

Chester F. Carlson Center for Imaging Science
Ph.D. Dissertation Defense

Susan M. Munn

**3D Head Motion, Point-of-Regard and Encoded Gaze Fixations in Real Scenes:
Next Generation Portable Video-Based Monocular Eye Tracking**

Advisor: Dr. Jeff B. Pelz

**Monday, 15 June 2009, 10:30 AM
Center for Imaging Science (Building 76), Room 1275**

Abstract

Portable eye trackers allow us to see where a subject is looking when performing a natural task with free head and body movements. These eye trackers include headgear containing a camera directed at one of the subject's eyes (the *eye camera*) and another camera (the *scene camera*) positioned above the same eye directed along the subject's line-of-sight. The output video includes the scene video with a crosshair depicting where the subject is looking – the *point-of-regard (POR)* – that is updated for each frame. This video may be the desired final result or it may be further analyzed to obtain more specific information about the subject's visual strategies. A list of the calculated POR positions in the scene video can also be analyzed. The goals of this project are to expand the information that we can obtain from a portable video-based monocular eye tracker and to minimize the amount of user interaction required to obtain and analyze this information. This work includes processing both the eye and scene videos to obtain robust 2D PORs in scene video frames, identify gaze fixations from these PORs, obtain 3D head motion and ray trace fixations through *volumes-of-interest (VOIs)* to determine what is being fixated, when and where (3D POR).

To avoid the redundancy of ray tracing the 2D POR in every video frame and to group these POR data meaningfully, a fixation-identification algorithm is employed to simplify the long list of 2D POR data into gaze fixations. In order to ray trace these fixations, the 3D motion – position and orientation over time – of the scene camera is computed. This camera motion is determined via an iterative structure and motion recovery algorithm that requires a calibrated camera and knowledge of the 3D location of at least four points in the scene (that can be selected from premeasured VOI vertices). The subject's 3D head motion is obtained directly from this camera motion. For the final stage of the algorithm, the 3D locations and dimensions of VOIs in the scene are required. This VOI information in world coordinates is converted to camera coordinates for ray tracing. A representative 2D POR position for each fixation is converted from image coordinates to the same camera coordinate system. Then, a ray is traced from the camera center through this position to determine which (if any) VOI is being fixated and where it is being fixated – the 3D POR in the world. Results are presented for various scenes along with a novel method for 3D visualization of the subject's PORs and motion with respect to VOIs.