Salvi, Fernandez, Pribanic _ Llado- STate of the Art in Structured Light Patterns for Surface Profilometry

Friday, March 5, 2021   5:05 PM
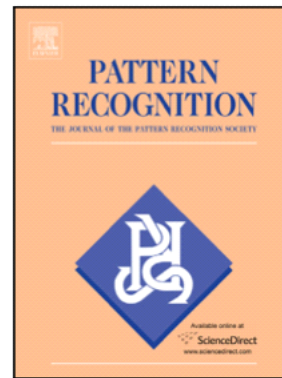
Salvi,
Fernande...

Author's Accepted Manuscript

A state of the art in structured light patterns for surface profilometry

Joaquim Salvi, Sergio Fernandez, Tomislav Pribanic, Xavier Llado

www.elsevier.com/locate/pr

Cite this article as: Joaquim Salvi, Sergio Fernandez, Tomislav Pribanic and Xavier Llado, A state of the art in structured light patterns for surface profilometry, *Pattern Recognition*, doi:10.1016/j.patcog.2010.03.004

A state of the art in structured light patterns for surface profilometry

Joaquim Salvi[a,*], Sergio Fernandez[a,**], Tomislav Pribanic[b], Xavier Llado[a]

[a]Institute of Informatics and Applications, University of Girona, Av. Lluis Santalo S/N, E-17071 Girona (Spain)
[b]Faculty of Electrical Engineering and Computing, University of Zagreb, Unska 3, HR-10000, Zagreb (Croatia)

Abstract

Shape reconstruction using coded structured light is considered one of the most reliable techniques to recover object surfaces. Having a calibrated projector-camera pair, a light pattern is projected onto the scene and imaged by the camera. Correspondences between projected and recovered patterns are found and used to extract 3D surface information. This paper presents an up-to-date review and a new classification of the existing techniques. Some of these techniques have been implemented and compared, obtaining both qualitative and quantitative results. The advantages and drawbacks of the different patterns and their potentials are discussed.

Key words: coded patterns, pattern projection, structured light, fourier transform profilometry, 3D measuring devices, active stereo, computer vision.

1. Introduction

Three dimensional measurement constitutes an important topic in computer vision, having different applications such as range sensing, industrial inspection of manufactured parts, reverse engineering (digitization of complex, free-form surfaces), object recognition, 3D map building, biometrics, clothing design and others. The developed solutions are traditionally categorized into contact and non-contact techniques. Contact measurement techniques have been used for a long time in reverse engineering and industrial inspections. The main problems of contact techniques are their slow performance and the high cost of

---

*Principal corresponding author. Phone: +34 972 41 8483. Fax: +34 972 41 8976
**Corresponding author. Phone: +34 972 41 9812. Fax: +34 972 41 8976
Email addresses: qsalvi@silver.udg.edu (Joaquim Salvi), sergiofn@eia.udg.edu (Sergio Fernandez), Tomislav.Pribanic@fer.hr (Tomislav Pribanic), llado@silver.udg.edu (Xavier Llado)

using mechanically calibrated passive arms [1]. Besides, the fact of touching the object is not feasible for many applications. Non-contact techniques were developed to cope with these problems, and have been widely studied. Non-contact techniques can be classified into two different categories: active and passive. In passive approaches, the scene is first imaged by video cameras from two or more points of view and correspondences between the images are found. It is important to mention that the cameras have to be previously calibrated [2]. The main problem experimented when using this approach is a sparse reconstruction since density is directly related to the texture of the object. This complicates the process of finding correspondences in presence of textureless surfaces [3]. Therefore, passive reconstruction is rather limited to reconstruct dense 3D surfaces, due to the problem of finding correspondences [4]. Methods based on structured light (active techniques) came to cope with this issue, creating correspondences and giving specific codewords to every unitary position in the image. In this approach one of the cameras is substituted by an active device (a projector), which projects a structured light pattern onto the scene. This active device is modeled as an inverse camera, being the calibration step a similar procedure to the one used in a classical stereo vision system [5]. The projected pattern imposes the illusion of texture onto an object, increasing the number of correspondences [6]. Therefore, surface reconstruction is possible when looking for differences between projected and recorded patterns.

In this paper, an exhaustive analysis of the different coding strategies used in active structured light is done, focusing on the advancements presented in the last years. A new classification regarding the strategy used to create the pattern is proposed, comparing some common characteristics between them. Feasibility and accuracy are analysed, giving both qualitative and quantitative results for the implemented techniques. This paper is structured as follows: section 2 presents a classification of the different techniques. Discrete pattern based codification is studied in section 3, while section 4 deals with the continuous ones. The results of implementing some of the most relevant techniques are showed in section 5, comparing their pros and cons. Finally, section 6 concludes with a discussion of the surveyed methods, pointing out advantages and disadvantages of the most relevant ones.
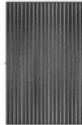
2

2. Classification

Coded structured light (CSL) systems are based on the projection of one pattern or a sequence of patterns that univocally determine the codeword of a pixel within a non periodic region. CSL has produced many works during the last decades and some recopilatory works can be found in the literature. This is the case of the surveys presented by Batlle et al. [4] and Salvi et al. [7], that analysed the different coded structured light techniques existing in temporal and spatial multiplexing domains until 1998 and 2004, respectively. Regarding frequency multiplexing, Su [8] reviewed the Fourier Transform Profilometry (FTP) techniques proposed until 2001. However, there is not any previous work comparing the three approaches together. Therefore, a classification extracting and analyzing attributes common in all the approaches is missing. This is overcome in the present survey, which also incorporates the most recent contributions done in CSL in the last years.

Table 1 shows a new classification of the existing pattern projection techniques. The main distinction has been done regarding the discrete or continuous nature of the pattern, rather than the codification process. Discrete patterns present a digital profile having the same value for the region represented by the same codeword. The size of this region largely determines the density of the reconstructed object. Besides, continuous patterns present a smooth profile where every pixel has a unique codeword within the non-periodicity region, assuring dense reconstruction. A posterior sub-classification is done regarding spatial, time and frequency multiplexing. Columns on the right indicate the value of some intrinsic attributes common to all the patterns. These attributes are:

- *Number of projected patterns*: determines whether the method is valid or not for measuring moving objects.
- *Number of cameras*: the method uses stereovision (2 or more cameras) coupled to a non calibrated pattern used only to get texture on the surface pattern, or a unique camera coupled to a calibrated projector.
- *Axis codification*: the pattern is coded along one or two axis.
- *Pixel depth*: refers to the color and luminance level of the projected pattern (B, G and C stands for Binary, Grayscale and Color respectively).
- *Coding strategy*: refers to the periodicity of the set of patterns projected on the surface (A stands for Absolute and P stands for Periodic).
- *Sub-pixel accuracy*: determines whether the features are found considering sub-pixel precision, thus providing better reconstruction results (Yes or No).

3

- *Color*: determines whether the technique can cope with colored objects (Yes or No).

4

Table 1: Proposed classification embracing every group of CSL.

			Author	Year	Shots	Cameras	Axis	Pixel depth	Coding strategy	Subpixel acc.	Color
DISCRETE	SPATIAL Multiplexing	De Bruijn	Beyer	1987	1	1	1	C	A	Y	N
			Salvi	1998	1	1	1	C	A	Y	Y
			Monks	1992	1	1	1	C	A	Y	N
			Pages	2004	1	1	1	C	A	Y	N
		Non formal	Forster	2007	1	1	1	C	A	Y	N
			Fechteler	2008	1	1	1	C	A	Y	N
			Tehrani	2008	1	2	1	C	A	N	Y
			Maruyama	1993	1	1	2	B	A	N	Y
			Kawasaki	2008	1	1	2	C	A	N	Y
			Ito	1995	1	1	2	G	A	N	Y
			Koninckx	2006	1	1	2	C	P	Y	Y
		M-Array	Griffin	1992	1	1	2	C	A	Y	Y
			Morano	1998	1	1	2	C	A	Y	Y
			Pages	2006	1	1	2	C	A	Y	N
			Albitar	2007	1	1	2	B	A	N	Y
	TIME Multiplexing	Binary codes	Posdamer	1982	>2	1	1	B	A	N	Y
			Ishii	2007	>2	1	1	B	A	N	N
			Sato	2006	>2	2	1	B	A	Y	Y
		N-ary codes	Caspi	1998	>2	1	1	C	A	N	N
		Shifting codes	Zhang	2002	>2	1	1	C	A	Y	N
			Sansoni	2000	>2	1	1	G	A	Y	Y
			Guhring	2001	>2	1	1	G	A	Y	Y
CONTINUOUS	FREQUENCY Multiplexing	Single Phase Shifting (SPS)	Srinivasan	1985	>2	1	1	G	P	Y	Y
			Ono	2004	>2	1	1	G	P	Y	Y
			Wust	1991	1	1	1	C	P	Y	N
			Guan	2004	1	1	1	G	P	Y	Y
		MultiplePhase Shifting(MPS)	Gushov	1991	>2	1	1	G	A	Y	Y
			Pribanic	2009	>2	1	1	G	A	Y	Y
		Single coding frequency	Takeda	1983	1	1	1	G	P	Y	Y
			Cobelli	2009	1	1	1	G	P	Y	Y
			Su	1990	2	1	1	G	P	Y	Y
			Hu	2009	2	2	1	G	P	Y	Y
			Chen	2007	1	1	1	C	P	Y	N
			Yue	2006	1	1	1	G	P	Y	Y
			Chen	2005	2	1	1	G	P	Y	Y
			Berryman	2008	1	1	1	G	P	Y	Y
			Gdeisat	2006	1	1	1	G	P	Y	Y
			Zhang	2006	1	1	1	G	P	Y	Y
			Lin	1995	2	1	1	G	P	Y	Y
			Huang	2005	>2	1	1	G	P	Y	Y
			Jia	2007	2	1	1	G	P	Y	Y
			Wu	2006	1	1	1	G	P	Y	Y
	SPATIAL multiplexing	Grading	Carrihill	1985	1	1	1	G	A	Y	N
			Tajima	1990	1	1	1	C	A	Y	N

5

3. Discrete coding methods

Discrete coding comprises all the methods where the pattern presents a digital profile. These methods are based on spatial or temporal multiplexing. Spatial multiplexing techniques code the pattern using the surrounding of a given feature, while temporal multiplexing creates the codeword by the succesive projection of patterns onto the object. In addition, some methods combine spatial and temporal information to take advantage of both techniques. Main contributions and latest proposals of both techniques are analysed herebelow.

3.1. Spatial multiplexing

Spatial multiplexing groups all techniques where the codeword of a specific position is extracted from surrounding points. Intensity or color variations are used to create the codeword. Three different coding strategies can be distinguished within this group: De Bruijn patterns, non-formal coding and M-arrays.

3.1.1. De Bruijn based techniques

De Bruijn sequences are a set of pseudo random values having specific properties between them. A k-ary De Bruijn sequence of order n is a circular sequence d_0, d_1, d_{n^k-1} (length $n^k - 1$) containing each substring of length k exactly once (window property of k). De Bruijn sequences can be constructed by taking a Hamiltonian or Eulerian path of an n-dimensional De Bruijn graph (see [9] for more details). This algorithm allows us to create univocal stripes sequences in the pattern, being able to extract the position by looking at the color of the stripes placed in the same window. Several proposals can be found using De Bruijn sequences, with both striped and multi-slit patterns. First proposals of De Bruijn-based striped pattern are found in the method developed by Boyer and Kak [10]. In this approach, RGB space was used to code the sequence of stripes. Being c_i^k the color of the stripe i of the sub-pattern k, the distance between two sub-patterns k and l is given by eq. (1):

$$d = \Sigma_{i=1}^{N} \delta_i \qquad (1)$$

where

$$\delta_i = \begin{cases} 0 \text{ if } c_i^k = c_i^l \\ 1 \text{ otherwise} \end{cases} \qquad (2)$$

The pattern proposed by Boyer and Kak [10] contains more than 300 stripes using three different colors. Color detection was done with a stripe indexing

6

algorithm preceded by a Hamming filtering. However, no color calibration was pursued to suppress the effect of different albedo, leading to some errors due to leackage from blue to green channels.

A different approach was followed by Monks et al. [11], where a multi-slit-based De Bruijn sequence was projected. A total of 6 colors were used to color the slits, separated by black gaps. The slit colors were chosen so that every subsequence of three colors appeared only once. Colors were chosen in the Hue channel (HSI space), despite the projection was performed in RGB and transformed back to HSI once the image was captured by the camera. Fully saturated and full intensity values were chosen in the other two channels. A previous color calibration step was performed by the authors in order to determine the transfer function of the optical system. Once the system was calibrated, captured colors were corrected before applying fringe detection. A minimum cost matching algorithm was used in the decoding step in order to find the most probable matching between projected and recovered patterns, provided that some slits appeared ocluded or not properly segmented [7].

To simplify the peak detection process, Salvi et al. [6] created a grid of horizontal and vertical colored slits. Every crossings of the two slits were extracted by simple peak intensity detection. Hue channel was used (in HSI space) to encode the colors. Three colors were assigned for the horizontal lines and other three for the vertical lines, using a De Bruijn sequence of order 3. The decoding step is done back in HSI space, showing negligible errors when scanning planar surfaces under scene light control. However, some problems were encountered due to the sensitivity of the Hue channel under different albedo of the illuminated object. Some years later, Pages et al. [12] [13] proposed an alternative approach to traditional striped or multi-slit-based pattern. They combined a striped pattern in the Hue channel with a multi-slit pattern in the Intensity channel (see Fig. 1), which defined dark and bright areas within the same color stripe. Therefore, the high resolution of classical striped patterns and the accuracy of multi-slit patterns were combined. The half illuminated stripes were colored according to a De Bruijn sequence for a sub-pattern of n stripes, while bright slits were colored equally within the same sub-pattern. In the experiments, a 128 striped pattern having 4 colors and a window property of 3 encoded stripes was applied. Using this codification, their approach doubled the resolution of traditional De Bruijn stripe-based techniques.
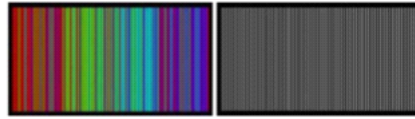
7

Figure 1: Pattern proposed by Pages et al. [12] (RGB pattern and luminance channel).

3.1.2. Non-formal coding

Non-formal coding comprises all the techniques having non-orthodox codification, in the sense that the authors create a specific pattern to fulfill some particular requirements. Both one-axis and two-axis encoding are suitable for these methods. One-axis coding methods are based on De Bruijn techniques. This is the case of Forster [14] and Fechteler and Eisert [15] proposals, which created color-based patterns in which two adjacent colors must differ in at least two color channels in the receptor device (red, green and blue). This condition is not usually accomplished in De Bruijn sequences. Forster used a striped pattern, while Fechteler and Eisert employed a multi-slit pattern. In Fechteler and Eisert a parabola was fitted in every RGB channel (or combination of channels for non-pure RGB colors, option selected by Forster). Optionally, surface color was acquired by projecting an extra white pattern. Tehrani [16] applied the idea of color slits to reconstruct images taken from two camera views, using 10 hue values to create the slit pattern (the difference between colors was maximal for adjacent slits).

There are also some proposals based on two-axis encoding. For instance, Maruyama and Abe [17] proposed a pattern of randomly cut black slits on a white background. In this approach, coding information was held in the length of the slits and their position within the pattern. Every recorded segment had its own length, which can be similar for several segments. The codeword corresponding to a segment was determined by its own length and the lengths of its 6 adjacent segments. The main drawback of this method is that the length of segments depended on the distances projector-object and object-camera, as well as on the camera optics, therefore reducing the reliability of the system. Another solution based on stripe lengths has been recently developed by Kawasaki et al. [3], who established a pattern of horizontal and vertical lines. In this work, the uniqueness of a specific location has been coded in the spacing between horizontal lines (in blue), whereas vertical lines (in red) have been equally

8

spaced. A peak detection algorithm was applied to locate the crossing points (dots) in the recovered image, and a posterior comparison with distances to neighboring dots determined their positions in the projected pattern. Ito and Ishii [18] did not use stripes or slits for coding, but a set of square cells (like a checkerboard), having one out of three possible intensity values. Every node (intersection between four cells of the checkerboard) was associated with the intensity values of the forming cells. In order to differenciate nodes having the same subcode, epipolar restrictions between the camera and the projector were employed. The idea of using epipolar restrictions was also applied in the work presented by Koninckx and Van Gool [19]. They proposed an adaptive system where green diagonal lines (named coding lines) were superimposed to a grid of vertical black lines (named base pattern). If a coding line was not coincident with an epipolar line, intersections created with the base pattern would all have laid on different epipolar lines on the camera image. This determines a unique point in the projected pattern, being able to perform the matching and the triangulation. A greater inclination of diagonal lines gave a higher density of the reconstruction, but a lower noise resistance. Therefore, the density of reconstruction could be chosen depending of how noisy the environment was, giving an adaptive robustness versus accuracy.

3.1.3. M-arrays

First presented by Etzion [20], M-arrays (perfect maps) are random arrays of dimensions $r \times v$ in which a sub-matrix of dimensions $n \times m$ appears only once in the whole pattern. Perfect maps are constructed theoretically having dimensions $rv = 2^{nm}$, but for real applications the zero submatrix is not considered. Thus, giving a total of $rv = 2^{nm} - 1$ unique submatrices in the pattern and a window property of $n \times m$. M-arrays represent in a two-dimensional space what De Bruijn patterns are in a one-dimensional space (see [20] and [21] for more details). Choosing an appropiate window property will determine the robustness of the pattern against pattern occlusions and object shadows for a given application. Morita et al. [22] proposed a two projection-based technique where an encoded matrix of black dots on a white background was projected, while in the second projection some black dots were removed according to a binary-encoded M-array. There are different approaches to represent non binary M-arrays, which are classified regarding the approach used to code the M-array: colored dots (color-based) or geometric features like circles and stripes (feature-based). For instance, Griffin et al. [23] generated an array of 18x66 features

9

using an alphabet of four words $1, 2, 3, 4$ comparing color and feature-based approaches. As the second approach is not color dependent, better results were obtained in presence of colored objects. Morano et al. [24] used a brute force (non-De Bruijn-based) algorithm to generate the pattern. An iterative algorithm adding one new code word and checking it against all the previous ones was performed. If all the distances between values were at least equal to the specified minimum Hamming distance, the new word was accepted and the next iteration was followed, until the pattern was created. The directions in which the pattern was created are indicated in Fig. 2.
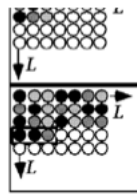

Figure 2: Code generation direction followed by Morano et al. [24].

This algorithm was posteriorly used by Pages et al. [25] to design a 20×20 M-array based on an alphabet of three symbols and a window property 3×3. A color approach was used for the dots codification, using Red, Green and Blue in order to separate them in the camera sensor. The decoding algorithm analysed the four neighbours of every dot. Once this was done, a comparison between all possible combinations of 8 neighbors was performed, in order to univocally locate the recorded dot in the projected pattern and perform the triangulation. A different approach has been followed by Albitar et al. [26], who used a 3×3 window property and three different symbols (black circle, circumference and stripe) to represent the codeword. As no color codification was employed, this solution presented robustness against colored objects. In the detection step, orientation of the projected pattern was extracted from the direction of the projected stripes. Once this is done, location of the symbols in the projected pattern was accomplished. Albitar et al. employed this method to create a 3D scan for medical imaging purposes (scanning of parts of the body), stating that this one-shot technique was robust against occlusions (up to a certain limit) and suitable for moving scenarios.

10

3.2. Time multiplexing

Time multiplexing methods are based on the codeword created by the succesive projection of patterns onto the object surface. Therefore, the codeword associated to a position in the image is not completely formed until all patterns have been projected. Usually the first projected pattern corresponds to the most significant bit, following a coarse-to-fine paradigm. Accuracy directly depends on the number of projections, as every pattern introduces finer resolution in the image. In addition, codeword basis tend to be small, providing higher resistance against noise. There are several approaches in discrete time multiplexing, which are exposed herebelow.

3.2.1. Temporal binary codes

These codes were firstly proposed by Posdamer and Altschuler [27] in 1982. A sequence of patterns having black and white stripes was projected onto the object. The number of stripes increased by two in every pattern, following a coarse-to-fine strategy. Therefore, the length of the codeword was given by 2^m bits, where m was the total number of projected patterns. The limit was given by the maximum resolution of both projector and camera devices. An edge detection algorithm was employed in the detection step to localize the transition between two consecutive stripes (black/white or viceversa). Moreover, Hamming distance between the codeword of two adjacent points could be maximized to reduce errors in the detection step, as was proposed by Minou et al. [28].

3.2.2. Temporal n-ary codes

Based on the use of n-ary codes, Caspi et al. [29] proposed a color based pattern where n^m stripes were coded in RGB space. The parameters to set were the number of colors to be used (N), the number of patterns to be projected (M) and the noise immunity factor alpha (α). For the calibration step, Caspi et al. proposed a reflectivity model given by eq. (3):

$$\underbrace{\begin{bmatrix} R \\ G \\ B \end{bmatrix}}_{\vec{C}} = \underbrace{\begin{bmatrix} a_{rr} & a_{rg} & a_{rb} \\ a_{gr} & a_{gg} & a_{gb} \\ a_{br} & a_{bg} & a_{bb} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} k_r & 0 & 0 \\ 0 & k_g & 0 \\ 0 & 0 & k_b \end{bmatrix}}_{K} \vec{P} \underbrace{\begin{Bmatrix} r \\ g \\ b \end{Bmatrix}}_{\vec{e}} + \underbrace{\begin{bmatrix} R_0 \\ G_0 \\ B_0 \end{bmatrix}}_{\vec{C}_0} \tag{3}$$

11

where $\vec{c}$ is the projected instruction for a given color, $\vec{P}$ is the non-linear transformation from projected instruction to actually projected intensities for every RGB channel, A is the projector-camera coupling matrix, K the reflectance matrix (constant reflectance in every RGB channel is assummed) and C_0 is the reading of the camera under ambient light.

3.2.3. Temporal hybrid codes

In order to reduce the number of projections, Ishii et al. [30] proposed a system where temporal and spatial coding were combined. The level of spatial or temporal dependance was given by the speed and accuracy requirements. For a given pixel $p(x, y)$ at time t of the projected pattern, the value was determined by using eq. (4):

$$I(x, y, t) = G(\lfloor \frac{x}{m} + t \rfloor \pmod{n}, y) \tag{4}$$

where

$$G(k, y) = G(\lfloor \frac{2^k y}{I_y} + \frac{1}{2} \rfloor \pmod{2}) \tag{5}$$

being G a binary image obtained from a camera at time t, n the space code size, m the light pattern width in the x direction, and I_y the image size in the y direction. There were n selectable code values for a pixel at time t, depending on the importance of temporal encoding or spatial encoding. As shown in Figure 3, combination of temporal and spatial information can be done from total temporal encoding (represented by $p = 1$) to total spatial encoding (given by $p = 8$). The parameter p is called the space coding weighter, as it provides an idea of how temporal or spatial the codification is.

Time multiplexing in stereo-vision

Time multiplexing can also be applied as an auxiliary support to stereo vision, as the matching process between the 2 camera images becomes easier when artificial features are created on the measuring surface. Sun et al. [31], proposed a mobile coded light-based stereo-vision system for measuring large surfaces, where a slit pattern is projected and subdivided in time, following a fine-to-coarse strategy.

3.2.4. Shifting methods in Discrete Time multiplexing

There are some discrete implementations that use the shifting of patterns to obtain dense reconstructions. This is the case of Sansoni et al. [32], Guhring [33]

12

(a) Spatio temporal representation of pattern.

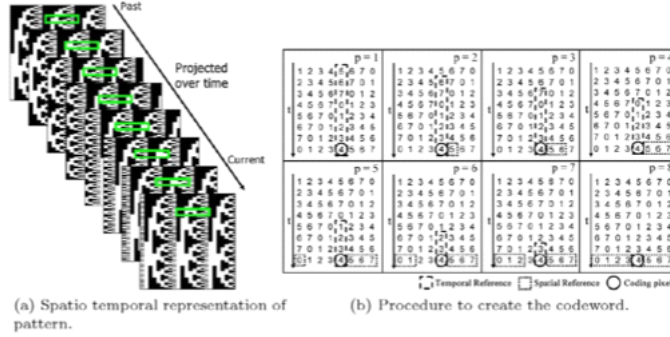(b) Procedure to create the codeword.

Figure 3: Spatio temporal algorithm proposed by Ishii et al. [30].

and Zhang et al. [34]. The proposals of Sansoni et al. and Guhring projected a set of black and white striped patterns (like in binary codes). Afterwards, the work of Sansoni et al. projected 4 shifted versions of the last pattern, while Guhring proposal projected shifted versions of a slit-based pattern covering every pixel in the image. Binary patterns provided an absolute location of the information given by shifted patterns, avoiding ambiguities in the decoding step. Using a different strategy Zhang employed color to project De Bruijn sequences, being smoothed and shifted versions of the same pattern. The smoothing process provided subpixel accuracy to this method. In order to avoid errors due to occlusions and discontinuities, multi-pass dynamic programming (a variance of the dynamic programming proposed by Chen et al. [35]) was employed to match observed to projected patterns.

4. Continuous coding methods

This group of methods is constituted by the set of patterns showing continuous variations on intensity or color throughout one or two axis. Among these methods, the use of periodic and absolute patterns can be found. Periodic patterns are used in time multiplexing phase shifting methods (PS) and in frequency multiplexing. Besides, absolute patterns are based on spatial grading.

13

4.1. Phase shifting methods

When projecting a sinusoidal grating onto a surface, every point along a line paralel to the coding axis can be characterized by a unique phase value. Any non-flat 3D shape will cause a deformation in the recorded pattern with respect to the projected one, which is recorded as a phase deviation. This phase deviation provides information about the illuminated shape. Matching the recovered image with the projected pattern, the object shape is recovered. The pattern must be shifted and projected several times in order to extract the phase deviation (this is not the case of frequency multiplexing approaches). Due to the grayscale nature of the projected patterns, they present advantages like resistance to ambient light and resistance to reflection variation. Depending on the number of frequencies used to create the pattern, we can distinguish between simple and multiple PS methods.

4.1.1. Single phase shifting methods

These techniques use only one frequency to create the sequence of patterns. In order to recover phase deviation the pattern is projected several times, every projection shifted from the previous projection by a factor of $\frac{2\pi}{N}$, being N the total number of projections, as shown in eq. (6) (super-index P indicates the projected pattern):

$$I_n^p(y^p) = A^p + B^p cos(2\pi f_\phi y^p - 2\pi n/N) \tag{6}$$

where A^p and B^p are the projection constants and (x^p, y^p) the projection co-ordinates, $n = 0, 1, ...N$. The received intensity values from the object surface, once the set of patterns is projected is:

$$I_n(x, y) = \alpha(x, y) [A + B cos(2\pi f_\phi y^p + \phi(x, y) - 2\pi n/N)] \tag{7}$$

As can be observed from eq. 7, it suffers of intensity and phase deviation, being necessary to cancel the effect of different albedo ($\alpha(x, y)$) to correctly extract the phase. This is shown in eq. (8):

$$\phi(x, y) = \arctan \left[\frac{\Sigma_{n=1}^N I_n(x, y) sin(2\pi n/N)}{\Sigma_{n=1}^N I_n(x, y) cos(2\pi n/N)} \right] \tag{8}$$

From a minimum of three projected shifted patterns is possible to create a relative phase map and to reconstruct the phase deviation caused by the object shape. However, the arctangent function returns values between the range

14

$(-\pi, \pi]$ and therefore a phase unwrapping procedure is necessary to work with a non-ambiguous phase value out of the wrapped phase. This is the reason why these patterns provide effective dense reconstruction only under the restriction of smoothed surfaces.

Phase shifting methods has been used in a variety of applications during the last years. For instance, Ono et al. [36] created the so called correlation image sensor (CIS), a device which generates temporal correlations between light intensity and three external reference signals on each pixel using phase shifting and a space-temporal unwrapping. Some approaches using phase shifting have also been developed from the work proposed by Srinivasan et al. [37].

One of the drawbacks of phase shifting methods is the necessity to project several patterns in time, which is more than the theoretic minimum of three patterns considered for real conditions. A solution to reduce the total time required in the projection step is to multiplex the patterns either in color space or in frequency. Following this idea, Wust and Capson [38] proposed a method which projected three overlapping sinusoidal patterns shifted 90 degrees, and coded in red, green and blue. Therefore, in this way the camera recorded phase deviation of every pattern in a different color channel and a normal phase extraction algorithm like the one shown in eq. (9) could be used:

$$\Phi(x,y) = arctan(\frac{I_r - I_g}{I_g - I_b})$$
(9)

where $\Phi(x,y)$ is the phase of a given pixel, and I_r, I_g and I_b are the red, green and blue intensities, respectively.

A different approach was proposed by Guan et al. [39], where the patterns were combined in frequency using the orthogonal dimension, as shown in Fig. 4. Basically, a traditional band pass filtering was performed to the recorded pattern, as it is theoretically done in communications for frequency multiplexing. This step filters noise without suppressing the information hold in the surroundings of the carriers. In particular, [39] used a maximally-flat magnitude Butterworth filter. Once this step was done, a normal phase extraction was performed over the obtained patterns. This method provided higher signal to noise ratio than color multiplexing approaches and it was not dependent on the surface color. However, some errors arose in presence of different albedo and abrupt shape variations.
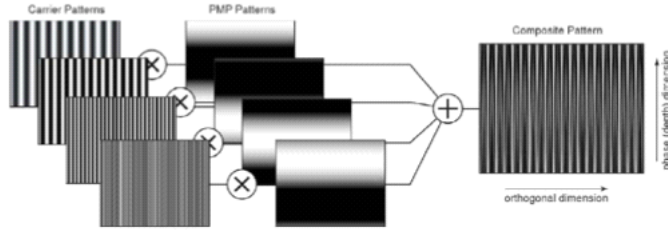
15

Figure 4: Composite pattern formed by tge multiplexation of modulated phase shifting profilometry (PMP) patterns using the algorithm of Guan et al. [39].

4.1.2. Multiple phase shifting (MPS) methods

The use of more than one frequency in phase shifting comes to cope with the uncertainty created in the extracted wrapped phase. As stated in the remainder theorem [40], an absolute phase map can be computed from two different relative phase maps having frequencies that are relative prime numbers. This principle was used by Gushov and Solodkin [41] for interferometry, where an interferometer able to deal with vibrations or relief parameters was constructed. More recently, Pribanic et al. [42] presented a multiple-phase shifting-based technique where only two patterns were used to create the relative phase maps. Two sinusoidal patterns were shifted and projected in time, in order to recover phase deviation (see Fig. 5). From these sets of images was possible to obtain two relative phase maps, using normal phase shifting decoding algorithms (as shown in eq. (8)). Having this, the absolute phase map was recovered. This map can be directly compared to the ideal phase-shifting map, providing correspondences for the triangulation step. The algorithm was tested for different pairs of frequencies over a flat surface. Finally the reconstruction of a footprint and a face were pursued, providing small 3D reconstruction errors.

4.2. Frequency multiplexing

Frequency multiplexing methods group all the techniques where phase decoding is performed in the frequency domain rather than in the spatial domain. There are different approaches depending on the frequency analysis performed to the image. Traditionally Fourier methods have been used for this purpose, although wavelet-based methods have also been studied.
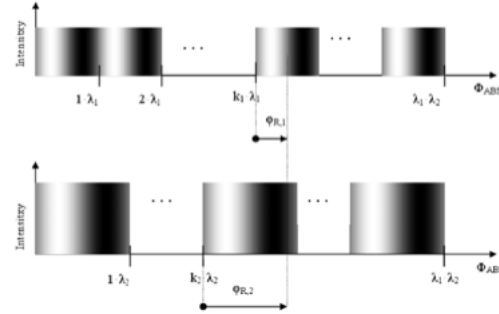
16

Figure 5: Pair of projected sinusoidal patterns, having two different frequencies (k is the number of periods).

4.2.1. Fourier Transform Profilometry

Fourier Transform Profilometry (FTP) was introduced to solve the necessity of having a phase-shifting-based method for moving scenarios. FTP was first proposed by Takeda and Mutoh [43], who extracted depth from one single projected pattern. A sinusoidal grating was projected onto the object, and the reflected deformed pattern was recorded. The projected signal for a sinusoidal grating was represented in eq. (10):

$$I_n^p(y^p) = A^p + B^p cos(2\pi f_\phi y^p) \qquad (10)$$

Once reflected onto the object, phase component was modified by the shape of the object, thus giving an intensity value expressed in eq. (11):

$$I(x,y) = \alpha(x,y)[A + Bcos(2\pi f_\phi y^p + \phi(x,y))] \qquad (11)$$

Phase component must be isolated to extract shape information. This was achieved performing a frequency filtering in the Fourier domain. The background component was suppressed and a translation in frequency was done to bring the carrier component (which holds the phase information) to zero frequency axis. Applying the sequence of equations showed below, the phase can be extracted from the signal. First, the input signal was rewritten as shown in eq. (12):

17

$$I(x,y) = a(x,y) + c(x,y)e^{2\pi i f_\phi y^p} + c*(x,y)e^{-2\pi i f_\phi y^p} \qquad (12)$$

where

$$c(x,y) = \frac{1}{2}b(x,y)e^{i\phi(x,y)} \qquad (13)$$

Finally, the phase component was extracted from the imaginary part of eq. (14):

$$log[c(x,y)] = log[(\frac{1}{2})b(x,y)] + i\phi \qquad (14)$$

The obtained phase component ranges from $(-\pi, \pi]$, being necessary to apply an unwrapping algorithm in order to obtain a continuous phase related to the object. Once the phase was unwrapped, the relative depth information was extracted using eq. (15):

$$h(x,y) = L \cdot \frac{\Delta\phi(x,y)}{(\Delta\phi(x,y) - 2\pi f_0 d)} \qquad (15)$$

where L is the distance to the reference plane and d the distance between the camera and the projector devices. FTP has been widely used in industrial applications. For instance, Cobelli et al. [44] used FTP for global measurement of water waves. In their work, two sources of noise were considered in the filtering step. The first one was related with illumination inhomogeneities of background variations over the field of view, which remains present as an additive variation. The second one was due to the local surface reflectivity. As this reflection varies much slower than the sinusoidal modulation impinged on the surface, it can also be treated as background noise. Thus, both sources of noise can be suppressed using the background component filtering procedure proposed by Takeda et al. [43]. Due to the periodic nature of the projected pattern, this method was constrained by the maximum reconstructible slope given by eq. (16):

$$|\frac{\partial h(x,y)}{\partial x}|_{MAX} < \frac{L}{3d} \qquad (16)$$

In order to increase this slope limitation, Su et al. [45] proposed the so called π-phase shifting FTP. Two sinusoidal patterns were projected using this method, being the second one a half-period shifted version of the first one. This solution multiplies by three the detectable range in depth slope. This principle was used by Hu and He [46] to scan moving objects having uniform velocity (like in an assembly line). In their work two scan line cameras were used, and one single

18

pattern was projected. The distance between the two cameras corresponded to half the period of the grating. As the velocity of the object was known, matching two scannings of the same point in different instants could be done. This procedure saves from projecting two patterns, and takes advantage of the movement present in assembly lines.

There are some proposals that combine both π-phase shifting FTP patterns in one single projected pattern using color or frequency multiplexing. For instance, Chen et al. [47] used color space to project a bi-color sinusoidal fringe pattern consisting of the sum of π-phase shifting FTP patterns, represented by blue and green patterns. Another approach was considered by Yue et al. [48]. In this work the same principle used by Guan for phase shifting was developed for FTP. Appropriate carrier frequencies were chosen regarding the characteristics of the projector and camera used, assuming that the Nyquist sampling theorem was satisfied. These frequencies were kept away from zero frequency as much as possible. When analyzing the results, standard deviation error is slight lower than for normal FTP, while accuracy remains unaltered.

In case of scanning coarse objects where discontinuities and speckle-like structures can appear, two dimensional FTP filtering must be used [8], as it permits better separation of the desired information from noise when dealing with coarse objects. This is due to the fact that noise is normally 2D distributed in a fringe pattern, having a spectra scattered in a 2D frequency domain. For instance, Hung and more recently Lin and Su [49] proposed a method for 2D FTP scanning where the filtering step, aimed to prevent from frequency spreading, was performed using a 2D Hanning window. However, some other filters having similar characteristics can also be used. This is the case of Chen et al. [50], who applies a Gaussian filter to perform this step. 2D FTP filtering has been used by Berryman et al. [51] to create a low cost automated system to measure three dimensional shape of the human back, obtaining an accuracy of $\pm 1 mm$.

The problem of phase unwrapping

Phase unwrapping respresents a crucial step in frequency multiplexing techniques. In absence of noise, if all phase variation between neighbouring pixels is less than π, the phase unwrapping procedure can be reduced to add the corresponding multiple of 2π when a discontinuity appears. Unfortunately, noise, local shadows, under-sampling, fringe discontinuities and irregular surface brightness make the unwrapping procedure much more difficult to solve. Plenty of approaches have been presented ([52], [53]). For instance, phase un-

19

wrapping based on modulation follows an iterative algorithm, starting from the pixel with higher intensity value and comparing it to the pixels inside a 3×3 surrounding square region. The comparison step is done one by one, queuing the affected pixels from maximum to minimum intensity. This method can also be applied when dealing with moving objects, substituting the searching area to a $3 \times 3 \times 3$ voxel. Besides, Wu and Peng [52] presented a phase unwrapping algorithm based on region growing. The phase was unwrapped from the smoothest area to the surroundings, according to a linear estimation. In order to decrease the error, a quality map was used to guide the unwrapping. The map can be defined in different ways as far as it provides quality information. For instance, second-order partial derivative can be used to determine the pixels to unwrap; that is, those pixels having this value lower than a specified threshold. Statistical methods can also be used considering the variance within a mask, for every pixel. Finally, Gorthi and Lolla [54] projected an extra color-coded pattern, which can be univocally identified once the image was captured, thus giving a rough information about the required phase to add or subtract in the unwrapping step. A further explanation of different unwrapping methods used in profilometry can be found in [55].

4.2.2. Wavelet transform Profilometry

The task of suppressing the zero component and avoiding the frequency overlapping between background and data has also been studied using other frequency-based approaches. For instance, Chen et al. [47] proposed the use of windowed Fourier transform (Gabor transform) to eliminate the zero spectrum. However, as was demonstrated by Gdeisat et al. [56], Chen's technique was not able to eliminate the zero spectrum neither in fringe patterns that have large bandwidths, nor in cases where the existence of large levels of speckle noise corrupts the fringe patterns. To cope with these problems, another frequency-based approach has been proposed based on wavelets analysis in the last years. This method used wavelets decomposition to solve the problem of frequency overlapping; that is, to extract and suppress the zero component automatically. Gdeisat et al. [56], applied a two dimensional wavelet function to the recovered image. Rotation and scale were considered jointly with x and y coordinates resulting in a four dimensional wavelet transform. To apply the transformation, the mother wavelet $\psi(x, y)$ must satisfy the admissibility condition. Under this condition Gdeisat used a Differential of Gaussian as the mother wavelet, while Zhang [57] employed a 2D complex Morlet wavelet. Four sub-images were

20

created at one iteration of the wavelet decomposition algorithm, corresponding to the low and high frequencies in both axis. Phase component was extracted from the ridge information present in the corresponding high frequency sub-image. The task of choosing appropriate values for rotation and scale parameters determined the results of filtering and phase extraction. Related to this, a novel method for choosing the adaptive level of discrete wavelet decomposition has been proposed by Zhang et al. [57]. They have achieved higher accuracy in the principal frequency estimation and low frequency energy suppression against traditional zero suppression algorithms used in FTP.

Alternatives to sinusoidal grating

Not all frequency transform methods use sinusoidal fringes for the projected pattern. As Huang et al. [58] stated, structured light techniques based on sinusoidal phase-shifting methods have the advantage of pixel level resolution, large dynamic range and few errors due to defocusing. However, the arctangent computation make them relatively slow. As an alternative, they used three 120 degrees phase-shifted trapezoidal fringe patterns. The phase deviation was extracted from the so-called intensity-ratio image, shown in eq. (17).

$$r(x,y) = \frac{I_{med}(x,y) - I_{min}(x,y)}{I_{max}(x,y) - I_{min}(x,y)} \tag{17}$$

where $I_{min}(x,y)$, $I_{med}(x,y)$ and $I_{max}(x,y)$ are the minimum, median, and maximum intensities of the three patterns for the image point (x,y). Image defocus does not cause major errors when using sinusoidal pattern, as it is still sinusoidal when the image is defocused. However, errors caused by bluring have to be taken into account when dealing with trapezoidal patterns. Modeling these errors as a Gaussian filtering, Huang et al. tests yielded defocusing errors not bigger than 0.6%. More recently, another approach using triangular patterns has been proposed by Jia et al. [59]. This approach used only two triangular patterns shifted half the period, making it more feasible to be implemented in real time applications. Ronchi grating has also been used in pattern projection as an alternative to sinusoidal grating. This is the case of Lin and Su [49], who proposed an algorithm where only one pattern was needed. Phase information was obtained taking the imaginary part of eq. (18):

$$\Delta\Phi(x,y) = log[\hat{I}(x,y)\hat{I}_0^*(x,y)] \tag{18}$$

21

where $\hat{I}(x,y)$ and $\hat{I}_0(x,y)$ are the recorded illuminance from the setup and the reference plane, respectively. A Ronchi grating was also used by Spagnolo et al. [60] in real applications, in order to recover 3D reconstructions of artwork surfaces.

4.3. Continuous spatial grading methods

Grading methods refer to all techniques containing the entire codeword for a given position only in its pixel value. Therefore, the resolution can be as high as the pixel resolution of the projector device is. However, these methods suffer from high sensitivity to noise and low sensitivity to surface changes, due to the short distances between the codeword of adjacent pixels. This is the reason why some authors use these methods introducing temporal redundancy, projecting the same pattern several times. As a drawback, note that restriction to static scenarios is imposed when projecting more than one pattern. There are two main techniques based on grading methods: grayscale-based patterns and color-based patterns. Regarding grayscale based methods, Carrihill and Hummel [61] proposed a linear grasycale wedge spread going from white to black, along the vertical axis. The authors achieved a mean error of $1cm$, due to the high sensitivity to noise and non-linearity of the projector device. In color-based patterns, the pixel is coded using color instead of grayscale values. As a drawback, color calibration is required. Tajima and Iwakawa [62] presented a rainbow codified in the vertical axis. In order to project this spectrum, a nematic liquid crystal was used to diffract white light. Two images were projected to suppress the effect of colored surfaces.

5. Experimental results

In order to test the effectiveness of the different strategies proposed in the literature a set of 6 representative techniques of table 1 have been implemented and compared. These methods are presented in table 2:

Three discrete coding and three continuous coding techniques have been chosen and implemented. It is important to mention that all the methods presented here have been implemented directly from the corresponding papers (original code was not available), and the parameters have been set in order to obtain optimal reconstruction results. Among discrete coding spatial multiplexing, one axis coding was chosen as it presents an easier decoding algorithm than two axis coding. Among them, Monks et al. [11] technique presents a color slits pattern

22

Table 2: Selected methods, with their main attributes.

	Group	Method	Characteristics
DC	Spatial m.	Monks et al. [11]	De Bruijn slits pattern. 6 Hue colors (1 pattern)
DC	Time m.	Posdamer et al. [27]	Stripes patterns. 7 bits Gray code (24 patterns)
DC	Time m. (PS)	Guhring [33]	Time multiplexing + shifting (16 patterns)
CC	Time m. (PS)	Pribanic et al. [42]	Multiple Phase Shifting (18 patterns)
CC	Frequency m.	Su et al. [45]	Sinusoidal pattern, π-shifting (2 patterns)
CC	Spatial m.	Carr. Hummel [61]	Grading grayscale pattern (1 pattern)

based technique that provides bigger vocabulary than grayscale approaches as well as easier detection and matching than stripes patterns techniques. Among discrete coding time multiplexing, Posdamer algorithm [27] was selected for being a well known effective technique in time multiplexing. Discrete time multiplexing using shifting codes are proposed by Sansoni et al. [32] and Guhring [33] to obtain dense reconstruction. Between them, Guhring method was selected because it uses slits shifting, easier to segment than the fringes shifting used by Sansoni et al. Among continuous coding approaches, the technique presented by Pribanic et al. [42] was selected for being the latest time multiplexing technique using multiple phase shifting. In continuous frequency multiplexing, π-phase shifting FTP method proposed by Su et al. [45] provides higher resistance to slopes than the traditional FTP of Takeda and Mutoh [43], without the necessity to perform Wavelet filtering or having to deal with bluring associated to non-sinusoidal patterns. Chen et al. [47] and Yue et al. [48] use the same π-phase shifting FTP multiplexing the patterns into one single projection. However, the main idea remains unaltered, and therefore the simpler solution proposed by Su et al. is still a good representative to evaluate the performance of these techniques. Finally, the greyscale spatial grading proposed by Carrihill and Hummel [61] was chosen against the rainbow pattern implemented by Tajima and Iwakawa [62] which employs a nematic liquid crystal.

The setup used for the tests was composed of an LCD video projector (Epson EMP-400W) with a resolution of 1024×768 pixels, a camera (Sony 3CCD) and a frame grabber (Matrox Meteor-II) digitizing images at 768×576 pixels with 3×8 bits per pixel (RGB). Both camera and video projector were calibrated using the projector camera calibration method developed by Zhang [63]. The baseline between camera and projector was about $1m$. The results and time estimates were computed using a standard Intel Core2 Duo CPU at 3.00GHz

23

and 4GB RAM memory. The algorithms were programmed and ran in Matlab 7.3.

5.1. Qualitative results

The reconstruction of a real object permits to analyse the performance of the programmed techniques in terms of accuracy and noise sensitivity. The reconstructed object used to perform the qualitative analysis of the results is a statue of a white horse of dimensions $21 \times 15 \times 10cm$ placed at a distance of about $80cm$ to the camera. In order to show the results, both 3D cloud of points and surfaces are used. The surface has been generated performing a 2D Delaunay triangulation over (x,y) coordinates. It is important to mention that no smoothing step has been applied to the 3D points or surfaces. Therefore the results are compared without a postprocessing step.

As can be observed in Fig. 6 and Fig. 7, the best results are obtained by shifting approaches, both in discrete coding (the case of Guhring [33]) and continuous coding (as shown by Pribanic et al. [42]). These techniques obtain the best accuracy results in time multiplexing, providing also dense reconstruction. Furthermore, both algorithms perform well in presence of surface slopes, as can be observed in some details of the reconstructed object (see for instance the ears of the horse). The holes presented in the surface of Fig. 7.b, 7.c correspond to discarded 3D points, due to dark regions or errors in the unwrapping step. As a drawback, it is important to mention the number of projections necessary to reconstruct the object, which make them unable to cope with moving scenarios. The original time multiplexing algorithm proposed by Posdamer et al. [27] has also been implemented in order to compare it to the other techniques. Despite the fact that the resolution obtained is higher than the other discrete coding techniques, it suffers some noise in the recovered cloud of points. This is mainly caused by unlinearities of the camera, which produces some leackage from white to black fringes that can lead to some errors in the position of the recovered edges. Moreover, time multiplexing techniques are also limited to static scenarios. Besides, Monks algorithm [11] employed De Bruijn color coding to obtain a dynamic discrete reconstruction. This algoritm provides good accuracy at expense of a low resolution. It is important to mention that this resolution is doubled in the approach of Pages et al. [12], which employed a more complex codification algorithm. In spite of using 6 color values in the Hue channel, previous color calibration is required. This can be simply done by comparing projected and recovered colors and working under light controlled

24

conditions (i.e. dark scenario), or performing a full color calibration step using a colorimeter, in order to obtain a model like the one shown in eq. (3). Frequency multiplexing is represented using the π-phase shifting approach proposed by Su et al. [45]. Recent proposals combined the two patterns into one single projection ([48], [47]). However, they followed the same initial ideas proposed by Su et al. As can be observed, dense reconstruction is obtained using only two patterns, which makes it able to work in dynamic scenarios. However, high frequencies are lost in the filtering step and the intensity distortion caused by the camera, loosing information in the surface details. Moreover, this technique can work only on smooth surfaces having slopes not exceeding three times the value given in eq. (16). The method chosen for phase unwrapping employed a qualitative map to determine the region where the unwrapping should start. Finally, the grading technique proposed by Carrihill and Hummel [61] resulted highly sensitive to noise and low sensitive to changes in depth, caused by the low range existing between adjacent pixels.

5.2. Quantitative results

Quantitative results have been analysed reconstructing a white plane at a distance of about $80cm$ to the camera. Principle Component Analysis (PCA) was applied to obtain the equation of the 3D plane for every technique and for every reconstruction. This technique is used to span the 3D cloud of points onto a 2D plane defined by the two eigenvectors corresponding to the two largest eigenvalues. The results of the experiment are shown in table 3. Observe that the algorithm of Su et al. [45] is conceived to measure deviation of smooth surfaces with respect to the reference plane, therefore a plane is not conceived to be reconstructed by depth deviation.

Table 3: Quantitative results. The headings are: author's name of the technique; average deviation of the reconstructing error; standard deviation of the reconstructing error; number of 3D points reconstructed; number of projected patterns.

Technique	Average (mm)	Stdev (mm)	3D Points	Patterns	Time (s)
Monks et al.	1.31	1.19	13899	1	45.29
Posdamer et al.	1.56	1.40	25387	14	32.18
Guhring	1.52	1.33	315273	24	158.22
Pribanic et al.	1.12	0.78	255572	18	165.65
Su et al.	—	—	—	1	—
Carr.and Hummel	11.9	5.02	202714	1	150.57

25

As can be observed, continuous coding approaches present a high number of reconstructed points, also achieved using a shifting technique in discrete coding (as shown by Guhring [33]). Among the techniques projecting more than one pattern (able to work only in static scenarios) the method developed by Pribanic et al. [42] gives the best results in terms of sensitivity to noise, as can be extracted from the values of average error and standard deviation. Similarly, the algorithm developed by Monks et al. [11] achieves the best performance among one-shot techniques (able to work in dynamic moving scenarios). Precision (expressed in terms of average error and standard deviation) keeps around $1mm$ in both cases, despite Pribanic et al. and Guhring algorithms provide visually better results due to their greater resolution. Finally, the poorest results are obtained by the spatial grading developed by Carrihill and Hummel [61] due to the low variance existing between adjacent pixels in the projected pattern. Regarding the computing time it can be observed that methods obtaining dense reconstructions (the case of Guhring, Pribanic et al., Su et al., and Carrihill and Hummel) need to compute more 3D points, requiring higher computational time. Among methods providing discrete reconstruction the color calibration step makes Monks et al. algorithm slower than Posdamer et al., despite it preserves the same order of magnitude. Finally, the computational cost of the algorithm proposed by Su et al. was tested using the dataset of previous section (qualitative results). The computational time was 108.52s, the same order of magnitude achieved in the other algorithms providing dense reconstructions.
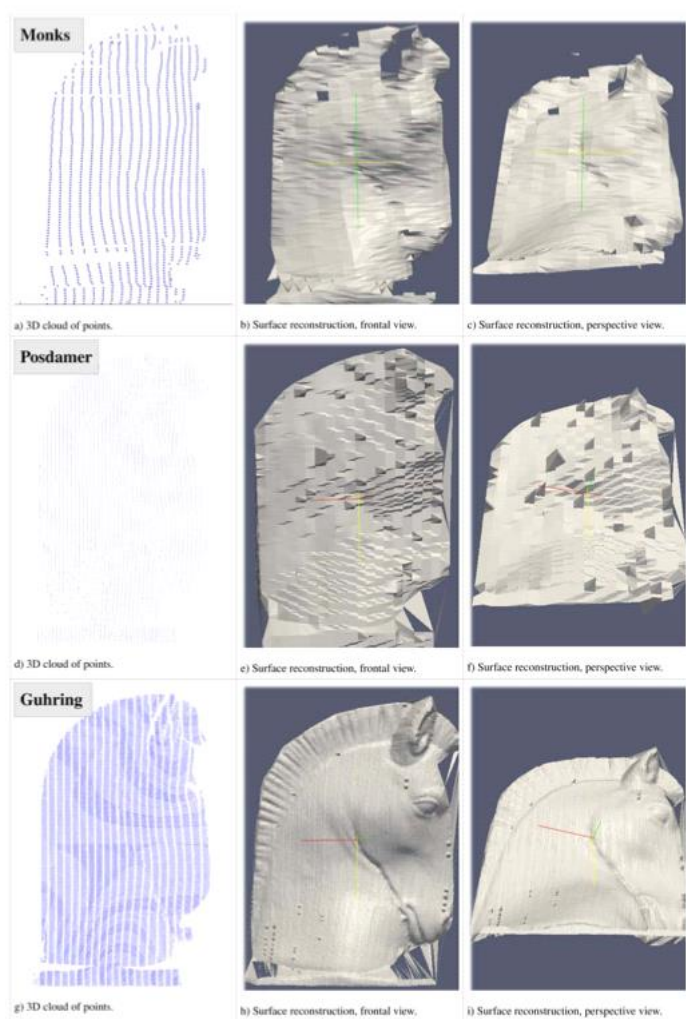
26

a) 3D cloud of points.
b) Surface reconstruction, frontal view.
c) Surface reconstruction, perspective view.

d) 3D cloud of points.
e) Surface reconstruction, frontal view.
f) Surface reconstruction, perspective view.

g) 3D cloud of points.
h) Surface reconstruction, frontal view.
i) Surface reconstruction, perspective view.

Figure 6: Results of Monks et al. [11], Posdamer et al. [27] and Guhring [33], respectively.
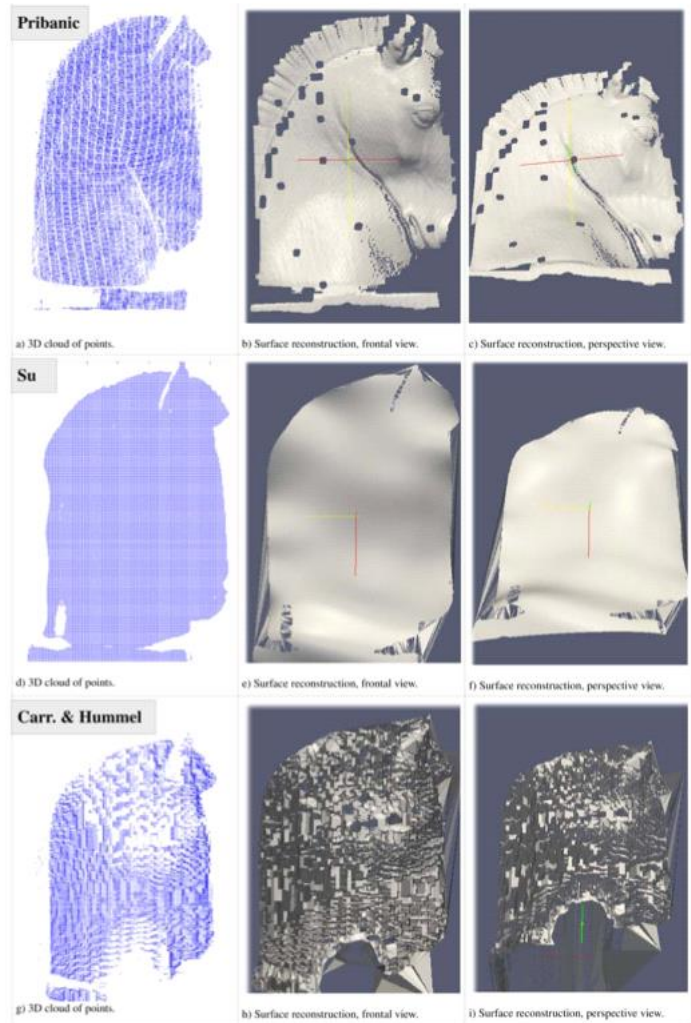
27

Figure 7: Results of Pribanic et al. [42], Su et al. [45], and Carrihill and Hummel [61], respectively.

28

6. Discussion

A new classification of the different CSL techniques has been proposed embracing and updating the spatial, temporal and frequency multiplexing strategies existing in the literature. Common attributes to all the techniques have been analysed and compared. Moreover, an update of the contributions done during the last years has been performed. Two main groups have been distinguished depending on the discrete or continuous nature of the projected pattern.

Discrete coding is created using stripes or slits in a unique axis codification, or geometric features (circles, checkerboard) or multi-slit for two axis codification. The final resolution depends on the number of encoded features present in the pattern. Among discrete coding techniques, spatial multiplexing and temporal multiplexing are distinguished. In the former, the codeword is determined by the pixel value in the pattern and the values of its surrounding pixels. In the later, the codeword is created by the sequence of patterns projected onto the surface. Spatial multiplexing needs only one or few patterns to create the codeword, being usually suitable for moving scenarios. However, they present lower spatial resolution than time multiplexing techniques, as all the information must be condensed in less projections. Among them, techniques based on De Bruijn codes, M-arrays and non-formal codification can be distinguished. De Bruijn codes create a striped or multi-slit based color pattern where the position in the image is determined by the color of a stripe (or slit) and the color of surrounding fringes. They present a trade-off between the resolution of the system and the window property (related with the separation between similar colors). The matching process must take the repetivity of the sequence into account, in order to avoid errors caused by occlusions. Multi-slit patterns present higher accuracy than striped patterns, at expense of lower resolution. Algorithms using slit-based patterns were proposed by Monks et al. [11] and Salvi et al. [6] in one and two axis coding, respectively. Besides, Pages et al. [12] proposed a combination of striped pattern in Hue channel with a multi-slit pattern in Intensity channel (in the HSI space), obtaining the advantages of both approaches. Other techniques create the color pattern having different values in at least two channels for adjacent stripes, in order to increase the resistance to errors ([14] and [15]). Regarding M-arrays (perfect maps), some efficient techniques have been proposed [24], [25], [26]. In spite of being difficult to generate, M-arrays take advantage of coding both axis to include higher degree of redundancy. There are two ways to represent M-arrays in a pattern: using color dots or using an

29

alphabet of symbols. As stated in [23] and [12], features based implementations are more robust against colored objects and can be segmented more easily. The coded points can be easily found in grid techniques using tracking algorithms, as they are placed in the intersections of edges. Besides, time multiplexing was the first paradigm of coded structured light used to obtain 3D data from an unknown surface. Having an easy implementation, time multiplexing methods achieve higher spatial resolution and accuracy than spatial multiplexing (in general terms), with the constraint of having to project several patterns, not being able for moving scenarios. Binary codes and n-ary codes have been proposed within this group. N-ary codes have a greater alphabet, so the number of patterns to project is reduced compared to binary codes. Moreover, the distance between pixels having different codewords decreases. However, this does not imply a higher sensitivity to noise for the alphabet size used in practice, as can be observed in the work of Caspi et al. [29]. Hybrid techniques combining spatial and temporal information have also been proposed [30]. Shifting approaches are grouped within time multiplexing, as more than one pattern are necessary to create the codeword. In discrete coding, a discrete pattern is projected and shifted over time, until all the object is covered. These are the cases of Zhang et al. [34] Sansoni et al. [32] and Guhring [33], who projected and shifted a De Bruijn smoothed pattern, a stripe-based pattern, and a slit-based pattern respectively, obtaining good accuracy results and dense reconstructions.

Continuous coding strategies achieve dense reconstruction by creating a pattern that changes the value between adjacent pixels. The depth of a given point in the image is determined by the deviation of its gray value with respect to the projected pattern. Among them, phase shifting techniques use the same principle mentioned above, but considering the information hold in the phase. Using Ronchi or sinusoidal patterns, phase deviation provides information of the surface shape when compared with the ideal phase map. These methods show good resistance to ambient light and to reflection variation, due to the grayscale nature of the projected patterns. For time multiplexing techniques is necessary to project at least three shifted patterns in order to suppress the effect of the albedo and to recover the phase deviation, in the so called phase shifting approaches [37], [36]. Combinations of shifted patterns into one single composite pattern have been proposed to achieve real time ([38], [39]). However, periodicity of the pattern imposes the assumption of smooth surfaces, as the presence of slopes would yield to some reconstruction errors. This problem is overcome in Multiple Phase Shifting approaches, which create an absolute phase map

30

from two relative phase maps, according to the ideas proposed by [40]. Pribanic et al. [42] used this principle to create dense reconstruction of surfaces having slopes. For frequency multiplexing techniques, phase decoding is performed in the frequency domain rather than spatial domain. Fourier methods have been traditionally used. From the first proposal of Takeda and Mutoh [43], an evolution using two projected patterns instead of one was proposed by Su et al. [45] in order to suppress the effect of background illumination in the phase extraction. This principle was preserved in Yue [48] and Chen et al. [47] approaches, who combined the two patterns in one single projection multiplexing them in the orthogonal axis or in different color channels, respectively. Other approaches used trapezoidal grating instead of sinusoidal grating [58], stating that the post processing becomes faster as there is no need to compute the arctangent function. However, errors arise due to defocusing, which does not affect the sinusoidal patterns. Finally grading methods, which belong to spatial coding methods, project one single pattern where the codeword is hold only in every projected point for that pixel. They achieve high spatial resolution with few projected patterns [61], [62]. However, sensitivity to errors is extremely high in these methods, as the received pattern can be affected by the resolution of the projector and the sensor device, or the reflectivity of the surface.

7. Conclusion

In this survey, an up-to-date review and a new classification of the different techniques existing in structured light has been proposed. The classification has been done regarding the continuous or discrete nature of the projected pattern, and group together the three different coding/decoding approaches used in CSL. A selection and implementation of representative techniques of every group has been done, and qualitative and quantitative comparisons have been performed extracting advantages and drawbacks of each technique. The results show the good performance of shifting approaches, which obtain dense reconstruction and excellent accuracy. However, they are only valid for static scenarios. Frequency multiplexing methods provide dense reconstruction for moving scenarios, but present high sensitivity to non linearities of the camera reducing the accuracy and sensitivity to details in the surface. Spatial multiplexing can also work in moving scenarios. Hence, the use of De Bruijn codes gives good accuracy in the reconstruction at expense of having discrete reconstruction and high sensitivity to changes in color surface or background illumination.

31

Summarizing the main contributions done in structured light in the last years, it is important to mention that most of the works have been concerned into frequency multiplexing approaches, trying to increase the robustness in the decoding step and the resistance to slopes under the constraint of moving scenarios ([48], [47], [56], [57]). Multiple phase shifting has arised also to overcome the problem of slopes in the objects [42]. Hybrid techniques have been the main contribution in both time multiplexing and spatial multiplexing approaches ([14], [15], [3], [25], [30]), preserving the principles proposed in previous works. Finally, continuous spatial coding has not experienced any contribution during the last years, due to its high sensitivity to noise.

Acknowledgments

References

[1] J. Vanherzeele, P. Guillaume, S. Vanlanduit, Fourier fringe processing using a regressive Fourier-transform technique, Optics and Lasers in Engineering 43 (6) (2005) 645–658.

[2] J. Salvi, X. Armangue, J. Batlle, A comparative review of camera calibrating methods with accuracy evaluation, Pattern recognition 35 (7) (2002) 1617–1635.

[3] H. Kawasaki, R. Furukawa, R. Sagawa, Y. Yagi, Dynamic scene shape reconstruction using a single structured light pattern, in: IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2008, pp. 1–8.

[4] J. Batlle, E. Mouaddib, J. Salvi, Recent progress in coded structured light as a technique to solve the correspondence problem: a survey, Pattern Recognition 31 (7) (1998) 963–982.

[5] S. Zhang, P. Huang, Novel method for structured light system calibration, Optical Engineering 45 (2006) 083601.

[6] J. Salvi, J. Batlle, E. Mouaddib, A robust-coded pattern projection for dynamic 3D scene measurement, Pattern Recognition Letters 19 (11) (1998) 1055–1065.

[7] J. Salvi, J. Pages, J. Batlle, Pattern codification strategies in structured light systems, Pattern Recognition 37 (4) (2004) 827–849.

[8] X. Su, W. Chen, Fourier transform profilometry: a review, Optics and Lasers in Engineering 35 (5) (2001) 263–284.

[9] H. Fredricksen, A survey of full length nonlinear shift register cycle algorithms, SIAM Review (1982) 195–221.

[10] K. Boyer, A. Kak, Color-encoded structured light for rapid active ranging, IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (1) (1987) 14–28.

32

[11] T. Monks, J. Carter, C. Shadle, Colour-encoded structured light for digitisation of real-time 3D data, in: Proceedings of the IEE 4th International Conference on Image Processing, 1992, pp. 327–30.

[12] J. Pages, J. Salvi, J. Forest, A new optimised de bruijn coding strategy for structured light patterns, in: 17th International Conference on Pattern Recognition, ICPR 2004, Vol. 4, 2004, pp. 284–287.

[13] J. Pages, J. Salvi, C. Collewet, J. Forest, Optimised De Bruijn patterns for one-shot shape acquisition 23 (2005) 707–720.

[14] F. Forster, A High-Resolution and High Accuracy Real-Time 3D Sensor Based on Structured Light, in: Proc.3th International Symposium on 3D Data Processing, Visualization, and Transmission, 2006, pp. 208–215.

[15] P. Fechteler, P. Eisert, Adaptive color classification for structured light systems, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008, pp. 1–7.

[16] M. Tehrani, A. Saghaeian, O. Mohajerani, A New Approach to 3D Modeling Using Structured Light Pattern, in: Information and Communication Technologies: From Theory to Applications, 2008. ICTTA 2008. 3rd International Conference on, 2008, pp. 1–5.

[17] M. Maruyama, S. Abe, Range sensing by projecting multiple slits with random cuts, IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (6) (1993) 647–651.

[18] M. Ito, A. Ishii, A three-level checkerboard pattern (TCP) projection method for curved surface measurement, Pattern Recognition 28 (1) (1995) 27–40.

[19] T. Koninckx, L. Van Gool, Real-time range acquisition by adaptive structured light, IEEE transactions on pattern analysis and machine intelligence 28 (3) (2006) 432–445.

[20] T. Etzion, Constructions for perfect maps and pseudorandom arrays, IEEE Transactions on information theory 34 (5 Part 1) (1988) 1308–1316.

[21] F. MacWilliams, N. Sloane, Pseudo-random sequences and arrays, Proceedings of the IEEE 64 (12) (1976) 1715–1729.

[22] H. Morita, K. Yajima, S. Sakata, Reconstruction of surfaces of 3-d objects by m-array pattern projection method, in: Computer Vision., Second International Conference on, 1988, pp. 468–473.

[23] P. Griffin, L. Narasimhan, S. Yee, Generation of uniquely encoded light patterns for range data acquisition, Pattern Recognition 25 (6) (1992) 609–616.

[24] R. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, J. Nissano, Structured light using pseudorandom codes, IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (3) (1998) 322–327.

[25] J. Pages, C. Collewet, F. Chaumette, J. Salvi, S. Girona, F. Rennes, An approach to visual servoing based on coded light, in: IEEE Int. Conference on Robotics and Automation, ICRA, Vol. 6, 2006, pp. 4118–4123.

[26] C. Albitar, P. Graebling, C. Doignon, Design of a monochromatic pattern for a robust structured light coding, in: IEEE Int. Conf. Image Process. ICIP, Vol. 6, 2007, pp. 529–532.

[27] J. Posdamer, M. Altschuler, Surface measurement by space-encoded projected beam systems, Computer Graphics and Image Processing 18 (1) (1982) 1–17.

33

[28] M. Minou, T. Kanade, T. Sakai, A method of time-coded parallel planes of light for depth measurement, Trans. IECE Japan 64 (8) (1981) 521–528.

[29] D. Caspi, N. Kiryati, J. Shamir, Range imaging with adaptive color structured light, IEEE Transactions on Pattern analysis and machine intelligence 20 (5) (1998) 470–480.

[30] I. Ishii, K. Yamamoto, K. Doi, T. Tsuji, High-speed 3D image acquisition using coded structured light projection, in: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2007, pp. 925–930.

[31] J. Sun, G. Zhang, Z. Wei, F. Zhou, Large 3D free surface measurement using a mobile coded light-based stereo vision system, Sensors & Actuators: A. Physical 132 (2) (2006) 460–471.

[32] G. Sansoni, M. Carocci, R. Rodella, Calibration and performance evaluation of a 3-D imaging sensorbased on the projection of structured light, IEEE Transactions on instrumentation and measurement 49 (3) (2000) 628–636.

[33] J. Gühring, Dense 3-D surface acquisition by structured light using off-the-shelf components, Videometrics and Optical Methods for 3D Shape Measurement 4309 (2001) 220–231.

[34] L. Zhang, B. Curless, S. Seitz, Rapid shape acquisition using color structured light and multi-pass dynamic programming, in: 3D Data Processing Visualization and Transmission,, 2002, pp. 24–36.

[35] C. Chen, Y. Hung, C. Chiang, J. Wu, Range data acquisition using color structured lighting and stereo vision, Image and Vision Computing 15 (6) (1997) 445–456.

[36] K. T. Ono N., Shimizu T., A. S., Real-time 3-D imager based on spatio-temporal phase unwrapping, Vol. 21, 2004, pp. 437–440.

[37] V. Srinivasan, H. Liu, M. Halious, Automated phase-measuring profilometry: a phase mapping approach, Applied Optics 24 (1985) 185–188.

[38] C. Wust, D. Capson, Surface profile measurement using color fringe projection, Machine Vision and App. 4 (3) (1991) 193–203.

[39] C. Guan, L. Hassebrook, D. Lau, Composite structured light pattern for three-dimensional video, Optics Express 11 (5) (2003) 406–417.

[40] P. Ribenboim, Algebraic numbers, R. Courant, L. Bers, J.J. Stoker. John Wiley and Sons. New York, 1972.

[41] V. Gushov, Y. Solodkin, Automatic processing of fringe patterns in integer interferometers, Opt. Lasers Eng 14 (4-5) (1991) 311–324.

[42] T. Pribanic, H. Dapo, J. Salvi, Efficient and Low-Cost 3D Structured Light System Based on a Modified Number-Theoretic Approach, EURASIP Journal on Advances in Signal Processing, In Press.

[43] M. Takeda M, Mutoh, Fourier transform profilometry for the automatic measurement of 3-D object shapes, Appl. Opt 22 (1983) 3977–3982.

[44] P. Cobelli, A. Maurel, V. Pagneux, P. Petitjeans, Global measurement of water waves by Fourier transform profilometry, Experiments in Fluids 46 (6) (2009) 1037–1047.

[45] J. Li, X. Su, L. Guo, Improved Fourier transform profilometry for the automatic measurement of three-dimensional object shapes (Journal Paper), Optical Engineering 29 (12) (1990) 1439–1444.

34

[46] E. Hu, Y. He, Surface profile measurement of moving objects by using an improved π phase-shifting Fourier transform profilometry, Optics and Lasers in Engineering 47 (1) (2009) 57–61.

[47] W. Chen, P. Bu, S. Zheng, X. Su, Study on Fourier transforms profilometry based on bi-color projecting, Optics and Laser Technology 39 (4) (2007) 821–827.

[48] H. Yue, X. Su, Y. Liu, Fourier transform profilometry based on composite structured light pattern, Optics and Laser Technology 39 (6) (2007) 1170–1175.

[49] J. Lin, X. Su, Two-dimensional Fourier transform profilometry for the automatic measurement of three-dimensional object shapes, Optical Engineering 34 (1995) 3297.

[50] W. Chen, X. Su, Y. Cao, Q. Zhang, L. Xiang, Method for eliminating zero spectrum in Fourier transform profilometry, Optics and Lasers in Engineering 43 (11) (2005) 1267–1276.

[51] F. Berryman, P. Pynsent, J. Fairbank, S. Disney, A new system for measuring three-dimensional back shape in scoliosis, European Spine Journal 17 (5) (2008) 663–672.

[52] L. Wu, Q. Peng, Research and development of fringe projection-based methods in 3D shape reconstruction, Journal of Zhejiang University-Science A 7 (6) (2006) 1026–1036.

[53] A. Baldi, F. Bertolino, F. Ginesu, On the performance of some unwrapping algorithms, Optics and Lasers in Engineering 37 (4) (2002) 313–330.

[54] S. Gorthi, K. Lolla, A new approach for simple and rapid shape measurement of objects with surface discontinuities, in: Proceedings. SPIE, Vol. 5856, 2005, pp. 184–194.

[55] T. Judge, P. Bryanston-Cross, A review of phase unwrapping techniques in fringe analysis, Optics and Lasers in Engineering 21 (4) (1994) 199–240.

[56] M. Gdeisat, D. Burton, M. Lalor, Eliminating the zero spectrum in Fourier transform profilometry using a two-dimensional continuous wavelet transform, Optics Communications 266 (2) (2006) 482–489.

[57] Q. Zhang, W. Chen, Y. Tang, Method of choosing the adaptive level of discrete wavelet decomposition to eliminate zero component, Optics Communications 282 (5) (2008) 778–785.

[58] P. Huang, S. Zhang, F. Chiang, et al., Trapezoidal phase-shifting method for three-dimensional shape measurement, Optical Engineering 44 (2005) 123601.

[59] P. Jia, J. Kofman, C. English, Two-step triangular-pattern phase-shifting method for three-dimensional object-shape measurement, Optical Engineering 46 (2007) 083201.

[60] G. Spagnolo, G. Guattari, C. Sapia, D. Ambrosini, D. Paoletti, G. Accardo, Contouring of artwork surface by fringe projection and FFT analysis, Optics and Lasers in Engineering 33 (2) (2000) 141–156.

[61] B. Carrihill, R. Hummel, Experiments with the intensity ratio depth sensor, Computer Vision, Graphics, and Image Processing 32 (3) (1985) 337–358.

[62] J. Tajima, M. Iwakawa, 3-D data acquisition by rainbow range finder, in: Pattern Recognition, 1990. Proceedings., 10th International Conference on, Vol. 1, 1990.

[63] Z. Zhang, A flexible new technique for camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (11) (2000) 1330–1334.

35