VAP / VAT: Video Analytics Platform and Testbed for testing and deploying Video Analytics

SPIE “Defense, Security, and Sensing” Conference
Track on Visual Analytics for Homeland Defense and Security
5 - 9 April 2010, Orlando

Dr. Dmitry O. Gorodnichy
Video Surveillance & Biometrics Section
Science and Engineering Directorate
1. The evolution of Video Surveillance
   - What is Video Analytics (VA) and what it can do
   - VA Technology Readiness Level / Customization req.-d

2. VA deployment challenges and 3-Phase Roadmap
   - Introducing Video Surveillance & Biometrics Section
   - VAP /VAT concepts and architecture
     - EventCapture & EventBrowser: software & results
     - Doing VA evaluation (inc. FaceRec): metrics & demos
   - Running Pilots in the “Field”

3. Next steps
What is Video Analytics?

- **Video Analytics**: Computational Analysis of Video Data and Extraction of Intelligence from Video. (aka Intelligent Video)
- Better quality, more affordable

- **IP cameras**
- **Digital**
- **Analog**

- Video Technology (VT) of the 1990’s was primarily concerned with video capture (Cameras and Recorders)
- Today, VT is Video Capture + Video Analytics

- Animated image example:
  The entire 17:00-24:00 activity is summarized into a few annotated snapshots.
Without Video Analytics - as it is now

Area of Unexploited Information!
With Video Analytics - what we can do

With ACE Surveillance™ software that we have developed within Video Recognition Systems project at National Research Council of Canada (2000-2007) - www.videorecognition.com/vrs.html

DATA COLLECTED

Area of Unexploited Information!

DATA USED
Monitoring Tasks Performed by Human (Status quo)

Two modes: a) real-time monitoring, b) post-event analysis
Monitoring Tasks Performed by Human & Software

Two modes: a) real-time monitoring, b) post-event analysis

Main condition - Open Architecture:
To be able to tap into (video signal) input and (data) output.
ACE Surveillance™ VA Pilot (2006-2008) [Gorodnichy, NATO-2008]

Outdoor, wireless, eye-level
Outdoor, webcam, overview
Indoor with sunlight, CCTV

Indoor w/o sunlight, CCTV

Camera / setup
Annotated CES
ACE daily summarization

Outdoor, wireless, eye-level
Outdoor, webcam, overview
Indoor with sunlight, CCTV
Phase 1. Building Business, Infrastructure foundations; R&D capacity

- From Knowing the Art of Possible to Making it Possible
  - Dealing with Stereotypes / Misconceptions
  - Not to over-estimate or under-estimate what VA can do
  - Technology Readiness Level vs. Customization required

Phase 2. Developing VA solution: Technological Challenges

- R&D programming + Knowledge of Operational Tasks
- Dealing with “Closed Architectures”
- Selecting/Building a solution

Phase 3. Piloting & deploying the solution in the “field”

- Knowing Clients needs & Educating/Training the Client
- Customizing, stress-testing (in Mock-up and real setups)
What Video Analytics is NOT?

- “Brightness change detection” marketed by industry as “motion-detection”, which is NOT.

- NOT a “Magic bullet”:
  - Just as with ANY image recognition (inc. Biometrics), there will be “False Hits” and “False Misses”.
    - However, their Rates can be minimized to acceptable for operational needs – by evaluation and customization.

- NO “one size fit all” solution (esp. in Non-cooperative scenarios)
  - Different VA codes required for each setup, environment, task.
    - However, experts may use the same library to write these codes.

- “High resolution / quality” do NOT assume “high intelligence”.

- It is NOT expensive with proper (unique) skills and planning.
  - In fact, it (significantly) reduces the entire cost operation, though the optimized equipment build-up and efficient data analysis.
1. Video Data Capture
2. Transfer
3. Storage
4. Protection and security

5. Integration with other sensors / software
   - Motion, heat sensor, audio, Video Analytics

6. Video data management
   - Indexing, visualization, retrieval of data
   - Data = video + associated Meta-data (Annotations) obtained with Video Analytics

7. Video analytics for automation and filtering:
   - Real-time event detection / recognition
   - Analysis of archived video data

Total cost = Hardware + Software/Programming/Testing
Video Analytics expertise minimizes the cost of both components.
Traditionally performed by Humans, many of these Monitoring Tasks can now be facilitated with VA software

<table>
<thead>
<tr>
<th>TYPE 1: Real-time monitoring tasks</th>
<th>Customization, testing req.-d</th>
<th>Technical readiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1* - &quot;Face extraction/tagging&quot;</td>
<td>Little</td>
<td>5</td>
</tr>
<tr>
<td>2* - &quot;Wrong direction detection (Run-away alarm)&quot;</td>
<td>Little</td>
<td>5</td>
</tr>
<tr>
<td>3 - &quot;Loitering alarm&quot;</td>
<td>Major</td>
<td>4</td>
</tr>
<tr>
<td>4 - &quot;Object-left behind or abundant object alarm&quot;</td>
<td>Major</td>
<td>4</td>
</tr>
<tr>
<td>5 - &quot;Tripwire (trespassing) alarm&quot;</td>
<td>Little</td>
<td>5</td>
</tr>
<tr>
<td>6 - Other events (door opening, car parking etc) alarm</td>
<td>Major</td>
<td>4</td>
</tr>
</tbody>
</table>

**General Tracking / Detection of people in multiple streams**

**TYPE 2: Post-Event (Archival) monitoring tasks**

<table>
<thead>
<tr>
<th>TYPE 2: Post-Event (Archival) monitoring tasks</th>
<th>Customization, testing req.-d</th>
<th>Technical readiness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - Summary of detected events &amp; statistics (trends)</td>
<td>Little - Medium</td>
<td>5</td>
</tr>
<tr>
<td>2 - Searching for a object/person in stored streams</td>
<td>Little - Major</td>
<td>5</td>
</tr>
</tbody>
</table>

**General Summary / Search in unstructured environment**

**Special case tasks**

| LPR (License Plate Recognition) | None | 5 |
| Face Recognition | Little-Medium | 1-4 |
Building R&D Capacity

- Best synergy and impact is achieved by treating Video Surveillance together with Biometrics.
  - … because most Biometrics are image-based.
  - Same unique set of skills:
  - Face Recognition – where Biometrics meets Surveillance.

- VSB Section consolidated all IP & PR expertise in one place
  - To support both CBSA portfolios: in Biometrics and Video Surveillance
  - To become the prime GoC R&D center in Video Analytics & Biometrics
Evolution of Video Surveillance & Biometrics

Towards more collectable, unconstrained environments

Surveillance

Biometrics

Automated identification of people from video
Automated detection of events
Remote monitoring only

2020
2010
2000
1990
1980
1970

Faces from video
Soft biometrics (audio/visual)
Multi-modal fusion

Towards collecting intelligence / evidence

Level of automation and intelligent processing

D. Gorodnichy © 2009
Three foci of our R&D work:

Our objective: To find what is possible and the best

- in Video Analytics, Biometrics, Face Recognition
- for LAND and AIRPORT Points of Entry (POE)

to be in a position to build solutions to CBSA & OGD.

- Focus 1: Evaluation of Market Solutions
- Focus 2: In-house R&D
- Focus 3: Live Tests/Pilots in the Field

See also:
- “C-BET: Comprehensive Biometric Evaluation Toolkit” in SPIE Biometric Technology for Human Identification track – today at 4pm
- Federal Dept’s workshops on Video Technology for National Security (VT4NS) - www.videorecognition.com/vt4ns
1. Different tasks and scenarios require different VA codes to be written, and the customization of the VA codes can be properly done only by a Video Recognition expert, which in many cases is not available in an operation-driven agency. At the same time, VA customization requires strong knowledge of operational tasks as well as constant communication with the involved regions. As a result, a solution coming from outside is often very expensive and in many cases not reliable.

2. IP-cameras, which are now extensively replacing analog cameras, contain vendor-specific coding/encoding mechanisms, and getting a video-feed from these cameras requires customization in programming codes specific to each vendor. Furthermore, while some vendors provide functionality to perform direct capture of the video-feed from their cameras through the use of a dedicated SDK (Open Architecture cameras), others do not (Closed Architecture cameras).

3. Selecting a good (or the best) VA product requires the testing of different products for the purpose of measuring and comparing their performance for a given task. Most agencies cannot afford to perform such testing and have to fully rely on the vendors’ claims instead.
**VAP / VAT definitions & concepts**

**Definition:** *Event of interest* \((E_i)\) is an instance when certain conditions related to what is observed in video are met.

**Definition:** *Details of interest* \(D\{ I_a, M_0, \{I_j\}, \{M_j\} \} (E_i)\) is set of static images & associated metadata (annotations) extracted and saved from video when an *Event of Interest* happens

- one image with annotation \((I_a)\) is chosen to represent the Event

**VAP task:** To replace a continuous video-stream with a list of Details \(\{D_j\}\) that can be efficiently browsed and analyzed – by using a Video Analytic module that operates on the video-stream.
VAP / VAT Architecture

Consists of two components:

- EventCapture: uses VA to detect *Event* and extract *Details*
- EventBrowser: Convert *Details* into searchable web-applet
VAP / VAT Architecture

VAP / VAT: EventCapture + EventBrowser

Real-time mode of operation

Archival mode of operation

Video (Tb)

DETAILS (Gb)
Tapping into Closed & Open Architecture

Open Architecture
- Allows to directly access the video-feed
- Allows to control the video-feed quality/size

→ VAP IPCamCapture

Closed Architecture
- Does not …
- Does not …

→ VAP ScreenCapture
frame = Source->getCurrentFrame()  // Obtain frame I(t)
VAModule->processFrame(frame)     // Process I(t) and update the VA knowledge
if VAModule->eventConditionsMet()
    E++                            // internal variables triggered an event
    VAModule->saveDetails()        // All snapshots and metadata saved

Three steps:
1. Adding video-stream
2. Selecting VA module
3. Configuring VA module
VAP EventBrowser

- All Events displayed represented by Thumbnails
- Each Event can be expanded
- Event can be viewed by attributes
Comparing VA-based solutions to one another

- Using *EventBrowser* statistics module
Using VAP/VAT for VA evaluation

- Comparing VA-based solution to Manual screening:
  - Count Vehicles in 24/7 out-door settings.

<table>
<thead>
<tr>
<th></th>
<th>VAS</th>
<th>MS</th>
<th>Difference</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Events</td>
<td>98</td>
<td>105</td>
<td>7</td>
<td>6.67</td>
</tr>
<tr>
<td>Left</td>
<td>51</td>
<td>53</td>
<td>2</td>
<td>3.77</td>
</tr>
<tr>
<td>Right</td>
<td>47</td>
<td>52</td>
<td>5</td>
<td>9.62</td>
</tr>
<tr>
<td>Time (mins)</td>
<td>30</td>
<td>360</td>
<td>330</td>
<td>91.67</td>
</tr>
</tbody>
</table>
Once VAP is installed in the region:

- Client sends a request: describing what “EVENT” and “DETAILS” they want to detect (in plain English language)
- VSB provides a VA solution (in VAP script language):
  1. a) in-house built or b) by customizing a 3rd-party VA code (via SDK)
  2. + testing: with a) actual video-feed from regions or b) mock-up settings
Conclusions and Next Steps:

- Deploying VA in operations is shown to be possible
  - Without disruption to normal workflow
  - Cost-effective
  - Objectively measured performance improvement

- Methodological approach for selecting/testing VA is described

- In the 3-Phase Road-map for successful VA deployment
  - We have passed Phase 1 (Capacity) and Phase 2 (Solutions)
  - Moving into Phase 3 (Pilots & Outreach)

- Enables Testing of Face Recognition with existing CCTV video

- Next VT4NS workshop (June 3, 2010) themes:
  - “Deploying Video Analytics”, and “Faces in Video”.

Contact: Dmitry.Gorodnichy@cbsa.gc.ca