BMVA symposium on "Security and surveillance: performance evaluation"

**Visor**

**Video Surveillance Online Repository**

Roberto Vezzani and Rita Cucchiara

Imagelab – Information Engineering Department

University of Modena and Reggio Emilia, Italy
Aims of the Visor Project

• Gather and make freely available a repository of surveillance video footages

• Together with the videos, the repository contains metadata annotation, both manually annotated as ground-truth and automatically obtained by video surveillance systems

• Create an open forum to exchange, compare and discuss problems and results on many video surveillance and retrieval projects.
VISOR is within the VIDIVIDEO project

Interactive semantic video search with a large thesaurus of machine-learned audio-visual concepts

PARTICIPANTS:
• Intelligent Systems Lab (Amsterdam)
• Informatics and Telematics Institute (Greece)
• Institute for Systems and Computer Engineering (Portugal)
• Centre for Vision, Speech and Signal Processing (University of Surrey UK)
• Media Integration and Communication Centre (Università degli Studi di Firenze)
• Computer Vision Center (Barcelona)
• Sound & Vision, the Dutch national audio and video archive (The Netherlands)
• FDR, Fondazione Rinascimento Digitale (Firenze)
Visor contains a set of videos grouped into categories.
Video Surveillance Online Repository

Visitor contains a set of videos grouped into categories:

- 6 videos for shadow detection
- 14 videos for smoke detection
- 20 videos for consistent labeling problems on a set of partially overlapped cameras
- 10 videos of fall detection
- And much more…
Visor contains a set of **videos** grouped into **categories**

For each video....
Video details

- Screenshot
- Video information (title, description, size, compression, etc…)
- Metadata (Author, location, date, keywords, etc…)
- Copyright statement
- Flash compressed version
  - (java flash player)
Video details

- Screenshot
- Video information (title, description, size, compression, etc...)
- Metadata (Author, location, date, keywords, etc...)
- Copyright statement
- Flash compressed version
  - (java flash player)
- Camera information
- Keywords
- Related files
- Annotations!
Temporal / semantic clip segmentation

- Each video can be segmented into clips
- Clip level web interface
- Clips of a video:
People, fixed and mobile objects (e.g., trees, vehicles, buildings, and so on…)

Actions by or interaction between people; events

Metadata (e.g.: author of the video, camera information, calibration data, location, date and time, and so on…)
Set of surveillance classes and concepts

<table>
<thead>
<tr>
<th>Class</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person</td>
<td>PhysicalObject</td>
</tr>
<tr>
<td>BodyPart</td>
<td>PhysicalObject</td>
</tr>
<tr>
<td>GroupOfPeople</td>
<td>PhysicalObject</td>
</tr>
<tr>
<td>MobileObject</td>
<td>PhysicalObject</td>
</tr>
<tr>
<td>ActionByAPerson</td>
<td>Action/Event</td>
</tr>
<tr>
<td>ActionByPeople</td>
<td>Action/Event</td>
</tr>
<tr>
<td>ObjectEvent</td>
<td>Action/Event</td>
</tr>
<tr>
<td>GenericEvent</td>
<td>Action/Event</td>
</tr>
<tr>
<td>Video</td>
<td>Context</td>
</tr>
<tr>
<td>Clip</td>
<td>Context</td>
</tr>
<tr>
<td>Location</td>
<td>Context</td>
</tr>
</tbody>
</table>

List of “IS-A” concepts (static concepts)
e.g.: male, girl, civilian, soldier

LSCOM and MediaMill concepts are included

List of “HAS-A” concepts (dynamic concepts)
e.g.: position, contour, name, color, appearance
XML based
- **Config and data are separated**
- Widely diffused
- Easy to use
- Clear
- Documentation
- Compact
- Extensible
- Annotation Tool (VIPER-GT)
Annotation Format – Viper

- XML based
- **Config** and **data** are separated

- **Annotation Tool** (VIPER-GT)
Viper-GT – a free annotation tool

• A free online tool to generate annotation data for videos

• Written in java (high portability)

• Visor provides the configuration file for VIPER-GT

• Visor can import annotation data generated with Viper-GT
Annotation

• Two different types:
  – Ground truth annotation (manually generated)
  – Automatic annotation (i.e., system outputs)

• Used for:
  – Video search within the repository
  – Ground truth annotation data for training or evaluation purposes
  – Automatic annotation for comparisons and challenges

• To this aim we want to introduce performance metrics and automatic evaluation modules (like ETISEO)
Concept-based video retrieval

Search by Concept: Adult

<table>
<thead>
<tr>
<th>Video</th>
<th>Screenshot</th>
<th>Author</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>cam1_120405A1</td>
<td><img src="image1.png" alt="Screenshot" /></td>
<td>v_amorus</td>
<td><img src="image1.png" alt="Annotation" /></td>
</tr>
<tr>
<td>cam1_120405B</td>
<td><img src="image2.png" alt="Screenshot" /></td>
<td>v_amorus</td>
<td><img src="image2.png" alt="Annotation" /></td>
</tr>
<tr>
<td>cam2_120405B</td>
<td><img src="image3.png" alt="Screenshot" /></td>
<td>v_amorus</td>
<td><img src="image3.png" alt="Annotation" /></td>
</tr>
<tr>
<td>cam3_120405A1</td>
<td><img src="image4.png" alt="Screenshot" /></td>
<td>v_amorus</td>
<td><img src="image4.png" alt="Annotation" /></td>
</tr>
</tbody>
</table>
Privacy issues

- Restricted access granted only to registered users
- User registration requires moderator grant
- Video processing in order to make it privacy compliant
- “Research only” videos
- Responsibility of the uploader of the video content (like Youtube)

- **Collaboration with the Law faculty of Modena**
- **CRIS (Inter department security research center – Modena - Italy)**
Privacy issues

Automatically blurred faces using Viola&Jones face detector
User group and Forum

- User group of researchers working on video surveillance topics
- Custom Forum Topics on
  - Generic video surveillance issues
  - The VISOR system (e.g., call for videos, comments, suggestions, etc…)
  - One topic for each video in order to share comments and votes
We invite you to

• Contribute to the growing and population of both the video repository and the annotation set
• Join to the community using the forum
• Sending us comments and suggestions…

http://imagelab.ing.unimore.it/visor

Ing. Roberto Vezzani
Dipartimento d’Ingegneria dell'informazione, Modena
roberto.vezzani@unimore.it
Tel +39-059-2056270

Prof. Rita Cucchiara
Dipartimento d’Ingegneria dell’informazione, Modena
rita.cucchiara@unimore.it
Tel +39-059-2056136
Thank you for your attention!!

Visor
Video Surveillance Online Repository

http://imagelab.ing.unimore.it/visor

Ing. Roberto Vezzani
Dipartimento d’Ingegneria dell’informazione, Modena
roberto.vezzani@unimore.it
Tel +39-059-2056270

Prof. Rita Cucchiara
Dipartimento d’Ingegneria dell’informazione, Modena
rita.cucchiara@unimore.it
Tel +39-059-2056136
System Features – Storage

• Video
  – Different video formats are supported
  – Set of tags/keywords
  – Metadata (*authors*, *creation date*, …)
  – Camera information
    • *Camera type and model, Infrared, static/moving camera, calibration information* ⚡

• Annotations
  – Stored in the same DB
  – Two categories: automatic generated and manual ground-truth
  – Completely customizable structure
System Features – Management

• Video
  – Video preview (flash)
  – Single Screenshot
  – Summary (Set of keyframes)
  – Protection based on User level
  – Query by keyword, concept, or metadata

• Annotation
  – Import/export in Viper format (MPEG7 and custom formats are in progress)
  – Partial annotation selection is allowed
  – Merge of different annotations (in progress)
End