Classification of Co-articulated and Partly-occluded Hand Gestures

Atid Shamaie Alistair Sutherland
School of Computer Applications, Dublin City University, Dublin 9

A robust complete system for the recognition of hand gestures is the main goal of this research. One of the most difficult tasks for such a system is the recognition of co-articulated and partially occluded hand gestures. “Co-articulation” is the influence of one gesture on the next in a temporal sequence. Each gesture has a “canonical” form, which is the form of the gesture when it is made in isolation. In a sequence of gestures each gesture will deviate from the canonical form. It will be influenced by the preceding and following gestures in that the shape of the hand is changed gradually from one gesture to the next. This phenomenon is known as “co-articulation”.

In this research we extract different canonical gestures from a co-articulated sequence. First we make a model of each canonical gesture in the vocabulary of our system by grabbing image sequences of different variations of the gesture. We calculate the covariance matrix from the pixel values and form a subspace by using the first three eigenvectors. By projecting the sequence of images into the eigenspace, each image is mapped onto a point in the eigenspace. We form a three dimensional mesh from the set of points. The mesh represents all possible variations of the canonical gesture (Figure 1).

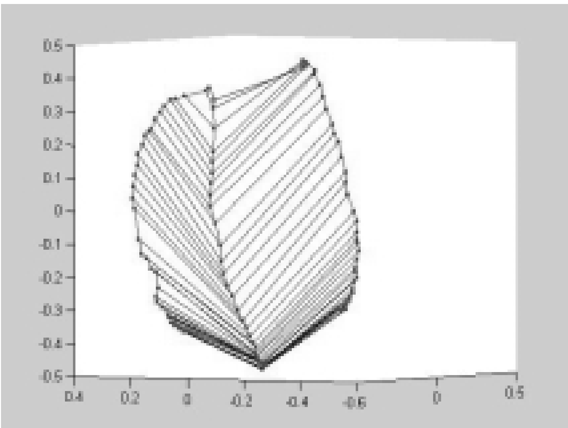

Figure 1. A 3D mesh representing all possible variations of a canonical gesture

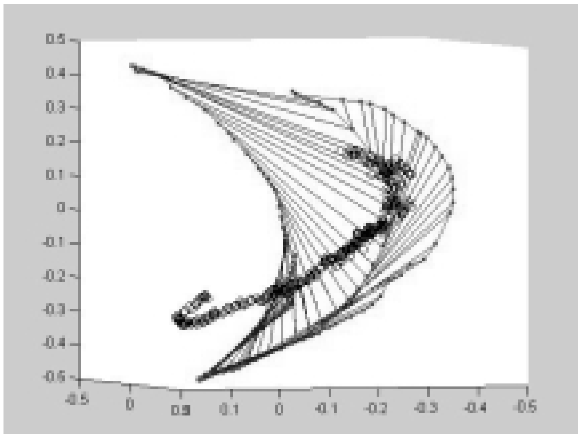

Figure2. Projection of a co-articulated gesture into a subspace

In Figure 2 we show the mesh of a canonical gesture and we show the trajectory of a co-articulated version of that gesture (shown by small circles). The trajectory starts some distance away from the mesh and moves towards it and eventually joins on. By considering a level of confidence, a three dimensional “envelope” is formed around every mesh. A well-chosen level of confidence enables us to find the part(s) of a trajectory, which is close to the mesh and falls inside the envelope. So the trajectory of a co-articulated sequence of, say, 2 canonical gestures is close to 2 different meshes in different parts.

We now consider only that section of the trajectory which falls within the envelope. We have developed a graph-matching algorithm (based on finding the shortest edges of a complete bipartite graph) which computes a distance between the trajectory and the mesh. The advantage of this algorithm is its ability to recognize partially occluded trajectories. Even if there is a gap in the new gesture sequence due to temporary occlusion it is still possible to compute a match. Gaps in the trajectory may also be caused by noise which can force certain points outside the envelope. Even in this case it is still possible to compute a match.

Using this distance measure we can find those canonical gestures which lie closest to the co-articulated sequence. Experimental results and measurements of accuracy based on a set of gestures will be presented.