

МОДЕЛЬ ПОДАННЯ МУЛЬТИМОДАЛЬНИХ ДАНИХ ДЛЯ КОМПЛЕКСНОГО ОПИСУ ОБ'ЄКТІВ СПОСТЕРЕЖЕННЯ

¹Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»

Розглянуто проблему комплексного подання мультимодальних даних про об'єкт спостереження, характеристики якого вимірюються та досліджуються з урахуванням часу та взаємозв'язку між даними різних модальностей. Ця проблема є актуальною для багатьох інженерних задач. Зокрема, комплексне подання даних про фізичний об'єкт є важливим у технології цифрових двійників, яка останнім часом набуває ширшого застосування. Виходячи з того, що цифровий двійник визначається поведінковою моделлю та візуальною моделлю, у статті запропоновано модель даних, що ґрунтується на понятті мукселью, який є мінімальним елементом мультимодальних даних. Ця модель дозволяє комплексне та несуперечливе подання даних про об'єкт спостереження.

Математична модель комплексного подання мультимодальних даних ґрунтується на поняттях агрегату та мультиобразу, визначених у алгебраїчній системі агрегатів. Проаналізовано зв'язок між комплексним поданням та обробленням мультимодальних даних, визначених у часі за допомогою апарату алгебраїчної системи агрегатів та розглядом наборів агрегованих даних як функцій багатьох змінних. Таким чином, оброблення даних муксельної моделі може виконуватись, як із застосуванням логічних операцій, операцій впорядкування та відношень, визначених у алгебраїчній системі агрегатів, так й операцій і підходів, визначених у інших математичних концепціях.

Оскільки муксельна модель представляє дані великого обсягу, доцільно оптимізувати спосіб подання даних. Цього можна досягти ущільненням даних, модифікацією структури даних або комбінацією першого та другого підходу. У першому підході доцільно застосувати ущільнення даних муксельної моделі на основі методу RLE. У другому підході пропонується застосовувати структури даних, аналогічні розрідженому воксельному октодереву, що застосовується для зменшення обсягів даних моделей у воксельній графіці. Практична реалізація оброблення мультимодальних даних, представлених за допомогою муксельної моделі, може бути виконана за допомогою мови програмування ASAMPL.

Ключові слова: модель подання даних, мультимодальні дані, дані великого обсягу.

Вступ

Для розв'язання різноманітних задач у галузі інженерії, в тому числі медичної інженерії, важливим є точне, повне та несуперечливе подання даних про об'єкт спостереження. При цьому для деяких задач принциповим є комплексний підхід, коли дані різних модальностей є взаємопов'язаними та визначеними у часі. Особливо важливим це є у випадках, коли потрібно проаналізувати динаміку зміни різних параметрів об'єкта одночасно. Прикладами можуть слугувати системи моніторингу технологічних процесів у фармацевтичній, автомобільній, нафтопереробній та інших галузях виробництва, зокрема й системи моніторингу у складі «розумного підприємства» [1]. Іншим важливим прикладом можуть слугувати моделі об'єктів за технологією цифрових двійників, яка наразі активно розвивається [2]—[5]. Технологія цифрових двійників розроблена, в першу чергу, для поведінкового та візуального моделювання критичних станів фізичного двійника — об'єкта, для якого створено цифровий двійник [2], [4]. На сьогодні немає відносно універсальних моделей подання даних цифрових двійників, отже, задача розроблення моделей подання мультимодальних даних, які могли би бути застосовані для визначення цифрових двійників об'єктів спостереження, є актуальною задачею, що потребує розв'язання.

Оскільки цифровий двійник включає як поведінкову модель, так і візуальну модель, то однією з модальностей у таких застосунках є інформація про зовнішній вигляд тривимірного об'єкта.

Найгрунтовнішу тривимірну графічну інформацію про об'єкти дозволяють зберігати структури даних на основі вокселя (*voxel*, від *volume element*) — об'ємного елемента одиничного розміру [6]. Воксельна структура даних лежить в основі воксельної графіки. Воксель є розширенням поняття пікселя (*pixel*, від *picture element*) — одиничного елемента (точки) у растровій графіці. Так само, як і піксель, воксель характеризується деяким кольором. Воксельна графіка використовується для візуалізації медичної інформації (наприклад, об'ємна візуалізація результату МРТ-дослідження), візуалізації наукових даних, комп'ютерній анімації тощо. Подальшим розвитком та узагальненням понять «піксель» та «воксель» є доксель (*doxel*, від *dynamic element*) — воксель, визначений у часі [7]. Ідею збереження «об'ємної» графічної інформації, реалізовану у воксельній та доксельній графіці, доцільно поширити на збереження інформації інших модальностей, таких як аудіо-інформація, інформація про фізичні властивості та параметри об'єкта (температура, вологість, густина) тощо.

Таким чином, мультимодальну інформацію про об'єкт спостереження можна зберігати на основі об'ємного елемента найбільш узагальненого типу, який зберігає інформацію всіх можливих модальностей, представлену у часі. За аналогією будемо називати такий елемент «муксель» (*mixel*, від *multimodal element*) [8].

Метою дослідження є розроблення моделі подання мультимодальних даних, яка ґрунтується на понятті мукселя та математичному апараті алгебраїчної системи агрегатів [9], [10] для оброблення мультимодальних даних, визначених у часі.

Математична модель подання мультимодальних даних

Математична модель комплексного подання мультимодальних даних ґрунтується на поняттях агрегату та мультиобразу, визначених у алгебраїчній системі агрегатів [9], [10].

Визначення 1. Агрегат A — це кортеж довільних кортежів, елементи кожного з яких належать до певної наперед визначеної множини

$$A = \left\| M_j \mid \left\langle a_i^j \right\rangle_{i=1}^{n_j} \right\|_{j=1}^N. \quad (1)$$

Розглянемо об'єкт спостереження S , який виявляє свою сутність через набір властивостей F_1, F_2, \dots, F_N , що можуть бути вимірні. В багатьох випадках ці властивості є взаємопов'язаними, оскільки вони визначають стан того самого об'єкта та, фактично, є різними проявами його поведінки у часі. Наприклад, згідно з дослідженнями фізіології людини [11], за підвищення температури тіла пацієнта, його пульс також зростає. Це означає, що комплексне, взаємопов'язане подання даних вимірювання властивостей об'єкта спостереження сприяє кращому розумінню загальної картини поведінки цього об'єкта.

Представимо виміряні значення властивостей об'єкта спостереження S як агрегат A_S , компонентами якого є кортежі значень $\left\langle f_i^j \right\rangle_{i=1}^{n_j}$ властивостей F_j ($j = 1 \dots N$) об'єкта S

$$A_S = \left\| M_1, M_2, \dots, M_N \mid \left\langle f_1^1, f_2^1, \dots, f_{n_1}^1 \right\rangle, \left\langle f_1^2, f_2^2, \dots, f_{n_2}^2 \right\rangle, \dots, \left\langle f_1^N, f_2^N, \dots, f_{n_N}^N \right\rangle \right\|. \quad (2)$$

Якщо вимірювання значень різних властивостей є одночасним, то всі кортежі агрегату A_S мають однакову довжину. В цьому випадку, агрегат A_S може бути розглянутий як послідовність мультикомпонентних значень (рис. 1).

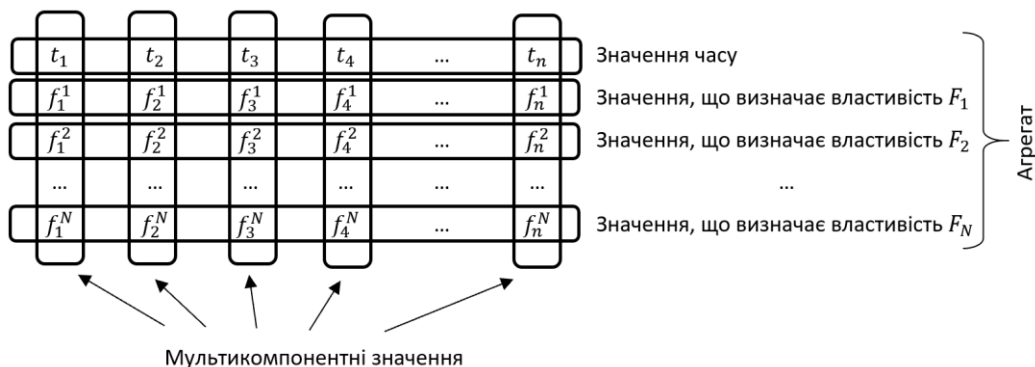


Рис. 1. Агрегація мультимодальних даних, визначених у часі

Ця послідовність мультикомпонентних значень може бути розглянута як функція багатьох змінних $y = x(t, F_1, F_2, \dots, F_N)$. Такий погляд на агреговані дані, дозволяє застосовувати до них апарат числення багатьох змінних у деяких інженерних задачах, в яких є доречним подання та оброблення даних за допомогою операцій та відношень, визначених в АСА, тобто застосування апарату АСА не заперечує використання інших математичних концепцій.

Якщо дані про об'єкт спостереження отримані з урахуванням часу спостереження або вимірювання, то у агрегаті має бути присутнім кортеж відповідних значень часу, які відповідають моментам часу, коли значення кортежів мультимодальних даних були отримані. Будемо називати такий агрегат мультиобразом об'єкта спостереження.

Визначення 2. *Мультиобраз* — це непорожній агрегат виду

$$I = \left\| T, M_1, \dots, M_N \left\langle t_1, \dots, t_\tau \right\rangle, \left\langle a_1^1, \dots, a_{n_1}^1 \right\rangle, \dots, \left\langle a_1^N, \dots, a_{n_N}^N \right\rangle \right\|, \quad (3)$$

де T — це множина значень часу; $\tau \geq n_i, i \in [1, \dots, N]$.

Мультиобраз є математичною моделлю представлення даних про об'єкт спостереження, на якій ґрунтується структура даних муксельної моделі.

Муксельна модель

Модель даних цифрового двійника об'єкта спостереження може бути визначена через поняття мукселю як одиничного елемента впорядкованої сукупності мультимодальних даних.

Муксель може бути описаний як гетерогенна структура даних — фрейм, який містить дані різних модальностей як слоти. Для надання повного опису стану деякої точки простору об'єкта в певний момент часу муксельна структура даних може містити такі дані:

- просторові дані (Spatial Data), задають координати мукселя у просторі сцени;
- часові дані (Time Data), задають момент часу, у який фіксується мультимодальна інформація;
- візуальні дані (Visual Data), задають колір та прозорість відповідної певному мукселю точки у просторі об'єкта або у просторі сцени;
- акустичні дані (Auditory Data), задають значення характеристик акустичного сигналу у відповідних певному мукселю точці простору та моменті часу;
- фізичні дані (Physical Data), задають тип та стан речовини у відповідній певному мукселю точці простору; стан речовини задається її характеристиками, що можуть бути виміряні, як-то, температурою, тиском, вологістю, густиною тощо.

Для певних галузей застосування, можливе розширення цих базових типів даних. Наприклад, для інженерних задач, що вирішуються у хімічній промисловості, може виникнути потреба у формальному описі хімічних речовин, що наразі виконується за допомогою відповідного програмного забезпечення [12], [13]. Такий формальний опис може бути застосований для визначення даних певної додаткової модальності у муксельній моделі.

Залежно від властивостей об'єкта спостереження деякі дані можуть не визначатись, а замінитись спеціальними мітками:

- «By Default», використовується для стандартних даних, що задаються за замовчуванням (у цьому випадку відповідне чисельне значення береться з наперед визначеної бази стандартних даних), або якщо визначення даних певного типу є необов'язковим;
- «Not Applicable», використовується, коли визначення даних певної модальності не має сенсу;
- «Unknown», використовується, коли визначення даних певної модальності є неважливим для певного об'єкта;
- «Repeated», використовується, коли значення повторюється для деякої послідовності мукселів або повторюється для одного мукселя для певного проміжку часу.

Кожен муксель може бути у чотирьох станах: інформаційному, невизначеному, квазівизначеному та у стані повторення. Муксель у інформаційному стані визначається явно у всіх модальностях — конкретними даними або мітками «By Default». Муксель у невизначеному стані визначається мітками «Unknown» у всіх модальностях. Цей стан означає, що дані про муксель або невідомі, або неважливі для певної задачі, тому не визначаються. Муксель у квазівизначеному стані визначається явно лише у деяких модальностях, а у всіх інших задається мітками «Unknown». Якщо дані всіх типів, окрім часових, певного мукселя збігаються з відповідними даними попереднього мукселя, то цей муксель може визначатись, як такий, що знаходиться у стані повторення. В цьому випадку замість даних відповідної модальності задається мітка «Repeated». Стан повторення потрібен для

ущільнення муксельних даних.

Частота фіксації зміни стану мукселів у часі залежить від частоти зміни найбільш змінюваних даних. Набір слотів даних залежить від конкретної задачі, що розв'язується із застосуванням мультимодального представлення даних. Точність опису об'єкта спостереження визначається розміром мукселя (так само, як і у разі воксельного опису), вибір якого залежить від точності визначення мультимодальних даних.

Формат файлу муксельної моделі передбачає наявність у файлі службової частини та інформаційної частини. Службова частина містить метадані та таблиці локальної бази даних моделі:

- 1) таблиці службових даних, зокрема й таблицю часових міток;
- 2) таблиці візуальної моделі, зокрема й таблицю кольорів;
- 3) таблиці поведінкової моделі, зокрема й, таблицю матеріалів, таблицю станів тощо.

Оскільки муксельна модель містить дані усіх модальностей про кожну точку простору об'єкта на деякому часовому проміжку, то обсяг цих даних є значним. Отже, муксельна модель належить до моделей категорії Big Data (дані великого обсягу) [14]. Тому доцільно оптимізувати муксельну модель з точки зору способу подання даних. Це можна досягти ущільненням даних, модифікацією структури даних або комбінацією першого та другого підходу.

Розглянемо перший підхід. Якщо частина мукселів знаходиться у невизначеному стані, квазівизначеному стані або у стані повтору, то слоти сусідніх мукселів набувають однакового значення, в результаті чого утворюються довгі ланцюжки однакових даних. В цьому випадку до даних муксельної моделі доцільно застосувати ущільнення даних на основі методу RLE [15]. Так, якщо послідовність мукселів знаходиться у стані повтору, то отримуємо набір однотипних фреймів:

```
<Ідентифікатор фрейму> { <значення часу>;
  repeated;
  repeated;
  repeated;
  repeated;
  repeated; }
```

які можуть бути замінені на пару <кількість повторень, екземпляр фрейму>, а сам екземпляр фрейму може бути записаний у такому вигляді:

```
<Ідентифікатор фрейму > { by default;
  5 repeated; }
```

Розглянемо другий підхід. Він може полягати в тому, що дані зберігатимуться у оптимізованій структурі даних, аналогічній розрідженому воксельному октодереву (Sparse Voxel Octree) [16], що застосовується для оптимізації воксельної моделі. Розріджене воксельне октодерево являє собою структуру даних, що зберігає воксельну модель у вигляді октодерева (простір сцени поділяється на 8 підпросторів, а кожний підпростір в свою чергу також поділяється на 8 підпросторів і так далі до досягнення потрібного рівня деталізації), більшість елементів, тобто вокселів, якого є порожніми.

Муксельна модель може бути порівняна з узагальненою воксельною моделлю [17], яку створено для розширення можливостей 3D візуалізації томографічних зображень органів людського тіла. Узагальнена воксельна модель дозволяє зберігати, крім графічних даних вокселів, також метадані, а саме, ознаку належності вокселя певному органу. На відміну від узагальненої воксельної моделі, муксельна модель передбачає збереження не лише візуальних даних, а також часових даних, просторових даних та будь-яких мультимодальних даних, що розширює галузь її застосування.

Зазначимо, які можливості надає муксельна модель. Оскільки муксельна модель містить повну інформацію про об'єкт, то ця модель робить можливим «проекціонування» інформації кожної модальності відносно потрібної позиції спостереження та отримання «зрізів» цієї інформації. Наприклад, вона дозволить отримати статичне растрове зображення під певним кутом погляду, визначити рівень акустичного сигналу у певній точці сцени з урахуванням наявних перешкод на шляху розповсюдження акустичних хвиль, визначити наявність та інтенсивність певної властивості речовини, з якої складається об'єкт або його компоненти тощо. Ця інформація може бути використана для подальшого аналізу та пошуку у мультимодальній базі даних подібних об'єктів у задачах діагностики, моніторингу, прогнозування тощо.

Для перевірки запропонованої моделі використовувався відкритий набір мультимодальних даних бази даних CREATE [18], яка містить мультимодальні дані, отримані в результаті 14 годин запису даних з датчиків мобільної платформи iRobot Create, а саме: координати платформи, швид-

кість руху, стерео RGB зображення, стерео аудіо, атмосферний тиск, температуру, напруженість магнітного поля, тощо. Специфікація даних [19] визначає ієрархію збереження набору даних у форматі HDF (Hierarchical Data Format), що дозволило використовувати їх у муксельній моделі. Зокрема, у експериментальній частині дослідження використано дані, показані на рис. 2.

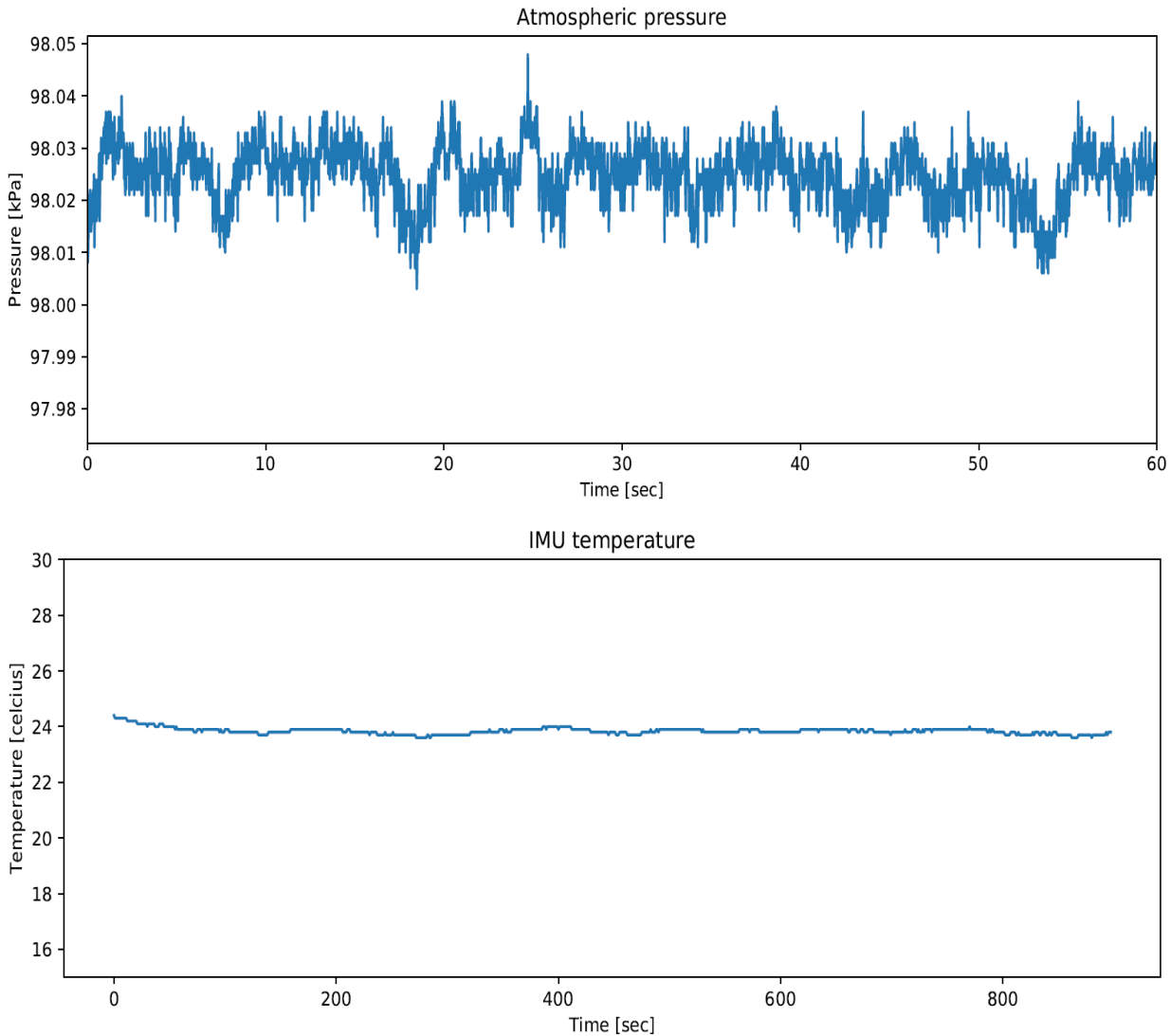


Рис. 2. Мультимодальні дані бази даних CREATE [18], [19]

У створеній на основі даних бази мультимодальних даних CREATE моделі використовувались такі дані: часові дані, просторові дані, дані про атмосферний тиск, дані про температуру повітря. Проведений експеримент показав можливість конвертації мультимодальних даних з формату HDF у формат муксельної моделі.

У подальшому запропонована модель даних має бути реалізована у вигляді сховища даних категорії NoSQL. Це зумовлено тим, що, по-перше, оброблення та збереження муксельних даних має відбуватись у розподілений спосіб, що пояснюється їх великими обсягами; по-друге, NoSQL бази даних забезпечують гнучкий механізм оперування та обміну даними. Зокрема, NoSQL бази даних підтримують агрегатний підхід до оброблення даних, де під агрегатом розуміється одиниця маніпулювання даними та управління їх узгодженістю [20]. По відношенню до муксельної моделі під агрегатом слід розуміти сукупність даних різних модальностей, що відносяться до того самого мукселя або до блока однакових мукселей. Таким чином, до даних муксельної моделі можуть бути застосовані операції алгебраїчної системи агрегатів, зокрема логічні операції та операції впорядкування, а також відношення, зокрема, відношення між дискретними інтервалами [9], [10]. Практична реалізація оброблення мультимодальних даних, представлених за допомогою муксельної моделі, може бути виконана за допомогою мови програмування ASAMPL [21].

Висновки

Комплексне подання мультимодальних даних, визначених у часі, є складовою процесу оброблення даних про об'єкт дослідження у багатьох інженерних задачах. Застосування запропонованої моделі дозволить забезпечити повний та несуперечливий опис об'єкта спостереження. Оскільки муксельна модель ґрунтується на математичній моделі мультиобразу, то до неї можуть бути застосовані логічні операції, операції впорядкування та відношення, визначені у алгебраїчній системі агрегатів.

Муксельна модель представляє дані великого обсягу, тому у статті пропонуються два підходи до оптимізації способу подання даних. У першому підході передбачається застосування ущільнення даних муксельної моделі на основі методу RLE. У другому підході пропонується застосовувати структури даних, аналогічні розрідженому воксельному октодереву, що застосовується для зменшення обсягів даних моделей у воксельній графіці. Практична реалізація оброблення мультимодальних даних, представлених за допомогою муксельної моделі, може бути виконана за допомогою мови програмування ASAMPL.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] І. Г. Цмоць, та ін., «Система моніторингу технологічних процесів «розумного підприємства»,» *Вісник Національного університету «Львівська політехніка» «Інформаційні системи та мережі»*, вип. 887, с. 10-17, 2018.
- [2] M. Grieves, and J. Vickers, “Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems,” in *Transdisciplinary Perspectives on Complex Systems*. Switzerland: Springer International Publishing, 2018, pp. 85-113. https://dx.doi.org/10.1007/978-3-319-38756-7_4.
- [3] B. A. Talkhestani, N. Jazdi, W. Schlögl, and M. Weyrich, “Consistency check to synchronize the Digital Twin of manufacturing automation based on anchor points,” in *Proc. 51st CIRP Conf. on Manufacturing Systems*, Stockholm, 2018, pp. 159-164. <https://dx.doi.org/10.1016/j.procir.2018.03.166>.
- [4] A. Madni, et al., “Leveraging Digital Twin Technology in Model-Based Systems Engineering,” *Systems*, no. 7 (7), pp. 1-13, 2019. <https://dx.doi.org/10.3390/systems7010007>.
- [5] L. Hu, et al., “Modeling of Cloud-Based Digital Twins for Smart Manufacturing with MT Connect,” *Procedia Manufacturing*, vol. 26, pp. 1193-1203, 2018. <https://dx.doi.org/10.1016/j.promfg.2018.07.155>.
- [6] D. L. Hill, C. Studholme, and D. J. Hawkes, “Voxel similarity measures for automated image registration,” *SPIE Proceedings “Visualization in Biomedical Computing,”* vol. 2359, no. 205, 1994. <https://dx.doi.org/10.1117/12.185180>.
- [7] J. Carnero, D. Diaz-Pernil, J. L. Mari, and P. Real, “Doxel: Towards a Software for Processing and Visualizing Topology Computations in Doxel-based 3D+t Images,” in *Proc. 16th International Conf. on Applications of Computer Algebra ACA'10*, Vlor, 2010.
- [8] Ye. Sulema, I. Dychka, and O. Sulema, “Multimodal Data Representation Models for Virtual, Remote, and Mixed Laboratories Development,” in *Lecture Notes in Networks and Systems*, vol. 47, Springer Cham, 2018, pp. 559-569.
- [9] I. Dychka, and Ye. Sulema, “Logical Operations in Algebraic System of Aggregates for Multimodal Data Representation and Processing,” *KPI Science News*, vol. 6, pp. 44-52, 2018.
- [10] I. Dychka, and Ye. Sulema, “Ordering Operations in Algebraic System of Aggregates for Multi-Image Data Processing,” *KPI Science News*, vol. 1, pp. 15-23, 2019.
- [11] P. Davies, and I. Maconochie, “The Relationship Between Body Temperature, Heart Rate and Respiratory Rate in Children,” *Emergency Medicine Journal*, vol. 26 (9), 2008. <https://dx.doi.org/10.1136/emj.2008.061598>.
- [12] В. А. Свідерський, та ін., «Програмне забезпечення технології портландцементу,» *Строительные материалы и изделия*, № 1, с. 16-17, 2014. [Електронний ресурс]. Режим доступу: http://nbuv.gov.ua/ujrn/smii_2014_1_7. Дата звернення: Січ. 29, 2020.
- [13] В. А. Свідерський, та ін., «Програмне забезпечення технології низькотемпературних в'язучих матеріалів,» *Строительные материалы и изделия*, № 1-2, с. 22-24, 2017. [Електронний ресурс]. Режим доступу: http://nbuv.gov.ua/ujrn/smii_2017_1-2_11. Дата звернення: Січ. 29, 2020.
- [14] J. Zhu, and A. Wang, “Data Modeling for Big Data,” *CA Technology Exchange. Insights from CA Technologies*, vol. 3 (2), pp. 77-82, 2012.
- [15] G. E. Blelloch, *Introduction to Data Compression*. Pittsburgh, Pennsylvania, USA: Carnegie Mellon University, 2001.
- [16] S. Laine, and T. Karras, “Efficient Sparse Voxel Octrees – Analysis, Extensions, and Implementation,” in *Vidia Technical Report NVR-2010-001*, 2010.
- [17] K. H. Höhne, A. Pommert, M. Riemer, and U. Tiede, “3D Visualization of tomographic volume data using the generalized voxel model,” *The Visual Computer*, vol. 6, pp. 28-36, 1990.
- [18] CREATE: Multimodal Dataset for Unsupervised Learning and Generative Modelling of Sensory Data from a Mobile Robot, in *IEEE DataPort*, 2018. [Electronic resource]. Available: <https://iee-dataport.org/open-access/create-multimodal-dataset-unsupervised-learning-and-generative-modeling-sensory-data>. Accessed: Feb. 10, 2020.
- [19] S. Brodeur, S. Carrier, and J. Rouat, “CREATE: Multimodal Dataset for Unsupervised Learning, Generative Modelling and Prediction of Sensory Data from a Mobile Robot in Indoor Environments,” 35 p., 2018. [Electronic resource]. Available: https://iee-dataport.s3.amazonaws.com/docs/2537/specifications.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAJOHYI4KJCE6Q7MIQ%2F20200211%2Fus-east-1%2Fs3%2Faws4_request&X-Amz-Date=20200211T212311Z&X-Amz-SignedHeaders=Host&X-Amz-Expires=86400&X-Amz-Signature=25f1356de0fed1f805bb45d1d144a84f0ede96521c095a25eea8d0d6e1acd360. Accessed: Feb. 10, 2020.

[20] P. Atzeni, F. Bugiotti, L. Cabibbo, and R. Torlone, "Data modeling in the NoSQL world," *Computer Standards & Interfaces*, vol. 67, pp. 1-11, 2020.

[21] Ye. Sulema "ASAMPL: Programming Language for Mulsemmedia Data Processing Based on Algebraic System of Aggregates," in *Advances in Intelligent Systems and Computing*. Springer, Cham, vol. 725, pp. 431-442, 2018. https://dx.doi.org/10.1007/978-3-319-75175-7_43.

Рекомендована кафедрою автоматизації та інтелектуальних інформаційних технологій ВНТУ

Стаття надійшла до редакції 03.02.2020

Дичка Іван Андрійович — д-р техн. наук, професор, декан факультету прикладної математики;
Сулема Євгенія Станіславівна — канд. техн. наук, доцент, доцент кафедри програмного забезпечення комп'ютерних систем, e-mail: sulema@pzks.fpm.kpi.ua.

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ

I. A. Dychka¹
Y. S. Sulema¹

Multimodal data representation model for complex description of observation objects

¹ Igor Sikorsky Kyiv Polytechnic Institute

The paper is devoted to solving the problem of the complex presentation of multimodal data about an object of observation, the characteristics of which are measured and investigated taking into account the time stamps and the relationship between the data of various modalities. This problem is relevant for many engineering tasks. In particular, the comprehensive presentation of data on a physical object is important in the technology of digital twins, which has recently become more widely used. Based on the fact that a digital twin is determined by both a behavioral model and a visual model, the paper presents a data model based on the concept of a multimodal data element. This model makes possible a comprehensive and consistent presentation of data about the object of observation.

The mathematical model of the complex presentation of multimodal data is based on the concepts of an aggregate and a multi-image defined in an algebraic system of aggregates. The article analyzes the relationship between the complex presentation and processing of multimodal data, which are determined in time, using the apparatus of the algebraic system of aggregates and the consideration of aggregated data as functions of several variables. Thus, the analysis of the data of the muxel model can be performed using logical operations, ordering operations, and relations defined in the algebraic system of aggregates as well as operations and approaches defined in other mathematical concepts.

Since the muxel model presents large amounts of data, it is advisable to optimize the way how data is presented. This can be achieved by data compression, modification of the data structure, or a combination of the first and second approaches. In the first approach, it is advisable to apply muxel model data compression based on the RLE method. In the second approach, it is proposed to apply data structures similar to the Sparse Voxel Octree, which is used to reduce data volume of models in voxel graphics. The practical implementation of the processing of multimodal data presented using the muxel model can be performed by means of ASAMPL programming language.

Keywords: data representation model, multimodal data, big data.

Dychka Ivan A. — Dr. Sc. (Eng.), Professor, Dean of the Department of Applied Mathematics;
Sulema Yevgeniya S. — Cand. Sc. (Eng.), Associate Professor, Associate Professor of the Chair of Computer Systems Software, e-mail: sulema@pzks.fpm.kpi.ua.

Модель представления мультимодальных данных для комплексного описания объектов наблюдения

¹Национальный технический университет Украины
«Киевский политехнический институт имени Игоря Сикорского»

Рассмотрена проблема комплексного представления мультимодальных данных об объекте наблюдения, характеристики которого измеряются и исследуются с учетом времени и взаимосвязи между данными различных модальностей. Эта проблема актуальна для многих инженерных задач. В частности, комплексное представление данных о физическом объекте является важным в технологии цифровых двойников, которая в последнее время приобретает все более широкое применение. Исходя из того, что цифровой двойник определяется поведенческой моделью и визуальной моделью, в статье предложена модель данных, основанная на понятии элемента мультимодальных данных. Эта модель делает возможным комплексное и непротиворечивое представление данных об объекте наблюдения.

Математическая модель комплексного представления мультимодальных данных основывается на понятиях агрегата и мультиобраза, определенных в алгебраической системе агрегатов. В статье анализируется связь между комплексным представлением и обработкой мультимодальных данных, которые определены во времени, с помощью аппарата алгебраической системы агрегатов и рассмотрением наборов агрегированных данных как функций нескольких переменных. Таким образом, анализ данных муксельной модели может выполняться как с применением логических операций, операций упорядочения и отношений, определенных в алгебраической системе агрегатов, так и операций и подходах, определенных в других математических концепциях.

Поскольку муксельна модель представляет данные большого объема, целесообразно оптимизировать способ представления данных. Это можно достичь сжатием данных, модификацией структуры данных или комбинацией первого и второго подхода. В первом подходе целесообразно применять сжатие данных муксельной модели на основе метода RLE. Во втором подходе предлагается применять структуры данных, аналогичные разреженному воксельному октодереву, которое применяется для уменьшения объемов данных моделей в воксельной графике. Практическая реализация обработки мультимодальных данных, представленных с помощью муксельной модели, может быть выполнена с помощью языка программирования ASAMPL.

Ключевые слова: модель представления данных, мультимодальные данные, данные большого объема.

Дичка Иван Андреевич — д-р техн. наук, профессор, декан факультета прикладной математики;

Сулема Евгения Станиславовна — канд. техн. наук, доцент, доцент кафедры программного обеспечения компьютерных систем, e-mail: sulema@pzks.fpm.kpi.ua