

**NUTQ SIGNALIGA DASTLABKI ISHLOV BERISHNING  
TAKOMILLASHTIRILGAN USUL VA ALGORITMLARI****Nuritdinov Nurbek Davlataliyevich****Muhammad al-Xorazmiy nomidagi TATU Axborot texnologiyalarining  
dasturiy ta'minoti kafedrasi, katta o'qituvchisi****nuritdinovnurbek85@gmail.com****tel: +99893-914-18-28.****Nabiyev Elchinbek****Muhammad al-Xorazmiy nomidagi TATU 3-bosqich talabasi****nabiyevelchoinbek55@gmail.com****tel: 94-488-05-16.****<https://doi.org/10.5281/zenodo.20700524>****ANNOTATSIYA**

Maqolada nutqni avtomatik tanib olish (Automatic Speech Recognition, ASR) tizimlarida nutq signaliga dastlabki ishlov berishning asosiy bosqichlari — old kuchaytirish (pre-emfaza), signalni kadrlarga ajratish va oynalash, ovozli faollikni aniqlash (Voice Activity Detection, VAD), shovqindan tozalash hamda signal amplitudasini normallashtirish ko'rib chiqilgan. Spektral ayirish, Viner filtri va veyvlet o'zgartirishi asosidagi takomillashtirilgan shovqindan tozalash algoritmlari qiyosiy tahlil qilinib, ularning nutqni tanib olish aniqligiga ta'siri baholangan. Mel-chastotali keprstral koefitsiyentlarni (MFCC) hisoblashdan oldingi dastlabki ishlov berish bosqichining ahamiyati asoslab berilgan.

**Kalit so'zlar:** nutq signali, dastlabki ishlov berish, old kuchaytirish (pre-emfaza), oynalash, ovozli faollikni aniqlash (VAD), spektral ayirish, shovqindan tozalash, veyvlet o'zgartirishi, MFCC, nutqni tanib olish.

**АННОТАЦИЯ**

В статье рассмотрены основные этапы предварительной обработки речевого сигнала в системах автоматического распознавания речи (Automatic Speech Recognition, ASR): предыскажение (pre-emphasis), разбиение сигнала на кадры и оконное взвешивание, обнаружение голосовой активности (Voice Activity Detection, VAD), шумоподавление и нормализация амплитуды сигнала. Проведён сравнительный анализ усовершенствованных алгоритмов шумоподавления на основе спектрального вычитания, фильтра Винера и вейвлет-преобразования, оценено их влияние на точность распознавания речи. Обоснована значимость этапа предварительной обработки перед вычислением мел-частотных кепстральных коэффициентов (MFCC).

**Ключевые слова:** речевой сигнал, предварительная обработка, предыскажение (pre-emphasis), оконное взвешивание, обнаружение голосовой активности (VAD), спектральное вычитание, шумоподавление, вейвлет-преобразование, MFCC, распознавание речи.

**ABSTRACT**

The article discusses the main stages of speech signal preprocessing in automatic speech recognition (ASR) systems: pre-emphasis, framing and windowing, voice activity detection (VAD), noise reduction, and amplitude normalization. A comparative analysis of improved noise-reduction algorithms based on spectral subtraction, the Wiener filter, and the wavelet transform is carried out, and their impact on speech recognition accuracy is assessed. The

importance of the preprocessing stage prior to computing mel-frequency cepstral coefficients (MFCC) is substantiated.

**Keywords:** speech signal, preprocessing, pre-emphasis, windowing, voice activity detection (VAD), spectral subtraction, noise reduction, wavelet transform, MFCC, speech recognition.

## KIRISH

Nutq insonlar o'rtasidagi eng tabiiy va keng tarqalgan muloqot vositasi bo'lib, so'nggi yillarda nutqqa asoslangan texnologiyalar — ovozli yordamchilar, avtomatik diktovka tizimlari, chaqiruv markazlaridagi xizmatlar, imkoniyati cheklangan shaxslar uchun yordamchi dasturlar hamda ko'p tilli tarjima tizimlari kundalik hayotning ajralmas qismiga aylanmoqda. Bunday tizimlarning asosini nutqni avtomatik tanib olish (ASR) tashkil etadi.

Nutqni tanib olish tizimi, odatda, bir nechta bosqichdan iborat bo'ladi: signalni qabul qilish (raqamlash), dastlabki ishlov berish, xususiyatlarni ajratib olish, akustik va til modellari yordamida tahlil qilish hamda yakuniy qaror qabul qilish (dekodlash). Mazkur bosqichlar ichida dastlabki ishlov berish (preprocessing) alohida o'rin tutadi, chunki keyingi barcha bosqichlarning samaradorligi aynan shu bosqichda olingan signal sifatiga bog'liq.

Amaliyotda nutq signali ideal sharoitlarda emas, balki turli shovqinlar, akslanishlar (reverberatsiya), mikrofon xususiyatlaridagi farqlar va yozib olish sharoitlarining o'zgaruvchanligi mavjud bo'lgan muhitda qayd etiladi. Bu omillar tanib olish aniqligini sezilarli darajada pasaytiradi. Shu sababli dastlabki ishlov berish usul va algoritmlarini takomillashtirish — zamonaviy nutqni qayta ishlash tizimlarining bardoshlilikini (robustness) oshirishning eng dolzarb yo'nalishlaridan biridir.

Mazkur maqolaning maqsadi — nutq signaliga dastlabki ishlov berishning asosiy bosqichlarini ko'rib chiqish, shovqindan tozalashning takomillashtirilgan algoritmlarini (spektral ayirish, Viner filtri, veyvlet o'zgartirishi) qiyosiy tahlil qilish va ularning nutqni tanib olish aniqligiga ta'sirini baholashdan iborat.

## TADQIQOT USULLARI

Nutq signaliga dastlabki ishlov berish bir nechta ketma-ket bosqichdan tashkil topadi. Quyida ushbu bosqichlar va ularda qo'llaniladigan takomillashtirilgan usullar batafsil bayon etiladi.

**1. Old kuchaytirish (pre-emfaza).** Nutq signalining spektrida yuqori chastotali tashkil etuvchilar ovoz manbai (glottal) va lablardan nurlanish ta'sirida taxminan  $-6$  dB/oktava miqdorida susayadi. Bu nomutanosiblikni bartaraf etish uchun signalga birinchi tartibli yuqori chastotali raqamli filtr qo'llaniladi:

$$y[n] = x[n] - \alpha \cdot x[n-1], \quad \alpha \in [0,95; 0,97]$$

bu yerda  $x[n]$  — kirish signali,  $y[n]$  — kuchaytirilgan signal,  $\alpha$  — pre-emfaza koeffitsiyenti. Old kuchaytirish yuqori chastotalarni kuchaytirib, spektrni tekislaydi va keyingi spektral tahlilning aniqligini oshiradi.

**2. Signalni kadrlarga ajratish (framing).** Nutq signali nostatsionar bo'lsa-da, qisqa vaqt oraliqlarida (20–40 ms) kvazi-statsionar deb hisoblanadi. Shu sababli signal odatda 20–25 ms uzunlikdagi kadrlarga bo'linadi; qo'shni kadrlar orasida 10 ms siljish (overlap) qoldiriladi. Kadrlarning qisman ustma-ust tushishi axborotning yo'qolishining oldini oladi va vaqt bo'yicha silliq o'zgarishni ta'minlaydi.

**3. Oynalash (windowing).** Har bir kadr chekkalaridagi keskin uzilishlar spektral “oqib chiqish” (spectral leakage) hodisasini keltirib chiqaradi. Buni kamaytirish maqsadida kadrqa oyna funksiyasi qo'llaniladi. Eng ko'p ishlatiladigan Hemming (Hamming) oynasi quyidagicha aniqlanadi:

$$w[n] = 0,54 - 0,46 \cdot \cos(2\pi n / (N-1)), \quad 0 \leq n \leq N-1$$

bu yerda  $N$  — kadr uzunligi (namunalar soni). Oynalash kadr chekkalaridagi qiymatlarni nolga yaqinlashtirib, spektral tahlildagi buzilishlarni kamaytiradi.

**4. Ovozli faollikni aniqlash (VAD).** Nutq yozuvlarida sukut (jimlik) va faqat shovqindan iborat qismlar mavjud bo'ladi. Ovozli faollikni aniqlash algoritmlari nutq mavjud bo'lgan kadrlarni nutqsiz kadrlardan ajratadi. Buning uchun qisqa muddatli energiya, noldan o'tishlar soni (zero-crossing rate, ZCR) va spektral entropiya kabi belgilardan foydalaniladi. VAD ortiqcha hisoblashlarni kamaytiradi va shovqin spektrini baholash uchun ishonchli ma'lumot beradi.

**5. Shovqindan tozalashning takomillashtirilgan algoritmlari.** Bu bosqich dastlabki ishlov berishning eng muhim qismi hisoblanadi. Quyida uchta keng tarqalgan takomillashtirilgan yondashuv keltirilgan.

**a) Spektral ayirish (spectral subtraction).** Bu usulda nutqsiz oraliqlarda shovqinning amplituda spektri baholanadi va keyin har bir kadr spektridan ayriladi. Usul oddiy va hisoblash jihatidan tejamli bo'lsa-da, natijada “musiqiy shovqin” (musical noise) deb ataladigan qoldiq artefaktlar paydo bo'lishi mumkin. Bu kamchilikni ortiqcha ayirish koeffitsiyenti va spektral pol (floor) kiritish orqali yumshatish mumkin.

**b) Viner filtri (Wiener filter).** Viner filtri o'rtacha kvadratik xatolikni (MSE) minimallashtirish mezoniga asoslanadi va signal hamda shovqinning spektral zichligi nisbatidan kelib chiqib chastotaga bog'liq kuchaytirish koeffitsiyentini hisoblaydi. U spektral ayirishga nisbatan silliqroq natija beradi va “musiqiy shovqin”ni sezilarli kamaytiradi, biroq shovqin statsionar bo'lganda eng samarali ishlaydi.

**v) Veyvlet o'zgartirishi asosida tozalash (wavelet denoising).** Veyvlet o'zgartirishi signalni bir vaqtning o'zida vaqt va chastota bo'yicha tahlil qilish imkonini beradi, bu esa uni nostatsionar shovqinlar uchun ayniqsa qulay qiladi. Algoritm signalni veyvlet koeffitsiyentlariga yoyadi, so'ngra qattiq yoki yumshoq chegaralash (thresholding) yordamida shovqinga mos kichik koeffitsiyentlarni yo'qotadi va teskari o'zgartirish orqali tozalangan signalni tiklaydi. Bu yondashuv signalning o'tkir o'zgarishlarini (masalan, undosh tovushlar chegaralarini) yaxshi saqlaydi.

**6. Normallashtirish (normalization).** Turli yozib olish sharoitlari va mikrofonlar tufayli signal amplitudasi har xil bo'ladi. Amplituda bo'yicha normallashtirish signal qiymatlarini muayyan oraliqqa keltiradi. Kepstral o'rtachani ayirish (cepstral mean normalization, CMN) esa kanal ta'sirini kamaytirib, xususiyatlarning barqarorligini oshiradi.

## MUHOKAMA VA NATIJALAR

Yuqorida keltirilgan shovqindan tozalash algoritmlarini qiyosiy baholash uchun ularning asosiy ko'rsatkichlari — signal/shovqin nisbatining (SNR) yaxshilanishi, nutqni tanib olish aniqligi va amaliy kamchiliklari tahlil qilindi. 1-jadvalda adabiyotlarda keltirilgan tipik natijalar umumlashtirilgan.

*1-jadval. Shovqindan tozalash algoritmlarining qiyosiy tahlili*

| Algoritm                    | SNR<br>yaxshilanishi,<br>dB | Tanib olish<br>aniqligi, % | Asosiy kamchiligi                            |
|-----------------------------|-----------------------------|----------------------------|----------------------------------------------|
| Spektral ayirish            | 5–8                         | 78–82                      | “Musiqiy shovqin” artefaktlari               |
| Viner filtri                | 7–10                        | 82–86                      | Asosan statsionar shovqin<br>uchun samarali  |
| Veyvlet tozalash            | 9–12                        | 85–89                      | Hisoblash murakkabligi<br>yuqoriroq          |
| Gibrid (veyvlet +<br>Viner) | 11–14                       | 88–92                      | Parametrlarni moslashtirishni<br>talab etadi |

*Eslatma: jadvaldagi qiymatlar adabiyotlarda keltirilgan o‘rtacha natijalarni aks ettiradi va eksperiment sharoitlariga qarab o‘zgarishi mumkin.*

Tahlil natijalari shuni ko‘rsatadiki, spektral ayirish o‘zining soddaligi tufayli real vaqt rejimidagi tizimlar uchun qulay, biroq past SNR sharoitlarida qoldiq artefaktlar yuzaga keladi. Viner filtri silliqroq natija berib, statsionar shovqinli muhitlarda yuqori samaradorlik ko‘rsatadi. Veyvlet asosidagi tozalash esa nostatsionar shovqinlar (masalan, ko‘chadagi shovqin, gavjum joylardagi fon) uchun afzal, chunki u signalning vaqt-chastota tuzilishini yaxshi saqlaydi.

Eng yaxshi natijalar bir necha usulni birlashtiruvchi gibrid yondashuvlarda kuzatildi: veyvlet bilan dastlabki tozalash va undan keyin Viner filtri qo‘llanilganda SNR yaxshilanishi va tanib olish aniqligi alohida usullarga nisbatan yuqori bo‘ldi. Shuningdek, so‘nggi yillarda chuqur o‘rganishga (deep learning) asoslangan dastlabki ishlov berish — masalan, denoising autoencoder va rekurrent neyron tarmoqlardan foydalanish — murakkab shovqin sharoitlarida an‘anaviy usullardan ustun kelmoqda, biroq ular ko‘proq hisoblash resursi va o‘qitish ma’lumotlarini talab qiladi.

Umuman olganda, dastlabki ishlov berish bosqichida shovqindan tozalash, ovozli faollikni aniqlash va normallashtirishning to‘g‘ri tanlangan kombinatsiyasi MFCC kabi xususiyatlarning sifatini oshiradi va nutqni tanib olish tizimining xato darajasini (Word Error Rate, WER) sezilarli kamaytiradi.

### **XULOSA**

O‘tkazilgan tahlil natijasida quyidagi xulosalarga kelindi. Nutq signaliga dastlabki ishlov berish nutqni avtomatik tanib olish tizimining sifatini belgilovchi muhim bosqich bo‘lib, undagi har bir qadam — old kuchaytirish, kadrlarga ajratish, oynalash, ovozli faollikni aniqlash, shovqindan tozalash va normallashtirish — yakuniy aniqlikka bevosita ta’sir ko‘rsatadi.

Shovqindan tozalashning takomillashtirilgan algoritmlari orasida veyvlet o‘zgartirishi nostatsionar shovqinlar uchun, Viner filtri esa statsionar shovqinlar uchun eng samarali ekanligi aniqlandi. Eng yuqori ko‘rsatkichlarga bir necha usulni birlashtiruvchi gibrid yondashuvlar orqali erishildi. Shu bois shovqin turini hisobga olgan holda mos algoritmni tanlash amaliy jihatdan maqsadga muvofiqdir.

Kelgusi tadqiqotlarning istiqbolli yo‘nalishi sifatida chuqur o‘rganishga asoslangan adaptiv dastlabki ishlov berish usullarini ishlab chiqish hamda ularni an‘anaviy raqamli signal

qayta ishlash algoritmlari bilan birlashtirish ko'rsatilishi mumkin. Bu o'zbek tili uchun mo'ljallangan nutqni tanib olish tizimlarining bardoshlilikini va aniqligini yanada oshirish imkonini beradi.

### **Adabiyotlar, References, Литературы:**

1. Rabiner L. R., Schafer R. W. Theory and Applications of Digital Speech Processing. — Upper Saddle River: Pearson, 2011. — 1056 p.
2. Rabiner L. R., Juang B. H. Fundamentals of Speech Recognition. — Englewood Cliffs: Prentice Hall, 1993. — 507 p.
3. Huang X., Acero A., Hon H. W. Spoken Language Processing: A Guide to Theory, Algorithm, and System Development. — Upper Saddle River: Prentice Hall, 2001. — 980 p.
4. Boll S. F. Suppression of acoustic noise in speech using spectral subtraction // IEEE Transactions on Acoustics, Speech, and Signal Processing. — 1979. — Vol. 27, No. 2. — P. 113–120.
5. Ephraim Y., Malah D. Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator // IEEE Transactions on Acoustics, Speech, and Signal Processing. — 1984. — Vol. 32, No. 6. — P. 1109–1121.
6. Davis S., Mermelstein P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences // IEEE Transactions on Acoustics, Speech, and Signal Processing. — 1980. — Vol. 28, No. 4. — P. 357–366.
7. Donoho D. L. De-noising by soft-thresholding // IEEE Transactions on Information Theory. — 1995. — Vol. 41, No. 3. — P. 613–627.
8. Benesty J., Sondhi M. M., Huang Y. (eds.). Springer Handbook of Speech Processing. — Berlin: Springer, 2008. — 1176 p.
9. Loizou P. C. Speech Enhancement: Theory and Practice. — 2nd ed. — Boca Raton: CRC Press, 2013. — 711 p.
10. Vaseghi S. V. Advanced Digital Signal Processing and Noise Reduction. — 4th ed. — Chichester: John Wiley & Sons, 2008. — 544 p.