

Theoretical and practical aspects of using artificial intelligence technologies in the field of sound design

Oleksandr A. Bobarchuk, Svitlana M. Halchenko, Serhii O. Hnidenko and Ivan P. Zavadetskyi

State Non-Commercial Company “State University “Kyiv Aviation Institute”, 1 Liubomyra Huzara Ave., Kyiv, 03058, Ukraine

Abstract

The theoretical and practical aspects of using artificial intelligence technologies in the field of sound design are considered. An analysis of modern technologies, their capabilities and limitations is conducted, the advantages and risks are examined, and the prospects for development in this field are outlined. The results of the research are aimed at increasing the understanding of the potential of AI in working with sound and determining ways to effectively implement these technologies in the creative process.

Keywords

sound design, artificial intelligence, sound creation for music, Suno AI, sound plugins, visual novels, AudioGen

1. Introduction

Modern artificial intelligence (AI) technologies are becoming an integral part of many spheres of human activity, including creative industries [1, 2]. One such area where AI demonstrates significant potential is sound design. This discipline combines art and technology to create sound compositions used in cinema, video games, advertising, music and other media. The use of AI in sound design opens up new possibilities for process automation, sound generation and interactive sound accompaniment, changing traditional approaches to working with sound.

Despite significant interest in the use of artificial intelligence in creative industries, the topic of AI application in sound design is not yet fully explored in modern scientific literature. Most studies focus on specific aspects such as sound synthesis, audio signal processing or adaptive sound systems for interactive environments. However, a holistic analysis of the theoretical foundations, practical applications, and the impact of these technologies on the industry as a whole remains fragmented.

Some works highlight the technical aspects, describing the algorithms and methods used to generate or process sound. Others focus on applied cases, such as the integration of AI in the production of music or sound effects for cinema and games. Meanwhile, a comprehensive approach that would take into account both creative and technical challenges, ethical aspects and development prospects is still lacking.

This indicates the need for deeper research that would create a general concept of using AI in sound design. This article attempts to fill this gap by analysing not only existing technologies, but also their impact on the process of sound creation, as well as outlining future prospects for this field.

2. Transformation of modern sound design

In the classical sense, sound design is the process of obtaining (generating), editing, and implementing sound elements (samples) in a multimedia composition [3]. It covers a wide range of applications,

AREdu 2024: 7th International Workshop on Augmented Reality in Education, May 14, 2024, Kryvyi Rih, Ukraine

✉ a.bobarchuk@interactiveklass.com (O. A. Bobarchuk); smgalchenko@gmail.com (S. M. Halchenko);

serhii.hnidenko@npp.nau.edu.ua (S. O. Hnidenko); ivan.zavadetskyi@npp.nau.edu.ua (I. P. Zavadetskyi)

ORCID 0000-0003-3176-7231 (O. A. Bobarchuk); 0000-0003-0531-1572 (S. M. Halchenko); 0009-0002-3215-8577 (S. O. Hnidenko);

0000-0002-6854-3971 (I. P. Zavadetskyi)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

including cinema, theatre, video games, advertising, the music industry and even architectural design of sound environments. The main principles of classical sound design include the following aspects [4]:

- **Realism and authenticity.** The main principle underlying the classical approach is the creation of realistic sounds that correspond to the visual or dramatic context.
- **Technical skill.** Sound designers rely on traditional methods of recording sound using microphones, field recorders, analogue and digital processing tools.
- **Foley art.** A special place in classical sound design is occupied by the art of creating sound effects manually, using real objects and materials to imitate various sounds. Real sounds are recorded and processed by appropriate means of artistic sound processing, such as reverberation, echo, chorus, etc., to achieve the desired result.
- **Composition and editing.** The sound designer combines sounds into a sound composition, using editing to achieve the desired rhythm, harmony and dramatic impact. That is, a synergistic combination of sound and dynamic change of images is performed.

The classical approach laid the fundamental principles of sound design, which still remain relevant. However, changes in the digital landscape pose new challenges. Classical methods of sound design have their limitations (for example, lack of adaptability, instrumental and technical limitations, time and resource costs).

The gradual development of digital technologies has also changed the principles and means of sound design. With the advent of VST (Virtual Studio Technology) and AU (Audio Units), sound designers gained access to thousands of digital instruments, simulators of analogue and digital synthesisers, classical musical instruments and effects [5, 6]. This significantly reduced equipment costs and expanded the possibilities for experimentation. The development of the gaming industry stimulated the emergence of adaptive audio systems, where sound changes depending on the player's actions or environment. Wwise (figure 1) and FMOD technologies have become the standard in the field of interactive sound design [7].

Digital sound libraries gradually began to appear – sets of pre-recorded samples that can be quickly applied to one's own projects. The use of ready-made sounds was not a new practice in itself – the 1950s-60s saw the creation of the first commercial libraries storing sounds of gunshots, natural phenomena,

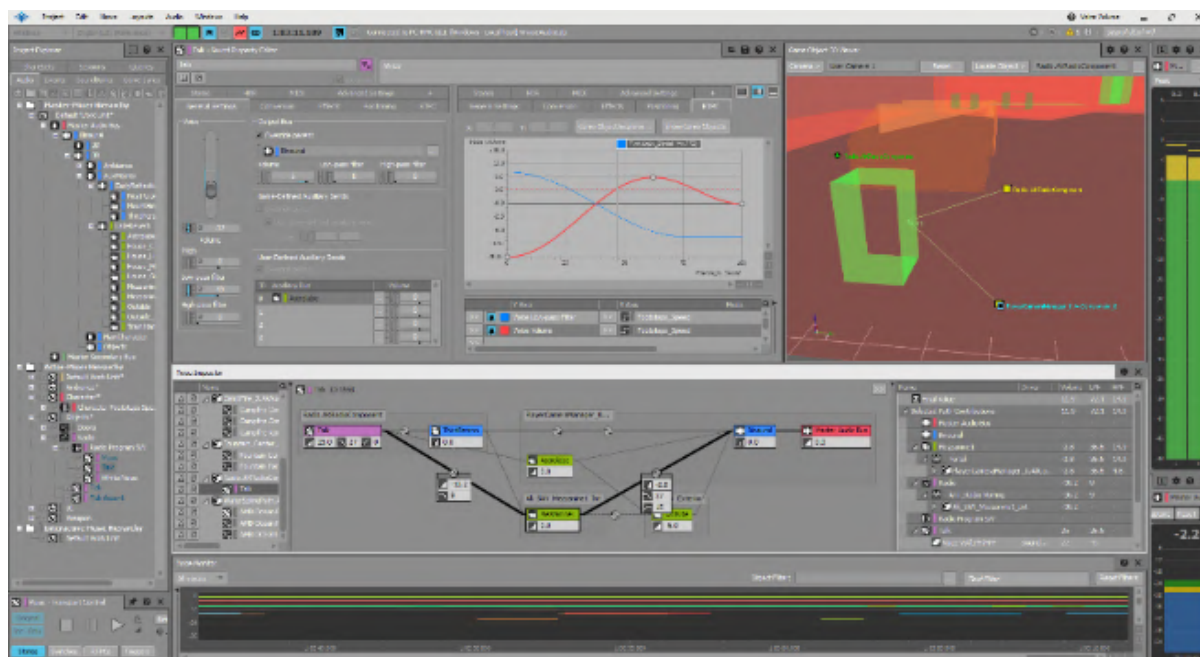


Figure 1: Wwise Software – a tool for creating sound for interactive media and video games.

transport, etc. [8]. They were recorded on analogue media (e.g. magnetic tape) and used in cinema and television.

Modern sound design has reached a level where technologies allow the creation of high-quality sound for various types of media – from cinema and video games to advertising and virtual reality. Audio processing tools have become more powerful, and access to large sound libraries, virtual instruments and modern technical means for recording and processing sound have greatly simplified the process of creating sound content. But the demand for constant improvement remains unchanged. That is why the question arises: how to adapt and use the possibilities of artificial intelligence, which continues to develop comprehensively, in the sound design industry. And how expedient is such use?

3. Artificial intelligence in sound design: main directions

The use of artificial intelligence for sound generation has become one of the most promising areas in the field of sound design. At a basic level, this process is based on the ability of algorithms to learn from large volumes of audio data, analyse them and create new sound textures that can be used in music, cinema, video games and virtual reality [9].

In the early stages of development, artificial intelligence algorithms worked primarily with existing sounds. They could restore audio, remove noise or imitate the sound character of specific instruments. However, with the development of machine learning [10] and neural networks [11], AI has become capable of creating completely new soundscapes that did not exist before. For example, generative adversarial networks (GANs) allow systems to synthesise sounds that have a natural timbre, and recurrent neural networks (RNNs) learn to predict the next sound segments, creating a continuous audio stream.

One of the most striking examples of AI applications is the creation of sounds for music. Algorithms analyse thousands of music tracks, extracting patterns and harmonies, and then generate melodies or rhythms. Programs like AIVA (Artificial Intelligence Virtual Artist) are capable of creating entire compositions in various genres, providing composers with a foundation for further work [12]. In the field of electronic music, AI is often used to create unique samples or synthetic textures that can be integrated into compositions.

At the beginning of 2024, the Suno AI network gained a high level of popularity [13]. Suno AI is an innovative platform that uses artificial intelligence to generate music based on text prompts. The user enters a description of the desired song, specifying the style, genre or theme, and the system creates a corresponding composition. This process takes about two minutes, after which the user receives two versions of the track: one with vocals, the other instrumental.

Suno AI technology is based on artificial intelligence models such as Bark and Chirp, which are capable of generating not only instrumental music but also adding vocal parts to songs. The algorithm analyses the entered text, determines its rhythmic and semantic features, and then synthesises a melody and harmony that match the given description. The vocals are synthesised taking into account the rhythm and intonation of the text, giving the song a natural sound.

The system uses an approach similar to large language models, such as ChatGPT [14]: it splits the text into individual segments (tokens), studies millions of usage variants, styles and structures, and then reconstructs them on request. However, creating audio, especially music, is a more complex task, as it requires taking into account many parameters such as melody, harmony, rhythm and timbre.

Of course, the tracks generated by both Suno AI and other platforms and models have noticeable subjective flaws, which most often manifest themselves in distortions in vocal parts, sharp changes in volume or banal misunderstanding or neglect of prompts. Artificial intelligence can work much more accurately with small sounds. Classic elements of sound design are various sound transitions, for example, a gradually increasing sound (rise), a sound of impact (hit), a sound of cutting air (whoosh), etc.

Artificial intelligence is able to generate and process such short sound effects with high accuracy due to its ability to analyse thousands of samples and extract key sound characteristics. These effects have a

clear structure and predictable dynamics, making them ideal material for algorithm work. AI models can create variations of hits, rises or noises based on text descriptions or user settings, providing precise adjustment of the duration, frequency spectrum and amplitude of each sound.

In addition, thanks to machine learning technologies, artificial intelligence can automatically select sounds for different scenes, creating smooth transitions and adapting them to the visual content. For example, AI can generate a whoosh sound of varying intensity depending on the speed of an object in the frame or synchronise impact effects with moments of climax.

There are already services that provide the ability to create cinematic sounds with a text prompt. But such sounds are only a small part of sound design. For us, sound design is primarily a complex sound landscape, an immersive global environment. Is artificial intelligence capable of forming something like this?

Immersive environment sound design requires not only layering sounds on top of each other, but also fine-tuning spatial acoustics, dynamics and the emotional content of each layer. In the real world, sounds interact with each other in unpredictable ways – echoes in space, gradual fading or swelling, the influence of textures of materials and objects that create or reflect sound. It is difficult for algorithms to reproduce this chaos and versatility of the sound environment in the same way as human hearing and perception.

Currently, artificial intelligence does an excellent job of reconstructing real environments through recordings and spatial analysis, but creating completely fictional sound landscapes that have no analogues in reality requires creative intuition. A human sound designer works not only with sounds as such, but with a concept – they create a story through sound, using audio as a tool to evoke emotions and build atmosphere.

However, there are also positive aspects. Artificial intelligence algorithms are becoming increasingly effective in creating procedural sound landscapes. They are able to analyse visual sequences or text descriptions and generate corresponding sound environments, automatically adding necessary elements: the sound of wind, raindrops, city bustle or any other simple ambient.

Artificial intelligence cannot fully construct a multi-layered sound environment. But it can be used as a tool that provides a certain foundation to work with. For example, it is possible to create simple patterns of classical instruments in a given key and rhythm, perform their gradual processing using classical means in any sound editing environment, mix the tracks, supplement them with various generated sounds, and further integrate the created composition into complex multimedia environments.

Another equally important aspect is the integration of artificial intelligence technologies into various plugins for working with sound. Often these are a wide variety of tools for different tasks, which have appeared quite a lot recently. AI assistants are used to perform general mastering (such as the built-in assistant in Izotope Ozone 10/11), for individual tasks such as compression, limiting, saturation, equalisation, etc. AI plugins are trained on large volumes of audio data. Developers train the neural network based on various recordings manually processed by professional sound engineers. The model analyses how classical tools for saturation, compression or limiting work, and learns patterns of effect application depending on the type of sound, genre or processing style.

When the user loads the plugin, the algorithm performs a multivariate analysis of the audio signal – analyses the frequency spectrum, dynamics, harmonics and noise level. AI models are able to detect problem areas or potentially weak zones and suggest processing parameters. Practical experience shows that these parameters are often not optimal, but can be used as a basis for further work with sound.

Much more interesting from the point of view of sound design are plugins such as Synplant 2. The Genopatch technology (figure 2) built into the plugin allows generating a variety of new sounds based on a single loaded sample [15]. The unique capabilities and interface of Synplant 2 promote experiments in the field of intuitive sound design, allowing to explore how non-standard methods of interaction with technology can influence the creative process.

Considering all of the above, we can note that artificial intelligence is already transforming the field of sound design today, opening up new possibilities for creativity and automation. Despite significant achievements, AI technologies in sound design face certain significant challenges: limited emotional depth of generated sounds, complexity of creating complex immersive environments, and various

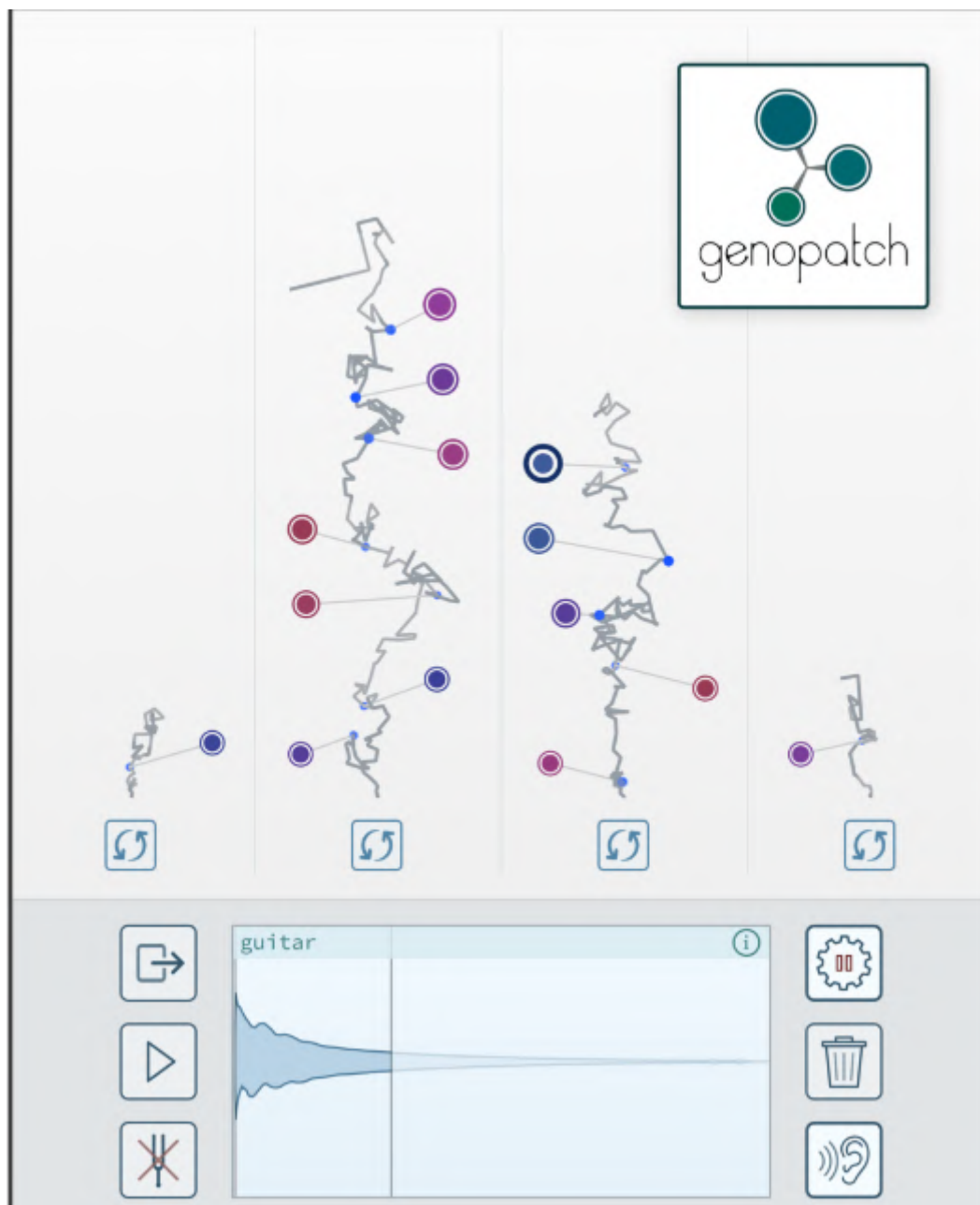


Figure 2: Working in the Genopatch editor.

technical defects that can distort the perception of the overall picture. However, these limitations stimulate the development of the industry and create space for improving algorithms, integrating new approaches and synergy with human creative ideas. AI does not replace sound engineers, but becomes a powerful tool that helps accelerate the workflow and expand creative horizons.

Now, let's demonstrate the possibilities and ways of applying artificial intelligence technologies in sound design through practical experience.

4. Practical aspects of using artificial intelligence technologies in sound design

In this part, as an example of the possibilities of using AI in sound design, we will form a simple sound design for several scenes that are planned to be used in a visual novel. Visual novels (VN) are a genre of

interactive games where the main emphasis is on the plot and characters, and sound plays an important role in creating an emotional response. The scenes themselves were created using Dalle-E 3 and refined in Adobe Photoshop (figure 3).



Figure 3: Prepared scenes.

To begin with, we break down the scenes into components and determine the overall mood and which sounds we need (figure 4).

Dark ambient;
 Low frequency noise;
 Wind noise;
 Melancholic classical instruments.



Humming of wires Cracking of branches

Figure 4: Determining sounds for the scene.

Now let’s determine the necessary artificial intelligence tools. For forming the general landscape, the aforementioned Suno AI from the previous section will work. Using the following prompt, we generate two compositions for download: Violin and piano, melancholic style, slow tempo, dark ambient. We transfer the downloaded result to the FL Studio environment and perform the following sequential processing: slowing down, equalisation, adding reverb and echo effects using the Crystalize granular generator from the developer SoundToys (figure 5).

It is worth noting that the recording generated in Suno AI without further processing would not fully correspond to the general concept of sound design for this project. As already mentioned above, Suno AI often does not interpret prompts very accurately, which leads to problems with generating music in less well-known genres. However, artistic processing tools allow to significantly change and improve the nature of the input sound and adapt it to the needs of the project.

Unlike Suno AI, the AudioGen AI instrument handled the prompt more accurately, generating distant humming of electric wires and cracking of branches with a rather short request (figure 6).

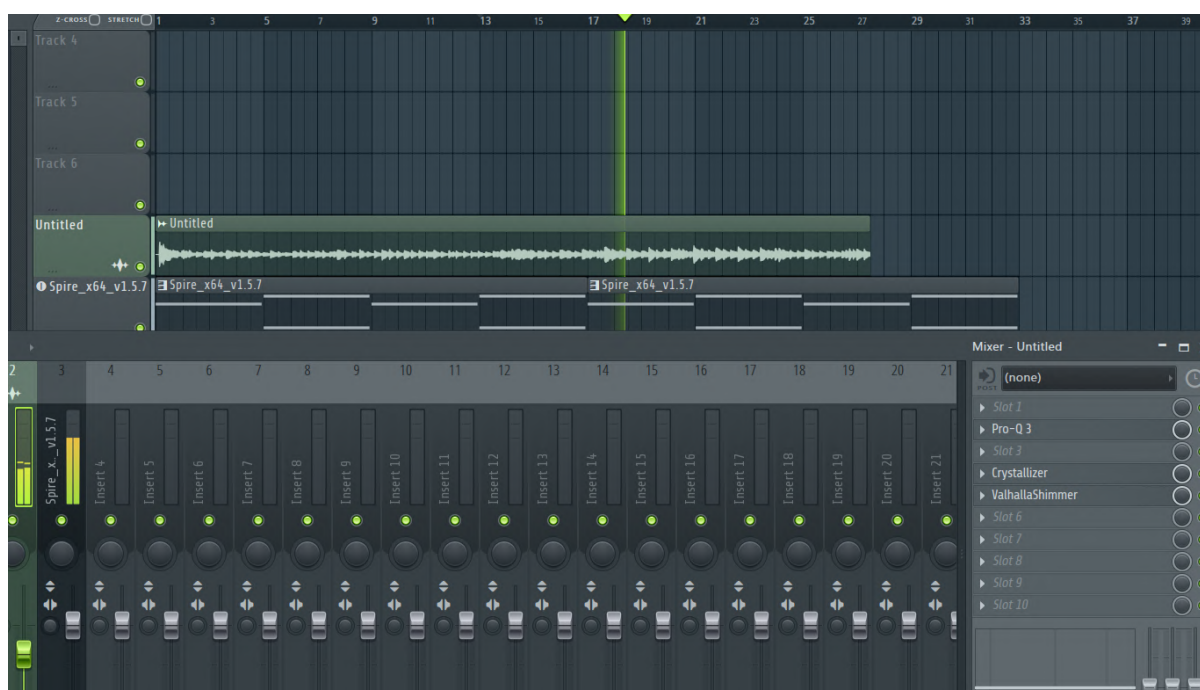


Figure 5: Work in the FL Studio environment.

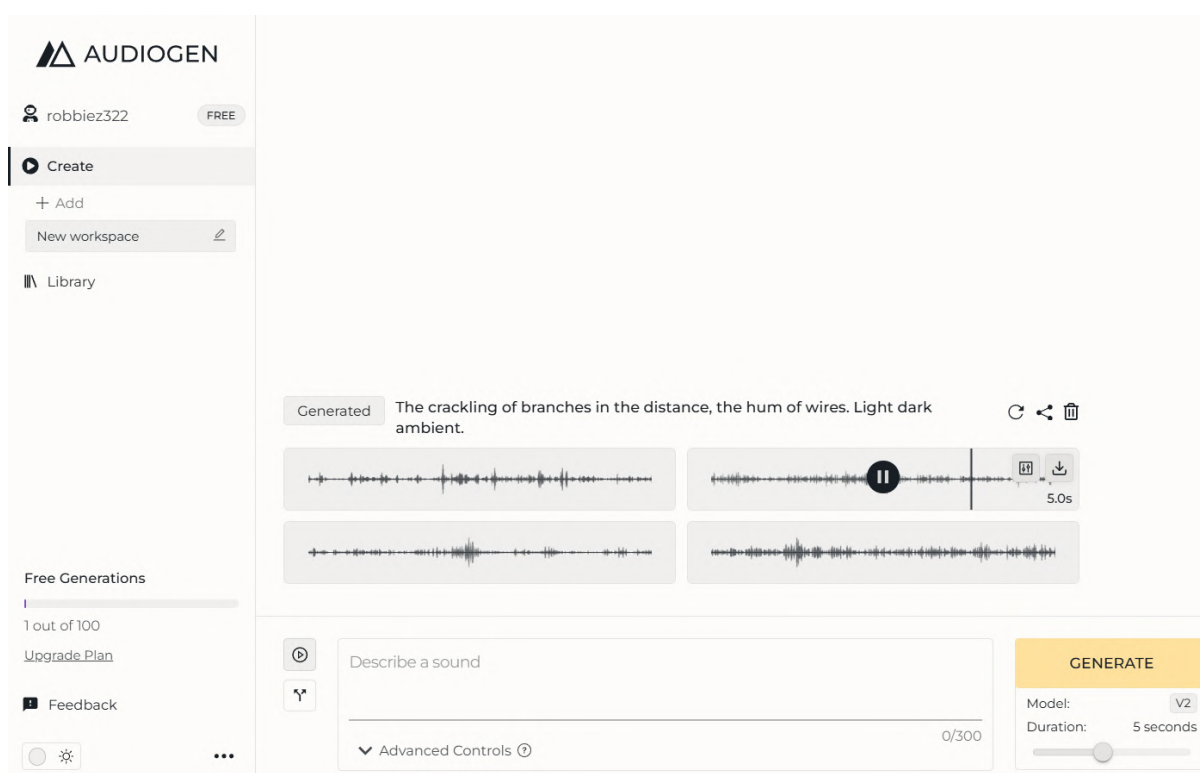


Figure 6: Work with the AudioGen tool.

To create various variations of the sample generated using AudioGen, we will use Synplant2 and its Genopatch technology. We load the sample into the plugin environment, after which Synplant2 automatically generates new sound samples based on the provided one (figure 7).

After combining all the generated sounds, we have as a result a simple but subjectively quite high-quality sample of sound design for scenes from a visual novel. Thus, based on practical experience,

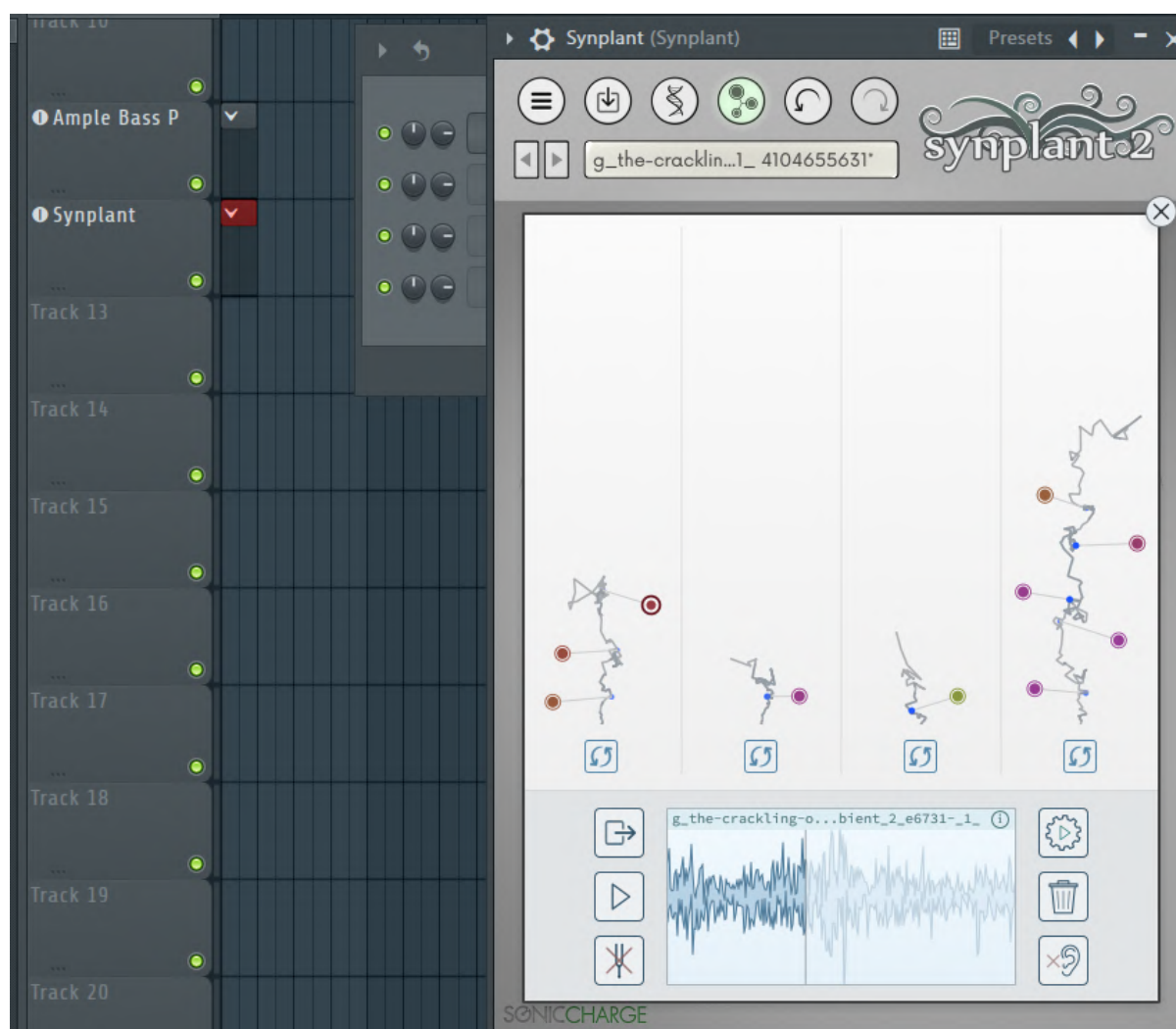


Figure 7: Generation of new sounds using Synplant2.

the feasibility of using artificial intelligence technologies as a tool for quickly obtaining the necessary sound samples for their further processing and combining into a coherent composition was confirmed.

5. Conclusions

The article provides a thorough analysis of the theoretical and practical aspects of using artificial intelligence technologies in the field of sound design. The current state of the industry is highlighted, considering its technical capabilities and creative challenges. A study of classical and modern tools for forming sound environments is conducted.

It has been proven that artificial intelligence significantly changes the traditional approach to sound creation, providing process automation, time savings, and expanded possibilities for experimentation. For the first time, a detailed analysis of the main trends and directions of the impact of artificial intelligence on sound design is presented. Special attention is paid to tools such as Suno AI, AudioGen, Synplant2, which demonstrate significant potential for generating sound textures and integration into creative projects.

The practical aspect of the research is based on the example of creating a sound accompaniment for visual novels, where artificial intelligence was used to generate musical compositions and sound effects. These materials, after further processing, can become the basis for high-quality full-fledged sound design. It is important to emphasise that although AI tools provide speed and adaptability in

working with sound, their results often require refinement to match creative ideas.

In this article, for the first time, an integrated approach to the use of various artificial intelligence tools for creating sound design is proposed. This approach takes into account both technical capabilities and creative needs. The study outlines the advantages of modern AI algorithms, such as efficiency in creating short sound effects, as well as their limitations, including difficulties in forming complex immersive sound landscapes.

Also, for the first time, the article presents a methodology for selecting artificial intelligence tools for specific tasks in the context of sound design for multimedia projects. For example, it is determined that tools such as Suno AI are appropriate for creating music and musical effects, AudioGen for generating sounds of certain environments, and Synplant2 for editing sounds. This methodology is formed on the basis of practical work with these tools and subjective evaluation of the generation results.

The overall results and prospects for further development of this problem can be defined as follows:

- A study of the main directions of using artificial intelligence in sound creation has been conducted, including the generation of musical compositions, short sound effects, and procedural sound landscapes. It is shown that tools like Suno AI, AudioGen, and Synplant2 are able to effectively perform sound generation and processing tasks, which greatly simplifies the complex process of creating sound design;
- The article presents an example of creating sound design for visual novels, which illustrates the capabilities of AI for quickly obtaining basic sound textures. It is shown that artificial intelligence can be used to automate sound creation with further processing and refinement, which allows achieving high-quality final results;
- An integration approach is proposed, which consists in using various artificial intelligence tools for different tasks that may include short musical compositions, simple sound landscapes, and short sounds. Subjective evaluation of the quality of the created samples shows that they are quite suitable for use in various multimedia projects.

The practical significance of the obtained results lies in increasing the efficiency of sound design creation processes through the integration of artificial intelligence technologies. In particular, the proposed methods allow automating routine tasks, such as generating basic sound textures and creating simple sound or musical effects. This reduces the time and resources required for work and allows designers to focus on the creative aspects of projects.

Prospects for further research are primarily related to improving algorithms for creating immersive sound environments, deepening the synergy of AI and human creativity, and more active integration of generated sounds into multimedia projects. These prospects demonstrate the potential for further transformation of the sound design industry, expanding the capabilities of creative professionals and stimulating the development of innovations in the use of artificial intelligence.

The study confirmed the practical value of artificial intelligence in transforming sound design, expanding the toolkit for creating sound compositions and opening up new horizons in creative industries.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

- [1] M. V. Marienko, S. O. Semerikov, O. M. Markova, Artificial intelligence literacy in secondary education: methodological approaches and challenges, *CEUR Workshop Proceedings* 3679 (2024) 87–97.
- [2] I. Mintii, S. Semerikov, Optimizing Teacher Training and Retraining for the Age of AI-Powered Personalized Learning: A Bibliometric Analysis, in: E. Faure, Y. Tryus, T. Vartiainen, O. Danchenko, M. Bondarenko, C. Bazilo, G. Zaspá (Eds.), *Information Technology for Education, Science, and Technics*, volume 222 of *Lecture Notes on Data Engineering and Communications Technologies*, Springer Nature Switzerland, Cham, 2024, pp. 339–357. doi:10.1007/978-3-031-71804-5_23.

- [3] K. Zizza, Sound Design, in: *Game Audio Fundamentals: An Introduction to the Theory, Planning, and Practice of Soundscape Creation for Games*, Focal Press, London, 2023, pp. 142–163. doi:10.4324/9781003218821-11.
- [4] E. Miranda, Computer sound synthesis fundamentals, in: *Computer Sound Design: Synthesis techniques and programming*, 2 ed., Routledge, New York, 2012, pp. 19–36. doi:10.4324/9780080490755-7.
- [5] A. Rosiński, The Use of Virtual Musical Instruments in Timbre Recognition Training, *International Journal of Learning and Teaching* 9 (2023) 256–257. doi:10.18178/ijlt.9.3.256-260.
- [6] T. Suzuki, A. Nakabayashi, Virtual Studio, *The Journal of The Institute of Image Information and Television Engineers* 61 (2007) 657–659. doi:10.3169/itej.61.657.
- [7] A. Zecevic, G. Durity, *Handbook of Game Audio Using Wwise*, Taylor & Francis Group, 2020.
- [8] M. Katz, Music in 1s and 0s: The Art and Politics of Digital Sampling, in: *Capturing Sound: How Technology has Changed Music*, University of California Press, Berkeley, 2004, pp. 137–157. URL: https://ia600409.us.archive.org/29/items/mat-bib_201710/Capturing-sound-how-technology-has-changed-music.pdf.
- [9] K. Saraf, M. D. Amritphale, H. Akhand, K. Vijayvargiya, Music AI, *International Research Journal of Modernization in Engineering Technology and Science* 6 (2024) 11174–11177. doi:10.56726/irjmets54679.
- [10] P. V. Zahorodko, S. O. Semerikov, V. N. Soloviev, A. M. Striuk, M. I. Striuk, H. M. Shalatska, Comparisons of performance between quantum-enhanced and classical machine learning algorithms on the IBM Quantum Experience, *Journal of Physics: Conference Series* 1840 (2021) 012021. doi:10.1088/1742-6596/1840/1/012021.
- [11] S. Semerikov, H. Kucherova, V. Los, D. Ocheretin, Neural network analytics and forecasting the country’s business climate in conditions of the coronavirus disease (COVID-19), *CEUR Workshop Proceedings* 2845 (2021) 22–32.
- [12] Aiva Technologies SARL, AIVA, the AI Music Generation Assistant, 2025. URL: <https://www.aiva.ai/>.
- [13] Suno, Inc., Suno, 2025. URL: <https://suno.com/>.
- [14] R. Liashenko, S. Semerikov, The Determination and Visualisation of Key Concepts Related to the Training of Chatbots, in: E. Faure, Y. Tryus, T. Vartiainen, O. Danchenko, M. Bondarenko, C. Bazilo, G. Zaspá (Eds.), *Information Technology for Education, Science, and Technics*, volume 222 of *Lecture Notes on Data Engineering and Communications Technologies*, Springer Nature Switzerland, Cham, 2024, pp. 111–126. doi:10.1007/978-3-031-71804-5_8.
- [15] NuEdge Development, Sonic Charge - Synplant, 2025. URL: <https://soniccharge.com/synplant>.