

# Tracking and Understanding Human Motion

David Fleet  
<http://www.cs.toronto.edu/~fleet>

Department of Computer Science,  
University of Toronto

The estimation and interpretation of human motion from video are key enabling technologies for myriad applications. But despite more than a decade of focused research, the general problem in natural environments, from monocular observations, remains challenging. I will outline the nature of the problem, and describe some recent advances in modeling human motion, 3D pose estimation, the inference of human attributes, and on the inference of peoples' interactions with their environments, based on principles of kinematics, biomechanics and mechanics[1, 2, 3, 4, 5, 6, 7].

- [1] Marcus A. Brubaker, Leonid Sigal, and David J. Fleet. Video-based people tracking. In H. Nakashima, H. Aghajan, and J.C. Augusto, editors, *Handbook of Ambient Intelligence and Smart Environments*. Springer, 2009.
- [2] Marcus A. Brubaker, Leonid Sigal, and David J. Fleet. Estimating contact dynamics. In *Proc. IEEE International Conference on Computer Vision*, Kyoto, 2009.
- [3] Marcus A. Brubaker, David J. Fleet, and Aaron Hertzmann. Physics-based person tracking using the anthropomorphic walker. *International Journal of Computer Vision*, 87(1):140–155, 2010.
- [4] Martin de La Gorce, David J. Fleet, and Nikos Paragios. Model-based 3D hand pose estimation from monocular video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(9), 2011.
- [5] David J. Fleet. Motion models for people tracking. In T. Moeslund, A. Hilton, V. Krueger, and L. Sigal, editors, *Guide to Visual Analysis of Humans: Looking at People*. Springer, 2011.
- [6] Leonid Sigal, David J. Fleet, Niko Troje, and Micha Livne. Human attributes from 3D pose tracking. In *Proc. of European Conference on Computer Vision (ECCV)*, Heraklion, Greece, 2010.
- [7] Raquel Urtasun, David J. Fleet, and Pascal Fua. 3D people tracking with Gaussian process dynamical models. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume II, pages 238–245, New York, NY, 2006.

**David Fleet Biography:** David Fleet is Professor of Computer Science at the University of Toronto. He received the PhD in Computer Science in 1991, after which he was on faculty at Queen's University, and then with the Palo Alto Research Center (PARC). His main research interests and publications span the areas of computer vision, image processing, visual perception, and visual neuroscience.

In 1996 he was awarded an Alfred P. Sloan Research Fellowship for his research on biological vision. He has won paper awards at ICCV 1999, CVPR 2001, UIST 2003, and BMVC 2009. In 2010, he was awarded the Koenderink Prize for his work with Michael Black and Hedvig Sidenbladh on human pose tracking. He has served as Area Chair for numerous computer vision and machine learning conferences. He was Program Co-Chair for CVPR 2003, and will be Program Co-Chair for ECCV 2014. He has been Associate Editor (2000-2004), Associate Editor-In-Chief (2004-2008), and is currently on the Advisory Board for IEEE Transactions on PAMI. He is Fellow of the Canadian Institute for Advanced Research.