DELOS
NETWORK OF EXCELLENCE ON DIGITAL LIBRARIES

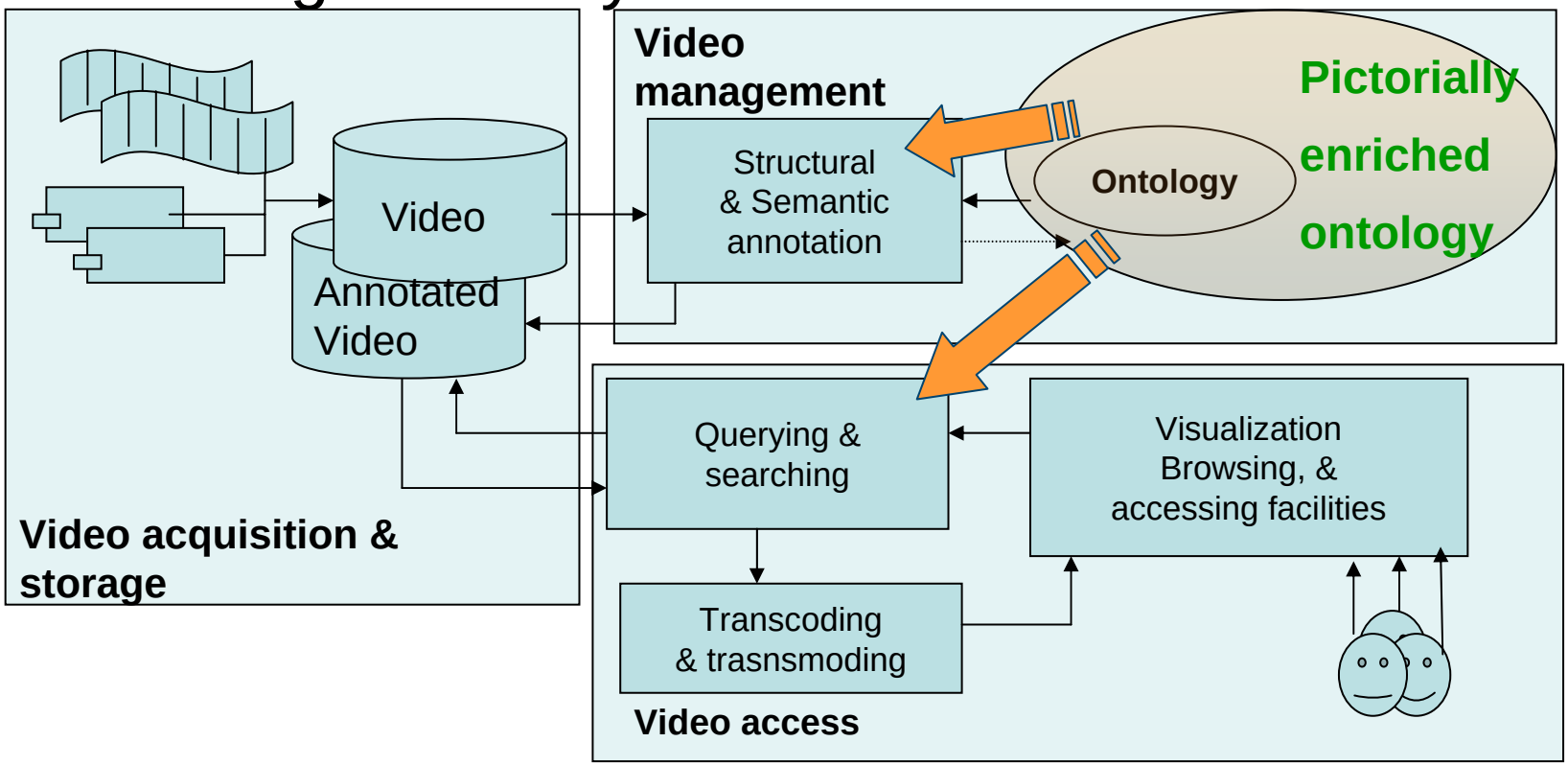
DELOS Task 3.6:
VAPEON : Video Annotation with Pictorially Enriched Ontologies

University of Florence (UNIFI-MICC), Italy
Technical University of Chrete (TUC), Greece
University of Modena and Reggio Emilia (UNIMORE), Italy
University of Amsterdam (UVA), Netherlands

Costantino Grana Imagelab, UNIMORE

Task objectives

- Video digital library



The goal is to develop **new ontology models** to support **retrieval by content** and **semantic annotation** of **video data**

DELOS
NETWORK OF EXCELLENCE ON DIGITAL LIBRARIES

Delos Vision & Task goals

- research directions defined in the DELOS brainstorming meeting
- ***Functionality specification for new objects***:
- ***Integrated and unified access to digital repositories***:
- ***DLMS specification, design, development***:
- ***Evaluation, metrics, benchmarks..***
- ***System integration, interoperability***:

- subsystems and deliverables of the task
- New solutions for video annotation and querying with PictEn ontologies
- PictEn ontology generalize concepts of ontology for accessing to multimedia DLs
- Visual feature definition, clustering techniques, design and development of the processes to annotate and access in Video DL
- Evaluation of computational aspects and of the added value with Picten ontology
- PE Ontologies, based on OWL, RDF and MPEG7 standards, are explicitly created to allow system interoperability

Task activities

- A – **METHODOLOGIES**
  - A1. Ontology definition
  - A2. Visual Features
- B **– IMPLEMENTATION**
  - B1. Creation of PictEn ontology
  - B2. Automatic annotation of video clips
  - B3. Experiments and tests.

A1. Ontology definition

- **Extension of the GraphOnto** component to support graphically the linking of pictorial information to linguistic ontologies (TUC)*
- Definition of **a query API** that allows utilizing the PE Ontologies in content-based queries (TUC)
- **Formal definition of PE ontologies** (UNIFI and UNIMORE)

- to support pictorial concepts and their visual features
  - **RDF extension** (UNIFI)
  - **MPEG-7 extension** (UNIMORE)

*A. Karanastasi, F. Kazasis, S. Christodoulakis: "A Natural Language Model and a System for Managing TV-Anytime Information in Mobile Environments", **Special Issue on ACM International Journal of Personal and Ubiquitous Computing**, Volume4, 2005

A2 Visual feature

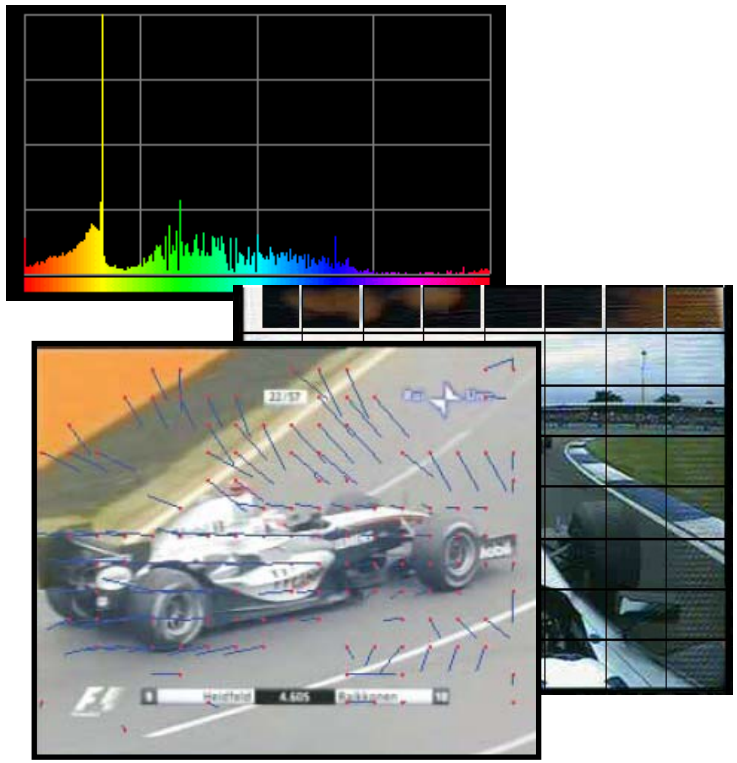
- Perceptual features for general purpose DLs (UNIMORE)
- Semantic features to domain specific (soccer) (UNIFI)
- parallel implementation of perceptual and semantic features for fast retrieval (UvA)

A2 visual feature (cont)

1. The **color histogram**, in 256 bins of HSV color space.
2. The **64 DCT coefficients for the frame texture**:
3. The **64 spatial color distributions**: the mean YCbCr color is computed for each area

4 The **four main motion vectors**:

computed as the median of the MPEG motion vectors, extracted in each quarter of frame.

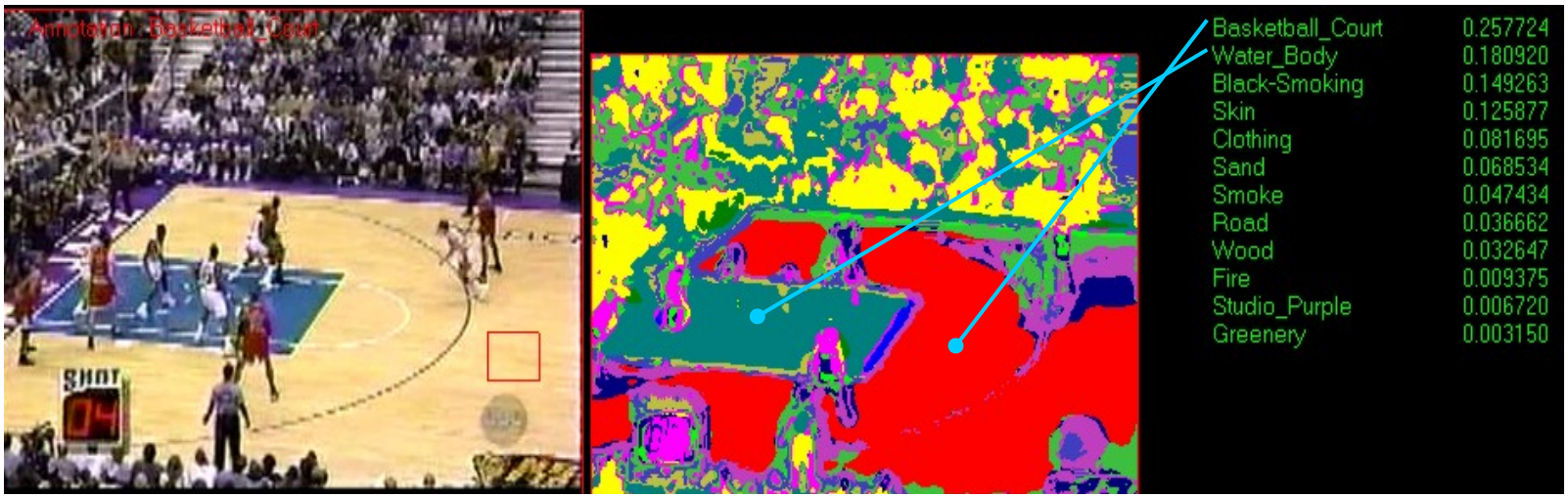


General purpose features,

used in Formula 1 and in broadcast TV (TRECVID)

A2 visual feature (cont)

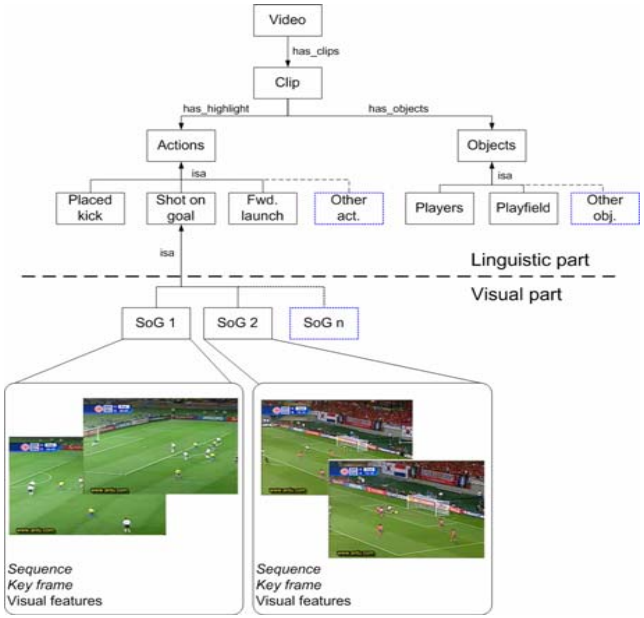
- Parallelization of processes for fast analysis at frame level (UvA)



–F.J. Seinstra, C.G.M. Snoek, D. Koelma, J.M. Geusebroek, and M. Worring. "User Transparent Parallel Processing of the 2004 NIST TRECVID Data Set", Proceedings of the 19th **International Parallel & Distributed Processing Symposium** (IPDPS 2005), Denver, Colorado, USA, April 4-8, 2005.

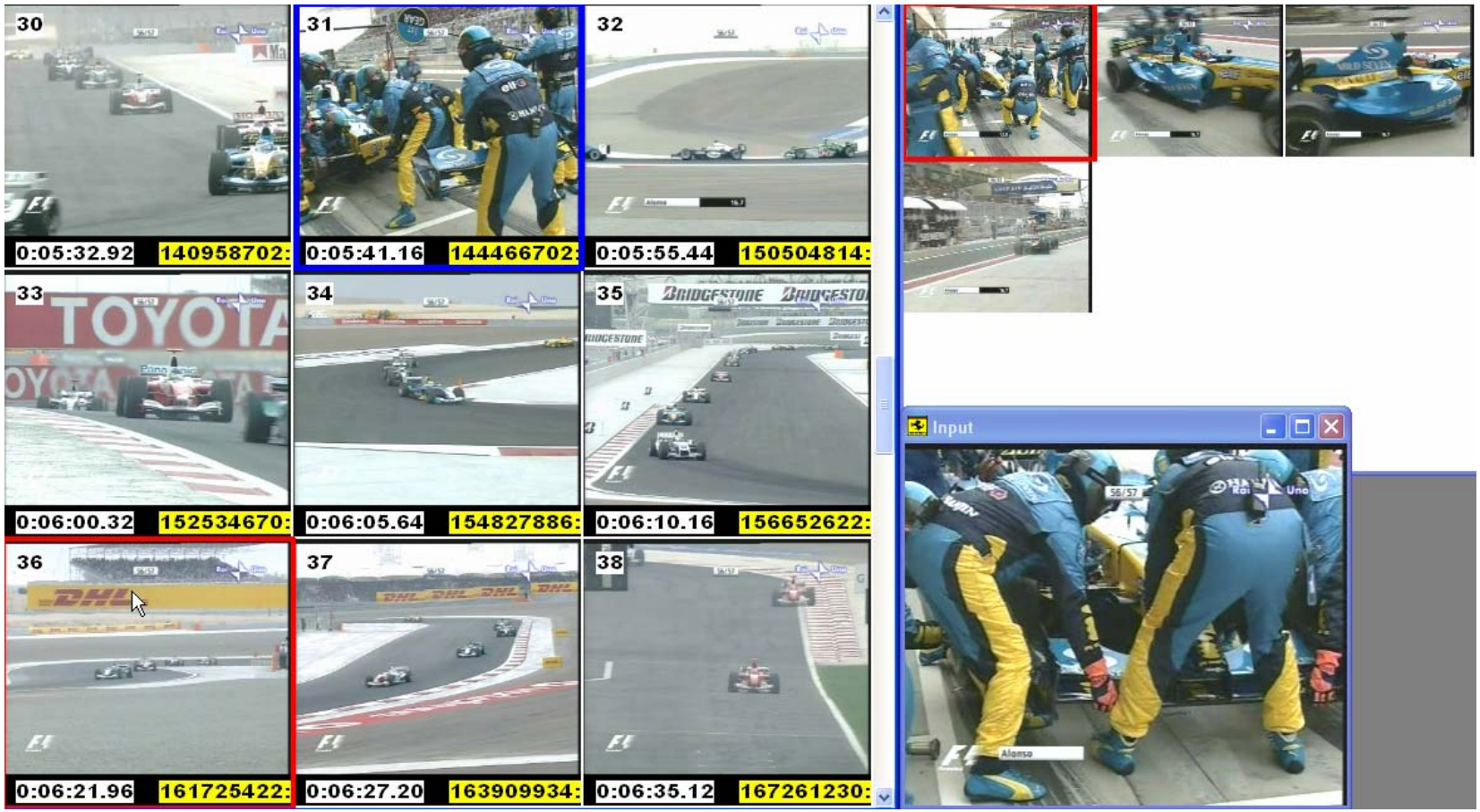
B: implementation

- 1) Formal definition of PE ontologies *
- 2) structural decomposition of video in shot subshot and clips
- 3) tools for manual annotation
- 4) tools for automatic annotation in specific context (soccer)
- 5)Tools for automatic extraction of pictorial concepts
- 6) automatic annotation with pictorial concepts and evaluation



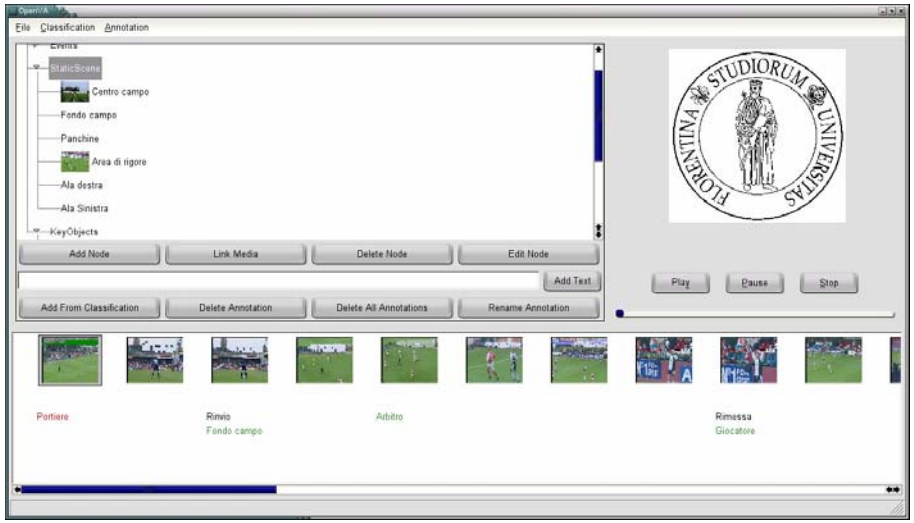
*M. Bertini, A. Del Bimbo, R. Cucchiara, C. Torniai, "Video Annotation with Pictorially Enriched Ontologies",
IEEE **International Conference on Multimedia and Expo** (ICME), Amsterdam (NL), July 6-8, 2005

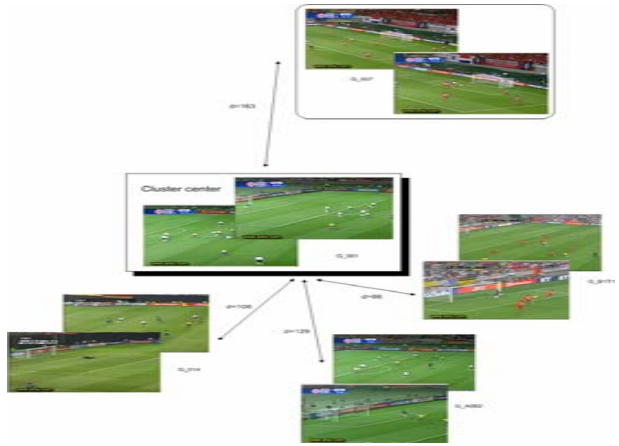
Decomposition of video in shot subshot and clips



–C. Grana, G. Tardini, R. Cucchiara, "MPEG-7 Compliant Shot Detection in Sport Videos"
–on Proceedings of **IEEE International Symposium on Multimedia** (ISM2005), Irvine,
–California, USA, December 12-14, 2005

Tools for manual annotation



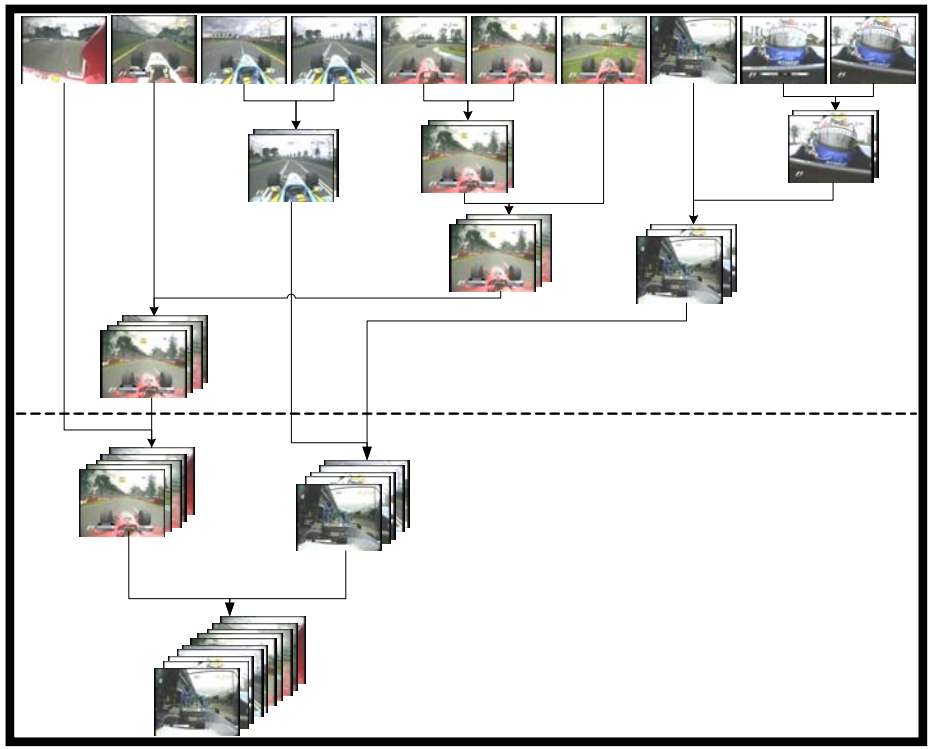


Manual annotation with natural languages and graphonto

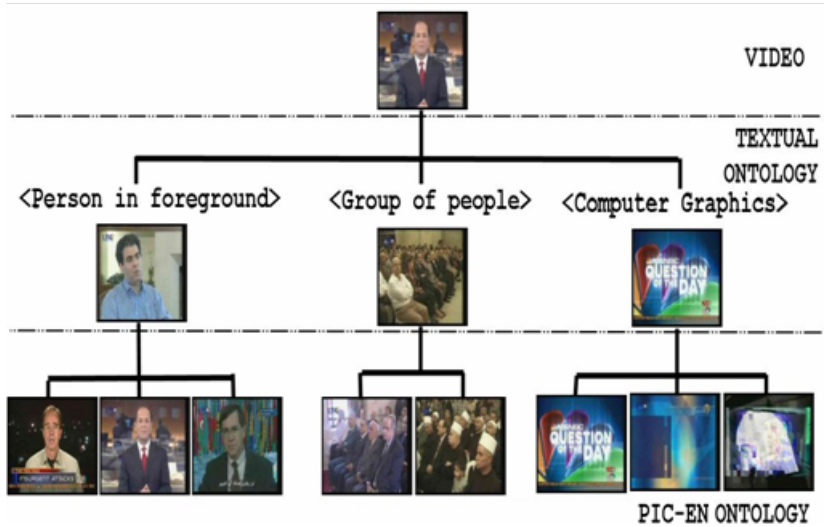
Manual annotation in RDF

Manual Annotation in MPEG-7

Extraction of Pictorial Concepts



Formula 1 Video DL



Broadcast Material TRECVID2005

(annotation in collaboration

UNIMORE + Univ. of south Florida, USA)

Automatic annotation with pictorial concepts

- Some tests and evaluation are started:
- Example: 340 clips were employed to create the ontology, with a total duration of 18 minutes (27800 frames). The tests were conducted on a 90-minutes video, with 1340 clips (135289 frames).

	CLASSIFICATION							
	Camera car		**E**xternal car view		**S**pectators		**P**eople	
C	40	45%	20	22%	0	0%	30	33%
E	10	1%	1050	95%	30	2%	20	2%
S	0	0%	0	0%	20	100%	0	0%
P	10	8%	10	8%	0	0%	100	84%

Confusion matrix for the test on Formula 1 videos.

Using ONLY

Perceptual

Pictorial concepts

No contxt-based features

Workshop, conferences and research exchange

- 2 researchers meetings – UNIFI / WINS-UvA
- 3 researchers exchanges – UNIFI / UNIMORE / TUC

- **organization** and participation to the “International DELOS Workshop on Audio-Visual Content and Information Visualization in Digital Libraries” (AVIVDiLib'05) –
- **organization** and participation to the 1st ImageLab Short Course on Computer Vision, Pattern Recognition, and Multimedia
- **organization** and participation to the International Conference on Multimedia and Expo 2005
- **organization** of the forthcoming DELOS summer school **–in collaboration with Noe Muscle**

- participation to
- “1st Italian Research Conference on Digital Library Management Systems"
- “2st Italian Research Conference on Digital Library Management Systems“
- 13th International Conference on Image Analysis and Processing,Cagliari Italy, 6-8 Sept, 2005
- European Semantic Web Conference 2005
- Summer School on Ontological Engineering and the Semantic Web
- 9th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2005) ACM Multimedia Conferences Nov 2005, Singapore
- IEEE Symposium of Multimedia San Diego Dec 2005

DELOS
NETWORK OF EXCELLENCE ON DIGITAL LIBRARIES

Publications 2005

- Tsinaraki C., Polydoros P., Christodoulakis S., “GraphOnto: A Component and a User Interface for the Definition and Use of Ontologies in Multimedia Information Systems”. In **Proc. of AvivDiLib** 2005, pp. 99-102, Cortona, Italy, May 2005.
- Tsinaraki C., Polydoros P., Kazasis F., Christodoulakis S., “Ontology-based Semantic Indexing for MPEG-7 and TV-Anytime Audiovisual Content”. In **Special issue of the Multimedia Tools and Application** Journal , August 2005.
- A. Karanastasi, S. Christodoulakis: "OntoNL: An Ontology-based Natural Language Interface Generator for Multimedia Repositories", **Proc of AVIVDiLib'05**, May 2005
- A. Karanastasi, F. Kazasis, S. Christodoulakis: "A Natural Language Model and a System for Managing TV-Anytime Information in Mobile Environments", **Special Issue on ACM International Journal of Personal and Ubiquitous Computing**, Volume4, 2005
- F.J. Seinstra, C.G.M. Snoek, D. Koelma, J.M. Geusebroek, and M. Worring. "User Transparent Parallel Processing of the 2004 NIST TRECVID Data Set", **Proceedings of the 19th International Parallel & Distributed Processing Symposium** (IPDPS 2005), Denver, Colorado, USA, April 4-8, 2005.
- G. Tardini, C. Grana, R. Marchi, R. Cucchiara, "Shot Detection and Motion Analysis for Automatic MPEG-7 Annotation of Sports Videos" in **Proceedings of the 13th International Conference on Image Analysis and Processing** (ICIAP 2005), Cagliari, Italy, pp. 653-660, 6-8 Sept, 2005
- R. Cucchiara, C. Grana, G. Tardini, "Shot Detection for Formula 1 Video Digital Libraries" in **Proceedings of AvivDlib**, Cortona, Italy, pp. 131-140, 4-6 May, 2005
- C. Grana, G. Tardini, R. Cucchiara, "Adaptation and Annotation of Formula 1 Sport Video" in **Proceedings of 1st Italian Research Conference on Digital Library Management Systems**, Padova, Italy, Jan 28, 2005

Publications 2005 (cont.)

- J.M. Geusebroek and F.J. Seinstra. "Object Recognition by a Robot Dog Connected to a Wide-Area Grid System". **Proceedings of the International Conference on Multimedia & Expo** (ICME 2005), Amsterdam, The Netherlands, July 6-8, 2005.
- C.G.M. Snoek, D. Koelma, J. van Rest, N. Schipper, F.J. Seinstra, A. Thean, and M. Worring. "MediaMill Video Analysis and Search Engines". **Proceedings of the International Conference on Multimedia & Expo** (ICME 2005), Amsterdam, The Netherlands, July 6-8, 2005.
- C.G.M. Snoek, M. Worring, J.M. Geusebroek, D. Koelma, and F.J. Seinstra. "On the Surplus Value of Semantic Video Analysis beyond the Key Frame". **Proceedings of the International Conference on Multimedia & Expo** (ICME 2005), Amsterdam, The Netherlands, July 6-8, 2005.
- M. Bertini, A. Del Bimbo, R. Cucchiara, C. Torniai, " Video Annotation with Pictorially Enriched Ontologies", **Proceedings of International Conference on Multimedia and Expo** (ICME), Amsterdam (NL), July 6-8, 2005
- M. Bertini, A. Del Bimbo, R. Cucchiara, C. Torniai, "Ontologies enriched with visual information for video annotation", **Multimedia and the Semantic Web Workshop** (ESWC), Heraklion (GR), May 29, 2005
- M. Bertini, A. Del Bimbo, W. Nunziati, “Soccer Videos Highlight Prediction and Annotation in Real Tim", **13th International Conference on Image Analysis and Processing** (ICIAP), Cagliari (I), September 6-8, 2005 (Springer LNCS)
- M. Bertini, A. Del Bimbo, W. Nunziati, “Players Identification in Sports Videos”, **Proc. Of AVIVDiLib'05**, Cortona, Italy,May 4-6, 2005
- C. Grana, G. Tardini, R. Cucchiara, "MPEG-7 Compliant Shot Detection in Sport Videos" on **Proceedings of IEEE International Symposium on Multimedia** (ISM2005), Irvine, California, USA, December 12-14, 2005

Current and future work

- Activities over the June 2005-June 2006 JPA2 period:
  1. Ontology definition: completed
  2. Visual features extraction: completed
  3. Automatic detection of relevant semantic and visual concepts (partially completed)
  4. Automatic annotation of video clips based on PE ontologies (under development)
  5. Analysis of computational aspects, experiments and tests (under development)

	Jun 05	Jul 05	Aug 05	Sep 05	Oct 05	Nov 05	Dec 05	Jan 05	Feb 06	Mar 06	Apr 06	May 06
1)												
2)												
3)												
4)												
5)												

Current and future work (cont.)

- Extension in JPA3:

- Extension with other pictorial concepts (e.g. objects, moving objects, color shapes)
- Using natural languages for defining concepts and querying
- Integration of all works in a unique tool
- Interfaces for interactive annotation
- User-based tests
- Generalization for other Video DLs (e.g. Video Art DLs, broadcast TV)