Passive Vision: What Problems Become Easier Over Time?

Robert Pless, Washington University in St. Louis.

Many cameras do not move, and continually watch a specific scene -- an ATM, an airport security desk, a traffic intersection, a mountain or a beach for years at a time. Over the last 5 years my research has focused on how exploit the regularities and redundancies in video captured from a single viewpoint to improve many parts of video understanding. These regularities are captured in local statistics of the video (summarized through the distribution of the spatio-temporal image derivatives) and in global statistics such as the PCA decomposition of a large set of images. Together, these local and global statistics offer novel and robust algorithms for anomaly detection, scene annotation, camera geo-localization and geo-orientation. Because these tools tend to be very robust, they lend themselves to sensor network applications, and we have and continue to explore applications that are possible on the AMOS (Archive of Many Outdoor Scenes) database, which has been capturing images from more than 1000 webcams around the world every 30 minutes for the last 3 years.

Local spatio-temporal statistics offer tools to understand motion within a scene. Because of the relationships between spatio-temporal image derivatives and image motion, the statistics of image derivatives at each pixel characterizes the motions visible at that pixel --- allowing one to characterize moving background for anomaly detection [pless2003], and to annotate scene features such as roads that have consistent motions [pless2006]. Characterizing the motion patterns within a scene also gives very strong, scene specific priors to make tracking more efficient and robust [jacobs2008]. Because these algorithms are based entirely on simple measurements such as image derivatives, they are excellent candidates for surveillance networks.

Beyond charactering motion, the statistics of global image variations provide novel and robust algorithms for different classes of problems. We have shown that looking at the components of the PCA decomposition of images captured over time give automated tools to segment images by scene structure and surface normal [jacobs2007a], and looking at the patterns of the coefficients of the PCA decomposition, and how they vary as a function of the time of day, offers an approach to geo-localization of arbitrary camera feeds [jacobs2007b]. We have been exploring the use of these algorithms to support very large camera networks, through analysis of images from a world-wide camera network captured within the AMOS dataset [jacobs2007a]. We, and others, have been using this data set to study the scalability of visual algorithm in large scale visual sensor networks, especially addressing problems in lighting variation, camera calibration, and object recognition.

[pless2003] R. Pless, J. Larson, S. Siebers, and B. Westover. *Evaluation of local models of dynamic backgrounds*. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2003.

[pless2006] Robert Pless. *Detecting roads in stabilized video with the spatio-temporal structure tensor. Geoinformatica*, 10:1, 39-56, 2006

[jacobs2007a] Nathan Jacobs, Nathaniel Roman, and Robert Pless. *Consistent temporal variations in many outdoor scenes*. In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis,MN, June 2007.

[jacobs2007b] Nathan Jacobs, Scott Satkin, Nathaniel Roman, Richard Speyer, and Robert Pless. *Geolocating static cameras*. In Proc. International Conference on Computer Vision (ICCV), Rio De Janiero, Brazil, October 2007.

[jacobs2008] Nathan Jacobs, Michael Dixon, and Robert Pless. *Location-specific transition distributions for tracking*. In Proc. IEEE Workshop on Applications of Computer Vision, 2008.