Intelligent Video Surveillance

Rogerio Schmidt Feris
Research Scientist, IBM Master Inventor
IBM T. J. Watson Research Center
http://rogerioferis.com
IBM Smart Vision Suite (SVS)

**Video Capture/ Encoding & Management**
- DVR - records & streams video

**Real-time alerts**
- Perimeter violation
- Tailgating attempt
- Red car on service road

**User driven queries**
- Find red cars
- Find tailgating incidents involving this person

**Analytics & Framework**
- Watches the video for alerts & events
  - Analytics modules:
    - Object tracking and classification
    - Face capture and recognition
    - License Plate Recognition
    - Many others
  - Gathers event meta-data & makes it searchable
  - Provides plug and play framework for analytics

**Sensors & Transactions**

**IBM Smart Vision Suite (SVS)**
Mode of Operation: Real-time Alerts

Examples of user configurable real-time alerts

Tripwire
- Triggers on the cat crossing the blue line

Directional Motion
- Triggers on right-turns, when the cars move in the direction of the arrow

Removed Object
- Triggers when object outlined in blue is removed from its position
Mode of Operation: Search After the Fact

→ “Show me all large vehicles with yellow color that crossed this road in the past 5 days” [Finds DHL delivery trucks]

→ “Show me all events with duration greater than 30 seconds” [Finds people loitering]
Technology

- Traditional: Blob-based systems

Frontiers in the field:

- Detector-centric video analytics: shifting from blobs to objects

- Attribute-centric video analytics: shifting from “naming” to “describing”
Traditional Pipeline: Blob-based Analytics

Background Subtraction: Moving Object Detection (Blobs)

Most existing smart surveillance systems in the market rely on blob-based video analysis. They are efficient and work well in low-activity scenarios.
Blob-based Analytics: Limitations / Challenges

- Dealing with Crowded Scenes
  - Objects close to each other are clustered into a single blob

- Environmental Conditions
  - Rain, snow, reflections, and shadows cause spurious blobs
Shifting from blobs to object-centric architectures

- Object Detection
- Object Tracking
- High-Level Processing

Sliding Window Approach

Training examples

Feature extraction

Car/non-car Classifier

e.g., [A. Krizhevsky et al, NIPS 2012] [P. Felzenszwalb et al, PAMI 2010]
Object-Centric Video Analytics

Car Detection in Crowded Scenes

Training Data is the Oil!

- IBM vehicle detector utilizes a huge training dataset - nearly **one million** images (50+ cameras)!

- Different weather conditions
- Multiple vehicle poses
Synthetic Data Helps - Occlusion Generator
Monolithic Detector versus Multiple Models

[Feris et al, IEEE Transactions on Multimedia, 2012]
Object-Centric Video Analytics: Limitations / Challenges

- Detector accuracy: dealing with appearance variations
  - Different object poses, lighting changes, etc.

- Detector efficiency / cost
  - State-of-the-art approaches usually run at low frame rates

- How many object classes are needed?
Attribute-Centric Video Analytics

Shifting from “naming” to “describing”

- Focus on modifiers rather than (or in addition to) nouns
- Attributes: shared properties among objects
Suspect Description Form

PENNSYLVANIA CAPITOL POLICE
SUSPECT DESCRIPTION

<table>
<thead>
<tr>
<th>SEX</th>
<th>RACE</th>
<th>AGE</th>
<th>HEIGHT</th>
<th>WEIGHT</th>
<th>TYPE OF WEAPON</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>HAIR/FACIAL HAIR</th>
<th>GLASSES (type)</th>
<th>TATTOOS</th>
<th>COMPLEXION</th>
<th>SCARS/MARKS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HARRISBURG EMERGENCY
DIAL 1-911
POLICE FIRE MEDICAL
NON-EMERGENCY
717-787-3199

DON'T HANG UP

FACIAL APPEARANCE

Hair Style
Wrinkles
Cheeks (Full or Sunken)
Shapes of Eyebrow
Ear Size and Shape
Mustache or Beard
Chin Clefts

Hair Texture
Size & Shape of Eye

Shape of Nose
Sideburns
Mouth & Lips
Neck & Adams Apple

WRITE BELOW SPECIFIC FACIAL DETAILS ONLY WHAT YOU DEFINITELY REMEMBER

WHAT DID SUSPECT SAY?

AUTO MAKE MODEL, COLOR
LICENSE NUMBER DIRECTION OF ESCAPE TIME OF DEPARTURE

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
System Overview

Video from camera

Analytics Engine

Background Subtraction → Face Detection & Tracking → Attribute Detectors → Database Backend

Suspect description form (query specification)

Search Interface

Result – thumbnails of clips matching the query
Suspect #1

Beard
Blue Shirt
Sunglasses

See video at: http://rogerioferis.com/demos/PeopleSearch_BlurredFaces.wmv
Privacy Concerns

Video Re-rendering according to different levels of authorization

Meta-data Masking
Conclusions and Future Directions

- Smart Surveillance Technology: Shift from Blob-based to Object-Centric and Attribute-Centric Analytics

- Computer Vision Meets Big Data: Robust object classifiers trained with large amounts of data

- What will a surveillance camera look like in the next 10 years?
Thank You !

See more at http://rogerioferis.com