Mulsemedia Data Representation Based on Multi-Image Concept

Yevgeniya Sulema
Igor Sikorsky Kyiv Polytechnic Institute, Ukraine
`sulema@pzks.fpm.kpi.ua`

Abhishek Bhattacharya
Athlone Institute of Technology, Ireland
`a.bhattacharya@research.ait.ie`

Niall Murray
Athlone Institute of Technology, Ireland
`nmurray@research.ait.ie`

Abstract—MulseMedia technology enables operation with multimodal data sets that requires their complex representation. In this paper, we propose a mathematical background of MulseMedia data representation. The application of both Algebraic System of Aggregates and Multi-Image Concept allows efficient representation and processing of MulseMedia data in computer systems as well as encourages development of new algorithms for data processing, including data compression and data modelling.

Keywords—mulsemedia, multimodal data, multi-image, algebraic system of aggregates

1 Introduction

The recent rapid development of new hardware technologies is enabling recording, processing, and reproduction of multiple sensorial media (MulseMedia) experiences. This requires development of new data representation models and new algorithms for multimodal data processing. Such new approaches can allow more efficient representation and processing of human-perceived types of information and, in this way, create new opportunities for widening the scope of MulseMedia application.

In spite of the general technological readiness in certain fields of MulseMedia and immersive technologies being relatively high [1], there is a lack of common approaches to MulseMedia data representation. Although, a systematic approach for MulseMedia content representation was defined in MPEG-V standard [2], [3], most existing standards are focused on individual media components and they do not consider the combination of data modalities.

There is a gap between physical nature of data (modality) and the algorithms and methods of data processing. In our opinion, this gap should be addressed upon mathematical principles for MulseMedia and multimodal data representation and processing.

2 Related Work

Since MulseMedia is a promising advancement of multimedia, related works concerning different aspects of MulseMedia concept, MulseMedia applications, etc. are continuously growing.

The overview of MulseMedia, its applications, and hardware is presented in [4], [5], [6], [7], [8], [9]. In particular, [6] gives the detailed information on human sensorial system, sensory data representation and discusses quality of service and quality of sensory experience. Hardware overview [7] shows the scope of technical possibilities for MulseMedia-based applications, particularly in education. Whilst [8] and [9] consider available hardware approaches as well as an overview of research methodologies and proposals of same for understand MulseMedia Quality of Experience (QoE). Similarly in [10], sensory effects and their influence on the quality of user's perception are discussed. Human Factors influence on user experience of MulseMedia is presented in [11] and [12]. Haptic technology and haptic devices are investigated in [13], [14], [15], particularly, the classification of haptic devices is given in [15]; advantages and disadvantages of different types of haptic devices are pointed out as well.

Considering some communication network related conditions [12], [16], [17], [18] analysed multimodal data synchronization issues. In particular, the results of a study aimed at clarification of the temporal boundaries within which video, haptic, and air-flow components can be successfully synchronized are presented in [9]. In a similar context, Quality of Service (QoS) and Quality of Experience (QoE) have been investigated in [19], [20], [21]. Particularly, in [21] the authors propose the adaptive MulseMedia framework for delivering scalable video and sensorial data to users over constrained networks. The readiness of the World Wide Web to sensory effects representation is addressed in [22]; in particular, the ways of embedding sensory effect metadata within Web content are proposed.

The analysis of these and other related research allows us to conclude that available approaches to MulseMedia data representation do not use a strong mathematical background. Thus, in this paper we give an overview of a mathematical apparatus based on both the Algebraic System of Aggregates and the Multi-Image Concept [23], [24], [25] and we demonstrate how this mathematical approach can be used for representation of different data modalities in MulseMedia applications.

3 Theoretical Background

3.1 Algebraic System of Aggregates

Since observation is a time domain based process and typically involves multimodal data sets, the apparatus of the Algebraic System of Aggregates [23], [24], [25] can be used to obtain a multi-image [23], [24], [25] which can be considered as a mathematical background for a digital twin of the object (subject, process, event, etc.). Algebraic System of Aggregates (ASA) is an algebraic system, a carrier of which is an arbitrary set of specific structures called aggregates [23], [24], [25]. An aggregate A is a tuple of arbitrary tuples, elements of which belong to certain sets:

$$A = \llbracket M_j | \langle a_i^j \rangle_{i=1}^{n_j} \rrbracket_{j=1}^{N} = \llbracket \{A\} | \langle A \rangle \rrbracket \tag{1}$$

where $\{A\}$ is a tuple of sets M_j ; $\langle A \rangle$ is a tuple of elements tuples $\langle a_i^j \rangle_{i=1}^{n_j}$ corresponding to the tuple of sets ($a_i^j \in M_j$).

This means that for defining the aggregate we need to indicate the sets, elements of which belong to the aggregate within appropriate tuples, and next we need to place the tuples of the elements belonging to these sets in the same order because there is a strict relation between them: the elements of the first tuple belong to the first set, the elements of the second tuple belong to the second set, etc.

Since the ASA is an algebraic system [26], [27], [28], it consists of sets ($\mathscr{M}$, $\mathscr{F}$, $\mathscr{R}$), where $\mathscr{M}$ is a non-empty set (carrier), elements of which are elements of the system; $\mathscr{F}$ is a set of operations; $\mathscr{R}$ is a set of relations.

The carrier of ASA is an arbitrary set of specific structures called aggregates. Tuple elements in an aggregate can be both strict and fuzzy values.

Aggregates can be compatible, quasi-compatible or incompatible. Aggregates A_1 and A_2 are compatible ($A_1 \doteqdot A_2$) if they have equal lengths and both the type and sequence order of these aggregates are the same. Aggregates A_1 and A_2 are quasi-compatible ($A_1 \doteq A_2$) if the type and sequence order of these aggregates coincide partly. There is no requirement of the equality of aggregates lengths in this case. Otherwise, aggregates A_1 and A_2 are incompatible ($A_1 \circeq A_2$). Besides, aggregates A_1 and A_2 can be hiddenly compatible: A_1 $(\doteqdot)$ A_2 if $A_1 \doteq A_2$ or $A_1 \circeq A_2$ but at the same time $\{A_1\} \equiv \{A_2\}$ which means that both aggregates have the same set of sets but the order of these sets differs.

The basic relations in ASA are: Is Equal (=), Is Less (<), Is Greater (>), Is Equivalent ($\equiv$), Includes ($\supset$), Is Included ($\subset$), Precedes ($\prec$), Succeeds ($\succ$).

Operations on aggregates in ASA include logical operations, ordering operations, and arithmetical operations.

The logical operations on aggregates are: Union ($\cup$), Intersection ($\cap$), Difference ($\setminus$), Symmetric Difference ($\triangle$), and Exclusive Intersection ($\neg$) [23]. The result of logical operations in ASA depends on aggregates compatibility. For example, an

Intersection of two aggregates A_1 and A_2 is the aggregate A_3 which includes only common components of both aggregates and is formed according to the following rule:

1. If $A_1 \doteqdot A_2$ then A_3 includes elements of both aggregates, which are common for them, in every tuple:

$$\begin{gathered}
A_1 = \llbracket M_1, M_2, \dots, M_N | \langle a_1^1, a_2^1, \dots, a_l^1 \rangle, \langle b_1^1, b_2^1, \dots, b_m^1 \rangle, \dots, \langle w_1^1, w_2^1, \dots, w_n^1 \rangle \rrbracket \\
A_2 = \llbracket M_1, M_2, \dots, M_N | \langle a_1^2, a_2^2, \dots, a_r^2 \rangle, \langle b_1^2, b_2^2, \dots, b_q^2 \rangle, \dots, \langle w_1^2, w_2^2, \dots, w_p^2 \rangle \rrbracket \\
A_3 = A_1 \cap A_2 = \left\llbracket M_1, M_2, \dots, M_N | \left\langle a_{l_1}^1, \dots, a_{l_\alpha}^1, a_{r_1}^2, \dots, a_{r_\beta}^2 \right\rangle, \right. \\
\left. \left\langle b_{m_1}^1, \dots, b_{m_\gamma}^1, b_{q_1}^2, \dots, b_{q_\delta}^2 \right\rangle, \dots, \left\langle w_{n_1}^1, \dots, w_{n_\lambda}^1, w_{p_1}^2, \dots, w_{p_\mu}^2 \right\rangle \right\rrbracket
\end{gathered} \tag{2}$$

where $a_{l_i}^1 \in \overline{a^1}, a_{l_i}^1 \in \overline{a^2}, i \in \langle 1, \dots, \alpha \rangle; a_{r_j}^2 \in \overline{a^1}, a_{r_j}^2 \in \overline{a^2}, j \in \langle 1, \dots, \beta \rangle;$
$b_{m_k}^1 \in \overline{b^1}, b_{m_k}^1 \in \overline{b^2}, k \in \langle 1, \dots, \gamma \rangle; b_{q_s}^2 \in \overline{b^1}, b_{q_s}^2 \in \overline{b^2}, s \in \langle 1, \dots, \delta \rangle;$
$w_{n_u}^1 \in \overline{w^1}, w_{n_u}^1 \in \overline{w^2}, u \in \langle 1, \dots, \lambda \rangle; w_{p_y}^2 \in \overline{w^1}, w_{p_y}^2 \in \overline{w^2}, y \in \langle 1, \dots, \mu \rangle.$

2. If $A_1 \circeq A_2$ then A_3 is a null-aggregate:

$$\begin{gathered}
A_1 = \llbracket M_1^1, M_2^1, \dots, M_N^1 | \langle a_1, a_2, \dots, a_l \rangle, \langle b_1, b_2, \dots, b_m \rangle, \dots, \langle w_1, w_2, \dots, w_n \rangle \rrbracket \\
A_2 = \llbracket M_1^2, M_2^2, \dots, M_K^2 | \langle c_1, c_2, \dots, c_r \rangle, \langle d_1, d_2, \dots, d_q \rangle, \dots, \langle z_1, z_2, \dots, z_p \rangle \rrbracket \\
A_3 = A_1 \cap A_2 = \llbracket \emptyset | \langle \emptyset \rangle \rrbracket = A_\emptyset.
\end{gathered} \tag{3}$$

3. If $A_1 \doteq A_2$ then A_3 includes elements of both aggregates, which are common for them, only in tuples of common sets, thus, the number of sets shortens:

$$\begin{gathered}
A_1 = \llbracket M_1, M_2^1, \dots, M_x, \dots, M_N^1 | \langle a_1^1, a_2^1, \dots, a_l^1 \rangle, \langle b_1, b_2, \dots, b_m \rangle, \dots, \\
\langle f_1^1, f_2^1, \dots, f_t^1 \rangle, \dots, \langle w_1, w_2, \dots, w_n \rangle \rrbracket \\
A_2 = \llbracket M_1, M_2^2, \dots, M_x, \dots, M_K^2 | \langle a_1^2, a_2^2, \dots, a_r^2 \rangle, \langle d_1, d_2, \dots, d_q \rangle, \dots, \\
\langle f_1^2, f_2^2, \dots, f_v^2 \rangle, \dots, \langle z_1, z_2, \dots, z_p \rangle \rrbracket \\
A_3 = A_1 \cap A_2 = \left\llbracket M_1, \dots, M_x | \left\langle a_{l_1}^1, \dots, a_{l_\alpha}^1, a_{r_1}^2, \dots, a_{r_\beta}^2 \right\rangle, \dots, \right. \\
\left. \left\langle f_{t_1}^1, \dots, f_{t_\rho}^1, f_{v_1}^2, \dots, f_{v_\omega}^2 \right\rangle \right\rrbracket
\end{gathered} \tag{4}$$

where $a_{l_i}^1 \in \overline{a^1}, a_{l_i}^1 \in \overline{a^2}, i \in \langle 1, \dots, \alpha \rangle; a_{r_j}^2 \in \overline{a^1}, a_{r_j}^2 \in \overline{a^2}, j \in \langle 1, \dots, \beta \rangle;$
$f_{t_e}^1 \in \overline{f^1}, f_{t_e}^1 \in \overline{f^2}, e \in \langle 1, \dots, \rho \rangle; f_{v_h}^2 \in \overline{f^1}, f_{v_h}^2 \in \overline{f^2}, h \in \langle 1, \dots, \omega \rangle.$

Ordering operations include: Sets Ordering $\vDash$; Ascending Sorting $\uparrow$; Descending Sorting $\downarrow$; Singling $\parallel$; Extraction $\ltimes$; Insertion $\rtimes$ [24].

For example, Ascending Sorting enables reordering of all tuples according to ascending order of a certain tuple (called a primary tuple) among all tuples of the aggregate. Thus, if $A_1 = \left\llbracket M_j | \langle a_i^j \rangle_{i=1}^{n_j} \right\rrbracket_{j=1}^{N}$ and $\exists k$ such as $1 < k < N, k \neq 2$ and $n_1 >$

$n_k > n_N, n_2 = n_k$, then the result of Ascending Sorting operation of A_1 according elements of tuple $\bar{a}^k$ is the aggregate A_2 such as:

$$A_2 = A_1 \uparrow \bar{a}^k = [\![M_1, M_2, \ldots, M_k, \ldots, M_N | \langle a^1_\alpha, a^1_\beta, \ldots, a^1_\nu, \ldots, a^1_\omega, a^1_{n_k+1}, \ldots, a^1_{n_1}\rangle, \\ \langle a^2_\alpha, a^2_\beta, \ldots, a^2_\nu, \ldots, a^2_\omega\rangle, \ldots, \langle a^k_\alpha, a^k_\beta, \ldots, a^k_\nu, \ldots, a^k_\omega\rangle, \ldots, \langle a^N_\alpha, a^N_\beta, \ldots, a^N_\nu\rangle]\!] \tag{5}$$

where $a^k_\alpha < a^k_\beta < \ldots < a^k_\nu < \ldots < a^k_\omega$, $a^j_m \in \langle a^j_i\rangle^{n_j}_{i=1}$, $j = 1\ldots N$, $m \in [\alpha, \beta, \ldots, \nu, \ldots, \omega]$, $1 \le m \le n$, and $n = n_k$ if $n_j \ge n_k$ or $n = n_j$ if $n_j < n_k$.

The ordering operations have a specific importance in the ASA because they allow us to compose and to operate with complex data structures called multi-images [23], [24], [25].

3.2 Multi-Image Concept

A multi-image is a complex representation of multiple data sets describing an object (subject, process, etc.) of observation which is obtained (measured, generated) in the course of time. In mathematical sense, the multi-image is an aggregate, the first data tuple of which is a non-empty tuple of time values. These values can be natural numbers or values of any other type which can be used for evidence and monosemantic representation of time. Thus, the multi-image can be defined in the following way:

$$I = [\![T, M_1, \ldots, M_N | \langle t_1, \ldots, t_\tau\rangle, \langle a^1_1, \ldots, a^1_{n_1}\rangle, \ldots, \langle a^N_1, \ldots, a^N_{n_N}\rangle]\!] \tag{6}$$

where T is a set of time values; $\tau \ge n_i, i \in [1, \ldots, N]$.

If there is a tuple, elements of which are constant in time, in reference to the multi-image, the corresponding set in the tuple of sets is marked as $\dot{M}$ and the tuple of identical elements is defined by one element which is supposed to be valid for every time value.

The precondition of a multi-image construction is that obtained (generated, measured) data sequences are of different modalities and they are recorded (generated, measured) with respect to time.

Since a multi-image is an aggregate, to process its data, we can use any operations defined in ASA. In particular, the logical operations of ASA can be used for preparing complex data representation of a multi-image and the ordering operations enable processing of multi-image data with respect to time [23], [24], [25].

4 Mulsemedia Data Representation

Let us apply mathematical approach of the Multi-Image Concept to MulseMedia data sequences. For this purpose, at first, we need to consider every data modality related to MulseMedia and give its formal representation.

Thus, monophonic digital audio signal is a tuple defined as follows:

$$\bar{S}_{mono} = \langle s_k \rangle_{k=0}^{K} = \langle s_0, s_1, \dots, s_K \rangle \tag{7}$$

where a_k is a physical value such as sound pressure level; $a_k \in R$.

Then stereophonic digital audio signal can be represented as the following tuple of tuples:

$$\bar{S}_{stereo} = \langle\langle s_k^1, s_k^2 \rangle\rangle_{k=0}^{K} = \langle\langle s_0^1, s_0^2 \rangle, \langle s_1^1, s_1^2 \rangle, \dots, \langle s_K^1, s_K^2 \rangle\rangle \tag{8}$$

where s_k^1 is a sound pressure level value for the left channel; s_k^2 is a sound pressure level value for the right channel; $s_k^{1,2} \in R$.

Finally, multichannel digital audio signal [4] can be formalized as the following tuple:

$$\begin{gathered}\bar{S}_{multi} = \langle\langle s_k^1, s_k^2, \dots, s_k^Q \rangle\rangle_{k=0}^{K} = \\ \langle\langle s_0^1, s_0^2, \dots, s_0^Q \rangle, \dots, \langle s_K^1, s_K^2, \dots, s_K^Q \rangle\rangle\end{gathered} \tag{9}$$

where s_k^q is a sound pressure level value for q-channel; $q \in [1, \dots, Q]$; $s_k^q \in R$.

Let us consider images. Then a 2D image of N×M pixels can be defined as a tuple of the following tuples:

$$\bar{C}_{2D} = \langle\langle \mathrm{c}_n^1, \mathrm{c}_n^2, \dots, \mathrm{c}_n^M \rangle\rangle_{n=0}^{N} = \langle\langle \mathrm{c}_0^1, \dots, \mathrm{c}_0^M \rangle, \dots, \langle \mathrm{c}_N^1, \dots, \mathrm{c}_N^M \rangle\rangle \tag{10}$$

where c_n^m is a tuple of pixel colour components; $\mathrm{c}_n^m \in Z^3$; $m \in [1, \dots, M]$.

The definition of c_n^m depends on a colour model. For example, if RGB model is used, c_n^m is defined as follows:

$$\mathrm{c}_n^m = \langle r_n^m, g_n^m, b_n^m \rangle \tag{11}$$

where r_n^m, g_n^m, b_n^m are colour components in RGB model; $r_n^m, g_n^m, b_n^m \in Z$.

The formal description of a 3D image depends on the technology of its reproduction [1], [7]. For example, if voxel graphics is supposed to be used for visualisation of the 3D image with the size of N×M×K voxels, then it can be defined as follows:

$$\bar{C}_{3D} = \langle\langle\langle \mathrm{c}_{n,k}^1 \rangle_{k=1}^{K}, \langle \mathrm{c}_{n,k}^2 \rangle_{k=1}^{K}, \dots, \langle \mathrm{c}_{n,k}^M \rangle_{k=1}^{K} \rangle\rangle_{n=1}^{N} \tag{12}$$

where $\mathrm{c}_{n,k}^m$ is a tuple of colour components, $\mathrm{c}_{n,k}^m \in Z^3$; $m \in [1, \dots, M]$.

If the 3D image is a stereo image, then its description is:

$$\bar{C}_{stereo} = \langle\langle\langle l_n^1, l_n^2, \dots, l_n^M \rangle, \langle r_n^1, r_n^2, \dots, r_n^M \rangle\rangle\rangle_{n=0}^{N} \tag{13}$$

where l_n^m is left frame data component; r_n^m is right frame data component.

A 2D video can be represented as an aggregate of an audio signal data tuple of a certain type (monophonic, stereophonic, multichannel) and a tuple of 2D images:

$$V_{2D} = \llbracket R^Q, Z^3 \mid \bar{S}, \bar{C} \rrbracket \tag{14}$$

where $\bar{S}$ is $\bar{S}_{mono}$, $\bar{S}_{stereo}$ or $\bar{S}_{multi}$; $\bar{C}$ is $\bar{C}_{2D}$ or $\bar{C}_{3D}$.

A 3D video can be defined as a combination of a stereo image sequence and spatial sound (Q-channel audio signal):

$$V_{3D} = [\![R^Q, Z^3 \mid \bar{S}_{multi}, \bar{C}_{stereo}]\!] \tag{15}$$

Haptic data received from a data glove [12] can be represented as a tuple of movement intensity for every finger. Then if every finger movement is registered by one sensor, we can define haptic data tuple as follows:

$$\bar{H} = \langle\langle h_k^1, h_k^2, h_k^3, h_k^4, h_k^5\rangle\rangle_{k=0}^K \tag{16}$$

where h_k^i is a movement intensity value; $h_k^i \in R$; $i \in [1, \ldots, 5]$.

Thermoceptional (environmental) data can be formalized as:

$$\bar{E} = \langle\langle e_k^T, e_k^H, e_k^W\rangle\rangle_{k=0}^K \tag{17}$$

where e_k^T is a temperature value, $e_k^T \in R$; e_k^H is a humidity value, $e_k^H \in R$; e_k^W denotes the wind effect intensity, $e_k^W \in R$.

Olfactory data can be formally defined as:

$$\bar{O} = \langle\langle o_k^T, o_k^I, o_k^d, o_k^s\rangle\rangle_{k=0}^K \tag{18}$$

where o_k^T denotes the scent types following the popular olfactory model in [6], [13], [16] $o_k^T \in M_O = \{"burnt", "flowery", "foul", \ \ "fruity", "resinous", "spicy"\}$; o_K^I denotes the olfaction intensity as in [28], $o_K^I \in M_I = \{low, medium, high\}$; $o_k^d \in R$ denotes the duration of olfactory stimuli, and $o_k^s \in R$ denotes the synchronization skew of olfaction which is proved to be crucial [16], [18].

Finally, gustatory data can be formalized as:

$$\bar{G} = \langle\langle g_k^T, g_k^I, g_k^d\rangle\rangle_{k=0}^K \tag{19}$$

where $g_k^T \in M_G = \{“sour”, ”bitter”, ”salty”, ”sweet”\}$ denote the gustation stimulus types [26], $g_k^I \in M_I = \{low, medium, high\}$ denote the gustation intensity [29], and $g_k^d \in R$ denote the duration of gustation stimuli [29].

In this paper, we assume for simplicity that every data structure defined in formulas from (3) to (15) is recorded (measured) simultaneously. Then, finally, we can compose a multi-image of the object of observation, such as:

$$\begin{aligned} I = [\![T, R^Q, Z^3, R^5, R^3, \langle M_O, M_I, R^2\rangle, \langle M_G, M_I, R\rangle | \langle t_1, .., t_\tau\rangle, \\ \bar{S}_{multi}, \bar{C}_{stereo}, \bar{H}, \bar{E}, \bar{O}, \bar{G}]\!] \end{aligned} \tag{20}$$

In the next section, we analyse how this mathematical approach can be applied in real conditions.

5 Results and Discussion

One of the possible use cases for Multi-Image Concept application is a MulseMedia-based remote lab (Fig. 1). Such lab can be used for remote study as well as immersive demonstration of a certain phenomenon to remote learners [30], [31], [32], [33]. A real-world object (subject, process, event, and phenomenon) can be characterized by multiple sequences of human-perceived information.

We assume that the data sequences are obtained from several sensors. Every data sequence characterizes a certain aspect of the object's nature and behaviour. At the same time, in terms of mathematics every data sequence is a tuple of values belonging to a certain data set. As mentioned above, for simplicity we assume that in one session of work in the lab all data sequences are recorded (measured) simultaneously. The result is that they can be composed in one complex data structure, namely, an aggregate. Since the object of observation exists in time, the aggregate can be supplemented with a tuple of time values which correspond to the time moments when values of other modalities are measured (recorded).

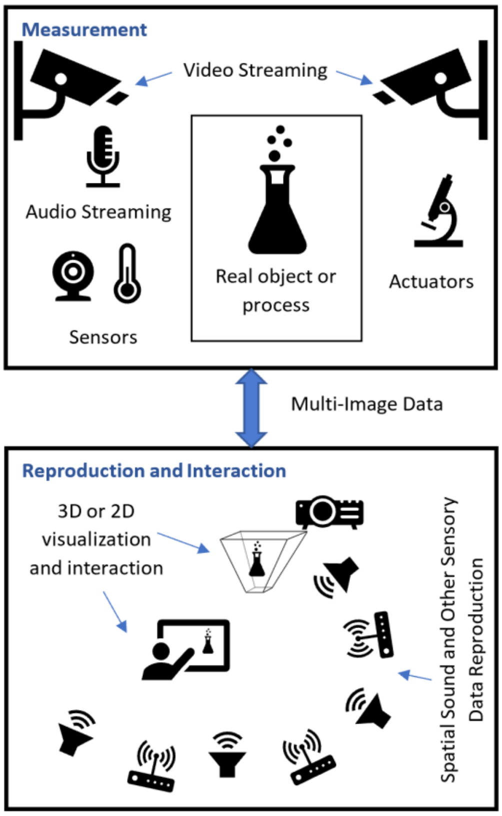

Fig. 1. Basic idea of complex representation of MulseMedia data as a multi-image of a real-world object

The multi-image of the object defined with respect to a certain period of time. In the next session of work with the lab, we obtain the next multi-image of the same object and so on. After a number of observations and measurements, we will possess a family of multi-images of the object of study. Thus, we come to the task of multi-images analysis. Such analysis based on logical operations [23], ordering operations [24], and relations defined in the Algebraic System of Aggregates according to the Multi-Image Concept. The main advantage of this concept is that we can operate with multiple values defined in a certain time moment as with a complex value of a function of several variables. Thus, we always have an overall view on the object's nature and behaviour in every moment of time. It allows us to compare multi-images, predict further state and behaviour of the object, model it, etc.

Moreover, having the object's multi-image for a certain time duration, we can reproduce the object virtually by using hardware for 3D visualization, spatial sound and other sensory data reproduction.

Let us consider an example of multi-image representation and processing. Let $\bar{V}$ be a short video fragment recorded in a time interval $\bar{T} = \langle 1, \dots, 10\rangle$ (in seconds) and presented as a data aggregate accordingly to (10).

Note that for simplicity we assume that the quantity of 2D frames and the quantity of audio signal samples in this recorded fragment are equal as well as we do not specify values in tuples but use their formal notations $\langle\langle s_k^1, s_k^2\rangle\rangle_{k=0}^{10}$ and $\langle\langle \mathrm{c}_{0,k}^1, \dots, \mathrm{c}_{0,k}^M\rangle, \dots, \langle \mathrm{c}_{N,k}^1, \dots, \mathrm{c}_{N,k}^M\rangle\rangle_{k=0}^{10}$ instead.

Then we can obtain a multi-image I_1 which represents video data with respect to time:

$$I_1 = [\![T, R^2, Z^3 \mid \langle 1,2,3,4,5,6,7,8,9,10\rangle, \langle\langle s_k^1, s_k^2\rangle\rangle_{k=0}^{10}, \\ \langle\langle \mathrm{c}_{0,k}^1, \dots, \mathrm{c}_{0,k}^M\rangle, \dots, \langle \mathrm{c}_{N,k}^1, \dots, \mathrm{c}_{N,k}^M\rangle\rangle_{k=0}^{10}]\!] \tag{21}$$

Besides, let us assume that we need to enrich this video fragment by adding olfactory effects, namely, flowery scent of low intensity starting with the 3rd second of the video and lasting 3 seconds (all numbers in this example are invented) as well as fruity scent of medium intensity starting with the 7th second and lasting 4 seconds, with zero skew in both cases. Then it can be expressed in the following multi-image I_2:

$$I_2 = [\![T, \langle M_O, M_I, R^2\rangle \mid \langle 3,7\rangle, \langle "flowery", low, 3, 0\rangle, \\ \langle "fruity", medium, 4, 0\rangle]\!] \tag{22}$$

Thus, we have two multi-images obtained from different sources: the first one is received from external device (video camera) and the second one is composed artificially (for example, by using a certain MulseMedia editor). Now we need to synchronize them and combine in one multi-image describing some immersive scene. For this purpose, we apply the following operations of the ASA: Union (∪), Sorting (↑) and Singling (‖) [23], [24]. Since these multi-images are quasi-compatible [23], [24] the result of this complex operation is as follows:

$$
\begin{aligned}
I = \big((I_1 \cup I_2) \uparrow \bar{t}\big) \parallel \bar{t} = [\![T, R^2, Z^3 \,|\, \langle 1,2,3,4,5,6,7,8,\\
9,10\rangle, \langle\langle s_k^1, s_k^2\rangle\rangle_{k=0}^{10}, \big\langle\langle \mathrm{c}_{0,k}^1, \dots, \mathrm{c}_{0,k}^M\rangle, \dots,\\
\langle \mathrm{c}_{N,k}^1, \dots, \mathrm{c}_{N,k}^M\rangle\big\rangle_{k=1}^{10}, \langle \emptyset, \emptyset, \langle "flowery", low, 3, 0\rangle,\\
\emptyset, \emptyset, \emptyset, \langle "fruity", medium, 4, 0\rangle, \emptyset, \emptyset, \emptyset\rangle]\!]
\end{aligned}
\tag{23}
$$

The multi-image *I* defined by (19) gives a full description of the immersive scene in its dynamics and with internal synchronization of multimodal data. It can be rendered by using a domain-specific language such as ASAMPL [25].

6 Conclusions

The Algebraic System of Aggregates is a tool for complex data representation. It enables multimodal data processing by using a range of operations and relations. In particular, the logical operations allow us to construct different compositions of multimodal data, which in turn enables complex data representation for compound description of objects and processes in different areas.

The Algebraic System of Aggregates is the basis for the Multi-Image Concept which enables overall description of an object (subject, process, phenomenon) of observation carried out in the course of time. In our opinion, the Multi-Image Concept creates a background for the development of new MulseMedia data models, new multimodal data compression and coding methods, new algorithms of MulseMedia data processing, and new protocols for MulseMedia data transmission. The formalization approach proposed by the Multi-Image Concept also enables better standardization for hardware and software to be used in immersive technologies.

Future work can be focused on the development of data synchronization methods and data compression algorithms.

7 Acknowledgement

This work has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) and is co-funded under the European Regional Development Fund under Grant Number 16/RC/3918.

8 References

1. Yevgeniya Sulema (2016), Mulsemedia Vs. Multimedia: State of the Art and Future Trends (Invited Paper), in Proceedings of the 23rd IEEE International Conference on Systems, Signals and Image Processing IWSSIP2016, Bratislava, Slovakia, pp. 19-23.
2. Jae Joon Han and Sang-Kyun Kim, Text of white paper on MPEG-V, Communication Group, http://mpeg.chiariglione.org/standards/mpeg-v
3. Kyoungro Yoon, Sang-Kyun Kim, Jae Joon Han, Seungju Han and Marius Preda (2015). MPEG-V: Bridging the Virtual and Real World. Academic Press, pp. 210. ISBN: 978-0-12-420140-8

4. Christian Timmerer, Karsten Müller (2010). Immersive Future Media Technologies: From 3D Video to Sensory Experiences, Proceedings of the International Conference on Multimedia (MM'10) (Alberto del Bimbo, Shih-Fu Chang, Arnold Smeulders, eds.), ACM, New York, NY, USA, pp. 1781-1782.
5. George Ghinea, Frederic Andres, Stephen R. Gulliver (2012). Multiple Sensorial Media Advances and Applications: New Developments in MulSeMedia. pp. 344, DOI: 10.4018/978-1-60960-821-7
6. Ghinea G, Timmerer, C., Lin, W. and Gulliver, S.R. (2014). Mulsemedia: State-of-the-Art, Perspectives and Challenges. ACM Trans. Multimedia Computing Communications and Applications, pp. 1-23. ISSN 1551-6865 doi: 10.1145/261799
7. Péter Tamás Kovács, Niall Murray, Gregor Rozinaj, Yevgeniya Sulema, Renata Rybárová (2015). Application of Immersive Technologies for Education: State of the Art. Proceedings of the 9th International Conference on Interactive Mobile Communication Technologies and Learning (IMCL2015), Thessaloniki, Greece
8. Niall Murray, Brian Lee, Yuansong Qiao, Gabriel-Miro Muntean (2016). Olfaction-Enhanced Multimedia: A Survey Application Domains, Displays and Research Challenges". In ACM Computing Surveys 48:4, article 56, 2016
9. Niall Murray, Oluwakemi A. Ademoye, George Ghinea and Gabriel-Miro Muntean (2017). A Tutorial for Olfaction-based Multisensorial Media Application Design and Evaluation" In ACM Computing Surveys vol. 50, issue 5, article 67.
10. Markus Waltl (2010). Enriching Multimedia with Sensory Effects, Saarbrücken, Germany, pp. 100
11. Niall Murray, Brian Lee, Yuansong Qiao and Gabriel Miro-Muntean (2016). The Influence of Human Factors on Olfaction based Mulsemedia Quality of Experience. In Special Session on User Factors in Multimedia Experiences at the 8th International Conference on Quality of Multimedia Experience (QoMEX 2016), pp. 1-6, doi: 10.1109/QoMEX.2016.7498975
12. Ghinea G., Ademoye O. (2010). A User Perspective of Olfaction-Enhanced Mulsemedia, Proceedings of the international conference on management of emergent digital ecosystems (MEDES'10), ACM, New York, pp. 277-280
13. Mihelj, M. & Podobnik, J. (2012). Haptics for Virtual Reality and Teleoperation, Intelligent Systems, Control and Automation: Science and Engineering. DOI: 10.1007/978-94-007-5718-9_2
14. Steinbach, E.; Hirche, S.; Ernst, M.; Brandi, F.; Chaudhari, R.; Kammerl, J.; Vittorias, I. (2012). Haptic Communications, Proceedings of the IEEE, pp. 937-956, Volume: 100, Issue: 4
15. Yevgeniya Sulema (2015). Haptic Interaction in Educational Applications. Proceedings of the 9th International Conference on Interactive Mobile Communication Technologies and Learning (IMCL2015), Thessaloniki, Greece.
16. N Murray, Y Qiao, B Lee, AK Karunakar, GM Muntean (2013), Subjective evaluation of olfactory and visual media synchronization, Proceedings of the 4th ACM Multimedia Systems Conference, pp. 162-171
17. Murray, N., Lee, B., Qiao, Y., and Muntean, G.-M. (2014). Multiple-scent enhanced multimedia Synchronization. ACM Trans. Multimedia Comput. Commun, pp. 20
18. Zhenhui Yuan, Ting Bi, Gabriel-Miro Muntean, Gheorghita Ghinea (2015). Perceived Synchronization of Mulsemedia Services. IEEE Transactions on Multimedia, Vol. 17, No. 7, p. 957-966
19. Christian Timmerer, Markus Waltl, Benjamin Rainer, Niall Murray (2014). Sensory Experience: Quality of Experience Beyond Audio-Visual, Chapter in Quality of

Experience: Advanced Concepts, Applications and Methods (Sebastian Möller, Alexander Raake, eds.), Springer, Heidelberg, pp. 351-365.

20. Zhenhui Yuan, Gheorghita Ghinea, Gabriel-Miro Muntean (2015). Beyond Multimedia Adaptation: Quality of Experience-Aware Multi-Sensorial Media Delivery. IEEE Transactions on Multimedia, Vol. 17, No. 1, p. 104-117
21. Zhenhui Yuan, Shengyang Chen, Gheorghita Ghinea, Gabriel-Miro Muntean (2014). User Quality of Experience of Mulsemedia Applications. ACM Transactions on Multimedia Computing, Communications, and Applications - Special Issue on MulSeMedia: Advances and Applications. Volume 11 Issue 1s, Article No. 15
22. Christian Timmerer, Markus Waltl, Hermann Hellwagner (2010). Are Sensory Effects Ready for the World Wide Web? In Proceedings of the Workshop on Interoperable Social Multimedia Applications (WISMA 2010), CEUR Workshop Proceedings (CEUR-WS.org), Aachen, Germany, pp. 57-60.
23. Ivan Dychka, Yevgeniya Sulema (2018). Logical Operations in Algebraic System of Aggregates for Multimodal Data Representation and Processing, Research Bulletin of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Vol. 6, pp. 44-52.
24. Ivan Dychka, Yevgeniya Sulema (2019). Ordering Operations in Algebraic System of Aggregates for Multi-Image Data Processing, Research Bulletin of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Vol. 1, pp. 15-23.
25. Sulema Yevgeniya (2018). ASAMPL: Programming Language for Mulsemedia Data Processing Based on Algebraic System of Aggregates, Advances in Intelligent Systems and Computing, Springer, Vol. 725, pp. 431-442, DOI: 10.1007/978-3-319-75175-7_43.
26. A. A. Fraenkel, et al. (1973), “Foundations of Set Theory”, Elsevier, 415 p.
27. A. I. Maltsev (1970), “Algebraic systems”, Nauka, 392 p. (in Russian)
28. B. A. Kulik, A. A. Zuenko, A. Ya. Fridman (2010), “Algebraic approach to intellectual processing of data and knowledge”, Izdatelstvo Politekhnicheskogo Universiteta, 235 p. (in Russian)
29. Steven T. Kelling, Bruce P. Halpern (1988). “Taste judgements and gustatory stimulus duration: taste quality, taste intensity, and reaction time”, Chemical Senses, Volume 13, Issue 4, pp. 559-586.
30. Wenshan Hu, Guo-Ping Liu, Hong Zhou (2013). Web-Based 3-D Control Laboratory for Remote Real-Time Experimentation. IEEE Transactions on Industrial Electronics, Vol. 60, No. 10, pp. 4673-4682
31. Mohamed Tawfik, et al. (2013). Virtual Instrument Systems in Reality (VISIR) for Remote Wiring and Measurement of Electronic Circuits on Breadboard. IEEE Transactions on Learning Technologies, Vol. 6, No. 1
32. Potkonjak V., et al. (2016). Virtual Laboratories for Education in Science, Technology, and Engineering: a Review. Computers & Education, Volume 95, pp. 309-327, 2016
33. F. Esquembre (2015). Facilitating the Creation of Virtual and Remote Laboratories for Science and Engineering Education. ScienceDirect, IFAC-PapersOnLine 48-29, 049-058.