Hypervideo and Annotations on the Web

Madjid Sadallah
Olivier Aubert
Yannick Prié
LIRIS - Claude Bernard University Lyon 1 / CERIST, Alger

MMWeb 2011 – Graz, Sept. 8th

LIRIS – Lyon Research Center for Images and Intelligent Information Systems
UMR 5205 – CNRS/INSA-Lyon/Université Lyon 1/Université Lyon 2/Centrale Lyon
http://liris.cnrs.fr/
Context

- Audiovisual metadata is essential

- (incomplete) definitions:
  - **(AV) Annotation**: any piece of data linked to an audiovisual fragment
  - **Augmented/annotated video**: video augmented with annotation data

- What for?
  - Search / retrieval
  - Linking
  - Navigation
  - Visualisation
Visualisation

- Variety of visualisation modalities
- Two different goals:
  - Find the most appropriate visualisation for the annotations **for the current task**
  - Do not too tightly bind the annotations and their visualisations (prevents reusability)

► empower users with the ability to define their own visualisations
Hypervideo

- Term used by Ted Nelson (1960s)
- A definition: *interactive video-centric hypermedia document built upon an audiovisual content augmented with data in a time synchronized way*
- Two dimensions:
  - Hypermedia
  - Video-centered
Hypervideo specificities

- Annotations mandatory to address/augment video content
- Variety of visualisation modalities
- Space/time disorientation more pregnant
- Cognitive load / time pressure
- Rhetorical and aesthetic challenges
Advene principle

Structure schema:
- Shot
- Episode
- Document
- Panel
- Related shot

Annotation structure

Annotations:
- shot 1
- shot 2
- shot 3
- shot 4
- shot 5
- shot 6

View definitions: templates, rules...

Queries

Resources

A summary of Nosferatu
The different parts of the film:

Reunion
(00:00:000 - 00:03:39:000)

Meet
(00:03:39:000 - 00:20:17:000)

End of the journey
(00:16:13:000 - 00:20:17:000)

Reception
(00:20:17:000 - 00:23:46:000)

Package

www.advene.org

MMWeb 2011 – Graz, Sept. 8th
Advene lessons

• Validated vision of hypervideo concepts and annotation usage

• But: poor bet on visualisation emergence – did not meet appropriate users
  – Need to provide bootstrap components/examples
  – With appropriate level of malleability / expressivity
CHM: Component-based Hypervideo Model

- Main goals:
  - Conceptual and implementable model
  - Explicit annotation decoupling
  - Expressivity / simplicity
General overview
Core concepts

• Timeline reference
  – Linked to a Media Player
  – Attributes: position / duration / status

• Components
  – Visual / non-visual
  – Timeline-based / non-timeline-based
Data model

- Advene / Cinelab model
  www.advene.org/cinelab

- Annotation:
  - Video reference
  - Start/end timecodes
  - Type (identifier)
  - Content
CHM Plain components
CHM High-level components

AnnotationReader

Component

Timeline
TableOfContent
Map
Transcription
...

MMWeb 2011 – Graz, Sept. 8th
WebCHM – an implementation

- Extended HTML with namespaced attributes
- Client-side javascript library
- Reuses libraries (mediaelement.js, timesheet)
- Extensible
Simple example: ToC

```html
<div chm:component="jsonreader" id="data"
    chm:src="data.json" >
</div>

<div chm:component="videoplayer" id="tr"
    chm:src="video.ogv" >
</div>

<div chm:component="toc"
    chm:src="data"
    title="Story parts"
    chm:filter="type=='Parts'"
    chm:content="${content}"
    chm:timelineref="tr" >
```
Conclusion

Documentation and prototype available at http://advene.org/chm/

- Future work:
  - Extend model/vocabulary
  - Improve visual/interaction design
  - Complete implementation
  - Authoring environment
  - Cognitive studies
The End

Thank you for your attention.
Hypervideo and Annotations on the Web

Madjid Sadallah
Olivier Aubert
Yannick Prié
LIRIS - Claude Bernard University Lyon 1 / CERIST, Alger

MMWeb 2011 – Graz, Sept. 8th
Context

- Audiovisual metadata is essential
- (incomplete) definitions:
  - **(AV) Annotation**: any piece of data linked to an audiovisual fragment
  - **Augmented.annotated video**: video augmented with annotation data
- What for?
  - Search / retrieval
  - Linking
  - Navigation
  - Visualisation

- importance: evidence
- ATM, metadata is mostly limited to resource metadata
- Bridge the semantic gap
- Search: first thing that comes to mind. Need to cross the semantic gap though, it is the aim of many projects
- Linking: in the LinkedData perspective, AV cannot directly be linked. Need additional layers (annotations)
- Visualisation: seems obvious, but often overlooked: once you have found the video fragments that interest you, how do you visualise them appropriately?
Visualisation

- Variety of visualisation modalities
- Two different goals:
  - Find the most appropriate visualisation for the annotations **for the current task**
  - Do not too tightly bind the annotations and their visualisations (prevents reusability)
  - empower users with the ability to define their own visualisations

- visualisation of augmented videos
- «for the task carried out» : essential criterion
- if we design annotation content with a specific visualisation in mind, we can prevent the reuse of annotations through other means
- if we design visualisations too tightly linked with annotations, it prevents reusability -> lost time.

Until know, I have spoken about visualisation in general
Will introduce a concept that we think is appropriate
Hypervideo

- Term used by Ted Nelson (1960s)
- A definition: interactive video-centric hypermedia document built upon an audiovisual content augmented with data in a time synchronized way
- Two dimensions:
  - Hypermedia
  - Video-centered

Additional property: video-centered
-> brings time
Hypervideo specificities

- Annotations mandatory to address/augment video content
- Variety of visualisation modalities
- Space/time disorientation more pregnant
- Cognitive load / time pressure
- Rhetorical and aesthetic challenges

HV : restriction of generic hypermedia

Given the combination of video+annotations ->
- variety : AV- or annotation-focused (subtitles vs transcription) / temporal / non-temporal (static) / overview / detailed / synthesis...
- disorientation : common in hypermedia, but exacerbated

Before going into our proposal of a model for HV, I will give some information about where we come from.
Advene principle

Structure schema:
- Shot
- Episode
- Document
- Panel
- Related shot

Annotation structure:

Annotations:
- Shot 1
- Shot 2
- Shot 3
- Shot 4
- Shot 5
- Shot 6
- title: 'Meeting'
- title: 'Mission'
- document 1.pdf

View definitions:
- templates, rules...

Queries

Resources

Annotations

Package
Advene lessons

• Validated vision of hypervideo concepts and annotation usage
• But: poor bet on visualisation emergence – did not meet appropriate users
  – Need to provide bootstrap components/examples
  – With appropriate level of malleability / expressivity

These components need to feature appropriate levels of malleability
CHM: Component-based Hypervideo Model

- Main goals:
  - Conceptual and implementable model
  - Explicit annotation decoupling
  - Expressivity / simplicity

- There are other hypermedia models that introduce time (AHM + NCM). Our approach is compatible with AHM (but with more focused)
HV is a restriction of hypermedia – profit from this constraint
HV: based on augmented (AnnotationStructure) video (TimedMedia)
HV visualises augmented video through components

Layout: very general concept, supposed to use underlying implementation layout mechanism
Core concepts

- Timeline reference
  - Linked to a Media Player
  - Attributes: position / duration / status

- Components
  - Visual / non-visual
  - Timeline-based / non-timeline-based

Visual artifact on screen
Non-visual (AnnotationReader, audio player)
Data model

- Advene / Cinelab model
  www.advene.org/cinelab

- Annotation:
  - Video reference
  - Start/end timecodes
  - Type (identifier)
  - Content
- basic components
- used to build more complex components, or can be used as-is
- emerged from the study of a number of existing hypervideos
WebCHM – an implementation

- Extended HTML with named spaced attributes
- Client-side javascript library
- Reuses libraries (mediaelement.js, timesheet)
- Extensible
Simple example: ToC

```html
<div chm:component="jsonreader" id="data" chm:src="data.json" >
</div>

<div chm:component="videoplayer" id="tr" chm:src="video.ogv" >
</div>

<div chm:component="toc" chm:src="data" title="Story parts"
    chm:filter="type=='Parts'"
    chm:content="${content}" chm:timelineref="tr" >
</div>
```
Conclusion

Documentation and prototype available at http://advene.org/chm/

- Future work:
  - Extend model/vocabulary
  - Improve visual/interaction design
  - Complete implementation
  - Authoring environment
  - Cognitive studies
The End

Thank you for your attention.