# Orientation histograms for hand gesture recognition

abstract:

We present a method to recognize hand gestures, based on a pattern recognition technique developed by McConnell \cite{McConnell86} employing histograms of local orientation. We use the orientation histogram as a feature vector for gesture classfication and interpolation. This method is simple and fast to compute, and offers some robustness to scene illumination changes. We have implemented a real-time version, which can distinguish a small vocabulary of about 10 different hand gestures. All the computation occurs on a workstation; special hardware is used only to digitize the image. A user can operate a computer graphic crane under hand gesture control, or play a game. We discuss limitations of this method. For moving or ``dynamic gestures'', the histogram of the spatio-temporal gradients of image intensity form the analogous feature vector and may be useful for dynamic gesture recognition.
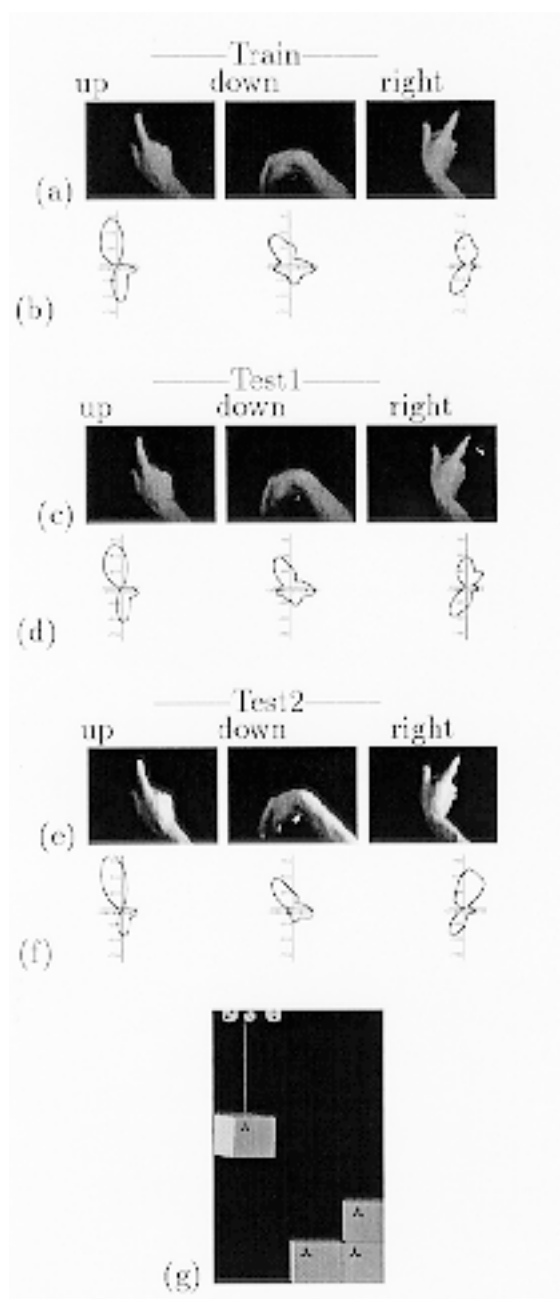
Subset of vocabulary of gestures used to control computer graphic crane. (a) shows the training set of gestures for the commands up, down and right. (c) shows a test set of the same gestures, under the same lighting conditions. (e) is a test set, made under different lighting conditions. (b), (d), and (f) are the corresponding orientation histograms. Note that the shapes look approximately the same as for the same hand

positions made under different lighting conditions, (b). An extension of this vocabulary of commands can control in real-time a computer graphic crane, (g).
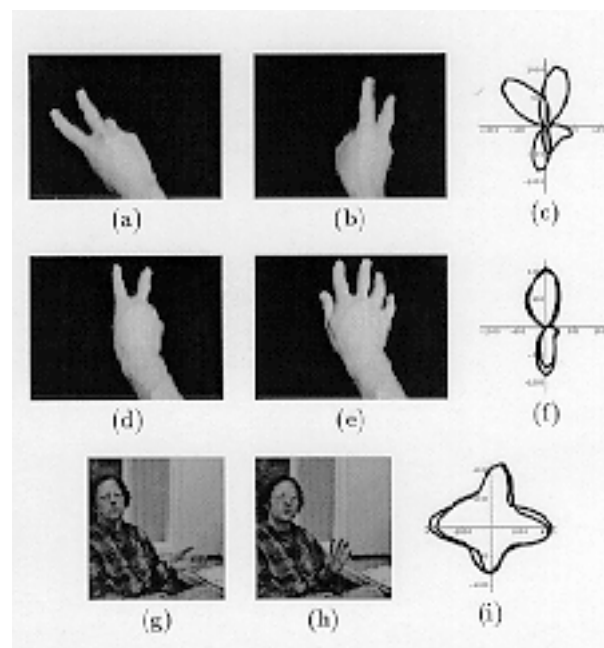
Problem images for the orientation histogram based gesture classifier. Users typically feel that (a) and (b) represent the same gesture, yet their orientation histograms are very different, shown overlaid in (c). A remedy for this problem is to provide training images of the gesture at various orientations. (Mathematical rotation of the feature vector is not sufficient; the corresponding orientation histograms are typically not simple rotations of each other.) Sometimes small changes in the image can cause large semantic differences, while changing the orientation histograms little. Users classify (d) and (e) as different gestures, yet their orientation histograms are nearly identical, (f). One has to construct a gesture vocabulary which avoids such gestures with similar orientation histograms. Finally, for this simple statistical technique to work, the hand must dominate the image. If it does not, then even large changes in the hand pose can cause negligible changes to the orientation histogram (g) -- (i).

References

Computer Vision for Interactive Computer Graphics, IEEE Computer Graphics and Applications, Vol. 18, No. 3, May-June 1998
W. T. Freeman, D. B. Anderson, P. A. Beardsley, C. N. Dodge, M. Roth, C. D. Weissman, W. S. Yerazunis, H. Kage, K. Kyuma, Y. Miyake, and K. Tanaka.
Also available as MERL-TR99-02.

W. T. Freeman and M. Roth, *Orientation histograms for hand gesture recognition*, Intl. Workshop on Automatic Face- and Gesture- Recognition, IEEE Computer Society, Zurich, Switzerland, June, 1995, pp. 296--301. MERL-TR94-03.