A real-time system for unobtrusive gaze tracking

Jeffrey B. Mulligan

An efficient method was demonstrated for estimating operator gaze (a product of eye and head position) in real time without constraining operators' head movements or requiring them to wear uncomfortable equipment. Gaze tracking is an important approach for unobtrusive performance monitoring in assessing the functional utility of complex display configurations, such as those used to display air traffic. Unfortunately, accurate gaze measurement usually requires either rigid positioning of the operator's head to allow a fixed camera to acquire a highly magnified image of the eye; or the use of a head-mounted camera system so that the eye image will not be lost when the operator makes small head movements. The latter approach is able to produce good data for someone seated naturally at a workstation but suffers from several drawbacks. In order to prevent slippage (and the consequent loss of calibration), the head mount must be mounted on the head so tightly that it is uncomfortable for extended use. There is always some obstruction of the field of view and it is not clear whether operators' behaviors will be completely natural when a bulky apparatus is attached to their head.

For these reasons, we have been developing a system that employs a steerable camera that can follow the movements of a user's head while maintaining a magnified image of one eye. To accomplish this, it was necessary to address the problem of keeping the eye's image centered within the image in real time to aid with steering. In the past, we had assumed that measurements of pupil position would be processed offline to determine gaze. The current system uses a pair of wide field cameras and an illumination method first developed at IBM Almaden Research Center to allow real-time localization of the pupils. The two cameras form a stereo pair that should allow recovery of the three-dimensional locations of the eyes, at least in principle. This allows, in turn, computation of the required steering control parameters.

It had been assumed that all of these calculations would require careful spatial calibration of the cameras. However, a camera-control method was developed that eliminates the need for precise calibration of the camera geometry. This was possible because it was unnecessary to recover the full three-dimensional position of the eye; it was sufficient to send only the pan and tilt parameters to the camera control unit. While the functional relationships among the measurements and required settings could be determined, given complete knowledge of the camera geometry, they would be sensitive to small errors in measurement of the camera positions and orientations. Thus, a method that did not require any knowledge of these parameters was required. A small laser pointer was mounted on the body of the steerable camera. It was aimed along the camera's line of sight and shone upon a screen placed at various positions in the volume of space within which the user's head would be tracked. The wide-field cameras image the entire space and can easily detect the small spot made by the laser beam. The pan and tilt positions corresponding to many positions of the spot within each of the wide-field images were obtained to generate a table of values relating pan and tilt settings to the corresponding measurements for the wide-field cameras. A variety of methods were used to generate pan and tilt settings for arbitrary measurement values based on this training set and it was found that a simple quadratic function was sufficient to produce pan and tilt values accurate to within one or two pixels.

This preliminary work employed a commercial pan-tilt camera designed for teleconferencing applications. Since the mechanism was too slow to follow the fastest natural head movements, work is underway to replace this component with a miniature camera whose line of sight is steered by a pair of galvanometer mirrors capable of much higher speeds.