

ПРОГРЕСИВНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

PROGRESSIVE INFORMATION TECHNOLOGIES

ПРОГРЕССИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

UDC 004.89

METHOD OF DEFINING FREE PLACES IN VIDEO CONTENT FOR IMPOSITION TYPHLOCOMMENTS

Demchuk A. B. – PhD, Assistant of Information Systems and Networks Department, Lviv Polytechnic National University, Lviv, Ukraine.

Lozynska O. V. – PhD, Assistant of Information Systems and Networks Department, Lviv Polytechnic National University, Lviv, Ukraine.

ABSTRACT

Context. The problem of accessibility of video content is one of the most pressing problem for people with visual impairments. To solve this problem, the methods and means of construction, editing and adaptation of video content for visually impaired persons are development.

Objective. The goal of the work is to develop the method of searching silent areas in the scale for the imposition of typhlocomments and improve the search modules in the software-algorithmic complex of adaptation of video content for visually impaired persons.

Method. The method of searching for places that free from dialogues and other important sounds in video content is implemented. These places of video content are used for inserting typhlocomments. The algorithms of scanning and filtration modules for arrays and searching places available for the imposition of the typhlocomments are developed. This will allow for additional smoothing of the spectrum. Smoothing runs in the forward direction, then in reverse. After calculating the correlation of these two smoothed directions, we see that almost all the short signals are deleted. For the useful signal, the values that smoothed in the forward and reverse direction overlap and therefore remain in the array. Next, the correlation between the smoothed arrays is compared with the set threshold, and if it does not match the set value, then this element of the array is 0. As a result of the algorithm, we have a list of places. For each item in the list, the beginning of the silence place and its length are specified.

Results. On the basis of the developed method of finding places to insert typhlocomments and improved modules, the testing of software-algorithmic complex was performed in standard configuration, and with the additional smoothing module. The first version of the software-algorithmic complex gave the next result: 120 useful pauses. Version with the additional module found 140 useful pauses.

Conclusions. The results of the experiment make it possible to evaluate the developed method and improved modules of scanning, filtration and smoothing. These modules giving a significant gain of results (about 13%), searching for the places for the imposition of typhlocomments, which improves the adapted video content for people with visual impairments.

KEYWORDS: audiodescription, typhlocomments, video content for visually impaired persons, scale, WaveLet transform of signals.

ABBREVIATIONS

HPF is a high-pass filter;
LPF is a low-pass filter;
SAC is a software-algorithmic complex.

NOMENCLATURE

F_i is a set of typhlocomments describing i -th plot;
 K is a threshold of signal energy;
 P_i is a number of plot;
 S^0 is a top-level subsystem;
 S^1 is a bottom-level subsystem;

$Sem(P_i)$ is a semantics of the i -th plot;
 $t(f_i^j)$ is a duration of the i -th plot;
 V is a video content;
 W is a set of information signal;
 X is a set of coordination signal;
 θ is a set of parameters (the position on the time axis, amplitude, width).

INTRODUCTION

More than 80 % of the information from the outside world, a person receives from the vision. Loss of vision

becomes a real information barrier for people with visual impairments throughout their lives. The availability of video content plays an important role in the social adaptation of such people. Nowadays, in Ukraine there are no laws, government programs that aimed at accessing the video content for this people.

As a result of the conducted researches it was found that in Ukraine more than 600 thousand of blind and visually impaired people and only 10% of them are aware of the possibility access to the heritage of culture, art, sports through viewing video content that is adapted for the visually impaired people.

Typhlocoment is a text that describes what is happening at that moment on the screen and read in special pauses with off-screen voice.

In general, typhlocomments are conceived to allow access to video content to people with visual impairments, but at present, potential users of video content with audio description can be anyone.

That's why it is necessary to develop specific software tools for providing video content for people with visual impairments. The scientists of the Information systems and networks department of Lviv Polytechnic National University are working on the developing software-algorithmic complex of access to video content for persons with visual impairments. They developed the method and improved the modules for finding places in the video content, which has allowed to improve the quality of the search places that available for the imposition of typhlocomments by 13%.

The object of study is the process of video content submission for visually impaired persons.

The subject of study is methods and tools for constructing, editing and adapting video content for visually impaired persons.

The purpose of the work is to develop and improve the method of searching silent areas in the scale for the imposition of typhlocomments.

1 PROBLEM STATEMENT

The research objective is to develop and improve SAC modules for the adaptation of video content for visually impaired persons.

To achieve this, the following tasks must be solved:

- 1) to analyze the developed software-algorithmic complex;
- 2) to develop the method of searching silent areas in audio data for the imposition of typhlocomments;
- 3) to construct the algorithms of the necessary modules for finding places in the scale;
- 4) to carry out an experimental comparison of the results of searching places for the insertion of typhlocomments of the software-algorithmic complex with the new places search module.

Any P_i plot is described by a set of typhlocomments $F_i = \{f_i^1, f_i^2, \dots, f_i^{n_i}\}$. We must select from the set F_i the following j -th typhlocoment determined by formula (1):

$$t(f_i^j) \leq t(P_i), \quad (1)$$

and then the corresponding semantics are almost equivalent by formula (2):

$$Sem(P_i) \cong Sem(f_i^j). \quad (2)$$

The realization of condition (1) required that the typhlocoment was not imposed on the dialogues of the plot, and the implementation of condition (2) required the expert assessments.

The i -th plot is divided into sub-plots $S_i \supseteq S_1 \cup S_2 \cup \dots \cup S_{m_i}$ without dialogues, so that the typhlocoment does not overlap with dialogues or other important to understand the plot sounds. Then $t(P)_i = t(P_1) + t(P_2) + \dots + t(P_{m_i})$.

Also required a set of conditions by formula (3):

$$t(f_k^j) \leq t(P_k), k = 1, 2, \dots, m_i. \quad (3)$$

From the set F_i we must select those elements for which condition (3) holds.

2 REVIEW OF THE LITERATURE

The problem of adapting video content for visually impaired persons has been studied from 80s of the last century. The scientists [1, 2] worked on information accessibility for visually impaired people in video format, but adaptation of the video content (finding places for insertion of typhlocomments) has always been performed with the person help.

There are many systems that help visually impaired people in everyday life. In [3] an indoor guidance system for helping blind people is presented. This system provided output instructions to the user by means of a speech synthesizer, was tested in a university building and has shown encouraging results.

In another work [4] system is designed to provide context-dependent guidance messages to blind people while they traverse local pathways. This system is tested under various local pathway scenes, and the results confirm its efficiency in assisting blind people to attain autonomous mobility.

The problem of adaptability of films for visually impaired people is considered in the works [5, 6].

A project DVS Theatrical [7] has been developed in the USA, where 25% of all big screen movies are accessible to people with visually impairments.

A new mobility assistive information and communication technologies for the visually impaired people are described in [8]. Methods for formative and summative evaluations of mobility devices are also discussed.

The most popular software for typhlocomments creation is Media Subtitler [9].

In work audio subtitling as a modality of audiovisual localisation which is positioned at the interface between subtitling, audio description and voice-over are described [10].

The author [11] presents a study aimed at assessing an application offering audio description for mobile devices. The results indicated positive ratings of the assessed features such as usability, utility and quality. The application can be used as a tool for providing improved access to cinema content for people with visually impairments.

The conception, development, and an initial usability evaluation of a software suite for audio description are presented in [12]. The proposal pursues to facilitate and widen a comprehension of videos for people with visual disabilities by using synthesized voice. The authors developed a video player (ADVPlayer) that synchronizes the original video with a second audio containing a text-to-speech version of an audio description script. The evaluation of usability of ADVPlayer revealed a high comprehension and acceptance in terms of satisfaction and confidence.

In [13] a prototype software (audio-vision manager) that uses many computer-vision technologies to automatically extract visual content, associate textual descriptions and add them to the audio track with a synthetic voice is presented. The technology that automatically processes a film offline, extracts visual information such as shot transitions, key-texts, key-places, key-faces and the action of actors, assigns a text to each visual element added to the audio band, and optionally reads it using a voice synthesizer has been implemented.

To search for places in the scale, free from important sounds (replicas of actors, noise mechanisms, etc.), spectral analysis and Wavelet transform are used.

The Wavelet transform of signals [14–16] is a generalization of spectral analysis, the typical representative of which is the classical Fourier transform. The bases used for this purpose are called Wavelet-functions of two arguments – scale and offset. They were introduced in the 80’s of the XX century, but later received fast theoretical development and significant application in various spheres of processing signals and images. In contrast to the traditional Fourier transform, Wavelet transform provides a two-dimensional representation of the investigated signal in the frequency domain in the frequency-position plane. In this case, the analog of frequency is the scale of the argument of the basic function (most often – the time), and the position is characterized by its offset. This allows splitting large and small details of the signals simultaneously to locate them on a timeline. In other words, Wavelet analysis can be described as a localized spectral analysis. Wavelet transform is used in the processing and coding of signals and images of different types (language, satellite images, radiographs of internal organs), image recognition, during studying the properties of crystals and nanobjects surfaces, and in many other cases [17, 18].

3 MATERIALS AND METHODS

Processing signal is the process of transformation and evaluation of its informative characteristics by any technical system. Any signal can be characterized by some generalized values – energy, power, moments or other function, for example, a signal auto-correlation function. The signal can also be presented as a sum of simple oscillations (for example, a sinusoid) and characterized by its set of numbers, called the spectrum, which determines the proportion of each oscillation in the signal. If the signal model is known, then it is characterized by the vector (set) of parameters θ of this model. Any signal parameter carries useful information about the studied object. Therefore, the task of processing is to select these parameters and evaluate their values, from which then the information about the investigated processes is extracted.

The process of inserting the typhlocoment is hierarchical. Hierarchy of layers is shown on Fig. 1. There are three components: top-level subsystem S^0 (audiodescription), bottom-level subsystem S^1 (choosing the place for typhlocoment) and video content V . For mathematical support of the audio description process the coordination theory is used.

The elements of the set X are the coordinating signals that subsystem S^0 uses to operate of bottom-level subsystem S^1 . Top-level subsystem S^0 called a coordinator because its outputs $x \in X$ are signals for coordinating system S^1 . The subsystem S^0 has one input – information ω , which is obtained from bottom-level subsystem using feedback. This input used to form a coordination action x . That is described by formula (5):

$$S^0 : W \rightarrow X. \quad (5)$$

The top-level subsystem defines where to insert the typhlocoment and starts the module of typhlocoment reading.

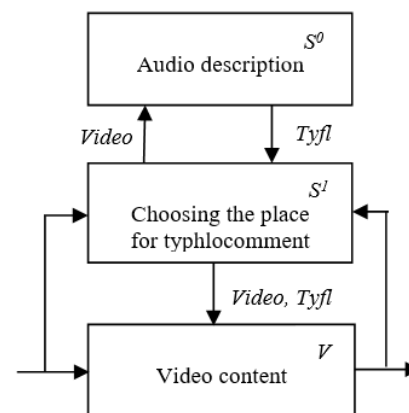


Figure 1 – The hierarchy of audio description process

The search process in the software-algorithmic complex of places for the imposition of typhlocomments includes the following parameters (Fig. 2):

1. $K, \%$ – is the threshold of signal energy, in case of exceeding of signal the search algorithm begins its operation.
2. Spectrum density, $\%$ – is the density spectrum threshold. If this threshold is exceeded, then the program passes to the next analysis. This parameter indicates the presence of explicitly expressed certain frequency bands in the spectrum relative to others.

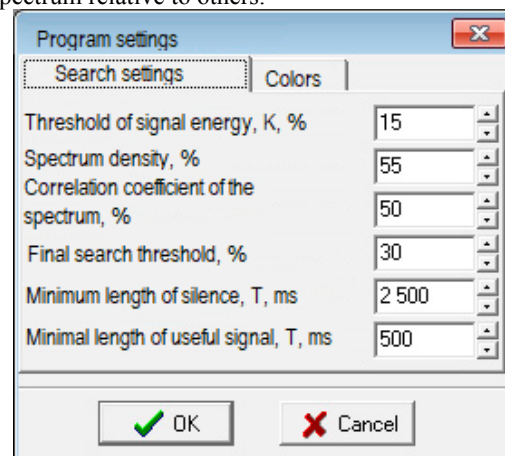


Figure 2 – Options settings window

3. Correlation coefficient of the spectrum, % – is the spectral correlation threshold in each section of the sound signal to the set standard spectrum.

4. The final search threshold, % – is the threshold on the resulting graph of the algorithm’s operation. When the signal is below this threshold, the program considers that corresponding area is a silent and lists it.

5. The minimum length of silence, T , ms – is the time that compares the length of the found area of silence.

6. Minimal length of the useful signal, T , ms – is the time that compares the interval between neighboring areas of silence.

The method of searching silent areas in audio data consists of four steps, which are applied consistently for each 10-ms plot. For this, the audio row preliminary broken up.

Step 1. For each 10 ms plot, the spectrum is calculated due to the WaveLet transform.

Step 2. For analysis, the program takes 1024 samples of audio data and divides into 10 parts of length 1, 2, 4, 8, 16, 32, 64, 128, 256 and 512 samples respectively using WaveLet recursive method. Each part corresponds to its frequency strip.

Step 3. After that, the energy is calculated for each part. Thus, get 10 values of level in accordance with frequency strips for each sound.

Step 4. Each section of the audio track is scanned in four stages:

1. Filtering of the sound section by means of the LPF and HPF passing the frequency range from 50 to 300 Hz. The energy of the received signal is calculated and compared with the passing threshold.

2. For each section of the audio track, if the corresponding element of the array is 255, checking is executed for the density of the spectrum.

3. For each section of the audio track, if the corresponding element of the array is 255, a correlation of its spectrum with the reference spectrum is performed.

If, as the result of the three stages, short signals that had the spectrum similar to the set passed verification, or the short pause between the words of the actor failed the check (there will be a value 0 in the array for this section), then the fourth step of verification is applied.

4. To delete sudden values of 0 or 255, smoothing the array is performed. Initially, the search is performed for sections of the array with values “0”, whose length does not exceed 500 ms. Such sections are considered false, and the value of the array in them is changed to 255. The search of section of array with values “255”, whose length does not exceed 150 ms, is executed. Such plots are considered false and the value of the array in them is changed to 0.

The algorithms of the scanning and filtration modules for arrays and searching places available for the imposition of typhlocomments are developed (Fig. 3–4). This will allow for additional smoothing of the spectrum.

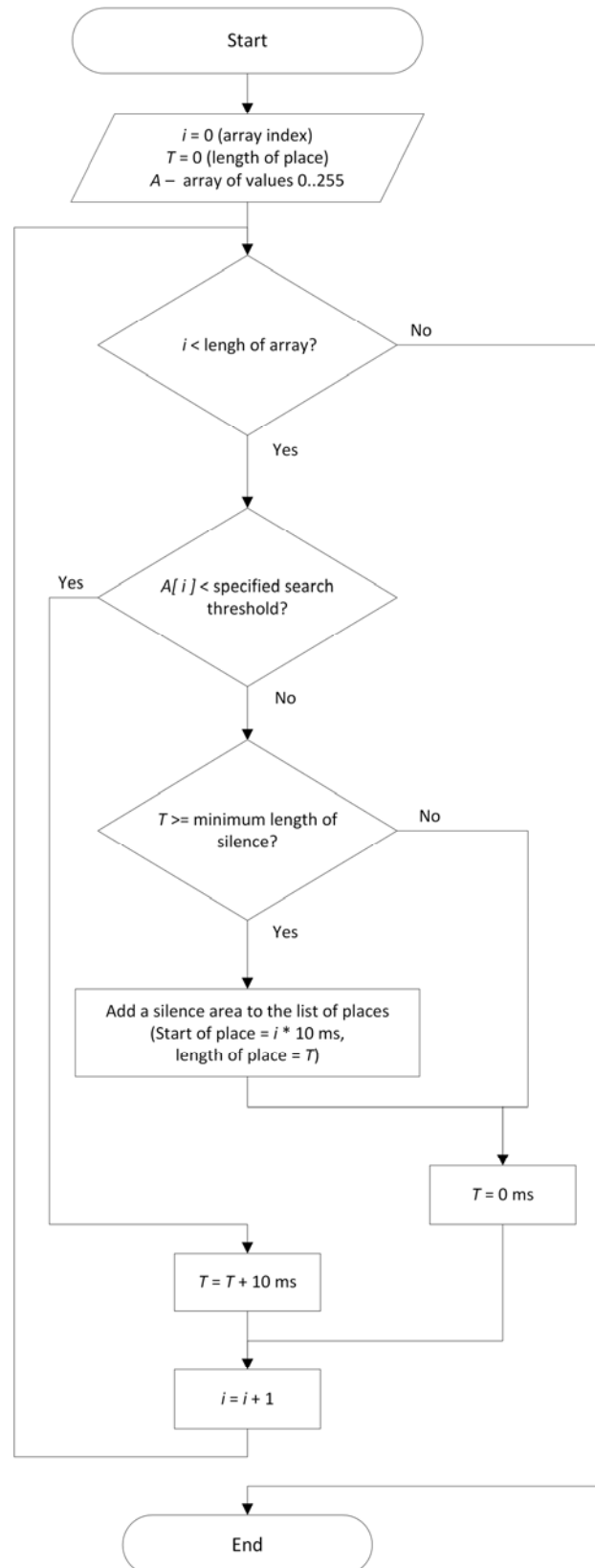


Figure 3 – The function of scanning the array and searching the places for inserting the typhlocomments

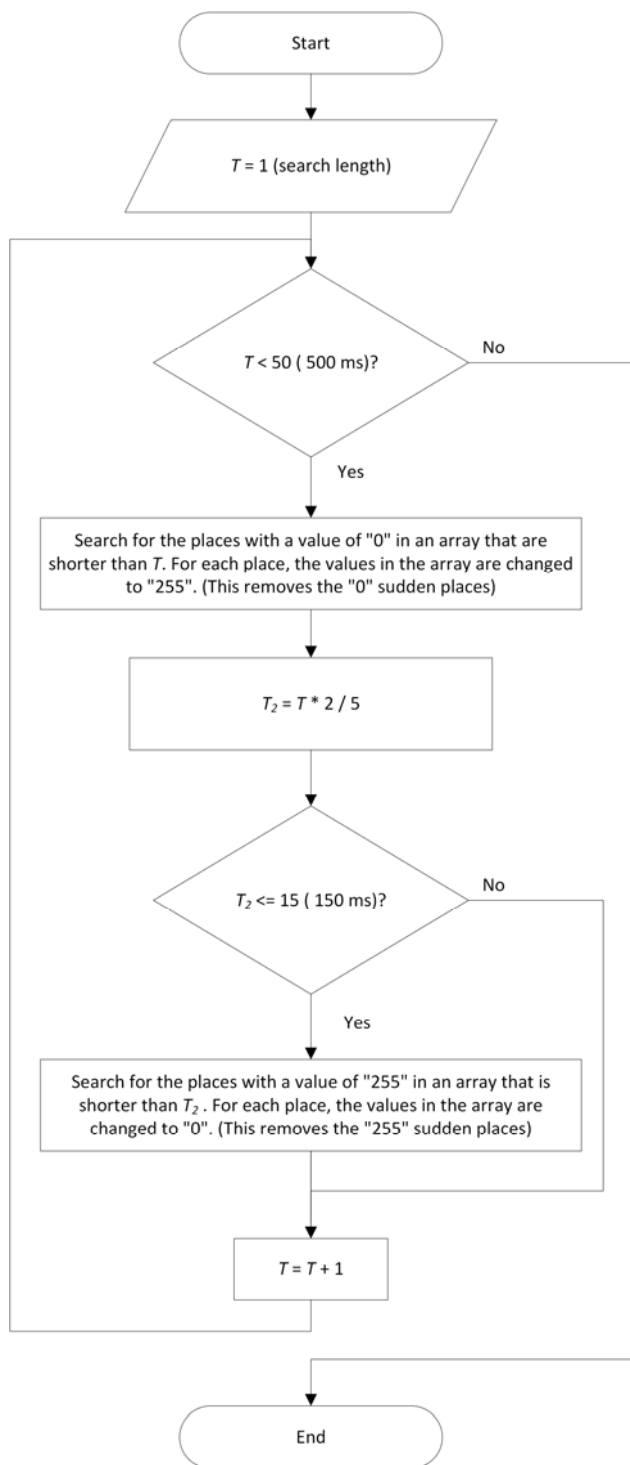


Figure 4 – The array filtration

Smoothing runs in the forward direction, then in reverse. After calculating the correlation of these two smoothed arrays, we see that almost all the short signals

are deleted. For the useful signal, the values that smoothed in the forward and reverse direction overlap and therefore remain in the array. Next, the correlation between the smoothed arrays is compared with the set threshold, and if it does not match the set value, then this element of the array is 0.

The work of the algorithm of SAC “Audio editor” (mode “Visualization of charts of searching places”) is graphically presented in Fig. 5–6.

After applying the search algorithm, we have an array of values: 0 – a silent / non-useful signal, for each section of 10 ms in length, or 255 – a signal that respond all requirements. Accordingly, the developed module of software-algorithmic complex scans the array and finds the parts of the array that are fully filled with a value of 0 and have a length of at least that set in the parameters (for example, 2000 ms, that corresponding to 200 elements of the array). After searching places with the value of 0, the last stage of the algorithm is executed – each place with the value of 255, that situated between neighboring places with the value of 0, is verified. If the length of the place with the value of 255 is shorter than specified in the parameters of the program (for example, 200 ms, that corresponding to 20 elements of the array), then such a place is considered accidentally entered into the useful signal, and it is equal to 0 and combined with neighboring places.

Consequently, as a result of the algorithm, we have a list of places. For each item in the list, the beginning of the silence place and its length are specified.

The module for writing a voice comment directly into an audio file of video content has been developed.

4 EXPERIMENTS

On the basis of the research, developed and improved the method of searching silent areas in audio data for the imposition of typhlocomments. The testing of software-algorithmic complex was performed in standard configuration, and with the additional smoothing module.

The check was carried out on the example of the film by Sergei Loznitsy “Donbass” (2018).

The first version of the software-algorithmic complex gave the next result: 120 useful pauses. Version with the additional module found 140 useful pauses.

The software-algorithmic complex detected pauses [19], and after that it is possible to use the module’s function to read out the typhlocomments immediately into the audio track file. This process is depicted in Fig. 7.

The window contains a stopwatch that shows hours, minutes, seconds, hundredth of the second, and progress that allows you to visually navigate within time, and also buttons of management the record. If the original sound in the silence place obstructs the overlay, you can reduce its volume.

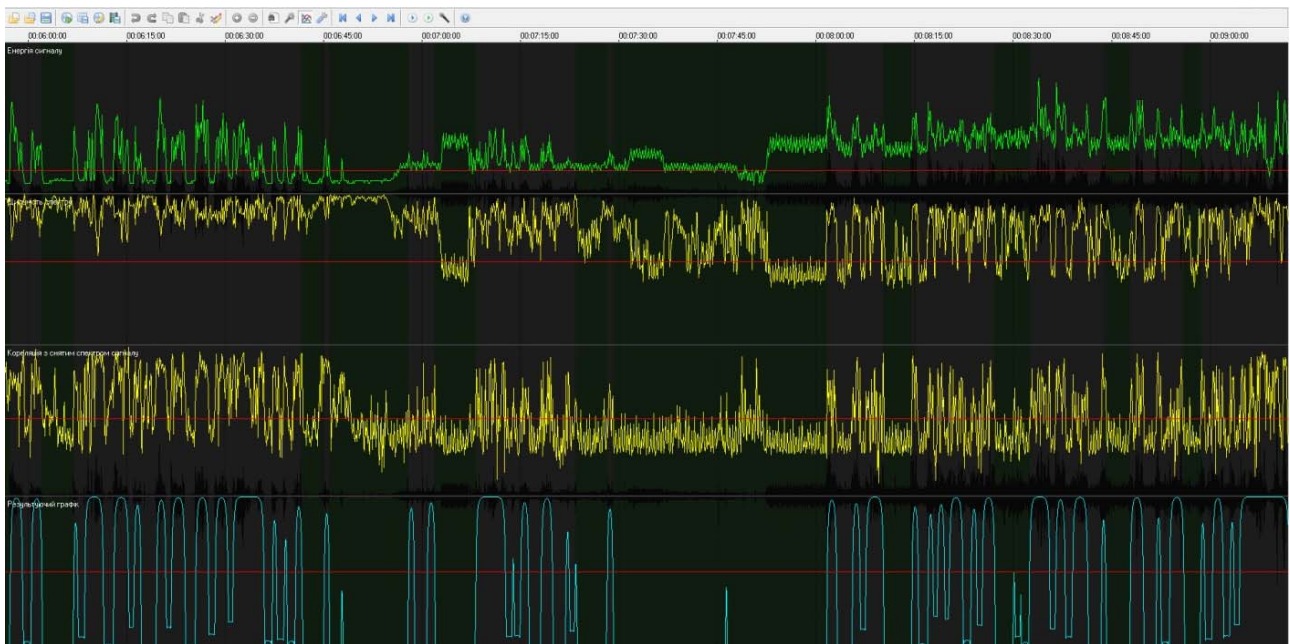


Figure 5 – Visualization of the search algorithm

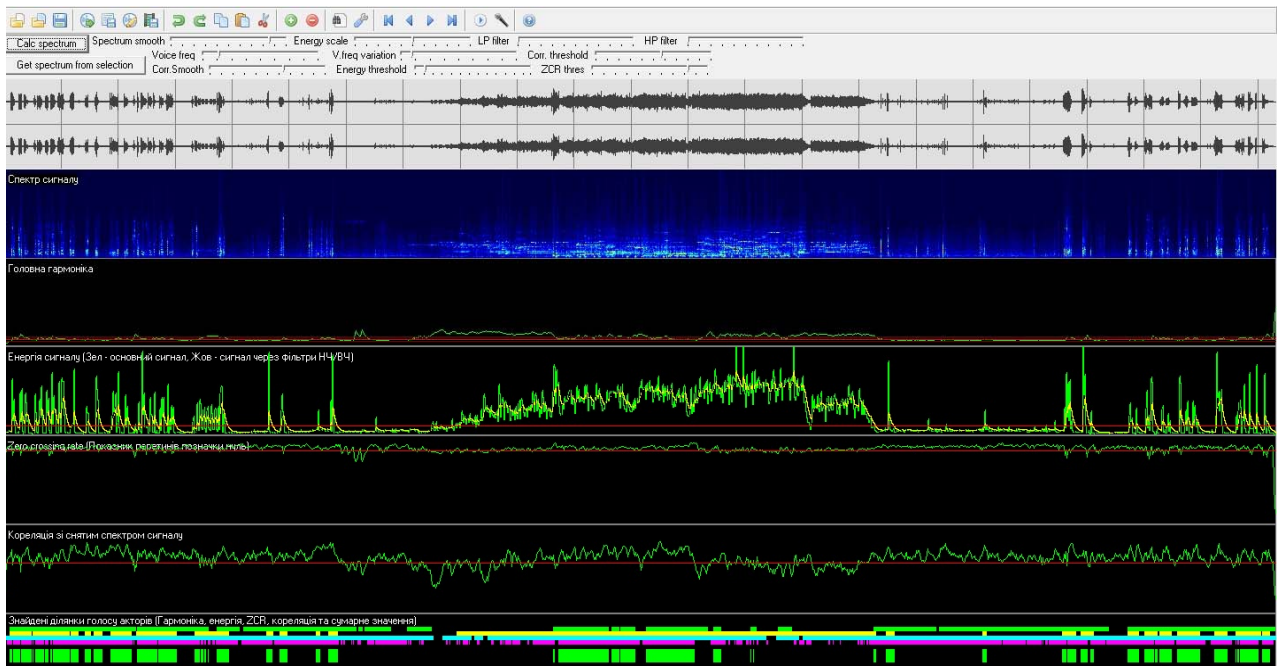


Figure 6 – Visualization of the search algorithm

5 RESULTS

The results of the experiment (see Fig. 8) make it possible to evaluate the developed and improved modules of scanning, filtration and smoothing. These modules giving a significant gain of results (about 13%), searching places for the imposition of typhlocomments, which improves the adapted video content for people with visual impairments.

6 DISCUSSION

Today, scientists around the world are developing information technologies and software that help people with

physical disabilities interact with other people and computers and provide themselves with the necessary information. The problem of information availability for the visually impaired people is caused by a poor quality presentation of information on sites and small-adapted modern television for the needs of such people.

There are no developed information systems in the world that would provide a full range of adaptation process of video content for visually impaired people. There are single systems that only perform some functions, but they are commercial and they have some disadvantages. This is due to the complexity of a daptation and integ-

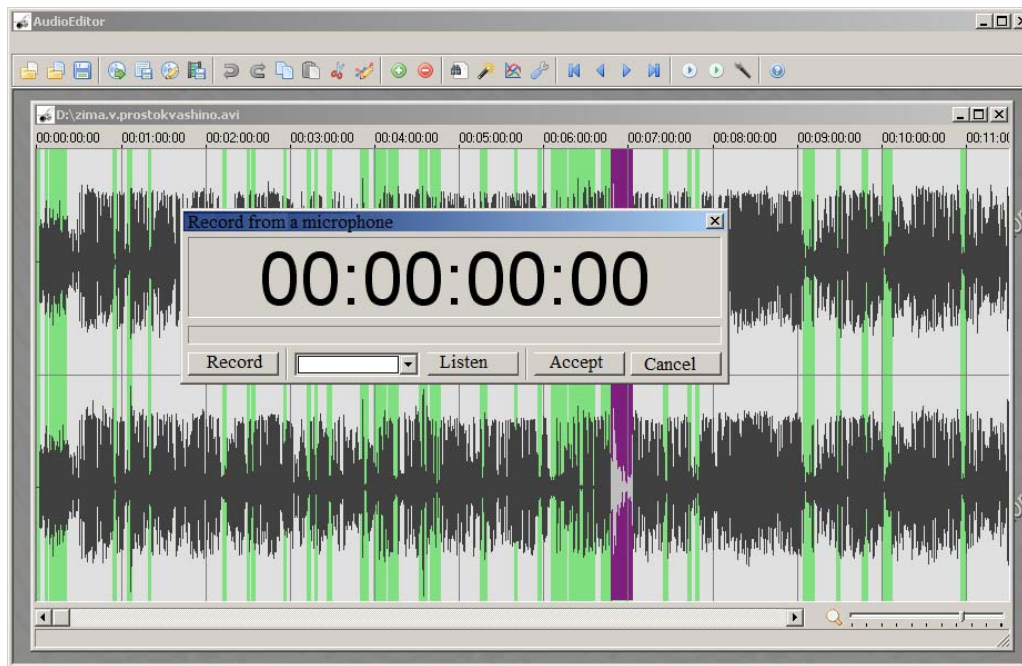


Figure 7 – Window of choice of the audio file area with the ability to record the typhlocomments

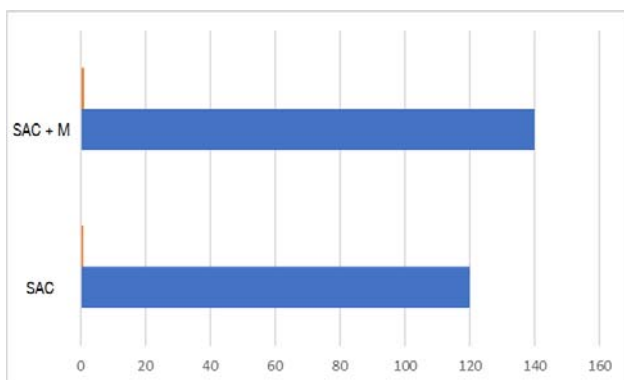


Figure 8 – Number of useful pauses that SAC can detected with the module and without it

rating different software and hardware into one system. As a result, there are no valuable software-algorithmic tools that would provide the solution to the formulated problem.

The experimental results, that have been achieved using the special module, show significant increase about 13% of search quality of the places available for the imposition of typhlocomments in comparison with results of software developed in [20]. This result was achieved by using the method of searching silent areas in audio data for the imposition of typhlocomments and the algorithms of the scanning and filtration modules for arrays. This will allow for additional smoothing of the spectrum.

Smoothing runs in the forward direction, then in reverse. After calculating the correlation of these two smoothed arrays, we see that almost all the short signals are deleted. Next, the correlation between the smoothed arrays is compared with the set threshold, and if it does not match the set value, then this element of the array is 0.

Consequently, as a result of the algorithm, we have a list of places. For each item in the list, the beginning of the silence place and its length are specified.

CONCLUSIONS

This article demonstrated efficient method and algorithms of the scanning and filtration modules for arrays and searching places available for the imposition of the typhlocomments. This method allowed to improve the quality of the search for available places for the imposition of typhlocomments by 13%.

Audio description, as an opportunity to adapt video content for people with visual impairments, is really a method of reporting the plot that is currently happening on the screen.

The scientific novelty of obtained results is that the method of searching places available for the imposition of the typhlocomments are firstly proposed.

The practical significance of obtained results is that the software-algorithmic complex of access to video content for persons with visual impairments realizing the modules for finding places in the video content. The check of SAC was carried out on the example of the film by Sergei Loznitsy "Donbass". Version with the additional module found 140 useful pauses.

Prospects for further research are to focus on implementation developed methods and algorithms for audio description of video content for people with visual impairments.

ACKNOWLEDGEMENTS

The work was supported by the state budget scientific research project of Lviv Polytechnic National University and Ministry of Education and Science of Ukraine "Mathematical and algorithmic modeling of translation

processes from gestures into text for specialized computer systems” (state registration number 0111U001222).

REFERENCES

1. Szarkowska A. Auteur Description: from the director’s creative vision to audio description, *Journal of Visual Impairment & Blindness*, 2013, Vol. 7, pp. 383–387. DOI: 10.1177/0145482X1310700507
2. Igune G. W. Inclusion of Blind Children In Primary Schools: A case study of teachers’ opinions in Moroto district. Uganda, Norway, University of Oslo, 2009, 108 p.
3. Mekhalfi M. L., Melgani F. An Indoor Guidance System for Blind People, *International Conference on Industry Academia Collaboration*. Cairo, Egypt, March 2014, proceedings, 2014, Vol. 1. DOI: 10.13140/RG.2.1.4192.5204
4. Qing L., Youngjoon H. A Context-Aware-Based Audio Guidance System for Blind People Using a Multimodal Profile Model, *Sensors (Basel)*, 2014, Vol. 14(10), pp. 18670–18700. DOI: 10.3390/s141018670
5. Remael A. Audio Description with Audio Subtitling for Dutch Multilingual Films: Manipulating Textual Cohesion on Different Levels, *June 2012 Meta: Translators’ Journal*, 2012, Vol. 57(2), pp. 385–407. DOI: 10.7202/1013952ar
6. Szarkowska A., Jankowska A. Audio describing foreign films, *The Journal of Specialised Translation*, 2015, Vol. 23, pp. 243–269.
7. Cronin B., King S. R. DVS Theatrical [Electronic resource]. Access mode: <http://www2.edc.org/NCIP/library/v&c/cronin.html>.
8. Chottin M. Mobility Technologies for Visually Impaired People Through the Prism of Classic Theories of Perception, *Mobility of Visually Impaired People*. Springer International Publishing, 2018, Section 3, pp. 77–108. DOI: 10.1007/978-3-319-54446-5_3
9. Media Subtitler [Electronic resource]. Access mode: <https://www.divxland.org/en/media-subtitler/>.
10. Sabine B., Braun S., Orero P. Audio description with audio subtitling – an emergent modality of audiovisual localization *Perspectives Studies in Translatology*, 2010, Vol. 18(3), pp. 173–188. DOI:10.1080/0907676X.2010.485687
11. Walczak A. Audio description on smartphones: making cinema accessible for visually impaired audiences, *Universal Access in the Information Society*, 2017, Vol. 17, pp. 833–840. DOI: 10.1007/s10209-017-0568-2
12. Façanha A. R., de Oliveira A. C., Lima M. V., Viana W., Sánchez J. Audio Description of Videos for People with Visual Disabilities, *International Conference on Universal Access in Human-Computer Interaction, 21 June 2016 : proceedings*. Cham, Springer, 2016, pp. 505–515. (Lecture Notes in Computer Science, Vol. 9739). DOI: 10.1007/978-3-319-40238-3_48
13. Gagnon L., Foucher S., Byrns D., Chapdelaine C., Turner J., Mathieu S., Laurendeau D., Nguyen N. T., Ouellet D. Towards computer-vision software tools to increase production and accessibility of video description for people with vision loss, *Universal Access in the Information Society*, 2009, No. 8(3), pp. 199–218. DOI: 10.1007/s10209-008-0141-0
14. Donoho D. Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data : project report : 437. Stanford University, Stanford, 1993, 34 p.
15. Graps A. An Introduction to Wavelets, *IEEE Computational Science and Engineering, Summer 1995: proceedings*. Los Alamitos, IEEE, 1995, pp. 1–18. DOI: 10.1109/99.388960
16. Meyer Y. Wavelets: Algorithms and Applications. Society for Industrial and Applied Mathematics, Philadelphia, 1993, 133 p.
17. Kaiser G. A Friendly Guide to Wavelets. Birkhauser, Boston, 2010, 321 p.
18. Polikar R. The wavelet tutorial. Second edition. Part I [Electronic resource]. Access mode: <http://web.iitd.ac.in/~sumeet/WaveletTutorial.pdf>
19. Lytvyn V. V., Demchuk A. B., Oborska O. V. Mathematical and software submission video content for visually impaired people, *Radio Electronics, Computer Science, Control*, 2016, Vol. 3, pp. 73–79. DOI: doi.org/10.15588/1607-3274-2016-3-9

Received 07.02.2020.
Accepted 22.09.2020.

УДК 004.89

МЕТОД ВИЗНАЧЕННЯ ВІЛЬНИХ ДІЛЯНОК У ВІДЕОКОНТЕНТІ ДЛЯ НАКЛАДАННЯ ТИФЛОКОМЕНТАРІВ

Демчук А. Б. – канд. техн. наук, асистент кафедри «Інформаційні системи та мережі», Національний університет «Львівська політехніка», Львів, Україна.

Лозинська О. В. – канд. техн. наук, асистент кафедри «Інформаційні системи та мережі», Національний університет «Львівська політехніка», Львів, Україна.

АНОТАЦІЯ

Актуальність. Проблема доступності відеоконтенту є однією з найактуальніших проблем для людей із вадами зору. Для вирішення цієї проблеми, розроблено методи та засоби побудови, редагування та адаптації відеоконтенту для осіб з вадами зору.

Мета роботи. Метою роботи є розроблення методу пошуку вільних від діалогів ділянок у звукоряді для вставки тифлокоментарів та вдосконалення модулів пошуку цих ділянок у програмно-алгоритмічному комплексі адаптації відеоконтенту для осіб з вадами зору.

Метод. Реалізовано метод пошуку вільних від діалогів та інших важливих звуків місць у відеоконтенті. Ці знайдені ділянки відеоконтенту використовуються для вставки тифлокоментарів. Розроблено алгоритми модулів сканування та фільтрації масивів та пошуку у них ділянок, доступних для накладання тифлокоментарів. Це дало змогу провести додаткове згладжування спектру.

Виконується згладжування в прямому напрямку, далі у зворотньому. Розрахувавши кореляцію цих двох згладжених масивів ми побачимо, що короткі сигнали практично всі видаляються, а для корисного сигналу, згладжені в прямому і протилежному напрямку значення перекриваються, і тому залишаються в масиві. Далі кореляція між згладженими масивами порівнюється з встановленим порогом, і якщо він не відповідає встановленому значенню, то цей елемент масиву дорівнює 0.

В результаті роботи алгоритму отримано список ділянок для вставки тифлокоментарів. Для кожного елемента списку задається початок ділянки тиші та її довжина.

Результати. На основі розробленого методу пошуку місць для вставки тифлокоментарів та вдосконалених модулів, було проведено тестування програмно-алгоритмічного комплексу у стандартній комплектації, та з модулем додаткового згладжування. Перша версія програмно-алгоритмічного комплексу дала результат у 120 корисних пауз. Версія з додатковим модулем – 140 корисних пауз.

Висновки. Результати експерименту дають змогу оцінити, що розроблений метод та вдосконалені модулі сканування, фільтрації масивів та згладжування, дають суттєвий приріст результатів (близько 13%) пошуку ділянок для накладання тифлокоментарів, що у свою чергу покращує створюваний адаптований відеоконтент для осіб з порушеннями зору.

КЛЮЧОВІ СЛОВА: тифлокоментування, тифлокоментар, відеоконтент для осіб з вадами зору, звукоряд, Вейвлет-перетворення сигналів.

УДК 004.89

МЕТОД ОПРЕДЕЛЕНИЯ СВОБОДНЫХ УЧАСТКОВ В ВИДЕОКОНТЕНТЕ ДЛЯ НАЛОЖЕНИЯ ТИФЛОКОММЕНТАРИЕВ

Демчук А. Б. – канд. техн. наук, ассистент кафедры «Информационные системы и сети», Национальный университет «Львовская политехника», Львов, Украина.

Лозинская О. В. – канд. техн. наук, ассистент кафедры «Информационные системы и сети», Национальный университет «Львовская политехника», Львов, Украина.

АННОТАЦИЯ

Актуальность. Проблема доступности видеоконтента является одной из самых актуальных проблем для людей с нарушениями зрения. Для решения этой проблемы, разработаны методы и средства построения, редактирования и адаптации видеоконтента для лиц с нарушениями зрения.

Цель работы. Целью работы является разработка метода поиска свободных от диалогов участков в звукоряде для вставки тифлокоментариев и усовершенствование модулей поиска этих участков в программно-алгоритмическом комплексе адаптации видеоконтента для лиц с нарушениями зрения.

Метод. Реализован метод поиска мест, свободных от диалогов и других важных звуков в видеоконтенте. Эти места видеоконтента используются для вставки тифлокоментариев. Разработаны алгоритмы модулей сканирования и фильтрации массивов и поиска в них участков, доступных для наложения тифлокоментариев. Это позволило провести дополнительное сглаживание спектра.

Выполняется сглаживание в прямом направлении, дальше в обратном. Рассчитав корреляцию этих двух сглаженных массивов, мы увидим, что короткие сигналы практически все удаляются, а для полезного сигнала, сглаженные в прямом и обратном направлении значение перекрываются, и поэтому остаются в массиве. Далее корреляция между сглаженными массивами сравнивается с установленным порогом, и если он не соответствует установленному значению, то этот элемент массива равен 0.

В результате работы алгоритма получено список участков для вставки тифлокоментариев. Для каждого элемента списка задается начало бесшумного участка и ее длина.

Результаты. На основе разработанного метода поиска мест для вставки тифлокоментариев и усовершенствованных модулей, было проведено тестирование программно-алгоритмического комплекса в стандартной комплектации, и с модулем дополнительного сглаживания. Первая версия программно-алгоритмического комплекса дала результат в 120 полезных пауз. Версия с дополнительным модулем – 140 полезных пауз.

Выводы. Результаты эксперимента позволяют оценить, что разработанный метод и усовершенствованные модули сканирования, фильтрации массивов и сглаживания, дают существенный прирост результатов (около 13%) поиска участков для наложения тифлокоментариев, что в свою очередь улучшает создаваемый адаптированный видеоконтент для лиц с нарушениями зрения.

КЛЮЧЕВЫЕ СЛОВА: тифлокомментирование, тифлокоментар, видеоконтент для лиц с нарушениями зрения, звукоряд, Вейвлет-преобразование сигналов.

ЛІТЕРАТУРА / ЛИТЕРАТУРА

1. Szarkowska A. Auteur Description: from the director's creative vision to audio description / A. Szarkowska // Journal of Visual Impairment & Blindness. – 2013. – Vol. 7. – P. 383–387. DOI: 10.1177/0145482X1310700507
2. Igune G. W. Inclusion of Blind Children In Primary Schools: A case study of teachers' opinions in Moroto

- district / G. W. Igune. – Uganda. Norway: University of Oslo. – 2009. – 108 p.
3. Mekhalfi M. L. An Indoor Guidance System for Blind People / M. L. Mekhalfi, F. Melgani // International Conference on Industry Academia Collaboration, Cairo, Egypt, March 2014: proceedings. – 2014. – Vol. 1. DOI: 10.13140/RG.2.1.4192.5204
 4. Qing L. A Context-Aware-Based Audio Guidance System for Blind People Using a Multimodal Profile Model / L. Qing, H. Youngjoon // Sensors (Basel). – 2014. – Vol. 14(10). – P. 18670–18700. DOI: 10.3390/s141018670
 5. Remael A. Audio Description with Audio Subtitling for Dutch Multilingual Films: Manipulating Textual Cohesion on Different Levels / A. Remael // June 2012 Meta: Translators' Journal. – 2012. – Vol. 57(2). – P. 385–407. DOI: 10.7202/1013952ar
 6. Szarkowska A. Audio describing foreign films / A. Szarkowska, A. Jankowska // The Journal of Specialised Translation. – 2015. – Vol. 23. – P. 243–269.
 7. Cronin B. DVS Theatrical [Electronic resource] / B. J. Cronin, S. R. King. – Access mode: <http://www2.edc.org/NCIP/library/v&c/cronin.html>.
 8. Chottin M. Mobility Technologies for Visually Impaired People Through the Prism of Classic Theories of Perception / M. Chottin // Mobility of Visually Impaired People. – Springer International Publishing, 2018. – Section 3. – P. 77–108. DOI: 10.1007/978-3-319-54446-5_3
 9. Media Subtitler [Electronic resource]. – Access mode: <https://www.divxland.org/en/media-subtitler/>.
 10. Sabine B. Audio description with audio subtitling – an emergent modality of audiovisual localization / S. Braun, P. Orero // Perspectives Studies in Translatology. – 2010. – Vol. 18(3). – P. 173–188. DOI:10.1080/0907676X.2010.485687
 11. Walczak A. Audio description on smartphones: making cinema accessible for visually impaired audiences / A. Walczak // Universal Access in the Information Society. – 2017. – Vol. 17. – P. 833–840. DOI: 10.1007/s10209-017-0568-2
 12. Audio Description of Videos for People with Visual Disabilities / [A. R. Façanha, A. C. de Oliveira, M. V. Lima et al] // International Conference on Universal Access in Human-Computer Interaction, 21 June 2016 : proceedings. Cham : Springer, 2016. – P. 505–515. – (Lecture Notes in Computer Science, Vol. 9739). DOI: 10.1007/978-3-319-40238-3_48
 13. Gagnon L. Towards computer-vision software tools to increase production and accessibility of video description for people with vision loss / L. Gagnon, S. Foucher, D. Byrns, C. Chapdelaine, J. Turner, S. Mathieu, D. Laurendeau, N. T. Nguyen, D. Ouellet. // Universal Access in the Information Society. – 2009. – № 8(3). – P. 199–218. DOI: 10.1007/s10209-008-0141-0
 14. Nonlinear Wavelet Methods for Recovery of Signals, Densities, and Spectra from Indirect and Noisy Data : project report : 437 / D. Donoho / Stanford University. – Stanford, 1993. – 34 p.
 15. Graps A. An Introduction to Wavelets / A. Graps // IEEE Computational Science and Engineering, Summer 1995: proceedings. – Los Alamitos: IEEE, 1995. – P. 1-18. DOI: 10.1109/99.388960
 16. Meyer Y. Wavelets: Algorithms and Applications / Y. Meyer // Society for Industrial and Applied Mathematics, Philadelphia, 1993. – 133 p.
 17. Kaiser G. A Friendly Guide to Wavelets / G. Kaiser. – Birkhauser, Boston, 2010. – 321 p.
 18. Polikar R. The wavelet tutorial. Second edition. Part I [Electronic resource] / R. Polikar. – Access mode: <http://web.iitd.ac.in/~sumeet/WaveletTutorial.pdf>
 19. Lytvyn V. V. Mathematical and software submission video content for visually impaired people / V. V. Lytvyn, A. B. Demchyk, O. V. Oborska // Radio Electronics, Computer Science, Control. – 2016. – Vol. 3. – P. 73–79. DOI: doi.org/10.15588/1607-3274-2016-3-9