Tracking People and Recognizing Content in Visual Data

Alexander C. Berg
Computer Science Department
Stony Brook University
http://acberg.com/

Some of the first pieces of information to extract from video of human activity are the location and trajectory of people. Even these simple observations can become difficult and nuanced when there are groups of people in the observed environment. Some of our work shows how to use cues about interaction between people moving together through an environment to make steps toward understanding group dynamics and to better track the motion. Approaching this as a recognition problem is one of the keys to an efficient solution. After position and trajectory, we can consider the motion of a person's limbs as they move on video. We show that it is possible to estimate action even for very low resolution video. The key to accomplishing this is a descriptor based on optical flow and effective analysis of spatial uncertainty. Surprisingly this works even for very low resolution video where individual limbs are difficult or impossible to discern. This simple technique is surprisingly powerful, allowing both analysis and synthesis of video showing human activity. Both the these techniques -- for tracking groups of people, and recognizing actions in low resolution video -- are in part based on previous work in still image recognition. Looking toward future work on understanding more general content in video we can take cues from the state of the art for large scale recognition in still images. To this end, we will present our recent work with ImageNet, the largest public, manually labeled collection of image data, and discuss how this may be relevant for future work on indexing visual content.

<u>Resources:</u>

"Who are you with and where are you going?"
Kota Yamaguchi, Alexander C. Berg, Tamara Berg, Luis Ortiz
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2011.

"Recognizing Action at a Distance"
Alexei A. Efros, Alexander C. Berg, Gregory P. Mori, Jitendra Malik
IEEE International Conference on Computer Vision (ICCV) 2003.

"Hedging Your Bets: Optimizing Accuracy-Specificity Trade-offs in Large Scale Visual Recognition"
Jia Deng, Jonathan Krause, Alexander C. Berg, Li Fei-Fei
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2012.

"Hierarchical Semantic Indexing for Large Scale Image Retrieval"
Jia Deng, Alexander C. Berg, Li Fei-Fei
IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2011.

"What does classifying more than 10,000 image categories tell us?"
Jia Deng, Alexander C. Berg, Kai Li, Li Fei-Fei
European Conference on Computer Vision (ECCV) 2010.