

МЕТОД КОНСОЛІДАЦІЇ МУЛЬСЕМЕДІЙНИХ ДАНИХ

Анотація: У статті пропонується метод, який дозволяє виконувати поєднання та синхронізацію мультимедійних даних із застосуванням принципів багатопотоковості. Застосування запропонованого методу вирішує проблеми, пов'язані з поєднанням у одному часі совому проміжку даних різних модальностей, що відрізняються форматами збереження. Цей метод призначений для застосування при розробленні нового класу програмного забезпечення – мультимедійних програмних систем.

Ключові слова: Програмне забезпечення, мультимедіа, темпоральні мультимодальні дані, метод синхронізації даних, багатопотоковість.

Постановка проблеми. Термін «мультимедіа» (англ. «multimedia» – MULTiple SEnsorial MEDIA) уперше був запропонований Георгієм Гінесом у роботі [1]. Концепція мультимедіа передбачає реєстрацію, зберігання, обробку, передачу та відтворення комп'ютерними засобами мультимодальної інформації про фізичні об'єкти, яка може сприйматися людиною через органи чуття. До такої інформації відносять: аудіовізуальну інформацію (зовнішній вигляд об'єкта спостереження, акустичні властивості або прояви функціонування об'єкта), тактильну інформацію (текстура поверхні, температура об'єкта), кінестетичну інформацію (вага, центр тяжіння об'єкта), інформацію про смак, запах тощо. Особливістю процесу сприйняття мультимедійної інформації людиною є те, що цей процес відбувається у плинні часу. Через це реєстрація мультимедійних даних має здійснюватися з фіксацією моментів часу, коли відповідна мультимедійна інформація існувала або її сприйняття мало сенс для людини-дослідника об'єкта спостереження, тобто мультимедійні дані є темпоральними.

Концепція мультимедіа є подальшим розвитком концепції мультимедіа, яка ґрунтується на записі, збереженні, обробці, передачі та відтворенні лише аудіовізуальної інформації. Принципи, підходи, формати, які дозволяють оброблювати мультимедійні дані, є уніфікованими та стандартизованими. Водночас, комп'ютерна обробка мультимедійних даних потребує формалізації та

подальшої стандартизації, що у свою чергу вимагає розроблення методів обробки мультимедійних даних та їхньої реалізації у вигляді мультимедійного програмного забезпечення.

Аналіз останніх досліджень і публікацій. Більшість з відомих методів синхронізації мультимодальних даних розроблені для вузькопрофільних застосувань. Прикладом такого спеціалізованого застосування є дослідження щодо оцінки частоти серцевих скорочень і електрошкіряної активності як метод об'єктивної оцінки користувацького досвіду для середовищ віртуальної реальності з ефектом занурення [2].

Як розвиток технологій застосування найпоширеніших медіа даних, таких як аудіо та відео, створюються та активно розробляються специфічні формати і структури даних, за допомогою яких можна зберігати та обробляти у комп'ютерний спосіб інформацію про смак, запах, тактильні відчуття, тощо. Проте наявних методів недостатньо для аналізу та поєднання складних мультимодальних компонентів [3]. Крім того, аналіз існуючих методів синхронізації мультимодальних даних не дозволяє виділити один універсальний метод для консолідації даних з широкого спектру модальностей, які задіюють три та більше людських органів чуття [1].

Проте, одночасне використання кількох органів відчуття в інтерактивних додатках стає все більш можливим завдяки поширенню комерційних пристроїв для створення імерсивних систем [4]. Прикладом таких пристроїв є Aroma Player та Aroma shooter [5]. Розвиток методів та систем поєднання мультисенсорних даних змінить підхід до роботи в деяких галузях, а також сприятиме створенню нового користувацького досвіду у споживанні медіаконтенту шляхом розширення спектру класичного аудіовізуального мультимедійного вмісту іншими модальностями, що дозволять задіяти інші органи чуття людини [6].

Аналіз відомих методів обробки та синхронізації мультимедійних даних дозволяє сформулювати вимоги, яким має відповідати універсальний метод консолідації таких даних:

- можливість обробки даних різних модальностей;
- формування єдиної часової шкали для мультимодальних даних, що консолідуються;
- використання особливостей метаданих файлів мультимодальних даних для кращої синхронізації.

Програмна реалізація такого універсального методу консолідації мультимедійних даних має передбачати:

- багатопоточність (обробка даних кожної модальності в окремому потоці);
- можливість використання мультимодальних даних, що зберігаються у розподілених або хмарних сховищах.

Оскільки жоден відомий метод у повній мірі не відповідає цим вимогам, задача створення універсального методу консолідації мультимедійних даних є актуальною.

Метою дослідження, результати якого представлені у цій статті, є розроблення основних засад комп'ютерної обробки мультимедійної інформації, а саме, розроблення методу, який дає змогу консолідувати мультимедійні дані про об'єкт спостереження з урахуванням мультимодальності та темпоральності цих даних, що вимагає їхньої синхронізації у межах єдиного часового проміжку.

Основні етапи процесу консолідації мультимедійних даних. Консолідація мультимедійних даних про об'єкт спостереження є процесом, що складається з таких основних етапів.

1. Аналіз метаданих файлів даних різних модальностей.
2. Формування часової шкали процесу дослідження об'єкта спостереження.
3. Визначення структури даних та створення файлу консолідованих мультимедійних даних.

Розглянемо кожен з цих етапів детально.

Аналіз метаданих. Окрім власно даних про об'єкт спостереження, які потрібно синхронізувати, вхідні дані, що організовані у вигляді файлу, включають метадані, а саме: тип файлу, його розмір, дата створення, дата останнього редагування, місцезнаходження у момент створення, технічна інформація про пристрій, яким файл було створено. Для відео та аудіо файлів до складу метаданих також входять: тривалість, бітрейт, кодек, кількість каналів, розмір кадру, тощо. Додатковими метаданими є назва файлу, його опис, інформація про авторів тощо.

Для коректної консолідації мультимедійних даних необхідно виявити метадані, які називатимемо темпоральними метаданими: дата створення файлу, дата його останнього редагування та тривалість запису. Аналіз темпоральних метаданих дозволяє отримати інформацію, яка потрібна для подальшої обробки даних, а саме, їх сортування, поєднання та синхронізації з метою комплексного подання мультимедійних даних [7], що всебічно описують об'єкт

спостереження. Наприклад, якщо файли з даними різних модальностей були створені у приблизно однаковий момент часу, то аналіз темпоральних метаданих дозволяє виявити, які дані почали надходити першими та визначити тривалість їхнього надходження, що дає змогу виконати синхронізацію цих даних з даними інших модальностей в одному часовому проміжку. У файлах різного типу темпоральні метадані можуть зберігатися у різних форматах, а також бути відсутніми. Крім того, мітки часу можуть зберігатись не у заголовку файлу, а в основній частині файлу, тобто не у метаданих, а у даних.

Для прискорення процесу консолідації мультимедійних даних доцільно забезпечити багатопотоковість алгоритму виявлення темпоральних даних для одночасної обробки файлів різних модальностей. Алгоритм багатопотокової обробки файлів для виявлення темпоральних метаданих наведений на рис. 1.

Формування часової шкали. Результатом виконання першого етапу консолідації мультимедійних даних є виявлення темпоральних метаданих – часових значень, що визначають тривалість та час початку і закінчення відтворення даних певної модальності. Проте ці часові значення мають бути визначені у будь-яких одиницях виміру часу: мілісекундах, годинах, місяцях тощо, що залежить від частоти отримання відповідних даних з сенсорів, камер та інших пристроїв. Для консолідації мультимедійних даних необхідно привести всі значення в одну систему виміру часу та визначити єдиний спільний часовий проміжок для узгодженого представлення даних різних модальностей – часової шкали процесу дослідження об'єкта спостереження. Ця шкала (рис. 2) формується на основі отриманого набору темпоральних метаданих, який є масивом часових міток, що позначають ключові моменти дослідження об'єкта спостереження та сприйняття людиною-дослідником мультимедійної інформації про цей об'єкт.

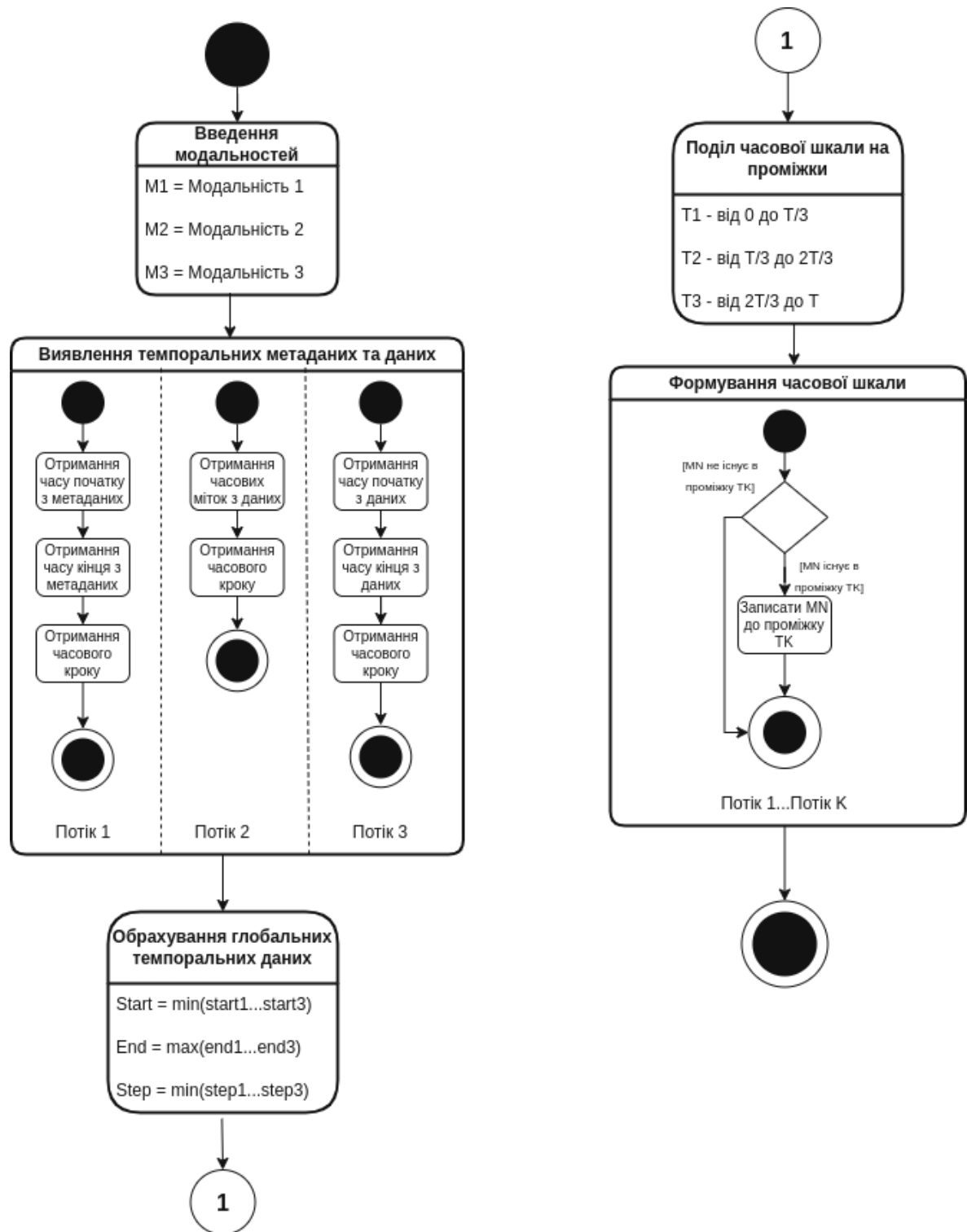


Рисунок 1 - Алгоритм багатопотокової обробки мультимедійних файлів для виявлення темпоральних метаданих

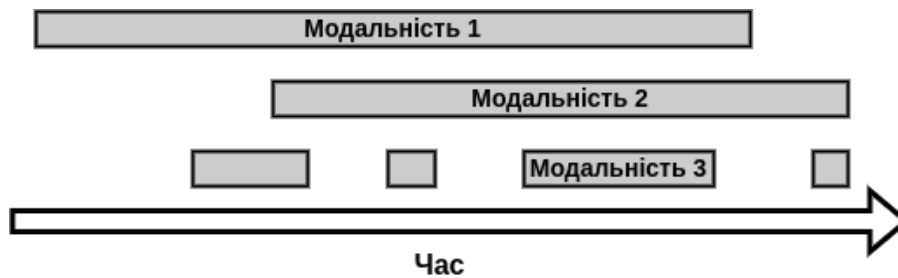


Рисунок 2 - Приклад часової шкали процесу дослідження об'єкта спостереження

Визначення структури даних та створення файлу консолідованих мультимедійних даних. Синхронізовані мультимедійні дані мають бути збережені у вигляді файлу. Можливим форматом такого файлу може бути JSON (JavaScript Object Notation) [8], оскільки він надає змогу структуровано зберігати мультимодальні дані. Проте мультимедійні дані є темпоральними, тому для забезпечення подальшого коректного оброблення файлу таких даних формат файлу має бути модифікований з урахуванням темпоральності даних. Назвемо такий модифікований формат Timeline JSON (TJSON). Приклад TJSON-об'єкту наведено у лістингу 1.

Лістинг 1. Приклад TJSON-об'єкту.

```
{
  "name": "назва мультимедійного файлу",
  "description": "опис мультимедійного файлу",
  "stepMeasurement": "формат вимірювання кроку (секунди, мілісекунди, тощо)",
  "step": розмір кроку (0.1, 0.01, тощо),
  "duration": загальна тривалість мультимедійного файлу в обраному форматі (секунди, мілісекунди, тощо)",
  "media_formats": [ масив використаних форматів (mp3, mp4, jpg, txt, тощо) ],
  "data": {
    "0.0": [ масив даних, що існують в проміжок часу 0.0 ],
    "0.1": [ масив даних, що існують в проміжок часу 0.1 ],
    ...
    "N": [ масив даних, що існують в проміжок часу N ],
  }
}
```

Після визначення TJSON-об'єкту, тобто структури даних для збереження мультимедійних даних, відбувається створення мультимедійного файлу. Тип файлу, що створюється, залежить від наявних модальностей і визначається виходячи з таких міркувань:

- якщо серед вхідних даних присутні відеодані, то вихідний файл має мати формат відео;
- якщо серед вхідних даних відсутні відеодані, проте присутні аудіодані та числові або текстові дані, то вихідний файл має мати формат відео;
- якщо серед вхідних даних відсутні відео- та аудіодані, проте присутні числові або текстові дані, то вихідний файл має мати формат .CSV, .PDF або спеціалізований формат, що визначається типом пристрою відтворення цих даних.

Вихідний файл консолідованих мультимедійних даних містить дані, впорядковані згідно з визначенням TJSON-об'єкту. Цей файл може бути відтворений за допомогою відповідного апаратного забезпечення [9, 10]. Також мультимедійні дані можуть бути попередньо оброблені відповідно до задачі дослідження. При цьому доцільно застосувати операції впорядкування або логічні операції на основі алгебраїчної системи агрегатів (ACA) [7], оскільки TJSON-об'єкт є програмною реалізацією агрегата – математичної абстракції, визначеної в ACA. Таким чином, консолідовані мультимедійні дані можна представити у вигляді набору агрегатів, над якими виконуються логічні операції та операції впорядкування. До логічних операцій над агрегатами відносять об'єднання, переріз, різниця, симетрична різниця та виключний переріз. До операцій впорядкування агрегатів відносять розміщення, сортування, проріджування, видалення та вставлення.

Якщо операції впорядкування можуть виконуватися в межах одного набору даних, то для логічних операцій наборів даних має бути два, тому логічні операції можуть виконуватися або між двома наборами із різних, можливо, сусідніх часових проміжків, або при наявності іншого такого TJSON-об'єкту. У тому випадку, якщо необхідно виконати операції по всьому TJSON-об'єкту або між двома TJSON-об'єктами, можна застосувати багатопотоковий підхід для збільшення ефективності обчислень. При цьому для забезпечення багатопотоковості часова шкала поділяється на фрагменти. Мультимедійні дані, що визначені у певному фрагменті шкали, оброблюються в окремому потоці.

Ефективність паралельних обчислень у запропонованому методі. Для дослідження доцільності багатопотоковості було виконано реалізацію запропонованого методу та протестовано розроблений програмний код на експериментальних даних. Метою експерименту було визначення швидкості виконання обробки мультимедійних даних при різних кількостях потоків.

Вхідними даними у цьому експерименті були файли наступних типів: відео, аудіо та CSV-файл. Тривалість відео складала 61 секунду. Тривалість аудіо – 84 секунди. CSV-файл містив таблицю двох значень «ключ»-«значення», де «ключем» є часова мітка, а «значенням» – числове значення, що надходить із сенсора, подане у текстовому форматі.

Експеримент був виконаний декілька разів зі застосуванням різної кількості потоків. Результати визначення швидкості виконання обробки мультимедійних даних при різних кількостях потоків наведені у табл. 1.

Таблиця 1

Результати експерименту з паралельних обчислень

	Один потік	Два потоки	Три потоки	Чотири потоки
Час, с	211,06	202,8377	191,3445	176,8694

Аналіз цих результатів дозволяє підтвердити припущення, що зі збільшенням кількості потоків швидкість оброблення мультимедійних даних зростає.

Висновки. Мультимедійні дані являють собою складну структуру даних, що може містити дані довільних модальностей. Наразі існує проблема коректного злиття даних різних модальностей для їхньої консолідованої обробки та зберігання. Запропонований у статті метод дає змогу вирішити проблему поєднання даних різних модальностей, а також їхньої синхронізації у межах єдиної часової шкали, що робить можливим створення мультимедійних файлів для подальшої програмної обробки. Для збільшення ефективності запропонованого методу доцільно використати багатопоточність на етапі аналізу метаданих та етапі створення вихідного файлу.

Запропонований метод може бути розширений іншими підходами та технологіями. Наприклад, на етапі опрацювання даних можна застосовувати методи штучного інтелекту для оцінки кореляції даних різних модальностей.

ЛІТЕРАТУРА

1. A User Perspective of Olfaction-Enhanced Mulsemmedia. Proceedings of the international conference on management of emergent digital ecosystems / Ghinea G., Ademoye O. // MEDES'10. – 2010. – P. 277–280.
2. Semi-automatic synchronization of sensory effects in mulsemmedia authoring tools / Abreu, R., Mattos, D., Santos, J. A. D., & Muchaluat-Saade, D. C. // In Proceedings of the 25th Brazilian Symposium on Multimedia and the Web. – 2019. – P. 201-208.
3. Beyond multimedia adaptation: Quality of experience-aware multi-sensorial media delivery / Yuan, Z., Ghinea, G., & Muntean, G. M. // IEEE Transactions on Multimedia. – 2014. – №17(1). – P. 104-117.
4. An evaluation of Heart Rate and ElectroDermal Activity as an objective QoE evaluation method for immersive virtual reality environments / Egan, D., Brennan, S., Barrett, J., Qiao, Y., Timmerer, C., & Murray, N. // In 2016 eighth international conference on quality of multimedia experience (QoMEX) IEEE. – 2016. – P. 1-6.
5. Towards a reference ontology on mulsemmedia system / Saleme, E. B., Santos, C. A., Falbo, R. A., Ghinea, G., & Andres, F. // In Proceedings of the 10th International Conference on Management of Digital EcoSystems. – 2018. – P. 23-30.
6. Development of directional olfactory display / Kim, D. W., & Ando, H. // In Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry. – 2010 – P. 143-144.
7. Temporal Multimodal Data-Processing Algorithms Based on Algebraic System of Aggregates / Pester, A., Sulema, Y., Dychka, I., & Sulema, O. // Algorithms. – 2023. – №16(4). – P 186.
8. Foundations of JSON schema / PEZOA, Felipe, et al. // In: Proceedings of the 25th international conference on World Wide Web. – 2016. – P. 263-273.
9. Multiple Sensorial Media Advances and Applications / GHINEA, George; ANDRES, Frederic; GULLIVER, Stephen. // New Developments in. – 2012.
10. Mulsemmedia vs. Multimedia: State of the art and future trends / Y. Sulema // International Conference on Systems, Signals and Image Processing (IWSSIP), Bratislava, Slovakia. – 2016. – P. 1-5.

REFERENCES

1. A User Perspective of Olfaction-Enhanced Mulsemmedia. Proceedings of the international conference on management of emergent digital ecosystems / Ghinea G., Ademoye O. // MEDES'10. – 2010. – P. 277–280.

2. Semi-automatic synchronization of sensory effects in mulsemedia authoring tools / Abreu, R., Mattos, D., Santos, J. A. D., & Muchaluat-Saade, D. C. // In Proceedings of the 25th Brazillian Symposium on Multimedia and the Web. – 2019. – P. 201-208.
3. Beyond multimedia adaptation: Quality of experience-aware multi-sensorial media delivery / Yuan, Z., Ghinea, G., & Muntean, G. M. // IEEE Transactions on Multimedia. – 2014. – №17(1). – P. 104-117.
4. An evaluation of Heart Rate and ElectroDermal Activity as an objective QoE evaluation method for immersive virtual reality environments / Egan, D., Brennan, S., Barrett, J., Qiao, Y., Timmerer, C., & Murray, N. // In 2016 eighth international conference on quality of multimedia experience (QoMEX) IEEE. – 2016. – P. 1-6.
5. Towards a reference ontology on mulsemedia system / Saleme, E. B., Santos, C. A., Falbo, R. A., Ghinea, G., & Andres, F. // In Proceedings of the 10th International Conference on Management of Digital EcoSystems. – 2018. – P. 23-30.
6. Development of directional olfactory display / Kim, D. W., & Ando, H. // In Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry. – 2010 – P. 143-144.
7. Temporal Multimodal Data-Processing Algorithms Based on Algebraic System of Aggregates / Pester, A., Sulema, Y., Dychka, I., & Sulema, O. // Algorithms. – 2023. – №16(4). – P 186.
8. Foundations of JSON schema / PEZOA, Felipe, et al. // In: Proceedings of the 25th international conference on World Wide Web. – 2016. – P. 263-273.
9. Multiple Sensorial Media Advances and Applications / GHINEA, George; ANDRES, Frederic; GULLIVER, Stephen. // New Developments in. – 2012.
10. Mulsemedia vs. Multimedia: State of the art and future trends / Y. Sulema // International Conference on Systems, Signals and Image Processing (IWSSIP), Bratislava, Slovakia. – 2016. – P. 1-5.

Received 17.11.2022.

Accepted 22.11.2022.

Mulsemedia data consolidation method

The synchronization of multimodal data is one of the essential tasks related to mulsemedia data processing. The concept of mulsemedia (MULTiple SEnsorial MEDIA) involves the registration, storage, processing, transmission and reproduction by computer-based tools of multimodal information about a physical object that humans can perceive through their senses. Such information includes audiovisual information (object's appearance, acoustic properties, etc.), tactile information (surface texture, temperature), kinesthetic information (weight, object's centre of gravity), information about

its taste, smell, etc. The perception of mulsemmedia information by a person is the process that exists over time. Because of this, the registration of mulsemmedia data should be carried out with the fixation of the moments of time when the relevant mulsemmedia information existed or its perception made sense for a human who supervises the object as mulsemmedia data is temporal.

This paper presents a method that enables the consolidation and synchronization of mulsemmedia data using the principles of multithreading. The universal method was designed to support combining data of different modalities in parallel threads. The application of the proposed method solves problems associated with integrating data of different modalities and formats in the same time interval. The effectiveness of applying this method increases by using multithreaded distributed computing. This method is designed for use in the development of mulsemmedia software systems. The modified JSON format (TJSON – Timeline JSON) was proposed in the paper, as well. TJSON-object is a complex data structure for representing the synchronized mulsemmedia data and their further processing.

The proposed method can be further extended with other approaches and technologies. For example, artificial intelligence methods can be applied to assess the correlation between data from different modalities. This can help improve the method's accuracy and the output files' quality.

Рвач Дмитро Вячеславович – аспірант кафедри програмного забезпечення комп'ютерних систем, rvach.d@gmail.com Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ.

Сулема Євгенія Станіславівна – д-р техн. наук, доцент, завідувач кафедри програмного забезпечення комп'ютерних систем. Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ.

Rvach Dmytro – Post-Graduate Student of Computer Systems Software Department, rvach.d@gmail.com National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv.

Yevgeniya Sulema – DSc, Associate Professor, Head of Computer Systems Software Department. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv.