INO

CHALLENGING LIGHT
FOR OUR PARTNERS

WHITE PAPER

VIDEO ANALYTICS

October 2011

VIDEO ANALYTICS

Video offers great promise in terms of discriminating power that advanced tracking systems could exploit to reach a higher level of performance. Yet, performing real-time video object detection and tracking in the natural outdoor environment is a difficult task, placing it in the general category of multiple-target tracking problems. Multiple-target tracking systems are being developed to address a large number of applications, including air traffic control (ATC). Implementing a functional ground ATC system reporting good quality tracks from video requires overcoming difficulties related to dynamic background and lighting changes. These include the visual impact of trees swaying in the wind, as well as the changing position and intensity of cast shadows according to the position of the sun and clouds in the sky. Continuously tracking an outdoor scene involves adapting from daytime to nighttime conditions and reducing as much as possible the negative impact of fog, rain or snow on the visual quality of the images.

Hence, the greatest challenge in video multiple-target tracking in the outdoors is to process noisy raw data (aka images) to form consistent observations from frame to frame in real-time. The reliability of this initial stage of the tracker will strongly influence system performance figures such as detection and false alarm rates. For the sake of simplicity, we will limit the discussion to a single fixed camera providing images a constant frame rate. Having access to a full network of cameras provides the system with additional perspectives without affecting the generality of camera-based object detection.

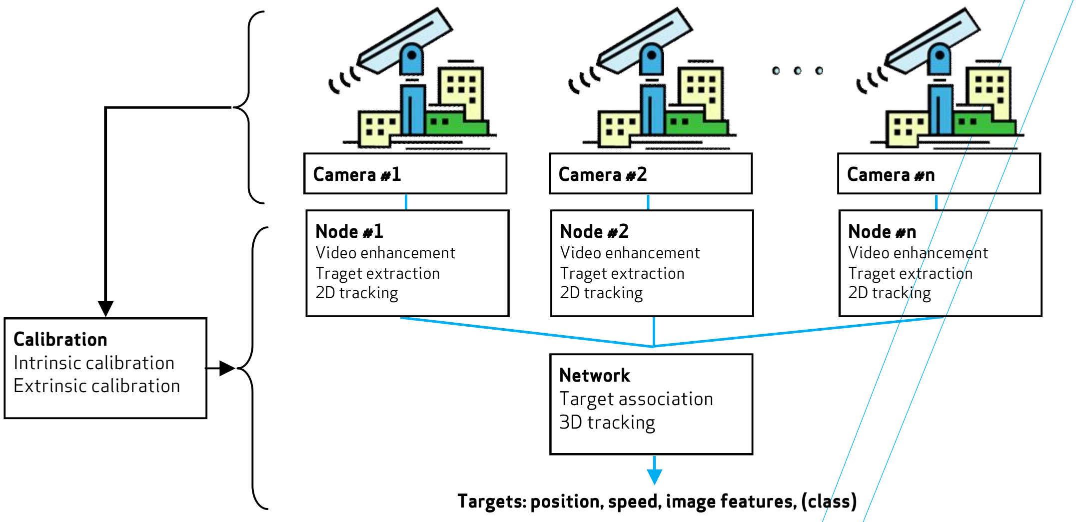

Figure 1: System architecture

Image processing starts with video enhancement. In general, this means that the image content will be analyzed and corrected for changes in illumination level, white balance or contrast enhancement to cope with lower visibility situations. One particular enhancement that allows for accurate tracking is image stabilization. Depending on their installation, outdoor cameras can be affected by strong winds and other vibrations. Image stabilization thus digitally remaps incoming images such that a series of stable landmarks in the field of view remain stationary in the enhanced video stream. In practice, this typically calls for different day and night processes, taking into account such considerations as cameras automatically switching to near-infrared night mode.

Once images are properly enhanced, the problem of object detection per se can be addressed. Several object detection techniques exist and designing better, more efficient object detection schemes is still a very active field of research. For fixed cameras, the most accurate and efficient object detection techniques are based on the background model subtraction scheme. Thus, a representation for the background has to be provided and dynamically updated. Comparing pixel -to-pixel the incoming image with the current state of the background model generates the foreground information. Objects are assumed to belong to the foreground information.

It is critical in real-time foreground-background segmentation to implement a very efficient model adapting to the changing environmental conditions at video rate. The implemented technique also has to support the presence of foreground objects in the scene during the background model initialization phase. The fastest—yet very accurate—techniques available fall into the category of non-parametric modeling. These implement a data-driven process to generate and maintain a pixel-wise, multi-modal representation of the background model. Practical design considerations also have to be taken into account to maximize performance and, in particular for color images, a proper choice of color transform yields a significant decrease in processing times.

Determination of the current foreground-background state of a pixel follows by estimating whether that pixel value belongs to any of the background modes or not. In cases where the pixel value represents an innovation with regards to the knowledge gathered in the background model, that pixel state is set to foreground. Next, a series of filtering steps have to be applied upon the foreground portions of the image in order to form groups of connecting pixels or blobs. The aim is also to eliminate foreground regions that resulted from spurious noise or could not be analyzed as meaningful objects in the application context, because certain constraints upon the attributes were not met, such as invalid location or size.

At that point, every blob now forms an observation. These are associated with a series of geometrical features such as location in the image; morphological attributes such as extent in pixels, mean intensity or color; and more generally, a series of statistical attributes or descriptors. Classification can also be performed at this point. The particular choice of observation attributes will have an important impact on overall system performance, and a compromise has to be made between the computational burden and the richness of the observation description. An important attribute of observations is real-world position. In the context of a relatively planar surveillance area, such as airport aprons and runways, a precise estimate of a ground object position can be obtained by computing the intersection of the projected camera's field of view with the plane corresponding to the ground altitude. This is called a monoscopic estimate of the position. Through this process, ground-touching observations are given an estimated world position right at the moment of detection.

However, image exploitation to render real-world position estimates to a high degree of precision has to be prepared beforehand through proper intrinsic and extrinsic calibrations of the camera. The object of intrinsic calibration is to obtain the camera model parameters. This is performed offline before the camera is installed. Once done, image processing algorithms can take into account real-world departures from the ideal camera such as misalignments of the optics with regards to the imaging electronics and optical aberrations causing image distortions. Finally, the ability to transform pixel coordinates into world position assumes the camera position and attitude with regards to the world reference frame has been properly estimated. This extrinsic calibration can be carried out once a series of world positioned reference landmarks appearing in the field of view of the camera have been identified.

From that point on, the tracker system is fed with good quality video observations and the subsequent tracking processes can take place. The remaining object tracking problem is the assignation of observations to tracks, and the description of a general practical multiple-target tracker system goes well beyond the scope of this discussion. Multiple-target tracking has fostered a huge amount of research and development activity, starting with radar tracking. Unlike radar tracking though, video tracking is richer, with a set of morphological features describing the extended objects being tracked.

Thus, video object tracking can be made more robust by exploiting the fusion of the geometrical and morphological features of the observations. This richer set of features provides the means to deal with partial or complete occlusion situations in the video stream where tracks meet one another, momentarily forming a "group." Upon later separation of the tracks, the morphological features are key to resolving ambiguities about the identity of the now separated tracks. These considerations support the claim that, upon successfully tackling the challenge of video tracking in the outdoors, one can unleash the inherent power of imaging to augment multiple-target tracking systems.

CONTACT INO

www.ino.ca / info@ino.ca

INO (HEADQUARTERS)

2740, Einstein Street
Quebec City, Quebec
G1P 4S4 Canada
418 657.7006 / 1 866 657.7406

ONTARIO

175 Longwood Rd. S
Suite 316 A
Hamilton, ON
L8P 0A1 Canada
905 529.7016

WESTERN CANADA

403 875.1521