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

Dmitry O. Gorodnichy† and Elan Dubrofsky
Video Surveillance and Biometrics Section
Science and Engineering Directorate, Canada Border Services Agency
14 Colonnade Rd, Ottawa, ON, Canada, K2E 7M6

ABSTRACT

Deploying Video Analytics in operational environments is extremely challenging. This paper presents a methodological approach developed by the *Video Surveillance and Biometrics* Section (VSB) of the Science and Engineering Directorate (S&E) of the Canada Border Services Agency (CBSA) to resolve these problems. A three-phase approach to enable VA deployment within an operational agency is presented and the *Video Analytics Platform and Testbed* (VAP/VAT) developed by the VSB section is introduced. In addition to allowing the integration of third party and in-house built VA codes into an existing video surveillance infrastructure, VAP/VAT also allows the agency to conduct an unbiased performance evaluation of the cameras and VA software available on the market. VAP/VAT consists of two components: *EventCapture*, which serves to Automatically detect a "Visual Event", and *EventBrowser*, which serves to Display & Peruse of "Visual Details" captured at the "Visual Event". To deal with Open architecture as well as with Closed architecture cameras, two video-feed capture mechanisms have been developed within the *EventCapture* component: *IPCamCapture* and *ScreenCapture*.

Keywords: Video Surveillance, Video Analytics, Cameras, CCTV, Face Recognition in Video

1. INTRODUCTION

As a result of the increasingly growing demand for security, many countries have been deploying *Video Surveillance Systems* (VSS) as an important tool for enhancing preventive measures and aiding post-incident investigations. Within the Canadian government, many federal departments heavily use VSSs including the Canada Border Services Agency (CBSA) who sees video surveillance as a key technological element in protecting the country's borders [1-4,7].

VSSs, which are often referred to as Audio/Video Systems or CCTV Systems, are used in one of two modes of operation: a) *Live mode* (or Real-time monitoring), and b) *Archival mode* (or post-event analysis through recordings). The Live mode of operation can be performed either through *Active monitoring*, which involves trained personnel who watch video streams at all times or through *Passive monitoring* which involves employees who watch video streams in conjunction with other duties.

While evaluating the utility of VSS, it has been realized and emphasized that currently deployed surveillance system and protocols are not fully efficient for either of described modes of operation [1,2,5,7]. In real-time monitoring mode, the problem is that an event may easily pass unnoticed due to false or simultaneous alarms and a lack of time required to rewind and analyze all potentially useful video streams. In archival mode, video data storage and manageability is the problem that complicates the efficiency of post-incident investigation the most. Due to the temporal nature of video data, it may take very long for a human to analyze it. Also, video files may be stored on different media and/or compressed using different proprietary algorithms, making it very difficult to share within the agency.

The only way to resolve these problems is to use automated video analysis tools, commonly referred to as Video Analytics, Video Recognition Systems or Intelligent Video.

The deployment of Video Analytics by agencies in operational environments however is extremely challenging.

---

* In Proceedings of SPIE Volume 7667: Conference on Defense, Security, and Sensing. - DS226: Visual Analytics for Homeland Defense and Security track. 5 - 9 April 2010, Orlando.
† Corresponding author (Email: dmitry.gorodnichy@cbsa-asfc.gc.ca, Tel: (613) 954 3785, Fax: (613) 960-5184).

- First, 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. At the same time, VA customization requires strong knowledge of the environment through constant dialog with those running the operations. As a result, a solution coming from outside is often very expensive and in many cases not reliable.
- Second, 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).
- Third, 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.

There are three phases in the roadmap for the successful deployment and use of VA in operational environment:

1) building the infrastructure foundation and the required R&D capacity
2) developing VA solutions (based on the in-house R&D expertise and/or the solutions available on the market)
3) testing and deploying VA solutions in operations.

In the following, we present how these challenges are being resolved by CBSA within the recently created VSB section. After outlining the main features of the video surveillance evolution (Section 2) and properly defining VA (Section 3), we define the main principles behind VAP/VAT technology and describe VAP EventCapture and EventBrowser software (Sections 4). Examples of using VAP/VAT for VA evaluation and selection are then presented (Section 5). Conclusions and future work wrap up the paper.

2. EVOLUTION OF VIDEO SURVEILLANCE

2.1 The role of cameras: Shift from remote monitoring to intelligence gathering.

In the past, cameras were used mainly for remote monitoring to serve as a tool for such security tasks as to: Deter, Detect, Respond, Investigate, and Reassure Community [1,2]. They were installed mainly in Restricted-access areas, where environment and behaviors are structured and controlled. Their deployment would normally be the responsibility of the departments who procure infrastructure and other types of security equipment.

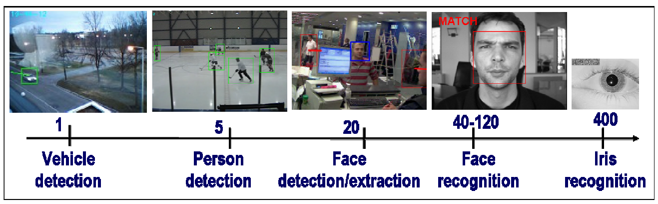
The transition from analogue cameras to IP-cameras has offered organizations three critical advantages over Closed Circuit Television (CCTV) systems used in the past:

1) ability to view cameras from distant locations - cameras in a region can be viewed from a centralized location in the headquarter and in a testing lab, without disrupting the normal workflow in the field;
2) use of high-resolution cameras - whereas in the past video images were on 320x240 resolution, now most video images can be expected to be of around 1600x1200 resolution;
3) remote control of video capture content, quality, and quantity, which can be conducted manually or *using video analytics*, so that only the important visual details are captured and stored.

These advantages along with the improved computer power, network bandwidth capabilities, camera quality, and video recognition algorithms opened the way to the new era of developing and using VSS solutions. - VSS are now used on a very large scale in public places, where environment and behaviors are normally not structured or controlled, for the purpose of collecting and managing the Intelligence. As a result, the use and deployment of VSS has now become the preoccupation of the departments that define operational programs and policies and deal with intelligence gathering and management.

2.2 Evolution towards biometrics and special interest in Face Recognition from Video

Over the years, computers have become faster and automated intelligent processing has become easier. As a result, surveillance systems are now increasingly applied towards collecting intelligence and recognition of identities. Of particular interest is the phenomenon of merging Biometrics and Video Surveillance, illustrated in Figure 1b (from [9]) and the arrival of such biometric technologies as *Biometric Surveillance*, *Soft Biometrics* and *Stand-off Biometrics*. There has also been an increased demand for *Face Recognition from Video*, which is where Biometrics meets Video Surveillance and which is seen as a golden solution to many operational needs.

a) 

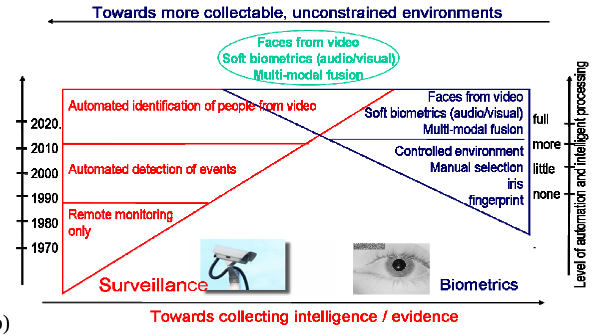
b) 

Figure 1: Video Analytics and Biometrics as two sides of the same Video Recognition problem (a) - numbers indicates the face size in pixels, and Evolution of Biometric and Video Surveillance systems: towards each other, with overlap in Face Recognition (b).

We also make a statement [10] that Video Analytics and image-based biometrics belong, in fact, to the same problem of object/person identification from video, but only applied at different scales (or distances to a person/object) as illustrated in Figure 2a. As will be discussed later, such treatment of VA helps to develop an objective methodology for evaluating VA technologies based on the concepts of False Hit / False Miss that are commonly used in evaluating Biometric solutions.

2.3 Video surveillance components and cost

The development of a VSS evolves around the following pillars: 1. Video Data Capture, 2. Transfer, 3. Storage, 4. Protection and security, 5. Video monitors, interfaces and controls, 6. Integration with other sensors / software (including motion, heat sensor, audio, Video Analytics), 7. Video data management, which includes indexing, visualization, retrieval of data (both video data and associated meta-data or annotations. 8. Video analytics for automation and filtering: both for Real-time event detection / recognition, and analysis of archived video data.

While components 1-5 can be considered technologically solved and include such parts as cameras ; encodes, decoders; transmitters and receivers; routers and multicast switches; network video recorders; storage media, the other components (6-8) are still being solved and require more exploring, evaluation, and testing.

The cost of VSS is no longer the cost of Hardware only, but is also the cost of Software, Software Tuning, and Software Testing. One of the most costly pieces of Software is Video Analytics. Being able to develop, tune and test Video Analytic Software in house could allow an organization to substantially reduce the cost of VSS, while insuring the continuity in its support, tuning and testing.

Another very important step to enable the development of task-efficient and cost-effective VSS solutions is to use *Open Architecture* camera systems, which are the systems that are modular in nature, where each module (component) of VSS can be easily replaced or added at a later time using not only a single provided (proprietary) solution. In addition to facilitating the competitive selection of the best components and the efficient upgrade of all VSS, open architecture systems also provide better support for the development of VA by allowing third party VA codes to access directly the video streams from the source cameras.

3. VIDEO ANALYTICS: DEFINITIONS AND EVOLUTION

To insure that a proper infrastructure and business foundation is set for VA deployment, it is important that Video Analytics be properly defined and explained to the stakeholders and users, and that common myths and misconceptions related to Video Analytics be properly dealt with. This has become one of the key objectives of our initial work [2,7,8] and the follow-up VT4NS'08 workshop that has been organized under the theme of "Future proofing the video technology". The main definitions and results from this work are summarized in this section.

Definition: *Video Analytics* is the Computational Analysis of Video Data conducted for the purpose of Automatically Extracting Intelligence from Video.

3.1 VA taxonomy

There are two types of Intelligence that can be extracted from video: 1) story telling or activity recognition (VERB), where temporal information is more important (less image resolution, but more continuous video streams), and 2) identity recognition (NOUN), where special information is more important (higher image resolution, but less or no video stream is required)

Generally, there are two types of recognition/identification used in intelligence gathering: 1) positive identification that proves that a VERB or NOUN has happened, and 2) negative identification that proves that a VERB or NOUN has not happened. While VA can be used for positive identification, it is important to appreciate that it cannot be used for negative identification.

Every environment is different, especially outdoors, and the following are seven factors that require VA solution customization: 1. Angles/Field of view, 2. Illumination : brightness, shadows, 3. object resolution in image (close vs. far), 4. Camera quality: dynamic range, exposure, aperture, 5.motion type (slow vs. fast) 6. Object position – with respect to the cameras and other objects (occlusions), 7. Object velocity (blur).

Additionally, different solutions may be required for different types of cameras, of which the following four types are available: 1. night-vision (IR or infra-red), 2. panoramic (fish-eye lense) cameras, 3. long focus (tele-photo) vs. short focus (wide angle), 4. PTZ vs. stationary.

3.2 VA Technology Readiness Level

Traditionally performed by Humans, many Monitoring Tasks can now be facilitated with VA software. As an illustration Table 1 lists some commonly contemplated uses of VA along with the level of required VA code customization and the technology readiness level. It is important to emphasize that while many VA modules may be available for these video recognition tasks, it is the quality of recognition that matters. Such quality can be measured using the False Hit and False Miss Rates, provided that it is possible to apply different modules to the same video data. This is further explored in Video Analytic Testbed, described in later sections.

Table 1: Video Analytics Technology Readiness Levels
(5 – ready, 4 – requires Evaluation only, 3/2– requires further Refining/Exploration, 1 – not yet ready)

TYPE 1: Real-time monitoring tasks	Customization required	Technical readiness
Face extraction/tagging	Little	5
Camera tampering detection	Little	5
Wrong direction detection (Run-away alarm)	Little	5
Loitering alarm	Major	4
Object-left behind or abundant object alarm	Major	4
Tripwire (trespassing) alarm	Little	5
Other events (door opening, car parking etc) alarm	Major	4
General Tracking / Detection of people in multiple streams	Little - Major	1
TYPE 2: Post-Event (Archival) monitoring tasks		
Summary of detected events & statistics (trends)	Medium	5
Searching for a object/person in stored streams	Little - Major	5
General Summary / Search in unstructured environment	Major	1
Special case tasks		
LPR (License Plate Recognition)	None	5
Face Recognition	Little-Medium	1-4

4. VAP: DEFINITIONS AND ARCHITECTURE

4.1 History

In 2005-2007, the first live pilot of a deployed Video Analytic solution within the Canadian federal environment was conducted by the National Research Council of Canada. It was developed by NRC and called ACE Surveillance [5-7]. The ACE surveillance pilot demonstrated that affordable and efficient automated video surveillance is achievable – both for indoor and outdoor environments and provided a reference baseline against which other intelligent video solutions can be measured such as those coming from industry [11,12].

ACE Surveillance has shown a critical advantage that VA offer to the users compared to the "motion detection" techniques based on illumination change detection sold on the market, which is the following. VA allows one to replace video files, which are difficult to manage and analyze, with static images accompanied by associated meta-data (or annotations) that are easy to manage and analyze by using standard database browsing and search techniques.

While ACE Surveillance was based solely on custom-developed codes and worked only with USB-enabled video sources such as webcams and USB video digitizers/grabbers, VAP is designed to offer a complete solution to an organization, by enabling the agency to choose, test, and tune *any* third party video analytic software with *any* video source: from any video-camera, video-player or video-file.

4.2 Key idea

The Video Analytics Platform (VAP) is conceived and developed with the objective to enable the efficient retrieval of intelligence in video data coming from existing operational video surveillance systems. It is based on the following understanding of the nature of the Video Surveillance.

> *Even though a camera captures video data non-stop 24/7, in reality it is only at certain moments of time, when particular Events of Interest happen, when the captured video data need to be analyzed and when particular Details of Interest related to the event need to be extracted.*

For a user, Event and the Details of Interest are described using the English language. For example, events include "Door opens", "Car entered", "Face/Person seen" and details include "images of all people who entered", "car size and colour", "face name" and "time and locations of persons". In the VAP/VAT lexicon, the concepts of Event and Details of interest are defined as follows.

Definition: *Event of interest*, designated as ***E***, is an instance when certain conditions related to what is observed in video are met.

Definition: *Details of interest*, designated as ***D{ Ia, M0, {Ij}, {Mj} } (E)***, is set of static images and associated metadata (annotations) that is extracted and saved from video when an Event of Interest happens, of which one image with annotation is chosen to represent the Event.

Based on these definitions, the main VAP task is the following:

Task: To replace a continuous video-stream with a list of Details ***{Dj}*** that can be efficiently browsed and analyzed – by using a Video Analytic module that operates on the video-stream.

To accomplish this task VAP is made of two components: the *EventCapture* component, which serves to automatically detect a "Visual Event" from a number of video streams, and the *EventBrowser* component, which serves to display & facilitate the perusal of "Visual Details" captured from the "Visual Events".

The *EventCapture* component is Windows MS program that has three functions. First, it taps into existing video-stream(s). Second, it embeds a VA module(software) that will process the video-steam(s). Third, it transfers the data that is extracted by a VA module to *EventBrowser*.

The *EventBrowser* component is a web application that is responsible for two main tasks. First, it stores all detected event details in a database. Second, it prepares an html-enabled web application that allows efficient viewing and analyzing of stored events. More details follow below.

4.3 EventCapture Component

The *EventCapture* component has to work with any video-source used in the operations. Therefore to deal with Open architecture as well as with Closed architecture cameras, two video-feed capture mechanisms have been developed within the *EventCapture* component: *IPCamCapture* and *ScreenCapture*. Additionally, video can be read from a file or from USB-enabled devices.

With *IPCamCapture*, an IP camera connects to a network using an Ethernet cable and can be accessed by any computer that has access to the network. Only open architecture cameras (those that allow access to their video stream) can be processed by *IPCamCapture. With Screen Capture – EventCapture* is set to view a region of a monitor and continually take screen shots of this region to form a video feed. This is used when a closed architecture IP camera is used. As long as the camera's feed is visible on the screen, EventCapture can process the feed by means of screen capture.

This is summarized in Figure 2, which also illustrates the deficiency of using Closed Architectures with Video Analytics. It does not allow one to access the image at its full resolution, and it can also cause the delay due to video image transfer.

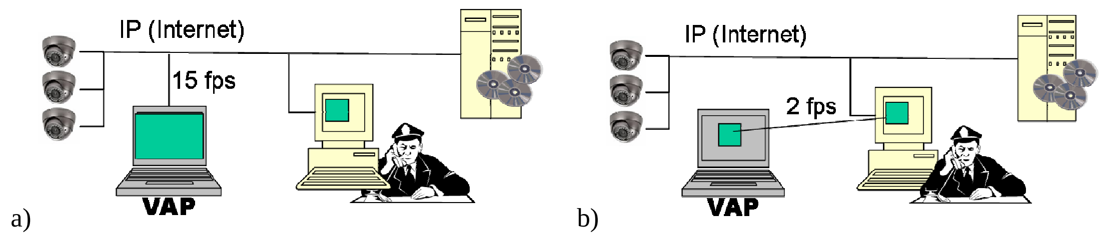


Figure 2: Deploying VAP with Open Architecture (a) and Closed Architecture (b) video surveillance systems.

The process of setting up EventCapture involves 3 steps (see Figure 3) - Step 1: Adding a video source. Step 2: Selecting a VA module, and Step 3: Configuring the settings of the VA module.

Once a video source is selected, a VA (video analytics) module must be selected to be associated with this specific video source instance.

A VA module is a program that runs video analytics on a given video source. It can be developed completely in house or can make use of 3rd party software libraries. For a VA module to be compatible with EventCapture it must contain code that specifies a) The conditions that determine whether an event has occurred and b) what details to output once an event has been detected. For example, a VA module named "Motion Detection VA module" could be programmed to continually check for object motion in the video stream. If motion is detected, the details outputted could include the frames in which the motion occurred along with the speed, location and direction of the object's motion.

A VA module will often require some settings to be configured in order to optimize its performance in a given environment. For example, a motion detection VA module would likely include a setting of how much motion is required in order for an event to be triggered.

Using pseudo-code the work of EventCapture for each Video Source with each VA module can be described as follows:

```
frame = Source->geCurrentFrame()  // Obtain frame I(t)
VAModule->processFrame(frame)     // Process  I(t) and update the VA knowledge (internal
                                  // variables,  buffers, frames etc)
if VAModule->eventConditionsMet()
   E++                            // A internal variables triggered an event
   VAModule->saveDetails()        // All snapshots and metadata representing event saved
```

We note that VA modules are used in both steps: in checking the conditions and extracting details.

The key to EventCapture that makes it so powerful is that it facilitates the use of any VA module with any type of video source. A VA module is always agnostic with regards to the type of video source that it is processing. All a VA module requires as input (aside from settings from the user) is the next frame of video to be processed at a given time. It is also important to note that the same video source can be added to EventCapture multiple times with a different VA module for each instance of the source. This is especially useful for the task of comparing the performance of different VA solutions which is discussed in the next section.

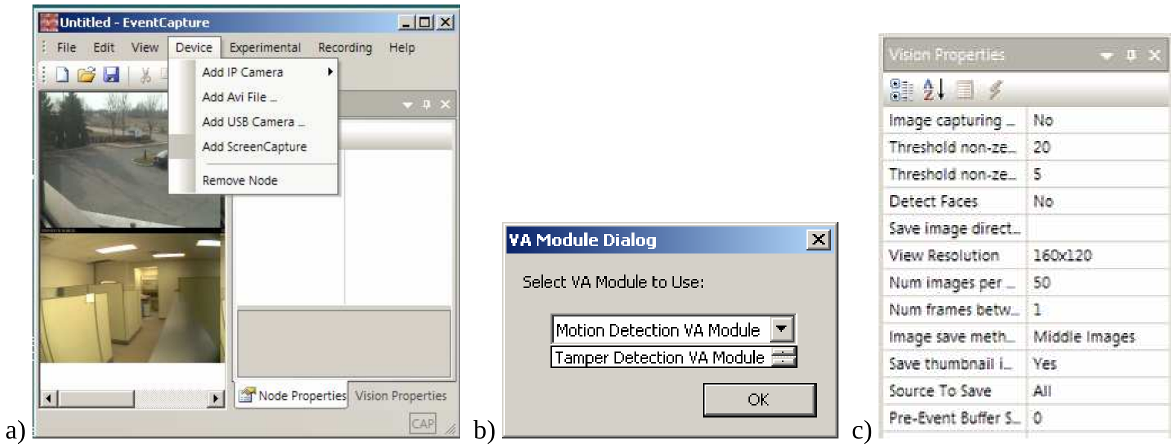


Figure 3: Three main steps of using VAP EventCapture: a) Adding a video source, b) Selecting a VA module in EventCapture, and c) Settings for a motion detection VA module.

4.4 EventBrowser Component

The *EventBrowser* component is a web application that is responsible for displaying the details of the events that are detected by EventCapture. Its prime function is convert the Event Details captured by the *EventCapture* component into *html* data readable by any Internet browser as a webpage. In order to do that, EventBrowser continuously monitors location on the hard-drive where EventCapture saves the event details. When a new event's details are found, EventBrowser adds an entry for the associated event to a database. This database is used to populate the EventBrowser web application with details of all of the Events that have been observed.

The purpose of EventBrowser is to allow a user to easily peruse the event Details of the Events that were captured by the EventCapture component. EventBrowser can be set to show all events that took place within a given time period, subject to the Event Details (eg. only show cars that were driving left). Figure 4a shows EventBrowser populated with many events. Each event is represented in EventBrowser by a rectangle that contains the time of the event and a snapshot representing the event. By clicking on the various buttons in an Event rectangle a user can browse through the thumbnail images that make up the event, view full resolution images of the event snapshots (Figure 4b), view additional details (attributes) describing the event and add additional comments that describe the event. The event rectangle also contains an "attributes" button which when pressed brings up an additional panel containing the non-snapshot Details from EventCapture (see Figure 4c). Up to 21 events can be shown on the screen at a time and a user can step through "pages" of events in order to browse through all events that are available.

EventBrowser provides functionality to view a summary of all events in a given time period. By clicking "Populate List" a user can get a list of every event that was captured ordered by the associated timestamps. This list can be loaded into a 3rd party application such as Excel for further processing.

EventBrowser also contains a statistics module (Figure 5) that can be used to view statistics of event frequencies. The user can set a time period and EventBrowser will display a graph showing the frequency of events in that period. For example, a user can select the time period to be the month of March in 2010 and the granularity to be "days" and the statistics module will show a graph representing the number of events that occurred on each day of March. Alternatively, the user can select a granularity of “hour” for a specific day to see how many events occurred in each hour of the day. These graphs can be used to gain valuable high level knowledge of operational conditions.

Using EventBrowser a user can browse through hundreds of events in a matter of minutes in order to find specific events of interest. This is orders of magnitude faster than watching a raw video stream. Once an event of interest is found the user can download the snapshots representing the event in order to provide evidence of the event in question.

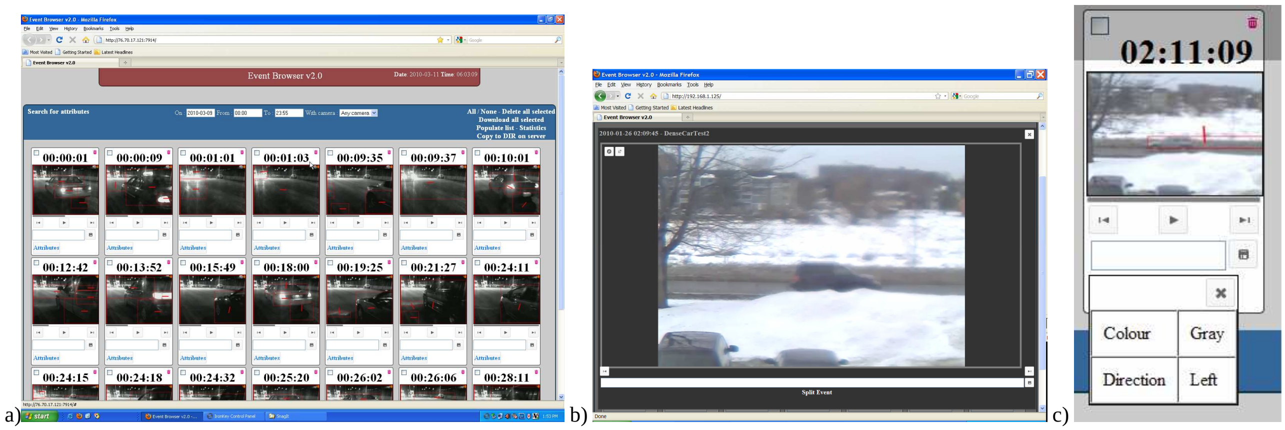


Figure 4: Key features of *EventBrowser*: a) *EventBrowser* populates the display with events represented by annotated thumbnails, b) Each thumbnail can be expanded by clicking on it to show all images associated with the event, as detected by the *EventCapture* VA module, c) Events can also be viewed/examined by associated metadata “attributes”.

5. ANALYTICS EVALUATION USING VAP/VAT

VAP can be used to test and evaluate VA modules and this is why we also refer to it as VAP/VAT (Video Analytics Testbed). By connecting several VA modules at the same time to the same live or pre-recorded video source, one can compare modules to one another. The visual statistics feature of the EventBrowser can be used to populate measurable and visual statistics (as shown in Figure 5).

In the following we provide video analytics evaluation examples performed by the VSB group using the VAT. The goal was to evaluate the performance of VA modules programmed to detect the event of a car driving by in a video stream where the camera is pointed at traffic. The outdoor environment in which these experiments were conducted as well as an representative annotated snapshot produced by a VA module are shown in Figure 4.

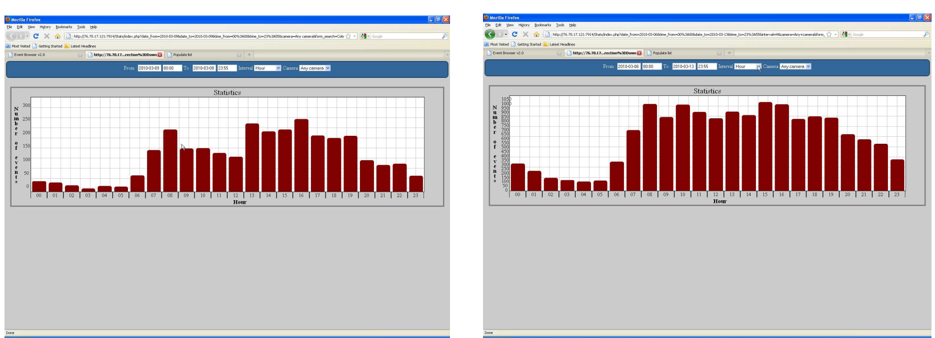


Figure 5: Using *EventBrowser* statistics module to compare twoVA modules

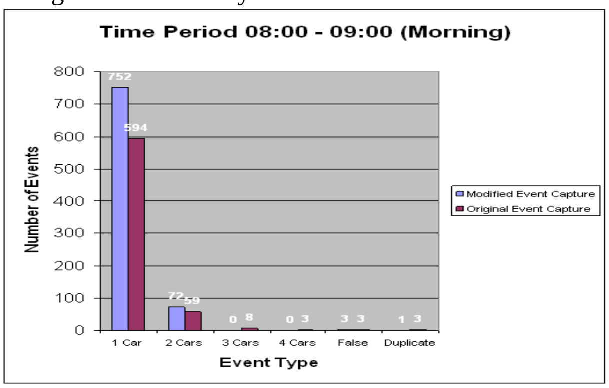
5.1 Comparing VA-based solutions to one another and to the manual screening

Two different traffic situations were considered: 1) *dense* traffic and 2) *sparse* traffic. Two different VA modules were compared to one another, each using a different algorithm for detecting “car driving by” events. The following are the steps that were taken in performing the evaluation of the two modules using VAT.:

1. Two IP cameras were set up: One pointing at a road where cars drive by at a very consistent pace (dense traffic), and one pointing at a parking lot where there are much fewer events of cars driving by (sparse traffic).
2. Four (Video Source, VA Module) pairs were added to EventCapture: (Dense Source, Module 1), (Dense Source, Module 2) , (Sparse Source, Module 1), (Sparse Source, Module 2).
3. Each (Video Source, VA Module) pair was set to run for 24 hours with each pair saving its Event Details to a different location. Also a 3rd party application was used to record the video from the two IP cameras for the 24 hour period.
4. After the 24 hour test was completed, the following evaluation tasks were performed:
    a. The number of "car driving by" events was manually counted using the recorded video for both the dense traffic and sparse traffic feeds.
    b. Each of the four sets of Event Details generated in step 3 was loaded into EventBrowser so that statistics could be generated.

At the end of this process, a report of how each VA Module performed compared both to the ground truth and to each other, is created. EventBrowser statistics module allows to generate valuable statistics such as how many events were detected in each hour. This allows us to investigate which VA module is the best for a given task.

One important thing to note is that the only step in the process that takes a significant amount of time is step 4a, i.e. the manual counting step. This step is required so that the performance of each VA module can be compared to the ground truth. Without this step the whole evaluation can be completed in the same time that the video sources are run, plus the negligible amount of time required to generate and analyze the statistics in EventBrowser.



a)

Figure 6: Two VA modules compared to one another on the same video data;

In a situation where manual counting is not feasible, VAT can still be a useful tool for VA evaluation. Even though no ground truth is available, if there are many VA Modules being compared to each other, it will become clear what the average/expected number of events captured should be. The event details of modules that deviate from this average can be analyzed using EventBrowser to see what is causing the deviation. For example, if one VA Module appears to be capturing many more events than the others, EventBrowser can be used to see if it is genuinely capturing more events or if it just registering a number of false positives.

5.2 Experimental results and VA performance report card

Using VAT, the following result has been obtained in a test that compared the in-house built video analytics solution (VAS) to the manual screening (MS) performed on a 24 hour video clip. The test involved counting the number of cars passing the region of interest from left to right and from right to left in one hour.

The time taken for MS to analyze the video was approximately 6 hours, whereas the time taken to do the same task with VAS was only 30 minutes; roughly 92% less than MS. Assuming that the accuracy of MS is 100%, VAS had an error of 6.67% where 7 events were missed out of a total of 105. Table 2 summarizes the results using the standardized report card.

For an organization that is considering whether to deploy a VA solution or not, VA performance evaluation results such as those that can now be obtained using VAT, an example of which is shown here, may become instrumental in making the decision.

Table 2: Counted Events (a) and Error of the Video Analytic Solution (b)

	VAS	MS	Difference	Percentage [1]
Total Events	98	105	7	6.67
Left	51	53	2	3.77
Right	47	52	5	9.62
Time (mins)	30	360	330	91.67

[1] Difference / MS

	Events	Misclassifications	Percentage
Cars	98	9	9.18 [2]
False Positives	7	-	6.60 [3]
Duplicates	1	-	0.94 [3]
Total	106	-	16.73

[2] Misclassifications / Cars, [3] Events / Total Events

6. CONCLUSION AND FUTURE WORK

This paper introduced the Video Analytics Platform and Testbed (VAP/VAT) that has been recently developed within the CBSA S&E directorate. VAP/VAT enables the development of in-house built VA codes as well as integration and testing of VA software available on the market with the existing video surveillance infrastructure.

VAP/VAT has been piloted in a number of indoor and outdoor environments and has shown that VA can be successfully deployed with existing equipment without any disruption to the normal workflow, which is very important for an operational agency. This opens a way to start developing task-specific VA solutions for each agency client and conducting thorough evaluations of the available solutions, thus approaching the second and third (final) phases in the VA deployment roadmap.

One of particular advances that has become possible with the development of VAP/VAT is the ability to embed Face Recognition software into existing Video Surveillance infrastructure and to conduct pilots related to the investigation of the feasibility of using Automated Face Recognition with video data, which has been one of the long-contemplated usages of facial biometrics in many agencies. This why for the next VT4NS workshop that will held in June 2010 [3] and where VAP/VAT will be presented to the GoC Community of Practice, the theme “Faces in Video” is chosen in addition to the main theme of the workshop “Deploying Video Analytics”.

ACKNOWLEDGEMENTS

The work of Bruce Rosenquist from the Electronics Section and the following students in implementing and testing VAP/VAT is gratefully acknowledged: Jean-Philippe Bergeron, Ehren Choy, Jeff Lalonde, Akarsh Malaya, Joshua Oh, Michael Smart, David Xu, Sabatini Bianchi and Timothy Wong.

REFERENCES

[1] Transport Canada CCTV Reference Manual for Security Application, December, 2007

[2] Second federal department’s workshop on Deploying Video Technologies for National Security (VT4NS’08). Theme: “Deploying Future-Proof Video Technology”. Organized by CBSA-S&E in coordination with DNI-VACE, Ottawa, 23 October, 2008 – www.videorecognition.com/vt4ns/08

[3] Third federal department’s workshop on Deploying Video Technologies for National Security (VT4NS’10). Themes: “Deploying Video Analytics” and “Faces in Video”. Organized by CBSA-S&E in coordination with DRDC-CSS, Ottawa, June 3, 2010.

[4] “Border Science: Going Beyond. 5-Year Vision/Strategy”. CBSA-S&E, May 2008

[5] D. Gorodnichy, “ACE Surveillance: The Next Generation Surveillance for Long-Term Monitoring and Activity Summarization”. First International Workshop on Video Processing for Security (VP4S-06), June 7-9, Quebec City, Canada. NRC 48493, 2006

[6] Dmitry O. Gorodnichy, Mohammad A. Ali, Elan Dubrofsky, Kris Woodbeck. “Zoom on the evidence with ACE Surveillance”. International Workshop on Video Processing and Recognition (VideoRec’07). May 28-30, 2007. Montreal , QC , Canada . NRC 49349. -
[7] D.Gorodnichy, T.Mungham, "Automated video surveillance: challenges and solutions. ACE Surveillance (Annotated Critical Evidence) case study". NATO SET-125 Symposium "Sensor and Technology for Defence against Terrorism", Mainheim, April 2008.
[8] Dmitry Gorodnichy. "Intelligent Surveillance: examples, myths and lessons", VT4NS’08 workshop, 23 October 2008
[9] D. Gorodnichy, "Image-based Biometric Technologies and their evaluation", Council on Security & Technology meeting on Biometrics, The Conference Board of Canada, Ottawa, 29 January 2009.
[10] Dmitry O. Gorodnichy, "Evolution and evaluation of biometric systems", Special Track on Applied Computational Intelligence in Biometrics, IEEE Symposium on Computational Intelligence for Security and Defence Applications (CISDA), 2009
[11] http://ipvideomarket.info/companies
[12] CBSA Video Surveillance GCPEDIA page : http://www.gcpedia.gc.ca/wiki/CBSA_Video_Surveillance. Last updated: 22 June 2009