11 SPATIOTEMPORAL SPECIFICATION & VERIFICATION OF MULTIMEDIA SCENARIOS

I. Kostalas, T.Sellis, M. Vazirgiannis

Abstract: A Multimedia Application (MAP) consists of a set of media objects ordered in the spatial and temporal domains. In this paper we present an authoring & verification methodology for MAP documents development. We capitalize on a solid theoretical model for spatiotemporal compositions. The tool may be used both for prototyping and verification of multimedia presentations or spatiotemporal compositions in general. Regarding the authoring phase, emphasis was put on the flexible definition of spatial and temporal relationships of the participating entities. The verification procedures are supported by multiple tools allowing designers to preview their applications, in various ways: spatial layouts of the application window, temporal layouts, indicating the temporal duration and relationships among the participating objects and animation of the application.

11.1 INTRODUCTION

A Multimedia Application (MAP) involves a variety of individual media objects, called *actors*, presented according to the MAP scenario. The term *scenario* covers two areas: i. the spatial and temporal ordering of actors within the application context and the relationships among them and ii. the way that the user will interact with the application as well as how the application will treat application or system events.

Another relevant issue is that the actors participating in a MAP, are transformed either spatially and/or temporally in order to be presented according to the author's requirements. For instance, we may want to present part of a video clip faster or slower at a bigger or smaller window.

The authoring procedure for complex MAPs, that involve a large number of actors, may be a very complicated task, having in mind the large set of possible events that may be encountered in the application context, the number of actors and relationships as well as the various potential combinations of these factors.

A MAP specification should describe both the temporal and spatial ordering of actors in the context of the MAP. In the past, the term "synchronization" has been widely used to describe the temporal ordering of actors in a MAP. The spatial ordering issues (i.e. absolute position and spatial relationships among actors) have not adequately been addressed up to now. We claim that the term synchronization is poor for MAPs. Instead we propose the term "*composition*" to represent both the temporal and the spatial composition of actors.

The potential high complexity of MAPs, results in substantial effort required for design and development. In real-life applications, usually only programmers are able to develop MAP scenarios since current authoring tools provide rather low level specification languages. Moreover, these languages are inadequate for the complete description of the scenario aspects mentioned before. Therefore, the lack of an integrated mechanism for high level complete specification of a MAP scenario arises as a main issue. Moreover, in current MAP development tools the MAP script is mixed with the application content (actors). This prevents the explicit reusability of scenarios in other MAPs with different content but similar functionality.

Another important aspect is the verification of MAP scenarios during MAP development. The term verification in this context implies the various procedures that will allow the author to review the result of their authoring effort prior to the production and execution of the MAP. This will enable revisiting the application and adjusting the spatiotemporal specification in order to align to the document style the authors has in mind or to fix specification errors that result in spatial and or temporal exceptions.

In this paper we present an authoring & verification methodology for MAP documents development supported by a full implementation. The MAP design is based on a theoretical model for spatiotemporal compositions in the context of multimedia presentations (Vazirgiannis et al, 1998). The tool may be used both for *prototyping* and *verification* of multimedia presentations or spatiotemporal compositions in general.

Regarding the authoring phase, emphasis was put on the flexible definition of spatial and temporal relationships of the participating entities. In other words, the authoring phase consists mainly of declarative specifications of the spatial and temporal ordering of participating multimedia objects based on their spatial and temporal relationships.

The *verification* procedures are supported by multiple tools allowing designers to preview their applications, in various ways: *spatial layouts* of the application window, *temporal layout* of parts (or the whole application), indicating the temporal duration and relationships among the participating objects and *animation (rendering)* of the application (i.e. what would the execution of the application like) in three modes (real time, manual and snapshots of the application at regular temporal intervals).

The paper is organized as follows. In the next section we present related research and industrial products in the area of multimedia application modeling and synchronization. In section three, we briefly present the theoretical model (Vazirgiannis et al, 1998) that served as the basis for the authoring and verification tool, which is presented in section four. There, we present the authoring procedures as spatiotemporal specification of multimedia objects and the verification tools provided by the system. In the last section, we conclude by presenting our contributions and discussing on further research and extensions of the tool.

11.2 RELATED WORK AND BACKGROUND

The specification of a multimedia application is essentially the definition of the composition of the participating media objects in space and time along with the appropriate transformations of the media object.

In the literature there is a confusion among the concepts of *synchronization* and *temporal specification*. They are used interchangeably. Existing multimedia document standards (MHEG, HyTime) propose Object-Oriented approaches for modeling the structure and behavior of multimedia documents. The proposed standards, however, do not provide means for the complete specification of spatial and temporal composition of actors. Moreover, the issues of storage, retrieval, execution and sharing of application scripts are not addressed, while event modeling and composition schemes are inadequate to cover the requirements for the variety of events that may occur in an MAP. Finally, there are no large scale implementations of these standards to prove their credibility and usability.

Regarding commercial authoring tools (Assymetrix/ToolBook, Macromedia/MacroMind Director) for MAP development, they adopt mostly Object Oriented authoring models, providing high-level script languages. These tools do not adequately fulfill requirements regarding:

- Declarative Spatiotemporal composition schemes and modeling of the actor transformations
- Event representation and composition: only a limited repertoire of events is supported, while no composition of events is feasible. They mostly manipulate events related to mouse and to the actors' state (start, stop, active, idle, etc.).
- Database support for the application scripts (multimedia data and scripts are bundled together in most cases)

Many existing models for temporal composition of multimedia objects in the framework of a multimedia application are based on Allen's relations (Allen, 1983). However, these relations are not suitable for composition description since the are descriptive (they do not reflect causal dependency between intervals), they depend on interval durations and they may lead to temporal inconsistency. More specifically, the problems that arise when trying to use these relations are the following (Duda et al, 1995):

The relations are designed to express relationships between intervals of fixed duration. In the case of multimedia applications it is required that a relationship holds independently from the duration of the related object (i.e. the relationship does not change when the duration changes). Their descriptive character does not convey the cause and the result in a relationship.

Other models for multimedia composition representation may be classified in two categories: point-based and interval-based. In point-based models, the elementary units are points in time and space. Each event in the model has its associated time point. The time points arranged according to some relations such as "precede", "simultaneous" or "after" form complex multimedia presentations. An example of the point-based approach is timeline. Interval based models consider elementary media entities as time intervals ordered according to some relations. Existing models are mainly based on the relations defined by Allen for expressing the knowledge about time.

An interesting mechanism for temporal composition is presented in (Duda et al, 1995). This work presents a model that takes into account the semantics of temporal relationships between objects. A set of operators is defined expressing the causal relations between intervals. In (Hirzala et al, 1995), a temporal model for interactive scenarios is presented. This model is based on the timeline approach and provides the primitives for specification of synchronous and asynchronous interactive and temporal multimedia compositions. The timeline approach is extended to a tree of timelines. Each branch of timelines represents the different scenarios that may be selected by the user.

Other approaches use Allen's relations to specify a multimedia database schema. In (Little et al, 1993) an OCPN (Object Composition Petri Nets) model equivalent to Allen relationships is proposed. This approach, though, does not take into account the possible unknown durations of intervals. Thus, in order to prepare an instantiated presentation the tree of interval relations must be traversed to get deadlines to be used in the presentation schedule. In (Iino et. al, 1994) a model for spatiotemporal multimedia presentations is presented. The temporal composition is handled in terms of Allen relationships whereas spatial aspects are treated in terms of a set of operators for binary and unary operations.

The model lacks the following features: i. There is no indication of the temporal causal relationships (i.e. what are the semantics of the temporal relationships between the intervals corresponding to multimedia objects); ii. The spatial synchronization essentially addresses only two topological relationships: overlap and meet, giving no representation means of the directional relationships between the objects (i.e. Object A is to the right of object B) and the distance information (i.e. object A is 10 cm away from object B); iii. The modeling formalism is rather oriented towards execution and rendering of the application rather than authoring. This means that the specifications are closer to the logic of the presentation scheduling rather that to the author's declarative requirements.

There are other approaches based on interval temporal logics (King, 1994). Although such formalisms have a solid mathematical background, the specification of multimedia presentations is awkward since the specification does not correspond explicitly to the author's perception of the multimedia composition.

Related work has been published in (Karmouch et al, 1996). Although it copes with multimedia objects, it models a smaller part of the MAPs that relate to multimedia documents and not to multimedia applications. Thus, taking for granted that the document will be based on textual resources, the model tries to make an interactive multimedia book containing some form of multimedia objects like images, sound and video. The book is divided in chapters and the screen layout is similar to the one of word processors, along with their temporal information. The temporal relationships are taken into account, but not the spatial ones since it is assumed that they are solved depending on the text flow on the page. An interesting survey on authoring models and approaches may be found in (Jourdan et al, 1998).

11.3 THE COMPOSITION MODEL

The "action" concept is related to the presentation of actors (multimedia objects) participating in MAPS. A multimedia application specification should describe both temporal and spatial ordering of actors in the context of the application. The spatial ordering (i.e. absolute positioning and spatial relationships among actors) issues have not been adequately addressed. We claim that the term "*synchronization*" is poor for multimedia applications and, instead, we propose the term "*composition*" to represent both the temporal and the spatial ordering of actors. In the current modeling effort we made a fundamental assumption by considering objects that appear and after some time they disappear from the presentation context without changing their position during their lifetime. Thus, here we do not consider motion. We also assume that the spatial features of the media objects are defined by the rectangle that bounds them.

11.3.1. Temporal Relationships

The topic of relations between temporal intervals has initially been addressed in (Allen, 1983). We exploit the temporal composition scheme as defined in (Vazirgiannis et al, 1998). We briefly introduce that scheme for representing the temporal composition of multimedia objects that also captures the causality of the temporal relationships. In that scheme the start & end points of a multimedia instance are exploited as events. The end of a multimedia object presentation is either *natural* (i.e. when the media objects finishes its execution) or *forced* (i.e. when an event explicitly stops the execution of a multimedia object). Moreover, the well-known *pause* (temporarily stop of execution) and *resume* procedures (start the execution from the point where the pause operation took place) are also taken into account.

An important concept is the temporal instance: we consider it as an arbitrary temporal measurement that is relative to some reference point (i.e. the application temporal starting point in our case, hereafter T).

Based on the above descriptions we define the following operators attached to the corresponding events:

Definition: Let A be a multimedia instance, then A> represents the start of the multimedia instance, A< the natural end of the instance, A! the forced stop, A|| the pause and A|> the resume actions.

Definition: Let A, B two multimedia instances, then the expression: Aop1 t Bop2 represents all the temporal relationships between the two multimedia instances, where op1∈{>,<,||,|>} and op2∈{>,!,||,|>} and t is a vacant temporal interval.

Definition : Let A be a multimedia instance, we define as t_{Aop} the temporal instances corresponding to the events Aop, where op ∈ {>,<,!,||,|>}.

Definition : Let A be multimedia instance, we define as d_A the temporal duration of the multimedia instance A.

11.3.2 Spatial Relationships

Another aspect of composition of multimedia objects in MAPs is related to the spatial layout of the application, i.e. the spatial arrangement and relationships of the participating objects. The spatial composition aims at representing three aspects:

- The topological relationships between the objects (*disjoint*, *meet*, *overlap*, etc.)
- The directional relationships between the objects (*left*, *right*, *above*, *above-left*, etc.)
- The distance characteristics between the objects (*outside* 5cm, *inside* 2cm, etc.)

We will exploit the scheme presented in (Vazirgiannis et al, 1998]. We will briefly present this scheme for clarity reasons.

Regarding directional relationships, there is a complete set of relationships defined in (Papadias et al, 1995). This set of 169 (13^2) relationships $R_{i\,j}$ (i = 1, ..., 13 and j = 1, ..., 13) arises from the exhaustive combination of the relationships defined in (Allen, 1983) regarding relationships between temporal intervals. It is evident that in the case of MAPs these relationships have to be grouped in sets of relationships to assist the authoring procedure.

11.3.3 Spatiotemporal composition model

In this section we present briefly the theoretical foundations we use for our spatiotemporal verification framework. We exploit the model introduced in (Vazirgiannis et al, 1998) aiming at representation of the spatiotemporal composition of media objects in the context of a MAP. The model we propose will translate spatiotemporal relationships among multimedia objects into minimal and uniform expressions, as imposed by the requirements for correct and complete representations.

For uniformity reasons, we define an object named Θ, that corresponds to the spatial and temporal start of the application (i.e. upper left corner of the application window and the temporal start of the application). Another assumption we make is that the objects that are included in the composition include their Spatiotemporal presentation characteristics (i.e. size, temporal duration, etc.)

Definition : Assuming two spatial objects A, B, we define the generalized spatial relationship between these objects as: $S\_R = (r_{ij}, v_i, v_j, x, y)$ where r_{ij} is the identifier of the topological-directional relationship between A and B , v_i, v_j are the closest vertices of A and B respectively (as defined in Vaz98) and x, y are the horizontal and vertical distances between v_i, vj.

Hereafter, we define a generalized operator expression to cover the spatial and temporal relationships between objects in the context of a multimedia application. It is important to stress the fact that in some cases we do not need to model a relationship between two objects but to declare the spatial and / or temporal position of an object relative to the application spatial and temporal start point Θ (i.e. object A to appear at the spatial coordinates (110, 200) on the 10th second of the application).

Definition: We define a composite Spatiotemporal operator that represents absolute spatial/temporal coordinates or Spatiotemporal relationships between objects in the application: `ST_R(sp_rel, temp_rel)`, where `sp_rel` is a spatial relationship (`S_R`) as defined above, and `temp_rel` is a temporal relationship.

The Spatiotemporal composition of a multimedia application consists of several independent fundamental compositions. The term "independent" implies that actors participating in them are not related explicitly (either spatially or temporally) except from their implicit relationship to the start point Θ. Thus, all compositions are explicitly related to Θ. We call these compositions *composition_tuples* and these include spatially and/or temporally related objects.

Definition: We define the composition_tuple in the context of a multimedia application as : `composition_tuple = Ai [{ ST_R Aj}]`, where

`Ai`, `Aj` are objects participating in the application, `ST_R` is a Spatiotemporal relationship (as defined above).

Definition: We define the composition of multimedia objects in the context of multimedia applications as a set of `composition_tuples`: `composition = Ci {,Cj}`, where `Ci`, `Cj` are `composition_tuples`.

The EBNF definition of the Spatiotemporal composition based on the above definition follows:

```
composition ==:
        composition_tuple{[,composition_tuple]}
composition_tuple ==:
        Θ {[spatio_temporal_relationship actor |
        composition]}
spatio_temporal_relationship ==:
    "[("[spatial_operator | spatial_instance")" ,
    "("temporal_operator | temporal_instance")]"
temporal_operator ==: Θ | t_event t_interval
                      TAC_operation
t_event ==: ">" | "<" | "!" | "|>" | "||"
spatial_operator ==: (rij, Vi, Vj, x, y)
x ==: INTEGER
y ==: INTEGER
```

where `rij` denotes a topological- directional relationship between two objects and `vi`, `vj` denote the closest vertices of the two objects (see definition above).

11.3.4 A Sample Multimedia Composition

In this section we describe a MAP corresponding to TV news clip in terms of spatio-temporal relationships as defined above. The high level scenario of the application is the following:

*"The **News** clip starts with presentation of image A (located at point 50,50 relatively to the application origin Θ). At the same time a background music E starts. 10 sec after a video clip B starts. It appears to the right side (18cm) and below the upper side of A (12cm). Just after the end of B, another MAP starts. This MAP(**Fashion_clip**) is related to fashion. The Fashion_clip consists of a video clip C that showing the highlights of a fashion show and appears 7cm below (and left aligned to) the position of B. 3 sec after the start of C, a text logo D (the designer's logo) appears inside C, 8cm above the bottom side of C, aligned to the right side. D will remain for 4 sec on the screen. Meanwhile, at the 10th sec of the News clip, the TV channel logo (F) appears at the bottom-left corner of the application window. F disappears after 3 sec. The application ends when music background E ends."*

The spatial composition (screen layout) appears in **Figure 11.1**, while the temporal one appears in Figure 11.2. The objects to be included in a composition tuple of a MAP are those that are spatially and/or temporally related. In our example (News clip), A and B and Fashion clip should be in the same composition tuple since A relates to B, B relates to Fashion clip. On the other hand, F is not related to any other object, nor spatially neither temporally, so it composes a different tuple.

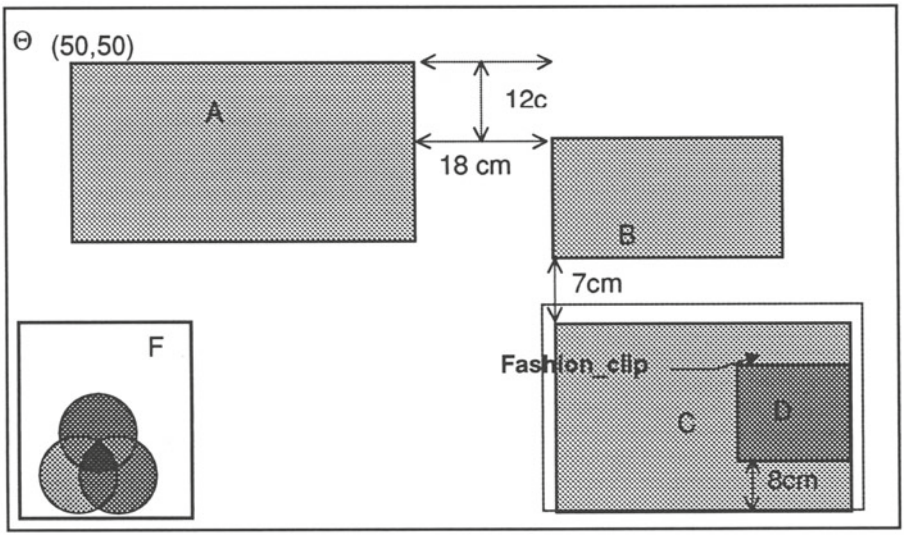


Figure 11.1: The spatial composition of the News MAP

The above spatial and temporal specifications defined by the author in a high level GUI are transformed into the following representation according to the model primitives defined:

```
// News Clip
    composition = {r1, r2}
    r1 =  Θ[(_, _, _, _,_), (>0>)]
          E [(_, _, _, _,_), (<0!)]
          News
    r2 =  Θ[(r1_1, _,v2, 5, 5), (>0>)]
          A [(r11_13, v3, v2,18,12),(>10>)]
          B [(r13_6, v1, v2, 0,-7), (>0>)]
          Fashion_clip
    r3 =  Θ [(_, _, v1, 0, 300), (>10>)]
          F
// Fashion clip
    composition = {r4}
    r4 =   Θ[(_, _,v2,0,0), (>0>)]
          C [(r9_10, v4, v4, 0, 8), (>3>)]
```

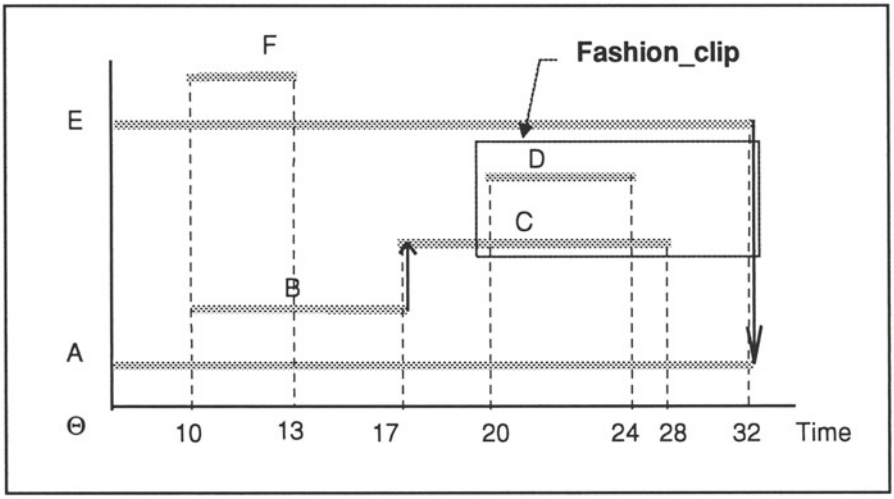


Figure 11.2: The temporal composition of the News MAP

11. 4. AUTHORING SPATIOTEMPORAL COMPOSITIONS FOR MAP DOCUMENTS

The authoring methodology we present hereafter is based on the model introduced in the previous section. A MAP document contains objects composed in space and time according to a set of spatial and temporal relationships. The scenario is build in a stepwise procedure. The first step is the specification of the objects that participate in the application along with their spatial and temporal transformations (i.e. which spatial and/or temporal part of the object will participate in the MAP, under what spatial/temporal scaling etc.). The second step is to define the actors' spatial and temporal position in the MAP document in terms of absolute or relative coordinates. The authoring tool transforms these specifications and produces the spatiotemporal scenario, i.e. when, where and for how long each object will be presented.

We distinguish the actors in four main categories: text, sound, image and video. We assume that each object (except sound) has some spatial extent so it can be represented by a rectangle in which the image text/video information is presented. On the other hand, sound objects have only temporal aspects.

11.4.1. Authoring environment

The authoring interface supports to a great extend visual definition of the multimedia scenario. Conceptually the scenario specification should start at the temporal start of the MAP. Thus, the authoring procedure starts at the beginning of the application (time = 0) and specifies the participating actors in a more or less increasing temporal order.

Composition Specifications. Initially the author(s) may insert or remove media objects from the actors list. Each actor object is further defined by assigning values to its spatial and temporal attributes.

Each actor includes its identification data such as name, media type and media file corresponding to the actor, the temporal and spatial coordinates of the object in the application context. The name of an actor must be unique name in the MAP context. Though, the tools allows the definition of different actors based on the same media object. The data that the user enters (*external*) are transformed into absolute coordinates (*internal*) in order to produce the scenario. The external data mostly relate the spatial and temporal coordinates of the new object to those of other objects, previously defined.

The methodology supports incremental authoring by adding new actors to the existing spatiotemporal composition and relating it to them spatially and/or temporally.

Temporal attributes specification. The *temporal attributes* of the object include its temporal start and end points (see
Figure **11.3**). The author may define the beginning of an object in relation to another object that is already, or will be, active. Thus, the starting time be relate to the temporal application origin **T**, or to the start/end point of another object (in this case, we do not distinguish between the *pause* and the *stop* events or between the *start* and the *resume* event). The same applies to the definition of the *end* time of the object as long as the object does not have an internal (natural) duration (like video and sound do).

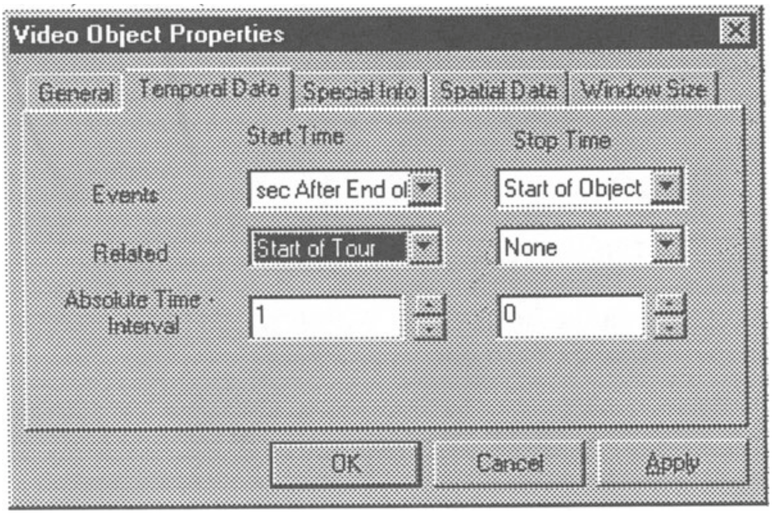


Figure 11.3: The Properties of an object to participate in the MAP (temporal data).

If the object has a natural length, then the user can either let it end naturally or force it to end before its natural end. That event can be a stop or a pause event. As mentioned before, the author may define the temporal data of the actor in question in relation to the starting/ending point of another actor. We have to distinguish here between the events *pause, restart* that are also included in our design. The pause event may be considered as a temporary *stop* event, whereas a *restart* event may be considered as a start event. This may be adjusted by the author using the property sheet appearing in Figure 11.4.

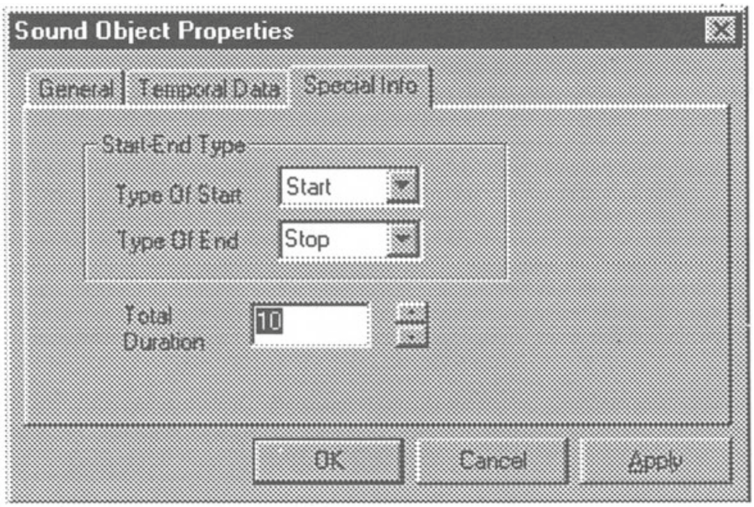


Figure 11.4: Specification of the start and stop events of an actor

Spatial attributes specification The issue of spatial features of actors in MAPs has been rather under-addressed in the multimedia literature. The authoring tools enable specification of the spatial features of an actor, either as absolute coordinates or in relation to other objects. We assume that each object is bounded by a rectangle (whose dimensions may be changed by the author) and the author may define its position. The actor's upper left corner is related either to the origin of the MAP (T) or to any vertice (Upper Left, Upper Right, Lower Left, Lower Right) of any other actor already defined. It is important to stress that it is possible to relate the actor under concern to different actors as regards the X and Y axis. In the case of absolute coordinates, we define the position of the upper left corner of the actor related to the top left corner of the application window (see Figure 11.5).

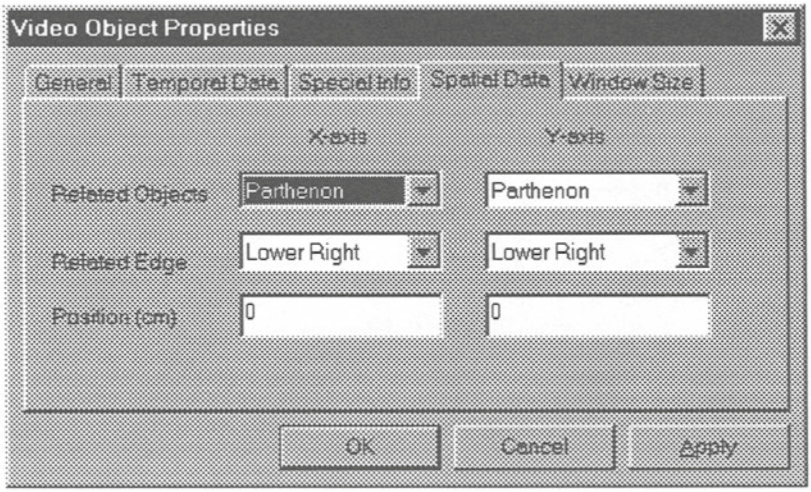


Figure 11.5: The spatial properties of an actor.

11.5. VERIFICATION OF MAP DOCUMENTS

During the authoring procedure it is probable/anticipated that the author(s) might want to query the scenario, especially if it is extended and complicated in terms of interactions and composed presentation actions. This can be helpful in order to:

- Inform the author(s) about the underlying spatiotemporal constraints of the scenario,
- Modify the scenario and, in this way, create, correct or improve the scenario in an incremental way.

In this last case, the author, depending on the answer of the query, can do the modification on the scenario, currently under creation. The queries may be classified into the following categories:

- *Point queries*, which are useful when the author is interested in finding relationships between events. For instance: "Does actor Starting Logo start before Tour video?" or "which objects appear at the position (50,50) before the 5th second of the MAP"?
- *Relationship queries*: In this category we are interested in relationships among the temporal intervals that represent the presentations of actors or alternatively the relationships between the spatial extents of the actors. Such a queries would be: "Are the presentations of objects "tour_video" and "agora"simultaneous?" or "Is "tour_video" overlapping with "agora"?".
- *Layout queries*. They result in graphical representations, depicting the spatial and temporal relationships among presented objects. Such kind of

queries is a necessity that is recognized among the authoring community (Vazirgiannis et al, 1998]. An example can be: "Show the temporal layout of the IMD between the 2nd and 10th second of the presentation".

It is evident that handling such queries can enhance the flexibility and quality of the authoring procedure through verification of the MAP under development. In the sequel we present the tools and methodologies we have developed to tackle the above set of requirements.

A multimedia scenario may be very complicated if we consider the multitude of objects and the spatiotemporal relationships among them. Thus, it is essential for a consistent MAP development procedure to allow the preview of several aspects of the application previous to the final implementation. In this section we introduce a set of techniques to verify a MAP document during authoring and prior to its execution. The verification is related to temporal & spatial aspects and may fulfill requirements such as viewing MAP snapshots at any temporal point, finding out the temporal relationships between actors, viewing an animation of the MAP, etc.

11.5.1. Temporal layout tool

The first verification tool is the *temporal layout* that displays, in a graphical form, the temporal order and the duration of the actors (see **Figure 11.6**). This facility gives an overview of the temporal configuration of the MAP and, moreover, provides support to queries of the type: "which objects are active at a specific time?" or "which objects are active at a specific period of time?" (i.e. during the temporal interval in which another object is active). The temporal layout may refer to a part of the temporal duration of the MAP (e.g. from the 5^{th} to the 20^{th} second of the MAP) or to its total duration. The list of objects (in absolute temporal coordinates) is sorted according to their starting time. This list is exploited in the rest of the verification tools, more specifically in the execution table and in the spatial layout tool that are described hereafter.

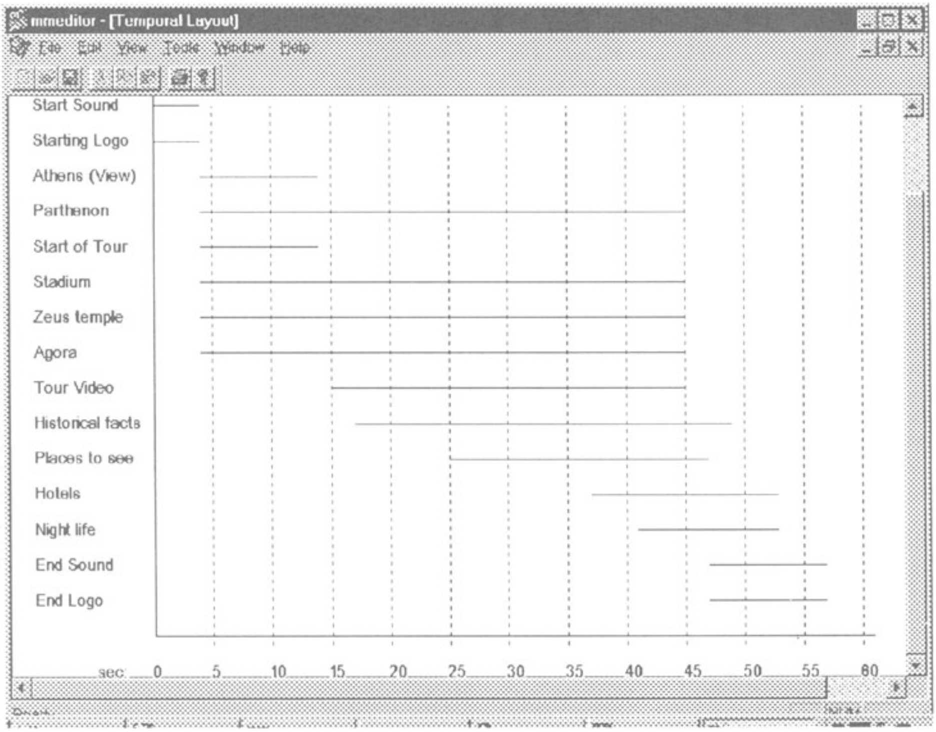


Figure 11.6: The temporal layout of the MAP

11.5.2. Spatial Layout

The term "spatial layout" implies the appearance of the MAP application window, conveying information about the position and dimensions of the actors participating in the MAP. It is important for MAP authors to be able to preview the MAP Spatiotemporal layout at any time during the development, so that the appropriate modification may take place. The spatial layout tool makes possible for the author to view how the application window will look like at any temporal point during a potential MAP execution (e.g. which objects and where on the application window, appear on the 20th sec. of the MAP). The temporal duration of each object is derived from the temporal layout.

The author may set the desired time point (see Figure 11.7) through the *timer* and then with *update display* option get the layout of the screen. Thus, the display may be checked before a new object is inserted and find *if* and *where* it should be placed.

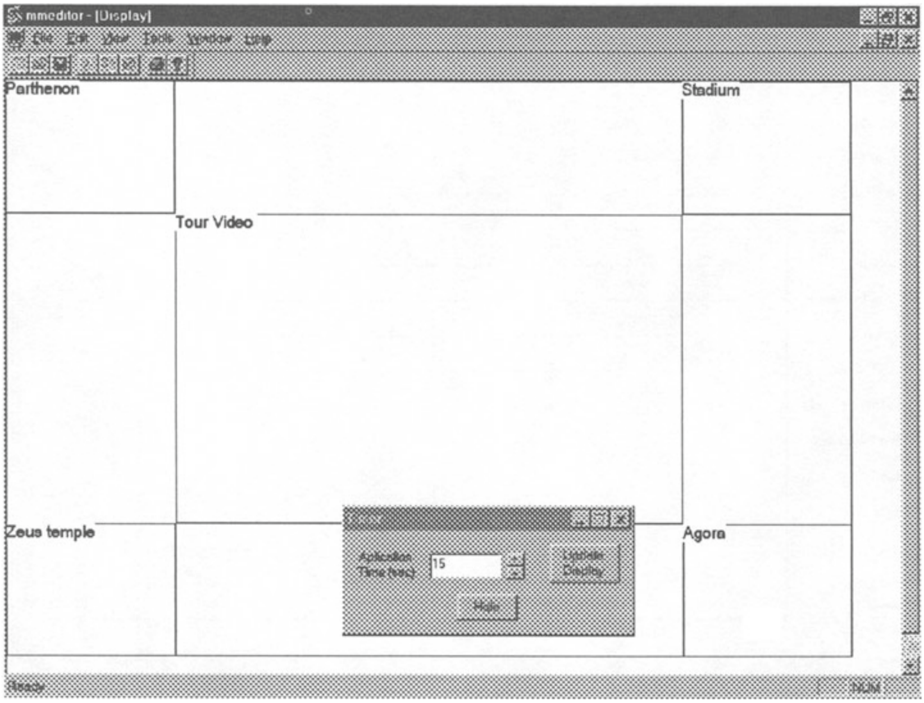


Figure 11.7: The spatial layout of the MAP at any time using the "Timer" tool.

11.5.3. Scenario animation tool

The most appealing feature of the verification tool is the opportunity to render (animate) the application. In other words, to have an animated view of the spatiotemporal specifications of an IMD in terms of spatial snapshots of the IMD evolving through time. This is accomplished through the "play" tool (see **Figure 11.8**). In this chapter we use the terms "rendering" and "animation" interchangeably. The animation is a simulation of an IMD execution session using consecutive snapshots of the application updated at regular time intervals. The author may change the value of this interval. The default temporal granularity is one second. This kind of animation enables the author notice any mistakes or misplacements and to mend them later.

The author has the opportunity to change the time limits of the animation and animate only a temporal part of the MAP (i.e. for the 5^{th} to the 25^{th} second). The animation may be interrupted at any time, while changes in the scenario may be done and then resume again into animation to verify the changes. The default values of the time limits are: 0 (for "Start Time") and the last second of the MAP (for the "End Time").

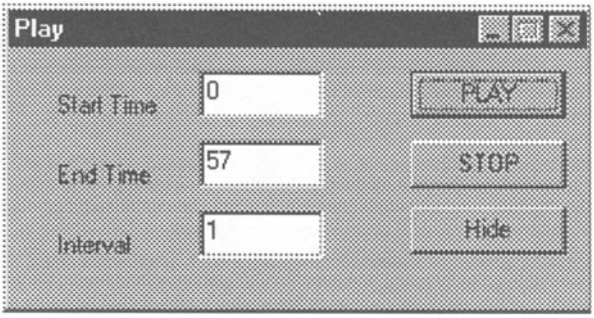


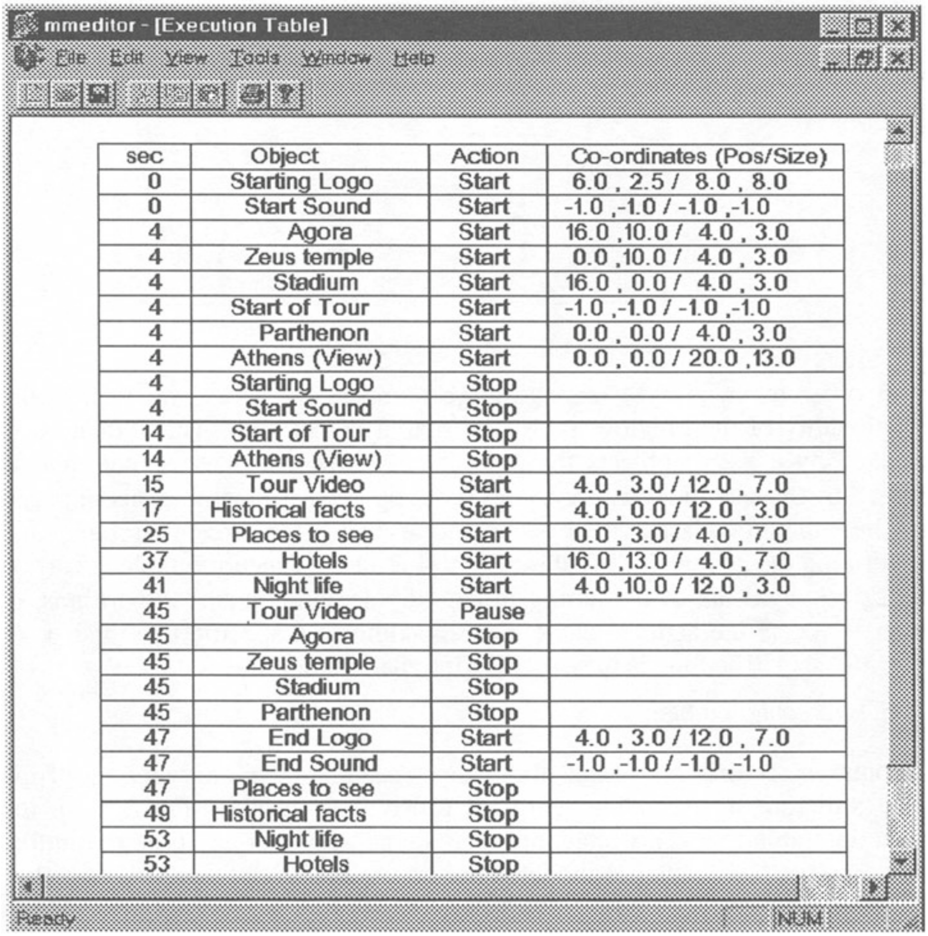
Figure 11.8: Scenario rendering tool.

In order to be able to manage large numbers of actors, we exploit the functionality of the previous tools (the spatial layout and). The animation is initiated by a list of objects that appear on the screen at start time, and is created by the spatial layout tool. This list is updated at time intervals imposed by the time granularity set by the author using the execution table and performing the actions it describes. Searching in the execution table is carried out by an algorithm of the family of the "divide and conquer" algorithms, in order to locate the actions of the corresponding to a specific temporal point (e.g. 10^{th} sec). The time is measured by internal system timers.

11.5.4. Execution table

A common requirement is that the authors need to have an overview of the IMD structure in ascending temporal order. This need is fulfilled by the execution table tool. This table, that may be generated at any time during the authoring session, includes the temporal and spatial coordinates of each start and stop event in the IMD (see **Figure 11.9**). The table contents are listed in ascending temporal order (i.e. start or stop events for actors along with their position and size in the IMD). The execution table is filled with the appropriate data resulting from the sorted list of objects, and sorted again depending on the time that each event occurs (two events, start/stop for each object). It is feasible to have the execution table in text file by pressing the save button when the execution table window is active.

The MAP becomes persistent by saving it to a file. The file is in binary format and can be interpreted only by the tool. Another type of output is the "script". This is adopted for compatibility between authoring tools. The script, however, contains the declarative script representing the spatiotemporal relationships among the participating actors as the user has defined them.



sec	Object	Action	Co-ordinates (Pos/Size)
0	Starting Logo	Start	6.0 , 2.5 / 8.0 , 8.0
0	Start Sound	Start	-1.0 ,-1.0 / -1.0 ,-1.0
4	Agora	Start	16.0 ,10.0 / 4.0 , 3.0
4	Zeus temple	Start	0.0 ,10.0 / 4.0 , 3.0
4	Stadium	Start	16.0 , 0.0 / 4.0 , 3.0
4	Start of Tour	Start	-1.0 ,-1.0 / -1.0 ,-1.0
4	Parthenon	Start	0.0 , 0.0 / 4.0 , 3.0
4	Athens (View)	Start	0.0 , 0.0 / 20.0 ,13.0
4	Starting Logo	Stop	
4	Start Sound	Stop	
14	Start of Tour	Stop	
14	Athens (View)	Stop	
15	Tour Video	Start	4.0 , 3.0 / 12.0 , 7.0
17	Historical facts	Start	4.0 , 0.0 / 12.0 , 3.0
25	Places to see	Start	0.0 , 3.0 / 4.0 , 7.0
37	Hotels	Start	16.0 ,13.0 / 4.0 , 7.0
41	Night life	Start	4.0 ,10.0 / 12.0 , 3.0
45	Tour Video	Pause	
45	Agora	Stop	
45	Zeus temple	Stop	
45	Stadium	Stop	
45	Parthenon	Stop	
47	End Logo	Start	4.0 , 3.0 / 12.0 , 7.0
47	End Sound	Start	-1.0 ,-1.0 / -1.0 ,-1.0
47	Places to see	Stop	
49	Historical facts	Stop	
53	Night life	Stop	
53	Hotels	Stop	

Figure 11.9: The MAP execution table, presentation actions in temporal ascending order including spatial information (position and size)

11.6. CONCLUSIONS

In this paper we presented an implemented authoring & verification methodology for MAP documents development. The MAP design is based on a theoretical model for spatiotemporal compositions in the context of multimedia presentations (Vazirgiannis et al, 1998). The tool may be used both for *prototyping* and *verification* of multimedia presentations or spatiotemporal compositions in general. As for the authoring phase, emphasis was put on the flexible definition of spatial and temporal relationships of the participating entities. The authoring phase consists mainly of declarative specifications of the spatial and temporal ordering of participating multimedia objects based on their spatial and temporal relationships.

The *verification* procedures give multiple tools to designers, preview their applications, in various ways: *spatial layouts* of the application window, *temporal layout* of parts (or the whole application), indicating the temporal duration and relationships among the participating objects and *animation (rendering)* of the application (i.e. what would be the execution of the application like)..

The *advantageous features* of the proposed tools are the following:

Regarding authoring:

- Declarative & visual authoring methodology
- Integrated Spatiotemporal MAP specification, taking into account the temporal and spatial features of the participating objects along with their spatiotemporal relationships. The methodology is based on a sound theoretical framework.
- Relative actor positioning in spatial and temporal domains. Moreover, the object under concern may be related to different objects (i.e. "Object A to appear 10cm to the right of the bottom right corner of object B and 4 cm above the upper-left corner of object C" or "Object A to start 10sec after object B and 4 sec before object C").

As for *verification*, we provide a set of tools that enable the authors verify multiple aspects of their scenario and answer queries related to the spatiotemporal configuration of the MAP. The tools include spatial/temporal layouts and MAP animation tools.

The most important advantage of our design is that the authoring and the verification process are well integrated and interleaved so as the author is able to verify the authoring specification within the authoring procedure and environment.

The tool presented above may be extended towards the following directions:

- *Interactive scenario integrity checking:* The scenario integrity issue is crucial when designing a complex interactive scenario. The multitude of events and the related interactions that may occur in a MAP session may create a lot of potential evolution paths which are difficult to follow at authoring time, considering all potential implications.
- *Connection to other authoring tools.* It will be feasible that the specification will be exported to various authoring tools scripts (i.e. Assymetrix/Toolbook, Mac/Director, Hytime etc.). This would require translation of the tool proprietary script into the authoring tool script language. The specific mappings would be required.

References

J. F. Allen. (1983). "Maintaining Knowledge about Temporal Intervals". *Communications of the ACM*, 26(11)

A. Duda, C. Keramane. (1995)."Structured Temporal Composition of Multimedia Data", in the proceedings of the *IEEE-MMDBMS '95 workshop.*

N. Hirzalla, B. Falchunk, A. Karmouch. (1995). "A Temporal Model for Interactive Multimedia Scenarios". *IEEE Multimedia Magazine, Fall 1995.*

M. Iino, Y. F. Day, A. Ghafoor. (1994). "An Object - Oriented Model for Spatiotemporal Synchronization of Multimedia Information", *in the proc. of the IEEE Multimedia Conference*, pp. 110-119.

M. Jourdan, N. Layaïda, C. Roisin. (1998). A survey on authoring techniques for temporal scenarios of multimedia documents, *to be published in Handbook of multimedia , CRC Press.*

Karmouch. A, Emery J. (1996), "A playback Schedule Model for Multimedia Documents", *IEEE Multimedia, v3(1), pp. 50-63.*

P. R. King. (1994). "Towards a temporal logic based formalism for expressing temporal constraints in multimedia documents". *Technical Report 942, LRI, Universite de Paris-Sud, Orsay*, France.

T. Little, A. Ghafoor, (1992)."Interval-Based Conceptual Models for Time Dependent Multimedia Data", *IEEE Transactions on Knowledge and Data Engineering, vol. 5(4), pp. 551-563.*

D. Papadias, Y. Theodoridis, T. Sellis, M. Egenhofer. (1995). "Topological Relations in the World of Minimum Bounding Rectangles: a Study with R-trees", *Proceedings of ACM SIGMOD '95.*

M. Vazirgiannis, Y. Theodoridis, T. Sellis. (1998)."Spatiotemporal Composition and Indexing for Large Multimedia Applications", *to appear in ACM/Springer-Verlag Multimedia Systems Journal.*