

UO'K: 004.93'14:159.942

NUTQ SIGNALIDAN SHOVQINLARNI KETMA-KET BARTARAF ETISH ALGORITMI

Niyozmatova Nilufar A'loxanovna¹ – texnika fanlari bo'yicha falsafa doktori (PhD), dotsent,
E-mail: n_niyozmatova@mail.ru

Yuldashev Yusuf Sheralievich² – texnika fanlari bo'yicha falsafa doktori (PhD), dotsent,
ORCID: 0000-0001-9708-2687, E-mail: yusuf_yuldoshev@mail.ru

Jalelov Qunish Moyatdinovich³ – doktorant (PhD), E-mail: kjalelov007@gmail.com

Samijonov Abdurashid Narzullo o'g'li³ – talaba, E-mail: an_samijonov@mail.ru

Abdullaeva Barno Mo'ydinjon qizi⁴ – doktorant (PhD), E-mail: shoxruxbarno@gmail.com

¹“Toshkent irrigatsiya va qishloq xo'jaligini mexanizatsiyalash muhandislari instituti” Milliy tadqiqot universiteti huzuridagi Fundamental va amaliy tadqiqotlar instituti,
Toshkent sh., O'zbekiston

²Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universiteti,
Toshkent sh., O'zbekiston

³“Toshkent irrigatsiya va qishloq xo'jaligini mexanizatsiyalash muhandislari instituti” Milliy tadqiqot universiteti, Toshkent sh., O'zbekiston

⁴Namangan davlat universiteti, Namangan sh., O'zbekiston

***Annotatsiya.** Nutq signalini shovqindan tozalash muammosi nutq tanish tizimlari, mobil aloqa qurilmalari va ovozli interfeyslar uchun dolzarb vazifa hisoblanadi. Mazkur tadqiqotda Self-Organizing Map (SOM) va Spectral Subtraction usullarini birlashtirgan yangi yondashuv taklif etilgan. Bu yondashuv shovqinli klasterlarni aniqlashda SOM neyron tarmog'idan foydalanib, energiya va chastota xususiyatlarini kombinatsiya qilish orqali nutqni samarali tozalaydi. Tadqiqot jarayonida shovqin taxmini uchun Minimum Statistics Noise Estimation usuli, xususiyatlarni adaptiv tanlash strategiyasi qo'llanilgan va shovqinni turli darajalarida (1% dan 25% gacha) tajribalar o'tkazilgan. PESQ metrikasi yordamida baholangan natijalar taklif etilgan yondashuv an'anaviy Veyvlet va Spectral Subtraction kabi usullarga nisbatan yuqori samaradorlikka egaligini ko'rsatgan. Ushbu yondashuvni afzalligi shovqinli klasterlarni aniqlashni takomillashtirilgan algoritmi va nutq tabiiyligini saqlab qolish uchun optimallashtirilgan post-processing bosqichida namoyon bo'lgan. Tadqiqot natijalari ko'rsatishicha, SOM va Spectral Subtraction usullari kombinatsiyasi shovqinli muhitda nutq signallarini tozalashda samarali yechim hisoblanadi.*

***Kalit so'zlar:** Self-Organizing Map, Spectral Subtraction, Mel-Frequency Cepstral Coefficients, Log-Mel spektrogrammasi, Perceptual Evaluation of Speech Quality.*

УДК: 004.93'14:159.942

АЛГОРИТМ ПОСЛЕДОВАТЕЛЬНОГО ПОДАВЛЕНИЯ ШУМОВ ИЗ РЕЧЕВОГО СИГНАЛА

Ниёзматова Нилюфар Алохановна¹ – доктор философии по техническим наукам (PhD), доцент

Юлдашев Юсуф Шералиевич² – доктор философии по техническим наукам (PhD), доцент

Джалелов Куаныш Моятдинович³ – докторант (PhD)

Самиджонов Абдурашид Нарзулло угли³ – студент

Абдуллаева Барно Муйдинжоновна⁴ – докторант (PhD)

¹Институт фундаментальных и прикладных исследований при Национальном исследовательском университете «Ташкентский институт инженеров ирригации и механизации сельского хозяйства», г. Ташкент, Узбекистон

²Ташкентский университет информационных технологий имени Мухаммада аль-Хоразми,
г. Ташкент, Узбекистон

³Национальный исследовательский университет «Ташкентский институт инженеров
ирригации и механизации сельского хозяйства», г. Ташкент, Узбекистон

⁴Наманганский государственный университет, г. Наманган, Узбекистон

Аннотация. Проблема очистки речевого сигнала от шума является актуальной задачей для систем распознавания речи, мобильных коммуникационных устройств и голосовых интерфейсов. В данном исследовании предлагается новый подход, объединяющий методы *Self-Organizing Map (SOM)* и *Spectral Subtraction*. Этот подход эффективно очищает речь, используя нейронную сеть *SOM* для идентификации шумных кластеров и комбинируя энергетические и частотные характеристики. В ходе исследования использовался метод *Minimum Statistics Noise Estimation* для оценки шума, а также стратегия адаптивного выбора признаков, и были проведены эксперименты при различных уровнях шума (от 1% до 25%). Оцененные с помощью метрики *PESQ* результаты показали, что предложенный подход превосходит по эффективности традиционные методы, такие как вейвлет-преобразование и *Spectral Subtraction*. Преимущество данного подхода заключается в усовершенствованном алгоритме идентификации шумных кластеров и оптимизированной постобработке, направленной на сохранение естественности речи. Результаты исследования показывают, что комбинация методов *SOM* и *Spectral Subtraction* является эффективным решением для очистки речевых сигналов в зашумленных условиях.

Ключевые слова: *Self-Organizing Map, Spectral Subtraction, Mel-Frequency Cepstral Coefficients, лог-мел спектрограмма, Perceptual Evaluation of Speech Quality.*

UDC: 004.93'14:159.942

SEQUENTIAL NOISE SUPPRESSION ALGORITHM FROM THE SPEECH SIGNAL

Niyozmatova, Nilufar Alokhanovna¹ – Doctor of Philosophy in Technical Sciences,
Associate Professor

Yuldashev, Yusuf Sheralievich² – Doctor of Philosophy in Technical Sciences,
Associate Professor

Jalelov, Kuanish Moyatdinovich³ – Doctoral Student (PhD)

Samijonov, Abdurashid Narzullo ugli³ – Student

Abdullaeva, Barno Muydinjon kizi⁴ – Doctoral Student (PhD)

¹Institute of Fundamental and Applied Research under the National Research University “Tashkent Institute of Irrigation and Agricultural Mechanization Engineers”, Tashkent city, Uzbekistan

²Tashkent University of Information Technologies named after Muhammad al-Khwarizmi,
Tashkent city, Uzbekistan

³National Research University “Tashkent Institute of Irrigation and Agricultural Mechanization Engineers”, Tashkent city, Uzbekistan

⁴Namangan State University, Namangan city, Uzbekistan

Abstract. The problem of denoising speech signals is a relevant task for speech recognition systems, mobile communication devices, and voice interfaces. This study proposes a new approach that combines *Self-Organizing Map (SOM)* and *Spectral Subtraction* methods. This approach effectively denoises speech by using the *SOM* neural network to detect noisy clusters and combining energy and frequency features. During the study, the *Minimum Statistics Noise Estimation* method was used for noise estimation, an adaptive feature selection strategy was applied, and experiments were conducted at various noise levels (from 1% to 25%). The results, evaluated using the *PESQ* metric, showed that the proposed approach outperforms traditional methods such as *Wavelet* and

Spectral Subtraction. The advantage of this approach lies in the improved algorithm for detecting noisy clusters and the optimized post-processing stage aimed at preserving speech naturalness. The research results indicate that the combination of the SOM and Spectral Subtraction methods is an effective solution for denoising speech signals in noisy environments.

Keywords: *Self-Organizing Map, Spectral Subtraction, Mel-Frequency Cepstral Coefficients, Log-Mel spectrogram, Perceptual Evaluation of Speech Quality.*

Kirish

Nutq signallarini shovqinlardan tozalash masalasi axborot texnologiyalari va raqamli signallarni qayta ishlash sohasida ustuvor yoʻnalishlardan biri sifatida eʼtirof etilgan [1]. Zamonaviy axborot-kommunikatsiya vositalarida, ayniqsa mobil qurilmalar, sunʼiy intellektga asoslangan assistentlar, maxsus qoʻllanmalar va tibbiy jihozlarda shovqinli muhitda nutqni aniq qayta tiklash muhim ahamiyat kasb etmoqda [2]. Statistik maʼlumotlar koʻrsatishicha, shovqinli muhitda nutqni tanish aniqligi 20-40% gacha kamayishi kuzatilgan, bu esa mazkur muammoni dolzarbligini yana bir bor tasdiqlaydi.

Nutq signallarini shovqinlardan tozalash uchun turli xil yondashuvlar ishlab chiqilgan boʻlib, ular orasida Veyner, Kalman filtri, veyvlet asosidagi algoritmlar va chuqur oʻqitishga asoslangan usullar mashhurdir. Biroq, mavjud yechimlarni turli darajadagi shovqinlarda barqaror natijalar bermasligi, hisoblash resurslarini yuqori talabiga javob berolmasligi, nutq tabiiyligini yoʻqotishi va real vaqt tizimlarida samarasizligi kabi kamchiliklari aniqlangan. Ayniqsa yuqori shovqin sharoitida (SNR<5dB) anʼanaviy usullar samaradorligi sezilarli darajada pasayib ketadi.

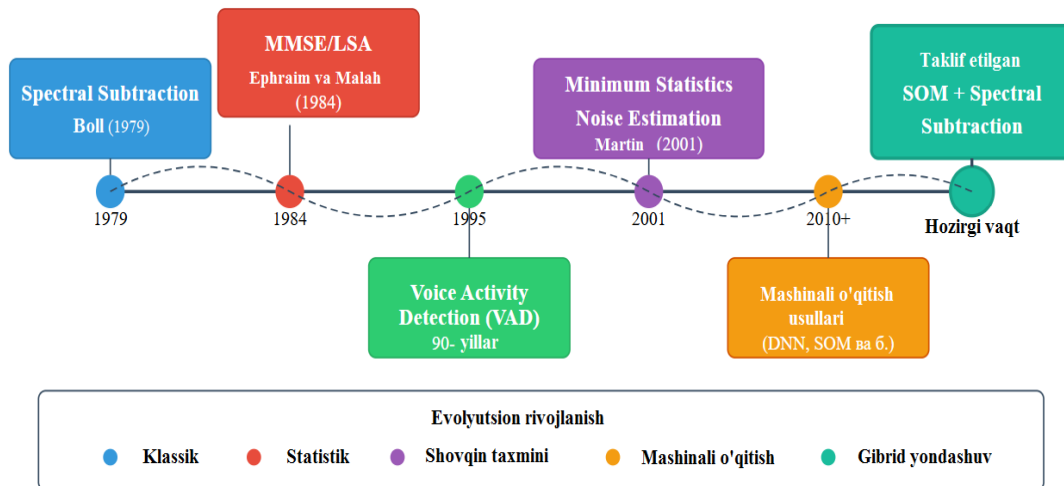
Mazkur tadqiqotda nutq signallarini shovqinlardan tozalashni innovatsion yondashuvi taklif etilgan boʻlib, u Self-Organizing Map (SOM) va Spectral Subtraction usullarini kuchli tomonlarini birlashtirgan. SOM oʻz-oʻzini tashkillashtiruvchi neyron tarmoqlar boʻlib, ular shovqin va nutq signallarini chastota-energetik xarakteristikalarini avtomatik ravishda aniqlash orqali nutq va shovqinni klasterlarga ajratish imkonini beradi. Spectral Subtraction usuli esa aniqlangan shovqin komponentlarini samarali olib tashlashga yordam beradi. Bu ikki usulni integratsiyasi orqali taklif etilgan algoritmlar turli shovqin darajalarida (1% dan 25% gacha) sinovlardan oʻtkazilgan. Tadqiqotni oʻziga xos jihati shundaki, Minimum Statistics Noise Estimation usuli yordamida shovqin taxmini takomillashtirilgan va shu orqali nostatsionar shovqinlarda ham ishonchli natijalar olingan. Shuningdek, shovqinni turli darajalariga mos keluvchi xususiyatlar toʻplamini adaptiv tanlash mexanizmi ishlab chiqilgan. Bunda past shovqin darajalarida MFCC (Mel-Frequency Cepstral Coefficients) xususiyatlari, yuqori shovqin darajalarida esa Log-Mel spektrogrammasi qoʻllanilgan. Bu yondashuv turli shovqin muhitlarida nutqni aniqroq ajratib olish imkonini beradi.

Adabiyotlar tahlili

Nutqni qayta ishlash va shovqinlardan tozalash sohasida olib borilgan ilmiy tadqiqotlar natijasida turli xil yondashuvlar ishlab chiqilgan boʻlib, ularni evolyusion rivojlanish tarixi quyidagi 1-rasmda koʻrsatilgan.

Bu rasmda keltirilgan usullar boʻyicha maʼlumotlar quyida batafsil bayon etilgan.

Spectral Subtraction usuli nutqni shovqindan tozalashni dastlabki va eng keng tarqalgan usullaridan biri sifatida eʼtirof etilgan. Boll tomonidan 1979-yili taklif qilingan bu usul shovqin spektrini signal spektridan ayirish tamoyiliga asoslangan [3]. Biroq, Spectral Subtraction "musiqiy shovqin" deb ataladigan artefaktlarni keltirib chiqarishi aniqlangan, bu esa nutq tabiiyligini sezilarli darajada pasaytiradi [4]. 1984-yili Ephraim va Malah tomonidan taklif qilingan statistik modellarga asoslangan shovqindan tozalash usullari Minimum Mean Square Error (MMSE) va Log-Spectral Amplitude (LSA) kabi yondashuvlarni oʻz ichiga oladi [5]. Bu usullar shovqin va nutqni ehtimollik taqsimotlarini hisobga olgan holda yaxshiroq natijalar beradi, biroq bu usulni baʼzi kamchilik jihatlari ham mavjud.



1-rasm. Nutqni shovqinlardan tozalash usullarini rivojlanish tarixi.

Shovqin taxmini nutqni tozalash samaradorligini belgilovchi muhim omillardan biri hisoblanadi. An'anaviy Voice Activity Detection (VAD) usullari ko'p yillar davomida qo'llanilgan bo'lsa-da, yuqori shovqin darajalarida nutq va shovqinni aniq farqlashda jiddiy to'siqlarga duch keladi [6]. Martin tomonidan 2001-yilda taklif qilingan Minimum Statistics Noise Estimation usuli shovqin taxmini uchun samarali muqobil yechim sifatida tan olingan [7]. Bu usul nutq signalining spektral komponentlari vaqt o'tishi bilan kamida bir marta minimumga yetishini taxmin qiladi. Shu sababli, har bir chastota diapazoni uchun vaqt oralig'ida kuzatilgan minimal qiymatlarni kuzatib, shovqin taxmini amalga oshiriladi.

SOM algoritmi Kohonen tomonidan taklif qilingan bo'lib, ma'lumotlarni klasterlash va vizualizatsiya qilish uchun keng qo'llaniladigan neyron tarmoq arxitekturasi hisoblanadi [8]. SOMning nutqni qayta ishlashda qo'llanilishi bo'yicha so'nggi yillarda qator tadqiqotlar amalga oshirilgan [9]. Odatda SOM shovqinli klasterlarni aniqlashda energiyaga asoslangan mezonlardan foydalanadi, biroq bu yondashuv nutqning past energiyali qismlarini (masalan, pauzalar yoki jarangsiz tovushlar) noto'g'ri shovqin sifatida belgilashiga olib kelishi mumkin [10]. Shuning uchun, tadqiqot ishida SOM va Spectral Subtraction usullarini birlashtirish g'oyasi taklif etilgan.

Yuqorida keltirib o'tilgan usullarni yutuq va kamchilik jihatlari quyidagi 1-jadvalda keltirilgan.

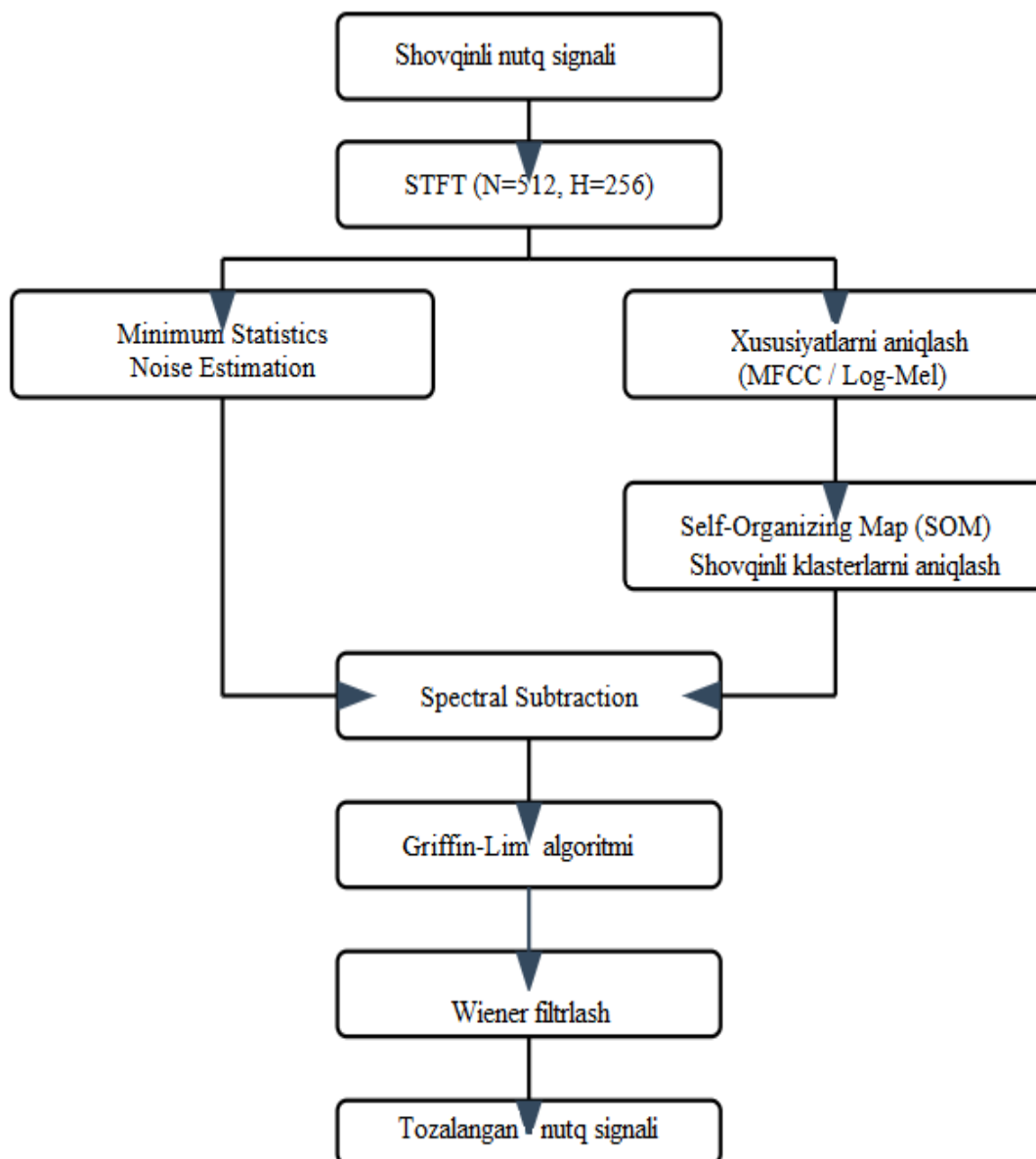
1-jadval

Nutqni shovqinlardan tozalash usullarini taqqoslash natijalari

Usul nomi	Mualliflar	Yutuqlar	Kamchiliklar	Samaradorlik
Spectral Subtraction [1]	Boll (1979)	- Soddaligi - Hisoblash samaradorligi Keng qo'llash doirasi	- "Musiqiy shovqin" artefaktlari - Yuqori SNR talabi	O'rtacha (PESQ: 2.0-2.5)
MMSE/LSA [3]	Ephraim va Malah (1984)	- Statistik optimallik - Kamroq aptefaktlar	- Yuqori hisoblash talablari - Shovqin taxminiga bog'liqligi	Yaxshi (PESQ: 2.3-2.8)
Voice Activity Detection (VAD) [4]	Turli Mualliflar	- Soddaligi - Real vaqt qo'llanilishi	- Yuqori shovqinda ishlamasligi - Ozgaruvchan shovqinda muammolar	O'rtacha -past (PESQ: 1.8-2.3)
Minimum Statistics Noise Estimation [5]	Martin (2001)	- Yuqori shovqinda barqarorlik - Nostatsionar shovqinda yaxshi	- Ba'zi shovqin turlarida xatolar - Kechikish	Yaxshi (PESQ: 2.4-3.0)

Tadqiqot metodologiyasi

Ishda taklif etilgan nutq signalini shovqinlardan tozalash algoritmi quyidagi sxemada ko'rsatilgan qadamlarni o'z ichiga oladi (2-rasm).



2-rasm. Nutq signalini shovqinlardan tozalash algoritmi.

Tadqiqot doirasida olib borilgan ishlar tavsifi quyida batafsil keltirib o'tilgan.

1. Ma'lumotlar to'plami. Tadqiqotda TIMIT ma'lumotlar bazasidan olingan toza nutq namunalaridan foydalanilgan bo'lib, ushbu ma'lumotlar 16 kHz chastotada yozib olingan va har bir audio faylni davomiyligi o'rtacha 3-5 soniyani tashkil etgan. Bunda shovqinli signal $x(t) = s(t) + \lambda n(t)$ formulasi asosida yaratilgan, bunda $s(t)$ -toza nutq signali, $n(t)$ - oq shovqin, λ - shovqin darajasi bo'lib, 1%, 5%, 10%, 15%, 20%, 25% ko'rinishida o'zgartirilgan.

2. Shovqin taxmini. Shovqin taxmini uchun an'anaviy VAD usuliga alternativa sifatida Minimum Statistics Noise Estimation usuli qo'llanilgan. Ushbu usul signalni magnituda spektrogrammasida har bir chastota diapazoni uchun minimal qiymatlarni topish orqali shovqin taxminini aniqlaydi.

$|S(f, t)|$ magnituda spektrogramma qisqa muddatli Fure almashtirish yordamida hisoblangan:

$$S(f, t) = \text{STFT}(x(t)), |S(f, t)| = \sqrt{\text{Re}(S(f, t))^2 + \text{Im}(S(f, t))^2},$$

bu yerda STFT - qisqa muddatli Fure almashtirish bo'lib, ishda STFT uzunligi 512, ramkalar orasidagi qadam 256 deb olingan.

Shovqin taxmini quyidagi qadamlar orqali amalga oshirilgan:

1) Har bir f chastota uchun minimal qiymatlar topilgan:

$$\hat{N}(f, t) = \min_{t' \in [t-W/2, t+W/2]} |S(f, t')|, W = 20$$

2) Silliqlashtirish qo'llanilgan:

$$\hat{N}_{\text{smoothed}}(f, t) = 0.9 \cdot \hat{N}_{\text{smoothed}}(f, t - 1) + 0.1 \cdot \hat{N}(f, t),$$

bu yerda 0.9- silliqashtirish koeffitsienti sifatida olingan.

3) O'rtacha shovqin taxmini hisoblangan:

$$\hat{N}(f) = \frac{1}{T} \sum_{t=1}^T \hat{N}_{\text{smoothed}}(f, t)$$

Minimum Statistics usulida 20 ramkadan iborat oyna ($W = 20$) foydalanilgan bo'lib, shovqin taxminini silliqashtirish uchun 0,9 koeffitsientli eksponensial silliqashtirish qo'llanilgan. Ushbu yondashuv yuqori shovqin darajalarida barqaror shovqin taxminini ta'minlaydi.

3. Xususiyatlarni tanlash. SOM algoritmi uchun xususiyatlar shovqin darajasiga qarab tanlangan. Past shovqin darajalarida ($\text{SNR} \geq 5$) MFCC xususiyatlari foydalanilgan, chunki ular past shovqinli muhitda nutqni yaxshi ifodalaydi. Yuqori shovqin darajalarida ($\text{SNR} < 5$) esa Log-Mel spektrogrammasidan foydalanilgan, chunki u shovqinli muhitda nutq xususiyatlarini yaxshiroq saqlaydi.

4. Shovqinli klasterlarni aniqlash (SOM). Ishda taklif etilgan usulni arxitekturasi va ishlash tamoyili 3-rasmda tasvirlangan. Bunda shovqinli nutq signalidan ajratib olingan xususiyatlar SOM neyron tarmog'i orqali klasterlangan, so'ngra energiya va chastota mezonlari asosida shovqinli klasterlar aniqlanib, Spectral Subtraction protsessi yordamida shovqin olib tashlangan.

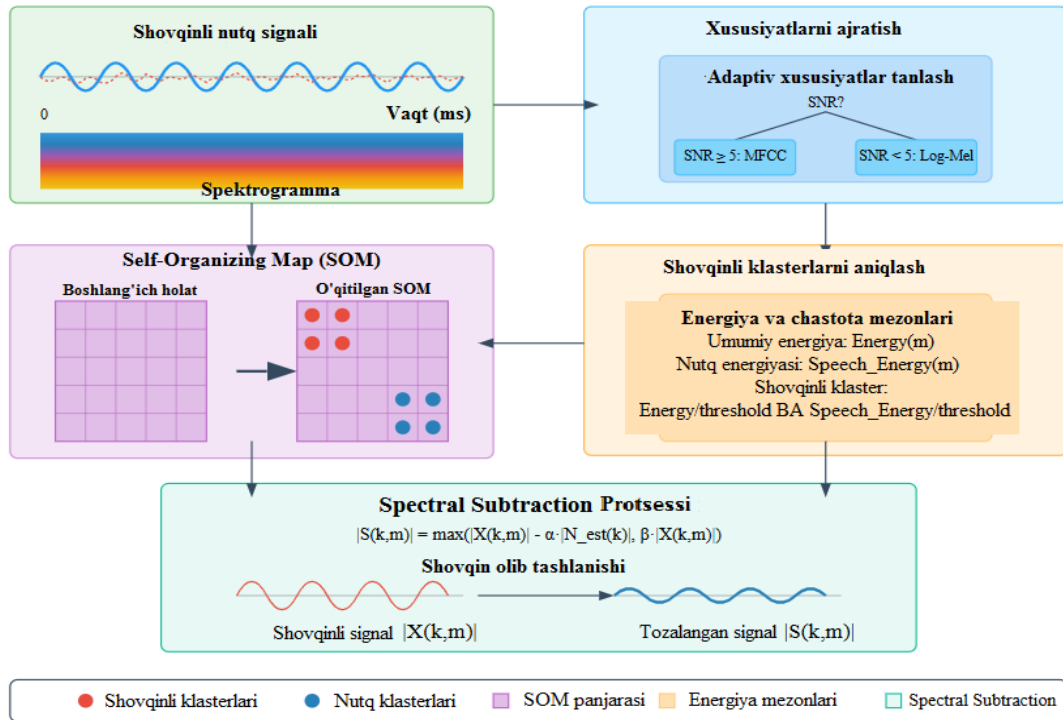
Shovqinli klasterlarni aniqlash uchun SOM ni o'qitish qoidasi quyidagi ko'rinishda ifodalangan:

$$w_j(t + 1) = w_j(t) + \eta(t) h_{j,j^*}(t) (x - w_j(t))$$

bu yerda, $\eta(t)$ - o'qitish tezligi, $w_j(t)$ - neyronni vazn vektori, x - kirish vektori, $h_{j,j^*}(t)$ - Gaussian qo'shnilik funksiyasi.

SOMni (G_x, G_y) panjara o'lchami, $\eta(t)$ o'quv tezligi va σ parametri shovqin darajasiga qarab dinamik ravishda tanlangan:

$$\begin{aligned} \text{SNR} < -5: & (15,15), \sigma = 4.0, \eta = 0.8; \\ -5 \leq \text{SNR} < 0: & (12,12), \sigma = 3.0, \eta = 0.6; \\ 0 \leq \text{SNR} < 5: & (10,10), \sigma = 2.0, \eta = 0.4; \\ \text{SNR} \geq 5: & (5,5), \sigma = 1.0, \eta = 0.2; \end{aligned}$$



3-rasm. SOM va shovqin klasterlash jarayoni.

Shovqinli klasterlarni aniqlash uchun SOM ni o'qitish qoidasi quyidagi ko'rinishda ifodalangan:

$$w_j(t+1) = w_j(t) + \eta(t)h_{j,j^*}(t)(x - w_j(t)),$$

bu yerda, $\eta(t)$ - o'qitish tezligi, $w_j(t)$ - neyronni vazn vektori, x - kirish vektori, $h_{j,j^*}(t)$ — Gaussian qo'shnilik funksiyasi.

SOMni (G_x, G_y) panjara o'lchami, $\eta(t)$ o'quv tezligi va σ parametri shovqin darajasiga qarab dinamik ravishda tanlangan:

$$\begin{aligned} SNR < -5: & (15,15), \sigma = 4.0, \eta = 0.8; \\ -5 \leq SNR < 0: & (12,12), \sigma = 3.0, \eta = 0.6; \\ 0 \leq SNR < 5: & (10,10), \sigma = 2.0, \eta = 0.4; \\ SNR \geq 5: & (5,5), \sigma = 1.0, \eta = 0.2. \end{aligned}$$

SOM 5000 iteratsiya davomida o'qitilgan. O'qitilgandan so'ng, shovqinli klasterlarni aniqlash uchun quyidagi mezonlar qo'llanilgan:

$$\text{Umumiy energiya: } E(t) = RMS(x(t)) = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$$

$$\text{Energiya chegarasi (threshold): } E_{th} = \text{percentile}(E(t), p)$$

Nutq chastota diapazonidagi energiya:

$$E_{\text{speech}}(t) = \sum_{f \in [300, 3400]} |S(f, t)|, E_{\text{speech,th}} = \text{percentile}(E_{\text{speech}}(t), 40)$$

Ramka shovqinli deb hisoblangan, agar quyidagilar bajarilsa:

$$E(t) = E_{th} \cap E_{\text{speech}}(t) < E_{\text{speech,th}}$$

Shovqinli klasterlarni aniqlashda energiyaga asoslangan mezon chastota xususiyatlari bilan birlashtirilgan. Har bir ramka uchun energiya va nutq chastota diapazoni (300 Hz dan 3400 Hz gacha) hisobga olinib, ushbu diapazondagi energiya alohida hisoblangan (speech_energy). Agar ramkani umumiy energiyasi va nutq diapazonidagi energiyasi threshold-dan past bo'lsa, u shovqinli klaster

sifatida belgilangan. Ushbu yondashuv nutqni past energiyali qismlarini (pauzalarni) shovqin sifatida noto‘g‘ri belgilash ehtimolini kamaytirgan.

5. Shovqindan tozalash. Shovqindan tozalash uchun Spectral Subtraction usuli qo‘llanilgan. Shovqinli klasterlar uchun:

$$|S_{\text{denoised}}(f, t)| = \max(|S(f, t)|) - \alpha \cdot \hat{N}(f), \beta \cdot \hat{N}(f),$$

bu yerda α - shovqin pasaytirish koeffitsiyenti, β - spektral daraja.

$$\begin{aligned} SNR < -5: \alpha &= 3.5, \beta = 0.005; \\ -5 \leq SNR < 0: \alpha &= 3.0, \beta = 0.003; \\ 0 \leq SNR < 5: \alpha &= 2.5, \beta = 0.002; \\ SNR \geq 5: \alpha &= 1.5, \beta = 0.001; \end{aligned}$$

Nutqli klasterlar uchun yumshoq shovqin pasaytirish qo‘llanilgan:

$$|S_{\text{cleaned}}(f, t)| = 0.8 \cdot |S(f, t)|, \quad f \notin [300, 3400] \text{ uchun } 0.5 \cdot |S(f, t)|$$

6. Signalni qayta tiklash. Tozalangan magnituda spektrogrammasidan signalni qayta tiklash uchun Griffin-Lim algoritmi foydalanilgan:

$$x_{\text{reconstructed}}(t) = \text{Griffin-Lim}(|S_{\text{denoised}}(f, t)|, \text{phase}, n_{\text{iter}} = 1000, \text{hop_length} = 256)$$

Bu jarayonda 1000 iteratsiya ($n_{\text{iter}} = 1000$) va $\text{hop_length} = 256$ parametri qo‘llanilgan.

Griffin-Lim algoritmi fazani qayta tiklashda yaxshi natijalar bergan. Biroq, signal tabiiyligini yanada oshirish uchun Post-processing jarayoni amalga oshirilgan.

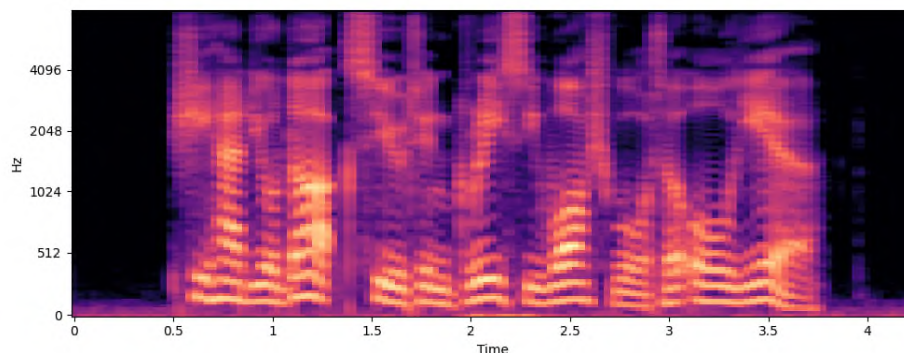
7. Post-processing. Ushbu jarayonda Veyner filtrlash qo‘llanilgan:

$$x_{\text{final}}(t) = \text{Wiener}(x_{\text{reconstructed}}(t), \text{size} = 7).$$

8. Baholash mezonlari. Taklif etilgan usul samaradorligi $\text{PESQ}(s(t), x_{\text{final}}(t))$ metrikasi yordamida baholangan. PESQ nutq sifatini inson eshitish tizimiga mos ravishda baholaydi va -0,5 dan 4,5 gacha bo‘lgan diapazonda qiymatlar beradi, bunda yuqori qiymatlar yaxshi sifatni ko‘rsatadi. Har bir shovqin darajasi uchun shovqinli signal, SOM bilan tozalangan signal va Spectral Subtraction bilan tozalangan signalni PESQ qiymatlari hisoblangan.

Natijalar

Mazkur tadqiqotda nutq signalini shovqindan tozalash uchun Veyvlet almashtirish, Spectral Subtraction, SOM, SOM bilan Spectral Subtraction usullarini birlashtirgan yondashuv asl nutq signali (4-rasm) va uni turli shovqin darajalarida (1%, 5%, 10%, 15%, 20%, 25% oq shovqin) sinovlardan o‘tkazilgan. Natijalar PESQ metrikasi yordamida baholangan (5-rasm).



4-rasm. Asl nutq signali.

Shovqin foizi

Shovqinli nutq signali spektrogramma

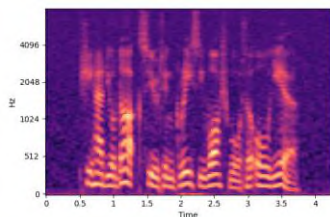
Veyvlet almashtirish bilan tozalangan

Spectral Subtraction bilan tozalangan spektrogramma

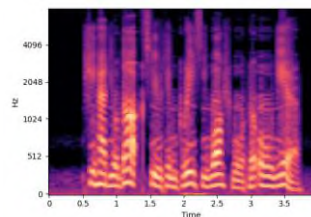
SOM bilan tozalangan

SOM bilan va Spectral Subtraction bilan tozalangan spektrogramma

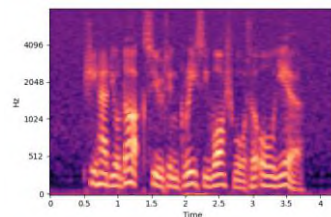
1%



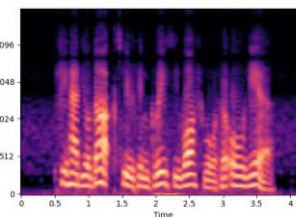
1,484082937



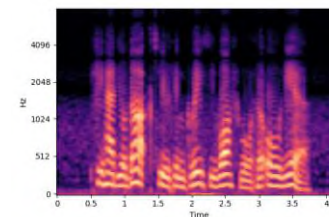
1,693255544



1,605126381

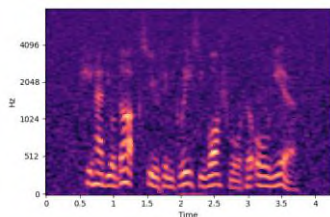


2,412364244

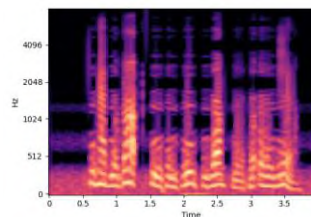


2,425626993

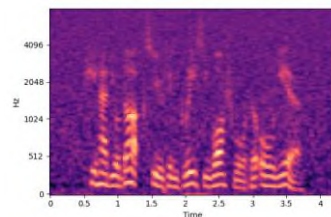
5%



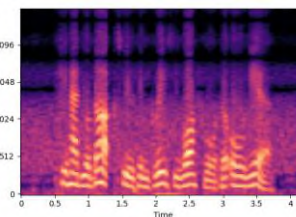
1,066443324



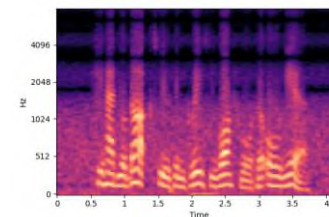
1,157373428



1,080800295

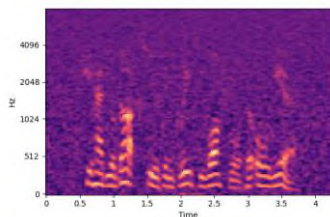


1,588101625

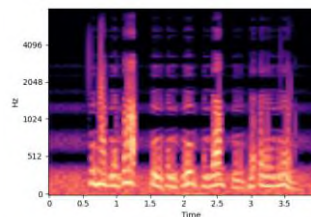


1,614011526

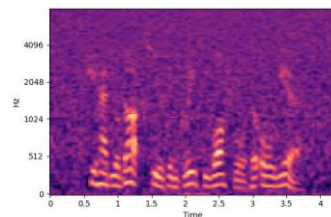
10%



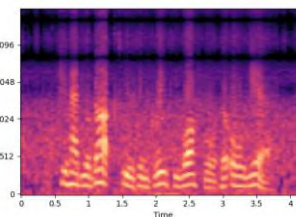
1,040338159



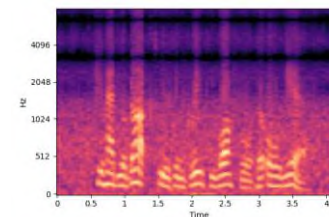
1,090646505



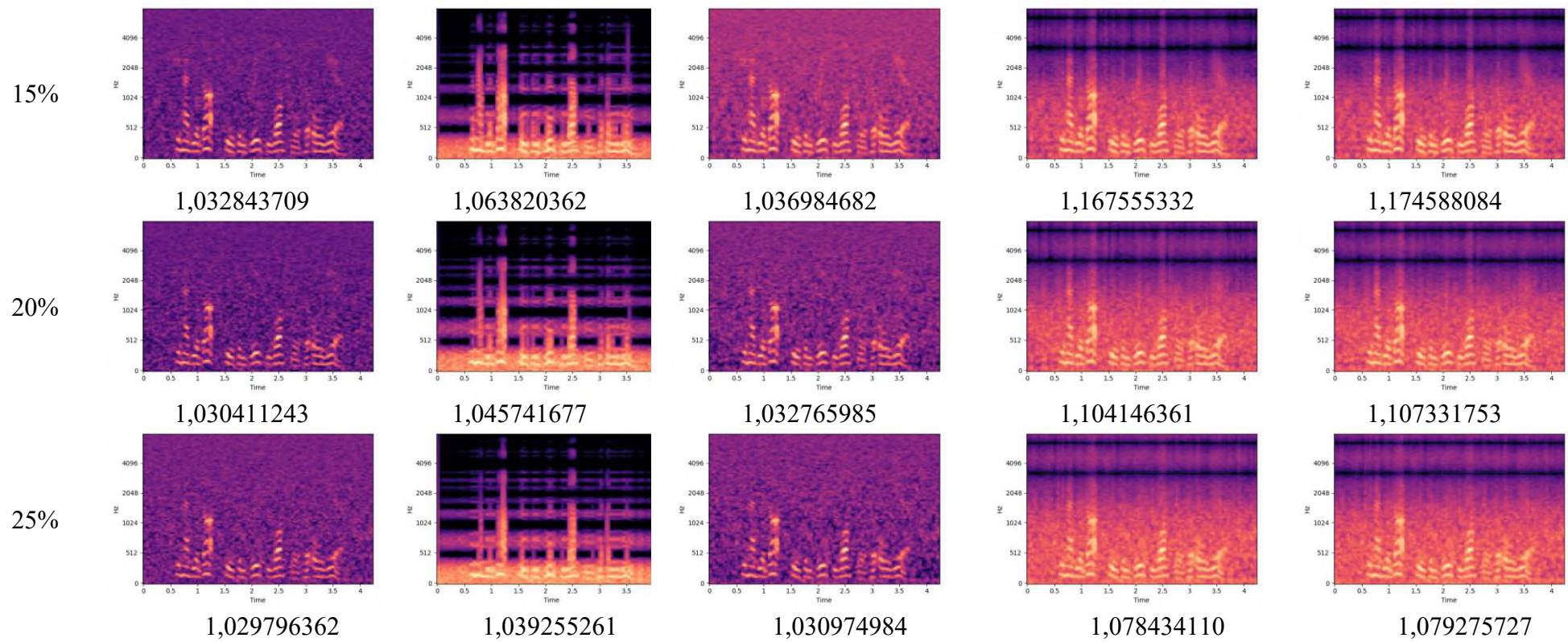
1,049759507



1,300831079

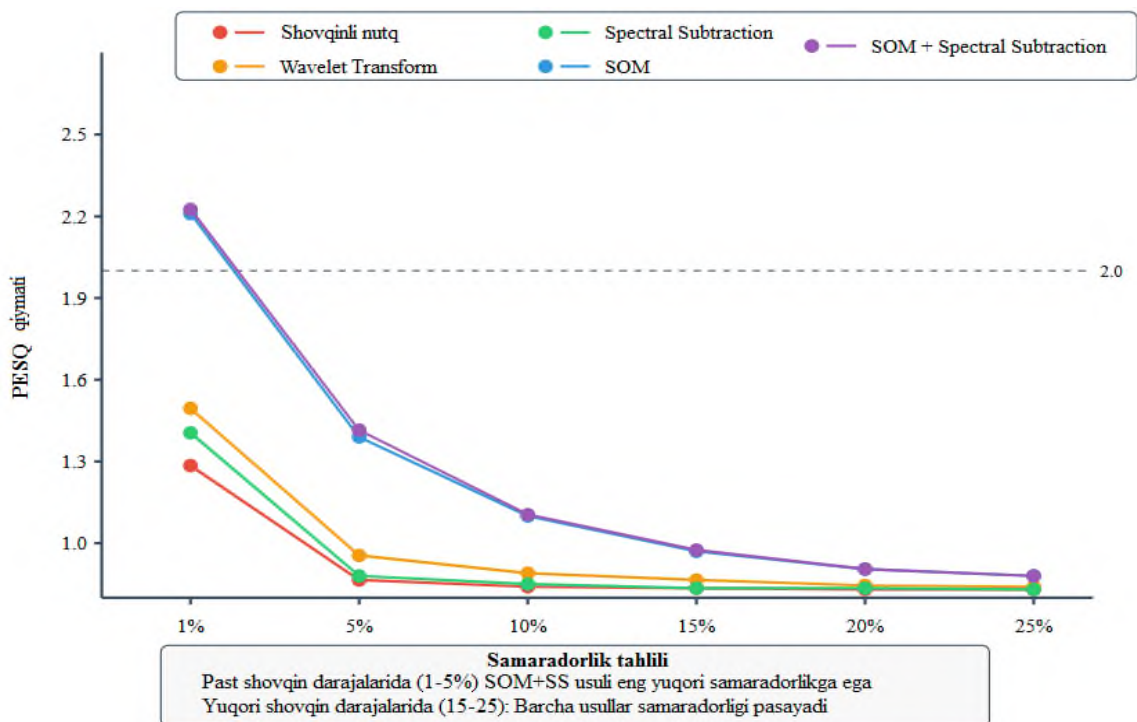


1,306622624



5-rasm. Shovqin pasaytirish usullarini PESQ qiymatlari.

Quyidagi 6-rasmda turli usullar yordamida tozalangan nutq signallari PESQ qiymatlarining taqqoslama tahlili keltirilgan.



6-rasm. Turli usullar yordamida tozalangan nutq signallarini PESQ qiymatlari.

Bu grafik shovqinni turli darajalarida (1% dan 25% gacha) har bir tozalash usulini samaradorligini ko'rsatadi. Grafikdan ko'rinib turibdiki, SOM va SOM+Spectral Subtraction usullari boshqa an'anaviy usullarga nisbatan sezilarli ustunlikka ega, ayniqsa past shovqin darajalarida ularni samaradorligi yuqori.

Xulosa

Mazkur tadqiqotda nutq signalini shovqindan tozalash uchun SOM va Spectral Subtraction usullariga asoslangan yangi yondashuv taklif etildi. Shovqinni turli darajalarida tajribalar o'tkazildi va natijalar PESQ metrikasi orqali baholandi. Olib borilgan tadqiqotlar asosida quyidagi xulosalar shakllantirildi:

1. Taklif etilgan usul shovqinli klasterlarni aniqlashda energiyaga asoslangan mezonni chastota xususiyatlari bilan birlashtirish orqali nutq sifatini sezilarli darajada yaxshilashini ko'rsatdi.

2. SOM va Spectral Subtraction usullari kombinatsiyasi shovqinni barcha darajalarida eng yuqori samaradorlikni ko'rsatdi. Past shovqin darajalarida (1% va 5%) PESQ qiymatlari mos ravishda 2.426 va 1.614 ga teng bo'ldi. Yuqori shovqin darajalarida (15-25%) esa bu yondashuv boshqa usullarga nisbatan barqaror natijalarni ko'rsatdi.

3. Xususiyatlarni adaptiv tanlash yondashuvidan foydalanish shovqinni turli muhitlarida algoritmlar barqarorligini ta'minlashi qayd etildi.

4. Minimum Statistics Noise Estimation usulini qo'llash orqali shovqin taxminini aniqligi oshirilgan, bu esa yuqori shovqin darajalarida ham barqaror natijalar olish imkonini berdi.

Umuman olganda, taklif etilgan usul shovqinni turli darajalarida nutq signallarini shovqinlardan tozalashda yuqori samaradorlikni namoyish etdi.

Adabiyotlar

[1] Mamatov, N., Niyozmatova, N., & Samijonov, A. (2021). Software for preprocessing voice signals. *International Journal of Applied Science and Engineering*, 18(1), 1-8.

- [2] Niyozmatova, N. N., Jalelov, N. K., Samijonov, N. B., & Madrahimova, N. M. (2024). Eliminating noise from a speech signal based on a pair of filters. *International Journal of Science and Research Archive*, 13(2), 401–410. <https://doi.org/10.30574/ijstra.2024.13.2.2058>
- [3] Boll, S.F. (1979). Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 27(2), 113-120.
- [4] Lim, J.S., & Oppenheim, A.V. (1979). Enhancement and bandwidth compression of noisy speech. *Proceedings of the IEEE*, 67(12), 1586-1604.
- [5] Ephraim, Y., & Malah, D. (1984). Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 32(6), 1109-1121.
- [6] Ramirez, J., Górriz, J.M., & Segura, J.C. (2007). Voice activity detection: fundamentals and speech recognition system robustness. In M. Grimm & K. Kroschel (Eds.), *Robust Speech Recognition and Understanding* (pp. 1-22). I-Tech Education and Publishing.
- [7] Martin, R. (2001). Noise power spectral density estimation based on optimal smoothing and minimum statistics. *IEEE Transactions on Speech and Audio Processing*, 9(5), 504-512.
- [8] Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1), 59-69.
- [9] Zhang, X.L., & Wu, J. (2013). Deep belief networks based voice activity detection. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(4), 697-710.
- [10] Ramírez, J., Segura, J.C., Benítez, C., De La Torre, A., & Rubio, A. (2004). Efficient voice activity detection algorithms using long-term speech information. *Speech Communication*, 42(3-4), 271-287.