Image and Video Analytics Technology

Image and video analytics have been around in their present form since at least the mid-1980s. The field is now quite broad, very advanced and definitely viable commercially, with numerous business and government applications currently deployed worldwide.

A typical current application of video and image analytics is intrusion detection. Large facilities, such as airports, cover such wide terrain that it's not feasible for a person, or even many people, to monitor the entire perimeter, but a camera, together with video analytics can do a good job of alerting security personnel when a vehicle or person is exiting or entering an area at a time or place that could be suspicious.

The following list of current, commercially successful applications for video and image analytics tells the story of how widely dispersed these applications are today:

- Detect fraudulent signatures on checks, and account applications, as well as of physician signatures for prescriptions

- Analysis of medical x-rays and images from fluorescence and electron microscopy

- Recognition of people such as known or suspected terrorists, wanted criminals or authorized personnel

- Traffic management and smart roadway systems
• Recognition of vehicle license plates, in varying weather conditions and at night. Recognition of numbers on railroad cars & cargo containers.

• Visual inspection and quality control for fabrics on a manufacturing line, to detect dirt spots, weaving faults, missing, warped or torn fibers.

• Ultra high speed light flicker event detection optimized for spark-plug analysis

• Multiple military target acquisition, tracking and guidance

• Terrain/object recognition and surveillance

Image Analytics for Social Networking Sites

There are estimated to be over 200 social networking websites and online communities worldwide, not counting niche (vertical) social network sites. These sites are used by more than one hundred million people worldwide and are garnering more page views than even top web portals. Registered users of these sites can post their content in the form of blogs, pictures, slide shows and videos and in this sense they become, in effect, publishers of online content. As a natural consequence, administrators of these sites face the problem of users who sometimes attempt to upload inappropriate content onto the site. Currently, the primary solution for this is to hire or contract persons to review all uploaded images one-by-one prior to release. As the user base and usage grow, the costs associated with this approach grow as well in a linear way. Image analytics offer an attractive alternative to this approach.
One of the most compelling benefits of image analytics is speed. Although the time it takes to process images and video through a manual review may vary, it certainly is not instantaneous. If analytics were fully implemented, then a significant number of the approved images and videos uploaded by users could appear on the site as soon as they were sent by the end user, thus increasing the confidence of new users that the upload process had been successful, increasing satisfaction and engagement with the site and increasing share of user due to the instant positive reinforcement, right when the user is actually engaged with the website. This could translate into significant economic benefits and brand differentiation for early adopters of image analytics.

Beyond this, administrators will like the fact that analytics have the ability to be trained, which means that webmasters can constantly fine-tune the selection/exclusion criteria, based on specific, real-life examples or changing rules. The system will constantly learn from example.

**How It Works**

Image analytics looks for patterns that have been seen before. Initially, the approach will be to train the application to spot benign content, such as pictures of nature, pictures of cars, pictures of pets and group photos with a row of smiling faces. It is logical that a significant portion of all approved images fall into definable groups such as these. Such images are uploaded into the image analytics engine to train the application to recognize patterns in these categories – just as commercially available software recognizes signatures that, although not exactly the same, are within a tolerance to one of the three signatures on file, for example.
This approach is focused on the green channel – benign images. This is low-hanging fruit, both from the standpoint of programming and from a cost savings perspective, since quick operational efficiencies can already be gained in this way.

The next stage in the process will focus on training the application to recognize objectionable content. In this case, anything that has a statistically relevant chance of being objectionable will be sent to the red channel for review by a team of analysts. It should be noted that human review is expected to continue even after full implementation of an image analytics solution. The difference is that humans will be asked to look at close calls and hard cases, not at thousands of pictures of smiling, happy faces.

**Project Work Plan**

**Phase One**

- Load the database with 1,000 – 10,000 images, including a mix of approved and rejected images. Close call images, identified as such, would be very useful as well.

- This phase concludes with a proof-of-concept, approved scope of work, proposal and implementation timeline.

**Phase Two**

- Parallel processing via the application and by agents. Track Type I and II error by the application and/or the
agents. Use images that generated classification error in order to further train the application.

- Initiate weekly calibration meetings. (See below.)

- Begin to feed end-user metadata into the mix of predictive variables. Fine-tune the application, using these secondary parameters as well.

- This phase concludes with commencement of release of the lowest-risk images directly from the application into the green channel without agent review.

Phase Three

- Continued refinement of image processing through training of the application using both inappropriate and close call content. Continued refinement of the decision rules through calibration.

- Initiate assessment of video, using the previously-developed image bank. Obtain examples of inappropriate videos that were rejected based on website guidelines as well as close calls. Analyze the sound tracks for objectionable material and train the application to recognize inappropriate audio content.

- Parallel processing of video content via the application and by agents. Track Type I and II error by application and/or agent. Use video content that generates classification error to further train the application.

- Add video content to calibration calls.
This phase concludes with attainment of full-automation of green channel. Human review, if desired, could be continued as a quality assessment function, targeted at the “close call” approved content.

**Phase Four**

- Release of video into green channel without agent review.
- Calibration and quality assessment continue
- Continuous improvement continues for both image processing and video processing.

**Calibration**

The process of calibration between a client and their outsource provider is very familiar to any firm that manages quality assessments in the outsourcing arena. For quality assessment, the function of calibration is to ensure that the outsourced quality analysts and supervisors are scoring the quality assessment form in the same way as the client would.

For this project, the purpose of calibration would be slightly different, but similar. The client will have established guidelines about inappropriate content. In the review process – whether manual or automated – those decision rules need to be interpreted in regard to specific user uploads, including close calls, i.e., uploads that push the limits of the site’s standards or that raise new questions that had not previously been considered. This is like the difference between the wording of a piece of legislation, and the case
law that takes shape around that legislation, as judges and courts try to interpret and apply the intent of the law in specific situations. There might also be cases where the client may have an evolving point of view about how to interpret the policy, and the calibration discipline will help manage and document all of that in an organized way.

Building up a rich, validated, logical and well-documented context for making consistent decisions about enforcement of the client policy is the goal of calibration and therefore we envision that calibration would be an ongoing part of the dialog between the service provider and Polaris Vanguard even after the initial implementation is complete.

**Economic Value**

There are several sources of economic value that can be created by the application of image analytics to content review of uploads to social networking websites and online communities. The most obvious source is the ability to automate the review of uploaded content. In addition to creating human resource economies, this can result in better utilization of existing investments in call center and BPO service centers that may be able to scale to handle review of images during non-peak hours, but that would have to invest in new facilities to perform those services if 100% of the work had to be performed by agents. Thus the ability to automate or semi-automate the assessment of image uploads could make the difference in whether sunk costs on existing facilities and IT infrastructure can be leveraged closer to their full capacity. Finally, as noted before, the near instant approvals associated with analytics would likely improve the user experience, engagement with the site and adoption, helping to differentiate the community from competitors.
Polaris Vanguard recently modeled the likely economic impact of use of image analytics in place of 100% manual review and we reached the conclusion that a 12-18 month ROI would be quite feasible for a community that uses 50-60 seats. ROI would be even shorter for applications with more seats and/or a growing user base.

A more precise estimate of ROI would be possible after an initial study and proof of concept is completed, using actual client data. This will allow empirical validation of percentages of images for which review can be fully automated. This information can then be documented in a service-level agreement.

**Next Step – Proof of Concept**

In order to facilitate evaluation of image analytics for web content review, we recommend a proof of concept. To ensure a valid test result, we recommend that an FTP site be set up with between 2,500 and 10,000 images – the closer to 10,000 the better – sorted into approved and objectionable categories. To ensure randomness of the sample, these images should represent 100% of a stream of work flow for a specific time frame, rather than having been selected by any other method.

From this, we will create an image topology that describes the percentage of images that fall into various machine-trainable classes, such as portrait photos, etc. At the conclusion of this stage, we will present a mini-application that recognizes benign content. The POC process will also help to refine the estimated time to deliver a production version of the application.
About Polaris Vanguard

Polaris Vanguard, is dedicated to speech, text, image and video analytics. We specify and implement analytics projects. With a focus on call center and BPO companies, our scope of work includes responsibility that the total solution will be used effectively on an ongoing basis by highly-trained experts – yours or ours, you decide – to constantly generate economic value. PV Services get you started with your business analysis so you can focus on results, not details.