

Система візуалізації музичної інформації

Ярошенко Р. О., ORCID [0000-0002-5360-5885](https://orcid.org/0000-0002-5360-5885)

Кафедра акустичних та мультимедійних електронних систем ames.kpi.ua
Національний технічний університет України
«Київський політехнічний інститут ім. Ігоря Сікорського», ROR [00syn5v21](https://ror.org/00syn5v21)
Київ, Україна

Анотація—Системи візуалізації музичної інформації поширені в основному у вигляді програмного забезпечення для персональних комп'ютерів. У роботі представлена система, яка обробляє музичні дані, що надходять із зовнішнього джерела музики через канал Bluetooth, а саме візуалізує співвідношення амплітуд спектральних складових й має фіксовану прив'язку частоти до кольорів. Розглянуто технологію обробки аудіо сигналу пристроєм, та відповідні компоненти приладу. Для забезпечення швидкості обробки інформації використано 32-х бітний мікроконтролер та семисмуговий графічний еквалайзер.

Ключові слова — музична інформація; система візуалізації; спектр; FFT.

I. ВСТУП

Системи візуалізації музики є актуальними на сьогоднішній день, оскільки є нерозривно пов'язаною із медійною складовою й створюють особливий аудіо-візуальний продукт. Завдяки візуалізації музики є можливість зрозуміти ритмічний рисунок музики, що у свою чергу створює більш емоційну атмосферу, й покращує сприйняття музики аудиторією.

Візуалізація музики – функція, котру можна знайти в програмному забезпеченні медіаплеєрів, на різних операційних системах. Ця функція відображає анімовані зображення з прив'язкою до музики. Зображення зазвичай відтворюються в режимі реального часу та синхронізуються з аудіо доріжкою, яка відтворюється. Музика давно поєднується з візуалізацією в різних напрямках мистецтва, таких як: опера, балет, музична драма та фільми. Причина, по якій використовується співзалежність між слуховим та зоровим відчуттями, пов'язана зі складною структурою музики та підвищенням емоційного сприйняття для звичайного слухача [1].

Для візуалізації необхідно виділити певні характеристики музикальних творів використовуючи методи аналізу аудіо інформації. Щоб використати певні параметри, такі як частотний діапазон, та амплітуда, дана інформація має впливати на інтенсивність та колористику або частоту зміни анімацій візуалізації, яка в свою чергу наочно впливає на спостерігача.

Візуальна інформація отримується через прив'язку зображення до певної музики. У цьому випадку зорове відчуття стимулюється слуховим, коли хтось, слухаючи той чи інший аудіозапис, може пов'язати його з певним образом. [1].

У системах, що зараз активно просуваються, використовують декілька інструментів для персональних

комп'ютерів, такі як . After Effects – The Audio Spectrum Effect, VSDC Video Editor Free – Audio Spectrum Visualizer, Magic Music Visuals [2].

Наведене вище програмне забезпечення, має недолік: використання потокового відео неможливе із надходженням одночасно аудіо й потребує обробки та рендерингу отриманого відео ряду.

Мета роботи визначити особливості спектрального аналізу музичної інформації та враховуючи обробку даних в реальному часі. Запропонувати варіант системи візуалізації музичної інформації, у якій відображається спектральний склад музики й амплітуди окремих гармонік, та заповнення LED-матриці відповідним кольором в залежності співвідношення амплітуд частот аудіо сигналу, з можливістю бездротової передачі сигналу від джерела музики до пристрою візуальних ефектів.

II. ОБґРУНТУВАННЯ ВИБОРУ ТЕХНОЛОГІЇ РОЗПІЗНАВАННЯ МУЗИКИ

Для даного проекту обрано технологію частотного аналізу спектру з оцінюванням амплітуди складових спектру музичних даних, що надходять на пристрій [3].

В основі методу лежить аналіз спектру у виділених смугах частот, що в свою чергу спрощує виконання функції пошуку максимумів на різних частотах.

Перетворення Фур'є – математичний алгоритм для розкладання сигналів на синусоїдальні коливання. Наприклад, якщо деякий сигнал $x(t)$ безперервний і нескінченний за часом, то його можна відобразити у вигляді інтеграла Фур'є:

$$x(t) = \int_0^{\infty} X \cos(\omega t + \varphi) d\omega,$$

де $x(t)$ – сигнал, що має нескінченну кількість синусоїдальних складових. X – амплітуда, ω – частота, φ – фаза.

Оскільки, система працює із цифровими сигналами, краще використати дискретне перетворення Фур'є, бо аналізуючи спектр аудіо сигналу ми відбираємо вузькі фрагменти у декілька секунд, протягом усієї музичної композиції

$$x(n) = \sum_{m=0}^{N/2} X_m \cos \frac{2\pi m(n + \varphi_m)}{N},$$

де N – кількість відліків дискретного сигналу $x(n)$ на інтервалі часу від 0 до $N-1$, сума якого дорівнює числу синусоїдальних коливань з амплітудами X_m і фазами φ_m . Частоти синусоїд дорівнюють mF/N , де F – частота дискретизації сигналу, а N – число відліків вихідного сигналу $x(n)$ на згаданому інтервалі. Сукупність коефіцієнтів X_m – амплітудний спектр сигналу. Частоти синусоїд, на які розкладається сигнал, рівномірно розподілені від 0 (постійна складова) до $F/2$, тобто до половини максимально можливої частоти цифрового сигналу.

Для швидшого обрахунку краще використати у аналізі аудіо інформації швидке перетворення Фур'є Fast Fourier Transform (FFT) – алгоритм швидкого обчислення дискретного перетворення Фур'є, який надає можливість аналізувати спектр звукових сигналів в реальному часі [4].

Розглянемо роботу FFT-аналізатора. На вхід йому надходить цифровий аудіо сигнал. Аналізатор вибирає з сигналу послідовні інтервали («вікна»), на яких буде обчислюватися спектр, і FFT виконується в кожному вікні для отримання амплітудного спектра X_m . Далі сигнал з часової області переводиться в частотну область й обчислюється спектр сигналу. Для цього короткий фрагмент аудіо сигналу перемножується на функцію вікна та обчислюється швидке перетворення Фур'є. Використання «прямокутного» вікна призводить до значних спотворень в обчисленні спектру, тому використовуємо вікно Хана (а також інші поширені типи вікон, наприклад, Хемінга або Блекмана).

Наступний фрагмент аудіо сигналу береться зі зсувом у часі, що дорівнює половині або 1/4 ширини вікна, тобто фрагменти перекриваються. Величина зсуву між фрагментами (або «стрибка») безпосередньо впливає на точність алгоритму – чим вона менша, тим вища «роздільна здатність» алгоритму, але більше потрібно пам'яті та часу на функціонування програми [5].

Обчислений спектр, як правило, відображається у вигляді графіка залежності амплітуди від частоти та часу и називається спектрограмою [6]. У запропонованій системі використовується схожа методика, оскільки є оцінка спектру сигналу, амплітуди складових спектру з певним з мінімальним періодом.

Слід зазначити, що подібні методики використовуються для оцінки якості звучання вокалістів в різних студіях звукозапису. Усереднені на великих

часових інтервалах спектри досить успішно застосовуються для аналізу розбірливості мовних сигналів [7].

У запропонованій системі реалізовано той же підхід, що і при створенні спектрограми, тільки колір світлодіодної матриці змінюється не від амплітуди спектральних складових, а від їх частоти.

III. ВИМОГИ ДО СИСТЕМИ ВІЗУАЛІЗАЦІЇ

Система візуалізації аудіоінформації має виконувати такі функції:

- Зміна інтенсивності світіння, як окремих, так і масиву світлодіодів в залежності від гучності;
- Зміна кольору світлодіоду в залежності від співвідношення частот частоти;
- Залежність кількості світлодіодів одного кольору від співвідношення амплітуд частот аудіо сигналу;
- Отримання музичної інформації з різних пристроїв через технологію Bluetooth;
- Забезпечувати достатню швидкість обробку потоку музичних даних.

Забезпечувати зручність керування пристроєм

Таким чином до складу системи входять такі компоненти: LED-матриця, мікроконтролерний блок та модуль прийому музичного сигналу за технологією Bluetooth. Далі обґрунтовано вибір компонентів системи, спираючись на заданий функціонал.

IV. БЛОК СХЕМА І РОБОТА СИСТЕМИ ВІЗУАЛІЗАЦІЇ МУЗИЧНОЇ ІНФОРМАЦІЇ

Блок-схему системи візуалізації музичної інформації наведено на рис.1.

Система працює таким чином. Інформація передається із мобільного пристрою на Bluetooth модуль, де отриманий сигнал декодується й надходить на семи-смуговий еквалайзер. Останній у свою чергу розкладає потік інформації на сім смуг. Програма вибирає один із семи фільтрів, сигнал з виходу якого подається на аналоговий вхід мікроконтролера. Мікроконтролер обраховує дані, що надійшли, виділяє та запам'ятовує гармоніки з найбільшою амплітудою. Потім обробляє сигнал з наступної смуги.

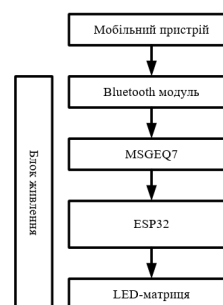


Рисунок. 1 – Блок-схема системи візуалізації музичної інформації.

Інтенсивність випромінювання LED-матриці залежить від амплітуди сигналу у спектрі, на колір випромінювання впливає частотний діапазон сигналу. Кількість світлодіодів які відображають сигнал окремої певної полоси частот пропорційне амплітуді частоті окремих полоси.

Кольори, що відображаються на матриці, пропорційно співвідношенні амплітуді з частотами у самому сигналі. В разі різкої зміни у звучанні фонограми LED-матриця повинна миттєво змінити, як яскравість, так й кольорову гаму випромінювання.

V. СЕМИСМУГОВИЙ ГРАФІЧНИЙ ЕКВАЛАЙЗЕР MSGEQ7

Семидіапазонний графічний еквалайзер - це чіп CMOS що ділить звуковий спектр на сім діапазонів, а саме на: 63 Гц, 160 Гц, 400 Гц, 1 кГц, 2,5 кГц, 6,25 кГц та 16 кГц.

Сім частот виявляються піковими і мультиплексується на виході, щоб забезпечити постійне представлення амплітуди кожної смуги. Для вибору відповідного фільтра не потрібні зовнішні компоненти. Для вибору частоти тактового генератора на мікросхемі потрібні лише резистор і конденсатор.

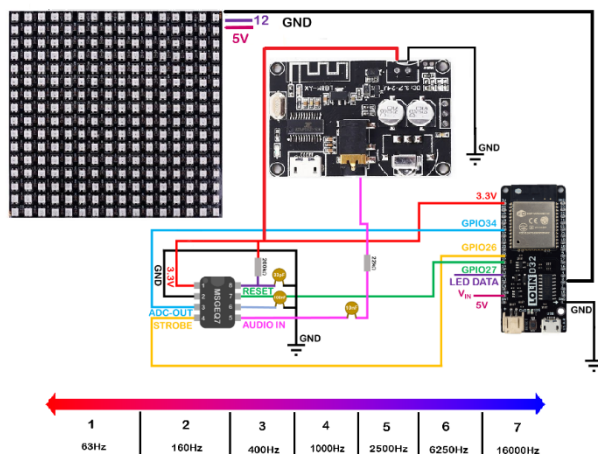
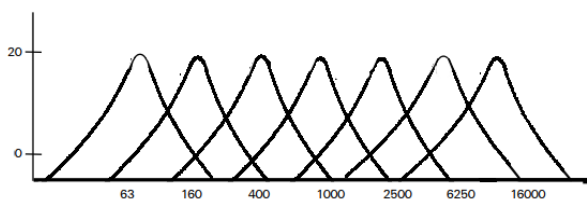


Рисунок 2 – Схема пристрою



MSGEQ7

Окрім роздільних конденсаторів та конденсаторів по колам живлення, ніякі інші зовнішні компоненти не потрібні. Живлення мікросхеми може становити від 2,7 до 5,5 вольт при 5 вольтах, забезпечуючи найкращі характеристики. Пристрій має дуже низький струм спокою (менше 1 мА, як правило), що важливо для портативних аудіопристроїв. Мультиплексор управляється за допомогою сигналів «Скидання» та «Строб», дозволяючи вибирати потрібний канал лише двома виводами. Скидання мультиплексора виконується сигналом високого рівня. Сигнал низького рівня розблоковує вивід «Строб».

Після переднього фронту першого строб-імпульсу до виходу підключається канал 63 Гц. Кожен додатковий передній фронт сигналу «Строб» просуває мультиплексор на один канал (63 Гц, 160 Гц, 400 Гц, 1 кГц, 2,5 кГц, 6,25 кГц, 16 кГц і т. д.), і це буде повторюватися нескінченно [8]

Технічні характеристики семисмугового графічного еквалайзера MSGEQ7:

- Вхідна напруга: 6 В
- Підсилення: 20 дБ

VI. КОНТРОЛЕР BLUETOOTH XY-WRBT

XY-WRBT модуль є декодером й відтворює музику після вибору музики на пристрої. Світлодіодний індикатор на модулі буде сповіщати мигтінням про встановлений Bluetooth зв'язок між пристроями. В разі подальших модифікацій є можливість під'єднати даний модуль через підсилювач до динаміків.

Модуль можна використовувати у портативних системах, що мають живлення від батареї, тоді Micro USB використовується для зарядки акумулятора [9].

Технічні характеристики контролеру Bluetooth XY-WRBT:

- Робоча напруга: 3,7 ~ 24 В
- Версія Bluetooth: 5.0
- Співвідношення сигнал / шум: ≥ 90 дБ
- Відстань передачі: 15 м
- Декодування аудіо: WAV, APE, FLAC, MP3



Рисунок 5 – Зовнішній вигляд XY-WRBT



Рисунок 6 – Зовнішній вигляд ESP 32 DEVKITV1

VII. ОПИС МІКРОКОНТРОЛЕРНОГО МОДУЛЮ

Технічні характеристики плати ESP 32 DEVKITV1

- Модуль: ESP32-WROOM з чіпом ESP32-D0WDQ6
- Частота бездротової передачі: 2,4 ГГц
- Тактова частота: до 240 МГц
- Flash-пам'ять: 448 КБ
- Зовнішня Flash-пам'ять: 4 МБ
- SRAM-пам'ять: 520 КБ
- Апаратні інтерфейси: 3 × SPI, 3 × UART, 2 × I²C і 2 × I²S
- Вхідна напруга через пін Vin: 5-14 В
- Розмір: 51 × 28 мм

Функції АЦП (аналого-цифровий перетворювач) і ЦАП (цифро-аналоговий перетворювач) призначені строго певним виводам, які контакти будуть відведені під інтерфейси UART, I²C, SPI, PWM і т.і. - визначаються програмістом. Це можливо завдяки функції мультиплексування чипу ESP32 [10].

ВИСНОВКИ

Запропонований варіант системи візуалізації музичної інформації забезпечує відображення на LED-матриці кольорів, які співвідносяться з частотами складових у музичній композиції. При чому кількість задіяних світлодіодів пропорційна співвідношенню амплітуд частотних складових сигналу. Потрібний результат досягається використанням швидкого перетворення Фур'є та вибір вікон Хана чи Хемінга для забезпечення кращих результатів аналізу спектру сигналу.

Додатково оцінюються амплітуди окремих складових спектру. при чому кожна полоса частот має свій колір. Робота системи полягає у аналізі складових

Надійшла до редакції 07 квітня 2021 р.

вих спектру та частоти музичної інформації. Дана інформація впливає на відображення кольорів на LED-матриці.

Використання 32-х бітного мікроконтролера забезпечує достатню швидкість обробки аудіо сигналу з мінімальними затримками.

Для підвищення точності і прискорення частотного аналізу звуковий діапазон розділено на сім смуг. Для чого використовується семисмуговий графічний еквайзер MSGEQ7.

Музична інформація передається в систему через канал Bluetooth., що значно спрощує вибір та підключення джерела музичних даних.

ПЕРЕЛІК ПОСИЛАНЬ

- [1] M. Vieira, "Interactive Music Visualization - Implementation, Realization, and Evaluation," Master Dissertation, Universidade da Madeira, Funchal, Madeira, Portugal, 2012, Sep. 2012, 115 p [Online]. Available: <https://core.ac.uk/download/pdf/62478193.pdf>
- [2] 11 Best Music Visualization Tools (Free)-Accessed on: Mar. 29, 2021. [Online]. Available: <https://inspirationfeed.com/music-visualizers/>
- [3] Florian Ramian – Implementation of Real-Time Spectrum Analysis at the Wayback Machine, p. 6, March, 2015, [Online]. Available: https://scdn.rohde-schwarz.com/ur/pws/dl_downloads/dl_application/application_notes/1ef77/1EF77_3e_Real-time_Spectrum_Analysis.pdf
- [4] Frigo, Matteo; Johnson, Steven G. "The Design and Implementation of FFTW3" (PDF). Proceedings of the IEEE. 93 (2): 216–231, 2005. [Online]. Available: <http://fftw.org/fftw-paper-ieee.pdf>. DOI: 10.1109/JPROC.2004.840301
- [5] V. S. Yatskov's'ky, D. M. Bruslik, "Teoriya ta metody obroblyennya syhnaliv [Theory and methods of signal processing]" Electronics and control systems. vol. 31, no. 1, 2012. [Online]. Available: <http://jml.nau.edu.ua/index.php/ESU/article/viewFile/924/906>
- [6] J. Smith, "Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications", Second Edition, 2007, ISBN 978-0-9745607-4-8.
- [7] A. B. Anan'yev, D. V. Prosvirov, "Spektral'ne porivnyannya muzychnykh tvoriv" Akustychnyy visnyk. 2004. vol. 7, no. 3. pp. 7-13. [Online]. Available: [http://hydromech.org.ua/content/pdf/av/av-07-3\(07-13\).pdf](http://hydromech.org.ua/content/pdf/av/av-07-3(07-13).pdf)
- [8] MSGEQ7 Datasheet (PDF) - Mixed Signal Integration Accessed on: Mar. 26, 2021. [Online]. Available: <https://pdf1.alldatasheet.com/datasheet-pdf/view/1132447/MSI/MSGEQ7.html>
- [9] Bluetooth Audio Receiver Board Bluetooth 5.0 mp3 Lossless decoder Board Wireless Stereo Music Module XY-WRBT Accessed on: Mar. 30, 2021. [Online]. Available: <https://www.amazon.com/MADONG-Bluetooth-Receiver-Lossless-Wireless/dp/B086SW1NQ1>
- [10] ESP32. Technical Reference Manual. Accessed on: Mar. 25, 2021. [Online]. Available: https://www.espressif.com/sites/default/files/documentation/esp32_technical_reference_manual_en.pdf

Musical Information Visualization System

R. O. Yaroshenko, ORCID [0000-0002-5360-5885](https://orcid.org/0000-0002-5360-5885)

Department of Acoustic and Multimedia Electronic Systems ames.kpi.ua

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", ROR [00syn5v21](https://ror.org/00syn5v21)
Kyiv, Ukraine

Abstract — The visualisation systems are spread widely as personal computer's software. The system, that are processing audio data are presented in this article. The system visualizes the ratio of spectrum amplitudes and has fixed frequency binding to colours. The technology of audio signals processing by the device and components of the device were considered. For the increasing information processing speed was used 32bit controller and graphic equalizer with seven passbands.

Music visualization it is function, that are spread widely in media player's software, on a different operation systems. This function shows animated images that are depends on music signal. Images are usually reproduced in the real time mode and synchronized with a played audio-track. Music and visualization are merges in the different kind of art: opera, ballet, music drama or movies. Dependencies of auditory and visual sensations are used for increasing the emotional perception for ordinary listeners. In the systems, that are currently being actively promoted, are used several tools for personal computers, such as: After Effects – The Audio Spectrum Effect, VSDC Video Editor Free – Audio Spectrum Visualizer, Magic Music Visuals. The software, that are mentioned above, has a one disadvantage: the using of streaming video is not possible with the simultaneous receipt of audio and requires processing and rendering of the resulting video series. The purpose of the work is to determine the features of spectral analysis of music information and taking into account real-time data processing. Propose a variant of the music information visualization system, which displays the spectral composition of music and the amplitude of individual harmonics, and filling the LED-matrix with the appropriate color depending on the amplitude of the audio signal, with the possibility of wireless signal transmission from the music source to the visual effects device. The technology of frequency analysis of the spectrum with estimation of amplitude of spectrum's components of the musical data, that is arriving on the device is chosen for this project. The method is based on the analysis of the spectrum in the selected frequency bands, which in turn simplifies the function of finding maxima at different frequencies. The proposed variant of the musical information visualization system provides display on the LED-matrix of colors that correspond to the frequencies spectrum's components in the musical composition. Moreover, the number of involved LEDs is proportional to the ratio of the amplitudes of the signal's frequency components. The desired result is achieved by using a Fast Fourier Transform and selecting Khan or Heming windows for providing a better analysis results of the signal spectrum.

The amplitudes of the individual components of the spectrum are estimated additionally and each frequency band has its own color. The work of the system is to analyze the components of the spectrum and frequency of musical information. This information affects the display of colors on the LED matrix.

The using of a 32-bit microcontroller provides sufficient speed of audio signal processing with minimal delays.

For the increasing the accuracy and speed up the frequency analysis, the sound range is divided into seven bands. For this purpose was used seven-band graphic equalizer MSGEQ7.

Music information is transmitted to the system via Bluetooth, which greatly simplifies the selection and connection of the music data source.

Keywords — musical information; visualization system; spectrum; FFT.

