

Information systems research

UDC 004.7

doi: <https://doi.org/10.20998/2522-9052.2022.1.08>

Olesia Barkovska, Vladyslav Kholiev, Vladyslav Lytvynenko

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

STUDY OF NOISE REDUCTION METHODS IN THE SOUND SEQUENCE WHEN SOLVING THE SPEECH-TO-TEXT PROBLEM

Abstract. The subject of this research is noise reduction methods in the sound sequence as a part of the proposed speech-to-text (STT) module for converting a verbal lecture or a lesson into a written text form on digital educational platforms. The goal is to investigate the influence of noise reduction methods on the operation of the acoustic signal recognition system. 3 methods of noise reduction were considered for integration in the proposed acoustic artifact recognition system and for the researching: spectral subtraction method; fast Fourier transform; Wiener filter with software modeling of every method. The obtained results: after testing the system with integrated noise reduction modules in it, based on the fast Fourier transform, Wiener filter and spectral subtraction method, it was concluded that the module using the Wiener filter improves the identification results by 25%, which is the highest result. However, performance testing has shown that fast Fourier transform is the fastest method. The practical significance of the work is – the identifying acoustic events system was developed, different noise reduction methods were integrated and researched into the module for converting a verbal lecture or a lesson into a written text form in a proposed system with the aim of increasing of speed and accuracy.

Keywords: speech; voice; processing; noise; reduction; filtering; learning management systems; digital educational platform.

Introduction

The period of the pandemic has significantly increased the relevance of the development and expansion of the functionality of various digital educational platforms [1, 2]. These information spaces are in demand at different levels of education in countries around the world – from primary schools to higher education institutions, as well as training courses in different areas of business, allowing students to be provided with teaching materials, communicating with teachers, and also providing an opportunity for remote control of the level of knowledge. This approach is not a new one [3], that is why now we have not the breakthrough, but the fur-

ther fast development [4, 5]. Thus, the virtual interaction of the distributed user community is organized.

The number of interactive possibilities of digital educational platforms is constantly expanding. New options are added to the basic functionality (access to training materials on cloud storage services, remote interaction with teachers). For example, testing blocks, knowledge control, access level management of training content depending on the user's role, analysis of statistical data (attendance), and some other special functionality like in [6,7]. The generalized structure of a digital educational platform with modules for learning educational material, checking gained knowledge and collecting statistics for monitoring progress is shown in Fig. 1.

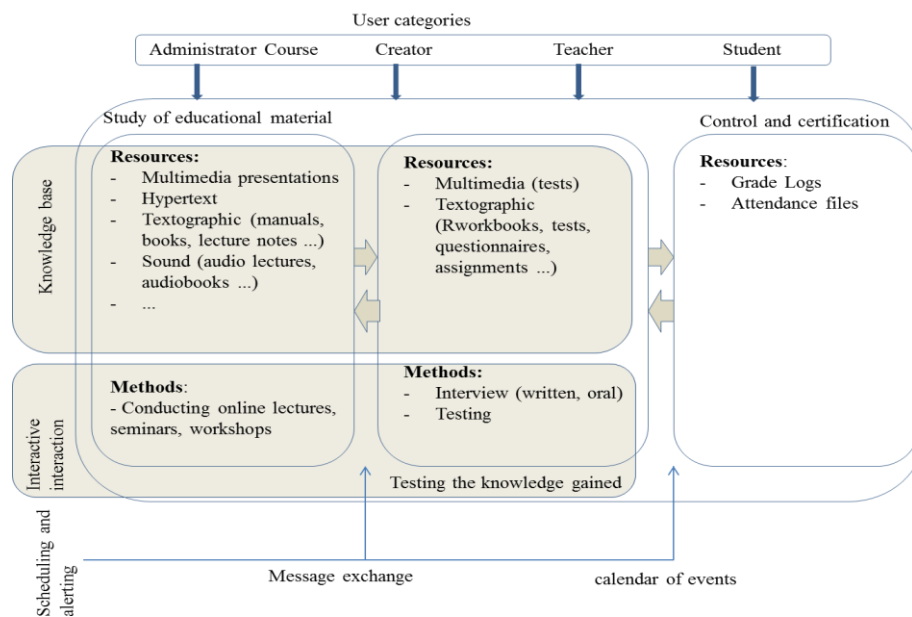


Fig. 1. Generalized structure of a digital educational platform

However an important part of the information space for remote education remains online lectures with experts, which can be conducted in a question-and-answer

mode and cover questions from course participants. Subsequent access to online lecture materials should be easy to understand and assimilate. The recording of the lecture

provides access to audio files that can be listened to, but are not intended for print reproduction.

The two platforms for schools, universities and other educational institutions from the great corporations Google [8, 9] and Microsoft are: G Suite for Education and Microsoft 365 Education. 'Microsoft 365' is the new name for 'Office 365'.

In both cases, the platforms run from the cloud, meaning you can use them through a standard internet browser – perfect for everything from a quick check of a document on a smartphone, or for a student to join a lesson on a desktop computer. And naturally, being browser-based and suitable for multiple devices, they are ideal for remote teaching and learning.

Software and services related to digital educational platform (DEP), which we have in a generalized structure of a DEP (Fig. 1):

- learning management systems (LMS) — integrations with LMS give users of digital learning platforms a streamlined way of sharing assignments with their students. Instructors can also use LMS integrations to track student progress on assignments;

- student information systems (SIS) — an institution's SIS harbors vast stores of student data, including contact information, demographic data, and their academic history. Integrations between these systems and digital learning platforms enable instructors to seamlessly transfer student rosters into digital learning platforms.

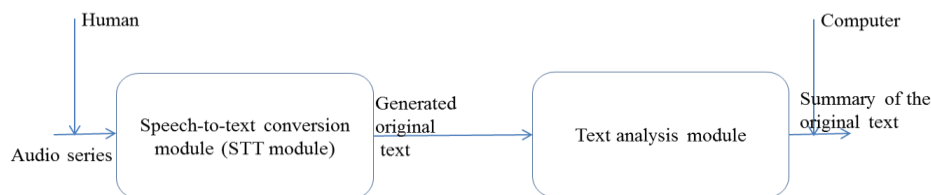


Fig. 2. Basic modules of the proposed system

Analysis of the literature. As the name suggests, a DEP is an online 'environment' comprising applications and tools for the education sector. It's used by teachers, administrative staff and students – and is designed specifically to accelerate digital teaching and learning for schools. A typical DEP contains tried-and-tested productivity software, such as those for word processing, spreadsheets and presentations – as well as email and calendar tools. No doubt you're already familiar with such software. For DEPs, however, these standard applications are bolstered with software for education. This includes virtual whiteboards; planning, assignment, marking and collaboration tools; and software to run lessons by live video.

Here's one example of how this might translate into the real world: a teacher plans the lesson and produce materials using the familiar productivity tools, as well as your school's own curriculum materials stored conveniently in the same platform. Then the mentor can run the lesson either remotely or in-class, using a presentation format, live video, or a mix of both. Based on the lesson, you could then issue an assignment to students, who would complete it remotely and return it through the platform. You could issue the assignment in question-and-answer format created with the platform's

They can also export student performance data from a digital learning platform assignment directly into an SIS gradebook;

- document creation software — digital learning platforms that enable instructors to create their own lessons often allow the user to upload files from document creation software.

Therefore, the expansion of existing digital educational platforms with the ability to form an annotation (resume, abstract) of a lecture and its presentation in the form of textographic materials for further use by course students on paper is an urgent task, since it can improve the quality of presentation of educational information and the conditions for working with it, and also to improve the assessment of the quality of the distance educational resource from the point of view of the content and methodological aspect.





The paper proposes converting audio files containing lecture material into text form, keeping only the important content in the form of a summary of the lecture (Fig. 2). So, it is well known speech-to-text (STT) problem in appliance to digital educational platforms. Since the reliability of the information contained in the educational resources of distance courses is one of the key requirements for digital educational platforms, reducing the appearance of false or distorted data when converting a sound sequence into text data for further semantic analysis is the goal of the work.

questionnaire tools. Or you might ask the student to submit a typed document – or even a handwritten response using a device stylus. For more practical subjects, the student could submit multimedia formats of their work – for instance, a photo of their drawing, or a video of a musical performance. Once these assignments are completed and submitted, you could mark the work using the marking software. These are tools that go beyond marking up documents with comments – they're intuitive and interactive, and can be set up so the data flows to a spreadsheet tracking the student's progress.

The benefits of using a digital education platform can be the next: teach classes as well as smaller groups of pupils; communicate with their pupils; set tasks for individuals as well as for larger groups; let pupils work together; give feedback; share useful links to digital learning resources; collaborate with their colleagues on lesson planning and related administrative tasks. Below is a ranking list of DEP with Higher Education capabilities (Table 1) [10]. The use of DEP in higher education systems is an important function of digital learning platforms. Some of them are: Canvas LMS, Nearpod, Khan Academy, Lumio by SMART, GoReact, Discovery Education Inc, Speexx, Pear Deck, Renaissance Accelerated Reader, McGraw-Hill and so

on. Comparison of various products is necessary in order to analyze the advantages and disadvantages of existing solutions and substantiate the relevance of the approach proposed in the work.

Table 1 – A ranking list of some of the digital educational platforms

				
Meets Requirements	8.6 (20 responses)	8.3 (23 responses)	8.6 (27 responses)	8.5 (266 responses)
Ease of Use	9.2 (20 responses)	8.7 (24 responses)	8.8 (27 responses)	8.5 (266 responses)
Ease of Setup	9.2 (6 responses)	7.6 (13 responses)	Not enough data available	8.4 (71 responses)
Ease of Admin	9.2 (6 responses)	7.2 (13 responses)	Not enough data available	8.6 (74 responses)
Quality of Support	9.5 (20 responses)	8.2 (23 responses)	9.3 (17 responses)	8.4 (226 responses)
Ease of Doing Business With	9.7 (6 responses)	8.5 (12 responses)	Not enough data available	8.9 (59 responses)
Reviewers' Industry				
	Higher Education 86.4%	Higher Education 59.4%	Education Management 40.6%	Education Management 39.0%
	Non-Profit Organization Management 4.5%	E-Learning 12.5%	Primary/Secondary Education 31.3%	Primary/Secondary Education 37.4%
	E-Learning 4.5%	Apparel & Fashion 6.3%	Higher Education 15.6%	Higher Education 9.7%
	Alternative Dispute Resolution 4.5%	Public Relations and Communications 3.1%	E-Learning 6.3%	E-Learning 5.7%
	Other, 0.0%	Maritime, 3.1%	Libraries, 3.1%	Writing and Editing, 0.6%
		Other, 15.6%	Other, 3.1%	Other, 7.5%

It can be seen that these platforms can satisfy many user needs, but still have minor disadvantages.

The analysis showed that according to such criteria as ease of use and meeting the requirements, the Alta system has the highest scores, however, its functionality does not include the possibility of converting a voice lecture or report into a printed form for further use both by students (when preparing for examinations) and by teachers (for example, for compiling work programs or writing a detailed lecture summary). Therefore, the development of a module for converting a verbal lecture or a lesson into a written text form while retaining only important information can be considered a relevant task.

The problem formulation. The quality and speed of converting audio sequence into text, which is a complex task, depend on such factors as the quality of the input signal and the characteristics of the equipment. The purpose of this study is to investigate the influence of noise reduction methods on the operation of the acoustic signal recognition system. To achieve this goal, the following tasks must be solved:

- research of the general parameters of acoustic systems;
- research of existing methods of noise reduction in the audio series;
- development of acoustic event recognition system;
- adaptation of noise reduction methods to different types of computing systems;
- assessment of the operating time of the developed system;
- assessment of the quality of acoustic event recognition;
- analysis of the obtained results.

Results and Discussion

1. Acoustic event recognition system. To study the methods of noise reduction in the audio sequences, a

computer system for recognizing acoustic events was developed, namely: music and its genres, animal sounds divided into classes, transport and its type, random events (knock, explosion, shot, etc.), human speech, natural phenomena, etc.

The system is designed to identify the type of event on a given audio sequences of any size and can be used as a ShotSpotter [11] analogue or to eliminate undesired artifacts when using software for audio and video conferencing [12, 13]. The system is able to work on the following three sources: pre-recorded audio file; capturing and processing microphone data; web-interface.

The system is written in Python, which allows it to be easily adapted and optimized. It uses a predefined model of a neural network of the CNN type. The model is based on the Google AudioSet dataset (the sample size is approximately 2 million segments from different video files, the test sample size is 20,000 segments from different video files), which is publicly available. Overall, the system uses the following external Python libraries: numpy; scipy; resampy; PyAudio; Tensorflow; Six; Devicehive. It should be noted that the neural network is trained on the VGGish model of the Tensorflow platform.

Due to the fact that GPUs are more suitable for training neural networks, and the Tensorflow platform provides the ability to use a graphics processor (with the above software), the following hardware was used: NVidia GeForce RTX 2070 Super 8 GB. Fig. 3 shows the general IDEF0 scheme of the system.

On the left of the diagram the input data to the system is shown, namely the audio sequence (audio file), which must be identified to one or more classes. The top shows the tools used by the system, namely: the TensorFlow platform, which is required for the neural network model to work; a collection of classes required to assign an input audio sequence to a class; YouTube-

8M model, interface for working with the neural network model; Python modules (libraries), required to run each part of the system (for example, os is used to run the system on local files).

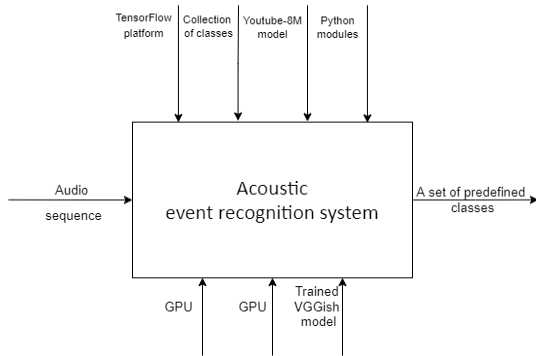


Fig. 3. Generalized scheme of work of acoustic artifact recognition system

The lower inputs are the tools and software that the system uses. The CPU and the VGGish-trained neural network model are mandatory components of the system, while the GPU is desirable but optional. VGGish is a pre-trained convolutional neural network developed by Google. The presence of GPUs in the system improves speed. Fig. 4 shows the functional model of the system.

The proposed system consists of modules: conversion of a file into an array of bits; conversion of an array of bits into a mel spectrogram; conversion of mel spectrogram into frames; obtaining key characteristics; processing of key characteristics; filtering predictions. The tasks of the module to convert a file into an array of bits include: reading a file into a stream; conversion of a stream into an array of bits; array type check. The module takes the input path to the file and outputs an array of bits.

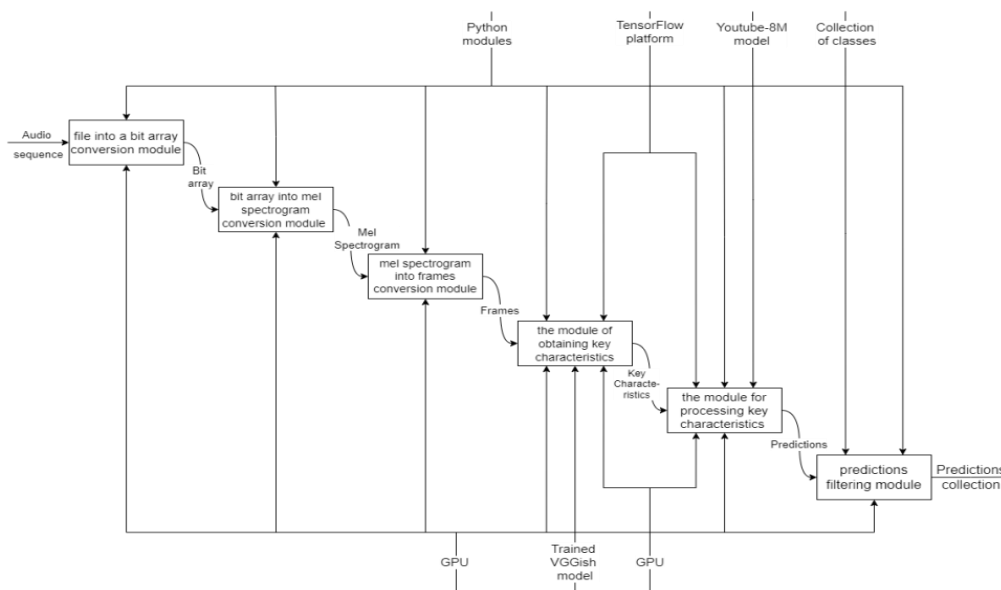


Fig. 4. Functional model of the system

Next is the module for converting an array of bits into a mel spectrogram. Its tasks include: conversion of 2 or more channel signals into mono; short-term (window) Fourier transform; obtaining a logarithmic amplitude mel-frequency spectrogram.

The short-term Fourier transform is performed to be able to obtain a stabilized spectrogram. Next, the obtained spectrogram is fed to the module for converting the mel spectrogram into frames. It has only one task – to divide the spectrogram into segments of 10 milliseconds. The obtained output is an array of frames (segments). The next step is the module of obtaining key characteristics. The essence of this module is to obtain tensors and key characteristics from a set of frames. The tasks of the second to last module for processing key characteristics are: pre-processing of key characteristics by the YouTube-8M module; obtaining predictions from key characteristics. And finally, the last module is prediction filtering. In this case, filtering means rejecting predictions that have not exceeded the specified threshold.

2. Adaptation of the proposed noise reduction methods for computer systems with different types of

memory organization. The basis of this adaptation is the integration of various noise reduction methods into the acoustic event recognition system in order to improve the recognition results. Fig. 5 shows a functional model of a system with an integrated noise reduction module. This module has a single task – pre-processing the audio sequences with one of the methods of noise reduction. An array of bits is fed to the input, and a processed array of the same type is obtained at the output.

This eliminates the need to modify other components of the system. Three methods of noise reduction were considered for integration: spectral subtraction method; fast Fourier transform; Wiener filter. Each of them represents its own version of the noise reduction module, which can be integrated into the system (Fig. 6).

Implementation of each method requires an understanding of the scope of the system.

That is, it needs to be known what kind of event the system needs to identify. For example, if the system needs to recognize a dog's bark, the "useful signal" is in the range of 300 to 3500 Hz, so everything else can be considered "noise".

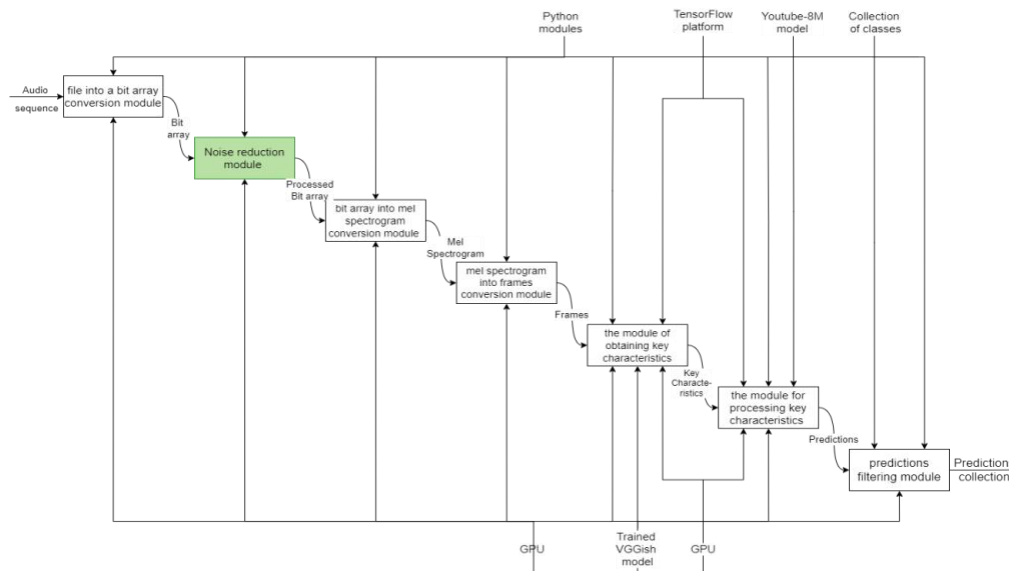


Fig. 5. Functional model of a system with an integrated noise reduction module

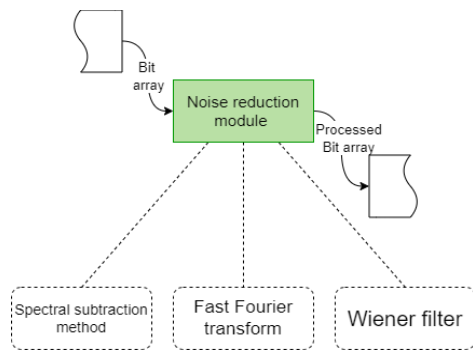


Fig. 6. Noise reduction module of different variations

3. Analysis of the recognition quality of the developed system depending on the methods of noise reduction. The system analysis was performed without an integrated noise reduction module and with three proposed noise reduction methods. The identification threshold was set at 0.8, and the noise type was a synthetically generated white noise of varying strength. Table 2 shows an example of comparing the results of event identification using different noise reduction methods. The value in the table is the identification factor by which the system recognized the event. An audio series of "dog barking" was taken as an example.

The test sample consists of 200 audio files, 50 of which are "dog barking". The overall evaluation of the system was performed using precision and recall metrics with a noise level of 0.5. The precision of a system within a class is the proportion of documents that actually belong to that class for all documents that the system has assigned to that class. Recall of the system is the fraction of the documents assigned by the classifier to the class, to all documents of this class in the test sample. These values are easy to calculate based on the contingency table compiled for each class separately. When using a system without the reduction, 32 samples were marked as "dog barking", of which 29 are correctly recognized and 3 do not belong to this class. Thus we have the following: TP = 29, FN = 21, FP = 3. Then, the precision will be 0.90625 and the recall will be 0.58.

Table 2 – Comparison of noise reduction methods depending on the white noise level

Noise reduction method used	Noise strength		
	0.1	0.5	1.0
No reduction	0,93	0,80	-
Spectral subtraction method	0,96	0,96	0,90
Fast Fourier transform	0,95	0,94	0,92
Wiener filter	0,97	1,00	1,00

Using the noise reduction module of the spectral subtraction method, TP = 46, FN = 4, FP = 1 were obtained. Therefore, the precision is 0.97872 and the recall is 0.92. Using the fast Fourier transform, TP = 45, FN = 5, FP = 3, precision – 0.9375, recall – 0.9. Finally, when using the Wiener filter, the following results were obtained: TP = 49, FN = 1, FP = 0. Therefore, the precision is 1 and the recall is 0.98. Table 3 presents the generalized test results. From Table 3, it can be concluded that the Wiener filter performs the task of noise reduction and improves the results of event identification the best.

Table 3 – Generalized results of event identification system evaluation

Noise reduction method	System evaluation			Metrics		Roe*
	TP	FN	FP	Precision	Recall	
No suppression	29	21	3	0,90625	0,58	-
Spectral subtraction method	46	4	1	0,97872	0,92	21%
Fast Fourier transform	45	5	3	0,9375	0,9	18%
Wiener filter	49	1	0	1	0,98	25%

Roe* – Relative optimization estimate

4. Analysis of system runtime for different noise reduction methods. The analysis of the system runtime was performed in parallel with the identification analysis. Therefore, the sample for testing is the same 200 audio files of the same duration. The countdown of the recognition time begins with the start of the first module – file to bits array conversion module, and the end point of the calculation after the operation of the

last module – the prediction filtering module. Fig. 7 shows the total graph of the time of identification of each sample with each method of noise reduction. The average time of each method was obtained from the graph and set of time metrics:

- no noise reduction method: 2,207 seconds;
- fast Fourier transform: 2,288 seconds;
- Wiener filter: 2,437 seconds;
- spectral subtraction method: 2,663 seconds.

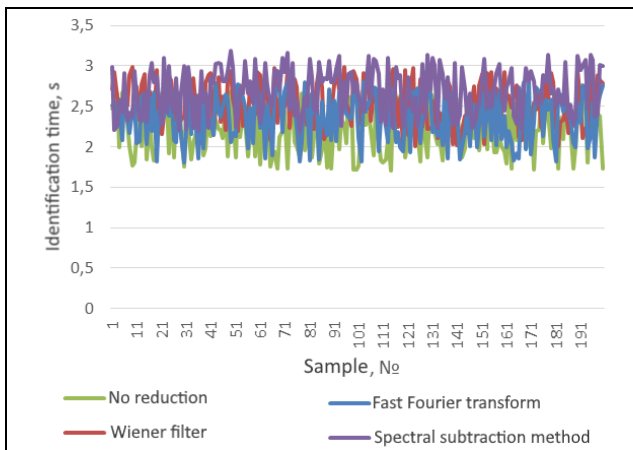


Fig. 7. System runtime comparison with different noise reduction methods

From the average values of the audio event identification time, it can be concluded that the fast Fourier transform is the fastest.

When choosing the noise reduction method for the system, both the processing time of the algorithm and the quality of identification should be taken into account. For example, if on a certain hardware the system processes the audio sequence fast enough that the delay is negligibly small, and the quality of recognition is important to the user, then the Wiener

filter should be chosen as the best according to the results of quality testing. If, for example, the noise level is low and the hardware is not productive enough (for example, in a case of no GPU), a simpler and faster method – fast Fourier transform – should be chosen.

Conclusions

This paper discussed the main applications of audio analytics, such as speech recognition and speech-to-text transformation. Many of them use noise reduction as a basis for improving the signal level, which was demonstrated in the paper as a proposed speech-to-text (STT) module for converting a verbal lecture or a lesson into a written text form on digital educational platforms.

The main methods and algorithms of noise reduction were considered with the presentation of examples of processing results, their computational complexity and the possibility of parallelization. The following methods and algorithms were considered:

- fast Fourier transform;
- Wiener filter;
- spectral subtraction method;
- adaptive filter;
- active noise control;
- use of neural networks.

In addition, a system for identifying acoustic events was developed during the work. The system was tested as is. Also, different noise reduction methods were integrated and researched into the module for converting a verbal lecture or a lesson into a written text form in a proposed system: fast Fourier transform; Wiener filter; spectral subtraction method.

After testing the system with the integration of each of the noise reduction modules, it was concluded that the module using the Wiener filter improves the identification results by 25%, which is the highest result. However, performance testing has shown that fast Fourier transform is the fastest method.

REFERENCE

1. Onyema, E. M., Eucheria, N. C., Obafemi, F. A., Sen, S., Atonye, F. G., Sharma, A., & Alsayed, A. O. (2020), "Impact of Coronavirus pandemic on education", *Journal of Education and Practice*, no. 11(13), pp. 108-121.
2. Arora, A. K. and Srinivasan, R. (2020), "Impact of pandemic COVID-19 on the teaching-learning process: A study of higher education teachers", *Indian journal of management*, no. 13(4), Prabandhan, pp. 43-56.
3. Alaneme, G. C., Olayiwola P. O. and Reju, C. O. (2010), "Combining traditional learning and the e-learning methods in higher distance education: Assessing learners' preference", *4th International Conference on Distance Learning and Education*, pp. 187-190, doi: <https://doi.org/10.1109/ICDLE.2010.5606008>.
4. Faltynkova, L., Simonova I. and Kostolanyova, K. (2020), "Utilization of distance education during COVID_19 crisis", *6th IEEE Congress on Information Science and Technology*, pp. 59-62, doi: <https://doi.org/10.1109/CiSt49399.2021.9357233>.
5. Gorshenin, A. (2018), "Toward Modern Educational <https://doi.org/10.1109/ICUMT.2018.8631229> IT-ecosystems: from Learning Management Systems to Digital Platforms", *10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pp. 1-5, doi: [1109/ICUMT.2018.8631229](https://doi.org/10.1109/ICUMT.2018.8631229).
6. Titov, I. and Titov, E. (2013), "Delivering video content in educational platforms", *International Conference on Interactive Collaborative Learning (ICL)*, pp. 726-727, doi: <https://doi.org/10.1109/ICL.2013.6644692>.
7. Jiang, H. (2018), "Design of Innovative Experiment Platform for Network Courses of Modern Educational Technology Based on Virtual Reality", *11th International Conference on Intelligent Computation Technology and Automation (ICICTA)*, pp. 97-100, doi: <https://doi.org/10.1109/ICICTA.2018.00030>.
8. Kansal AK, Gautam J, Chintalapudi N, Jain S and Battineni G. (2021), "Google Trend Analysis and Paradigm Shift of Online Education Platforms during the COVID-19 Pandemic", *Infectious Disease Reports*, no. 13(2), pp. 418-428, doi: <https://doi.org/10.3390/idr13020040>.
9. Kirange, S., Sawai, D. & Director, I. M. (2021). "A Comparative Study Of E-Learning Platforms And Associated Online Activities", *The Online Journal of Distance Education and e-Learning*, no. 9(2).
10. Cavus, N. and Sekyere-Asiedu, D. (2021), "A comparison of online video conference platforms: Their contributions to education during COVID-19 pandemic", *World Journal on Educational Technology*, no. 13(4), pp. 1162-1173.

11. Watkins, C., Mazerolle, L. G., Rogan, D. and Frank, J. (2002), "Technological approaches to controlling random gunfire: Results of a gunshot detection system field test Policing", *International Journal of Police Strategies & Management*, no. 6, 25 p.
12. Barkovska O., Kholiev V., Pyvovarova D., Ivaschenko G. and Rosinskiy D. (2021), "International system of knowledge exchange for young scientists", *Advanced Information Systems*, Vol. 5, No. 1, pp. 69-74, doi: <https://doi.org/10.20998/2522-9052.2021.1.09>.
13. Barkovska O., Pyvovarova D., Kholiev V., Ivashchenko H. and Rosinskiy D. (2021), "Information Object Storage Model with Accelerated Text Processing Methods", *Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021)*, vol. 1, pp. 286-299.

Received (Надійшла) 10.12.2021

Accepted for publication (Прийнята до друку) 02.02.2022

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Барковська Оlesia Юрійвна – кандидат технічних наук, доцент, доцент кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Olesia Barkovska – Candidate of Technical Sciences, Associate Professor Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: olesia.barkovska@nure.ua; ORCID ID: <http://orcid.org/0000-0001-7496-4353>.

Холєв Владислав Олександрович – аспірант кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Vladyslav Kholiev – postgraduate student at the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: vladyslav.kholiev@nure.ua; ORCID ID: <http://orcid.org/0000-0002-9148-1561>.

Литвиненко Владислав Сергійович – студент кафедри Електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Vladyslav Lytvynenko – students of the Department of Electronic Computers, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: vladyslav.lytvynenko@nure.ua; ORCID ID: <http://orcid.org/0000-0002-4376-7629>.

Вивчення методів зниження шуму в звуковому ряді для вирішення задач перетворення мови в текст

О. Ю. Барковська, В. О. Холєв, В. С. Литвиненко

Анотація. Предметом дослідження є методи шумоподавлення у звуковій послідовності у складі пропонованого модуля перетворення мови в текст (STT) для представлення усної лекції або уроку в текстовому вигляді на цифрових освітніх платформах. **Мета** – дослідити вплив методів шумозаглушення на роботу системи розпізнавання акустичного сигналу. Для інтеграції в запропоновану систему розпізнавання акустичних артефактів та для дослідження було розглянуто три методи зменшення шуму: метод спектрального віднімання; швидке перетворення Фур'є; Фільтр Вінера з програмним моделюванням кожного методу. **Отримані результати:** після тестування системи з інтегрованими в неї модулями шумозаглушення на основі швидкого перетворення Фур'є, фільтра Вінера та методу спектрального віднімання було зроблено висновки, що модуль із застосуванням фільтра Вінера покращує результати ідентифікації акустичних артефактів на 25%, що є найвищим результатом. Проте тестування продуктивності показало, що швидке перетворення Фур'є є найшвидшим методом. **Практична значущість роботи** полягає в тому, що запропоновано та розроблено модулі системи ідентифікації акустичних подій, інтегровано та досліджено різні методи шумозаглушення в модуль для перетворення усної лекції чи уроку в письмову текстову форму в запропонованій системі з метою збільшення швидкості і точності результату.

Ключові слова: мовлення; голос; обробка; шум; скорочення; фільтрація; системи управління навчанням; цифрова освітня платформа.

Изучение методов понижения шума в звуковой последовательности при решении задач преобразования речи в текст

О. Ю. Барковская, В. О. Холєв, В. С. Литвиненко

Аннотация. Предметом исследования являются методы шумоподавления в звуковой последовательности в составе предлагаемого модуля преобразования речи в текст (STT) для представления устной лекции или урока в текстовом виде на цифровых образовательных платформах. **Цель** – исследовать влияние методов шумоподавления на работу системы распознавания акустического сигнала. Для интеграции в предложенную систему распознавания акустических артефактов и для исследования было рассмотрено три метода уменьшения шума: метод спектрального вычитания; быстрое превращение Фурье; Фильтр Винера с программным моделированием каждого способа. **Полученные результаты:** после тестирования систем с интегрированными в нее модулями шумоподавления на основе быстрой трансформации Фурье, фильтра Винера и метода спектрального вычитания был сделан вывод, что модуль с применением фильтра Винера улучшает результаты идентификации акустических артефактов на 25%, что является наивысшим результатом. Тем не менее, тестирование производительности показало, что быстрое преобразование Фурье является самым быстрым методом. **Практическая значимость работы** заключается в том, что предложены и разработаны модули системы идентификации акустических событий, интегрированы и исследованы различные методы шумоподавления в модуль для преобразования устной лекции или урока в форму письменного текста в предложенной системе с увеличением скорости и точности получения результата.

Ключевые слова: речь; голос; обработка; шум; сокращение; фильтрация; системы управления обучением; цифровая образовательная платформа.