From video information retrieval to hypervideo management

Olivier Aubert
Yannick Prié

CORIMEDIA
International workshop on multidisciplinary image, video and audio retrieval and mining
25-26 oct 2004 – Sherbrooke

LIRIS Lyon Research Center for Images and Intelligent Information Systems
FRE 2672 – CNRS/INSA-Lyon/Université Lyon 1/Université Lyon 2/Centrale Lyon

Overview

- Audiovisual information systems and information retrieval
- Hypervideos
- The advene model and application
- Discussion

AVIS audiovisual information systems

- Several activities...
  - Video indexing and retrieval
    - high / medium level features
  - Query results selection
    - video skim, surrogates
  - Query results exploitation
    - watch, reuse video
- ... different tools and descriptors

AV information usage

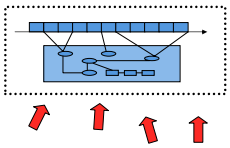
- Video usage
  - from simple visualisation to video reuse in other documents
  - retrieval is determined by video usage
- We should
  - seek for interoperability of systems
  - integrate video usage in AVIS
- cf. Web innovation
  - easiness of use
  - interoperability of tools
    - document visualisation, search, refering, manipulation, etc.

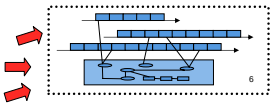
AV descriptors

- Video documents do not provide minimal AV units such as characters and words
- AV descriptors are the key to
  - retrieval
  - visualization
- We need to find useful descriptors
- A way of doing it : build systems
  - that are fully descriptor-based
  - that integrate AV information usage as a whole
  - that facilitate the emergence of new ideas

Hypervideos

- Annotated AV documents
  - AV document
  - *plus* annotation structure
- View ➡
  - « way of visualizing » an annotated AV document
- Hypervideo
  - A view that uses information from both the document and the annotation structure, giving access to the document as a stream
  - Hyper → random access
  - Video → temporal stream
- Generalization for several AV documents





Views and hypervideos: some examples

- Annotated AV document
  - Movie + shot/sequence decomposition
- Non-hypervideo views
  - The movie itself
  - Table of contents (with or without images from the movie)
- Hypervideo views
  - Movie with
    - shot number inserted on the stream
    - link to « next shot » / « next sequence »
  - Shot/sequence tree-view with links to the film at each shot transition
- Hypervideos in the wild

7

Advene

- « advene: to be added to something or become a part of it, though not essential » (Webster 1913)
- Advene: Annotate DVds, Exchange on the Net
- Objectives
  - Be fully descriptor-based
  - Allow new usages for audiovisual documents amongst different communities
    - Cinema / language teachers
    - Humanities scientists
    - Movies goers (sharing comments)
  - Better define and study hypervideos
    - Graphical interfaces
    - Documents and sharing

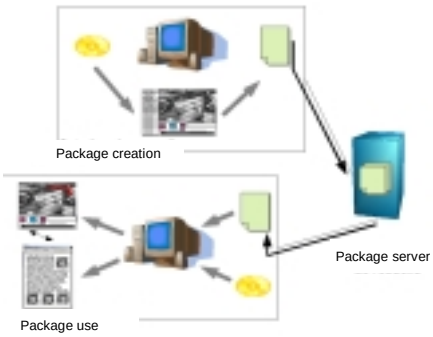
8

Advene principles

- Objective
  - facilitate the development of new ideas and usages
- Development principles
  - simplicity : wide audience
  - open source : movie-goers community developpers
- AV Documents (films) on DVD
  - alleviate rights problems
  - wide accessible corpus
- Package
  - Self-contained document with annotation structure, annotation definitions, views.
  - Hypervideo generation from package + DVD

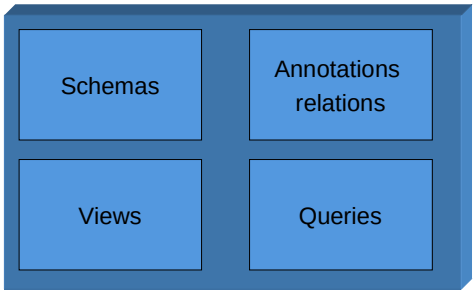
9

Advene overview



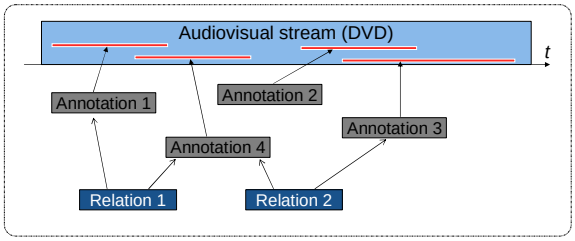
10

Advene packages



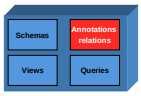
11

Advene annotated AV document



12

Advene
annotations and relations



- Annotation
  - type (ex : shot, character, episode...)
  - content (text, image, sound...)
  - fragment (stream-id, temporal interval, other...)
  - context (subtitles, language, angle...)
- Relation
  - members (two or more annotations)
  - type
  - content (text, image, sound...)

13

Advene
schemas

- Annotation type
  - name
  - constraints on the content
    - MIME types
    - Structured XML types (XML-Schema)
- Relation type
  - name
  - constraints on the types of the annotations belonging to the relation
  - constraints on the content (if needed)
- Schema
  - useful set of annotation and relation types
- Package imports
  - schema reuse
  - annotation / relation reuse

14

Advene
queries

- Dynamic extraction of significant Advene elements (annotations, types, schemas,…)
  - Applies on all the elements of a package
  - The result is a set of items
- For the moment
  - Not a full-featured query language (limited set of conditions)
  - Rather an implementation of a filtering
    - Fragment duration, annotations content, etc.

15

Advene
views

- Ways of visualizing and interacting with an element or set of elements from the annotated AV document
- Three main types in Advene
  - standard web navigator view (static view)
  - enhanced video player view
  - ad-hoc interface view
- Possibility to switch from one view to another

16

Advene
navigator views

- Detemporalized views
  - UTBV (user-time based view)
  - Mainly images + text in XHTML
  - Images extracted from the stream
- Template attribute principles
  - X(HT)ML file
  - TAL Syntax
    - element replacement
    - iterations
  - TALES expressions
    - queries on the package, expressed as paths
    - Ex : /package/annotations/annotation121/type

17

Advene
augmented video player views

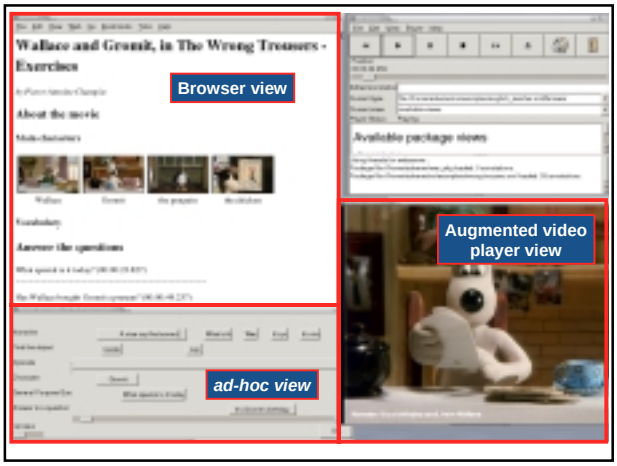
- Views in relation with a classical video player
  - Extended and automated actions
  - Control of the played stream
  - Augmentation of the stream
- First exploration in Advene
  - Rule-based model
    - ECA = event – condition – action
  - View = set of rules
- Example
  - If the player arrives to the beginning of a fragment associated with an annotation (event), if the annotation type is *character* (condition), then print on the stream ‘Character: character_name‘ (parameterized action)

18

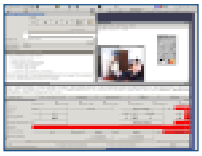
Advene
ad-hoc views

- Complex views
  - Explicit programming
  - Apply to a set of package elements
  - Examples :
    - Timeline
    - Tree-like view
    - Text synchronised with video

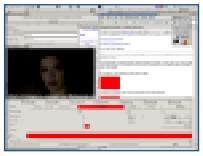
19



More examples



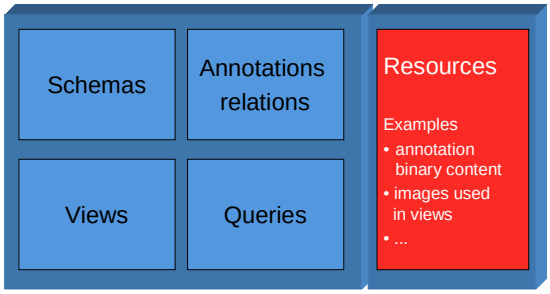
ELAN



Mulholland drive

21

Advene
package related resources



22

Conclusion

- AVIS : not only retrieval, but video usage
- Hypervideo is a concept for considering new usages
- Video usage is based on descriptors
- We need to build integrated, descriptor-based systems that facilitate the emergence of new usages/descriptors/interfaces
- The Advene project aims at it, with
  - Minimal modeling
  - Simplicity / extensibility
  - Easiness to integrate exterior knowledge (ontology, thesaurus)
  - Interaction with existing video retrieval tools.

23

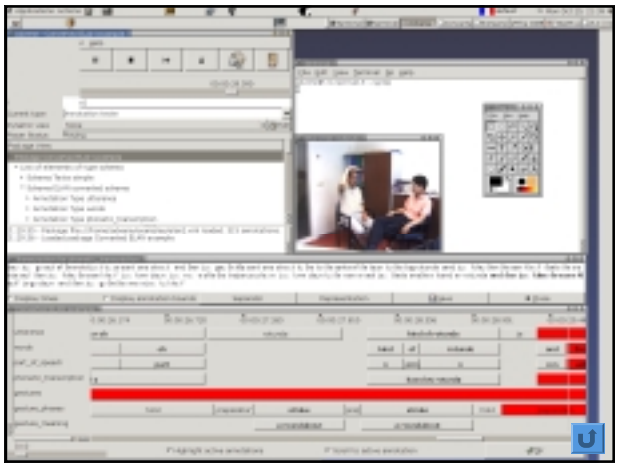
Ongoing work

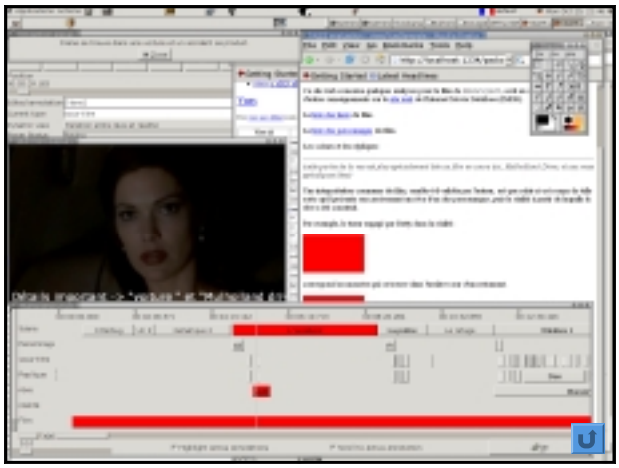
- Theoretical analysis
  - Audiovisual temporalities, detemporalisations
  - Hypervideo perception: cognitive studies and interfaces
  - Document theory (package, film references, etc.)
- Prototype
  - Finalization for open-source release
  - Various developpements
- Projects
  - DVD oriented project (cinema teaching)
  - Humanities oriented project (interaction analysis)
- We need money!
  - piloting the Open Source development

24

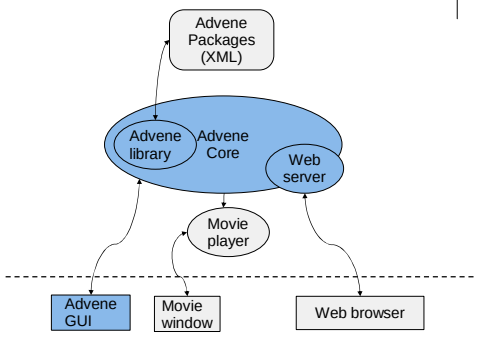
Thank you...

25





Advene architecture



28

TAL / TALES
generalities

- Developped for the Zope application server
- TAL : Template Attribute Language
  - Objective : to describe view in X(HT)ML + iteration and replace instructions int the attributes
- TALES : TAL Expression Syntax
  - Web requests for accessing elements
  - Path-like expression (cf. URL)

29

TAL / TALES
example

TAL Template

```
<h1>Studying <em  tal:content='here/title'>The title</em></h1>
Annotations:
<ul>
  <li tal:repeat='annot here/annotations'>
    <strong tal:content='annot/content/data'>The text</strong>
    (duration:
     <span tal:replace='annot/fragment/formated/duration'>xxx</span>
    ) </li> </ul>
```

Template pre-visualisation

Studying *The title*

Annotations:
The text (duration : xxx)

Visualisation

Studying *The Wrong Trousers*

Annotations :
Episode one: Gromit's Birthday (duration: 00:00:02.123)
A nice cup of tea (duration: 00:00:01.453)
...

30