ERCIM "Alain Bensoussan" Fellowship Scientific Report

Fellow:	Peter Nillius
Visited Location :	FORTH, Heraklion, Greece
Duration of Visit:	21/11/05 - 20/08/06

I - Scientific activity

Research

The overall topic for my research has been human motion estimation. For this nine-month period I have focused on developing shape detectors based on shading. The idea is to detect the arms, legs of the human body as they are approximately cylindrical. The developed shape detectors potentially also have a much wider use as low-level feature detectors for computer vision problems such as object recognition and face detection/recognition.

I developed, during my PhD studies, a framework for constructing low-dimensional models shading. For any given shape, a model is computed from the description of the variations in lighting and the variations in the surface reflectance properties (BRDF). The model consists of the principal components of the implicit set of images defined by the variations in lighting and BRDF. These components can be derived analytically making it possible to compute them rapidly. Figure 1 shows an example model, computed for a (slightly titled) cylinder.

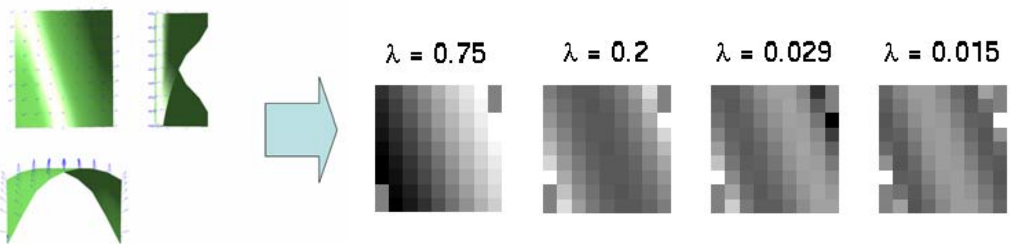

Figure 1. Shading model for a cylinder computed from a set of illumination maps and captured BRDFs. The eigenvalues indicate the captured variance of the basis function.

Such a model can be used to detect the shape simply by analyzing the residual variance after fitting the model to the image. This procedure detects the shape at a specific scale and pose. In this research project we were interested in recognizing shapes invariant of scale and pose. Therefore, the above mentioned framework had to be extended so that also shape variations could be encoded in the models.

To detect a shape at all scales and rotations the following is done. Large scale variations are handled by running the detector on a scale pyramid. Moreover, the models are trained to handle the scale variations interlinking the discrete scales in the pyramid. To handle pose variation a set of models is created to represent the shape in different poses, while each model is trained to handle the small pose changes in between models.

Detectors for spheres and cylinders were implemented. Applying the detectors results in a series of images showing the fit of the shape at different scales and rotations. Figure 2 summarizes these results for the cylinder detector, by showing the best rotation and scale at each position. Smooth cylinders are usually detected at the correct scale and direction, while textured cylinders are detected to a varying degree, depending on the amount of texture. In combination with edges these shading-based detectors could provide valuable cues in many vision tasks, e.g. object recognition and human motion estimation.

Figure 2. Results from the shading-based cylinder detector. Detected cylinders are colored according to their direction (left image) and size (right image).

Professional Activities
Apart from research I reviewed papers as a member of the program committee for *IEEE Computer Vision and Pattern Recognition (CVPR) 2006* and as a reviewer for the journal *IEEE Trans. on Pattern Recognition and Machine Intelligence (PAMI).*

Reading Group
During the fellowship I organized a reading group on Human Motion Estimation, in which members from the lab read and discussed papers on the topic.

<u>II- Publication(s) during your fellowship</u>

Multi-Target Tracking – Linking Identities using Bayesian Network Inference
Nillius, P., Sullivan, J. and Carlsson, S.
In *Proc. IEEE Computer Vision and Pattern Recognition* (CVPR06)
New York City, June 2006

Abstract
Multi-target tracking requires locating the targets and labeling their identities. The latter is a challenge when many targets, with indistinct appearances, frequently occlude one another, as in football and surveillance tracking. We present an approach to solving this labeling problem.

When isolated, a target can be tracked and its identity maintained. While, if targets interact this is not always the case. This paper assumes a track graph exists, denoting when targets are isolated and describing how they interact. Measures of similarity between isolated tracks are defined. The goal is to associate the identities of the isolated tracks, by exploiting the graph constraints and similarity measures.

We formulate this as a Bayesian network inference problem, allowing us to use standard message propagation to find the most probable set of paths in an efficient way. The high complexity inevitable in large problems is gracefully reduced by removing dependency links between tracks. We apply the method to a 10 min sequence of an international football game and compare results to ground truth.

III -Attended Seminars, Workshops, and Conferences
During the fellowship I attended the following events:

- IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 17-22, 2006, New York, USA
- Learning, Representation and Context for Human Sensing in Video, Workshop in connection with CVPR, June 22, 2006.
- The Onassis Foundation 2006 Lectures in Computer Science: Robots Intelligently Interaction with People, July 24-28, 2006, Heraklion, Greece