DESIGN OF CODED STRUCTURED LIGHT PATTERN FOR 3D FACIAL SURFACE CAPTURE

Charles Beumier

Signal and Image Centre, Royal Military Academy
Avenue de la Renaissance, 30, B-1000 Brussels, Belgium (Europe)
phone: +32 2 7376474, fax: +32 2 7376472, email: beumier@elec.rma.ac.be
web: http://www.elte.rma.ac.be/~beumier/index.html

ABSTRACT

In the context of 3D face recognition, facial surfaces are advantageously captured by a structured light acquisition system, which is typically quick, low cost and uses off-the-shelve components. The light pattern projected, a key aspect of the structured light approach, makes the major difference between developed systems. In most of them, elements of the light pattern must be identified by a property such as element thickness or colour. We present in this paper the design of projected patterns that led to the realisation of three 3D acquisition prototypes.

1. INTRODUCTION

3D Face recognition is an effective approach to address the problems of point of view and illumination dependence affecting the classical 2D face recognition methods. In order to capture facial surfaces, the structured light technique appears appropriate since it offers a sufficient resolution while being quick, low cost and possibly built from off-the-shelve components. However, hardware design and software developments are needed to calibrate the system and analyse the acquired images to extract 3D coordinates.

A structured light system uses a camera and a projector with a specific pattern of light to localise 3D points in the scene. The aspect of calibration, a step necessary to deliver accurate 3D measurements, is not described here but in [1].

Many structured light systems have been developed [2]. They mainly differ in the projected pattern. Our own developments started in 1996 when only a few and expensive systems were available. We later refined the prototype to improve the quality of reconstructed surfaces.

After a presentation of the principle of structured light in section 2, the paper concentrates on the design of the light pattern (section 3). In our specific designs related to *coded* structured light, the projected elements must be identified. Several implementations have been considered (section 4), following an evolution leading to three prototypes of increasing quality.

2. CODED STRUCTURED LIGHT

2.1 Principle

The structured light principle relies on the use of a projector illuminating the scene of interest with a specific light pattern. A camera grabs images of the scene from a point of view that is different from the illumination direction. The difference in point of view allows for triangulation to localise 3D points in space.

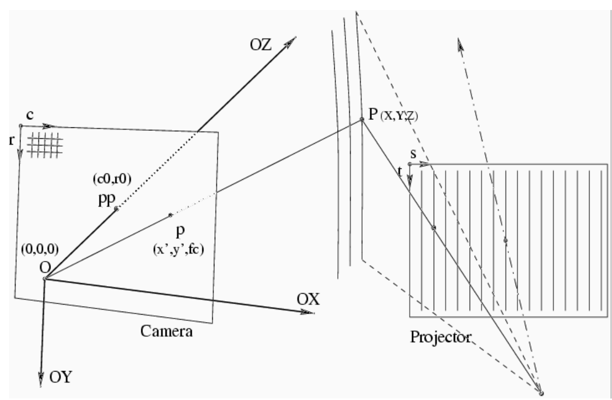

Figure 1: Structured light principle

More precisely, and referring to Figure 1, point P is localised at the intersection of two 3D lines, one emanating from the camera (Op) and the other from the projector. Both lines are estimated from the same captured image. The first line is obtained thanks to the detected image position (c,r) which is converted into position p (x',y') to take camera distortion into account:

$$x' = (c - c0) * (1 + k_r * d^2)$$
$$y' = (r - r0) * (1 + k_r * d^2)$$

where

(x',y') are the coordinates in the image plane
(c,r) are the pixel column and row,
(c0,r0) are the coordinates of the principal point,
$d^2 = ((c - c0)^2 + (r - r0)^2)$,
k_r is the first radial distortion coefficient.

Having the axes origin at the focal point of the camera, with OX parallel to pixel rows, OY parallel to pixel columns and OZ perpendicular to the image plane, point P (X,Y,Z) is on the 3D line passing by O (0,0,0) and p (x',y',f_c), with f_c the camera focal length. The second line, related to coordinates (s,t) of the projected pattern, can be obtained similarly, but in practice, it is preferable, as explained in section 2.2, to project lines (stripes, as in Figure 1), so that point P is at the intersection of line Op and a plane (neglecting projector distortion) emanating from the projector.

Mathematically, the projection planes correspond to

$$a_s X + b_s Y + c_s Z = d_s$$

with (a_s,b_s,c_s) the unit vector normal to the plane.
The intersection with line Op gives the equations:

$$\left.\begin{aligned} X &= d_s x' / (-a_s x' - b_s y' - c_s f_c) \\ Y &= d_s y' / (-a_s x' - b_s y' - c_s f_c) \\ Z &= d_s f_c / (-a_s x' - b_s y' - c_s f_c) \end{aligned}\right\} \quad \text{(Equations 1)}$$

2.2 Projected pattern

Structured light systems mainly differ in the type of patterns projected [2]. As suggested in [3], the systems can be characterised by the underlying assumptions made about the reflectance and the coherence in space and time of the scene. In [4], a reflectance assumption is made to benefit from colour projection to identify pattern elements. The spatial coherence is a necessary hypothesis if element identity is coded in the label of several neighbours [4]. The time coherence is necessary for systems projecting a sequence of patterns [3], which we found too restrictive regarding hardware needs and user cooperation.

Following the reflectance and spatial coherence assumptions, valid in most facial surface captures, the projected pattern has to serve several roles. First, the pattern elements must allow point localisation in the captured image to identify the 3D line Op (see Figure 1). Secondly, the pattern elements must have a label that allows their unique identification to determine the projected plane of light. This is not a strict rule since the plane ordering suffices in an orthographic projection and for a connected surface, as assumed in the work of Proesmans [5]. Finally, the projected pattern should not modify too much the surface texture appearance if the texture is to be acquired from the striped image.

As already mentioned, a common type of pattern element is the line or *stripe*. Stripes are precisely localised thanks to their continuity allowing for the average of localised points. Stripe identification is achieved by labelling the stripes with a property like thickness or colour. The continuity of the property along the stripe brings robustness in the labelling. A set of parallel stripes separated by white stripes gives a high density of points along the stripes while leaving achromatic illumination for texture measurements.

3. PATTERN DESIGN

3.1 Parallel stripes

The patterns that we have designed were all based on parallel stripes, localised either by their centre or by their edges.

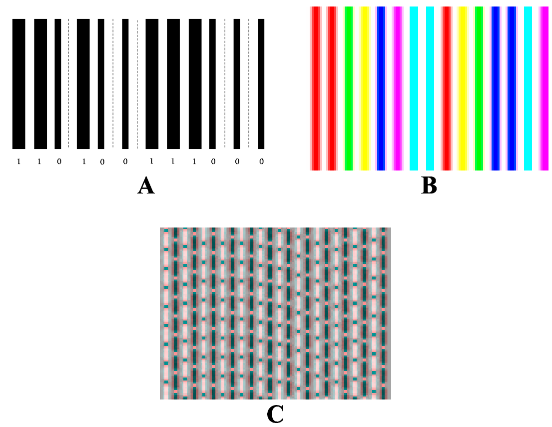

Figure 2: Partial designs for pattern A, B and C

Three different properties were investigated for stripe labelling: stripe thickness (A), stripe colour (B) and vertical position of dots (C) on the stripe. The unique identification of each stripe is achieved through the combination of labels of several stripe neighbours.

The intervals between the stripes are transparent or grey to measure the real colour texture of the surface.

3.2 Stripe localisation

For pattern A, localisation consists in detecting the left edges of the black stripes, which are regularly spaced in the horizontal direction. In pattern B, each colour stripe gives two colour stripe edges for localisation in at least one of the RGB fields. Pattern C provides good localisation along the black and the white stripe centres.

3.3 Stripe identification

The unique identification of each stripe is coded in the label of a few stripe neighbours.

In the case of pattern A, the labels are coded in the stripe thickness (thin or thick), and 7 stripe neighbours are used to get a unique identifier. This offers 128 possibilities, leading to a sequence of 128+7-1 = 134 stripes.

Pattern B has six possible colours (Red, Yellow, Green, Cyan, Blue, Magenta) that are well separated in the space of hues. An identification scheme based on two successive colour stripes has been created, leading to a sequence of 36 unique pairs of colours. This sequence has been repeated three times, with rotation of the Red, Green and Blue fields of the six colours. We thus dispose of a sequence of 108 colour stripes separated by transparent spaces, giving a total of

218 stripe edges with at least one transition in one of the RGB fields. Each pair of colours appears three times in the sequence, but sufficiently separated to avoid ambiguity for surfaces with limited depth discontinuities. The use of additional neighbours reduces possible ambiguities.

Pattern C uses six different vertical offsets between consecutive stripes. Three consecutive stripes bring two vertical offsets, leading to 36 possibilities. As the stripes are alternatively black and white, the repetition of the same sequence applied to stripes of the opposite colour allows a combined sequence of 71 possibilities with unique identification of any three consecutive stripes. This sequence has been repeated three times to reach 213 stripes. We argue here again that the smoothness of surfaces and the limited depth of field will prevent ambiguities between stripes separated by one third of the pattern width. Take note that the absolute vertical position of the dots can be used as label if we already dispose of a vertical reference. In that case, only two consecutive stripes suffice for identification.

3.4 Texture capture

Pattern A was not designed with texture recovery in mind. The presence of wide and dark stripes impairs the visibility of the surface colour.

Prototype B considered texture capture by direct measurements in the achromatic spaces and colour compensation in the colour stripes. Colours are compensated in each R,G,B field independently, based on the left and right achromatic spaces.

Prototype C compensates the R,G,B colour levels in the black and white stripes based on the left and right levels in the grey background. The compensation is based on the linear assumption of the levels in the horizontal direction across a stripe.

3.5 Binary sequence creation

For pattern A, a binary sequence has been created so that all the possible (128) 7-bit words appear once and only once.

Starting with any 7-bit word (e.g. 0000000), we enlarge the sequence by successively adding a ‘0’ or ‘1’ bit as long as the word consisting of the 7 trailing bits has not been used yet.

Once a trailing bit cannot be added, a second sequence is started with any unused 7-bit word and enlarged. After this second sequence has been completed, it can be shown that it is possible to insert it into the first sequence [1]. The insertion of other sequences into the first (growing) one finishes when all the words have been used.

We give below the obtained sequence. The solution is cyclic and not unique. Interestingly, the distribution of smaller words (1, 2, …, 6 bits) is rather balanced in the whole sequence. It can be used as pseudo random sequence.

Sequence for 7-bit words (134-bit long):
00000001111111011111001111000111 \
01001110000110111011011001101000 \
11000001011110101110010110001010 \
11010101001010000100110010010001000000.

3.6 6-uple sequence with unique pairs

The creation of n-tuple sequences follows the same principle but is more complex as the number of possibilities (n) to enlarge a partial sequence is larger. In the case of pattern B and C requiring 6 labels, **Table 1** shows the order of attribution of label pairs. The rows correspond to left labels of a pair and the columns correspond to right labels of a pair. The first pair of labels (index 0) in the sequence is 0-1, the second (1) is 1-2, etc, leading to the sequence:

“0123450213540314251043205241530”

It was obtained in one cycle. We finally add pairs 11, 22, 33, 44, 55 by inserting 11 where there is a 1, 22 where there is a 2, etc.

L\R	**0**	**1**	**2**	**3**	**4**	**5**
0	30	0	6	12	19	23
1	18		1	8	14	27
2	22	7		2	25	16
3	29	13	21		3	9
4	11	26	15	20		4
5	5	17	24	28	10	

Table 1: Index of attribution of label pairs

The complete sequence of labels is:

011234502133540314225104432055241530 0,

where the digits 0..5 are the six labels.

4. PROTOTYPES

4.1 Prototype A

Prototype A uses pattern A projected by a slide projector. This development was carried out in the framework of the M2VTS project [6] in 1997 [7]. The slide was realised by offprint on glass, which offers high contrast and rigidity against heat deformation. A database has been collected [8] and is available on request, free of charge:
(see http://www.sic.rma.ac.be/~beumier/DB/3d_rma.html).

Figure 3: 3D capture with prototype A

Although used with success in face recognition experiments, the acquired facial surfaces suffer from limited quality due to the camera resolution (768x576) and to the wrong localisation of edges implied by the digitisation board (edge shifting due to the low pass filter).

4.2 Prototype B

Prototype B [9] was developed in the framework of the BIOMET project [10] to benefit from the availability of a high-resolution colour digital camera (canon G2, 2272x1704 pixels). Pattern B was printed on a colour slide film stuck on a plastic support for rigidity and planarity. A database [10] has been collected and should be soon available through ELDA (Evaluations and Language resources Distribution Agency). The quality of the acquired surfaces is better than with Prototype A. However, the limited contrast of the slide and the reduced contrast of yellow and cyan stripes on many faces limited the quality.

Figure 4: 3D and texture capture with Prototype B

4.3 Prototype C

Prototype C, still in the development stage, reconsidered the projection of black and white stripes (pattern C) for maximal contrast, with the high-resolution digital camera canon G2. Until now, a video projector has been used to project the pattern, which limits the depth of field and hence the calibration performance. Compared to prototype B, prototype C benefits from a more precise localisation and better stripe visibility allowing for the capture of a larger facial surface area.

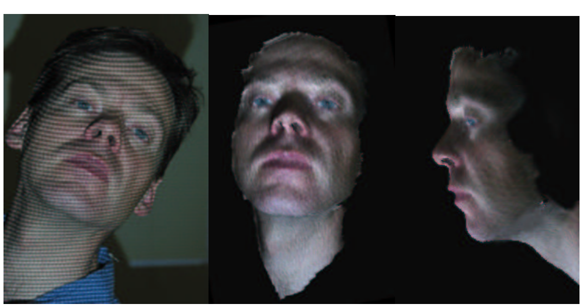

Figure 5: 3D and texture capture with Prototype C

5. CONCLUSIONS

This paper presents the design of projected patterns for 3D acquisition by coded structured light. Three patterns were described in terms of pattern elements and labels used for stripe identification. The way to build up a label sequence for unique identification was presented.

Prototype C offers the best quality thanks to the large contrast of the black and white stripes and the precision of a high-resolution digital camera. This implies a precise localisation of the stripes and the capture of a larger area of the facial surface.

REFERENCES

[1] C. Beumier, "Authentification d'identité par analyse tridimensionnelle du visage", PhD dissertation, ENST, Paris, 2003.

[2] J. Battle and E. Mouaddib, "Recent progress in coded structured light as a technique to solve the correspondence problem: a survey", *Pattern Recognition*, vol. 31, No 7, pp. 963-982, 1998.

[3] O. Hall-Holt and S. Rusinkiewicz, "Stripe Boundary Codes for Real-Time Structured Light Range Scanning of Mo-ving Objects", in *ICCV 8th Int. Conference on Computer Vision*, Vancouver, Canada, pp. 359-366, 2001.

[4] K. Boyer and A. Kak, "Color-Encoded Structured Light for Rapid Active Ranging", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, No 1, pp. 14-28, Jan. 1987.

[5] M. Proesmans and L. Van Gool, "Reading between the lines - a method for extracting dynamic 3D with texture", in *ACM Symposium on virtual reality software and technology, VRST 97*, Lausanne, Switzerland, Sept. 1997, pp. 95-102.

[6] "Multi-Modal Verification for Telesurveillance and Security Application", European project, ACTS programme (AC102), 1995-1998.

[7] C. Beumier and M. Acheroy, "3D Facial Surface Acquisition by Structured Light", in *International Workshop on Synthetic-Natural Hybrid Coding and Three Dimensional Imaging,* Santorini, Greece, Sept. 1999, pp. 103-106.

[8] C. Beumier and M. Acheroy, "SIC_DB: Multi-Modal Database for Person Authentication", in *Proceedings of the 10th International Conference on Image Analysis and Processing*, Venice, Italy, 27-29 Sept. 1999, pp. 704-708.

[9] C. Beumier, "Facial Surface Acquisition with Colour Striping", in *Proc. ICANN/ICONIP 2003*, Istanbul, Turkey, June 25-29 2003, pp. 548-551.

[10] S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J. Leroux-les-jardins, Jan Lunter, Yang Ni, and D. Petrovska-Delacrétaz, "BIOMET: A Multimodal Person Authentication Database Including Face, Voice, Fingerprint, Hand and Signature Modalities", in *Proc. AVBPA 2003*, Guildford, UK, June 9-11, 2003.