IRM PRESS
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.irm-press.com
ITB10092

Chapter IX

Automatic 3D Face Model Adaptation with Two Complexity Modes for Visual Communication*

Markus Kampmann
Ericsson Eurolab Deutschland GmbH, Germany

Liang Zhang
Communications Research Centre, Canada

Abstract

This chapter introduces a complete framework for automatic adaptation of a 3D face model to a human face for visual communication applications like video conferencing or video telephony. First, facial features in a facial image are estimated. Then, the 3D face model is adapted using the estimated facial features. This framework is scalable with respect to complexity. Two complexity modes, a low complexity and a high complexity mode, are introduced. For the low complexity mode, only eye and mouth features are estimated and the low complexity face model Candide *is adapted. For the*

high complexity mode, a more detailed face model is adapted, using eye and mouth features, eyebrow and nose features, and chin and cheek contours. Experimental results with natural videophone sequences show that with this framework automatic 3D face model adaptation with high accuracy is possible.

Introduction

In the last few years, virtual humans and especially animated virtual faces (also called talking heads) have achieved more and more attention and are used in various applications. In modern computer games, virtual humans act as football players or Kung Fu fighters. In movies, highly realistic animated virtual humans are replacing real actors (e.g., in the science fiction movie "Final Fantasy"). On the Internet, animated virtual faces are acting as news announcers or sales agents. In visual communication applications, like video telephony or video conferencing, the real faces of the participants are represented by virtual face clones of themselves. If we take a closer look at the technology behind these animated faces, the underlying shape of a virtual face is often built from a 3D wireframe consisting of vertices and triangles. This wireframe is textured using textures from a real person's facial image. Synthetic facial expressions are generated by animating the 3D wireframe. Usually, the face is animated by movement of the wireframe's vertices. In order to produce natural looking facial movements, an underlying animation structure (providing rules for animation) is needed, simulating the behavior of a real human face.

The creation of such an animated face requires generating a well-shaped and textured 3D wire-frame of a human face, as well as providing rules for animation of this specific 3D wireframe. There are different ways to create an animated face. One possibility is that an animated face is created manually by an experienced 3D modeler or animator. However, an automatic approach is less time consuming and is required for some applications. Dependent on the specific application and its requirements, different ways for the automatic creation of an animated face exist.

For 3D modeling of the shape of the head or face, i.e., for generation of the 3D wire-frame, techniques that are common for the 3D modeling of objects in general could be used. With a 3D scanner, a laser beam is sent out and reflected by the object's surface. Range data from the object can be obtained and used for 3D modeling. Other approaches use range data from multi-view images (Niem, 1994) obtained by multiple cameras for 3D modeling. All these techniques allow a very accurate 3D modeling of an object, i.e., a human head or face. However,