

Hand Tracking with Flocks of Features

Mathias Kölsch
Computer Science Department
Naval Postgraduate School
Monterey, CA 93943

Matthew Turk
Computer Science Department
University of California
Santa Barbara, CA 93106

Description

Tracking hands in live video is a challenging task: the hand appearance can change too rapidly for appearance-based trackers to work, and color-based trackers (that do not rely on geometry) have to make limiting assumptions about the background color.

This video shows the results of hand tracking with “Flocks of Features,” a tracking method that combines motion cues and a learned foreground color distribution to achieve fast and robust 2D tracking of highly articulated objects. Many independent image artifacts are tracked from one frame to the next, adhering only to local constraints. This concept is borrowed from nature since these tracks mimic the flight of flocking birds – exhibiting local individualism and variability while maintaining a clustered entirety.

The method’s benefits lie in its ability to track objects that undergo vast and rapid deformations, its ability to overcome failure modes from the motion cue as well as the color cue, its speed (5-15ms processing time per 640x480-sized frame), and its robustness against background noise. Tracker performance was demonstrated on hand tracking with a non-stationary camera in unconstrained indoor and outdoor environments. When compared to a CamShift tracker on the highly varied test data, Flocks of Features tracking yields a threefold improvement in terms of the number of frames of successful target tracking.

The Flocks of Features method differs from particle filters (condensation or particle swarm optimization) in three aspects. First, the features in a flock react to local observations only and do not have global knowledge as the samples or particles in Monte Carlo methods do. Second, the features’ domain is the two-dimensional image data (optical flow), not a higher-level model space. Third, they move in a deterministic way, rather than probabilistically sampling their state space.

Hand tracking has important applications for interaction with wearable computers, for intuitive manipulation of virtual objects, for detection of activity signatures, and much more. Tracking with Flocks

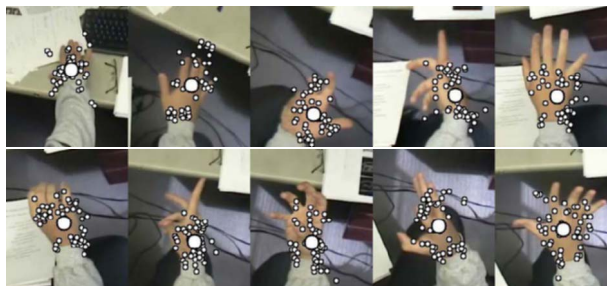


Figure 1: These images are taken from individual frames of a video sequence with highly articulated hand motions. For better visibility in print, areas with 200x230 pixel were cropped from the 720x480-sized frames. The cloud of little dots represents the flock of features, the big dot is their mean.

of Features is not limited to hands - any articulated or appearance-changing object can benefit from this multi-cue tracking method.

The Flocks of Features tracking method was first presented at the RTV4HCI workshop at CVPR 2004 [1], and it is described to more detail in [2]. We have integrated Flocks of Features into an open-source software library that implements a complete hand gesture recognition system and provides an of-the-shelf vision-based interface. For more information, please google *HandVu* or visit <http://www.cs.ucsb.edu/~matz/HGI/HandVu.html>.

References

- [1] M. Kölsch and M. Turk. Fast 2D Hand Tracking with Flocks of Features and Multi-Cue Integration. In *IEEE Workshop on Real-Time Vision for Human-Computer Interaction (at CVPR)*, 2004.
- [2] M. Kölsch and M. Turk. Flocks of Features for Tracking Articulated Objects. In B. Kisačanin, V. Pavlović, and T. Huang, editors, *Real-Time Vision for Human-Computer Interaction*, chapter 9. Springer-Verlag, 2005.