

Watching People Making Errors: Vision Architectures for Monitoring Task Performance

Marten den Uyl

VicarVision, Amsterdam, The Netherlands. denuy@smr.nl

Abstract

People at work sometimes make errors, which can be of grave importance in some circumstances –e.g. medical, traffic- and bring nuisance and economic costs in other areas, e.g. order picking for web shops. It would be of value then, to have systems that can observe task performance and signal when –possibly- errors are being made, and thus may help to save lives and money. The Patria2 project (EURegio 2010) develops a system for observing task performance in medical and other settings, using primarily camera systems, and evaluates to what extent such a system could help in reducing avoidable errors by signaling increased risk of error. This presentation focusses on the challenges and problems of achieving adequate monitoring of task performance through the integration of results, in real time, from a number of vision systems under development at VicarVision.

A human supervisor could see increased risk of error in the way a subject performs a task from observing two kinds of deviations from standard or low risk behavior: 1) deviation by subject from standard or prescribed task procedures; and 2) deviation from optimal affective and motivational work attitude, e.g. stress, frustration or boredom may have negative effects on task performance. To see whether the subject deviates from standard procedures may require detailed knowledge of these procedures for the task at hand, as well as the ability to perceive specific details of task performance, hand and arm movements, direction of gaze. To perceive mental states and emotional responses in a subject may require even more attention to details of facial actions and body movements. And while performing a task, a subject may move outside the limited scope of a single camera. Thus, multiple camera systems will be required just to register task performance in sufficient detail. And for task monitoring to be relevant in most tasks, it will be necessary that the monitoring system responds in real time to an increased risk of error observed.

The Patria2 system involves three vision subsystems. FaceReader, a system for analyzing emotional expressions on the face that is available as a commercial product for behavioral research since almost 5 years www.noldus.com/human-behavior-research/products/facereader. GazeTracker, for determining direction of gaze from standard camera images, e.g. a webcam, and BodyReader, for the analysis of body pose in motion from video, are both under development for some years, but have not yet achieved maturity as a product. A typical set up for the Patria2 system uses six cameras.

The presentation will give an overview of the vision systems comprising the Patria2 architecture, status, performance and challenges. A number of special issues will be discussed:

- Alternative architectures for processing and integrating six or more camera streams in real time;
- Use of stereo and mono camera configurations;
- Integration of standard infrared eye tracking systems;
- Use of active depth sensors (Kinect);
- Problems of measuring mental states from video;
- Problems in identifying the precise steps in task performance.