Discovery of activities via statistical clustering of fixation patterns

Jeffrey B. Mulligan, NASA Ames Research Center

Jeffrey B. Mulligan, NASA Ames Research Center

Human behavior often consists of a series of distinct activities, each characterized by a unique signature of visual behavior. This is true even in a restricted domain, such as piloting an aircraft, where patterns of visual signatures might represent activities like communicating, navigating, and monitoring. We propose a novel analysis method for gaze-tracking data, to perform blind discovery of these activities based on their behavioral signatures. The method is in some respects similar to recurrence analysis, but here we compare not individual fixations, but groups of fixations aggregated over a fixed time interval. The duration of this interval is a parameter that we will refer to as τ. We assume that the environment has been divided into a set of N different areas-of-interest (AOIs). For a given interval of time of duration τ, we compute the proportion of time spent fixating each AOI, resulting in an N-dimensional vector. These proportions can be converted to counts by multiplying by τ divided by the average fixation duration (another parameter that we fix at 280 milliseconds). We compare different intervals by computing the chi-square statistic. The p-value associated with the statistic is the likelihood of observing the data under the hypothesis that the data in the two intervals were generated by a single process with a single set of probabilities governing the fixation of each AOI.

We have investigated the method using a set of 10 synthetic "activities," that sample 4 AOIs. Four of these activities visit 3 of the 4 AOIs, with equal probability; as there are four different ways to leave-one-out, there are four such activities. Similarly, there are six different activities that leave-two-out.

Sequences of simulated behavior were generated by running each activity for 40 seconds, in sequence, for a total of 6.7 minutes. The figure to the right shows the matrix of chi-square statistics, using a value of 2.8 seconds for τ, corresponding to 10 fixations. Low values (dark) indicate poor evidence for activity differences, while high values (bright) indicate strong evidence. The dark squares along the main diagonal each correspond to the forty second intervals in which the activity was held constant; the 4x4 block at the lower left corresponds to the four leave-one-out activities, while the 6x6 block in the upper right corresponds to the leave-two-out activities. (The anti-diagonal pattern of white squares indicates those activity pairs that share *no* AOIs.) The chi-square values can be binarized by choosing a particular significance level; we are interested in grouping bins that represent the same activity, effectively accepting the null hypothesis. Therefore, we may adopt a relatively lax criterion; for example, choosing a p-value of 0.2 means that two behaviors that have only a 1-in-5 chance of being produced by a single activity might nevertheless be clustered together.
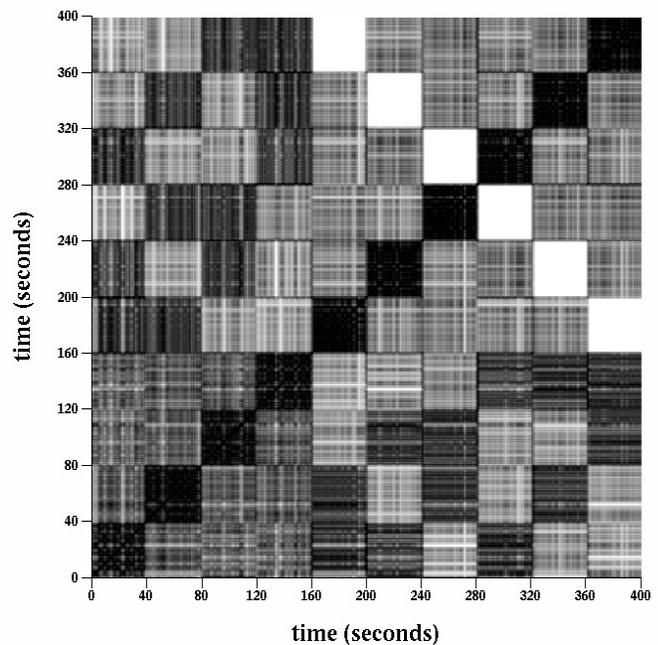
We have explored several methods to perform clustering on the data and solving for the activity probabilities. Greedy methods begin by selecting the time bin that is similar to the most (or least) other bins, and then forming a cluster from it and all other non-discriminable bins. These methods show mediocre performance, as they do not take into account temporal contiguity. Preliminary results indicate that methods that "grow" clusters in time from seed points perform better.