Advanced Sound Manipulation in Interactive Multimedia Environments

Yannis Deliyiannis[1], Andreas Floros[2] and Christos Tsakostas[3]

[1] Dept. of Audiovisual Arts, Ionian University, Corfu, Greece, yiannis@ionio.gr

[2] Dept. of Audiovisual Arts, Ionian University, Corfu, Greece, floros@ionio.gr

[3] Holistiks Engineering Systems, Athens, Greece, tsakostas@holistiks.com

Abstract — **Multimedia standards, frameworks and models are already well established for generalized presentations. However, the situation is much less advanced for systems, which require the combination of advanced sound-oriented features and capabilities similar to those used to interact with highly demanding visual content. In this respect, current commercial presentation applications have been found lacking, often revealing multifaceted presentation limitations, including lack of sound control and delivery. They rarely offer cross-platform compatibility, provide limited programmability, are restrictive on data-interaction, and only support static WWW-based delivery. To overcome the above-stated deficiencies, a number of innovations are proposed, including the presentation of a combined multimedia framework, supported by a model that describes content-connectivity and stream-synchronization, enabling interaction for all audiovisual data-types included in the system.**

I. Introduction

Today most commercial multimedia applications permitting little interaction and control of sound, which is usually limited to background use, while enabling highly-interactive features for content such as images, video, animation and 3D environments. When exploring the majority of applications, we see that interaction is limited to pre-programmed event-driven actions and their combinations. Only a few environments such as MAX/MSP combined with Jitter and PD can support fully-interactive audiovisual scenarios, as they allow system-customization through user-programming, and this may be the main reason for their wide use in interactive installations and computer music. Substantive goal of this research has been to deliver comprehensive multimedia-system functionality featuring advanced sound synchronization, control and advanced interactive capabilities, both at the stand-alone system level and over the Internet, via a single, multi-purpose implementation [15].

When examining commercial multimedia development practices, M. Lang and C. Barry reported: “there is no uniform approach to multimedia systems development and the approaches being prescribed by *the literature are not being used in practice*” [1]. Professional multimedia companies used to employ frameworks developed in-house [28], more particularly to accommodate specific client-needs. Today this continues to be the case. One may attribute this state-of-affair to the knowledge-background of the developer, and less to the distinct nature of the multimedia case studies and end-applications, particularly when these involve complex sound environments which have to be built to order.

To date and for practical reasons, multimedia system (MMS) development normally commences once the developer has received all prerequisite data. This offers limited system functionality, as only a few input sources are then made available to the application, mainly due to compatibility requirements. We see in various applications the sound input to be limited to a single low-resolution microphone line, which may be extended to a stereo input for sound processing in some cases. Propriety multimedia authoring environments such as Adobe Director [35] utilizing Shockwave technology [30] require external additional code to be loaded before they can fully utilize and process sound. This is also the case with state-of-the-art 3D multimedia authoring environments [39], which typically support dynamic rendering but limit sound use to background purposes, requiring advanced programming from the developer’s side.

As a result, much effort has been directed towards system-functionality, often disregarding significant issues, such as quality software development practices (*quality assurance, risk management, validation, verification and configuration management*), other related guidelines, and state-of-the-art web and multimedia development technologies [1], [29].

A principal difference between the proposed methodology presented in this paper and those used commercially, lies in the fact that here, MMS-development may be initiated after the specification stage. This removes the need for data provision upfront. As data may be treated in a modular fashion (plugged-in later), one may extend this approach somewhat further, leaving data insertion to the user. Other departures in methodology from the norm include direct use of the model for automated MMS-development and verification (being rule-based).

Both proposed framework and model are independent of the implementation platform and sound is represented as any other stream of data, reducing the complexity and allowing uniform interaction features and synchronization methods between different content-types. In fact the same multimedia methodologies may be used for applications focusing on audio as well as video, given of course that the appropriate tools are provided for each content type. This is actually proved today as typical music-based environments are used today for the development of interactive audiovisual installations. In that respect Multimedia Environments (MME) may also be utilized for the implementation of audio-based environments and dynamic sound-manipulation, as they offer flexibility for streams with various content-types, and support modular unit-construction (frames). An alternative approach would be to implement the target-system in JAVA, with multimedia-extensions.

II. Comments on Other Frameworks and Models

In the literature, a framework is used to describe interaction between the various parties involved, from the construction of a presentation to its actual delivery. A model describes, in detail, the construction/presentation components. This would include content-organization mechanisms and other implementation-level issues. Prior discussion has illustrated the need for new content- and presentation-organization mechanisms, with respect to complex audiovisual content. The weakness in existing frameworks, standards, models and architectures is exposed through implementation, with actual complex data-content such as multiple input and output sound streams. When examining the formulation of existing frameworks, a common observation is that their suitability is not generalized, but problem-related. The frameworks discussed earlier need various adaptations before they may be used as educational tools in the field of audio and music-education. Take for example the case of educational frameworks, where these are formulated to support courseware delivery, separating multimedia development from the expert (researcher/teacher) and target audience (student, academic, commercial). When such frameworks are used over varied content-types, they fail to specify explicitly the relationship between developers, users and target audience [34], [17].

Some have recognized there is a need to evaluate current educational presentation-frameworks, see for example “Ciao!”, Open University [6]. This is a higher-level evaluation framework, itself designed to classify educational construction mechanisms. It attempts to classify notions such as “Rationale”, “Data” and “Methods”, and evaluate the same for “Context”, “Interaction” and “Outcome”. Some generalized frameworks are tested in practice, through general-purpose implementations, such as the “Authoring on the fly” system [22]. Although quite general, the implementation offers features such as stream-playback and timing. It becomes apparent, that the underlying framework may be employed for courseware delivery (amongst other uses). A further example is MDF (Multimedia Description Framework) [16], based on the MPEG-7 standard, classified as a standardization of content description for video data/documents. A similar approach is discussed in [11], which addresses the difficulties encountered when different content-domains are introduced, namely the “*applicability problem*”. Only recently has the need been recognized for advanced-level presentations, through non-linear organization aspects [12].

Passing from content-based to user-centered design [32], [38] content-organization concerns become less prominent as user-interface construction issues take precedence. In contrast to the above, under user-centered design the issue is to automatically construct user-interfaces by querying the content. In this context, one encounters frameworks and models intended for user-interface design. These provide appropriate user-interface guidelines and structures [13], relating to underlying content. Examples would include frameworks, which permit particular interaction models/levels over only visual data and its representation or some educational models (termed Model-Based User-Interface Development Environments MB-UIDEs). A characteristic instance, cited as “TELLACH” [33], [23] describes a modular tool that supports automatic interface development via a code generator. This tool is employed to design a library system via a database-oriented environment. Other models/frameworks include MOBI-D [25]. When automatic generation is employed, some systems are left partially functional. Here, variation in content-domains inhibits a “one-for-all” approach and a “one-off” user-interface construction process. In this context, an adaptive user-interface is required. Hence, the advent of knowledge-elicitation systems and decision trees [37], [26]. A more complete evaluation of the main characteristics of the many models available (ADEPT, AME, FUSE, GENIUS, HUMANOID, JANUS, ITS, MASTERMIND, MECANO, *MOBI-D*, TADEUS, *TELLACH*, TRIDENT and UIDE) is addressed under the Unified Modeling Language for interactive applications project (UMLi) [31]. Commonly, the complexity of such systems is a factor that limits their usability, utilizing custom Object-Oriented construction techniques [7] (non-compliant to a standard). Nevertheless, effective user-interface construction necessitates Object-Oriented techniques, so that re-usability and automatic lower-level system response may be implemented consistently. In contrast, other approaches, such as the use of Interaction Object Graphs (IOC’s) [2], offer a transition-diagram aid to user-interface development. These may easily be understood and implemented, particularly for straightforward problems (say, when tackling ambiguity using a WWW-browser, for example). Nevertheless, complexity does not permit their application in larger more complex data-domains. A similar, but earlier approach, used Petri-networks formally to specify human-software interaction [24]. Yet, another possibility is to employ trajectories in graphical user-interfaces. This notion combines and describes parameter variation and relationships using agents, in two-dimensional space [5]. In addition, there are frameworks proposed within international standards, an example being the draft-version ISO 14915-1. In this manner, concrete underlying specification is avoided, if implementation is at a sufficiently high-level. Such abstraction is deemed necessary, as the purpose of the standard is intended to provide general guidelines.

Overall, frameworks and models are either “too general and abstract”, or have been designed with “simplistic”, “uniform” and/or “easily-ordered” data in mind. For example, under most frameworks the roles of developing parties are distinct, separating the “performer” who controls the system from the “audience”, failing to describe the interactive nature and the mixed role of the user.

In general, when content is concerned, particular models are found to perform well with certain types of data, while they under-perform with others (usually non-linear). This introduces several disadvantages, when the data goes beyond this realm, and non-linear presentation demands are made, only to be met by a combination of linear structures. The introduction of non-linear content-organization and the advanced demands of complex interactive audiovisual content, necessitates novel organization methods. This is achieved through the introduction of a new multimedia presentation hierarchy, consisting of two interrelated parts: a framework and a model, design to specify/design and develop MMSs respectively. The proposed framework recognizes and represents interdisciplinary expertise between the parties involved in the development/use of complex-content systems. In addition, it may easily be adjusted to cover dynamically changing development and presentation requirements. Subsequently, the proposed model effectively deals with variable content-complexities, providing non-linear interaction without data re-organization of any type. These parts tackle various data-complexities, from simplistic linear-like structures, to dynamically changing scenarios such as live multi-channel audio.

A typical interactive multimedia example is described, in an attempt to present the complexity of the task in hand. In an abstract audiovisual system for content-exploration, the same interface constructs may be used for navigation, with the limitation the user should not be allowed to enter the same virtual-space twice. This particular task requires dynamic link re-programming, permitting the system to dynamically adjust the interface constructs, while providing full exploration capabilities to the user, for the remaining undiscovered audiovisual space. More complex scenarios include dynamic re-programming of a wide range of parameters, transforming internal content-context interrelations under user or event-driven scenarios.

The proposed model supports this type of functionality, allowing interactive re-programming for all input streams that may originate from devices, controllers, buttons, audio and video streams. Extensions to propriety environments such as MAX/MSP with Jitter may be implemented based on the iMM model, providing advanced system-functionality for complex interactive implementations, which are not limited to specific scenarios and parameter combinations.

III. Proposing a Suitable Model and Framework

The need for a suitable multimedia presentation mechanism is initiated, by the complexity of the task in hand. It is observed that system-usability becomes ever more problematic with escalating content-complexity. Here complexity is not described purely by the number of say audiovisual sources, but also the interaction with each of them. Here, various content-domains (with diverse target-audiences and increasing complexity) are described formally, which reveals the deficiencies behind existing frameworks and models.

First, a combined multimedia *organization, construction and presentation framework* is presented. Subsequently, a *model* that describes *content-connectivity* is composed. The *model* transforms the system-specification provided by the *framework*, into a complete Interactive Multi-Media System (iMMS). This is designed to accommodate the advanced presentation requirements commonly introduced with complex content used within interactive sound environments, allowing dynamic acquisition from various sources, processing and synchronized output, possibly combined with other forms of content such as static and interactive audio, video and animation streams. In such a fashion, the presentation of multimedia-content is managed within meaningful and easily accessible modes.

A. Interactive Multimedia Framework (iMMF)

The proposed (iMMF) framework, describes interaction between three particular parties, involved in the construction and use of a multimedia system. This includes developers, users (field-experts/scientists) and observers (audience). One may utilize an overlapping set-diagram representation, to describe the relationships between the various groups (Fig. 1). Framework-areas may be colored, to indicate individual and overlapping party-roles. Arrows may indicate group-interaction across the system. Set unions (A, B, C, D) indicate higher-order interaction, say between groups in construction-mode for a single-MMS, or indeed, a super-MMS. Such a system is an implied product of the developers' domain.

The first instance of interaction considered is area-exclusive. This implies only limited interaction between the groups, where there are clearly defined expertise-boundaries. Each group describes their requirements, according to their individual needs and constraints. A common agreement across groups, specifies the resulting MMS. From the individual perspective of each group, this is achieved without a detailed understanding of the mechanics of other groups. This is a typical scenario when, for example, a multimedia company is employed to design a MM-system. Line-diagrams may be used to specify relationships, in detail, between these clearly defined groups.

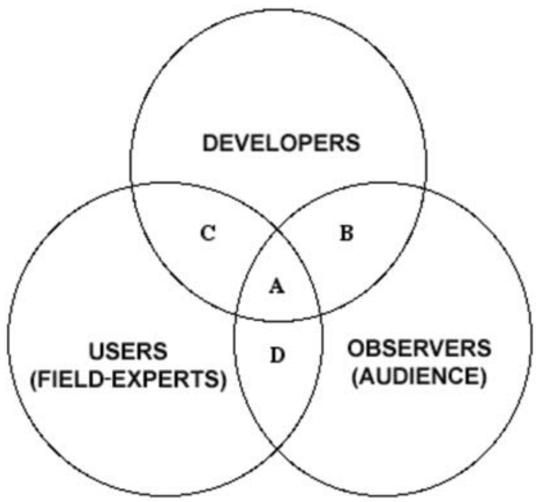

Fig. 1. The Interactive Multimedia Framework

A second setting describes the instance where only "A" is colored. Here, the development group possesses knowledge that transcends all three groups. This would encompass the single-developer instance: where one has understanding of MM-development, expertise on the subject matter and appreciates audience expectations.

In passing to more specific instances, a third interaction setting is where "C" and "D" only are highlighted (Fig. 2). This scenario implies that users/field-experts assume an active role in development (indicated by arrow), and commute between roles as user/audience. This accurately describes the development scenario of a live interactive audiovisual performance. Bi-directional relationships may be indicated by double-arrow (often omitted). In this manner, interaction between the parties involved may be described in a highly informative manner.

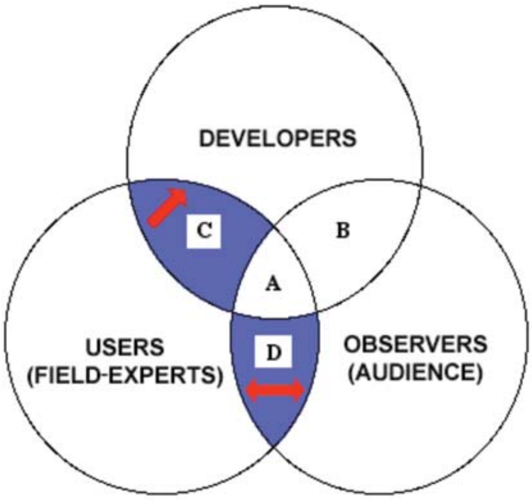

Fig. 2. Active Role for Users and Field-Experts

The fourth instance is of considerable significance, within the context of the current body of research. This describes the complex setting, for current MMS construction, combining users (field-experts), developers and various audience-types (both versed and unversed in the subject matter). Here, the proposed framework achieves a higher-order description of the relationship between the parties involved. Varying levels of expertise may be described: users/field-experts (aware of development constraints), developers (aware of audiovisual and artistic context), and their interrelations. In such respect, traditional frameworks have been found lacking. By design, they were never intended to describe diverse configurations, in terms of expertise, content and development for an MMS. Instead, they assign further workload to a single party (usually developer), to evaluate the quality of information. To interpret Carter [14] "…developers must understand the intended meaning of the data from the information providers, in order to evaluate its quality". This is a particularly pertinent issue when considering research-content. One notes that the developer may not always be able to perform this task, commonly due to a lack of expertise.

Clearly, one may observe a variation of scenarios emerging under this framework. For example, the users may be performers, striving to advance the boundaries of presentation requirements using multiple live and streamed input sources such as microphones, midi information, live video feeds and various types of sensors. Yet another scenario is where the MMS is developed by a user sub-group, via interaction with multimedia experts. Here, a novice (student) may develop the MMS with the aid of a field-expert, with courseware firmly in mind. Other intermediate scenarios are not precluded. Such a framework covers multi-varied requirements extensively. Additionally, at a higher level-of-abstraction, other frameworks may be emulated.

For the instance one may observe direct interaction between the multimedia system and the parties involved. Under the proposed framework, construction of the MMS is attributed to the developers. Also, information providers or information recipients may be involved in this task. This may be identified clearly under the new framework, by inclusion of sets "C" and "B", respectively. In fact, this is perceived as a redundancy of some frameworks, when information from the various parties is transferred to the MMS only through the developers.

A complicated case is described below, where secondary information providers are included. These may be represented as sub-groups within the user/field-experts region. Notice that information flow passes from information providers to developers (hence, into the MMS), in place for delivery to information recipients. Requirements, on the other hand, are directed from all groups to developers. An additional group may be introduced, that of sponsors, associated with the developers-group only (Fig. 3) displays the same relationships, through the new model representation. Arrows are utilized to denote direction of information (red) and requirements (black), for each individual group/sub-group. Notwithstanding associated complexity, this scenario is represented comprehensively through the new framework. This delineates clearly how information is processed and the requirements path. Notice here, that all groups are distinct, and their expertise does not overlap. In such an instance, representation of interaction between groups using alternative frameworks would require a significant amount of "additional" linking and explanation.

B. *Interactive Multimedia Model (iMM)*

The proposed model addresses the issue of content- and context-connectivity. This is achieved through a higher-level of abstraction in the representation of content. Indexed encapsulating frames are used, instead of considering individual media components and defining interrelationships between them. Each frame may contain one or more media elements. Static content includes text fields, buttons, icons, animations, sounds and images. Interactive content includes multiple live and even streamed input sources over the Internet, rendering this framework ideal for the development of Internet Art. At this abstracted level, interrelated content instances may appear within the same frame (where further connectivity may not be necessary), or across distinct frames. Recently, similar ideas have been cited in the literature, utilizing a complex rule-based approach [10].

To emulate connectivity across lower-level content (such as single-streams), frames containing only a single media-element may be employed. This approach poses multiple advantages, when compared to other cited in the literature [20]. First, low-level content connectivity is abstracted, as direct comparison is activated between two or more streams. In-turn, the volume of links is reduced. A second advantage is that individual media-components may be referenced separately, despite being organized within a more general structure. In addition, their ability to form further combinations is not reduced. This may be achieved separately, by being linked to other frames. Furthermore, a frame containing a title, and a set of streams provides immediate identification and categorization of content. This introduces an advanced indexing mechanism, consisting of a frame-index, media-indexes (contained within each frame), and a set-of-states (in which a frame and a set of media are active). Content linking is addressed using a "media-element to frame" referencing procedure. Each media-element may be employed as a user-interface construct. Fixed links may be programmed directly, or copied across other media-elements, reducing re-programming. Additionally, case-based links may be programmed locally (within frames of a single-MMS), unless external frame-access is necessary. Use of a knowledge-based approach (to dictate link-behavior) is supported within the model via scripting.

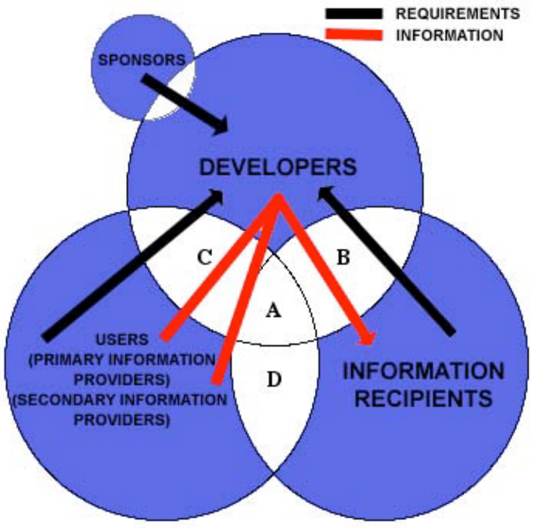

Fig. 3. Complex Interaction Scenarios

Mathematical representation is used to formally describe the proposed model. Various factors render the mathematical representation advantageous over other algorithmic approaches. From the computer science perspective, it is common to use mathematical constructs in system specification, automated verification and development [18]. Typical formal specification tools employ algebras, signatures and rules [27]. As a consequence, the set/functional representation chosen may be modeled directly through computer algorithms, to enable automated presentation-construction and verification. It is therefore important to use an appropriate model-representation, to aid in future system-development.

C. Formal Definition and Operations of iMM

Frame positioning, indexing and linking is implemented via an underlying linear structure. This structure, termed 't', is defined as:

$$\{\text{ordered finite set of states } t_m \mid m \in N^+\} \qquad (1)$$

This may be viewed as any appropriate linear indexed-structure, which offers discrete indexing for frames (from 1 to m, see Fig. 4).

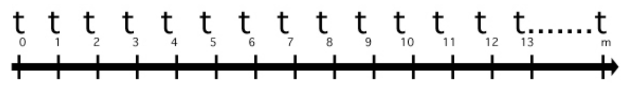

Fig. 4. Linear Frame Indexing (Timeline)

Fig. 5. Adobe Director Frame Indexing (Timeline)

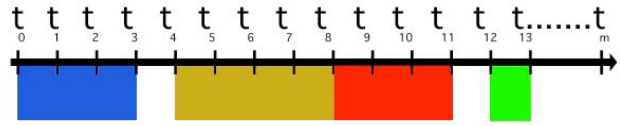

Fig. 6. Linear Representation of iMM

As shown in Fig. 5, the timeline used within the authoring environment of Adobe Director is a time-based example. Subsequently, frames spanning across this structure may be introduced. A complete MMS (a set of MM-frames, including a super-MMS) may be represented using the following construct:

$$\{\textit{finite set of frames } F_n \mid n \in N^+ \text{ and } n \leq m\} \qquad (2)$$

All frames span across the finite set of states 't', occupying continuous segments (Fig. 6). The condition on the subscript 'n' ensures that the number of frames may not exceed the maximum number of states. Each frame F_n is a *tuple* defined as:

$$F_n(t_{a,b}, \{S_i(t_{c,d}, l_r(k))\}) \mid a<b<m,\ a<c<d<b,\ i \in N^+,\ i \leq m,\ r \in N^+; \text{ set } k \in t_m + \{\varnothing\} \qquad (3)$$

'∅' denotes an empty link (undefined). The subscript indexes of the first term (a,b) indicate the span of a particular frame across 't'. Any frame, may constitute all states together, or be a single-state. The range of 'a' and 'b' is between 0 and m. As indicated above, a multimedia presentation consists of a finite set of such frames, each containing a set of media-components (text, image, sound, animation). All frame-components are described in the second parameter-term of F, included within the set notation '$\{S_i\}$'. Each component may be used as a linking-mechanism to other frames. Once a frame component is accessed, if no link is actioned, a transition is performed to the subsequent frame. Delay for user-action is imposed in the following manner: any frame-state (for example, last within a frame), may be programmed as a link to the start of the frame. This creates a loop, and does not involve "system-waiting", where the system must be instructed to idle for a given period. This is how one may accomplish user-interaction through system waiting.

Having accomplished frame indexing, frame-contents must be specified and linking performed. The second parameter-term of F is, itself, a finite set-of-tuples, that relates to the linking of individual media-components to other frames:

$$S_i(t_{c,d}, l_i(k) \mid i \in N^+,\ i \leq m,\ r \in N^+, \text{ set } k \in t_m + \{\varnothing\}. \qquad (4)$$

This tuple describes linking for a media-component contained within a frame F_n. The subscript indexes 'c' and 'd' indicate the length of each media component, limited only by the number of states. 'l_i' is the set of all links for the i^{th} media component, and 'k' is the link launched, upon activation of 'S'. If 'k' consists of a single link, it is followed by default. In the instance where two or more links are listed, external evaluation is necessary to determine which link to follow. For an empty link '∅', no action is taken. To describe transient behavior of multimedia frames, a set of discrete finite states is utilized, that a single-frame may occupy. These states may be mapped onto other conventions (such as time-based systems), via a mapping function. Transient links may be modeled appropriately over the range of programmed links for a specific frame-component. Here, only one link may be selectable at any one time. Various factors classify this model as highly flexible, proving its general utility. It supplies the indexing mechanism required to describe transitions between frames, forming a default linear structure. This is convenient, as modern authoring environments utilize the time-line metaphor. Therefore, they may be used in conjunction with the proposed model, for the creation of complex MMS.

Also, one may view this as a procedure to transform linear structures, to higher connectivity-orders, without re-organization. Normally, this would involve only the addition of links to other frames. The following schematic representation displays a characteristic case of two MMS, with internal and external connectivity (Fig. 7).

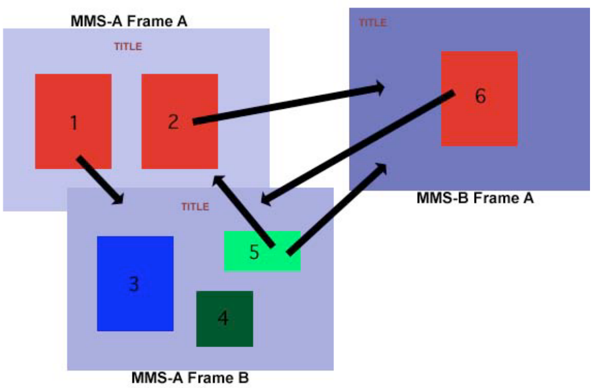

Fig. 7. Internal and External Connectivity for iMM

Here, MMS-A Frame A consists of a title and two components. This may be specified as:

$$F_1(t_{a,b},\{S_1(t_{a,b},l_1(c));S_2(t_{a,b},l_2(e))\}) \quad (5)$$

Component-A is connected internally to MMS-A, frame-B:

$$F_2(t_{c,d},\{S_3(t_{c,d},l_3(\emptyset));S_4(t_{c,d},l_4(\emptyset));S_5(t_{c,d},l_5(\{a,e\}))\}) \quad (6)$$

This connectivity is indicated through 'c', the starting point of F_2. Component-B to the external MMS-B, frame-A, is specified as: $F_3(t_{e,f},\{S_6(t_{e,f},l_3(c))\})$. External MMS may be linked under the proposed model, by inclusion within the set-of-frames referenced. This may require a finite extension to include additional frames. Frame-B contains a somewhat more complex linking scenario, indicated via the dual linking of component-E. This is a case-based scenario. Depending on certain conditions, a selection must be performed between frame-A, of either MMS-A or MMS-B. This is achieved by employing an external function, to determine k, under equation (3).

A further advantage of the proposed model (beyond content-connectivity) is its extendibility to accommodate content-based, user-interface construction. In this respect, two basic user-interface construction-types may be employed: media-elements themselves, and WIMP interface-constructs. From the perspective of an audio-interface, sound may be represented as a stream by displaying the beginning, any number of intermediate points, the end-point the graphical representation of the This specifies the functionality of the interface only, not their placement on-screen. Layout issues are determined externally, typically by field-experts/developer, without excluding the audience, in the case of rapid user-interface prototyping. In fact, one may extend the model to cover two or three-dimensional placement, by adding coordinate-fields and object-size. In general, user interface technologies are advancing continually. Therefore, it would be limiting to consider a fixed two- or three-dimensional approach, as this would be inappropriate for future interface development. One may quote examples, where there is departure from two-dimensional displays to advanced interaction-modes. Computer Aided Virtual Environments (CAVE), modern technologies utilizing Virtual Reality (VR) environments are encountered widely today. Other research avenues include sensualization-based and multi-modal interfaces [8], [3].

IV. Functionality & Examples

To give an example of how the proposed model complies with advanced navigation, a list of desirable navigation modes is provided, as described within draft ISO/DIS 14915-2. If appropriate to the task, the user should be able to perform the following actions: move backwards and forwards within a single-level, and upwards and downwards across levels; advance in larger steps, proceed to top-level, beginning- or end-of-structure, or a controlling-structure (such as, table-of-contents, index, or search facility). Having described typical MMS-level linking, frame linking indicates how the proposed model accommodates transition across internal sections.

The main development steps, in connection to the MMS herein, may be described as follows. First, field-experts organize the data into a suitable categorization. This may involve centralized or distributed-content. In the latter case, default MMS actions are planned in advance, to cover cases where externally accessible content is unavailable. Subsequently, content-connectivity is introduced across frames. This is a hierarchical iterative process that may be left partially complete at early development stages, to be built upon subsequently. In this fashion, a first version of the user-interface may be generated (with media-components positioned appropriately). Accordingly, developers/experts may extend upon the process at subsequent stages. Design and implementation of the multi-menu is completed for those characteristic frames constructed already. Connectivity is verified at this stage, and the multi-menu structure is embedded within frame-templates. Then, remaining content is inserted to fully linked frames, completing

MMS construction. In addition, connectivity may be extended through a range of mechanisms: cruise-control, knowledge-base access [36], or randomized algorithms and stratagems (as in game-theory).

A. Connectivity types

Principal organizational structures, that may be constructed using the above model include: *Single-node (type-A).* This may be expanded, across a single-state or multiple-states (as a single-frame). It supports a single-link, denoted either by a cyclic arrow (pointing to itself) or to another frame. *Linearly linked nodes (termed type-B)* consist of a combination of type-A nodes, linked with arrows in a line-structure. This structure supports linear, doubly linked linear, and combinations of the same. *Tree structure (type-C)* includes branching, in combination with B-type structures. *Enhanced tree/linear structure* (*type-D*) may comprise of type-A, -B and -C structures, with some direct links to non-adjacent nodes (categorized as sparse tree). *Graph structure (type-E)* contains graphs of various densities (including any of above combinations). *Transient dynamically adjusting graphs (type-F)* are used to describe complex structures that change their connectivity at runtime, for some or all nodes.

Advanced audio manipulation implementations may utilize all previously described connectivity types. In addition, one may incorporate dynamically changing streams in separate frames, complete with rules and code describing interaction responses. A typical example includes the use of appropriately positioned microphones, and distance sensors, which may be used to detect the users' location in space. The system can be programmed to adjust appropriately the direction of sound and video as if the presentation follows the user.

It is informative to compare the flexibility of the proposed model, to other examples cited in the literature. One may commence with a formal model, using finite-state automata [19], from which it becomes apparent that extensive programmability is required for relatively straightforward problems. Additionally, one may consider the "MORENA" (Multimedia Organization Employing a Network Approach) model, under a hypermedia authoring and browsing context [21]. "MORENA" is essentially an adaptation of the "Trellis" model [9]. Both utilize a Petri-net based structure. Typical disadvantages include the inelegant "waiting" technique imposed, which effectively instructs the system to "pause" for a given period. The currently proposed model addresses pausing via loops, allowing user-determined time-of-interaction within a frame. Alternatively (when automated presentations are considered), conditional linking is utilized to detect when all streams within a frame have terminated, instructing the MMS to proceed to a subsequent frame.

V. Conclusions

As shown above, the iterative development cycle supported by the framework and the model, furnished MMS-development with additional features. These relate directly to particular user-requirements for content-access under each audiovisual case study. As existing multimedia authoring environments may be utilised, flexibility is also extended in the development of the user-interface and interaction processes, in contrast to propriety systems, where such changes usually require re-programming of the user-interface [15].

User, content-expert and system interaction can be modeled using the proposed framework for most multimedia uses and applications, even under complex interactive audiovisual environments, or other applications where sound is the dominant modality. In addition, combination of the framework with the model described above, allows advanced interactive capabilities to be formally described and realized. The iMMS multimedia model may again be utilized when sound is the primary driver of the presentation.

Exporting and linking content (locally under a single computer, or over the World-Wide-Web) forming higher-order presentation structures is also supported, emphasizing the fact that MMEs facilitate expandability and content re-use. This practically supports the extension of existing systems internally and externally, incorporating other content-encapsulating formats into MMSs (other audiovisual sources, HTML-pages through external browsers, remotely-located files, multimedia databases using SQL queries over internet connections). Dynamic frame-linking and on-demand 'request-for-streams' have been found not to overload available modest hardware resource provision (typical memory 64Mb RAM). As a result, these MMS demonstrate the key advantage of reliability, over competing presentation software and technologies. All such MMS deployed may be delivered both in standalone form, or over Internet/WWW communication channels [4]. This particular ability in providing a single implementation, exportable across various systems without loss of functionality, has proved a most useful feature for various content-domains and we hope this to be the case for interactive sound manipulation multimedia environments.

References

[1] M. L. a. C. Barry, *Techniques and Methodologies for Multimedia Systems Development: A Survey of Industrial Practice*, in N. L. e. a. Russo, ed., *Realigning Research and Practice in Information Systems Development: The Social and Organizational Perspective, Proceedings of International Federation for Information Processing (IFIP) Working Group 8.2 Conference*, Kluwer Academic Publishers, Boston, Boise, Idaho, USA, 2001, pp. 77-86.

[2] D. Carr, *Interaction Object Graphs: An Executable Graphical Notation for Specifying User Interfaces. Formal Methods for Computer-Human Interaction*, Springer-Verlag, 1997.

[3] F. P. a. A. Celentano, *A Multimodal Approach for Orientation and Navigation in 3D Scenes, Report n. CS-99-25*, Dept. of Computer Science, Ca' Foscari University, 1999.

[4] I. Deliyannis, *Interactive Multimedia Systems for Science and Rheology*, *Computer Science*, Ph.D Thesis, University of Wales, Swansea, 2002, pp. 285.

[5] D. G. a. D. A. Duce, *The Notion of Trajectory in Graphical User Interfaces*, in J. C. T. M.D.Harrison, ed., *Design, Specification and Verification of Interactive Systems '97 DSV-IS'97*, SpringerWienNewYork, Granada, Spain, 1997.

[6] A. J. E. Scanlon, J. Barnard, J. Thompson, J. Calder, *Evaluating information and communication technologies for learning*, Educational Technology & Society, 3.

[7] W. Emmerich, *Distributed & Object-Oriented Software Development*, John Wiley & Sons, 2000.

[8] B. M. a. S. Feiner, *Future multimedia user interfaces*, Multimedia Systems (1996), pp. 250-268.

[9] P. D. S. a. R. Furuta, *Petri-net-based hypertext: Document structure with browsing semantics*, ACM Transactions on Information Systems, 7 (1989), pp. 3-29.

[10] J. P. T. M. Geurts, J. R. v. Ossenbruggen and H. L. Hardman, *Application-specific constraints for multimedia presentation generation*, *Eighth International Conference on Multimedia Modeling (MMM01)*, Amsterdam, Netherlands, 2001.

[11] F. M. N. a. H. L. Hardman, *Denotative and connotative semantics in hypermedia: Proposal for a semiotic-aware architecture*, CENTRUM VOOR WISKUNDE EN INFORMATICA (CWI) (2002).

[12] F. M. N. a. H. L. Hardman, *Towards a syntax for multimedia semantics*, CENTRUM VOOR WISKUNDE EN INFORMATICA (CWI) (2002).

[13] D. H. a. H. R. Hartson, *User Interface Development: Ensuring Usability through Product and Process*, John Wiley & Sons, New York, 1993.

[14] W. H. I. Heath, R. Crowder, G. Wills, J. Ballantyne, *Towards a New Authoring Methodology for Large-Scale Hypermedia Applications*, Multimedia Systems and Applications, 12 (2000).

[15] M. F. W. Ioannis Deliyannis, *WWW Delivery of Graph-Based, Multi-Level Multimedia Systems: Interaction over Scientific, Industrial and Educational Data*, *IADIS Int. Conf. WWW/Internet 2002*, Lisbon Portugal, 2002, pp. 607-612.

[16] M. J. H. a. Y. Jian, *Multimedia Description Framework (MDF) for Content Description of Audio/Video Documents*, *4th Int. Conf. Digital Libraries*, 1999.

[17] V. V. Kamat, *Framework for Development of Course Material for Teaching Computer Graphics*, *Graphics and Visualisation Education Workshop (GVE)*, Coimbra, Portugal, 1999.

[18] O. R. Martinez, *Design Dependencies within the Automatic Generation of Hypermedia Presentations*, Technical University of Catalonia, 2002.

[19] K. Morton, *A formal Model for Multimedia Presentation*, in W. H. R. I Damper, J. W. Richards, ed., *Multimedia Technologies and Future Applications*, Pentech Press, London 1994, pp. 191-204.

[20] R. A. B. a. D. Mossé, *The Capoeira Model for Distributed and Reconfigurable Multimedia Systems*, Computer Science Department, Universidade Federal Fluminense, Brazil, 1996.

[21] R. B. a. D. Mossé, *The MORENA Model for Hypermedia Authoring and Browsing*, IEEE Conf. Mult. Comp. Systems (1995).

[22] R. M. a. T. Ottmann, *The "Authoring on the Fly" system for automated recording and replay of (tele)presentations*, Multimedia Systems (2000), pp. 158-176.

[23] T. G. P. P. da Silva, N. W. Paton, *Generating User Interface Code in a Model-Based User Interface Development Environment*, in e. a. V. di Gesu, ed., *Advanced Visual Interfaces*, ACM Press, 2000.

[24] R. B. a. P. Palanque, *Petri Net Based Design of User-Driven Interfaces Using the Interactive Cooperative Objects Formalism*, in F. Paternò, ed., *Interactive Systems: Design, Specification, and Verification DSV-IS'94*, Springer-Verlag, 1994.

[25] A. R. Puerta, *A Model-Based Interface Development Environment*, IEEE Software, 14 (1997).

[26] J. E. a. A. Puerta, *Adaptation in Automated User-Interface Design*, *Intelligent User Interfaces*, ACM Press, New Orleans LA, 2000.

[27] D. L. L. Rees, K. Stephenson and J. V. Tucker, *Towards an Algebraic Structure of Interface Definition Languages and Architectures*, Computer Science Department, University of Wales Swansea, Swansea, 2000.

[28] C. S. a. T. P. Rout, *A Structured Methodology for Multimedia Product and Systems Development*, *ASCILITE '98*, Sydney, 1998.

[29] D. Sannella, *The Common Framework Initiative for Algebraic Specification and Development of Software: Recent Progress*, *WADT*, 2001.

[30] B. Schmitt, *Shockwave Studio: Designing Multimedia for the Web*, O'Reilly and Associates, Inc, 1997.

[31] P. P. d. Silva, *User Interface Declarative Models and Development Environments: A Survey. In Interactive Systems: Design, Specification and Verification*, *7th International Workshop on Design, Specification and Verification of Interactive Systems*, Springer-Verlag, Limerick, Ireland, 2000.

[32] A. G. Sutcliffe, *User-Centered Design for Multimedia Applications*, ICMCS, 1 (1999).

[33] P. J. B. T. Griffiths, N. W. Paton, J. McKirdy, J. B. Kennedy, P. D. Gray, R. Cooper, C. A. Goble, P. P. da Silva, *Teallach: a model-based user interface development environment for object databases*, Interacting with Computers, 14 (2001), pp. 31-68.

[34] F. Tamburini, *A multimedia framework for second language teaching in self-access environments*, Computers & Education (1999), pp. 137-149.

[35] Underdahl, *Macromedia Director MX Bible*, John Wiley, 2004.

[36] J. Vanderdonckt, *Advice-Giving Systems for Selecting Interaction Objects*, *User Interfaces to Data Intensive Systems*, Edinburgh, Scotland, 1999.

[37] N. O. B. a. S. Verjans, *Information Mapping Knowledge-Based Support For User Interface Design*, *CHI'95 Workshop on Knowledge-Based Support for the User Interface Design Process*, 1995.

[38] J. S. Willans, *Integrating behavioural design into the virtual environment development process*, *Department of Computer Science*, University of York, 2001.

[39] www.anark.com, *Anark Corporation Website*, 2006.