

SOUND FEATURE ANALYSIS FOR AUDIO SIGNAL RECOGNITION

IVAN GARVANOV, MAGDALENA GARVANOVA, GALINA TKACH AND BOYAN VASOVICH

***ABSTRACT:** This paper analyses sound features using spectral methods for audio signal recognition. For this purpose, sound signals from three different firearms are used. The studied characteristics can be used to recognize and identify different sources of sound signals generated by people, birds, objects, natural phenomena, etc.*

***KEYWORDS:** Signal processing, signal recognition, sound features*

***DOI:** <https://doi.org/10.46687/IEKL1591>*

АНАЛИЗ НА ЗВУКОВИ ХАРАКТЕРИСТИКИ ЗА РАЗПОЗНАВАНЕ НА АУДИОСИГНАЛИ *

ИВАН ГАРВАНОВ, МАГДАЛЕНА ГАРВАНОВА, ГАЛИНА ТКАЧ И БОЯН ВАСОВИЧ

***АБСТРАКТ:** В статията се прави анализ на звукови характеристики чрез спектрални методи за разпознаване на аудиосигнали. За целта се използват звукови сигнали от три различни огнестрелни оръжия. Изследваните характеристики може да се използват за разпознаване и идентифициране на различни източници на звукови сигнали, генерирани от хора, птици, предмети, природни явления и др.*

1 Въведение

Разпознаването на звукови сигнали е фундаментална научна задача, решението на която намира приложение както за цивилни, така и за военни цели. Разпознаването на вида на огнестрелни оръжия по техните звукови характеристики е от съществено значение при борба с тероризма и/или при военни операции. Своевременното разпознаване на вида на използваното оръжие би позволило да се предприемат адекватни действия за противодействие. В контекста на нарастващите предизвикателства пред обществото като терористични атаки, масови стрелби и други видове насилие, способността да се идентифицират и анализират звукови сигнали, генерирани от огнестрелни оръжия, е от критично значение за правоохранителните органи, службите за сигурност и други свързани институции.

Звуковите сигнали, генерирани от огнестрелни оръжия, се отличават с уникални акустични характеристики, които могат да бъдат анализирани и използвани за тяхното разпознаване [1, 2]. Спектралният анализ е един от най-мощните инструменти за изучаване на тези звукови сигнали. Той позволява да се разгледа честотния състав на звука и да се извлекат специфични характеристики като честотни пикове, спектрален центроид и спектрална ширина, които са ключови за идентифициране на звуците от изстрели. Чрез разбиване на звука на неговите основни честотни компоненти, спектралният анализ

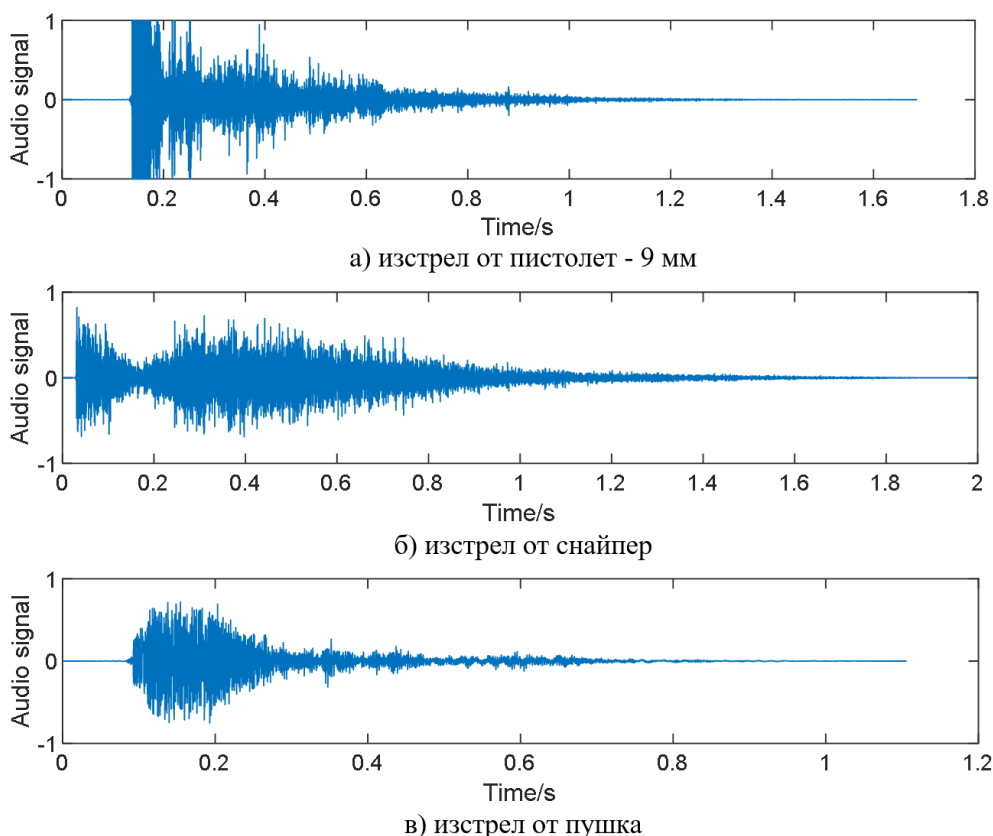
* Тази статия е реализирана по Национална научна програма „Сигурност и отбрана“, финансирана от МОН по договор № Д01-74/19.05.2022 г.

предоставя детайлна картина на звуковия сигнал, което го прави особено подходящ за задачата за разпознаване на огнестрелни оръжия [3].

Настоящата статия изследва и анализира звукови характеристики с цел разпознаване на източника на генериране на аудиосигнали. Целта е да се установят характеристиките които дават най-съществените различия в аудиосигналите с цел постигане на по-висока вероятност за разпознаване на източника на звук. За провеждане на изследването са използвани реални звукови сигнали генерирани от огнестрелни оръжия.

2 Аудиосигнали, генерирани от огнестрелни оръжия

Звуците, генерирани по време на работа с огнестрелно оръжие, са комплексни и включват множество различни елементи, като всеки от тях може да предостави ценна информация за състоянието на оръжието, характеристиките на изстрела и взаимодействието с околната среда. Изследването и анализът на тези звуци са важни за акустично разпознаване на вида оръжие, генериращо съответния звук. В статията са използвани звуци от три вида огнестрелни оръжия (пистолет – 9 мм, снайпер и пушка). Видът на сигналите във функция на времето са визуализирани на *Фигура 1*, като от нея се вижда, че всяко едно от оръжията има специфична амплитудно-времева характеристика.



Фигура 1 – Визуализация на звуци от огнестрелни оръжия

От фигурата се вижда, че аудиосигналите се различават, но при наличие на допълнителни шумове и смущения тези сигнали са трудни за разпознаване. За да се проектира система за автоматично разпознаване на аудиозвуци, е необходимо да се

използват различни звукови характеристики, оценени както във времевата, така и в честотната област на сигнала.

3 Звукови характеристики

3.1. Спектър на сигнала

Спектърът на сигнала се получава чрез Фурие преобразование, което трансформира сигнала от времевата в честотната област. Модулът на спектралната функция се нарича амплитуден спектър, а неговият аргумент – фазов спектър. Фурие преобразованието свързва времевия сигнал със спектралната му функция, запазвайки цялата информация. Представянето на сигнала в честотната област съдържа същото количество информация, както и оригиналният сигнал във времевата област [1].

3.2. Спектрограма на сигнала

Съвкупността от моментните спектри на даден сигнал във функция на времето представлява спектрограмата на сигнала. За да се получи спектрограмата, е необходимо времевият сигнал да се раздели на сегменти и да се приложи Фурие преобразованието за всеки един тях [1]. Наборът от спектри образува спектрограмата на сигнала. Тъй като звуковият сигнал, генериран при изстрел, е с по-голяма продължителност, както и поради различните му честотни характеристики във времето, е по-уместно да се анализира неговата спектрална диаграма [2].

3.3. Мел-честотен кепстрален коефициент (MFCC)

Мел-честотните кепстрални коефициенти (MFCC) са едни от най-популярните методи при извличане на спектрални характеристики. Освен че първоначално са получени за изследване на ехото в сеизмичния звук, те се използват и за моделиране на характеристиките на човешкия глас [3], както и в системите за разпознаване на реч.

Анализът на MFCC включва прилагане на бърза трансформация на Фурие (FFT) върху последователност от кадри, за да се получат определени параметри. Тези параметри преобразуват спектъра на мощност в спектър на Мелчестота, след което се прилага логаритъм и се извършва обратна трансформация на Фурие [4].

Мел-честотните кепстрални коефициенти описват цялостната форма на спектралната обвивка [5].

3.4. Гаматон кепстрални коефициенти (GTCC)

Гаматонните кепстрални коефициенти моделират човешката слухова реакция и спектралния анализ в кохлеята [6]. Някои автори предпочитат тези коефициенти за класификационни задачи. Те се използват главно при разпознаване на реч.

3.5. Спектрален центроид

Спектралният центроид показва върху коя честота е центрирана енергията на спектъра, подобно на средно претеглена стойност [3, 7].

3.6. Спектрална точка на спадане

Спектралната точка на спадане (spectral rolloff point) е граничната честота, под която е концентрирана 95% от мощността на преобразованието на Фурие [3, 8].

3.7. ZCR-скорост на преминаване през нулата

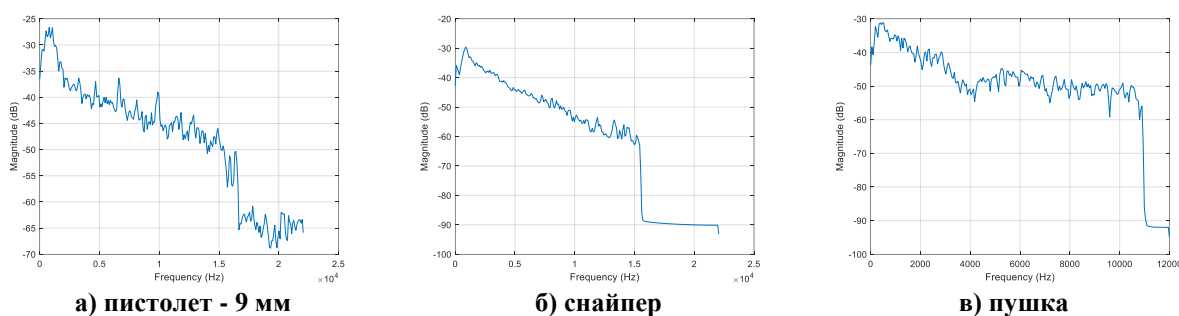
Скоростта на преминаване през нулата определя колко пъти сигналът преминава през нулевата стойност. Този параметър дава косвена информация за основната честота [3, 9].

3.8. Спектрална ентропия

Спектралната ентропия (spectral entropy) на сигнала е нелинеен параметър, който представя спектралното разпределение на неговата мощност. Тя е базирана на ентропията на Шанън или информационната ентропия в теорията на информацията [3]. За да се изчисли ентропията за разпределение на кратковременен спектър на говорен сигнал, той се изследва като вероятно разпределение на дискретна случайна величина.

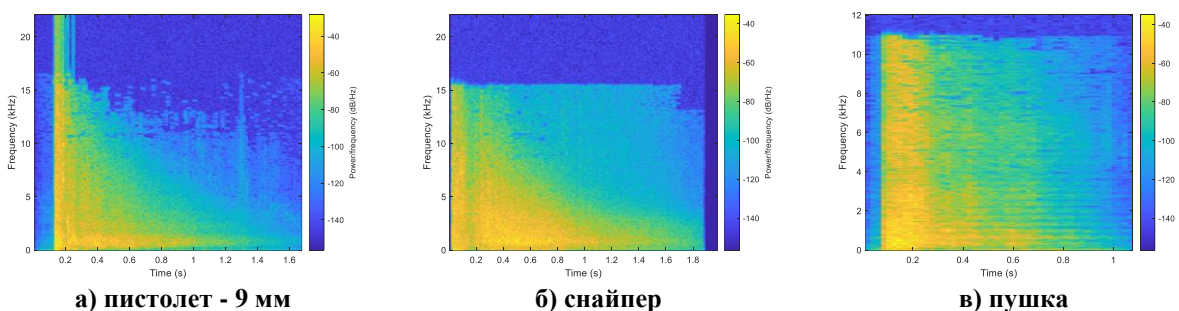
4 Резултати

За анализирането на звуковите сигнали от три вида огнестрелни оръжия са използвани едни от най-известните звукови характеристики в честотната и времевата област на сигнала. За разглеждането на честотното разпределение на сигналите са получени спектрите на сигналите от пистолет, пушка и снайпер (*Фигура 2*). Спектърът показва как се разпределя енергията на звука сред различните честоти. С други думи, спектърът на сигнала е като отпечатък, който разкрива съставните части на звука и дава информация, която не може да се получи само от времевата форма на сигнала.



Фигура 2 – Спектри на звукови сигнали, получени от огнестрелни оръжия

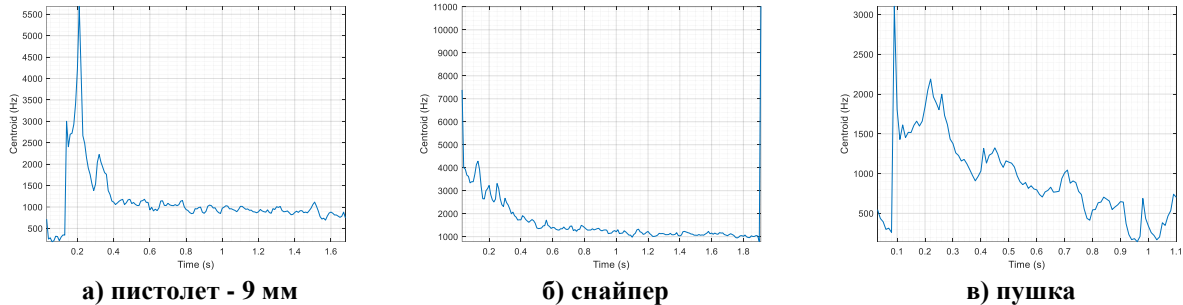
Визуалното представяне на спектъра от честоти на сигнала във функция на времето се нарича *спектрограма* и е подходящо средство за идентифициране на звукови сигнали. Спектрограмите на изследваните звуци са показани на *Фигура 3*, като от тях се вижда от какви честоти е съставен звуковият сигнал и как те се променят във времето, разкривайки съществени детайли от звука.



Фигура 3 – Спектрограми на звукови сигнали, получени от огнестрелни оръжия

Една от ключовите характеристики в анализа на звуковите сигнали, която описва „центъра на тежестта“ на честотния спектър на звука, е спектралният центроид. Той дава информация за това къде е съсредоточена основната част от енергията на звука. Ако спектралният центроид е разположен в по-ниския честотен диапазон, това означава, че

звукът е по-нисък, по-тежък и с повече бас. Ако е в по-високите честоти, звукът ще бъде по-остър, с повече високи тонове. Спектралният центроид на изследваните звуци са показани на *Фигура 4*.



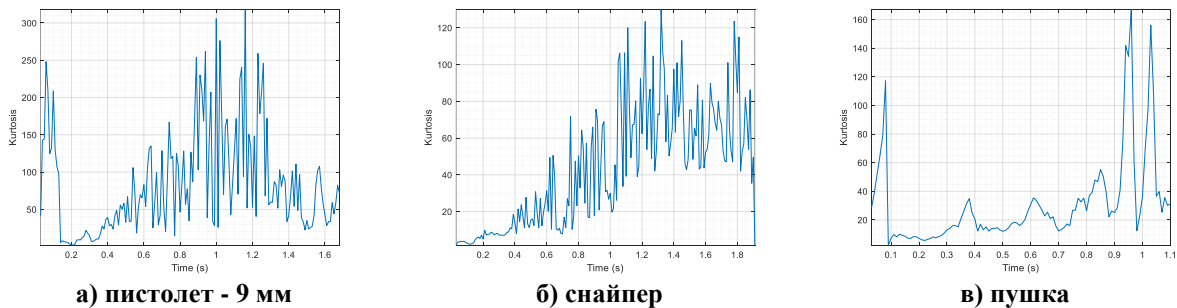
а) пистолет - 9 мм

б) снайпер

в) пушка

Фигура 4 – Спектрален центроид на звукови сигнали, получени от огнестрелни оръжия

Спектралният ексцес (spectral kurtosis) е статистическа мярка, която помага да се разбере колко „остри“ или „плоски“ са честотните компоненти на даден звуков сигнал. Той показва кой от звуците е по-особен или отклоняващ се от нормалното разпределение. Ако ексцесът е висок, това означава, че сигналът има силно изразени честотни компоненти или енергията е концентрирана в малко на брой честоти. Ако ексцесът е нисък, това означава, че енергията на звука е разпределена по-равномерно, което е характерно за звуци като бял шум или звуков фон, където няма ясно изразени честотни пикове. Спектралният ексцес на изследваните звуци са показани на *Фигура 5*.



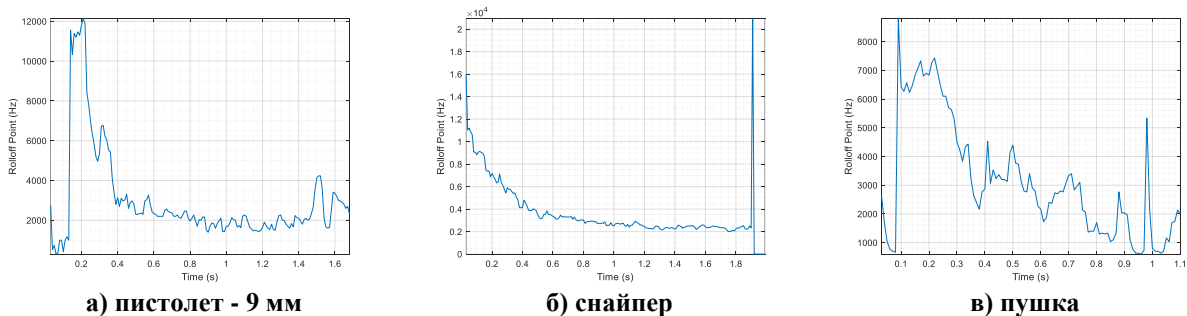
а) пистолет - 9 мм

б) снайпер

в) пушка

Фигура 5 – Спектрален ексцес на звукови сигнали, получени от огнестрелни оръжия

Спектралната точка на спадане е по-висока за глух звук (звук, богат на високи честоти), отколкото за гласен звук (енергия, концентрирана в по-ниски честоти). Следователно това измерване дава възможност да се характеризират гласните/беззвучните промени в речта. За музиката това измерване по същество е винаги едно и също, но за звук от изстрели има вида от *Фигура 6*.



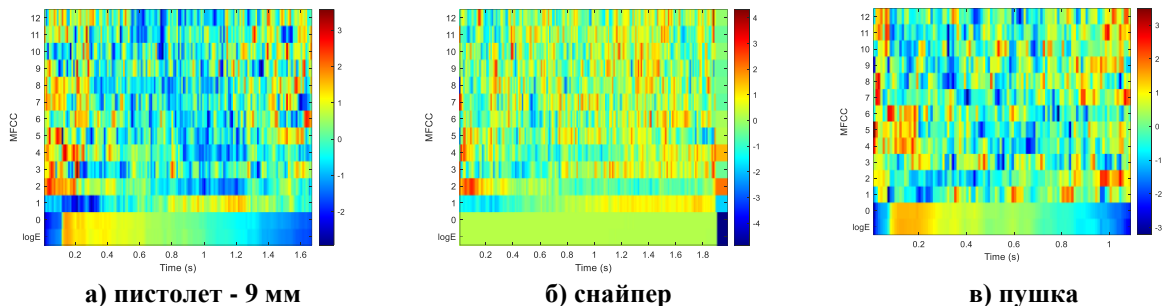
а) пистолет - 9 мм

б) снайпер

в) пушка

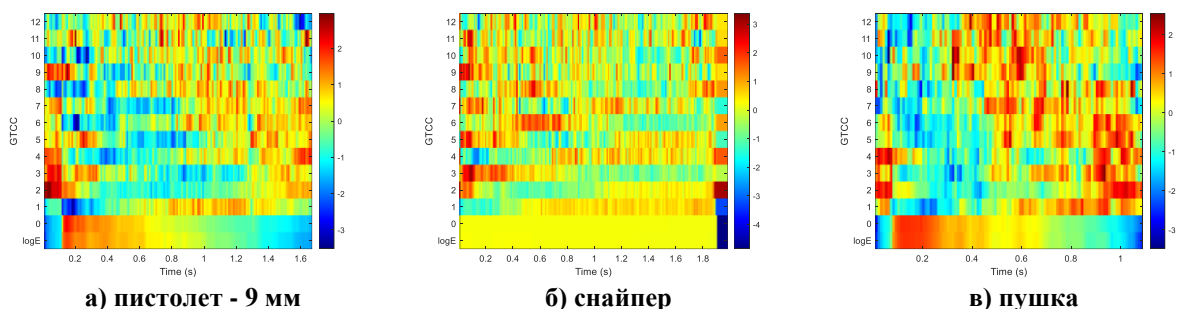
Фигура 6 – Спектрална точка на спадане на звукови сигнали, получени от огнестрелни оръжия

Мел-честотните кепстрални коефициенти са най-използваният и най-ефективен съвременен метод за автоматична обработка на реч и други звукови сигнали при автоматична идентификация на езици, разпознаване на звуци или класификация на реч/музика/шум. Мел-честотните кепстрални коефициенти на изследваните звуци са показани на *Фигура 7*.



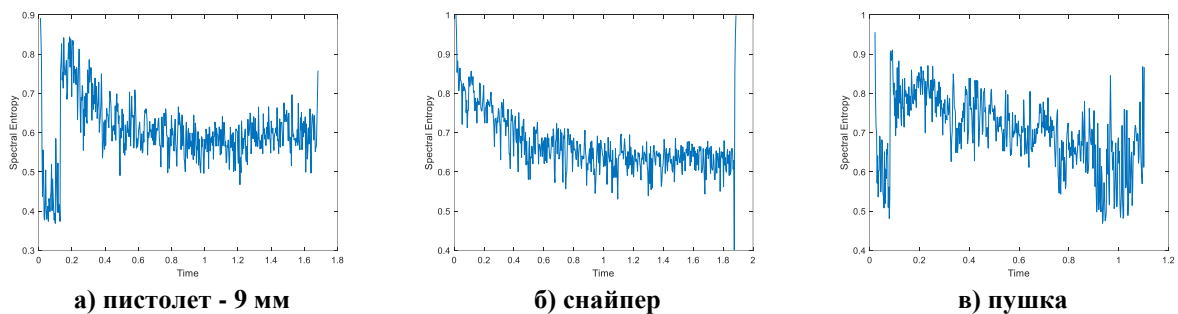
Фигура 7 – Мел-честотни кепстрални коефициенти на звукови сигнали, получени от огнестрелни оръжия

Гаматон кепстралните коефициенти са слухови характеристики, които се използват за разпознаване на реч, музика, подводни цели и др. Те се прилагат като подобрене на конвенционалните Мел-честотни кепстрални коефициенти. Гаматон кепстралните коефициенти на изследваните звуци от огнестрелни оръжия са показани на *Фигура 8*.



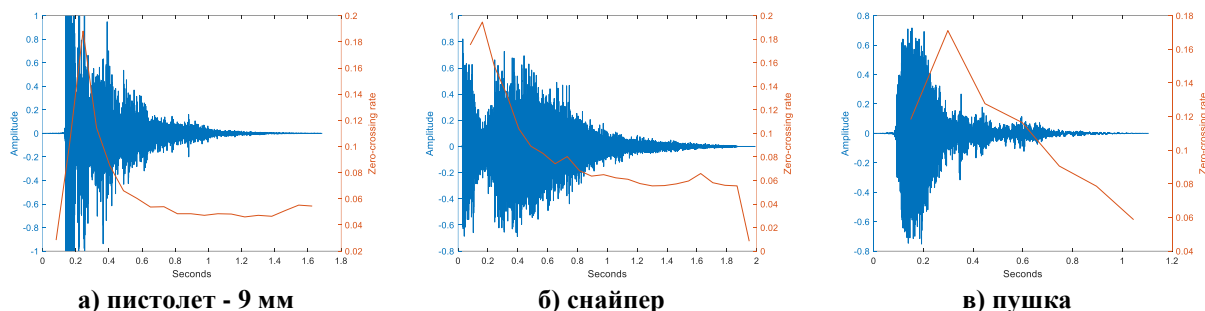
Фигура 8 – Гаматон кепстрални коефициенти на звукови сигнали, получени от огнестрелни оръжия

Спектралната ентропия на изследваните звуци от оръжия са визуализирани на *Фигура 9* и показват спектралното разпределение на мощността.



Фигура 9 – Спектрална ентропия на звукови сигнали, получени от огнестрелни оръжия

Скоростта на преминаване през нулата е скоростта, с която сигналът се променя от положителен към отрицателен или от отрицателен към положителен. Стойността му се използва широко както при разпознаване на реч, така и при извличане на музикална информация и е ключова характеристика за класифициране на ударни звуци. Скоростта на преминаване през нулата на изследваните звуци от огнестрелни оръжия са показани на *Фигура 10*.



Фигура 10 – Скорост на преминаване през нулата на звукови сигнали, получени от огнестрелни оръжия

Получените звукови характеристики нагледно показват различията в изследваните аудиосигнали, генерирани от огнестрелни оръжия. Тези разлики са от съществено значение при следващо разпознаване и класификация на типа оръжие. Изследваните характеристики могат да се използват както при разпознаване, така и за идентифициране на различни звукови сигнали, произвеждани от хора, птици, предмети, природни явления и други източници.

Заклучение

В статия са получени и анализирани някои от най-популярните звукови характеристики, които могат да се използват при разпознаването на звукови сигнали. Получаването на разнообразни характеристики от аудиосигнали е първа стъпка за всяка система за извличане на звук и последваща класификация. За разпознаването и последващо звуково индексирание бяха проучени различни основни компоненти на звука, позволяващи аудиовизуално структуриране и анализирание. Разработените алгоритми са реализирани по проект “Иновативни методи и алгоритми за откриване и разпознаване на подвижни обекти чрез интегриране на хетерогенни данни”, КП-06-Н 72/4/05.12.2023, финансиран от ФНИ-МОН.

ЛИТЕРАТУРА:

- [1] Гарванов, И., Пергелова, П. Разпознаване на изстрел от огнестрелно оръжие, използвайки спектрален анализ. Общество на знанието и хуманизмът на XXI век. XXI национална научна конференция с международно участие, София, 1 ноември 2023 г., АИ “За буквите – О писменехъ”, 2023, 336-346.
- [2] Гарванова, М., Гарванов, И. Акустични характеристики за откриване и разпознаване на дроне. Proceedings of the XXI International Congress “Machines. Technologies. Materials” – Winter Session, 6-9 March 2024, Borovets, Bulgaria, vol. 1, issue 1 (28), 2024, 98-101.
- [3] Pinquier, J. Indexation sonore: recherche de composantes primaires pour une structuration audiovisuelle. Human-Computer Interaction. Universite Paul Sabatier - Toulouse III, 2004, French.

- [4] Kurzekar, P., Deshmukh, R., Vishal, B., Shrishrimal, P. A Comparative study of Feature Extraction Techniques for Speech Recognition System. International Journal of Innovative Research, Engineering and Technology, 3(12):1–11, 2014, doi: 10.15680/IJRSET.2014.031203.
- [5] Prithvi, P., Kishore, T. Comparative Analysis of MFCC LFCC RASTA-LPP. International Journal of Scientific Engineering and Research, 4(5):1–4, 2016.
- [6] Utebayeva, D., Ilipbayeva, L., Matson, E. T. Practical Study of Recurrent Neural Networks for Efficient, Real-Time Drone Sound Detection: A Review. Drones, 7(1): 26, 2023, <https://doi.org/10.3390/drones7010026>.
- [7] Dwivedi, D., Ganguly, A., Haragopal, V. Haragopal Contrast between simple and complex classification algorithms. Statistical Modeling in Machine Learning, 2023.
- [8] Garvanov, I., Pergelova, P., Nurdaulet, N. Acoustic System for the Detection and Recognition of Drones. Telecommunications and Remote Sensing. ICTRS 2023. Communications in Computer and Information Science, vol. 1990. Springer, Cham, 2023, https://doi.org/10.1007/978-3-031-49263-1_8.
- [9] Гарванова, М., Гарванов, И. Акустична система за откриване на безпилотни летателни апарати. VII International Scientific Conference on Security “CONFSEC 2023” – 4-7 December 2023, Borovets, Bulgaria, vol. 1 (9), 2023, 104-107.

Иван Гарванов

Университет по библиотекознание и информационни технологии, София, България,
i.garvanov@unibit.bg

Магдалена Гарванова

Университет по библиотекознание и информационни технологии, София, България,
m.garvanova@unibit.bg

Галина Ткач

Казахски национален университет “Ал-Фараби”, Алмати, Казахстан,
Galinatkach66@gmail.com

Боян Васович

Академия за приложни изследвания, Прокупле, Сърбия,
bojan.vasovic@yahoo.com