni verifikatsiyalash tizimi uchun FAR gorizont o'qda, FRR esa vertikal o'qda joylashadi. Qaror qabul qilish q bo'sag'a qiymatini sozlash orqali FAR va FRR orasidagi bog'liqlikni DET egri chizig'ida ifodalash mumkin.



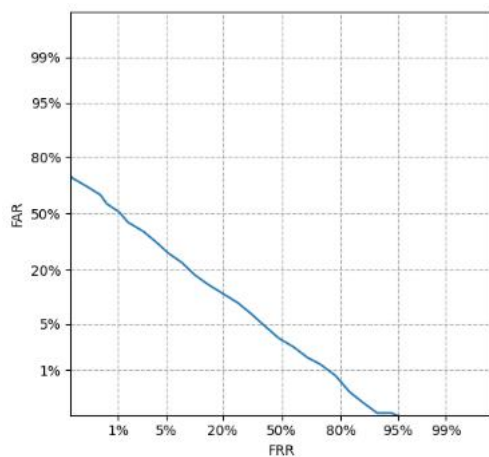
3.5-rasm. I darajali xatolik (FAR) va II darajali xatolik (FRR) orasidagi bog'liklik grafigi

DET egri chizig'ida FAR FRR ga teng bo'lgan nutqadagi xato darajasi teng xato darajasi (EER) deb ataladi. Shubhasiz, EER qanchalik past bo'lsa, tizim samaradorligi shunchalik yuqori bo'ladi, ya'ni tizim qanchalik yaxshi bo'lsa, egri chiziq nolga yaqinroq bo'ladi. Quyidagi rasmda DET egri chizig'iga misol keltirilgan.

Yopiq-to'plamli shaxsni identifikatsiyalash tizimi uchun ko'pincha identifikatsiya darajasi (IDR) baholash ko'rsatkichi sifatida foydalaniladi. IDR bu to'g'ri aniqlangan so'rovlar sonini ro'yxatdan o'tgan umumiy shaxslar to'plami soniga nisbati ko'rinishida baholanadi va quyidagicha ifodalanadi:

$$IDR = \frac{I_c}{I_c + I_i} \quad (3.14)$$

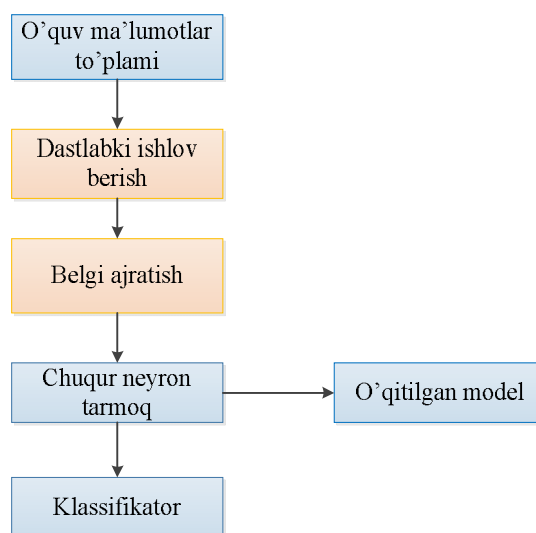
bu yerda I_c - va I_i - mos ravishda to'g'ri va noto'g'ri aniqlangan urinishlar soni.



3.6-rasm. DET egri chizig'i

Chuqur neyron tarmoq yordamida shaxsni ovozi asosida identifikatsiyalash modeli. Bunda model asl jumladan tasodifiy tanlangan qisqa segmentlar bo'yicha o'qitiladi. O'qitish jarayonini nazorat qilish va ortiqcha ko'p o'qitishni oldini olish uchun o'qitish to'plamining 5% validatsiya to'plami sifatida olindi. Tajribalar uchun ikki turdagi belgilar to'plamini ajratish funksiyalaridan foydalanilgan: mel-masshtablangan filtr banklar (A3 algoritmi) va mel-chastotali kepsral ko'effitsiyentlar (A4 algoritmi). Ma'lumotlarni qayta ishlashdan oldin 4 soniyadan qisqaroq audio segmentlar o'qitish ma'lumotlar to'plamidan chiqarib tashlandi. O'qitish va testlash ma'lumotlar to'plamidagi har bir jumla uchun bir xil qayta ishlash konveyeri qo'llanilgan.

Tajribalarning birinchi qismida har bir kadr uchun A3 algoritmi asosida 60 ta mel-masshtablangan filtr banklari va ikkinchi qismida esa har bir kadr uchun A4 algoritmi asosida 20 ta mel-chastotali kepsral ko'effitsiyentlar olingan. Chuqur neyron tarmoqni o'qitish asosida shaxsni identifikatsiyalash modelini yaratish jarayoni 3.7-rasmda keltirilgan.



3.7-rasm. Chuqur neyron tarmoqni o'qitish asosida tanib olish modelini yaratish jarayoni

Chuqur neyron tarmoqni o'qitish uchun tajribaviy tadqiqotlar Python dasturlash tilida olib borildi [129, 130; 28-31-b]. Har bir model GeForce GTX 3080 grafik prosessorida 200 ta epoxada o'qitildi. Taklif etilayotgan chuqur neyron

tarmoq arxitekturasi 3.3-jadvalda keltirilgan.

3.3-jadval

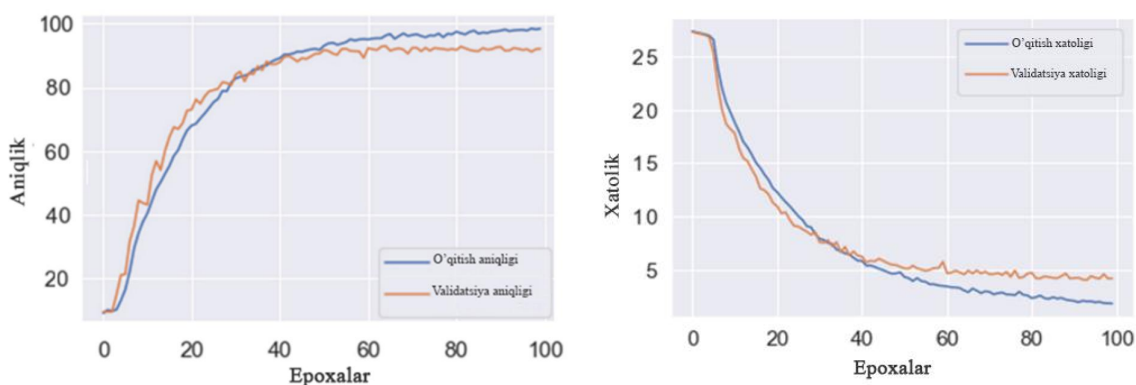
Chuqur neyron tarmoq arxitekturasi

Qatlam	O'lcham	kirish	chiqish
Blok 1	3x512	B x F x T	512 x F x T
Blok 2	2x512	512 x F x T	512 x F x T
Blok 3	2x512	512 x F x T	512 x F x T
SPOoling6	Mean	512 x F x T	3000
Conv1	3000	3000	512
Conv2	N	512	N
Fc7	512	N	512
Fc8	256	512	256

Blok tuzilmasi

Qatlam	o'lcham	kirish	chiqish
1D Conv	512 filter	B x F x T	512 x F x T
Batch Norm	-	512 x F x T	512 x F x T
ReLU	-	512 x F x T	512 x F x T
Dropout	0.5	512 x F x T	512 x F x T

Chuqur neyron tarmoqni o'qitish aniqlik grafigi 3.8-rasmda keltirilgan.



3.8-rasm. Chuqur neyron tarmoqni o‘qitishning aniqlik va xatoliklarni kamaytirish funksiyasi grafigi

3.4-jadval

Chuqur neyron tarmoqni o‘qitish natijalari

Diskretlash	Belgi shakllantirish	Modellashtirish usullari	Aniqlik (%)	Xatolik (Loss)
A1	A4	GMM	91,34	0,432
	A3	CNN	94,74	0,302
		DNN	96,34	0,224
A2	A4	GMM	92,01	0,409
	A3	CNN	92,62	0,359
		DNN	94,95	0,322

Chuqur neyron tarmoqni o‘qitish uchun VoxCeleb1 o‘quv ma’lumotlar to‘plamidan foydalanildi. Tajribaviy tadqiqotlar natijasi 3.6-jadvalda keltirilgan bo‘lib, chuqur neyron tarmoqni o‘qitishda taklif etilgan A1, A3 algoritmlari va chuqur neyron tarmoq modelidan foydalanganda 96,34% aniqlik va 0,224 xatolik, A2, A3 algoritmlari va chuqur neyron tarmoq modelidan foydalanganda 94,95% aniqlik va 0,322 xatolik bilan eng yaxshi natijaga erishildi. Ushbu tajriba tadqiqotlar natijasida, chuqur neyron tarmoq asosida shaxsni ovozi asosida identifikatsiyalash masalasi uchun belgilar to‘plamini shakllantirishda mel-masshtablangan filtr banklari (A3) algoritmi maqbul ekanligi aniqlandi. Eng yaxshi natijaga erishish orqali yaratilgan modeldan keyingi bosqichda chuqur neyron tarmoq asosida shaxsni ovozi asosida identifikatsiyalash tizimiga qo‘llash tavsiya etiladi.

3.4-§. So‘zlovchini matnga bog‘liq verifikatsiyalash

So‘zlovchini tanib olishning umumiy modeli. So‘zlovchini tanib olish masalasini klassik qo‘yilishining asosiy tushunchalari aniqlab olaylik [131, 132; 71-76-b].

W gipotetik to‘plam bu $w \in W$ tanib olish obyektlari to‘plamidir, bu yerda, mos ravishda, $w \in W$ element so‘zlovchiga tegishli nutq qismi.

Indikator funksiyasi deb W to‘plamni m ta kesishmaydigan W_1, \dots, W_m sinflarga ajratuvchi $g: W \rightarrow M$ funksiyaga aytiladi, bu yerda turli sinflar turli so‘zlovchilarga mos keladi, $M = \{1, \dots, m\}$ - so‘zlovchilar to‘plami. Indikator funksiyasi so‘zlovchini o‘qitish jarayonida quriladi.

Har bir $w \in W$ obyektga kuzatuvchi tomonidan bevosita idrok etiladigan o‘z $x(w) \in X$ timsolini mos qo‘yadigan $x: W \rightarrow X$ funksiya bo‘lsin. U holda X to‘plam kuzatuv fazosi deb ataladi.

$\hat{g}: X \rightarrow M$ hal qiluvchi qoida kuzatuvchiga X kuzatuvlar fazosida $w \in W$ obyektning $g(w)$ sinfini uning $x(w)$ timsoliga tayangan holda tanib olish imkonini beradi. Hal qiluvchi qoidaning sifati to‘g‘ri yechimlarning paydo bo‘lish chastotasi bilan o‘lchanadi. Odatda W obyektlar to‘plamiga qandaydir ehtimollik o‘lchovi berish orqali baholanadi, u holda so‘zlovchini tanib olish masalasi quyidagicha yoziladi $\min_{w \in W} P\{\hat{g}(x(w)) = g(w)\}$.

Identifikatsiyalash masalasi berilgan nutq signali bo‘yicha hisoblangan belgilar qiymatlari vektori asosida tizimga (kuzatuvchiga) ma‘lum bo‘lgan so‘zlovchilar to‘plamidan (to‘plam W) kuzatilgan ovoz bo‘yicha bitta so‘zlovchini tanlashdir. So‘zlovchini identifikatsiyalash uchun hal qiluvchi qoidani quyidagicha yozish mumkin $\hat{g}(x(w)) = \arg \min_{q=1, \dots, m} S(w, W_q)$.

Verifikatsiyalash masalasida so‘zlovchi ma‘lum deb hisoblanadi va taqdim etilgan nutq qismidan uning ushbu so‘zlovchiga tegishli ekanligi to‘g‘risida xulosa chiqarish kerak. Boshqacha qilib aytganda, avvaldan ma‘lum bo‘lgan W sinfga ko‘ra, va w obyektning (nutq qismi) $x(w)$ timsoliga ko‘ra, hal qiluvchi qoida

qarorlardan iborat bo‘ladi: tanib olish w obyektini W sinfga tegishli yoki aksincha. Faqat grammatik jihatdan bir xil nutq signalidan foydalanadigan so‘zlovchilarning matnga bog‘liq verifikatsiyasini ko‘rib chiqamiz, shuning uchun o‘qitish va verifikatsiyalashda so‘zlovchi bir xil jumla talaffuz qilishi kerak. X kuzatuv fazosining elementlari nutq signali asosida hisoblangan belgilar vektorlaridir[133; 102-111-b.].

Taklif etilgan so‘zlovchini verifikatsiyalash usulida nutq signali qisqa muddatli kesishmaydigan vokallashgan segmentlarga bo‘linadi, ular uchun individual xususiyatlar hisoblab chiqiladi. Obyektlar belgilarini $x^i \in V$ orqali belgilaymiz, $i=1, \dots, n$, bu yerda n - jumlaning segmentlari soni.

So‘zlovchini o‘qitish natijasida, W tanib olish obyektlarining barcha to‘plami kesishmaydigan W_1, \dots, W_m qism to‘plamlarga bo‘lingan bo‘lsin, ularning har biri tegishli so‘zlovchiga mos keladi.

Buning uchun tanib olinayotgan w obyekt hamda $W_q, q=1, 2, \dots, m$ sinf orasidagi yaqinlik o‘lchovi $S(w, W_q)$ - kattalikni ko‘rib chiqamiz. So‘zlovchini verifikatsiyalash uchun $\hat{g}: X \rightarrow M \otimes \{0, 1\}$ hal qiluvchi qoidani qayta aniqlaymiz, buning uchun dastlab kiritilgan metrikada w obyektini kuzatuv belgilar vektori kirishidan verifikatsiyalanuvchi so‘zlovchi W_q sinfigacha bo‘lgan masofani minamallashtirish mezonidan foydalaniladi.

$$\hat{g}(x(w), q) = \begin{cases} 1, & S(w, W_q) < l \text{ бۆлса;} \\ 0, & \text{акс холда,} \end{cases} \quad (3.15)$$

bu yerda l - talab etilayotgan birinchi va ikkinchi tur xatoliklarga bog‘liq holda aniqlanadigan chegaraviy qiymat.

So‘zlovchini o‘qitish natijasida W tanib olish obyektlarining barcha to‘plami kesishmaydigan W_1, \dots, W_m qism to‘plamlarga bo‘lingan bo‘lsin, ularning har biri tegishli so‘zlovchiga mos keladi[138; 256-260-b.]. Har bir sinfga tegishli obyektlar quyidagicha belgilanadi: $W_1 = \{w_{11}, w_{12}, \dots, w_{1k_1}\}, \dots, W_m = \{w_{m1}, w_{m2}, \dots, w_{mk_m}\}$.

Tanib olinayotgan w obyekt va o‘z obyektlari bilan berilgan $W_q, q=1, 2, \dots, m$

sinf o'rtasidagi $S(w, W_q)$ - yaqinlik o'lchovining $S(w, W_q)$ kattalikni ko'rib chiqamiz

$$S(w, W_q) = \frac{1}{|W_q|} \sum_{\{w\} \in W_q} d(w, w).$$

Yaqinlik yoki o'xshashlik o'lchovini aniqlash uchun $d(w, w')$ metrikani kiritish zarur, metrikani tanlash ixtiyoriy. Evklid metrikasi:

$$d^2(w_{pk}, w_{ql}) = \sum_{j=1}^n (x_{pk}^j - x_{ql}^j)^2 \quad (3.16)$$

$$p, q = 1, \dots, m; k = 1, \dots, k_p; l = 1, \dots, k_q,$$

bu yerda x_{pk}^j - p -sinf k -obyekti j -belgi qiymati, ya'ni w_{pk} obyektini; x_{ql}^j - q -sinf l -obyekti j -belgisi qiymati, ya'ni w_{ql} obyektini.

Evklid vaznli metrikasi yoki Maxalonobis metrikasi quyidagicha:

$$d^2(w_{pk}, w_{ql}) = \sum_{j=1}^n m_j (x_{pk}^j - x_{ql}^j)^2, \quad (3.17)$$

bu yerda m_j - maxsus aniqlangan vazn koeffitsiyentlari.

$$\text{Xemming o'lchovi: } d^2(w_{pk}, w_{ql}) = \sum_{j=1}^n y_j,$$

bu yerda

$$y_j = \begin{cases} 0, & x_{pk}^j \in [\min_{i=1, \dots, n} (x_{ql}^i), \max_{i=1, \dots, n} (x_{ql}^i)] \\ 1, & \text{акс холда.} \end{cases} \quad (3.18)$$

Shuningdek Chebishev o'lchovidan ham foydalanish mumkin:

$$d^2(w_{pk}, w_{ql}) = \max_{j=1, \dots, n} |x_{pk}^j - x_{ql}^j|. \quad (3.19)$$

Eng sodda metrikalardan biri l_1 - metrika hisoblanadi:

$$d^2(w_{pk}, w_{ql}) = \sum_{j=1}^n |x_{pk}^j - x_{ql}^j|. \quad (3.20)$$

$$p, q = 1, \dots, m; k = 1, \dots, k_p; l_1 = 1, \dots, k_q, \quad (3.21)$$

bu yerda x_{pk}^j - p -sinf k -obyekti j -belgisi qiymati, ya'ni w_{pk} obyektning; x_{ql}^j - q -sinf l -obyekti j -belgisi qiymati, ya'ni w_{ql} obyektini.

Birinchi tur xatolik deganda, (3.17) hal qiluvchi qoidaning natijasi 1 ga teng bo'lishi kerak bo'lsada, 0 qiymatiga teng bo'ladigan xatolar sonini tushuniladi.

Ya'ni, birinchi turdagi xato bir xil so'zlovchiga tegishli nutq signallari belgilar vektorlarini mos kelmasligi soni.

Ikkinchi tur xatolik deganda, (3.17) hal qiluvchi qoidaning natijasi 0 ga teng bo'lishi kerak bo'lsada, 1 qiymatiga teng bo'ladigan xatolar sonini tushunamiz. Ya'ni, ikkinchi turdagi xatolik turli so'zlovchilarga tegishli nutq signallari belgilar vektorlarini mos tushishlari soni.

So'zlovchini tanib olishni ishonchliligi deganda birinchi va ikkinchi turdagi xatolar darajasini tushuniladi: so'zlovchini verifikatsiyalash xatoligi soni qanchalik kichik bo'lsa, tizim shunchalik ishonchli.

Tadqiqotlar [139; 20-24-b] shuni ko'rsatdiki, eng oddiy l_1 - metrika so'zlovchini verifikatsiyalashni ishonchliligi nutqai nazaridan eng samarali hisoblanadi va shuning uchun uni so'zlovchini verifikatsiyalashda qo'llash tavsiya etiladi.

Hozirgi kunda chiziqli va nochiziqli ajratish funksiyalariga ega neyron tarmoqlardan foydalanadigan hal qiluvchi qoidalar keng qo'llanilmoqda, bu esa ushbu yondashuvdan foydalanish samaradorligini jiddiy o'rganishni talab qiladi. Dastlabki tadqiqotlar shuni ko'rsatadiki, bu yondashuv yuqori ishonchlilikka erishishi mumkin [140; 102-134 b, 141; 158 b] ishda hal qiluvchi qoidani qurishni ikki usuli tavsiflangan: statistik va algebraik. Tayanch vektorlar usuli tasniflash va regression tahlili masalalarini hal qilishda foydalaniladi [142; 20-24-b.].

So'zlovchini matnga bog'liq verifikatsiyalashning formant usuli.

So'zlovchini matnga bog'liq verifikatsiyalashning umumiy modelida nutq signalini umumiy qayta ishlash va parametrlashtirish mavjud bo'lgan o'qitish va verifikatsiyalash tartiblari ko'rsatilgan.

Qayta ishlash shovqinni tozalash, dastlabki va yakuniy to'xtalishlarni olib tashlash, signal darajasini normallashtirishdan iborat. Parametrlashtirishda birinchi navbatda nutq signalini segmentlash usulidan foydalaniladi. So'zlovchini tanib olishni mavjud tizimlarini aksariyati kadrlarni qayta ishlash standart usulini qo'llaydi, bunda signal ma'lum bir uzunlik va siljish qadami bilan kesishgan kadrlarga bo'linadi. So'ngra har bir segmentda belgilar vektori deb ataladigan unikal

signal belgilari hisoblab chiqiladi. O'qitish jarayonida foydalanuvchidan kalit iborani aytish so'raladi. Qayta ishlash va parametrlashdan so'ng, belgilar vektorlaridan iborat so'zlovchi etaloni yaratiladi va saqlanadi. Verifikatsiyalash jarayonida foydalanuvchidan identifikatorni taqdim etish va kalit iborani talaffuz qilish so'raladi. Qayta ishlash va parametrlashdan so'ng, verifikatsiyalashni yakuniy bosqichi amalga oshiriladi, ya'ni hal qiluvchi qoida yordamida ibora ma'lum bir so'zlovchiga tegishli yoki tegishli emasligi tekshiriladi[143; 121-167-b.].

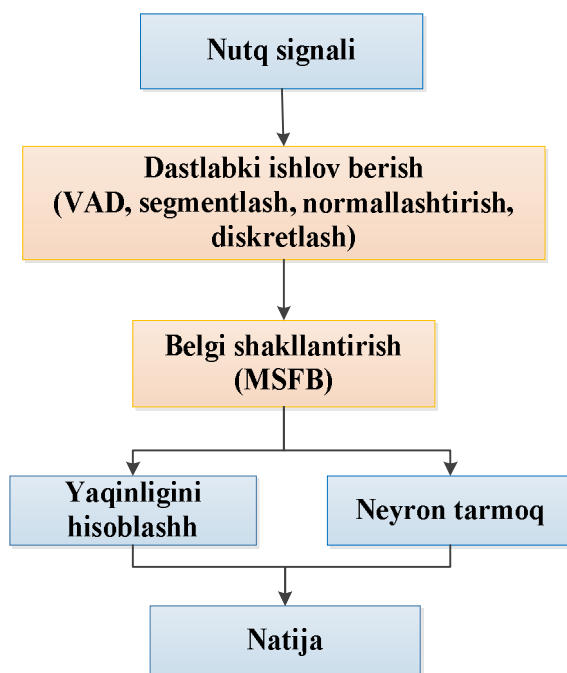
Belgilar vektorlarini hisoblash. Tajribalar orqali olingan ma'lumotlarga asoslanib, individuallikni asosiy fizik ko'rinishlarini nutq signali spektral yoki formantli belgilaridan izlash kerak. Formantlar to'plamini belgilar vektorlari sifatida belgilaylik.

"Formant" atamasi birinchi marta o'tgan asr o'rtalarida D.L.Flanagan tomonidan kiritilgan [133; 102-111-b.]. Fizikaviy tushunchalarni qoldirib, formant deganda silliqlangan spektr amplitudasi formanti amplitudasi deb ataluvchi lokal maksimum va unga mos formant chastotasi deb ataluvchi ekstremum nutqasi tushuniladi.

K² ta so'z algoritmi. Taklif etilgan modelni chuqur neyron tarmoq asosida shaxsni ovozi asosida identifikatsiyalash tizimida qo'llanilishi keltiriladi. Tizim arxitekturasi 3.9-rasmda keltirilgan bo'lib, u asosan ikkita qismdan iborat, ya'ni ro'yxatdan o'tkazish va identifikatsiyalashdan[142; 20-24-b.].

1. Shaxsni ovozi asosida ro'yxatdan o'tkazish va identifikatsiyalash jarayonida mikrofoniga qandaydir jummalarni gapirishi talab etiladi. Ushbu jarayon uchun alohida algoritmi taklif etilgan bo'lib, bunda tizim shaxsni ro'yxatdan o'tkazish jarayonida oldindan tayyorlab qo'yilgan k² ta so'zdan iborat matnni taklif qiladi.

2. Identifikatsiyalash jarayonida esa shaxsdan ushbu so'zlardan tashkil topgan M , ($M \in N$) satrni aytish talab etiladi. Bunda tizim tomonidan har bir ro'yxatdan o'tuvchi shaxs uchun takrorlanmaydigan va o'zaro kesishmaydigan jumla taqdim etiladi. Taklif etilgan algoritmi asosida o'tkazilgan tajribaviy tadqiqotlar natijalari quyidagi jadvallarda keltirilgan.



3.9-rasm. Chuqur neyron tarmoq asosida yaratilgan modelni shaxsni ovozi asosida identifikatsiyalashga qo‘llash jarayoni

Takrorlanmaydigan so‘zlar matrisasini hosil qilish algoritmi taklif etildi. Bu ikki omilli autentifikatsiya shartlarini bajarish orqali tizim ishonchligini oshirish imkonini beradi. Ikki omilli autentifikatsiya, xavfsizlikni ikki karra oshirishga xizmat qiladi, Birinchi omil sifatida shaxsning ovozi, ikkinchi omil sifatida esa aynan matritsadagi so‘zlar ketma-ketligi olingan.

Ikki omilli autentifikatsiya shartlarini inobatga olgan holda takrorlanmaydigan so‘zlar matrisasini hosil qilish algoritmi qo‘llash orqali 8 ta shaxs va 8 ta so‘zlar ketma-ketligida tadqiqotlar o‘tkazilgan. Natijaga ko‘ra taklif etilgan algoritm qo‘llanilganda 3,8% aniqlik oshganini ko‘rish mumkin.

3.5-jadval

Taklif etilayotgan usul qo‘llanilgandagi shaxsni identifikatsiyalash natijalari

	shaxs 1	shaxs 2	shaxs 3	shaxs 4	shaxs 5	shaxs 6	shaxs 7	shaxs 8
shaxs 1	86,5	36,5	47,4	43,4	42,9	43,6	34,3	42,1
shaxs 2	40,1	92,7	46,9	50,6	43,6	37,9	39,1	41,0

shaxs 3	48,5	49,2	94,8	43,5	36,7	40,8	30,5	40,7
shaxs 4	43,2	50,1	42,9	96,2	41,9	43,5	40,2	43,2
shaxs 5	40,5	44,7	37,9	43,7	93,3	43,5	43,8	45,8
shaxs 6	44,9	39,7	41,9	46,3	41,4	94,5	48,8	51,5
shaxs 7	34,7	35,7	33,9	36,9	44,9	50,2	95,9	40,5
shaxs 8	41,3	40,2	40,5	43,9	45,7	50,9	41,2	91,8

3.6-jadval

ikki omilli autentifikatsiya shartlarini inobatga olgan holda takrorlanmaydigan
soʻzlar matrisasi

kompyuter	gulzor	hudud	osmon	biznes	doʻst	banan	smartfon
smartfon	kompyuter	gulzor	hudud	osmon	biznes	doʻst	banan
banan	smartfon	kompyuter	gulzor	hudud	osmon	biznes	doʻst
doʻst	banan	smartfon	kompyuter	gulzor	hudud	osmon	biznes
biznes	doʻst	banan	smartfon	kompyuter	gulzor	hudud	osmon
osmon	biznes	doʻst	banan	smartfon	kompyuter	gulzor	hudud
hudud	osmon	biznes	doʻst	banan	smartfon	kompyuter	gulzor
gulzor	hudud	osmon	biznes	doʻst	banan	smartfon	kompyuter

IV BOB. SHAXSNI OVOZI ASOSIDA IDENTIFIKATSIYALASH VA AUTENTIFIKATSIYALASH DASTURIY TA'MINOTI VA UNI AMALIY MASALALARDA QO'LLANISHI

Ushbu bob monografiyada taklif etilgan usul va algoritmlarining tajribaviy tadqiqini o'tkazish uchun foydalaniladigan o'quv va nazorat tanlanmalari hamda belgilar fazosini shakllantirish uslubiyatini aks ettiradi. Bundan tashqari, mazkur bobda "Speaker.recognize.uz" dasturlar majmuasining tavsifi va uning amaliy tadbiri keltirilgan.

4.1-§. Shaxsni ovozi asosida identifikatsiyalash masalasining ma'lumotlar bazasi modellashtirish

Axborot texnologiyalari sohasi hozirda jadal sur'atlarda rivojlanayotgan tarmoqlardan biri hisoblanadi. Shuning barobarida axborot tizimlarining tarkibiy tuzilmalari tobora murakkablashib, ularni loyihalashga bo'lgan talab va ehtiyojlar ortib bormoqda. Texnik tizimlar tarkibiy qismlarining takomillashuvi, internet tarmog'ining keng tarqalishi, raqamli iqtisodiyot, elektron tijorat, mobil aloqa tizimlari kabi zamonaviy tuzilmalar va texnologiyalarni loyihalash murakkab arxitekturali ma'lumotlar bazalari (MB)ni modellashtirishga bo'lgan ehtiyojni oshirdi. Bugungi kunda ma'lumotlar bazasi axborot tizimlarining hamda dasturiy mahsulotlarning asosiy tarkibiy qismi hisoblanadi.

Tizim arxitekturasi, tarkibiy qismlarining tasvirlanishi va o'zaro munosabatlari, shuningdek, umumiylikdagi faoliyati jarayonlarini modellashtirishda IDEF dan tashqari UML, Chen va MIRIS kabi metodologiyalar ham mavjud bo'lib, ularning loyihalashdagi yondoshuvlari umumiy o'xshashliklarga ega. Faqat muayyan qismlarni va jarayon bosqichlarini ifodalashdagi belgilanishlar va metodologiyaga muvofiq holda loyiha arxitekturasiga moslik darajasidan kelib chiqilganda o'zaro farqli jihatlari ko'zga tashlanishi mumkin[134; 338-345-b.].

Yuqorida IDEF metodologiyasi va uning turlariga berilgan tavsiflarni

inobatga oladigan bo‘lsak, ma’lumotlar bazasi, dasturiy mahsulot, boshqa turli axborot texnologiyalari vositalari, texnik ta’minot va umumiy faoliyat jarayonlarining o‘zaro integratsiyalashuvini ifodalashda IDEF metodologiyasi nisbatan kengroq bo‘lgan imkoniyatlari orqali o‘zining tezkor ommalashuviga sabab bo‘ldi. Muayyan tizim loyihasini ishlab chiqish uchun avvalo xususiy imkoniyatlarini inobatga olgan holda eng ma’qul IDEF metodologiyasi turini tanlab olish zarur.

IDEF metodologiyasi turlarini qo‘llab-quvvatlaydi, lekin u IDEF0 va boshqa umumiy turlar uchun shablonni o‘z ichiga olganligi sababli, unda ma’lumotlar bazasi modelini chizish ham oson amalga oshiriladi.

4.1-jadval

Belgilarni solishtirish

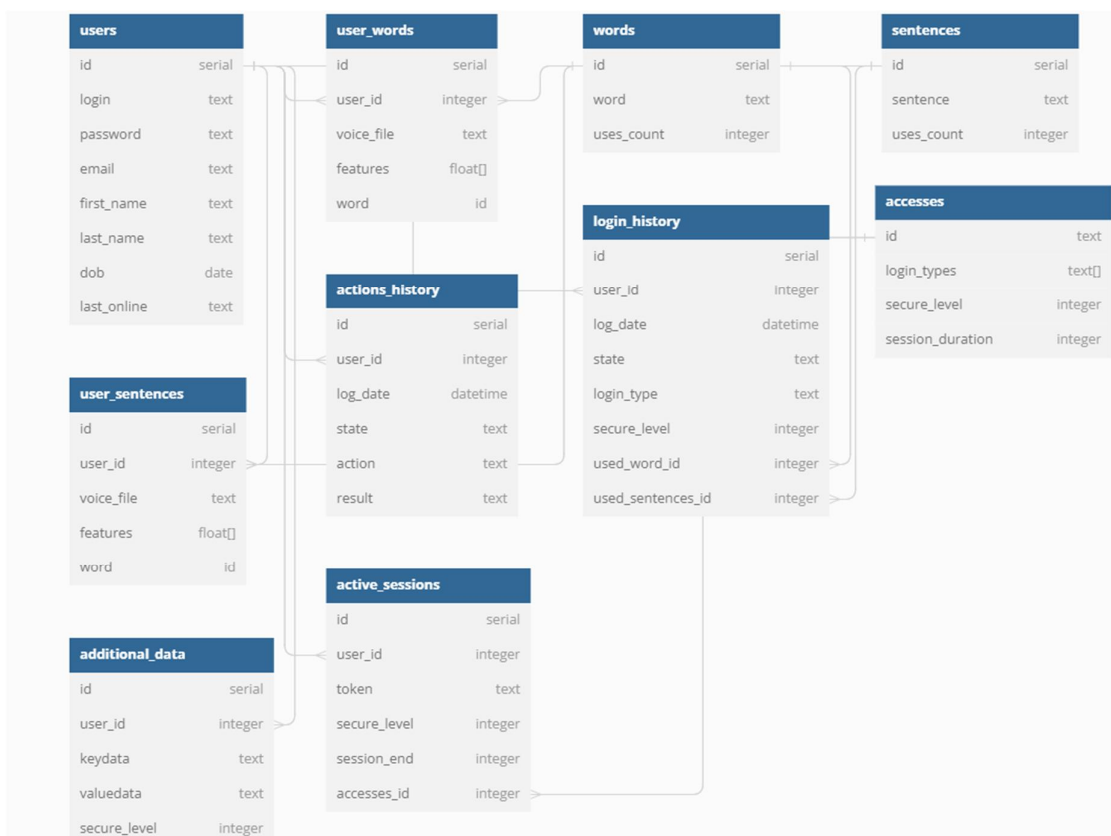
Tur nomi / mezonlari	IDEF0	IDEF3	IDEF5	IDEF1X
Chizish tezligi	5	4	4	5
Intuitivlik	4	3	3	4
tushunarligi	4	4	5	5
Qo‘llab-quvvatlash	3	4	4	5
Jami	16	15	16	19

Ma’lumotlar bazasini modellashtirishga mo‘ljallangan IDEF metodologiyasining to‘rtta turi (IDEF0, IDEF3, IDEF5, IDEF1X) taqqoslandi. Taqqoslashdan oldin har bir tur uchun tavsif berildi, batafsil o‘rganishlar olib borildi va ma’lumotlar bazasi modeli taqdim etildi. IDEF metodologiyasining barcha turlari chizmani hosil qilishga mo‘ljallangan bir xil dasturiy vosita orqali solishtirildi. Natijalar shuni ko‘rsatdiki, IDEF1X turi qolgan turlar ichida relyatsion ma’lumotlar bazasini loyihalashda qo‘llanilishi mumkin bo‘lgan eng maqbul metodologiya hisoblanadi.

Belgilar orasidagi eng katta farq - bu munosabatlarning bog‘langanlik ko‘lamida. IDEF1X da munosabatlarning o‘zaro bog‘langanlik ko‘lami nisbatan kattaroq.

IDEF1X ma’lumotli modellashtirish – bu obyekt munosabatlari modeliga asoslangan ma’lumotlar bazasini modellashtirish metodologiyasi. U ishlab chiqarish

tizimi yoki atrof-muhit funksiyalarini qo‘llab-quvvatlash uchun zarur bo‘lgan ma’lumotlarning tuzilishini ifodalovchi axborot modelini yaratish uchun ishlatiladi[135; 85-90-b.].



4.1-rasm. Shaxsni ovozi asosida identifikatsiyalash tizimining IDEF1X modeli

Yuqorida keltirilgan fikrlarni inobatga olgan holda shaxsni ovozi asosida identifikatsiyalash tizimi uchun nutq ma’lumotlar bazasini uchun model ishlab chiqish masalasini ko‘rib chiqamiz.

Maqsadga binoan tanlangan nutq ma’lumotlar bazasi tuzilmasida asosiy qism sifatida foydalanuvchi ma’lumotlari qaraladi. Ya’ni bunda nutq ma’lumotlar bazasini shakllantirishda berilgan matnli parchalarga mos ovoqli nutq bo‘laklarini taqdim qilgan har bir shaxs uchun alohida ma’lumotlar maydonlari hosil qilinadi. Ovozi asosida shaxsni identifikatsiyalash tizimida undan ro‘yxatdan o‘tgan va nutq ma’lumotlar bazasi uchun o‘zining nutq bo‘laklarini taqdim qilgan shaxslar uchun alohida profillar yaratiladi. Tizimning har bir a’zosi o‘zining profili orqali tizim imkoniyatlaridan foydalanishi hamda o‘zining turli holatlar va muhitlarda yozib

olgan nutq bo‘laklarini yuklashi mumkin.

4.2-§. Shaxsni ovozi asosida identifikatsiyalash va audentifikatsiyalash dasturiy ta‘minot strukturasi va interfeysi

Ushbu tadqiqot ishida yuqoridagi fikrlar hamda avvalgi boblarda keltirilgan algoritmlar tahlili bo‘yicha olingan xulosalar, tanlab olingan usullar va yondashuvlarga tayangan holda shaxsni ovozi asosida tanib olishga mo‘ljallangan dasturiy ta‘minot ishlab chiqilgan bo‘lib, undan muayyan shaxslarga oid jarayonlarning avtomatlashtirilgan holda, xavfsizlik va ishonchni ta‘minlash talab etiluvchi muhitlarda amalga oshirish vositasi sifatida foydalanish mumkin[145].

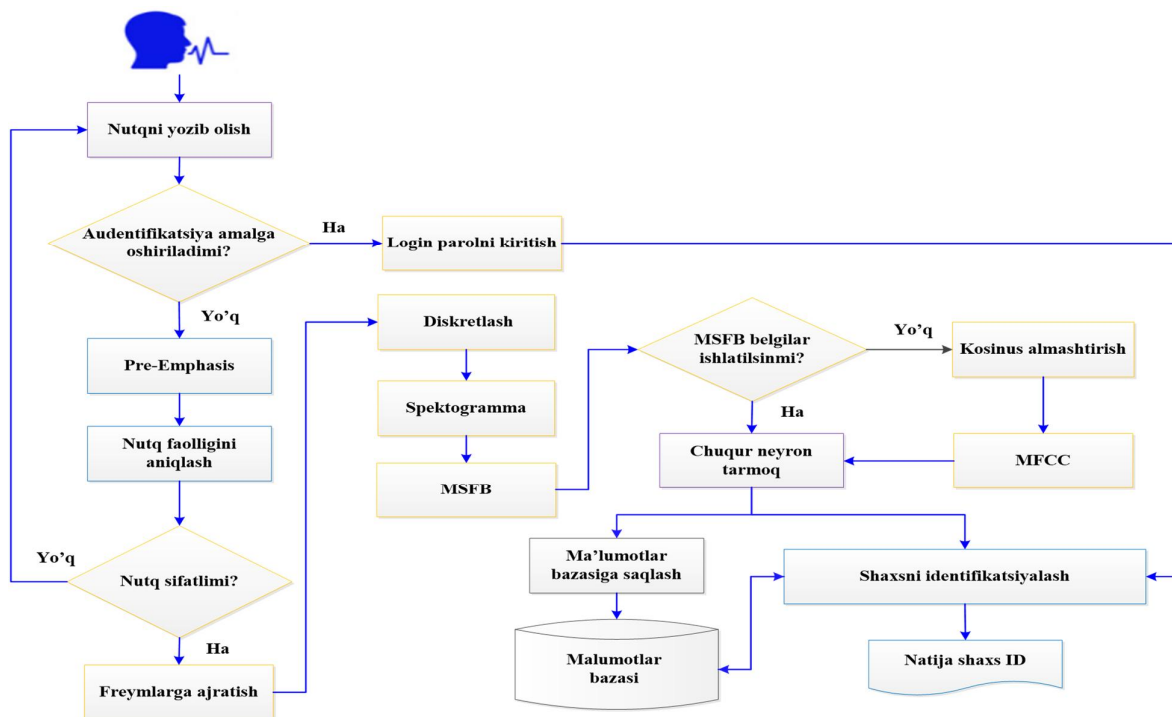
Shaxsni ovozi asosida tanib olish nutqni tanib olish singari kirish nutq signaliga dastlabki ishlov berish jarayonidan boshlanadi. Undan farqli tomoni shundaki, bunda shaxs nutqi emas, balki uning shaxsini aniqlash bilan bog‘liq belgilarni ajratib olish bosqichi amalga oshiriladi. Hozirda shaxsni ovozi asosida tanib olishning “Matnga bog‘liq holda tanish” va “Matnga bog‘liq bo‘lmagan holda tanish” deb nomlangan ikkita asosiy turlari mavjud. Matnga bog‘liq holda tanish tizimlari tekshirish bosqichida ham, o‘qitish bosqichida ham qat‘iy so‘z yoki jumalardan foydalanishni talab qiladi.

Matnga bog‘liq bo‘lmagan holda tanish tizimlarida esa avvaldan yozib olingan so‘z yoki jumalardan foydalanish majburiyati qo‘yilmaydi. Bu yondashuv qoidasiga ko‘ra ishlovchi tizim ixtiyoriy uzunlik va mazmundagi so‘z va jumalarni qabul qilish imkoniga ega bo‘ladi[146]. Dasturiy ta‘minot tadbiri qilingin tizimning xavfsizligini yanada oshirish uchun ovozli identifikatsiyalashni boshqa autentifikatsiyalash yondashuvlari bilan birgalikda qo‘llash maqsadga muvofiq bo‘ladi. Mazkur ishda aynan ana shunday yondashuvdan foydalanib dasturiy ta‘minot ishlab chiqilgan va uning batafsil bayoni quyida keltirilgan.

Dasturiy ta‘minotni loyihalashtirish mavjud va taklif etilgan model, usul va algoritmlar asosida, python dasturlash tilida amalga oshirildi. Dasturiy ta‘minotning foydalanuvchi interfeysini ishlab chiqishda PyQt5 kutubxonasidan foydalanildi. Ro‘yxatdan o‘tkazilgan shaxslar ma‘lumotlarini saqlash uchun dasturiy ta‘minotda ma‘lumotlar bazasi yaratilgan bo‘lib, buning uchun SQLite ma‘lumotlar bazasini

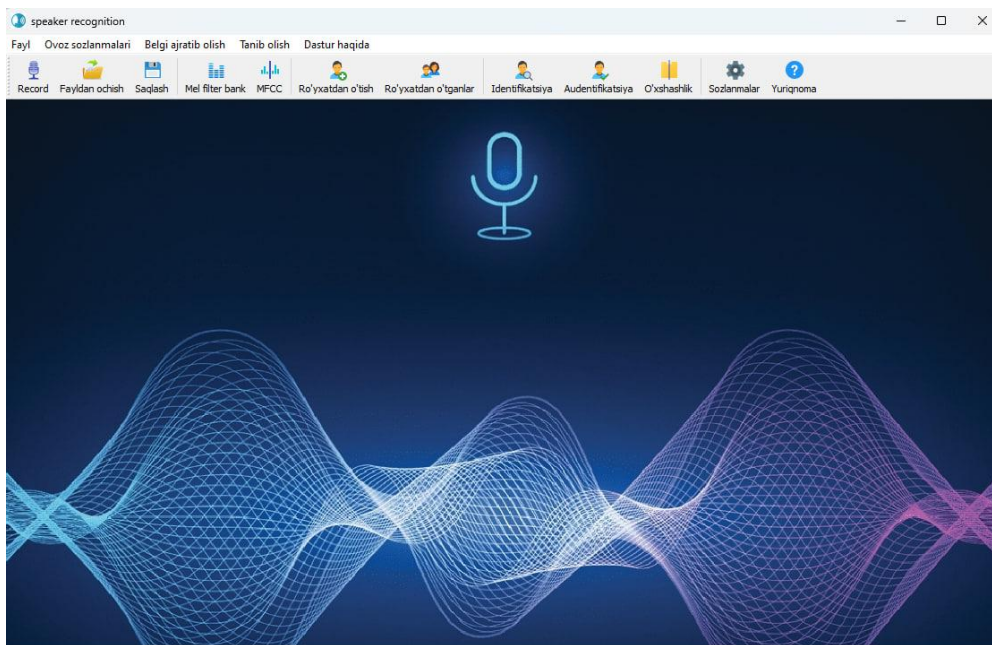
boshqarish tizimi foydalanildi. Bundan tashqari, chuqur neyron tarmoqlari bilan ishlash uchun PyTorch kutubxonasi qo‘llanilgan. Yaratilgan dasturiy ta‘minot 2 ta qismdan iborat bo‘lib, stasionar holatda va kliyent-server texnologiyasi asosida ishlashga mo‘ljallangan. Dasturiy ta‘minotning kliyent-server rejimiga oid kliyent qismi foydalanuvchi uchun qulay va sodda interfeysni taqdim etadi. Dasturiy ta‘minotning stasionar rejimga tegishli qismining tashkiliy tuzilmasi va ishlash tartibini quyida batafsil ko‘rib chiqiladi.

Dasturiy ta‘minot modulli dasturlash tamoyiliga asoslangan holda ishlab chiqilgan bo‘lib, u ko‘plab dasturiy modullarni o‘z ichiga oladi[147]. Unda ochiq kodli, yuqori darajali dasturlash tili qo‘llanilgan bo‘lib, ma‘lumotlarni qayta ishlash va tahlil qilish uchun ko‘p sonli kutubxonalar chaqirilgan. Dasturiy ta‘minotning tuzilmaviy sxemasini 4.2-rasmda ko‘rish mumkin.



4.2-rasm. Shaxsni ovozi asosida identifikatsiyalash dasturiy ta‘minotsining umumiy ishlash sxemasi

Yuqorida sanab o‘tilgan bosqichlarni amalga oshiruvchi dasturiy ta‘minotning foydalanuvchi interfeyslariga 4.3-4.7–rasmlarda misollar keltirilgan.



4.3-rasm. Dasturiy ta'minotning bosh oynasi

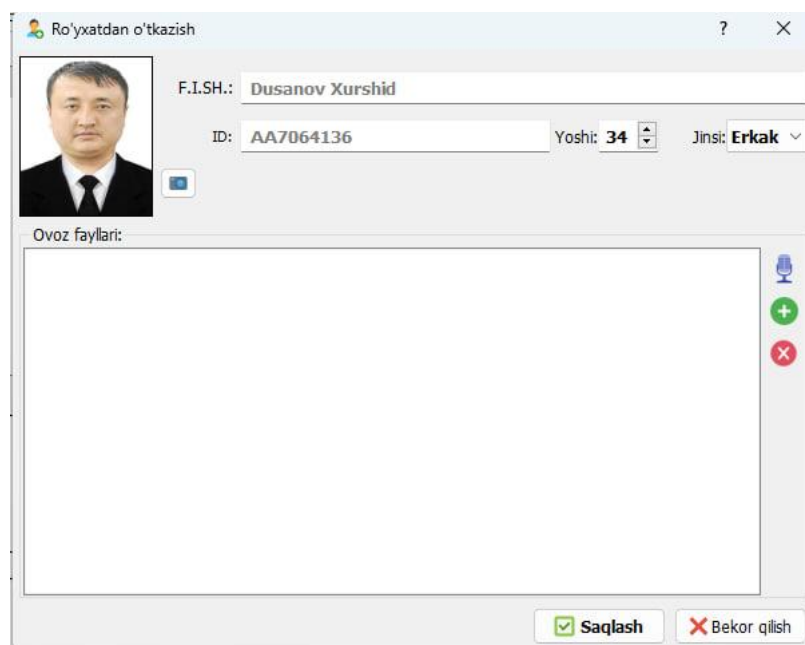
Dasturiy ta'minot bosh oynasi 4.3-rasmda keltirilgan bo'lib, unda dasturiy ta'minotni boshqarishga mo'ljallangan menyular qatorini ko'rish mumkin va u quyidagi menyulardan tashkil topgan:

- "Fayl" menyusida dasturiy ta'minotga nutq signallarini qabul qilish, qayta ishlangan nutq signallarini saqlash va dasturiy ta'minotdan chiqish amallarini bajarish mumkin;

- "Dastlabki ishlov" menyusida nutq signallariga dastlabki ishlov berish amallari bajariladi. Masalan: VAD orqali nutq signallarini jimlik va nutq bo'lmagan qismlardan tozalaydi[148];

- "Belgi ajratib olish" menyusida shaxs ovozini xarakterlovchi nutq signallaridan mel-masshtablangan filtr banklar va mel-chastotali kepstral koeffitsiyentlar usullari orqali belgilar to'plamini ajratib olish amallarini bajarish mumkin[149].

- "Dastur" menyusida tizim administratori tomonidan shaxs va foydalanuvchilarni ro'yxatdan o'tkazish, tahrirlash va o'chirish amallari bajariladi. Dastur menyusi quyidagi qismlardan iborat:



4.4-rasm. Shaxsni ro'yxatdan o'tkazish oynasi

- "Model qo'shish" bo'limida shaxslarni ro'yxatdan o'tkazish amalga oshiriladi. Ro'yxatdan o'tkazilgan shaxs ma'lumotlari va ovozi modeli ma'lumotlar bazasiga saqlanadi. 4.4-rasmda ro'yxatdan o'tkazish oynasi keltirilgan;

- "Modellar" bo'limida ro'yxatdan o'tkazilgan shaxslar ma'lumotlari bilan ishlash amalga oshiriladi. 4.5-rasmda ro'yxatdan o'tkazilgan shaxslar haqida ma'lumot beruvchi oyna keltirilgan;

Nº	Kod	F.I.S.H.	Yoshi	Jinsi	Rasmi	O'qitilgan	Sozlas
1	AB0304756	Abdullayev Sherzod	35	Erkak	Abdullayev_Sherzod.jpg	ha	<input checked="" type="checkbox"/>
2	AB1026115	Urinov Elmurod	33	Erkak	Elmurod(3x4)	ha	<input type="checkbox"/>
3	AA6262033	Xoshimov Axmadjon	32	Erkak	photo_2023-08-01_08-03-35.jpg	ha	<input type="checkbox"/>
4	KA0079845	Nurimov Paraxat	34	Erkak	Nurimov_Paraxat.jpg	ha	<input type="checkbox"/>
5	SA0000000	Notalish shaxs	0	-		yuq	<input type="checkbox"/>
6	AB 3574787	Asrayev Muxammadmullo	32	Erkak	11.jpg	yuq	<input type="checkbox"/>
7	AA7064136	Dusanov Xurshid	34	Erkak	Dusanov_Xurshid.jpg	ha	<input type="checkbox"/>
8	AB0823363	Sobirov Ro'zimboy	31	Erkak	Ruzimboy.jpg	ha	<input type="checkbox"/>
9	AA5382611	Jurayev Sherali	37	Erkak	photo_2023-08-01_08-02-45.jpg	yuq	<input type="checkbox"/>
10	AB6382348	Madaminjonov Akbar	28	Erkak	Akbar.jpg	ha	<input type="checkbox"/>
11	KA0077611	Jalelova Malika	26	Ayol	Malika.jpg	ha	<input type="checkbox"/>
12	AA3478767	Asrayev Muxammadmullo	32	Erkak	11.jpg	ha	<input type="checkbox"/>
13	AC8968129	Ibroximov Sanjar	30	Erkak	photo_2023-08-01_08-03-02.jpg	ha	<input type="checkbox"/>
14	KA0591614	Jalelov Kuanish	33	Erkak	Jalelov.jpg	ha	<input type="checkbox"/>

4.5-rasm. Ro'yxatdan o'tkazilgan shaxslar haqida ma'lumot beruvchi oyna

- “Identifikatsiya” bo‘limida shaxsni ovozi asosida identifikatsiyalash amalga oshiriladi;

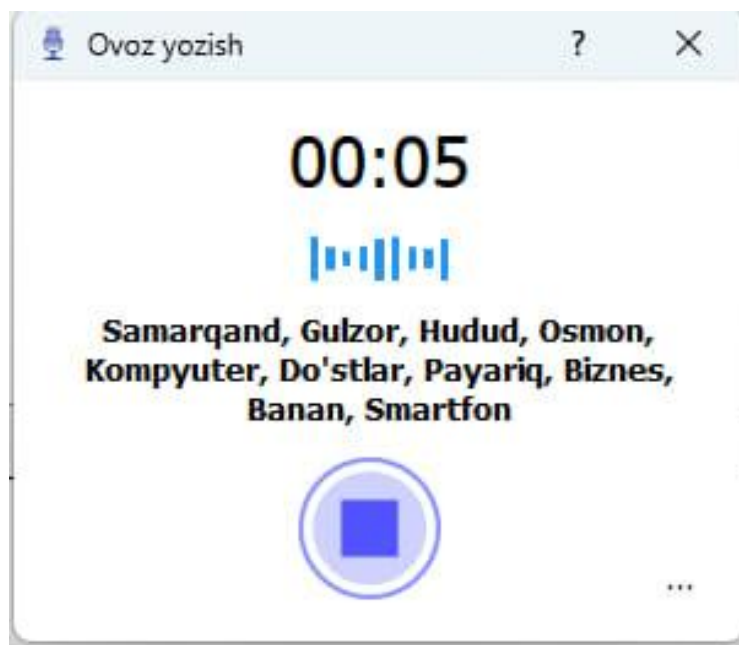
- “Autyentifikatsiya” bo‘limida shaxsni ovozi asosida verifikatsiyalash amalga oshiriladi;

- “Yo‘riqnoma” bo‘limi orqali dasturiy ta‘minotdan foydalanish va dastur haqida ma‘lumotlarni olish mumkin;

- “Sozlash” menyusi orqali dasturiy ta‘minot, mikrofon va modullar parametrlarini sozlash amallari bajariladi;

Bundan tashqari dasturiy majuada audio fayllarni eshitish, qayta ishlangan fayllarni qayta saqlash kabi qo‘shimcha funksiyalar mavjud.

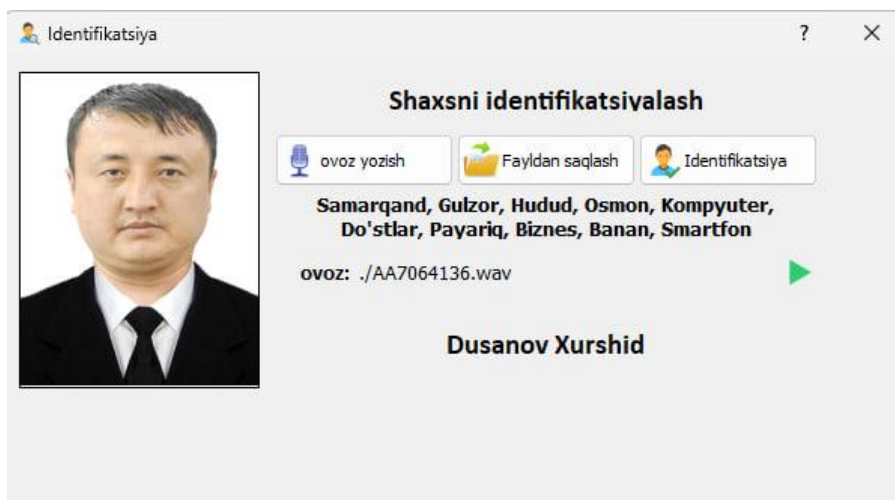
Shaxsni ovozi asosida identifikatsiyalashni amalga oshirish uchun “Identifikatsiya” bo‘limidan foydaniladi. Identifikatsiyalashni ikki xil usulda, ya’ni fayl va mikrofon orqali real vaqtda amalga oshirish mumkin. Ovoz yozib olish jarayoni 4.6-rasmda keltirilgan.



4.6-rasm. Ovoz yozib olish oynasi

Ovoz qayd etilgandan keyin, u audio fayl kurinishida dasturning asosiy oynasida joylashadi. “Identifikatsiya” tugmasini bosish orqali identifikatsiyalash jarayoni amalga oshiriladi va jadvalda audio faylga mos shaxs haqida ma‘lumot paydo bo‘ladi. 4.7-rasmda identifikatsiyalash jarayonini amalga oshirish natijasidan

olingan namuna va shaxs haqida to‘liq ma’lumot berish oynasi keltirilgan.



4.7-rasm. Identifikatsiyalashni amalga oshirish natijasi

4.3-§. Tajribaviy tadqiqot natijalari

Mazkur tadqiqot ishida shaxsni ovozi asosida identifikatsiyalash masalasini yechish uchun mavjud hamda ishda taklif etilgan model, usul va algoritmlarini tajriba-sinovdan o‘tkazish, tadqiq qilish orqali samaradorligi aniqlashda VCTK, TIMIT, VoxCeleb1, VoxCeleb2 kabi xalqaro ochiq turdagi nutq ma’lumotlar bazalari va ushbu tadqiqot ishi doirasida yaratilgan oddiy sharoitda yozilgan “NutqUzData” nutq ma’lumotlar bazasidan foydalanilgan. Ushbu nutq ma’lumotlar bazalari haqida to‘liq ma’lumot 4.2-jadvalda keltirilgan. Unga ko‘ra nutq signallarini diskretlash uchun taklif etilgan A1 va A2 algoritmlari samaradorligini baholash maqsadida NutqUzData nutqlar bazasida shaxsni ovozi asosida identifikatsiyalash masalasida tajribaviy tadqiqotlar o‘tkazilgan.

4.2-jadval

Nutq ma’lumotlar bazalari

Nutq ma’lumotlar bazasi	Shaxslar soni	Audio fayllar soni	O‘rtacha davomiyligi (sek.)
VCTK	110	44 000	4,2
TIMIT	630	6 300	3,6

VoxCeleb1	1 211	153 516	8,2
SITW	299	40 000	7,3
VoxCeleb2	5994	1 128 246	7,8
NutqUzData	6324	75888	4,6

Nutq signallaridan belgilar to‘plamini ajratib olish uchun LPCC, PLP, MFCC algoritmlaridan foydalanilgan bo‘lib, shaxsni ovozi asosida modelini yaratish uchun shablonli modellashtirish (Vektorli kvantlash-VQ) va stoxastik modellashtirish (Gauss aralashmasi modeli-GMM) usullari qo‘llanilgan. Modellarini solishtirish orqali qaror qabul qilish Evklid masofasi va k-yaqin qo‘shnilar usuli orqali amalga oshirilgan. Olingan natijalar 4.3-4.6 jadvallarda keltirilgan bo‘lib, unda delta diskretlash hamda maksimum-minimum diskretlash algoritmlarining qo‘llanilishdan oldingi va qo‘llanilgandan keyingi natijalari taqdim etilgan. 4.5-jadvalda identifikatsiyalash uchun sarflanadigan vaqtlarning o‘rtacha qiymatlari berilgan va diskretlash amalga oshirilmasdan oldin identifikatsiyalash uchun o‘rtacha 22,5 soniya vaqt sarflagan bo‘lsa, taklif etilgan diskretlash algoritmlari qo‘llanilgandan keyin o‘rtacha 2,9 soniya vaqt sarflaganini ko‘rish mumkin. Bu oldingi sarflagan vaqtdan qariyb 7,8 baravar kichikdir. Bundan tashqari, identifikatsiyalash uchun foydalaniladigan ma’lumotlar hajmining solishtirma tahlili 4.6-jadvalda keltirilgan bo‘lib, $m=4$ darajali diskretlash amalga oshirilgandan keyin ma’lumotlar hajmi qariyb 16 baravar kamaygan.

4.3-jadval

Aralash baza aniqlik darajasi foizlarda berilgan

Algoritm	VCTK + TIMIT	VoxCeleb1+ VoxCeleb2	TIMIT + VoxCeleb +SITW	NutqUzData
PLP+VQ	82,3	84,3	80,4	87,7
MFCC+VQ	83,2	86,2	83,1	88,8
PLP+GMM	84,8	85,8	82,7	90,1
MFCC+GMM	85,1	87,5	82,3	93,4

MSFB+GRU	94,5	95,2	90,6	96,3
MSFB+DNN	97,6	98,5	98,1	99,1

4.4-jadval

Delta diskretlash algoritmi asosida olingan natijalar. Aniqlik darajasi foizlarda berilgan. m-diskretlash darajasi

Algoritm	Baza nomi	Shaxslar soni	Tili	Aniqlik darajasi(%)
PLP+VQ	ICA based RASTA	18	ingliz	99,4
MFCC+VQ		44	Ingliz	97,2
PLP+GMM	TIMIT	Qisqa ovoz	Ingliz	79,7%
		Uzun ovoz	Ingliz	94,8%
MFCC+GMM	AiShell	400	Xitoy	96.7%
MSFB+GRU	AiShell	400	Xitoy	98.9
MSFB+DNN	VoxCeleb	7205	Ingliz	98.1

4.5-jadval

Minimum-maksimum diskretlash (A2) algoritmi asosida olingan natijalar. Aniqlik darajasi foizlarda berilgan.

L-freym uzunligi, K-belgilar soni.

Algoritm	Baza	energiyaga asoslangan VAD	nolni kesib o'tish chastotasiga asoslangan VAD
MFCC+GMM	93,4	92,7	90,9
MSFB+GRU	96,3	95,6	94,1
MSFB+DNN	99,1	97,9	96,3

Identifikatsiyalash uchun sarflangan o‘rtacha vaqtlarning solishtirma jadvali
(millisekundlarda).

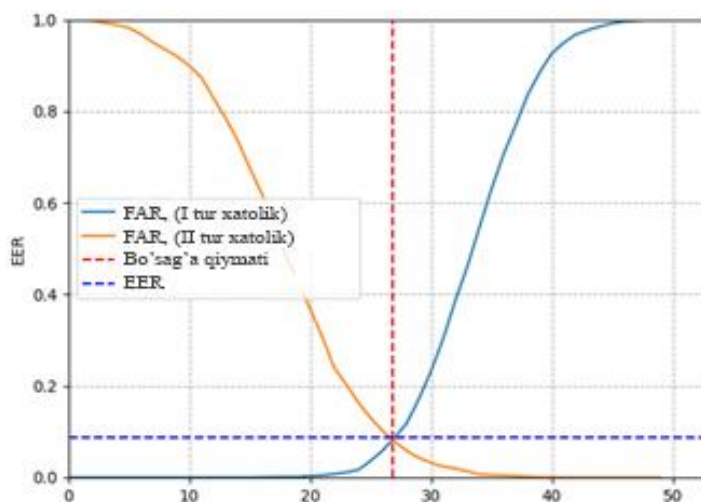
Algoritm	Baza	energiyaga asoslangan VAD	nolni kesib o‘tish chastotasiga asoslangan VAD
MFCC+GMM	54	48	43
MSFB+GRU	60	52	44
MSFB+DNN	80	65	48

Yuqoridagi natijalarni olishda foydalanilgan dasturiy va apparat vositalari ham muhim ahamiyatga ega. Bunda algoritm ish vaqtining qisqarishi hisoblash qurilmasining unumdorligiga ham bog‘liq. Dasturiy ta‘minotining tanib olish aniqligi va samaradorligi shaxs tomonidan turli muhitlarda yozib olingan ovoz fayllarining miqdori hamda apparat vositalari parametrlari bilan bog‘liq. Dasturiy ta‘minotni ishlab chiqishda va tadqiqotni sinovdan o‘tkazishda quyida keltirilgan parametrli hisoblash qurilmasi, texnologiyalaridan foydalanilgan:

- CPU: Intel i9 - 12900K
- GPU: 24GB 2x GeForce RTX3090Ti
- Memory: 128Gb
- Python version: 3.9.11
- Compiler: PyQt5, Pycharm
- OS: Windows 10
- PyTorch: 1.8
- CUDA: 11.8
- Model architecture: QuartzNet+ SpeakerNet
- Numpy: 1.18

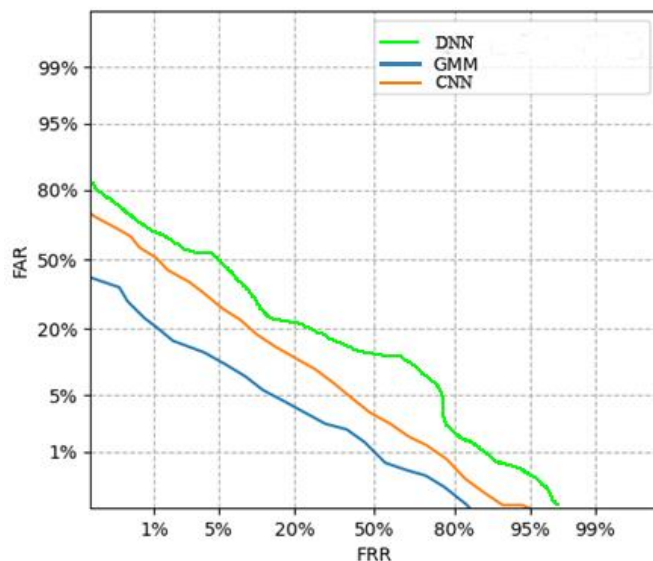
Ishni bajarish jarayonida chuqur neyron tarmoq asosida yaratilgan model samaradorligini baholash maqsadida Voxceleb1 bazadasida ochiq to‘plamli shaxsni ovozi asosida identifikatsiyalash masalasi uchun tajribaviy tadqiqotlar o‘tkazilgan.

Ochiq to'plamli identifikatsiyalashda tizim samaradorligi 1-tur (FAR) va 2-tur (FRR) xatoliklarning o'lchovi bilan belgilanadi. 4.8-rasmda chuqur neyron tarmoq asosida yaratilagn modelning 1-tur xatolik (FAR) va 2-tur xatolik (FRR) orasidagi bog'liqlik grafigi keltirilgan.



4.8-rasm. Chuqur neyron tarmoq asosida yaratilagn modelning 1-tur xatolik (FAR) va 2-tur xatolik (FRR) orasidagi bog'liklik grafigi

Voxceleb1(test) ma'lumotlar bazasida chuqur neyron tarmoq asosida yaratilgan modelning Gauss aralashma modeli (GMM) bilan taqqoslash uchun o'tkazilgan tajriba-sinovlari natijalari bo'yicha 1-tur va 2-tur xatoliklar bog'liqligining egri chiziqlari (DET) qurilgan. Olingan natijalardan 1-tur xatolikning oshishi, 2-tur xatolikning kamayishiga olib kelishini va aksincha ekanligini ko'rish mumkin. 1-tur va 2-tur xatoliklar orqali shaxsni ovozi asosida identifikatsiyalash tizimining umumiy samaradorligi belgilaniladi. 4.9-rasmda keltirilgan grafik orqali chuqur neyron tarmoq samaradorligini baholash mumkin.



4.9-rasm. Gauss aralashma modeli va chuqur neyron tarmoq modellari xatoliklarini baholash grafigi

4.7-jadvalda Voxceleb1(test) nutq ma'lumotlar bazasi bilan o'tkazilgan tajriba sinovi bo'yicha tanib olishning xatolik o'lchovi – ERR qiymatlari keltirilgan. Shaxsni ovozi asosida identifikatsiyalash tizimi aniqlik darajasi ochiq to'plamli identifikatsiyalashdagi xatolik o'lchovi –ERR yetarlicha kichik bo'lishi bilan belgilanadi. Olingan natijalar shaxsni ovozi asosida identifikatsiyalash tizimi uchun taklif etilgan chuqur neyron tarmoq modelining samarali ekanligini ko'rsatdi.

4.7-jadval

Gauss aralashma modeli va chuqur neyron tarmoq asosida yaratilgan modelining samaradorliklari solishtirma jadvali

Model	EER
Gauss aralashma modeli (GMM)	8,9%
GRU	4,3%
Chuqur neyron tarmoq modeli	3,2%

4.4-§. Dasturiy ta'minotning amaliy masalalarni yechishga tadbiqu

Hozirgi kunda inson uchun qulaylik va imkoniyatlar yaratish uchun sun'iy texnologiyalaridan foydalanish samarali usul bo'lib qolmoqda.

Kompyuter tarmoqlari rivojlanishi bilan tarmoq resurslari va obyektlarini himoya qilishning yanada mustahkam yondashuvlariga talab ortib bormoqda. Resurslar va ulardan foydalanuvchilar sonining oshishi esa firibgar va buzg'unchilar faollashuviga sabab bo'lmoqda. Ular g'arazli maqsadlarda foydalanuvchi akkauntlari, ilovalar, kirish-chiqish tizimlarini buzib kirishadi.

Ayni paytda buzg'unchilikning oldini olishga xizmat qiluvchi ko'plab yondashuvlari ishlab chiqilgan. Identifikatsiyalash va autentifikatsiyalash himoyalashning ana shunday yondashuvlaridan bo'lib, unda insonning turli biometrik ma'lumotlari keng qo'llaniladi. Jumladan, shaxsni yuz tasviri, barmoq izi, ovozi, qo'lyozmasi kabilarni shunday ma'lumotlar sirasiga kiritish mumkin. Bunday ma'lumotlar turlarining ko'pligi shaxsni tanib olish aniqligining oshishiga xizmat qiladi.

Tadqiqot ishi natijasida ishlab chiqilgan algoritmlar, taklif etilgan modellar asosida yaratilgan dasturiy ta'minot bank mijozlarini telefon orqali uning ovozi asosida identifikatsiyalash va ichki ishlar boshqarmasi faoliyatida ovozli xabarlarning kimga tegishi ekanligini aniqlash uchun identifikatsiyalash masalasini hal qilishda foydalanilgan.

Shaxsni ovozi asosida tanib olish dasturiy ta'minot Hamkorbank xodimlarining maxsus xonalarga kirish-chiqishini nazorat qilishda ikki omilli autentifikatsiyalash asosida qo'llanilgan bo'lib, u bino bo'ylab xavfsizlik darajasini oshirish imkonini bergan. Unga ko'ra dasturiy ta'minotga 50 nafar bank xodimining ovoz ma'lumotlari kiritilgan. Xodimlarni ro'yxatga olish bosqichida har bir xodim uchun dastur tomonidan takrorlanmaydigan 10 ta so'zli ketma-ketlik taqdim etilgan va identifikatsiyalash jarayonida xodimlardan taqdim etilgan so'zlardan kamida 5 ta ixtiyoriy so'zni aytish talab qilingan. Natijada telefon tarmog'i orqali bo'lgan murojaatlarni identifikatsiyalash va xodimlarni maxsus xonalarga kirish-chiqishini nazorat qilish uchun sarflanadigan vaqtni o'rtacha 10% ga qisqartirish imkonini yaratilgan (4.8-jadval).

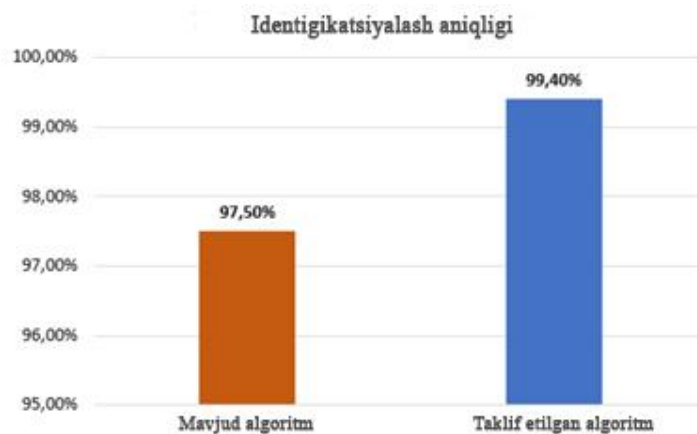
Hamkorbank xodimlarini maxsus xonalarga kirish-chiqishini nazorat
qilishda ikki omilli autentifikatsiyalash natijalari

Algoritm	Shaxslar soni	Ma'lumotlar hajmi (MB)	Identifikatsiyalash aniqligi	O'qitish uchun sarflangan vaqt (sek.)	Identifikatsiyalashga sarflangan vaqt (sek.)
Asosiy	50	113,8	95,7%	23,1	26,3
Delta diskretlash		28,5	98,1%	5,7	6,7
Maksimum minimum		28,5	98,4%	4,3	5,5

Hamkorbankning mijozlar bilan ishlash tizimi joriy qilingan bo'lib, bunda mijozni ikki omilli autentifikatsiya qilish amalga oshirilgan. Mijoz ovozi asosida uning shaxsi identifikatsiya qilingandan so'ng, o'ziga tegishli parol sifatida foydalaniladigan so'zlar ketma-ketligini talaffuz qilishi so'raladi. Bunda autentifikatsiyadan o'tish uchun mijoz 10 ta so'zlardan iborat ketma-ketlikni to'g'ri talaffuz qilishi talab etiladi. Shundan so'ng bank xodimi tomonidan dastur tavsiyasi asosida mijoz tomonidan so'ralgan shaxsiy profiliga kirish, o'tkazilgan operatsiyalarini ko'rish va hisob raqami to'g'risida ma'lumotlar beriladi yoki uni rad etishi mumkin bo'ladi. Bunda dasturiy ta'minotdan bank mijozlariga oid ma'lumotlar xavfsizligini ta'minlash uchun xodim tomonidan qabul qilinadigan qarorlarni qo'llab-quvvatlash tizimi sifatida foydalanilgan. Bundan tashqari mijozlar uchun masofadan turib foydalanish imkoni mavjud bo'lganligi tufayli qulay va ikki omilli, ya'ni ham biometrik-mijozning ovozi asosida, ham parolli-tegishli so'zlar ketma-ketligi asosida autentifikatsiyalash bosqichlari tashkil qilingan.

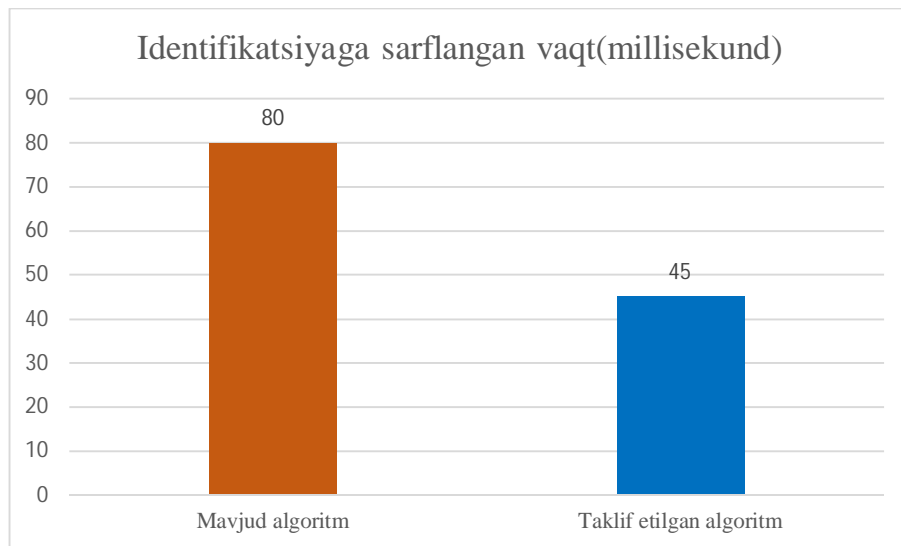
Dasturiy ta'minotning tadbqiq jarayonida jami 64 ta mijozlarning telefon orqali autentifikatsiya qilish va dastur tavsiyasi asosida shaxsiy profiliga kirish, o'tkazilgan

operatsiyalarini ko‘rish va hisob raqami to‘g‘risida ma’lumotlar olish ishlari amalga oshirildi. Bunda mijozlarning 64 nafari, ya’ni 9% qismida sog‘ligidagi muammo tufayli tizim tomonidan noto‘g‘ri rad etish (FRR) xatoligi kuzatildi, ammo noto‘g‘ri qabul qilish (FAR) xatoliklari kuzatilmadi. Parol-so‘zlar ketma-ketligini tanib olishda 3 ta mijozda aloqa sifati pastligi uchun xuddi shu xatolik kuzatildi. Qolgan 91% holatlar tizimdan foydalanish mumkinligini ko‘rsatdi. Bu amaliyot orqali mijozning bankka kelib-ketish uchun ketgan vaqt va pul resurslari tejaliishiga erishildi (4.10-rasm).



4.10-rasm. Xodimlarni maxsus xonalarga kirish-chiqishini nazorat qilishda identifikatsiyalash aniqligi

Identifikatsiyalash aniqligi va unga sarflanadigan o‘rtacha vaqt 4.9- va 4.11-rasmlarda grafik ko‘rinishida keltirilgan. Bunda xodimning ovozi asosida shaxsi identifikatsiyalanganidan so‘ng, o‘ziga tegishli parol sifatida foydalaniladigan so‘zlar ketma-ketligini talaffuz qilishi so‘raladi. Autentifikatsiyadan o‘tish uchun maxsus xonaga kirishda xodimdan 10 ta so‘zdan iborat ketma-ketlikni to‘g‘ri talaffuz qilishi talab etildi. Bunda xodimlardan 9% qismida sog‘ligidagi muammo tufayli tizim tomonidan noto‘g‘ri rad etish (False negative) xatoligi kuzatildi, noto‘g‘ri qabul qilish (False positive) xatoligi kuzatilmadi. Parol-so‘zlar ketma-ketligini tanib olishda xatolik kuzatilmadi.



4.11-rasm. Xodimlarni maxsus xonalarga kirish-chiqishini nazorat qilishda sarflangan o‘rtacha vaqt

Olingan natijalar mazkur soha mutaxassislari tomonidan yaxshi qabul qilindi. Xulosalar esa keyingi tadqiqotlarda nutqni matnga tezkor o‘tkazish va shaxsni tanib olish samaradorligini oshirishda informativ belgilarga nisbatan olib borilishi orqali tanib olish darajasini oshirish maqsadga muvofiq deb topildi. Taklif etilgan usul va algoritmlar orqali nutq texnologiyalarini rivojlantirish asosida amaliy tadqiqotlarda kelajak vazifalarini rejalashtirish ko‘plab soha mutaxassislari qulayliklar yaratdi va ortiqcha sarf-xarajatlarni kamaytirdi.

XULOSA

«Shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash algoritmlari hamda dasturiy ta'minotini ishlab chiqish» mavzusida olib borilgan monografiya tadqiqotining asosiy natijalari quyidagilardan iborat:

1. Shaxsni ovozi asosida identifikatsiyalash yondashuvlari afzalliklari, o'ziga xos jihatlari, yutuq va kamchiliklari aniqlandi. Olinishi zarur bo'lgan belgilar to'liq tahlil etildi. Bu foydalanuvchilarni avtomatik identifikatsiyalash va verifikatsiyalash tizimi himoyalanganligini oshirish, vaqt va manbalarni tejash orqali ruxsat berish jarayonida inson omilini aralashishiga yo'l qo'ymaydigan mukammal tizim yaratish imkonini beradi.

2. So'zlovchilarni tanib olish masalasi va texnologiyalari to'liq tadqiq etildi va unda yuzaga kelishi mumkin bo'lgan muammolar aniqlandi. Bu samaradorli yuqori bo'lgan shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash usul va algoritmlarini yaratishga asos bo'ladi.

3. Har bir shaxs ovozi o'ziga xosligi asoslandi. Bu so'zlovchini ovozi bo'yicha muvaffaqiyatli tanib olish imkoniyati mavjudligini tasdiqlab, shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash tizimini yaratishga asos bo'ladi.

4. Nutq signaliga dastlabki ishlov berish usullari tahlil etildi. Tahlil asosida mavjud usullardan samaradorligi yuqori bo'lgan usullar tanlab olindi va ular uchun algoritmlar ishlab chiqildi. Bu ishlash qobiliyati yuqori bo'lgan shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalash dasturiy taminotini yaratishga asos bo'ladi.

5. Nutq signallarini diskretlash algoritmlari takomillashtirildi va ular asosida belgilar to'plamini shakllantirish uchun algoritmlari ishlab chiqildi. Bu tezkorligi va xavfsizlik darajasi yuqori bo'lgan identifikatsiyalash tizimlarini ishlab chiqish imkonini beradi.

6. Kepstral silliqlik o'lchovini takomillashtirish orqali chegaralarni yuqori darajada aniqlash algoritmi taklif etildi. Bu nutqni tekis talaffuz qilish holatlarida o'tkazish ehtimolliki kamaytirish imkonini beradi.

7. Tanib olish masalasi uchun nutqli malumotlar bazasini shakllantirish

yondashuvi taklif etildi. Bu nutqli ma'lumotlar bazalarini rivojlatirishga xizmat qiladi.

8. Ovoz modelini shakllantirish uchun chuqur neyron tarmoq modeli taklif etildi. Bu tanib olish aniqligi yuqori bo'lgan tizimni yaratishga asos bo'ladi.

9. Takrorlanmaydigan so'zlar matritsasini hosil qilish algoritmi taklif etildi. Bu ikki omilli autentifikatsiya shartlarini bajarish orqali tizim ishonchliligini oshirish imkonini beradi.

10. Shaxsni ovozi asosida identifikatsiyalash va autentifikatsiyalashning mavjud va taklif etilgan algoritmlarini yagona interfeysga birlashtiruvchi dasturiy taminot ishlab chiqildi. Ishlab chiqilgan dasturiy ta'minot yopiq hudud va maxsus xonalarga kirish-chiqishni nazorat qilish, boshqarish, axborot va kompyuter tizimlarida foydalanuvchilarni tanib olish bilan bog'liq bir qator amaliy masalalarni yechish imkonini beradi.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI

1. Рабинер Л., Шафер Р. Цифровая обработка речевых сигналов. – М.: Радио и связь, 1981. – 496 с.
2. Абдураззақов Ф.Б, Юлдошев Ю.Ш., Нуримов П.Б. Нутқ сигналларига рақамли ишлов бериш назарияси ва технологияси, Республика илмий-техник анжуманининг маърузалари туплами, ТАТУ, Ташкент, 12-15 март 2019й. Б. 16-18
3. Лакин Г.Ф. Биометрия. 4-е изд., перераб. и доп.- М.: Высшая школа, 1990.-352 с.
4. Jain, A. K., & Ross, A. Introduction to Biometrics. Handbook of Biometrics, 1–22. doi:10.1007/978-0-387-71041-9_1
5. Матвеев, Ю.Н. Технологии биометрической идентификации личности по голосу и другим модальностям // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. — 2012. — № 3. — С. 46–61.
6. Кухарев, Г. А. Биометрические системы: методы и средства идентификации личности человека / Г. А. Кухарев. СПб.: Политехника, 2001. 240 с.
7. Сорокин, В. Н. Распознавание личности по голосу: анализ. обзор / Сорокин, В. В. Вьюгин, А. А. Танапыкпи // Информ, процессы. 2012. Т. 12, № 1. С. 1-30.
8. Мармарелис, П. Анализ физиологических систем (метод белого шума) / П. Мармарелис, В. Мармарелис. М.: Мир, 1981. 480 с.
9. Аграновский, А. В. Теоретические аспекты алгоритмов обработки и классификации сигналов / А. В. Аграновский, Д. А. Леднов. М.: Радио и связь, 2004. 164 с.
10. Уоссерман, Ф. Нейрокомпьютерная техника: теория и практика / Ф. Уоссерман. М.: Мир, 1992. 240 с.
11. Маматов Н.С., Нуримов П.Б., Самижонов А.Н. Нутқ сигналларида овоз фаоллигини аниқлаш алгоритмлари. «Ахборот коммуникация

технологиялари ва дастурий таъминот яратишда инновацион ғоялар»
Республика илмий-техник конференцияси 17-18 май 2021 йил.

12. D. Carlson, “Biometrics - Your Body as a Key,” 2014. [Online]. Available: <http://www.dynotech.com/articles/biometrics.shtml>. [Accessed: 04-May-2017].
13. K. Delac and M. Grgic, “A survey of biometric recognition methods,” 46th Int. Symp. Electron. Mar., 2004.
14. Нейросетевая защита персональных биометрических данных / В. И. Волчихин, А. И. Иванов, И. Г. Назаров и др. М.: Радиотехника, 2012. 158 с.
15. Айфичер, Э. Цифровая обработка сигналов: практический подход / Э. Айфичер, Б. Джервис. М.: Вильямс, 2004. 992 с.
16. J. Kaur and S. Kaur, “A Brief Review: Voice Biometric For Speaker Verification in Attendance Systems,” *Imp. J. Interdiscip. Res.*, vol. 2, no. 10, 2016.
17. D. Tran, M. Wagner, Y. W. Lau, and M. Gen, “Fuzzy methods for voice-based person authentication,” *IEEJ Trans. Electron. Inf. Syst.*, vol. 124, no. 10, pp. 1958–1963, 2004.
18. J.-L. Dugelay, J.-C. Junqua, C. Kotropoulos, R. Kuhn, F. Perronnin, and I. Pitas, “Recent advances in biometric person authentication,” in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, IEEE, 2002, vol. 4, p. 4060-4063.
19. P. Belin, S. Fecteau, and C. Bedard, “Thinking the voice: neural correlates of voice perception,” *Trends Cogn. Sci.*, vol. 8, no. 3, pp. 129–135, 2004.
20. H. Aronowitz, R. Hoory, J. Pelecanos, and D. Nahamoo, “New developments in voice biometrics for user authentication,” in *Twelfth Annual Conference of the International Speech Communication Association*, 2011.
21. P. Korshunov and S. Marcel, “Joint operation of voice biometrics and presentation attack detection,” *8th Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, ieeexplore.ieee.org, 2016.

22. S. Krawczyk and A. K. Jain, "Securing electronic medical records using biometric authentication," in *International Conference on Audio-and VideoBased Biometric Person Authentication*, 2005, pp. 1110– 1119.
23. L. M. Mazaira-Fernandez, A. Álvarez-Marquina, and P. Gómez-Vilda, "Improving Speaker Recognition by Biometric Voice Deconstruction," *Front. Bioeng. Biotechnol.*, vol. 3, Sep. 2015.
24. N. Scheffer, L. Ferrer, A. Lawson, Y. Lei, and M. McLaren, "Recent developments in voice biometrics: Robustness and high accuracy," in *International Conference onTechnologies for Homeland Security (HST)*, IEEE, 2013, pp. 447–452.
25. K. S. Bhogal, A. Rick, D. Kanevsky, and Pickover C A, "Using voice biometrics across virtual environments in association with an avatar's movements," US Patent, Google Patents, 2012.
26. N. Zhang, M. Paluri, Y. Taigman, R. Fergus, and L. Bourdev, "Beyond frontal faces:Improving person recognition using multiple cues," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 4804–4813.
27. M. Khitrov, "Talking passwords: voice biometrics for data access and security," *Biometric Technol. Today*, vol. 2, pp. 9–11, 2013.
28. M. Vatsa, R. Singh, and A. Noore, "Feature based RDWT watermarking for multimodal biometric system," *Image Vis. Comput.*, vol. 27, no. 3, pp. 293–304, 2009.
29. Singh N, Khan RA, Shree R. Applications of speaker recognition. *Procedia Eng* 2012;38(2012):3122–6. <https://doi.org/10.1016/j.proeng.2012.06.363>. ISSN 1877-7058.
30. Rosenberg, et al. L16: speaker recognition. Lect Slides 2007. Retrieved from, <http://research.cs.tamu.edu/prism/lectures/sp/116.pdf>.
31. Gish H, Schmidt M. Text-independent speaker identification. *Signal Process Mag*, IEEE 1994;11:18–32.
32. Gaikward SK, Gawali BW, Yannawar P. A review on speech recognition

- technique. *Int J Comput Appl* 2010;10(3):16–24.
33. Sturim, D.E., Campbell, W.M. & Reynolds, D.A. (2007). Classification methods for speaker recognition. *proceedings of speaker classification i: fundamentals, features, and methods*, pp. 278–97. 10.1007/978-3-540-74200-5_16.
 34. A. K. Jain, A. K., & Ross, A. Introduction to Biometrics. *Handbook of Biometrics*, 1–22. doi:10.1007/978-0-387-71041-9_1
 35. A. M. Bojamma, B. Nithya, and C. N. Prasad, “An Overview of Biometric System,” *Int. J. Comput. Sci. Eng. Inf. Technol. Res.* ISSN 2249-6831, vol. 3, no. 2, pp. 153–160, 2013.
 36. T. Reuters, “Global directory,” 2017. [Online]. Available: <https://blogs.thomsonreuters.com/answerson/biometrics-technology-convenience-data-privacy/>. [Accessed: 03-Oct 2018].
 37. D. Carlson, “Biometrics - Your Body as a Key,” 2014. [Online]. Available: <http://www.dynotech.com/articles/biometrics.shtml>. [Accessed: 04-May-2017].
 38. Sharma, A.M. (2019). Speaker recognition using machine learning techniques. (Master’s Projects). Retrieved from https://scholarworks.sjsu.edu/etd_projects/685.
 39. Biometrics: Authentication & Identification – 2020 Review. (2019). Retrieved from <https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/inspired/biometrics>.
 40. Zhang Z. Mechanics of Human Voice Production and Control. *J Acoust Soc Am* 2016;140(4):2614. <https://doi.org/10.1121/1.4964509>.
 41. Furui S. 40 years of progress in automatic speaker recognition. In: Tistarelli M, Nixon MS, editors. *Advances in biometrics*”. *ICB 2009. lecture notes in computer science*, 5558. Berlin, Heidelberg: Springer; 2009.
 42. Kershet J, Bengio S. Introduction. In: Margin L, Methods K, editors. *Automatic speech and speaker recognition*. West Sussex, United Kingdom: John Wiley & Sons Ltd; 2009.

43. Kikel, C. (2019, November 11). Difference between voice recognition and speech recognition. Retrieved from <https://www.totalvoicetech.com/difference-between-voice-recognition-and-speech-recognition/>.
44. Hanifa RM, Isa K, Mohamad S. Malay speech recognition for different ethnic speakers: an exploratory study. In: 2017 IEEE symposium on computer applications & industrial electronics (ISCAIE); 2017. p. 91–6.
45. Biometric Today. (2018). Retrieved from <https://biometrictoday.com/5-differences-between-voice-and-speech-recognition/>.
46. Sharma V, Bansal PK. A review on speaker recognition approaches and challenges. *Int J Eng Res Technol* 2013;2(5):1580–8.
47. Kaphungkui NK, Kandali AB. Text dependent speaker recognition with back propagation neural network. *Int J Eng Adv Technol (IJEAT)* 2019:1431–4.
48. T. Kohler, “The 2010 NIST Speaker Recognition Evaluation”, SLTC Newsletter, July 2010
49. N. Scheffer and R. Vogt, “On The Use of Speaker Superfactors For Speaker Recognition”, In Proc. of ICASSP2010, 2010, pp.4410-4413.
50. S. Kajarekar, N. Scheffer, M. Graciarena, E. Shriberg, A. Stolcke, L. Ferrer, and T. Bocklet, “The SRI NIST 2008 speaker recognition evaluation system”, In Proc. of the 2009 IEEE ICASSP2010, 2010, Washington, DC, USA, 4205-4208.
51. Campbell, W., Sturim, D., and Reynolds, D., “Support vector machines using GMM supervectors”, *IEEE Signal Processing Letters* 13, 5 (May 2006), 308–311.
52. D. Sturim, W. Campbell, Z. Karam, D. Reynolds, F. Richardson, “The MIT Lincoln Laboratory 2008 Speaker Recognition System”, *Interspeech 2009*, Brighton, UK, Sept. 6, 2009.
53. Schwartz, R., Shen, W., Campbell, J., Granville, R., *Measuring Typicality of Speech Features in American English Dialects: Towards Likelihood Ratios in Speaker Recognition Casework*, 5th European Academy of Forensics Science,

Glasgow, Scotland, Sept. 8, 2009.

54. N. Chen, W. Shen, J. Campbell, P. Torres-Carrasquillo, "Informative Dialect Recognition Using Context-Dependent Pronunciation Modeling", ICASSP 2011, Prague Czech Republic, May 2011.
55. C. Shaver and J. Acken, "Effects of Equipment Variation on Speaker Recognition Error Rates", IEEE, ICASSP2010, Dallas Texas, March 2010.
56. J. Ming, T. Hazen, J. Glass, D. Reynolds, "Robust Speaker Recognition in Noisy Conditions", IEEE Transactions on Audio, Speech, And Language Processing, VOL. 15, NO. 5, JULY 2007.
57. Zheng TF, Li L. Robustness-related issues in speaker recognition. Springer; 2017.
58. Singh N. A study on speech and speaker recognition technology and its challenges. In: Proceedings of national conference on information security challenges; 2014. p. 34–6.
59. Gomar MG. System and method for speaker recognition on mobile devices. Google Patents; 2015.
60. Muda L, Begam M, Elamvazuthi L. Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. J Comput 2020;2(3):138–43.
61. Tolba H. A high-performance text-independent speaker identification of Arabic speakers using a CHMM-based approach. Alexandr EngJ 2011 2011;50:43–7. <https://doi.org/10.1016/j.aej.2011.01.007>.
62. Sandouk, U. (2012). Speaker recognition: speaker diarization & identification. (Master's Thesis). Retrieved from https://studentnet.cs.manchester.ac.uk/resources/library/thesis_abstracts/ProjProgReptsMSc12/Sandouk-Ubai-ProgressReport.pdf.
63. Nakagawa S, Wang L, Ohtsuka S. Speaker identification and verification by combining MFCC and phase information. IEEE Trans Audio, Speech, Lang Process 2012;20:1085–95.

64. Doddington GR. Speaker recognition – identifying people by their voices. *Proc IEEE* 1985;73(11):1651–65.
65. Kotti M, Moschou V, Kotropoulos C. Speaker segmentation and clustering. *Signal Process* 2007;88(5):1091–124.
66. Kaphungkui NK, Kandali AB. Text dependent speaker recognition with back propagation neural network. *Int J Eng Adv Technol (IJEAT)* 2019;8(5):1431–4.
67. Suchitha TR, Bindu AT. Feature extraction using MFCC and classification using GMM. *Int J Sci Res Dev (IJSRD)* 2015;3(5):1278–83.
68. Gomar MG. System and method for speaker recognition on mobile devices. *Google Patents*; 2015.
69. Singh N, Agrawal A, Ahmad Khan R. A critical review on automatic speaker recognition. *Sci J Circ, Syst Signal Process* 2015;4(2):14–7.
70. Haohui (2019). Adversarial attacks in machine learning and how to defend against them. Notes from the keynote speaker speech by Professor Ling Liu at the 2019 IEEE big data conference. Retrieved from: <https://towardsdatascience.com/adversarial-attacks-in-machine-learning-and-how-to-defend-against-them- a2beed95f49c>.
71. Li Z, Shi C, Xie Y, Liu J, Yuan B, Chen Y. Practical adversarial attacks against speaker recognition system. In: *Proceedings of the 21st international workshop on mobile computing systems and applications (HotMobile '20)*, March 3–4; 2020. <https://doi.org/10.1145/3376897.3377856>. Austin, TX, USA. ACM, New York, NY, USA, 6 pages.
72. Yuan X, Chen Y, Zhao Y, Long Y, Liu X, Chen K, Zhang S, Huang H, Way X, Gunter CA. CommanderSong: a systematic approach for practical adversarial voice recognition. In: *Proceedings of the 27th USENIX security symposium*, August 15 – 17; 2018. p. 49–64. ISBN: 978-1-939133-04-5.
73. Zhang Z. Mechanics of Human Voice Production and Control. *J Acoust Soc Am* 2016;140(4):2614. <https://doi.org/10.1121/1.4964509>.
74. Афанасьев А.А. Создание интеллектуальных систем индивидуального

- низкоскоростного речевого кодирования / А.А. Афанасьев // Материалы 8-й Всероссийской научно-технической интернет-конференции. – Тула: Изд. ТулГУ, 2011. – С. 62–63.
75. Попов В.И. Основы сотовой связи стандарта GSM // В.И.Попов. – М.: Эко-Трендз, 2005. – 296 с
76. Огнев И.В., Парамонов П.А. Предварительная обработка речевого сигнала для построения базы произношений одиночных слов // Информационные средства и технологии: труды Международной научно-технической конференции (20 – 22 октября 2012 г.): в 3 т. – М.: МЭИ, 2012. – 1 т. – с. 53-58.
77. Юлдошев Ю.Ш. Нутқ сигналларини нормаллаштириш ва соҳаларини ажратиб олиш алгоритми // Муҳаммад ал – Хоразмий авлодлари, 3(13) /2020, Б. 24-30.
78. Mamatov N., Payazov M., Niyozmatova N., Yuldoshev Yu., Mamadalieva N. Definition of Speech Intelligibility of the Uzbek Language International Journal of Innovative Technology and Exploring Engineering (IJITEE)ISSN: 2278-3075, Volume-9 Issue-3, January 2020, DOI: 10.35940/ijitee.B6370.019320
79. Sukittanon S. Modulation-scale analysis for content identification // S. Sukittanon, L.E.Atlas, J.W. Pitton // IEEE Transactions On Signal Processing. – 2004. – Vol. 52. – No. 10. – P. 3023–3035.
80. Зилинберг А.Ю. Разработка и исследование временных и спектральных алгоритмов VAD (Voice Activity Detection) //А.Ю.Зилинберг, Ю.А.Корнеев // Российская школа-конференция «Мобильные системы передачи данных» / Зеленоград: МИЭТ, 2006. – С. 58–70.
81. Stejskal V. Empty speech pause detection algorithms' comparison // V.Stejskal, N.Bourbakis, A.Esposito // International Journal of Advanced Intelligence. – 2010. – Vol. 2. – No. 1. – P. 145–160.
82. Тупицин Г.С. Повышение качества закрытой текстонезависимой идентификации диктора в условия шумов с помощью бинарных масок /

- Г.С.Тупицин, М.В.Сагациян, С.А.Кравцов // доклад 15 всероссийской научно-практической конференции «Проблемы развития и применения средств противовоздушной обороны на современном этапе». – Ярославль, 2014. – С. 137–144.
83. Кравцов С.А. Алгоритм обнаружения речевой активности на основе моделей гауссовых смесей // С.А.Кравцов, Г.С.Тупицин // доклад 15 всероссийской научно-практической конференции «Проблемы развития и применения средств противовоздушной обороны на современном этапе». – Ярославль, 2014. – С. 39–44.
84. Hirsch, H. Noise estimation techniques for robust speech recognition // H.Hirsch, C.Ehrlicher // 1995 International Conference on Acoustics, Speech, and Signal Processing. – IEEE, 1995. – Vol. 1. – P. 153–156.
85. Jensen J. Spectral Magnitude Minimum Mean-Square Error Estimation Using Binary and Continuous Gain Functions // J.Jensen, R.C.Hendriks // IEEE Transactions on Audio, Speech, and Language Processing. – 2012. – Vol. 20. – № 1. – P. 92–102.
86. Li Y. On the optimality of ideal binary time-frequency masks // Y. Li, D. Wang // 2008 IEEE International Conference on Acoustics, Speech and Signal Processing. – IEEE, 2008. – P. 3501–3504.
87. Кравцов С.А. Алгоритм обнаружения речевой активности на основе моделей гауссовых смесей // С.А.Кравцов, Г.С.Тупицин // доклад 15 всероссийской научно-практической конференции «Проблемы развития и применения средств противовоздушной обороны на современном этапе». – Ярославль, 2014. – С. 39–44.
88. Kondo K. Subjective Quality Measurement of Speech: Signals and Communication Technology // K.Kondo. – Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. – P. 149.
89. Lu Y. Estimators of the Magnitude-Squared Spectrum and Methods for Incorporating SNR Uncertainty // Y. Lu, P.C. Loizou // IEEE Transactions on Audio, Speech, and Language Processing. – 2017. – Vol. 19. – № 5. – P.

1123–1137.

90. Huang X., Acero A. Spoken language processing: a guide to theory, algorithm, and system development, Prentice Hall. – 2001, 1008 p.
91. Николенко Л.А. Формирование признаков для дикторонезависимого распознавания фонем русского языка / Л.А.Николенко // Материалы Всероссийской научно-методической конференции «Повышение качества высшего профессионального образования» / Красноярск: ИПК СФУ, 2008. – Ч. 2. – С. 323–326.
92. Клименко Н.С. Разработка структуры текстонезависимой системы идентификации диктора //Н.С.Клименко // Искусственный интеллект. – 2017. – № 4. – С. 161–171.
93. Tan B.T. The use of wavelet transforms in phoneme recognition / B.T. Tan, M.Fu, A.Spray, P.Dermody // Proc. of International Conference on Spoken Language Processing (ICSLP). – 1996. – Vol. 4. – P. 2431–2434.
94. Dusan S. On the Relation Between Maximum Spectral Transition Positions and Phone Boundaries / S.Dusan, L.R.Rabiner // Proc. of ICSLP. – 2016.
95. K. Sreenivasa Rao, Sourjya Sarkar Robust Speaker Recognition in Noisy Environments, Springer, 2014.
96. Дорохин О.А. Старушко Д.Г, Федоров Е.Е. Сегментация речевого сигнала, Искусственный интеллект, 2000, № 3. Стр. 450-458.
97. Петров А.А. Выделение признаков речевого сигнала на основе вейвлет-анализа / А.А.Петров // Сборник трудов VI Всероссийской научно-практической конференции Молодежь и современные информационные технологии / Томск: ТПУ, 2008. – С. 135–136.
98. Yermolenko T.V. Segmentation of a speech signal with application of fast wavelet-transformation // T.V.Yermolenko // International Journal on Information Theories and Applications. – 2003. – Vol. 10. – No. 3. –P. 306–310.
99. Федоров В.М. Сегментация сигналов на основе дискретного вейвлет-преобразования // В.М.Федоров, П.Ю.Юрков // Информационное

- противодействие угрозам терроризма. – Таганрог: ЮФУ, 2009. – С. 138–146.
100. Вишнякова О.А. Автоматическая сегментация речевого сигнала на базе дискретного вейвлет-преобразования / О.А.Вишнякова, Д.Н.Лавров // Математические структуры и моделирование. / Омск: Ом. гос. ун-т, 2011. – Вып. 23. – С. 43–48.
 101. Basile P. The time-scale transform method as an instrument for phonetic analysis /P.Basile, F.Cutugno, P.Maturi, A.Piccialli // Visual representations of speech signals / Chicester, UK : John Wiley & Sons, 1993. – Chapter 13. – P. 169–174.
 102. Жуйков В.Я. Алгоритм автоматической классификации сегментов речи на основе автокорреляционных и энергетических характеристик // В.Я.Жуйков, Н.Н.Кузнецов, А.Н.Харченко // Электроника и связь. – Тематический выпуск «Электроника и нанотехнологии». – 2010. – № 5. – С. 83–89.
 103. Радзишевский А.Ю. Основы аналогового и цифрового звука // А.Ю.Радзишевский // Москва: Издательский дом «Вильямс», 2016. – 288 с.
 104. Ziolkо B. Wavelet method of speech segmentation // B.Ziolkо, S.Manandhar, R.Wilson, M.Ziolkо // Proceedings of 14th European Signal Processing Conference EUSIPCO // Florence, Italy. – 2006.
 105. Kronland-Martinet R. Analysis of sound patterns through wavelet transforms // R.Kronland-Martinet, J.Morlet, A.Grossmann // International Journal of Pattern Recognition and Artificial Intelligence. – 1987. – No. 1 (2). – P. 273–302.
 106. Gerhard D. Pitch extraction and fundamental frequency: history and current techniques // D.Gerhard. – Regina, Saskatchewan, Canada: University of Regina. – 2003. – 22 p.
 107. Азаров И.С. Алгоритм оценки мгновенной частоты основного тона речевого сигнала / Азаров И.С., Вашкевич М.И., Петровский А.А. //

Цифровая обработка сигналов. – 2012. – № 4. – С. 49–57.

108. Бочаров И.В. Распознавание речевых сигналов на основе корреляционного метода [Электронный ресурс] / И.В.Бочаров Д.Ю. Акатъев // Исследовано в России. – 2008. – № 6. – С. 1547–1557. – Режим доступа: <http://zhurnal.ape.relarn.ru/articles/2008/131.pdf>
109. Mamatov N., Samijonov A., Niyozmatova N., Yuldoshev Yu., Mamadalieva N. Algorithm For Identification Based On Voice International Journal of Innovative Technology and Exploring Engineering (IJITEE)ISSN: 2278-3075, Volume-9 Issue-2, December 2019, DOI: 10.35940/ijitee.B6368.129219.
110. Рахимов Н.О., Дусанов Х.Т. Интеллектуал тизимларда билимларни акс эттириш моделлари. “Замонавий тадқиқотлар, инновациялар, техника ва технологияларнинг долзарб муаммолари ва ривожланиш тенденциялари” мавзусидаги Республика миқёсидаги илмий-техник анжуман 2020 йил 10-11 апрель 393-395 бет.
111. Бабомурадов О. Ж., Бобоев Л. Б., Дусанов Х.Т. Матнинг кетма-кетлик модели. “Инновацион ёндашувлар илм-фан тараққиёти калити сифатида: ечимлар ва истиқболлар” мавзусидаги Республика илмий-амалий анжумани 2020 йил 8-10 октябр 196-201 бет.
112. Дусанов Х.Т., Қувондиқов Ж.Т. Маълумотларни интеллектуал таҳлиллашда тимсолларни аниқлаш масалалари. Современное состояние и перспективы применения цифровых технологий и искусственного интеллекта в управлении республиканская научно-техническая конференция 6-7 сентября 2021г. Стр. 189-192
113. Sang-Sik A. An improved statistical model-based VAD algorithm with an adaptive threshold // A.Sang-Sik, L.Yoon-Chang // Journal of The Chinese Institute of Engineers //Taipei. – 2006. – Vol. 29. – No. 5. – P. 783–789.
114. Шарий Т.В. Об одном методе автоматической сегментации речевых сигналов / Т.В.Шарий // Бионика интеллекта: научно-технический журнал. – 2009. – № 2 (71). – С. 61–65.

115. Mermelstein P. Automatic segmentation of speech into syllabic units // P. Mermelstein // Journal of the Acoustical Society of America. – 1975. – Vol. 58. – N.4. – pp. 880-883.
116. Н.С.Маматов, Ғ.Ғ.Пулатов, Х.Т. Дусанов Шахсни овози асосида таниб олиш усуллари. Рақамли трансформация ва сунъий интеллектимий журнали 2023/2. 90-95 бет
117. Mas Soro P.A. spectral estimator of vocal jitter/P.Mas Soro J.Schoentgen //Brussel, Belgium: Université libre de Bruxelles, 2017. – 108p.
118. Кодзасов С.В. Общая фонетика //С.В.Кодзасов, О.Ф.Кривнова. – М.: Рос. гос. гуманитар. ун-т, 2001. – 592 с.
119. Koestoer, N. P. Robust linear prediction analysis for low bit-rate speech coding //N. P. Koestoer //PhD dissertation.–Brisbane: Griffith University, 2002.–132 p.
120. Ахмад Х.М. Математические модели принятия решений в задачах распознавания говорящего / Х.М.Ахмад // Вестник ТГТУ. – 2008. – Т. 14. – № 1. – С. 19–32.
121. Н.С.Маматов, О.Ж.Бабомурадов, Х.Т. Дусанов Creation of a database for identifying persons by voice. International Journal of Theoretical and Applied Issues of Digital Technologies Vol. 5 No. 3 (2023), 25-32 b.
122. Х.Т. Дусанов The issue of recognizing a person based on his voice. Computer Science and Engineering Technology. Collection of materials of the international scientific and technical conference - Jizzakh: Jizzakh branch of UzMU, October 13, 2023. 140-143 bet.
123. Х.Т. Дусанов, А.А.Абдуваитов Алгоритмы интерактивной системы контроля достоверности текстов на основе словаря словоформ и баз данных. Машинасозликда замонавий материаллар, техника ва технологиялар халқаро илмий-техникавий анжумани 2016-йил 19-21 апрел.
124. Замалиев А.И., Кирпичников А.П., Ляшева С.А., Шлеймович М.П. Текстозависимая идентификация и верификация диктора по голосу в

- системе контроля и управления доступом // Вестник технологического университета. 2016. Т.19, №17. –С. 138 –143.
125. George Doddington, Speaker Recognition based on Idiolectal Differences between Speakers, Eurospeech 2001 – Scandinavia
126. N.A.Niyozmatova, N.S. Mamatov, X.T. Dusanov, B.N. Samijonov, A.N.Samijonov MFCC-GMM method for speaker identification by voice. International conference on ocean engineering technology and informatics 2023 (icoeti 2023) 4 - 6 december 2023.
127. N.S. Mamatov, O.J. Babomurodov, X.T. Dusanov Automatic speaker recognition through voice using deep neural networks. Journal of Computer Science Engineering and Information Technology Research (JCSEITR) ISSN (Print): 2250-2416; ISSN (Online): Applied Vol. 14, Issue 1; Jun 30 2024, 13-20.
128. X.T. Дусанов, Ж.Т.Кувондиқов Тимсолларни аниқловчи тизимларни яратиш масаласи ва уларни ечиш усуллари. Инновацион ёндашувлар илм-фан тараққиёти калити сифатида:ечимлар ва истикболлар. Республика миқёсидаги илмий-техник анжуман материаллари тўплами 2021 йил. 226-231 бет.
129. Н.С. Маматов, X.T. Дусанов Нутқ сигнали белгилар тўпламини шакллантиришнинг MFCC усули. Замонавий инновацион тадқиқотларнинг долзарб муаммолари ва ривожланиш тенденциялари: Ечимлар ва истикболлари. Республика миқёсидаги илмий-техник анжуман 2022 йил 13-14 май 179-182 бет.
130. X.T. Дусанов, А.А. Абдуваитов Системы обработки данных алгоритмы нейронечеткой слабоформализуемых процессов. Шестнадцатой международной конференции «информатика: проблемы, методология, технологии» 11-12 февраля 2016 года Воронеж стр. 28-31
131. Raximov N.O., Quvondiqov J.T., Daminova B.E., Dusanov X.T. As a mechanism that achieves the goal of decision management. International Conference on Information Science and Communications Technologies

- (ICISCT 2021) Tashkent – 2021. 3-5 november. Pp. 1-6.
132. Маматов Н.С., Абдуллаев Ш.Ш., Дусанов Х.Т., Самижонов Б.Н., Абдуллаев А.И. Нутқни автоматик таниб олишда QuartzNet модели. Ҳисоблаш ва амалий математика муаммолари Ўзбекистон журнали Тошкент-2022 й -Б. 71-76 бет.
 133. Н.С.Маматов, Ю.Ш.Юлдошев, Ш.Ш.Абдуллаев, Х.Т. Дусанов, А.Н.Самижонов Нутқ сигнали белгиларини шакллантириш. Ёнғин-портлаш хавфсизлиги илмий-амалий журнали. Тошкент-2022 й -Б.102-111 бет
 134. Н.С.Маматов, Х.Т. Дусанов IDEF ва UML билан биргаланганишлаб чиқариш вазифаларини моделлаш усуллари. Ёнғин-портлаш хавфсизлиги илмий-амалий журнали. Тошкент-2023 й -Б.338-345 бет.
 135. Н.С.Маматов, О.Ж.Бабомурадов, Х.Т. Дусанов Маълумотлар базасини моделлаштириш методологияси The scientific journal vehicles and roads, 2023 №3 85-90 бет.
 136. Ian Pointer, Programming PyTorch for Deep Learning, Creating and Deploying Deep Learning Applications, 2019.
 137. Joshua Willman, Beginning PyQt: A Hands-on Approach to GUI Programming, <https://www.apress.com/gp/book/9781484258569>, ISBN-10:1484258568, 2020.
 138. Дусанов Х.Т., Қувондиқов Ж.Т. Методы определения символов, основанных на прецедентов. Таълимда замонавий ахборот технологиялари мавзусидаги халқаро илмий-техникавий конференцияси 2021 йил 23 апрель 256-260 бет.
 139. N.S.Mamatov, O.J.Babomuradov, X.T. Dusanov Stages of speaker recognition based on voice. Proceedings of the 4th International Scientific and Practical Conference. Innovative development in the global science boston, Boston, USA, August 26-28, 2024. 122-128.
 140. Репалов, С. А. Разработка математических моделей и робастных алгоритмов идентификации дикторов по их речи: дис. канд. физ.-

- мат.наук / С. А. Реналов. Ростов н/Д.: Ростов, гос. ун-т, 2003.140 с.
141. Нейросетевая защита персональных биометрических данных / В. И. Волчихин, А. И. Иванов, И. Г. Назаров и др. М.: Радиотехника, 2012. 158 с.
 142. Иванов, А. Высоконадежная биометрическая аутентификация пользователя: последний дюйм первой мили / А. Иванов, А. Малыгин // Первая миля. 2007. № 2. С. 20-24.
 143. Burges, C. Tutorial on Support Vector Machines / C. Burges // Data Mining and Knowledge Discovery 2. 1998. P. 121-167.
 144. Журавлев, Ю. И. Об алгебраическом подходе к решению задач распознавания или классификации / Ю. И. Журавлев // Проблемы кибернетики. 1978. Т. 33. С. 5-68.
 145. MacKay, D. Information Theory, Inference, and Learning Algorithms / D. MacKay. Cambridge Univ. Press, 2003.
 146. Vapnik, V. N. The Nature of Statistical Learning Theory / V.N. Vapnik. New York : Springer, 1995.
 147. Бобамурадов О.Ж., Дусанов Х.Т., Туракулов О.Х. Маълумотларни интеллектуал таҳлил қилишда кластеризатсия масаласини ечиш дастури. Ўзбекистон республикаси адлия вазирлиги ҳузуридаги интеллектуал мулк агентлиги № DGU 15304 28.03.2022 йил
 148. Маматов Н.С., Нуримов П.Б., Дусанов Х.Т., Самижонов А.Н. Нутқ сигналларига дастлабки ишлов бериш дастури. Ўзбекистон республикаси адлия вазирлиги ҳузуридаги интеллектуал мулк агентлиги № DGU 23831 16.01.2023 йил
 149. Ниёзматова Н.А., Нуримов П.Б., Дусанов Х.Т., Самижонов А.Н. MFCC белгилар тўпламини шакллантириш алгоритми ва GMM модели ёрдамида шахсни таниб олиш дастури. Ўзбекистон республикаси адлия вазирлиги ҳузуридаги интеллектуал мулк агентлиги № DGU 23829 16.01.2023 йил

Belgilash va qisqartmalar

MSFB – Mel Scaled Filter Banks (Mel-masshtabli filtr banklar)

MFCC – Mel-frequency Cepstral Coefficient (Mel-chastotali kepral koeffisiyentlar)

LPC – Linear Predictive Coding (Chiziqli bashoratli kodi)

LPCC – Linear Prediction Cepstral Coefficient (Chiziqli bashoratli kepral koeffisiyentlar)

PLP – Perceptual Linear Predictive (Perseptual chiziqli bashorat)

FEC– Front End Clipping (to‘xtalishdan nutqga o‘tish chegarasini aniqlash xatoligi)

MSC – Mid Speech Clipping (nutqni to‘xtalish sifatida noto‘g‘ri aniqlash)

VQ – Vector Quantization (Vektorli kvantlash)

HMM – Hidden Markov Model (Yashirin Markov Modeli)

GMM – Gaussian Mixture Modeling (Gaus aralashmali modeli)

UBM –Universal Background Model (universal fon modeli)

JFA – Joint Factor Analysis (Birgalikta faktor tahlili)

PLDA – Probabilistic Linear Discriminant Analysis (Ihtimolli chiziqli diskriminat tahlil)

SVM – Support Vector Machine (Tayanch vektor mashinasi)

DFT – Discrete Fourier Transform (Diskret Fur‘ye almashtirishi)

FFT – Fast Fourier Transform (Tezkor Fur‘ye almashtirishi)

VAD – Voice Activity Detection (Nutq faolligini aniqlash)

EER – Equal Error Rate (Teng xatoliklar darajasi)

DET – Detection Error Trade-off (Aniqlash xatoligi bog‘liqlik egri chizig‘i)

FRR – False Reject Rate (Yolg‘ondan rad etish darajasi)

FAR – False Accept Rate (Yolg‘ondan qabul qilish darajasi)

IDR – Identification Rate (Identifikatsiyalash darajasi)

DNN – Deep Neural Network (Chuqur neyron tarmoq)

CNN – Convolutional Neural Network (O‘ramli neyron tarmoq)