EE398B project proposal

3D face model extraction from a single image

James P. Mammen
(jmammen@stanford.edu)

Motivation

In specific applications like video-conferencing, news telecast, etc, where most of the image area is covered by a human face, it is possible to achieve extremely low bit-rate video transmission by using 3D head models. At present tracking algorithms such as [1], [2], [3] etc exist, which, given a 3D model of the head are able to track it in a video sequence. As yet, no fully automatic system has been reported which extracts a 3D head model from the video and uses it for tracking. If 3D head models can be extracted from the first frame (or a first few frames) in a video sequence then it will become possible to build extremely low bit rate video coding systems for communicating head and shoulder scenes. 3D head models can also be used for synthesizing novel views and facial expressions, animating virtual characters, enhancing lighting etc.

Project goals

In this project we intend to develop parts of a completely automatic system for 3D head model extraction. The most impressive research in this area is reported in [4] and [5]. We plan to implement and if possible enhance the scheme proposed in [4]. This will required the creation of a morphable 3D head model from a dataset of 3D scans of human faces. Once a morphable model is developed, the scheme for extracting the 3D model from a single image will be implemented.

Thus the goal of the project is to obtain a 3D face model from a single frontal face image in which the segmented face is available.

Work plan

The following work is expected to be carried out:

Stage 1: Using a dataset of 3D scans of human faces, a morphable 3D model will be built. Building the morphable model will involve obtaining 3D correspondences using optic flow and bootstrapping of the model as explaine in [6]. (Expected duration: 2 weeks)

Stage 2: In order to obtain a 3D model of a face from an image, the morphable model will have to be matched to the image. For this purpose, the parameters of the 3D model will be optimized along with a set of rendering parameters such that they produce an image as close as possible to the input image. (Expected duration: 2 weeks)

References

[1] P. Eisert, T. Wiegand and B. Girod, "Model-aided coding: A new approach to incorporate facial animation into motion-compensated video coding",*IEEE Trans. on Circuits and Systems for Video Tech.*, 10(3):344-358, April 2000.

[2] H. Li, P. Roivainen and R. Forchheimer, "3-D motion estimation in model-based facial image coding", *IEEE Trans. on PAMI.* 15(6):545-555, June 1993.

[3] M. Brand and R. Bhotika, "Flexible flow for 3D nonrigid tracking and shape recovery", *In Proc. of IEEE CVPR2001*, vol I:315-322, 2001.

[4] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces", *In Proc. of SIGGRAPH99*, 1999.

[5] M. Brand, "Morphable 3D models from video", *In Proc. of IEEE CVPR2001*, vol II:456-463, 2001.

[6] T. Vetter, M. Jones and T. Poggio, "A bootstrapping algorithm for learning linear models of object classes", *In Proc. of IEEE CVPR*, 40-46, 1997.