3D Facial Surface Acquisition by Structured Light

Charles Beumier [1] , Marc Acheroy

Royal Military Academy, Signal & Image Centre (c/o ELEC)
Avenue de la Renaissance, 30 B1000 Brussels, Belgium
beumier@elec.rma.ac.be
Tel: +32 2 737 64 74, Fax: +32 2 737 64 72

Abstract

This paper presents a 3D acquisition system based on structured light and adapted to 3D facial capture. The design followed speed, low cost and resolution requirements. A single shot with correct posture is sufficient to capture nearly the whole facial surface in half a second. The system has been used to validate person verification from 3D facial analysis.

1 Introduction

Person verification from profile [2] has proven to be efficient in terms of speed, performance and simplicity. We attribute these qualities to the geometrical nature of the profile information, independent of illumination and involving rather rigid parts.

More geometrical clues are available from a full 3D facial representation [1]. This allows for translation, rotation and scale compensation of the point of view. Dealing with range information also simplifies face detection.

Among 3D acquisition solutions [3], a structured light system has been developed and optimized to benefit from a low cost and quick capture.

2 Motivations for structured light

Structured light acquisition systems use the projection of a known light pattern (in our case, parallel 'stripes') to recover 3D coordinates. Our choice has been motivated by the following considerations.

First, compared to a classical camera, the additional cost is limited to a projector and its slide. Secondly, a single image with the projected light pattern suffices to recover absolute 3D coordinates. This allows for sequence analysis or time integration. Thirdly, we are convinced that most of 3D facial information lie in the resolution offered by normal cameras and projectors. Fourthly, texture (without stripes) and 3-D data can be acquired in near registration by switching the projector on and off.

The main disadvantages of a structured light system with slide projection are the relative bulkiness of the camera/projector head and the limited depth of focus due to optical systems.

[1] This work has been partially supported by the European ACTS programme (AC102 "M2VTS")

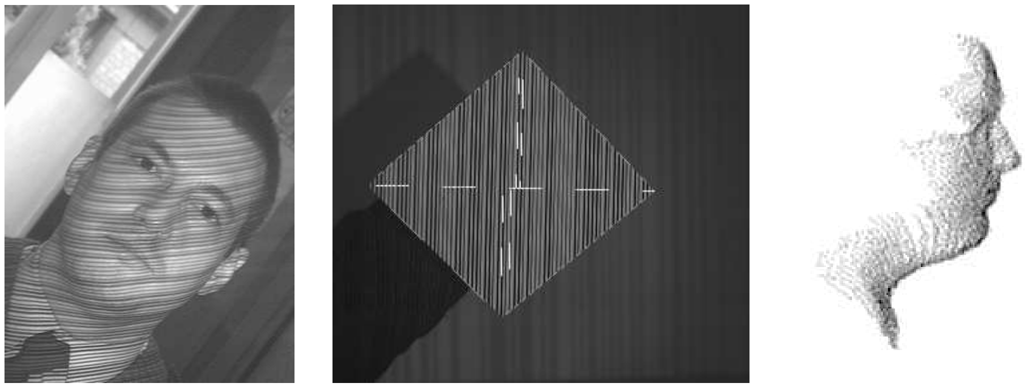

Figure 1: a) Striped image, b) Image for calibration, c) 3D from Fig.1a.

3 Hardware

A normal monochrome CCD camera and a 24x36 mm slide projector have been used. A 1-D stripe pattern consisting of parallel thick and thin stripes has been carefully designed. The index of each stripe is encoded in the binary thickness of its neighbours.

Optical lenses were chosen to achieve a sufficient field of view (40x30 cm) and depth of focus (30 cm) with limited projected light power, working at a distance of 150 cm. This is compatible with a cooperative scenario for a sitting attitude.

The camera was rotated (portrait) to benefit from the larger horizontal resolution. The 3D acquisition head was rotated of about 45^o (see Fig. 1.a) so that stripes are nor horizontal (avoiding stripe detection problems in eyebrows or mouth) neither vertical (avoiding a difference in stripe density between left and right).

4 Calibration

The camera and the projector have been fixed with their optical axis coplanar. In that situation and with the stripe pattern perpendicular to the optical axes plane, the number of parameters is reduced to 7. They concern the camera/projector distance, optical axis angles, pixel to angle and stripe to angle conversion factors.

The first calibration step consists in the collection of distances and optical values to derive rough estimates of the parameters. Then parameter refinement is carried out to bring the corners of a reference square object (see Fig. 1.b) at correct relative positions. For this, two criteria are used: the maximal planarity and the most compatible interdistances between corners.

5 3D extraction from striped image

The 3D coordinates of points along each stripe are derived from the position in the image and the index of the stripe. First, the stripes are detected from horizontal gradient and vertical continuity. Secondly, the stripe thickness (thick or thin) is estimated from the horizontal grey level profile. The thickness distribution of consecutive stripes allows for stripe labelling. Thirdly, individual labels are globally checked to solve local inconsistencies. Finally, median filtering smooths out the surface to clean local bumps incompatible with a 3D facial surface. The whole process, from image capture to coordinates delivery takes 0.5 second on a Pentium 200 MHz.

6 3D database and tests

Running the 3D reconstruction algorithm on 120 individuals (2 sessions of 3 shots each, with little orientation changes) made us confident in the overall quality of stripe following, labelling and background independence. However, it highlighted the problems encountered in bushy beards, glasses, nose and eyes, by order of importance. The quality of the 3D capture was later supported by recognition experiments.

7 Conclusions

The presented system for 3D acquisition by structured light has the advantages of delivering 3D coordinates quickly and automatically, without background influence. Tests with a large population confirmed those qualities as well as the limitations encountered for bushy beard and spectacles. The system proved adequate to deliver data valuable for 3D face comparison.

References

[1] C. Beumier, M.P. Acheroy, "Automatic Face Authentication from 3D Surface", In *British Machine Vision Conference BMVC 98*, University of Southampton, UK, 449-458, 14-17 Sep 1998.

[2] R. Chellappa, C.L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces", In *Proceedings of the IEEE*, The institute of electrical and electronics engineers, inc., vol. 83, No. 5, pp. 705-740, May 1995.

[3] R. Jarvis, "Range Sensing for Computer Vision", In *Three-Dimensional Object Recognition Systems*, Advances in Image Communication, Volume 1, A.K. Jain and P.J. Flynn (Editors), Elsevier Science Publisher 1993, pp 17-56.