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
ITB10090

Chapter VII

Analysis and Synthesis of Facial Expressions

Peter Eisert
Fraunhofer Institute for Telecommunications, Germany

Abstract

In this chapter, the state-of-the-art in facial animation and expression analysis is reviewed and new techniques for the estimation of 3-D human motion, deformation, and facial expressions from monocular video sequences are presented. Since illumination has a considerable influence on the appearance of objects in a scene, methods for the derivation of photometric scene properties from images are also addressed. For a particular implementation, the potential of these analysis techniques is illustrated for applications like character animation and model-based video coding. Experiments have shown that the usage of 3-D computer models allows video transmissions at bit-rates of a few kbit/s, enabling a wide variety of new applications.

Introduction

Facial expression analysis and synthesis techniques have received increasing interest in recent years. Numerous new applications in the areas of low-bit-rate communication, user-friendly computer interfaces, the film industry, or medicine

are becoming more available with today's computers. In this chapter, the state-of-the-art in facial animation and analysis is reviewed and new techniques for the estimation of 3-D human motion, deformation, and facial expressions from monocular video sequences are presented. The chapter starts with an overview of existing methods for representing human heads and facial expressions three-dimensionally in a computer. Algorithms for the determination of facial expressions from images and image sequences are reviewed, focusing on feature-based and optical-flow based methods. For natural video capture conditions, scene lighting often varies over time. This illumination variability has a considerable influence not only on the visual appearance of the objects in the scene, but also on the performance of the estimation algorithms. Therefore, methods for determining lighting changes in the scene are discussed for the purpose of robust facial analysis under uncontrolled illumination settings. After this overview, an example of a hierarchical, gradient-based method for the robust estimation of MPEG-4 facial animation parameters is given, illustrating the potential of model-based coding. This method is able to simultaneously determine both global and local motion in the face in a linear, low-complexity framework. In order to improve the robustness against lighting changes in the scene, a new technique for the estimation of photometric properties based on *Eigen light maps* is added to the system. The performance of the presented methods is evaluated in some experiments given in the application section. First, the concept of model-based coding is described, where head-and-shoulder image sequences are represented by computer graphics models that are animated according to the facial motion and deformation extracted from real video sequences. Experiments validate that such sequences can be encoded at less than 1 kbit/s enabling a wide range of new applications. Given an object-based representation of the current scene, changes can easily be made by modifying the 3-D object models. In that context, we will show how facial expression analysis can be used to synthesize new video sequences of arbitrary people, who act exactly in the same way as the person in a reference sequence, which, e.g., enables applications in facial animation for film productions.

Review of Facial Analysis and Synthesis Techniques

Facial Animation

Modeling the human face is a challenging task because of its familiarity. Already early in life, we are confronted with faces and learn how to interpret them. We