WHITE PAPER – Interaction of AD and Rules

Background

The majority of today's software that enable VCA on a large number of cameras for both real time event detection and post-event forensic search, use a "rule-based" approach where the user needs to specify exactly what he is looking for (i.e. alert when a person loiters in an area for more than 3 minutes). The method first analyzes the video to extract metadata from it which describes the location, type, colour, size and other attributes of all objects in the scene. Based on the rules defined by the user, the software then analyzes the metadata and generates appropriate alerts when it detects a rule violation.

Current challenges and problems

The challenge faced today is that for most cameras, the users do not know in advance what may be interesting to them or which rules to set to detect a violation. The result is that most cameras that are deployed are constantly generating video which is being recorded, archived and then discarded without anybody deriving any valuable information from them. This problem is compounded by the increasing number of cameras being deployed for surveillance in city centers and public places where thousands of cameras are recording video continuously 24x7.

In video surveillance, the dominating applications for VCA are the detection of specific events such as perimeter intrusion, traffic violations or unusual events such as fighting, loitering etc. VCA has also gained further popularity with the surge in Big Data trends.

Current commercial video surveillance tools come in the form of rules such as virtual trip wires or objects entering/leaving regions. Such rules are however limited in the detection of events in which the rules are built for. For example, it is virtually impossible to construct rules for detecting fighting accurately due to its innumerable variations.

Machine vision is sensitive to variations imperceptible to humans. Environmental factors such as illumination, the position of the sun, shadows, rain and reflection are challenges in practice limiting the range of operating conditions for vision based systems. As such, most video surveillance with some form of intelligence is mostly limited to indoor environments where such factors can be more easily controlled.

Dealing with external environmental conditions requires the human operator to come up with creatively placing masks, applying multiple trip wires or specifying more focused regions of interest. Doing so is not only time consuming, it does not necessarily solve the problems associated with these outdoor conditions. Rule-based systems are also prone to an unrealistically high rate of false alarms or miss-detections outdoors. Such intractable situations are relatively often encountered in reality which often contributes to the declining credibility and reliability of rule-based VCA systems.

There has been research into the area of introducing learning based methods so that these rules need not be pre-defined. The system is expected to learn the environment itself and from there identify abnormal events that can potentially present a threat to security or safety.

However, both the above methods (i.e. learning-based and rule-based) have their shortcomings and the purpose of this project is to propose a novel method that will not only address these problems but attempt to solve them. We first list the problems and challenges faced by these two methods:

Rule-based method:

1. High False Alarm Rates (FAR) in challenging environments. – Most VCA methods make use of some form of background masking to reduce false alarms. Constantly changing or moving backgrounds generate false alarms that are impossible to eliminate using these conventional methods.
2. Difficulty in calibration/configuration - For example, a vehicle entering an unauthorized region or illegal parking can be detected by manually marking specific image regions of the video frame and looking for disturbances in pixel values belonging to only those pre-specified regions. This approach requires enumerating all possible events of interest, modeling and fine-tuning them to every installation. The scale of difficulty increases when the scene of interest differs due to large number of video feeds from cameras in a city-wide surveillance system.
3. Limitations in event detection – In many real-life scenarios, we may not actually know when an event might pose a threat to security. We may know what is "normal" but we may not have an expectation beforehand what is "abnormal". For example, an "abnormal" vehicle behaviour, such as unexpected stops, deviation from standard routes, speeding, traffic violations etc may indicate threats and dangers related to terrorist plots, hijacking, drunk driving, bank heists etc. Because of the sheer number of abnormal events that could happen, it is not possible to define rules in order to effectively detect them.

Learning Based Method:

1. Long learning period – In order for a learning based system to "learn" and differentiate between normal and abnormal events, it needs to be "trained" by feeding the system with video data over a period of time. To eliminate infrequent normal events that would otherwise have been detected as abnormal, the system would have to learn for days or weeks before becoming operational. Learning based systems learn from data. However, they often ignore the fact that a human operator wishes to customize the system results for operational reasons. There should be a way for an operator to input human prior knowledge on top of data collection.
2. Inefficient for detecting simple events – In general, a learning based system may be able to identify abnormal events better than a rule-based system but it becomes very inefficient in detecting simple but known events like a person crossing a virtual line. In order to detect this, it needs to be trained on hours of videos showing the usual behaviour of people not crossing the virtual line before it can understand that crossing the virtual line is an abnormal behaviour. The difficulty in training the system often leads to very high false alarm rates and low positive detection rates. In contrast, a rule-based system can easily detect this kind of event without any training.

Integrating Rules and Anomaly Detection (AD)

We introduce our solution which is known as ARVAS. ARVAS is made up of the fusion of rule based system and AD. Doing so provides mutual benefits for both AD and rule based systems:

1. AD receives information on rules and integrates the rules. The rules can be seen as the representation of a priori expert human knowledge that can be injected into the AD database. The incorporation of expert human knowledge and machine learning will potentially improve the accuracy of the AD. Additionally, the injection of rules representing normal motion allows easy customization and overcome data collection deficiencies / difficulties encountered.
2. Rules can also consider information from the AD to decide whether to trigger an alarm. In certain cases, the consideration of AD in rules alarm can lead to a reduction of false alarms. This is particularly useful in situations where there is some sort of frequently occurring phenomenon that poses problems for rules, or bad camera position view for example.

Practical benefits to integration:

ARVAS is an integrated system which will use a combined AI and rule-based approach to eliminate the above shortcomings to achieve:

1. Reduced FAR in certain outdoor situations difficult for rule based approaches – Constantly changing and moving backgrounds can be better accounted for and filtered out by AD while positive detection of the desired event is performed by rules.
2. Enhanced detection capabilities – An integrated system can now detect a wide range of events (abnormal and normal) with greater accuracy and not just a few known events pre-defined by a rule-based system.
3. Human knowledge and customization - ARVAS allows the incorporation of human/expert prior knowledge and similarly the customization of what is considered normal. This is of importance in practice as data collected might not be representative and the context might dictate some motion patterns as uninteresting. A key advantage of ARVAS is in the balance between two aspects; the ease in specifying human intuition/knowledge on the obvious parts, while leaving the complexity of the scene to be automatically discovered using the AI.

How Integration Can Help

Consider the case of using only rules in surveillance application. In this case, events that occur in the scene are either detected or not detected by a rule based system.

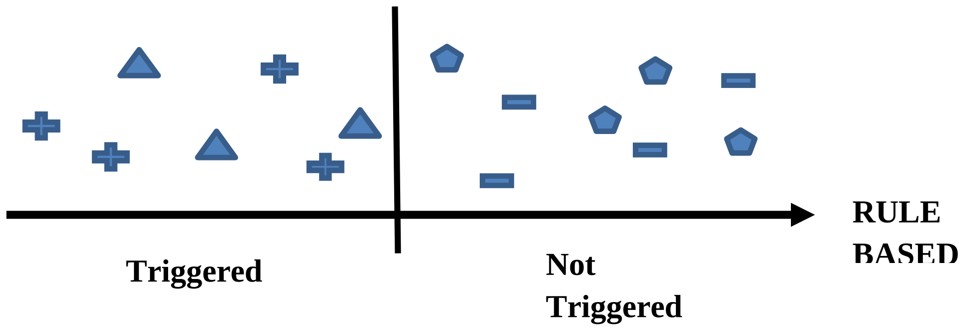

Figure 1. Representation of true positives (+ symbol), true negatives (- symbol), other "ambiguous cases" (triangle/pentagon symbols)

Integrating AD information into the picture is akin to adding another dimension that will under ideal circumstances allow a better separation of data such that events can be better differentiated (See Figure 2):

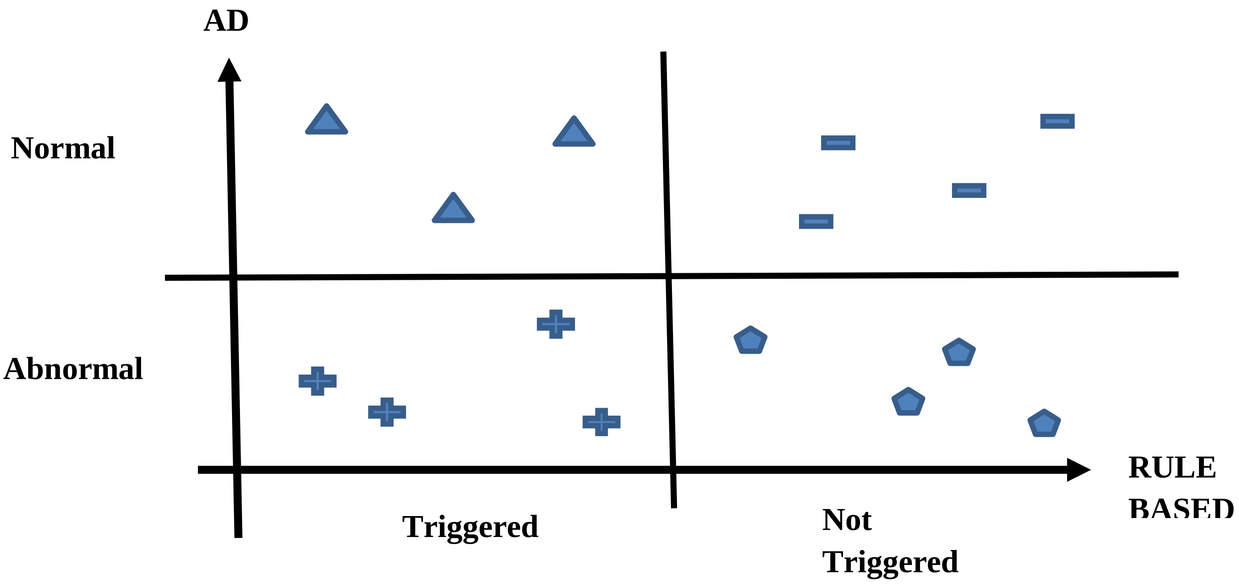

Figure 2: Integrating AD information adds an additional dimension and allows a better separation of event types. The graph illustrates the ideal case where true positives and negatives are well separated.

Taking outdoor environments as an example. Environmental influences like shadows or reflection from rain or sunshine for example tend to be systematic for most cases. These influences might be unnecessarily triggering an alarm detection intrusion for example rather frequently. These types of false alarms can be potentially filtered off by taking into consideration AD as it will be able to recognize these systematic environmental influences as normal after a period of time or after training.

An intrusion can be more reliably detected by the rule in this case by considering intrusion rule triggered when an abnormal event is happening. At the other end, we can be more reliably sure that there are no intrusions when it is mostly normal and that no intrusion rule is triggered. The other cases can be considered either as nuisance cases, or cases of lower priority to be separately handled, depending on the user alarm workflow.

About Vi Dimensions

Vi Dimensions was founded in 2015 with the simple idea that video analytics can be done in a much better and efficient way with the ultimate goal to revolutionize safe city surveillance harnessing thousands of cameras.

The company uses its patented algorithms and proprietary unsupervised Machine Learning techniques to derive meaningful information and actionable insights from live streaming video data. This translates to immediate value to the customer not only in terms of security and surveillance but also improves the organisation’s safety, operational and maintenance aspects.

Our advanced and innovative system analyses vast amounts of real-time streaming (or archived) data autonomously for abnormal behavior and events. It does not require human intervention to automatically discover dominant motion patterns which means that unlike conventional systems, it does not require a human to specify rules necessary for detection.

Contact

Vi Dimensions Pte Ltd
16 Ayer Rajah Crescent
Tempco Technominium
#05-04
Singapore 139965
Tel: +65 65702331
www.vidimensions.com
enquiries@vidimensions.com