

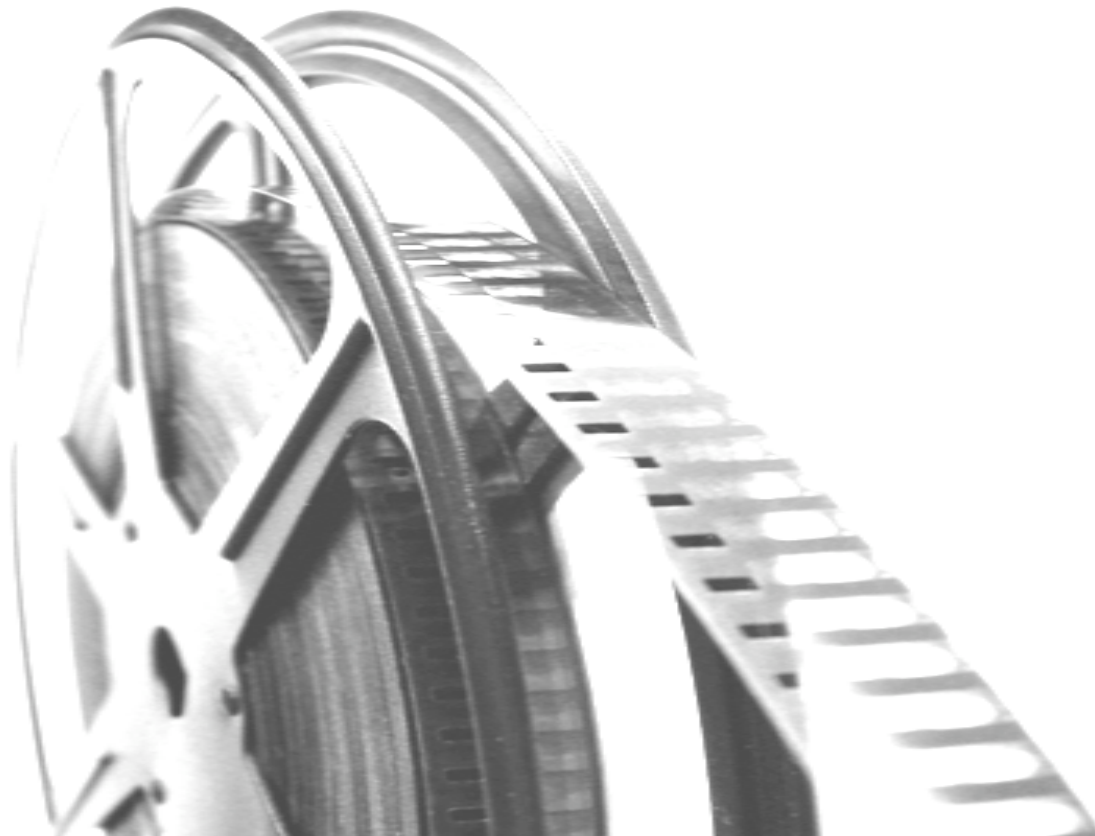
The background of the slide is a faded, light blue version of the famous Japanese woodblock print 'The Great Wave off Kanagawa' by Katsushika Hokusai. The image depicts a massive, curling wave about to crash over a small boat with several figures. In the distance, the snow-capped Mount Fuji is visible under a pale sky. The overall style is traditional Japanese art.

MIF14

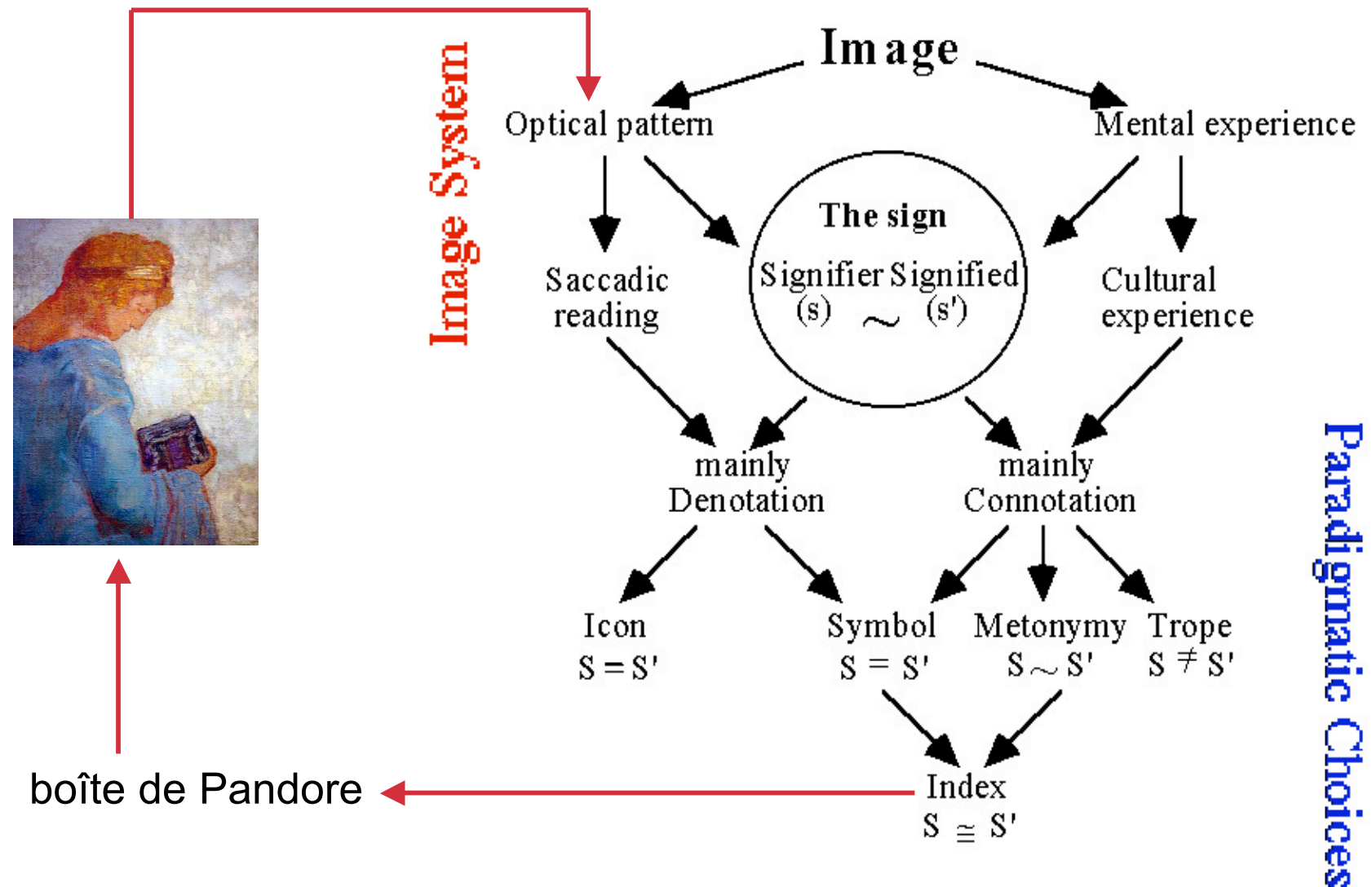
Anotation et Intelligence Artificielle

Yannick Prié et Frank Nack
LIRIS - SILEX

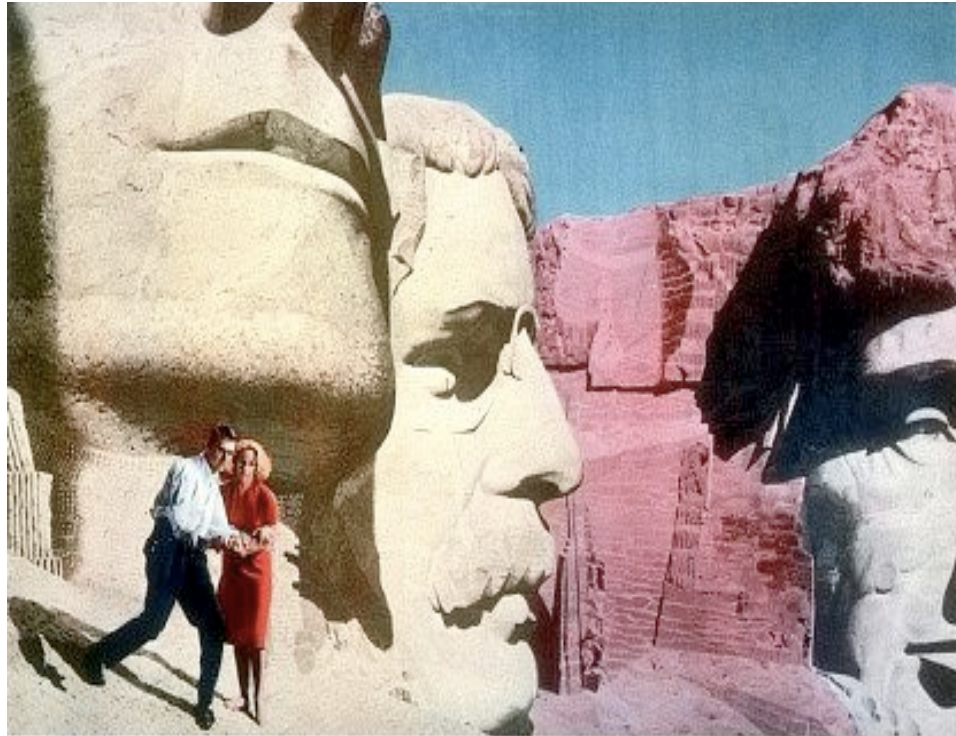
Annotation de média audio-visuel



Le système image - Résumé



Le système vidéo - Définition



North by Northwest (1959)

Un film ou une vidéo est une projection visuelle (des images représentant des objets) en mouvement, le plus souvent sonorisée.

Le système vidéo - Structure



Un **frame** est une des différentes images photographiques d'une vidéo. Normalement, il faut 24 frames pour constituer une seconde de film. Aux EU, il faut 25 images / secondes, les vieux films muets utilisent 18 images seconde. (voir aussi I-, P-, B- frames en compression vidéo)

Un **plan** est une prise de vue sans interruption. Un plan ne dure généralement que quelques secondes et constitue l'unité de base du langage cinématographique.

Une **séquence** est un terme du vocabulaire cinématographique désignant une série de plans continus qui forment une unité narrative. Un film peut comporter plusieurs séquences plus ou moins importantes pour la structure et la compréhension du film.

Toutes les image de 'North by Northwest' (1959)

Le système vidéo - Temps



Problème au niveau de la pellicule :

- Combien de frames sont utilisées pour représenter une seconde ?
- La vidéo est-elle compressée ou non ?
- Quelle zone de temps est utilisée comme point de référence ?

Problème au niveau du montage :

- Y'a-t-il des liens entre les prises (ponctuation : disparition, apparition, balayage, dissolution) ?
- D'autres changements temporels comme les frames vides ou les coupures ?

Problème au niveau de la narration :

- Flashbacks ou Flashforwards
- Temps de la narration = temps de la presentation = temps de l'histoire ou non?

Toutes les image de 'North by Northwest' (1959)

Le système vidéo - Sémantique



Metropolis

Dirigé par: Fritz Lang,

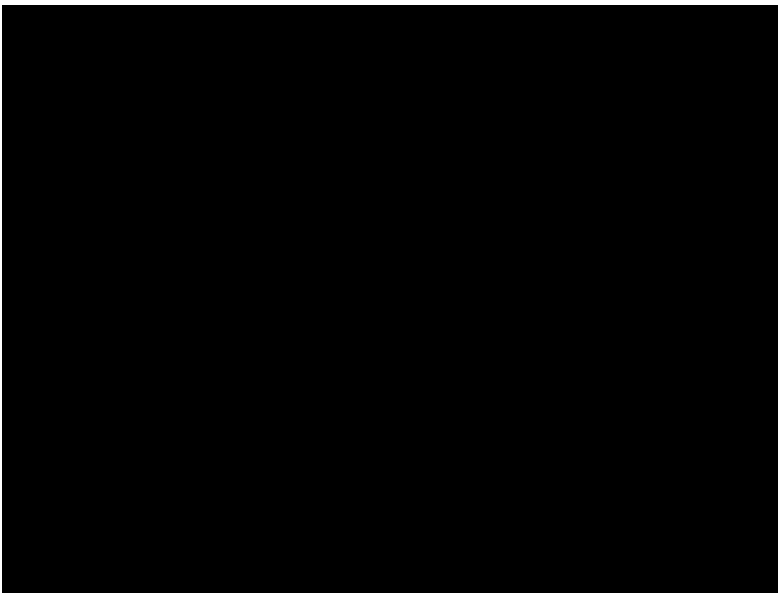
Pays: Allemagne

Année: 1927

Type: Feature Film

Introduction du conflit

Divertissement



Metropolis

Dirigé par: Unknown,

Pays: USA

Année: 2003

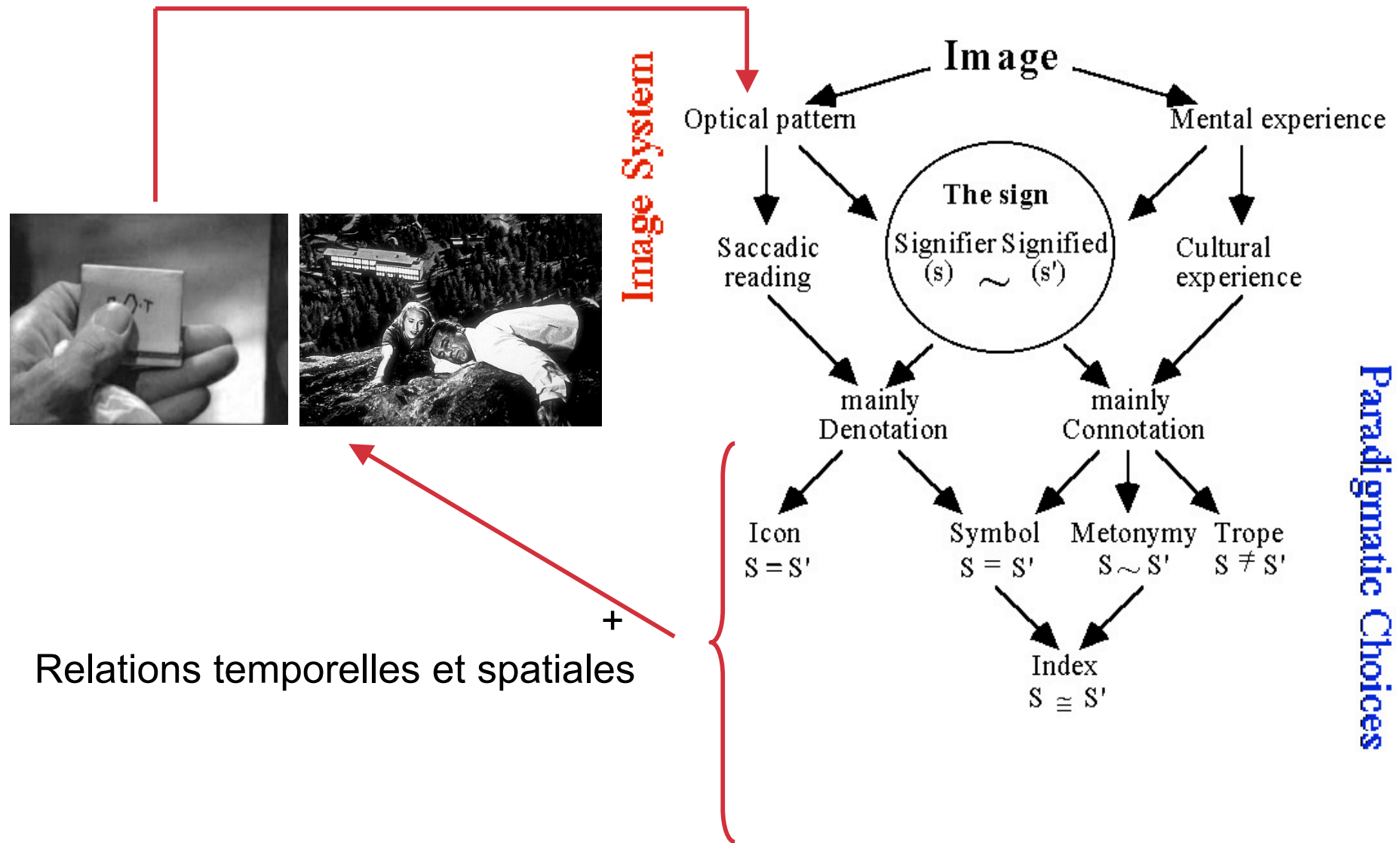
Type: Sequence

Episode

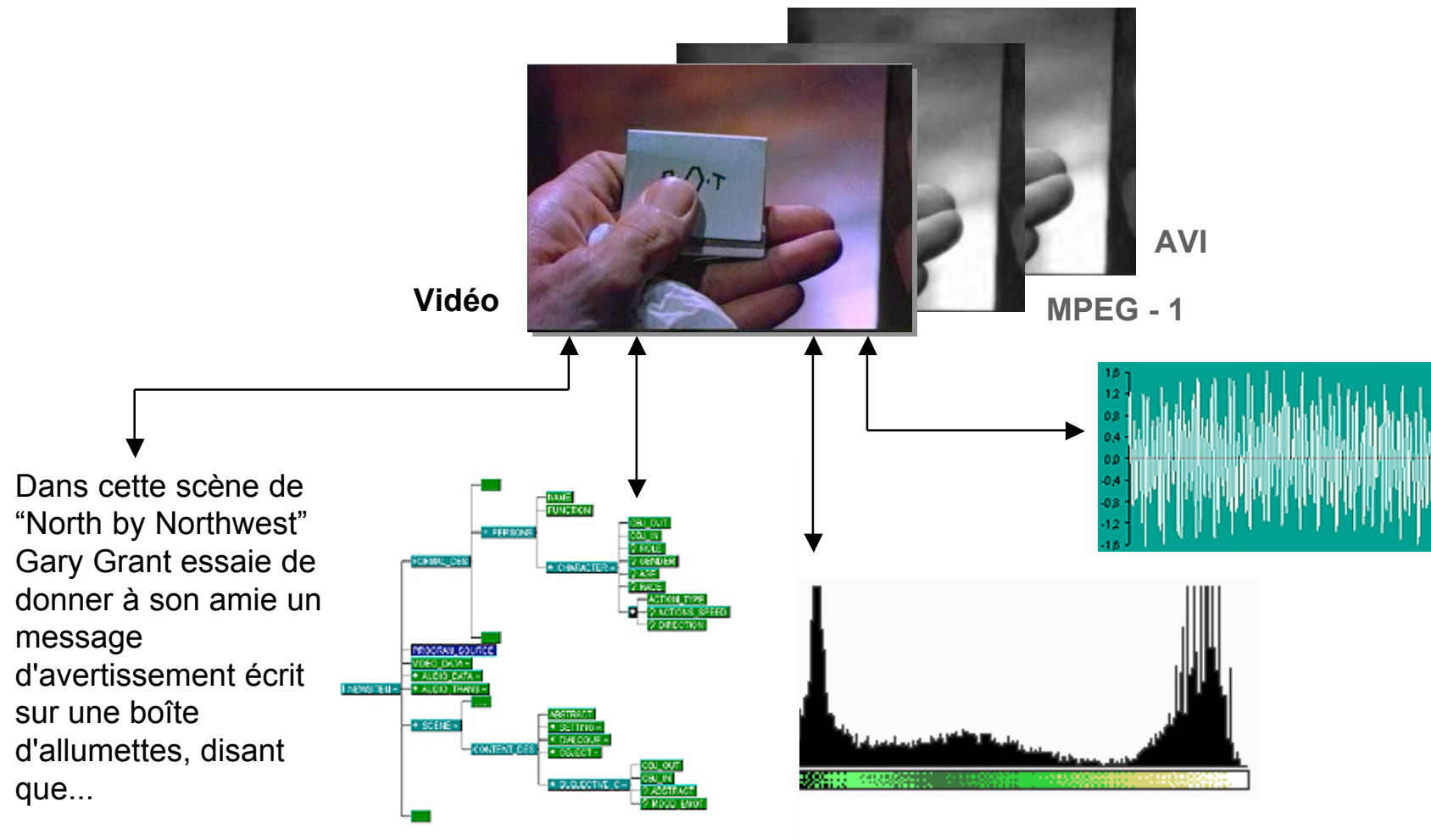
Resolution de conflit

Exercice d'édition

Le système vidéo - Résumé

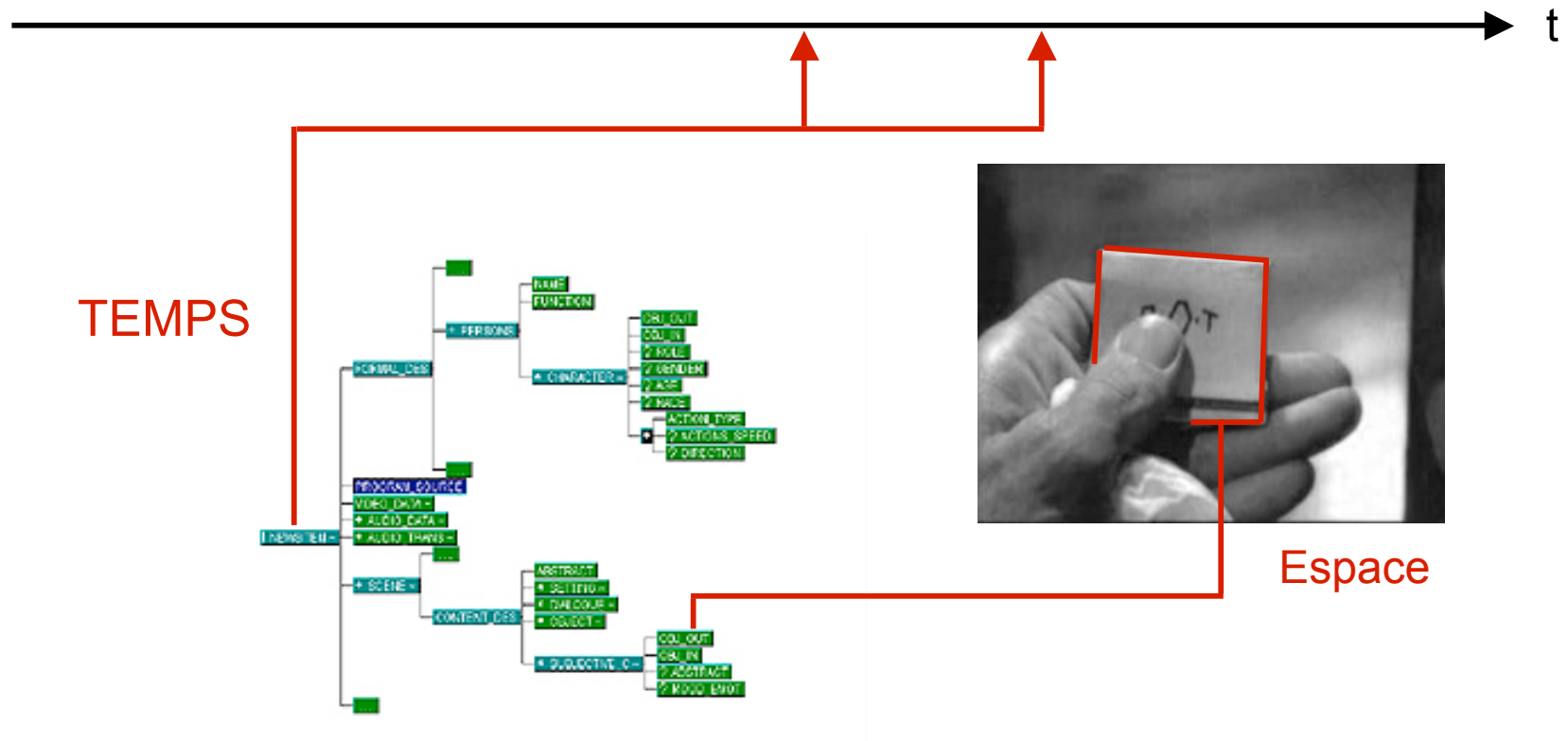


Annotation audiovisuelle



Toutes les image de 'North by Northwest' (1959)

Annotation audiovisuelle



Toutes les image de 'North by Northwest' (1959)

Annotation audiovisuelle – Ressources pertinentes

W3C

- RDF <http://www.w3c.org/RDF/>
- OWL <http://www.w3c.org/2004/OWL/>
- SW <http://www.w3c.org/2001/sw/>
- SMIL <http://www.w3c.org/AudioVideo/>

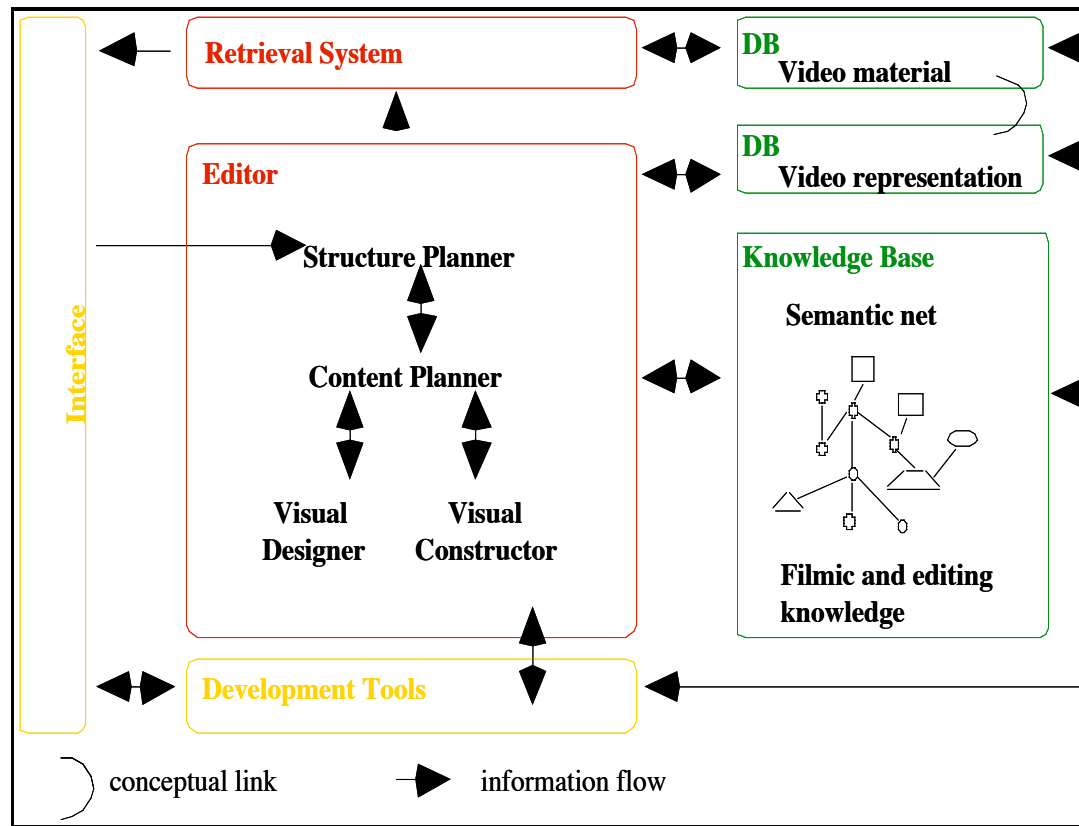
Synchronized Multimedia Integration Language (SMIL) est une spécification du W3C. L'objectif de SMIL est de permettre l'intégration de contenus multimédia diversifiés (images, sons, textes, vidéos, animations, flux de texte) en les synchronisant afin de permettre la création de présentations multimédias.

SMIL s'appuie sur le langage XML. La structure XML d'un document SMIL décrit le déroulement temporel et spatial des différents composants intégrés. En d'autres termes, SMIL permet d'indiquer le moment où un contenu sera affiché, pendant combien de temps et dans quelle partie de la fenêtre d'affichage.

ISO

- MPEG – 7 <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>
- MPEG – 21 <http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm>

Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)



AUTEUR (1996)

- Composants vidéo pré-définis
- Génération de séquences comiques d'après une base de clips
- Règles de continuité d'édition, de génération d'actions, de thèmes et de génération de présentation

Stratégies de l'humour

H-Strategy 1 *An action forms the most suitable subject for a joke, then an actor, then an object, and finally a location.*

H-Strategy 2 *if the action portrays an intention [goal], interrupt the action in a way that is unexpected by the character, so that the goal cannot be fulfilled and the character's mood is downgraded or he or she suffers in some way. (Mischief + Schadenfreude)*

H-Strategy 4 *If the action portrays an intention (goal), interrupt the action, in a way that is expected by the character, so that the goal is unfulfilled and the character's mood is downgraded or he suffers in some way. (Mischief + Schadenfreude + Superiority)*

Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)

Annotation - Personnage

Name	Description
Shot ID	
Startframe	
Endframe	
Identifier	Identifier for a character, e.g. a name or a number
Gender	male, female, hermaphrodite, artificial
Age	e.g. young, old, 25, etc.
Race	e.g. black, white, Asian, etc.
Appearance	a structure including: role e.g. lawyer, plumber, stewardess, etc. Costume kind e.g. business suit, apron dress, overall, etc. colour a doublet list providing the major colour for the top and bottom part e.g. [black, white] appeal e.g. casual, formal, etc.

Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)

Annotation - Action Personnage

Name	Description
Shot ID	
Startframe	
Endframe	
Identifier	Identifier for a character, e.g. name or number
Relative Position	(Screen position first frame, screen position last frame), e.g. (left, right), (left, middle), (right, right), etc.
Action	e.g. eat, drink, walk, read, etc.
Speed of action	e.g. slow, medium, fast
Direction of action	left, up-left, up, up-right, right, down-right, down, down-left, front, back, circular
Bodygesture	a structure containing: full body horizontal, vertical, left-diagonal, right-diagonal Head profile right, left, half-left, half-right movement up-down, left-right, up, down, left right, circle eyebrows up, down, straight, etc. line of sight left, right, straight, up, down, etc. mouth up, down, straight, open Hand left action/related object e.g.(tap/table) right action/related object e.g.(holding/head) Foot left e.g. tap, lift, etc. right e.g. tap, lift, etc.

Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)

Stratégies du montage

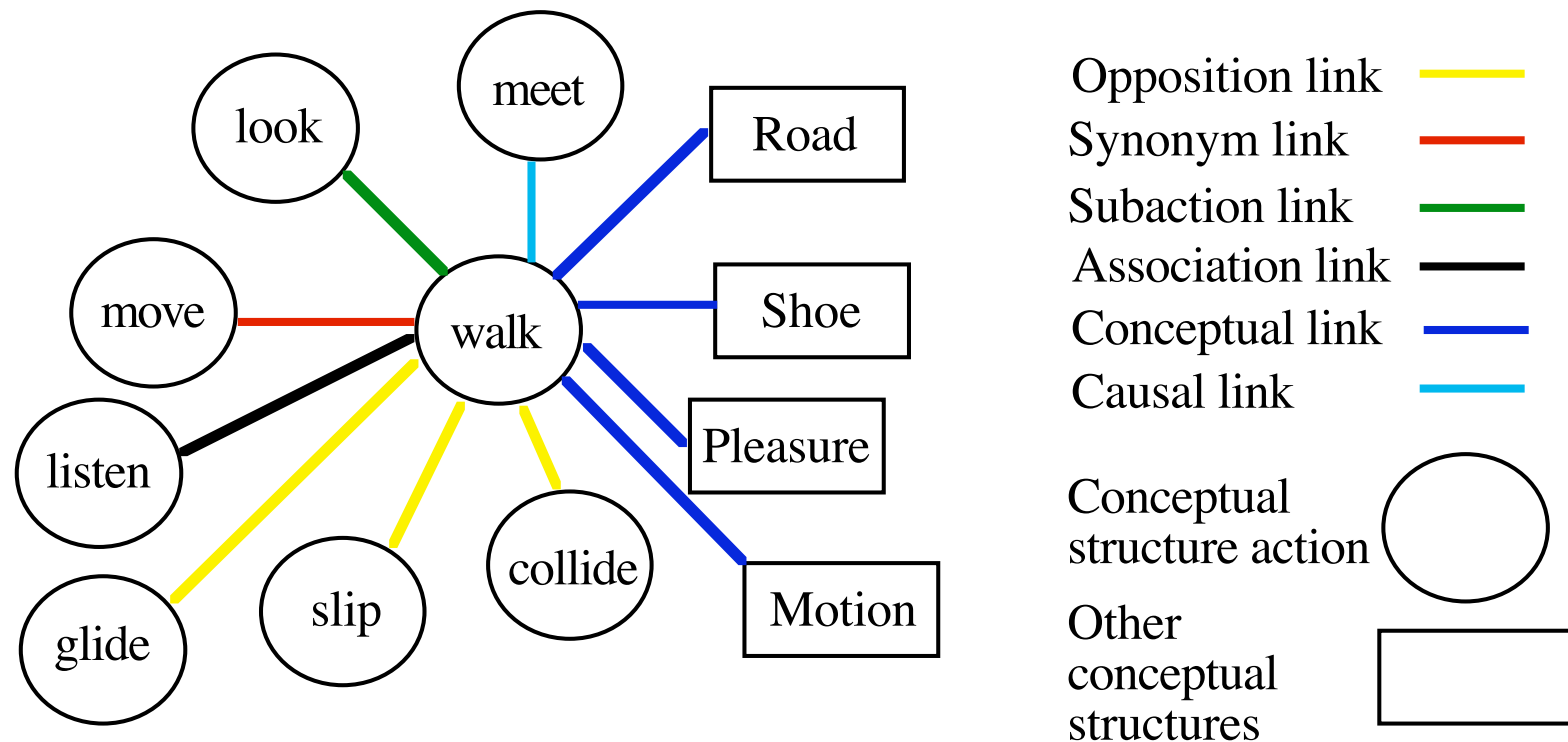
E-Strategy 1

*If sequence.kind = Motivation
then
Camera distance of Shot to be chosen is
long => location.function = outdoor
medium-long or medium => location.function = indoor*

E-Strategy 18

*If sequence.intention = relation and action and interaction
number of actors = 1 and
action of type moving
then
camera distance shot A $\leq$ camera distance shot B
line of sight character shot A = camera position in shot B
camera angle shot A = camera angle shot B
direction of action in shot A = camera movement in shot B
=> favour this shot as shot B*

Représentation sémantique de l'action 'walk'



Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)

Représentation conceptuelle de l'action 'slip'

Name	slip
Domain	motion
Nature of location	outdoors
Set of objects	[banana_peel, dog_shit, soap, ice]
Body part / related object	[shoe]
Location	[road]
Relation Location -> Object	under
Relation Object -> Body part	under
Intention	[unintentional]
Result actions	[sit, lie, kneel, shake, look_back]
Result mood	[anger, rage, astonishment]

Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)

Représentation des concepts abstraits

Abstract concept name	Representation structure [character/object, action]
time	[[shadows], [passing]]

Représentation des émotions

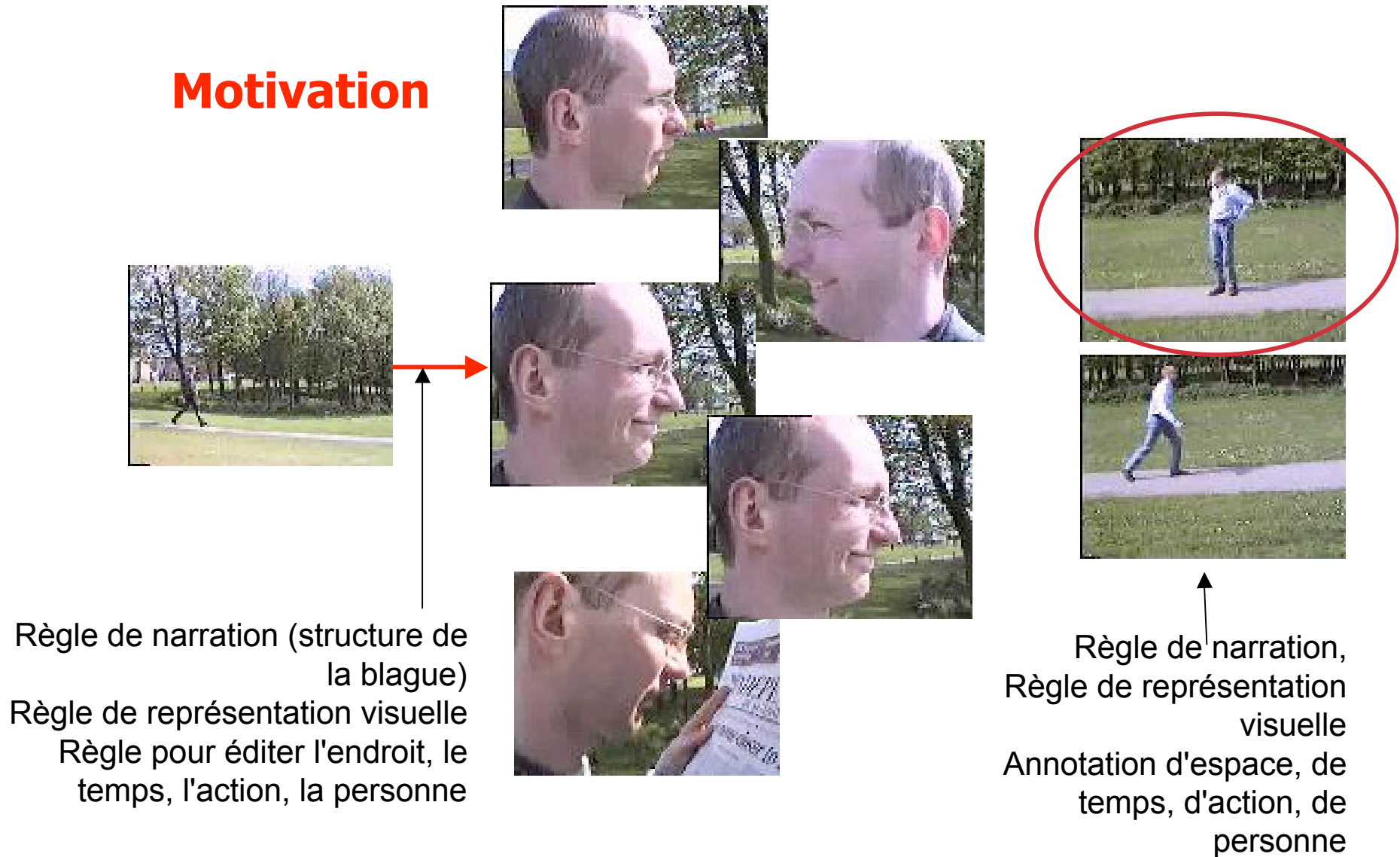
Type specification	Emotional token
Happiness	[Pleasure, Joy]

Représentation des concepts visuels

Emotional class name	Body part	Action
pleasure	Head	[lip, up]
	--	whistle

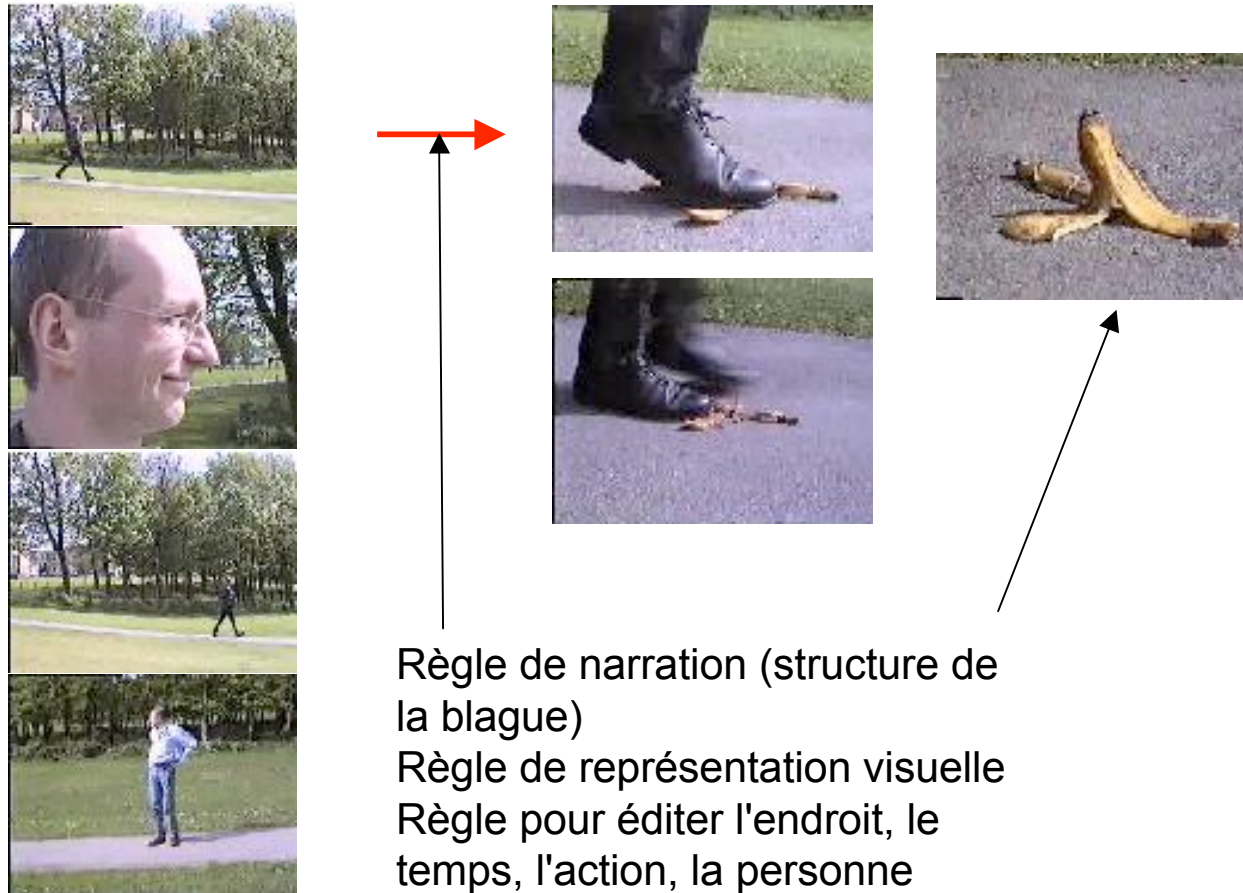
Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)

Motivation



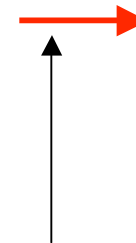
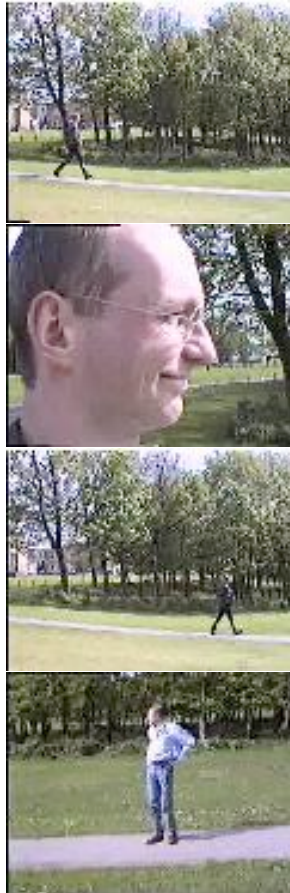
Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)

Realisation



Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)

Resolution



Règle de narration (structure de la blague)
Règle de représentation visuelle
Règle pour éditer l'endroit, le temps, l'action, la personne

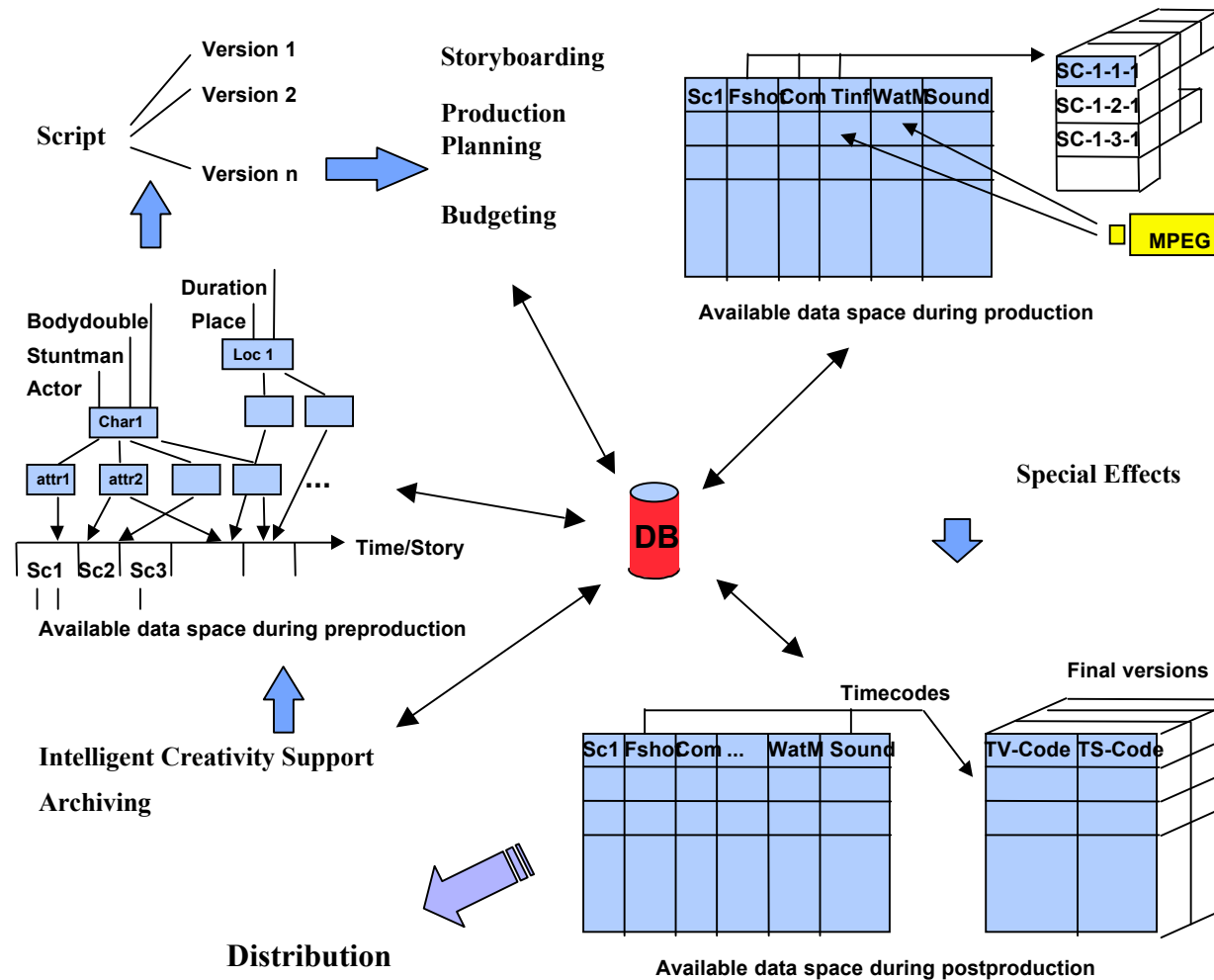
Annotation audiovisuelle - Exemple d'utilisation (AUTEUR)

Résultat



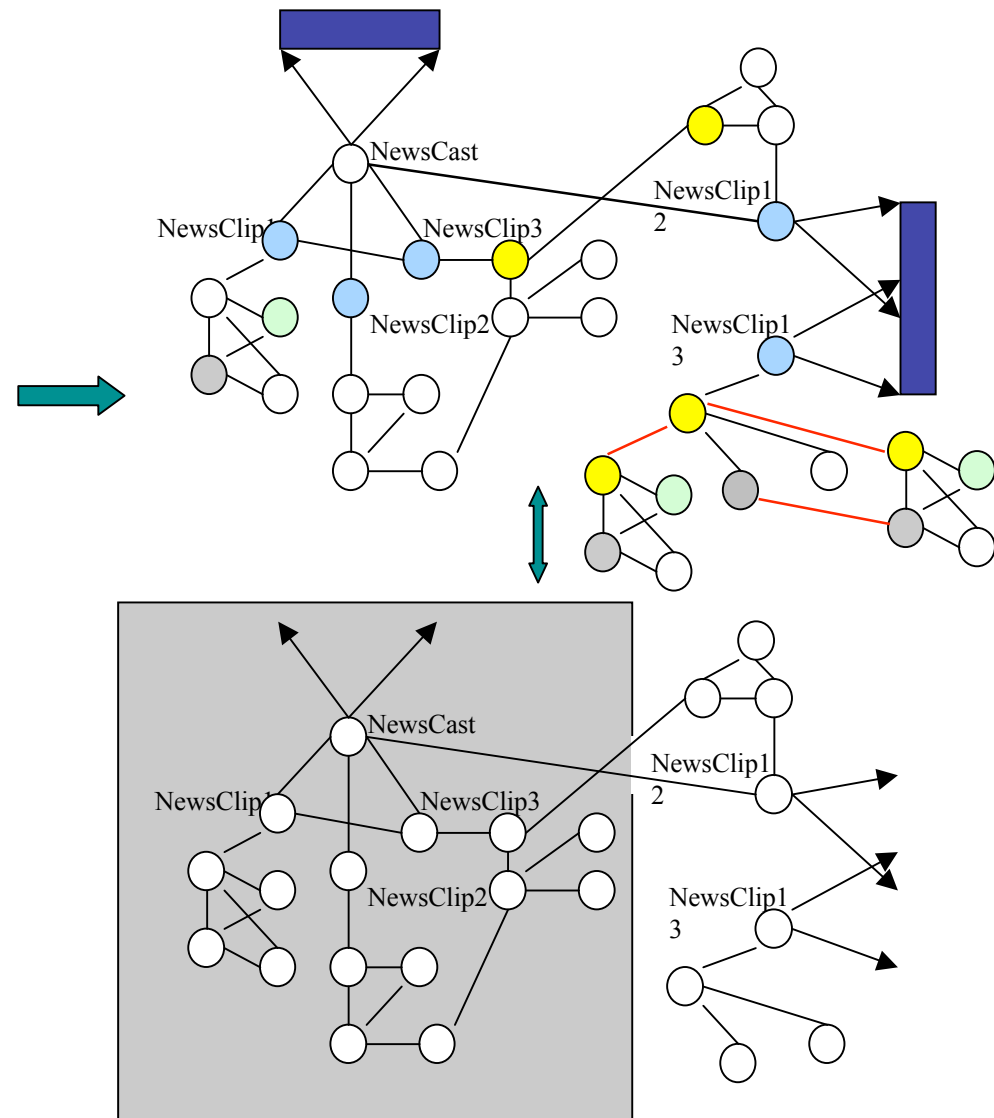
Annotation audiovisuelle – Exemple d'utilisation (A4SM)

A framework for digital media production



Annotation d'audio-visuel – Exemple d'utilisation (A4SM)

Archive	NewsCast
BP-Info	NewsClip
Person	Event
Formal_ds	Dialogue
	Character
	Setting
	Object
	Subjective C
	Media_Device

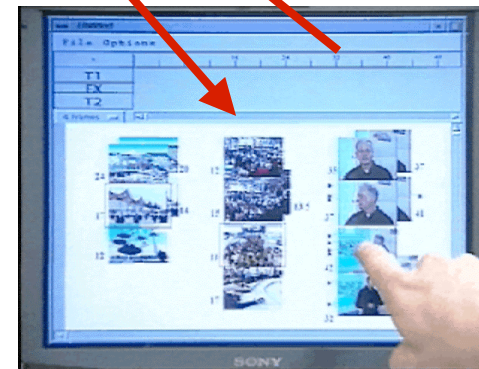
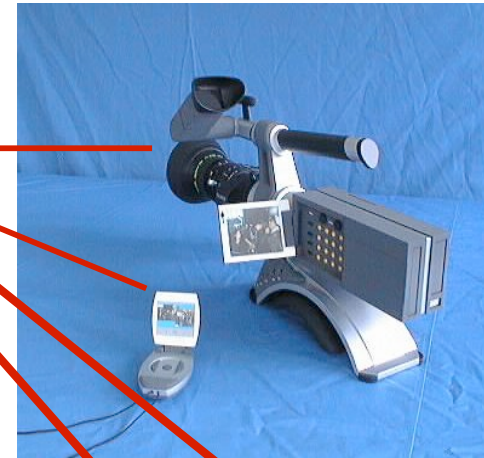


Annotation d'audio-visuel – Exemple d'utilisation (A4SM)



A4SM (Putz and Nack)

Common sense camera – Barbara Berry, MIT
Media Fabric (former Interactive Cinema Group)
Garage Cinema (Mark Davis), Berkeley



Annotation d'audio-visuel – Exemple d'utilisation (A4SM)



Matériel:

- Pentium 200 MMX
- Polihemus FastScan
- Sony EVID-30
- Videonics Phyton MPEG-1 encoder

Logiciel:

- Microsoft Windows 98
- Microsoft Visual C++
- Videonics Libraries

Camera:

- lens
- focus
- shutter
- colour

Tracker:

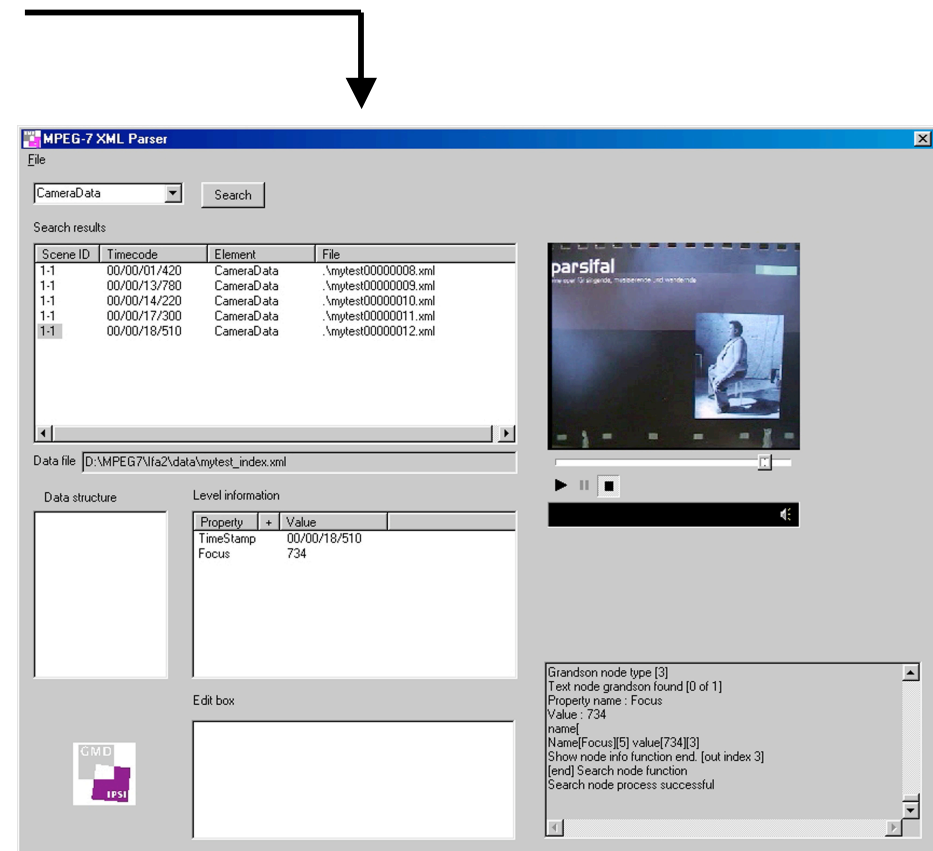
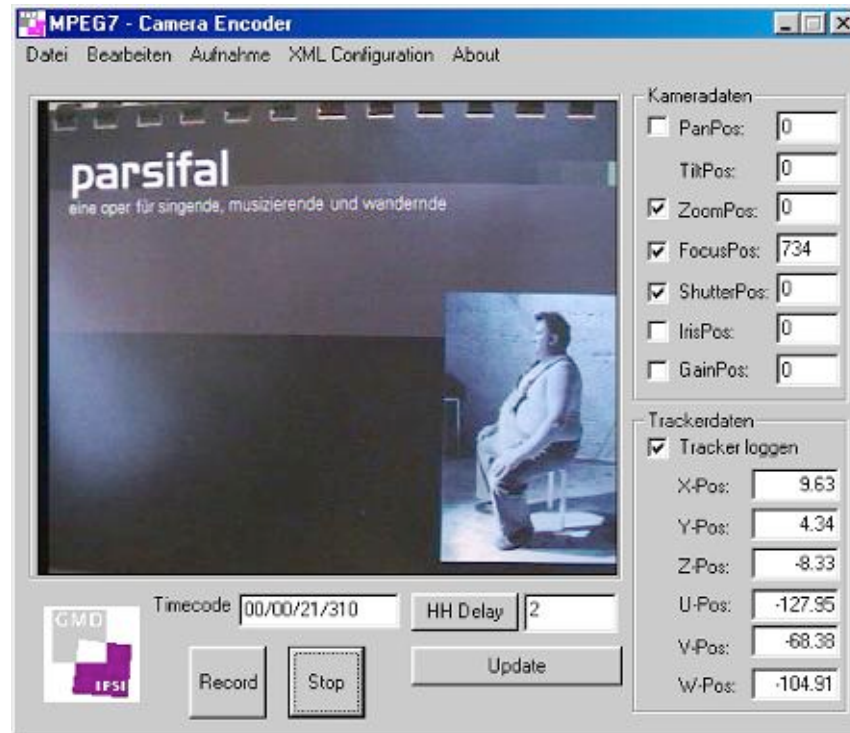
- location
- position
- roll (z)
- yaw (y)
- pitch (x)

Handheld:

- a-in
- a-out
- v-in
- v-out

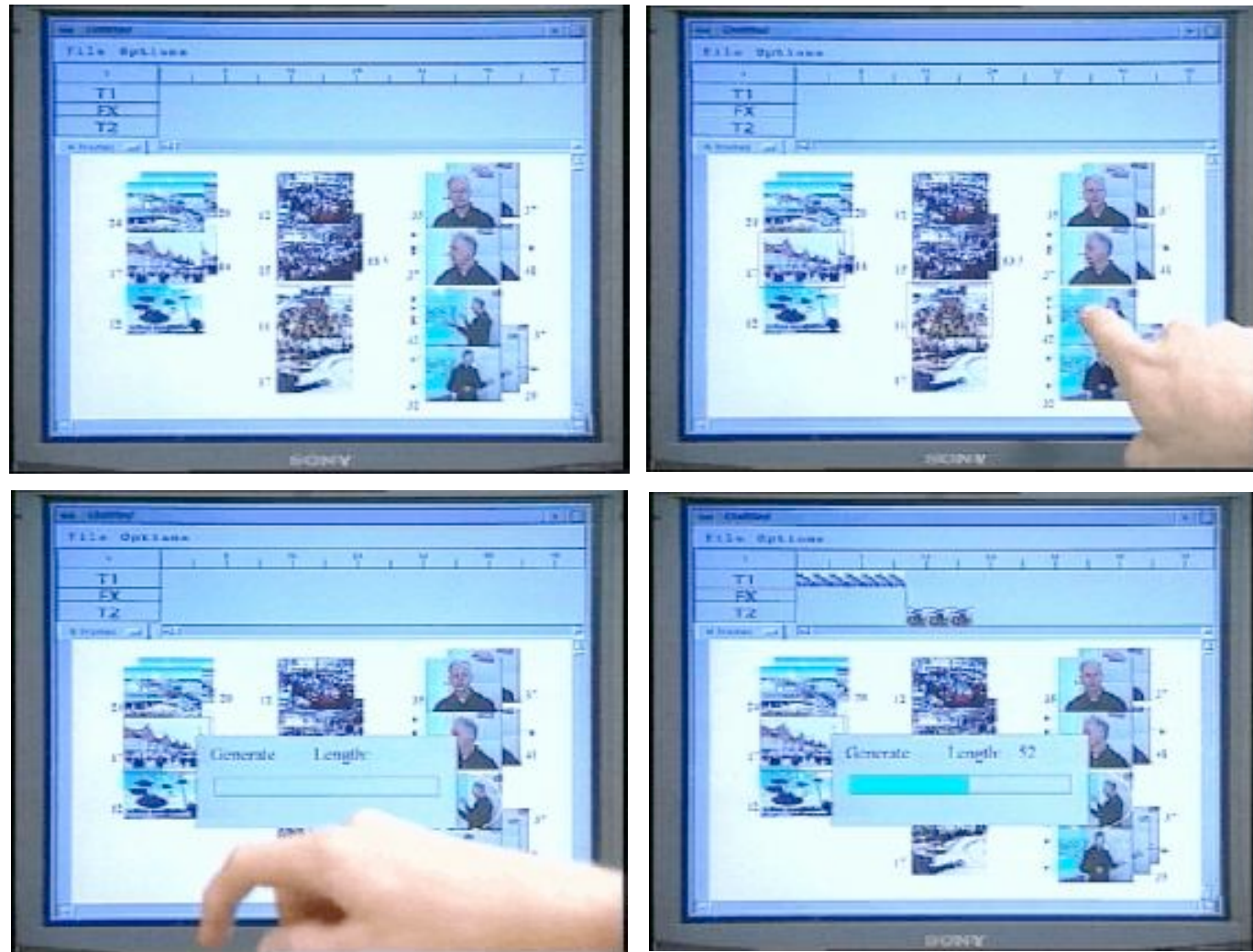
Annotation audiovisuelle – Exemple d'utilisation (A4SM)

Annotation semi-automatique (caméra)

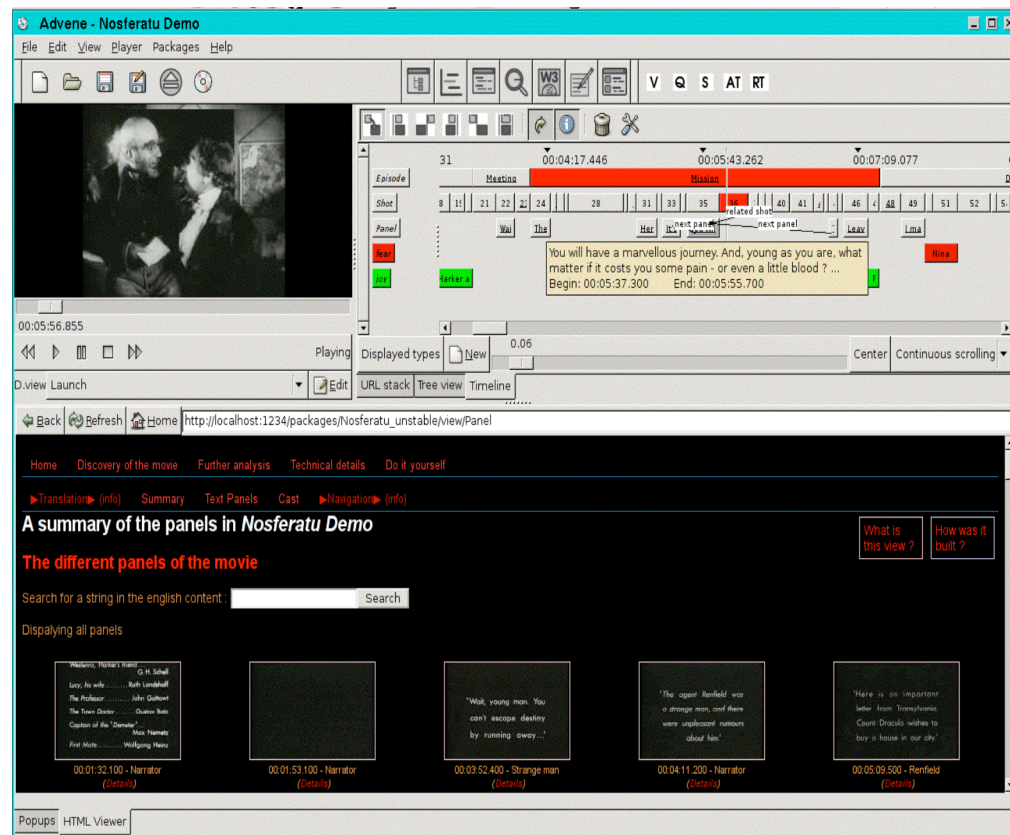


Annotation audiovisuelle – Exemple d'utilisation (A4SM)

Edition semi-automatique (utilisation des règles d'édition de AUTEUR)



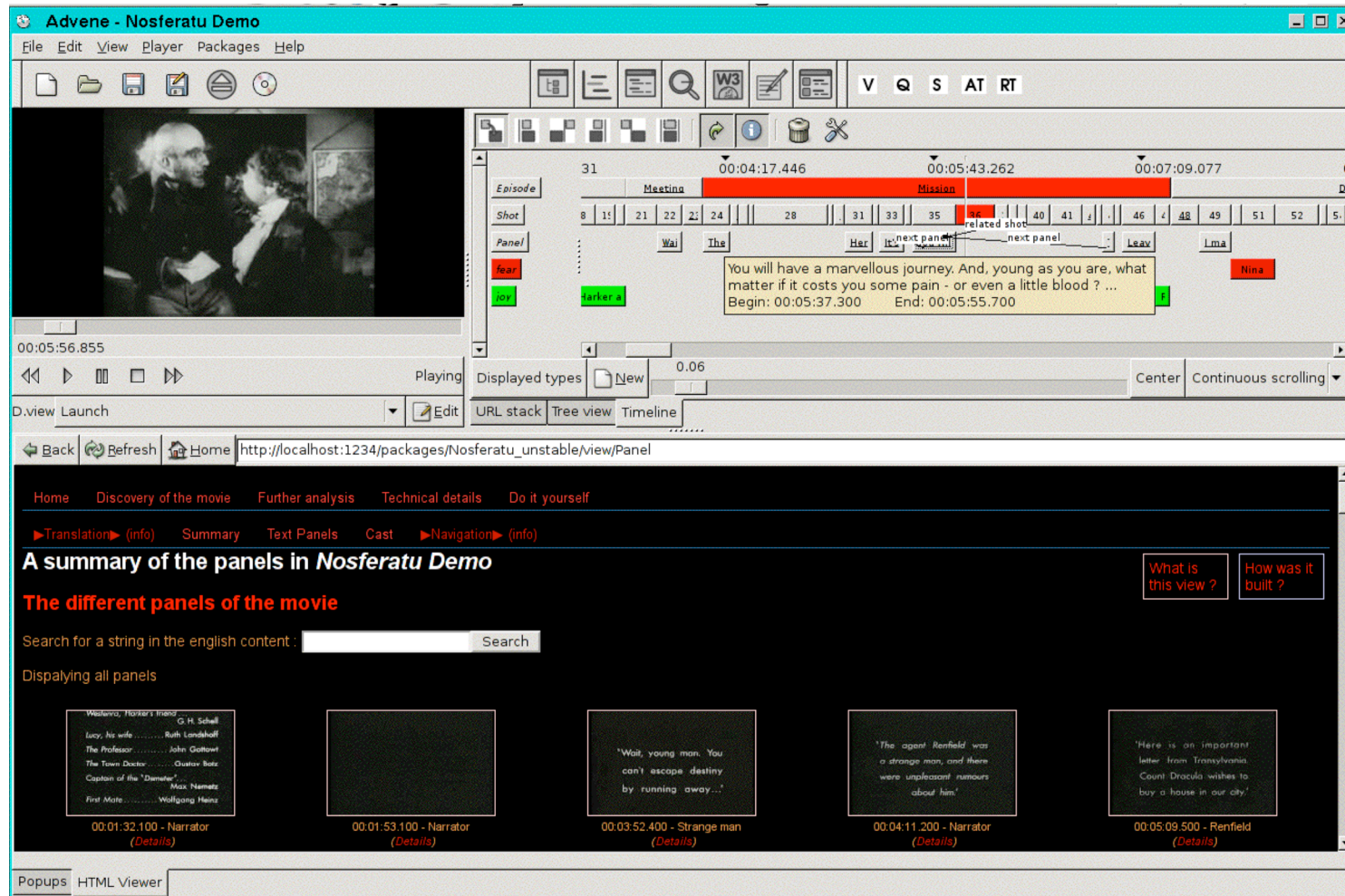
Annotation audiovisuelle - Exemple d'utilisation (ADVENE)



<http://liris.cnrs.fr/advene/>

L'objectif général du projet Advene (Annotations de vidéo numérique (Digital Video) échangées sur le NEt) est de fournir un modèle et un format permettant de partager des annotations de documents audiovisuels numériques (films, cours, conférence, etc. ainsi que des outils permettant d'éditer et de visualiser/jouer des hypervidéos générées à partir des informations liées aux annotations et aux documents audiovisuels.

Annotation audiovisuelle - Exemple d'utilisation (ADVENE)



Demo.