Pattern Analysis for an Automatic and Low-Cost 3D Face Acquisition Technique

Karima Ouji[1], Mohsen Ardabilian[1], Liming Chen[1], and Faouzi Ghorbel[2]

[1] LIRIS, Lyon Research Center for Images and Intelligent Information Systems, Ecole Centrale de Lyon. 36, av. Guy de Collongue, 69134 Ecully, France
[2] GRIFT, Groupe de Recherche en Images et Formes de Tunisie, Ecole Nationale des Sciences de l'Informatique, Tunisie

Abstract. This paper proposes an automatic 3D face modeling and localizing technique, based on active stereovision. In the offline stage, the optical and geometrical parameters of the stereosensor are estimated. In the online acquisition stage, alternate complementary patterns are successively projected. The captured right and left images are separately analyzed in order to localize left and right primitives with sub-pixel precision. This analysis also provides us with an efficient segmentation of the informative facial region. Epipolar geometry transforms a stereo matching problem into a one-dimensional search problem. Indeed, we employ an adapted, optimized dynamic programming algorithm to pairs of primitives which are already located in each epiline. 3D geometry is retrieved by computing the intersection of optical rays coming from the pair of matched features. A pipeline of geometric modeling techniques is applied to densify the obtained 3D point cloud, and to mesh and texturize the 3D final face model. An appropriate evaluation strategy is proposed and experimental results are provided.

Keywords: Active stereovision, Amplitude analysis, Biometry, Plastic surgery, Sub-pixel sampling, Dynamic programming, Face segmentation.

1 Introduction

3D acquisition techniques have been applied in several fields such as biometry and facial animation or for aesthetic and plastic surgery, in surgical operation simulation, medical tracking, digital design of facial conformer and custom prosthesis, for example. Among 3D scanning techniques, two main groups are to be distinguished: passive methods and active ones. The techniques in which special lights are not used are called passive methods [1,2,3]. They use only information contained in the images of the scene. Otherwise, methods which use a controlled source of light such as laser or a coded light to recover the 3D information are called active methods. Lights intervene to extract dimensional information and to provide more precise reconstruction than passive methods. Active approaches can be 3D non-contact scanners like laser triangulation based ones. Here, laser rays coming out of a light source hit the object and are captured by a camera

J. Blanc-Talon et al. (Eds.): ACIVS 2009, LNCS 5807, pp. 300–308, 2009.

at different angles, using a rotating mirror. These devices capture highly accurate reconstructions. However, they are expensive and their output is usually noisy, requiring manual editing. In addition, they suffer from the surfaces reflection which generates small unreal variations in the surface of the scanned object [4,5].

The second solution is a structured light based approach in which special light is directed onto the object to be scanned, like the techniques presented in [7,8,9]. This process helps to solve the correspondence problem. When one pattern of light is projected onto the scene to be measured, depth information is extracted by analyzing the pattern deformations. If a set of light patterns is projected, extracting the codeword assigned to each pixel in the images allows the range image to be formed[4]. In this paper we propose a fully automatic and low cost 3D acquisition approach based on active stereovision. It uses a calibrated binocular stereo system assisted by a structured light source to recover 3D geometry. The proposed approach has the advantage of acquiring many images per second without any special synchronization between light source and cameras. It is especially appropriate for the acquisition of a 3D video sequence of non-textured surfaces like the human face or body.

The remainder of the paper is organized as follows: Section (2) describes our 3D reconstruction technique. In Section (3), the automatic face segmentation and primitives extraction process is detailed. In the section (4), we emphasize our evaluation strategy and first results of the developed technique and section (5) concludes the paper.

2 Overview of the Proposed Technique

Our 3D acquisition technique uses a calibrated binocular stereo system assisted by a structured light source. It is a low cost technique which retrieves depth information through a fully automatic procedure. It employs a particular pattern projected with a non invasive light source to deal with non textured surfaces like the human face or body. This provides us with left and right primitives for a given face with sub-pixel precision. The proposed stereo sensor is illustrated in figure 1.

We first calibrate the left and right cameras in order to extract their optical characteristics and geometrical parameters. Second, two complementary patterns with alternate black and white stripes are projected onto the face. The images are rectified through epipolar geometry transformation. Thus corresponding points necessarily have the same Y-coordinate, in rectified images.

Epipolar geometry transforms a stereo matching problem into a unidimensional search problem. Indeed, we employ an adapted, optimized dynamic programming algorithm to pairs of primitives which are already located [4,10]. Continuity, uniqueness and ordering constraints are taken into account within our algorithm. Also, the maximum disparity range is computed regarding the maximum and minimum of depth and geometry of our stereo system. This represents a disparity limit constraint which improves the stereo matching process.

In this paper, we propose to resolve the matching problem through an automatic process for each epiline separately. An amplitude analysis of gray level