

# Tracking People and Recognizing Their Activities

Deva Ramanan  
EECS Division  
UC Berkeley  
USA

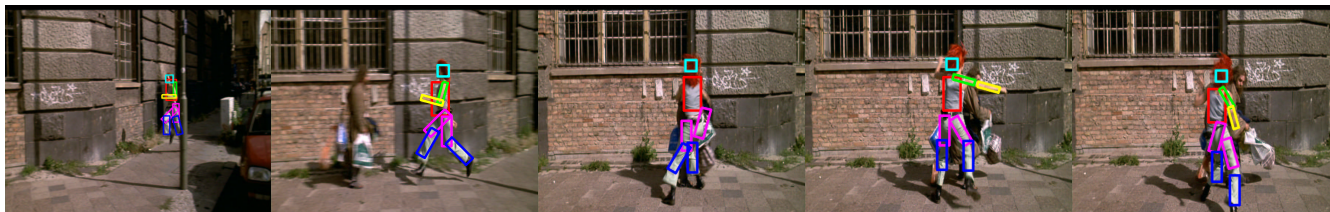
Email: ramanan@eecs.berkeley.edu

David Forsyth  
EECS Division  
UC Berkeley & UIUC  
USA

Email: daf@cs.berkeley.edu

Andrew Zisserman  
Engineering Science Dept  
Oxford University  
UK

Email: az@robots.ox.ac.uk



## DESCRIPTION

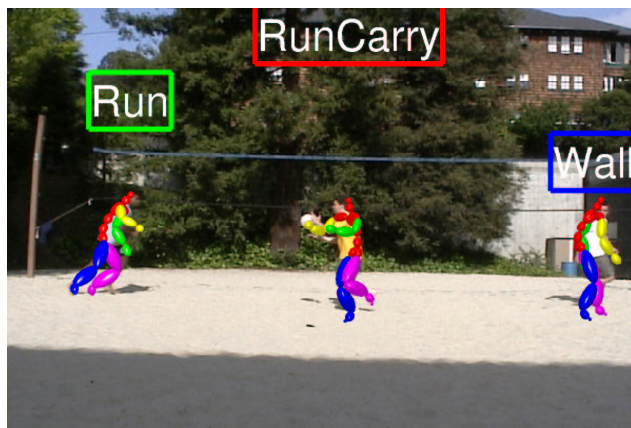
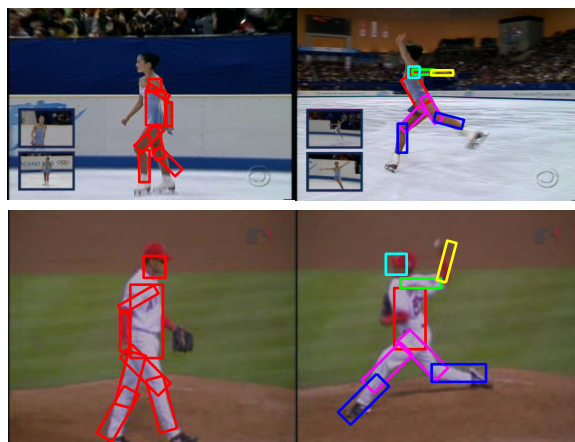
We present a system for automatic people tracking and activity recognition. This video includes the contribution of 3 papers over 2 years [1], [2], [3]. We show results on many sequences, including clips from a feature-length film and historical sports footage (all tracked completely automatically). We show results that utilize the tracker to obtain 3D reconstructions and activity descriptions (also done automatically).

Our basic approach to people-tracking is to build an appearance model for the person(s) in the video [1]. Since we do not know the appearance of a person *a priori*, we build the models on-the-fly. The video illustrates our method of using a stylized-pose detector (one that only finds people in lateral-walking configurations) [2]. Since people do not always walk laterally, we expect the detector to only fire in a few frames in a video sequence (shown in *red*). Our system builds a model of limb appearance from those sparse stylized detections (e.g., it might learn a baseball player has red arms and white legs). Our algorithm then re-processes the video, using the learned appearance models to find people in unrestricted configurations (shown in *color*).

We can use our tracker to recover 3D configurations and activity labels. We assume we have a motion capture library where the 3D poses have been labeled off-line with activity descriptions. We then synthesize 3D motions that look like our tracks by matching short clips of 3D motion to short clips of the 2D tracks [3]. The synthesized motions are constructed from the original pool of labeled motions and thus are labeled themselves.

## ACKNOWLEDGMENTS

The authors would like to thank the Computer Vision and Computer Graphics groups at UC Berkeley and the Visual Geometry Group at Oxford for many helpful discussions. This work was supported by NSF award no. 0098682 and by Office of Naval Research grant no. N00014-00-1-0890, as part of the MURI program. D. R. is supported by a NSF fellowship.



## REFERENCES

- [1] D. Ramanan and D.A. Forsyth. "Finding and Tracking People from the Bottom Up" *CVPR* June, 2003. Madison, WI.
- [2] D. Ramanan and D.A. Forsyth and A. Zisserman. "Strike a Pose: Tracking People by Finding Stylized Poses" *CVPR* June, 2005, San Diego, CA.
- [3] D. Ramanan and D.A. Forsyth. "Automatic Annotation of Everyday Movements" *NIPS* Oct., 2003, Vancouver, Canada.