

Structured Light Stereoscopic Imaging with Dynamic Pseudo-random Patterns

Pierre Payeur and Danick Desjardins

University of Ottawa, SITE, 800 King Edward,
Ottawa, ON, Canada, K1N 6N5
{ppayeur, ddesjard}@site.uottawa.ca

Abstract. Structured light stereoscopic imaging offers an efficient and affordable solution to 3D modeling of objects. The majority of structured light patterns that have been proposed either provide a limited resolution or are sensitive to the inherent texture on the surface of the object. This paper proposes an innovative imaging strategy that accomplishes 3D reconstruction of objects using a combination of spatial-neighboring and time-multiplexing structured light patterns encoded with uniquely defined pseudo-random color codes. The approach is extended with the concept of dynamic patterns that adaptively increases the reconstruction resolution. Original techniques are introduced to recover and validate pseudo-random codes from stereoscopic images, and to consistently map color and texture over the reconstructed surface map. Experimental results demonstrate the potential of the solution to create reconstructions with various densities of points and prove the robustness of the approach on objects with different surface properties.

Keywords: stereoscopic vision, structured light, 3D modeling, pseudo-random patterns.

1 Introduction

The objective of the present work is to develop an integrated 3D imaging and reconstruction system that operates from affordable off-the-shelf equipment, provides the possibility to scan at various resolutions on demand, and operates on a wide variety of objects with different reflectance characteristics. Several techniques can be employed to estimate the shape and visual appearance of objects. A valid solution should be able to gather accurate 3D points that can be interpreted readily, with different density levels, and while only sacrificing processing time when deemed necessary. Many techniques arguably produce very high quality 3D maps but either require a large amount of processing or expensive devices to achieve them. The proposed solution provides entire flexibility between accuracy and processing time. It uses active stereoscopic vision and counts on a projected pattern with unique encodings. This ensures a high reliability in the feature matching procedure and a low dependency on the reflective properties and colors of the objects.

Under its basic formulation, stereoscopy usually suffers from a lack of reliable feature points, especially in man-made environments, or from unreliable matching

between the views. As an alternative, active vision methods perform exceptionally well in areas where no features can be readily detected, and ensure higher reliability on matches. Beyond time-of-flight solutions that are suitable over larger distances [1], or active laser triangulation that provides very high resolution but requires expensive devices [2], structured lighting remains an affordable but robust strategy. The main difference between structured lighting approaches remains in the nature of the projected pattern. The latter can be classified in three main categories [3]: time-multiplexing, spatial neighboring, and direct coding.

Time-multiplexing patterns are temporal codes that are projected sequentially over the scene. The codeword for a certain position is composed of the illumination values of a particular pixel position throughout all the patterns and is unique for every location. Binary Gray codes [4], [5], and n -ary codes that use linear or periodical variation over the entire range of intensity values [6], [7], are well established examples of time-multiplexing approaches. The sensitivity to external lighting level and reflectance properties remain high, and solutions based on these techniques are often constrained to uniformly shaded surfaces.

Spatial neighboring encodes positions as spatially distributed codes. Because a single pattern is projected, precision is usually not as high as with the one obtained when employing time-multiplexing patterns. A variety of such patterns have been proposed, that encode the alphabet of the sequence in different ways [8], [9], [10]. Pseudo-random arrays offer a simplified way to encode spatially distributed maps that are uniquely defined. Lavoie *et al.* [11] encode pseudo-random grids as bi-dimensional binary arrays (PRBA) composed of sequences in which any subset (or window) is unique and defines a specific location. They also extend the technique to pseudo-random multi-valued sequences (PRMVS) where multiple values are substituted by colors from a predetermined color palette [12].

Direct codification involves a complete pattern that uniquely labels every visible pixel and is projected only once. A location is then coded as a precise intensity level or color. Carrihill and Hummel [13] developed a sensor where a linear wedge spreads along columns containing varying grey levels. A ratio is calculated between the intensity values of every pixel under the linear wedge and under a pattern of uniform illumination to find the correspondence of pixels. Tajima and Iwakawa [14] suggest a rainbow pattern, where two images of the scene are taken with different color filters and the ratio between the images reveals a unique code. This method can theoretically reach high resolution. But practically the extraction of the correct intensity value for each location is difficult.

The proposed imaging technique originally combines the strengths of time-multiplexing and spatial neighboring structured lighting patterns to form dynamic pseudo-random codes.

2 Dynamic Structured Light Imaging System

The proposed imaging approach uses pseudo-random codes to create a projected pattern which is progressively shifted horizontally and vertically to increase the number of 3D reconstructed points. As such, advantage is taken of the same uniquely coded feature points several times. A stereoscopic acquisition system is designed to acquire

clouds of 3D points from which colored 3D surface models of objects are computed. The characteristics of the sensor and projected pattern are defined here.

2.1 System Description and Calibration

The active stereo rig used for experimentation consists of two Lumenera LU135C color CCD cameras with a 1392x1040 resolution, equipped with fixed focal 8.5 mm lenses, and linked together by a rigid bracket. An Electrohome EPS1024 projector with a resolution of 1024x768 is also used, as in Fig. 1a.

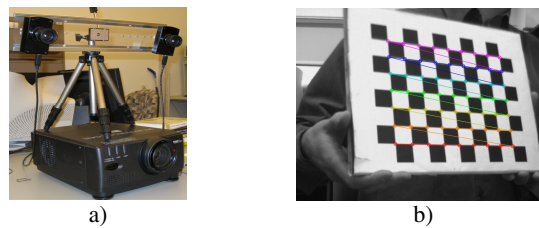


Fig. 1. a) Stereo structured lighting acquisition system, and b) calibration pattern

The use of a stereoscopic approach rather than a classical structured light system with only one camera [10] eliminates the need for calibration between the projector and the cameras. This solution gives access to focusing, zooming and brightness settings of the projector to adapt to various operating conditions and depth without influencing the calibration of the acquisition system. The respective intrinsic and extrinsic parameters of the two cameras are estimated successively using Zhang's calibration algorithm [15]. The calibration procedure is facilitated by a custom calibration routine with audio cues that guides the user to properly position the checkerboard pattern in the overlapping field of the view of the cameras as shown in Fig. 1b. Ten images of the checkerboard pattern are required to complete the calibration.

In the current configuration, the sensor's depth of view ranges from 693 mm to 3130 mm. The minimum depth is limited by the focusing capability of the projector, while the maximum is constrained by the power of the projector's lamp to ensure that features are accurately perceived by the cameras. At minimum and maximum depth, the sensor respectively achieves a resolution of 0.61 mm and 1.53 mm.

2.2 Pseudo-random Pattern Definition and Projection

A bi-dimensional pseudo-random (PR) pattern defined as a grid of square colored regions is projected on the surface of the object to be modeled. Color patches are selected instead of intersecting lines because they reduce the ambiguity in the color recognition and can be projected with high resolution using a LCD projector. Uniquely defined PR codes composed of a 3x3 neighborhood of square color regions ensure simplicity in the pattern generation and robustness in the image acquisition. They also provide enough information for the integration of a minimum Hamming distance criterion between codes to ensure higher discrimination.

The series of 3x3 codes that compose the PR array is generated off-line for a selected number of colors, k , following a pseudo-random iterative process [16]. Squares of a random color are successively added to the array while testing for the uniqueness of the resulting 3x3 color code throughout the grid [17]. The result is a grid of square color regions where every 3x3 sub-array appears at most once in the complete pattern of $N \times M$ color elements. Since the size of a color code is 3x3, a nine-element vector $V_{(ij)}$ defines every local color code such as:

$$V_{(ij)} = \{V_{(ij)1}, V_{(ij)2}, V_{(ij)3}, V_{(ij)4}, V_{(ij)5}, V_{(ij)6}, V_{(ij)7}, V_{(ij)8}, V_{(ij)9}\} \quad i=1, \dots, N, j=1, \dots, M \quad (1)$$

Every $V_{(ij)}$ must also differ by a defined minimum Hamming distance h in order to be valid. The Hamming distance between two codes $V_{(ij)}$ and $V_{(i'j')}$ is defined as:

$$H(ij, i'j') = \sum_{r=1}^9 (\delta_r) \quad \text{where: } \delta_r = \begin{cases} 0 \rightarrow v_{(ij)r} = v_{(i'j')r} \\ 1 \rightarrow v_{(ij)r} \neq v_{(i'j')r} \end{cases} \quad (2)$$

The PR pattern composed of $N \times M$ elements is smaller than the pixel resolution of the projector such that the extra pixels available are used to mark a delimitation of n black pixels around every color region. This makes color code segmentation and recovery more robust when scanning scenes with varying depth and sharp inclinations. As colored regions are skewed according to the orientation and distance to the surface, various colored region sizes tend to appear in the images. Each color element is projected as a color region of $P \times Q$ pixels, separated by rows and columns of n pixels. Moreover, increasing the number of colors, k , in the pattern enlarges the number of possible code permutations. Also, the k color components can be interchanged to provide adaptation to colors already present in the scene. Selected colors from the color palette should however remain far apart in the hue space.

For our experiments, P , Q and n , were all set to 9 pixels. The PR pattern contains 55x40 color codes with a palette size of $k=3$ colors, namely *Red*, *Green*, and *Blue*, as shown in Fig. 2a.

The projection of the pattern is dynamic in that the color grid is shifted horizontally and vertically several times, resulting in the projection of the same PR color pattern at several locations over the scene. Stereoscopic images are collected in synchronization for every location of the PR pattern. As a result, a large number of unique feature points are created. This leads to a dense 3D reconstruction of the objects, in spite of a relatively low number of codes required in the PR pattern.

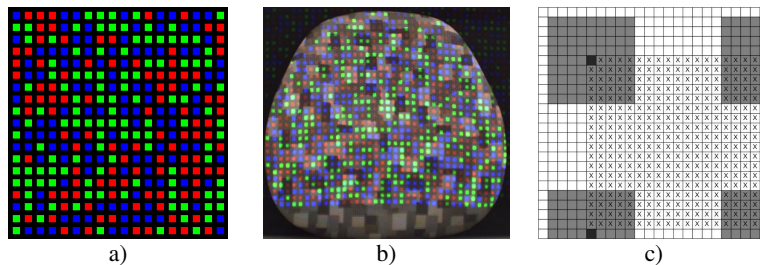


Fig. 2. a) Pseudo-random pattern, b) its projection over an object, and c) dynamic projection of the pseudo-random pattern to increase resolution

The pattern is successively shifted from left to right by Δx pixels for u increments and from top to bottom by Δy pixels for v increments. These shifting parameters can be adjusted on demand to suit the desired resolution of the reconstruction. The dynamic pattern projection is illustrated in Fig. 2c, where the X symbols represent the successive line and column positions of the center point (dark) of a given colored region (gray area). Because the color regions contain 9×9 pixels and are separated by black borders of $n=9$ pixels, a maximum of $u=18$ horizontal and $v=18$ vertical positions of the pattern, shifted by $\Delta x=\Delta y=1$ pixel each, are possible. This provides a maximum increase of the resolution by 324 times what is permitted with the initial 55×40 array of codes. Overall, a maximum of 712800 uniquely matching features can be created and their position estimated in 3D space.

In comparison with solely time-multiplexing solutions, the proposed dynamic PR codes do not intrinsically require multiple projections to estimate the shape of an object. A single pattern projection ensures a fair distribution of features points over the entire projection surface. The dynamic pattern approach augments the resolution on demand, and only if required. It represents an original combination of spatial neighboring and time-multiplexing structured light techniques. Accuracy can be set at variable levels by selecting the desired number of horizontal and vertical shifts of the dynamic pattern, depending on the acquisition and computation time available.

3 Image Processing and Code Recovery

The calibrated stereoscopic pair of cameras collects images of the scene with virtual features for every position of the PR pattern. The unique color codes extracted from the pattern ensure a reliable match between the left and right views. This leads to accurate 3D point estimation. The image processing performed on every pair of images and the validation of color codes are described in this section.

3.1 Color Regions Extraction and Grouping

The extraction of the PR codes from one pair of color images is performed in a series of steps. First, color regions are segmented from the black subdivision areas in the HSV color space. An intensity histogram is computed on the hue dimension to reduce the number of parameters required to properly segment an image. The peaks of the histogram determine the dominant colors in the image and isolate the corresponding regions that are mapped onto k separate binary sub-images. An evaluation of the peaks and valleys of the histogram with adaptive threshold is performed to discriminate the k dominant chromatic components. This approach reveals to be robust to large variations in the mapping of colors in the images.

Next, group labeling is performed on every sub-image to cluster consistent regions based on their geometrical proximity. However, given that the projected pattern tends to create color blobs with non-uniform intensity distribution, some color regions adopt variable shades of a color, and can therefore be incorrectly segmented, as shown in Fig. 3a.

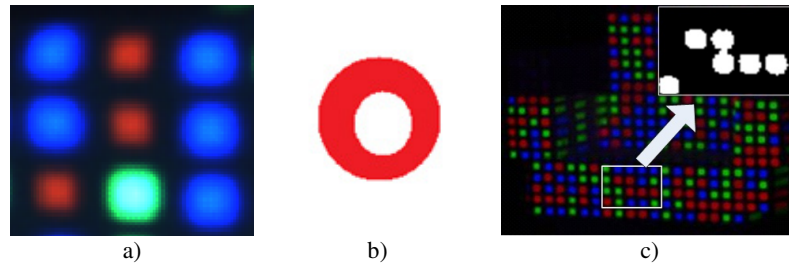


Fig. 3. a) Segmentation and code recovery over non-uniform color distribution, b) regions with holes, and c) combined color regions

To overcome this situation and ensure reliable color code extraction, the third step performs a statistical group analysis over the dimension of each labeled region, and either drops small blobs or tags groups that are larger than the average region for further subdivision. The topology of extracted regions is first analyzed to detect holes appearing within a colored region, as shown in Fig. 3b. Such empty regions are filled with the same color as their bounding pixels. The average size and standard deviation of the pixel count in every labeled region is calculated and a threshold is applied on the size of the regions. The regions that are larger than the average size plus one standard deviation are further segmented in the next step. The regions that are under 5 pixels are considered noise and are discarded.

The fourth step splits larger regions into smaller groups to eliminate merged color regions. The latter can appear in the images due to the skew created by variable surface orientations, as shown in Fig. 3c. Such situations are detected from the assumption that all color regions should be approximately of the same size over a local neighborhood in the image. Any regions that are significantly larger are further segmented. The average intensity, V channel of the HSV color space, and its standard deviation are calculated over the entire area of the merged color regions. The addition of the average intensity and its standard deviation is used as a threshold that is locally applied to eliminate lower intensity pixels. Such pixels being typically found on the borders of color regions, this procedure provides the desired disconnected regions.

Finally, a second group labeling operation is performed on the thresholded images and the detected color regions are added to a group list. The result is k lists of groups, one for every dominant color found, which are then used to recover the PR codes.

3.2 Pseudo-random Code Validation

Pseudo-random codes are recovered by traversing the list of color regions throughout the k color image masks. The nine closest geometrical regions to a given color region are determined by calculating the square distance between the current color region and all remaining color regions in the respective color masks. Given that a code is composed of 3x3 neighboring color regions, the nine color regions with the smallest square distance are considered as being part of a code. After being identified based on their geometrical proximity to a given color region, the nine closest color blobs are sorted vertically and horizontally to determine their correct spatial location.

Next, code validation is performed by computing a confidence map on the retrieved color codes. Every element in a 3x3 color code, besides the middle element, is part of other 3x3 codes. This means that a given code will also have eight neighboring codes containing at least one of its elements, except on the contours of the PR pattern. These neighboring codes are extracted from the captured image and compared to the known original projected pattern. The number of neighboring codes that can be locally matched between the image and the original pattern defines the confidence, $S(x)$, attributed to the given code, $C(x)$, which is defined as:

$$S(x) = \sum_{i=x-1}^{x+1} \sum_{j=x-1}^{x+1} [C_{image}(i, j) \wedge C_{original}(i, j)] \quad (3)$$

where \wedge represents the intersection logical operator.

Codes that are not present in the original projected pattern are discarded. Any duplication of codes is detected and the confidence level of such codes is analyzed. The codes with the highest confidence level are kept and the others are discarded.

4 3D Reconstruction and Texture Mapping

From the list of PR codes recovered that pass the confidence level criterion, every unique PR code found in the left and right images represents a reliable matching feature point artificially created over the scene. The final operation consists of computing the location of the corresponding physical points in 3D space and mapping their color and texture to provide an accurate model that can be readily interpreted.

In order to eliminate outliers and produce smoother 3D models, a random sample consensus (RANSAC) algorithm is applied on the resulting list of points corresponding to the recovered and validated color codes. The optimal polynomial reconstruction algorithm of Hartley *et al.* [18] which uses a triangulation method with a minimizing criterion based on the epipolar geometry is applied to reconstruct the 3D cloud of points corresponding to every matched color coded feature point. This final phase produces a list of 3D points that model the surface shape.

To ensure that the color and texture information mapped on the model is representative of the actual characteristics on the surface of the object, provision is taken to collect an image of the scene without any pattern being projected prior to the 3D acquisition. As the sensor is kept at a constant position and orientation throughout the entire acquisition phase, the color images collected a priori offer a pixel-wise correspondence with the series of images collected while the pseudo-random pattern is dynamically shifted over the scene. For every 3D reconstructed point, the corresponding pair of 2D pixel matches is queried for their associated RGB color information. The color parameters associated to a given 3D point is determined by averaging the corresponding pixels color from the left and the right images.

In the present work, an OpenGL viewer is developed that interpolates the color information between the defined vertices and fills the color over missing regions by interpolation. As a result, accurate color information is associated with every surface patch displayed along with the shape of the objects. This enables a visually consistent model, both in shape and texture, when compared to the original characteristics of the scene.

5 Experimental Evaluation

Experimentation with the proposed dynamic structured light stereoscopic sensor was conducted on several types of objects with different reflectance properties, surface complexity, color distribution and intensities. Due to space limitations, only a subset of 3D models is presented to demonstrate the quality of colored reconstruction.

The first object considered is a craft basket whose surface exhibits small variations in depth and contains a combination of matt and highly reflective areas, as shown in Fig. 4. This object was scanned at maximum resolution with 324 shifted patterns.

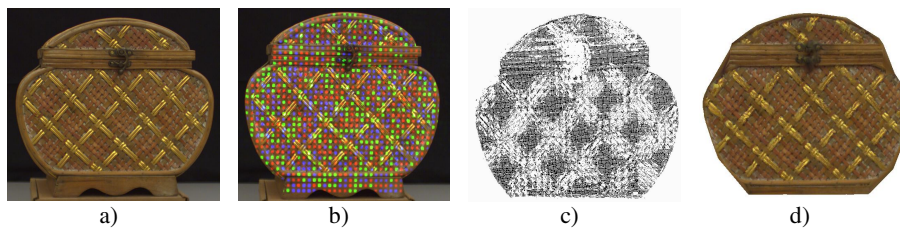


Fig. 4. 3D modeling of a basket with variable reflectance properties: a) color image of the basket, b) PR pattern projection, c) surface mesh reconstruction of the front face, and d) textured 3D model

The reconstruction is accurate considering the complexity of the textures present on the object. The density of reconstructed 3D points remains lower over the diagonal golden stripes that exhibit high reflectivity. But the successive projections of the dynamic PR pattern collect sufficient information over those areas. A high density of points is achieved over the non-reflective strands of the basket in spite of the non-uniformity and roughness of the surface. The color mapping over the 3D surface mesh is accurate and clearly renders the information about all regions of the basket's face.

The second example presented in Fig. 5a is a heavily textured chair back rest that exhibits non uniform colors and varying degrees of curvature over its surface. The back rest was scanned with only 25 shifts of the PR pattern previously defined.

Even though only 8% of the possible projections were used, the reconstruction of the back rest, which contains 35142 points, is very accurate. The curvatures of the object clearly appear in Fig. 5b where the front and lateral views of the surface mesh are presented. The patches of dark color present on the chair do not influence the reconstruction process and the model shows uniform density in spite of the strong variation in texture on the object. When colors are mapped on the surface mesh, the subtle details of the textures are crisp and clearly visible, as shown in Fig. 5c.

The comparison between the colored reconstructions on these two objects with the original images demonstrates that the proposed 3D imaging approach can provide dense reconstruction of surface shapes and yield colored representations that are very close to the original texture of the object, in spite of their inherent complexity. In terms of performance, the stereoscopic imaging system takes on average 13.9 seconds to perform the acquisition, image processing, code recovery, 3D reconstruction and texture mapping steps for one projection of the PR pattern. This represents less than

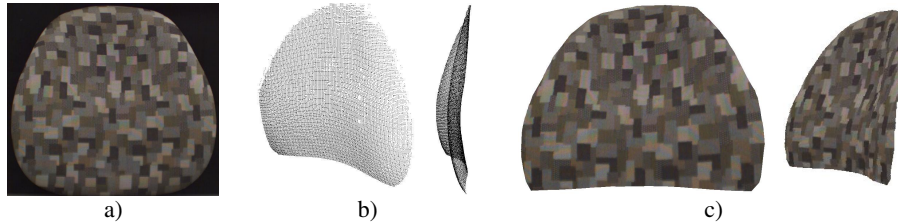


Fig. 5. 3D modeling of a chair back rest exhibiting heavy and non-uniform color patterns: a) color image of the back rest, b) surface mesh reconstruction from 2 perspectives, and c) textured 3D model from 2 perspectives

10ms per reconstructed 3D point, which is acceptable for a wide range of applications. The possibility to adjust the number of shifted patterns on demand also offers a maximum of flexibility on the desired resolution and processing time.

6 Conclusion

This paper presents an original structured light stereoscopic imaging technique that runs entirely from affordable off-the-shelf equipment, and combines spatial neighboring and time-multiplexing structured light approaches to build 3D colored model of objects with an adaptive resolution. The respective strengths of both structured lighting techniques are combined in an efficient and integrated way. The concept of dynamic projection of pseudo-random patterns is introduced to provide the necessary flexibility to achieve dense 3D model when required with no modification to the projected pattern and to the acquisition system. A robust approach to recover colored pseudo-random codes is presented that builds upon statistical code validation to minimize the occurrence of mismatches while remaining independent from the configuration of the objects, their reflectance properties and textures. The resulting models are full colored reconstructions of the objects that enable accurate representations and make interpretation easier. Experimental results on objects with complex surface properties demonstrate the quality and reliability of the models obtained with the proposed 3D imaging technique.

References

1. Marszalec, J., Myllylä, R.: Shape Measurements Using Time-Of-Flight-Based Imaging Lidar. In: SPIE Conference on Three-Dimensional Imaging and Laser-based Systems for Metrology and Inspection III, vol. 3204, pp. 14–15. SPIE Press, Pittsburgh (1997)
2. Blais, F., Taylor, J., Cournoyer, L., Picard, M., Borgeat, L., Godin, G., Beraldin, J.-A., Rioux, M., Lahanier, C.: Ultra High-Resolution 3D Laser Color Imaging of Paintings: the Mona Lisa by Leonardo da Vinci. In: Castillejo, et al. (eds.) Lasers in the Conservation of Artworks, pp. 435–440. CRC Press Taylor & Francis Group, London (2008)
3. Salvi, J., Pagès, J., Batlle, J.: Pattern Codification Strategies in Structured Light Systems. *Pattern Recognition* 37(4), 827–849 (2004)

4. Scharstein, D., Szeliski, R.: High-Accuracy Stereo Depth Maps Using Structured Light. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 195–202. IEEE Press, Madison (2003)
5. Sansoni, G., Corini, S., Lazzari, S., Rodella, R., Docchio, F.: Three-Dimensional Imaging Based on Gray-Code Light Projection : Characterization of the Measuring Algorithm and Development of a Measuring System for Industrial Applications. *Applied Optics* 36(9), 4463–4472 (1997)
6. Zhang, S., Huang, P.S.: High-Resolution, Real-Time 3D Shape Acquisition. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 28–38. IEEE Press, Washington (2004)
7. Hu, Y., Xi, J., Yang, Z., Li, E., Chicharo, J.: Study on Generalized Analysis Model for Fringe Pattern Profilometry. *IEEE Transactions on Instrumentation and Measurement* 57(1), 160–167 (2008)
8. Hügli, H., Maitre, G.: Generation and Use of Color Pseudo Random Sequences for Coding Structured Light in Active Ranging. *Industrial Inspection* 1010, 75–82 (1989)
9. Vuylsteke, P., Oosterlinck, A.: Range Image Acquisition with a Single Binary-Encoded Light Pattern. *Pattern Analysis and Machine Intelligence* 12(2), 148–163 (1990)
10. Molinier, T., Fofi, D., Gorria, P., Salvi, J.: 2D Virtual Texture on 3D Real Object with Color Structured Light. In: *SPIE Electronic Imaging – Image Processing: Machine Vision Applications*, vol. 6813, pp. 1–8. SPIE Press, Pittsburgh (2008)
11. Lavoie, P., Ionescu, D., Petriu, E.: 3-D Object Model Recovery from 2-D Images Using Structured Light. In: *IEEE Instrumentation and Measurement Technology Conference*, pp. 377–382. IEEE Press, Brussels (1996)
12. Lavoie, P., Ionescu, D., Petriu, E.: A High Precision 3D Object Reconstruction Method Using a Color Coded Grid and Nurbs. In: *IEEE International Conference on Image Analysis and Processing*, pp. 379–385. IEEE Press, Venice (1999)
13. Carrihill, B., Hummel, R.: Experiments with the Intensity Ratio Depth Sensor. *Computer Vision, Graphics and Image Processing* 32(3), 337–358 (1985)
14. Tajima, J., Iwakawa, M.: 3-D Data Acquisition by Rainbow Range Finder. In: *IEEE International Conference on Pattern Recognition*, vol. 1, pp. 309–313. IEEE Press, Atlantic City (1990)
15. Zhang, Z.: A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(11), 1330–1334 (2000)
16. Morano, R.A., Ozturk, C., Conn, R., Dubin, S., Zietz, S., Nissanov, J.: Structured Light Using Pseudorandom Codes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(3), 322–327 (1998)
17. Desjardins, D.: Structured Lighting Stereoscopy with Marching Pseudo-Random Patterns. MSc thesis, University of Ottawa, Ottawa, Canada (2008)
18. Hartley, R., Zisserman, A.: *Multiple View Geometry*. Cambridge University Press, Cambridge (2000)