

# Using SIVA XML and SMIL for Interactive Non-linear Videos: a Comparison

Emanuel Berndl, Britta Meixner,  
Harald Kosch



# Overview

- Introduction
- Related Work
- Comparison
- Requirements
- Feasibility Analysis
- Metrics
- Proposal for Extensions
- Conclusion



# Introduction - Definitions

- Nimmagadda et al.: “**Multimedia presentations** are collections of different media files [...] like text, images, videos, and animations with different resolutions, durations, and start-times. [...] The layout of multimedia presentations is defined by the locations and the start times of the objects.”
- “An **interactive non-linear video** is a digitally enriched form of video materials arranged for an overall concept. It presents additional information beyond the original content. Furthermore, it offers new forms of influence and navigation in the video and additional contents.”
- **Hypervideo** is defined as video based hypermedia that combines non-linear video structuring and dynamic information presentations. Video information is linked with different kinds of additional information (like texts, pictures, audio files, or further videos). Users can mouse-click on sensitive regions (having spatial and temporal characteristics) within the videos to access the additional information (heterogeneous hypervideo) or jump to other scenes (homogeneous hypervideo). Hyperlinks build a graph between main video scenes and additional information.



# Comparison

## Multimedia Presentation

- Viewer rather passive
- Basic interaction and navigation may be possible
- No main medium
- Standard description format: SMIL

## Hypervideo/Interactive Non-linear Video

- Viewer active
- Extended controls and navigational elements
- Main medium video, enhanced with additional information
- No standard description format



# Related Work

- SMIL 3.0
  - Standardized
  - Several extensions available
  - 12 modules of elements and attributes
  - 5 profiles
  - Described in a DTD
  - General, flexible, and not by default made to fit our use cases of interactive non-linear videos
- SIVA XML schema
  - Not standardized
  - 6 main elements (2 optional elements)
  - Described in an XML schema
  - Specific, focused, and limited approach

Not taken into account for this work: NCM/NCL, CHM, ZYX, HTML5



# Requirements I

- **Media, main video, and annotations:**
  - One video is displayed as the main video
  - Additional information (images, audio-files, videos, and text) may be shown with this video
- **Event-based timing model:**
  - Main video and annotations may be time dependent or time independent
  - Event-based timing model is preferred to a structured timing model due to the high level of interactivity mixed with fixed points in time where annotations are displayed or hidden
  - By keeping timing issues as local as possible, synchronization is realizable more easily
- **Temporal relationships between main video and annotations:**
  - Start and end point or durations of display between the main video and annotations
- **Spatial relationships between videos and annotations:**
  - Positioning of main video and single annotations or groups of annotations
  - Annotations may be displayed (areas, overlay, paths), automated arrangement of annotations in defined areas facilitates the authoring process



# Requirements II

- **Decision elements at forks in video flow:**
  - Different strands of scenes → selection elements are needed to select the next scenes which are displayed to the viewer
  - Selection elements may be buttons or links
- **Table of contents:**
  - Extended navigation
  - Tree structure linked with single scenes in graph structure
- **Keyword reference list:**
  - Extended navigation
  - Keywords linked with scenes or annotations
- **Extensibility:**
  - For new ways of interaction that should be mapped into the model
  - Changes in existing XML files should be kept as local as possible in the structure without changing bigger parts of the existing file
  - Scripting is not considered as useful with respect to the affordance of an easy to use authoring tool



# Feasibility Analysis

Requirement	SIVA	SMIL
Media, main video and annotations	++	+
Event-based timing model	++	+
Temporal relationships	++	++
Spatial relationships	++	0
Decision elements	++	+
Table of contents	++	0
Keyword reference list	++	--
Extensibility	++	++





# Spatial Relationships between Main Video and Annotations

- SIVA XML
  - All media and navigational elements can be placed individually
  - Annotations can be arranged in areas, overlays or on paths
  - Player adapts to too large elements by using techniques like scrollbars to supply the full range of accessibility for all elements
- SMIL
  - Each element has to be aligned with left-, top-, right-, bottom-, height-, and width attributes in order to determine its position
  - If a given set, for example a list of links or buttons, is too large, it cannot be displayed entirely → portion that is too large for the displaying area will be cut out
  - Links can not be sized according to their text width
    - too large area → activate a link by clicking into the “free” area at the right side of the link → may result in faulty activations of links
    - size the buttons to a fixed length → if a given text exceeds that boundary, the text will be cut off as well



# Table of Contents

- Table of contents contributes to the non-linear character of the interactive video
- Panel with links in a tree structure is presented to the user
  - by activating one of these links, corresponding scene will be played
- SIVA XML
  - Allows the addition of sub entries for each item of the panel
- SMIL
  - Table of contents is constituted by a list of clickable links
  - For every entry, a link will be created and then arranged in a top down manner in the specified area
  - Suffers the spatial problems described on previous slide



# Keyword Reference List

- Keyword search could not be established for the SMIL export as such functionality is not supported by the language
- SIVA XML supports all the requirements that are needed for a search
  - Search for strings as keywords
  - Keywords are linked with scenes or annotations
  - Keyword of a scene → the scene starts at its beginning
  - Keyword of an annotation → the video starts playback at the point where the annotation is displayed
  - Keywords can be searched while the interactive non-linear video is played



# Metrics

	Size	Structure Complexity	Depth	Fan-In	Fan-Out
SIVA XML	58	67	5	12	8
SMIL (w/o meta)	40	430	$\infty$	21	16
SMIL (with meta)	41	507	$\infty$	22	38

- SMIL
  - Has less elements but higher complexity
  - $\infty$  for depth because `<par>` and `<seq>` can be boxed repeatedly
- SIVA XML
  - Has lower complexity and a fixed maximum depth



# Proposal for Extension

- Jumps in XML file
  - `<goto>` and `<end>` elements
- Choices at forks
  - `<fork>` and `<choice>` elements

```
<seq xml:id="start">
  <!-- Any SMIL content -->
  <fork shape="circle" size="20"
region="main_region"
dur="30s" defaultPath="path1"
xml:id="fork">
  <choice xml:id="path1" after="#fork">
    <!-- Any SMIL content -->
  </choice>
  <choice xml:id="path2">
    <seq> <!-- Any SMIL content -->
      <goto to="#start"/>
    </seq>
  </choice>
  <choice xml:id="path3">
    <par> <!-- Any SMIL content -->
      <end begin="20s"/>
    </par>
  </choice>
</fork>
<!-- Any SMIL content -->
</seq>
```



# Conclusion

- SIVA XML
  - Shows advantages regarding to the usefulness for interactive non-linear videos
  - All of the requirements that are needed to fully implement an interactive non-linear video are met by the language
- SMIL
  - Can realize most of the functionality (without keywords)
  - Lacks in some details like spatial problems of decision boxes, the placement of subtitles or moving annotations
  - Extension of existing SMIL documents may be complex due to many parallel, sequential, and conditional elements which are stacked and interwoven
- SMIL more complex than the SIVA XML indicated by XML metrics
  - construction and understanding of SMIL documents harder
- SIVA XML is better suited for interactive non-linear videos, while SMIL can be used but is not meant or designed particularly to support interactive non-linear videos
- Extensions would make SMIL more usable for interactive non-linear videos
- Implementation of a SMIL exporter into the SIVA Producer



**THANK YOU FOR YOUR ATTENTION**

