MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

10—1 Adaptive Color Structured Light

Dalit Caspi Nahum Kiryati* Joseph Shamir
Department of Electrical Engineering
Technion, Haifa 32000, Israel

Abstract

A novel structured light method is described. The projection patterns are automatically adapted to the characteristics of the scene as part of the acquisition process, thus maximizing performance and minimizing the number of projection patterns. Color is used for light plane labeling. This allows to increase the dimension of the code space, hence the noise margins, without increasing the number of projection patterns and the acquisition period. Unlike previous color structured light techniques, few assumptions on smoothness and color neutrality of the scene are necessary and there are no compromises with respect to the spatial resolution and the immunity to noise. The theoretical results are supported by experiments. A significant reduction in the number of patterns with respect to the Gray code scheme is obtained.

1 Introduction

A triangulation-based structured light system [1] is similar to a passive stereo vision system with one of the cameras replaced by a projector. Suppose that a light source projects a plane of light that creates a narrow stripe on the scene. Since the intersection of a known illumination plane and a line of sight uniquely determines a point, the 3-D location of all points along the stripe that are visible by the camera can be simultaneously obtained. For dense reconstruction the scene must be scanned by the plane of light. In order to speed up the range sensing process, many light planes can be projected simultaneously. Then, however, the light planes have to be identified in the image.

High reliability identification of light planes with minimal assumptions on the nature of the scene can be achieved by sequentially projecting several patterns. Gray code can be used to label the light planes [7], where each illumination pattern represents one of the bit planes of the Gray code labels. The scene must remain static during the acquisition period, which is essentially proportional to the number of projection patterns used. Optimal design of projection patterns has been considered in [5].

Suppose that a unique color is given to each light plane. Had it been possible to reliably identify the projected colors in the image of the illuminated scene, range imaging using a single projection pattern would have been accomplished. However, unless the color of a given light plane is monochromatic, the spectral composition of the reflected light will depend on the color of the scene as well as on the projected color. Identification of the projected color will then be possible only if the scene is white or if its spectral response is spatially uniform and well known in advance. These are severely limiting assumptions.

In the rainbow range-finder [8] a bundle of monochromatic light planes is generated by dispersing a white illumination source. The wavelength of light reflected from the scene is then identical to that of the illuminating light plane. In principle, by using a color camera, the reflected colors can be identified and triangulation can be carried out. In practice, since the change of wavelength in the light plane bundle is gradual, the estimate of light plane position is sensitive to noise.

Boyer and Kak [2] assumed that the color of the scene is neutral and suggested to use spatial color sequences, i.e. blocks of colored light planes, to encode spatial position. This reduces the spatial resolution and requires assumptions on the smoothness of the scene. Improvements in that approach have been proposed by Hügli and Maître [6].

In this paper we suggest a new approach to the use of color in structured light. We show that the color of an impinging light plane *can* be identified from the image of the illuminated scene. This can be accomplished even if the scene is colorful, so the assumption on color neutrality of the scene can be relaxed. The identification is local and does not rely on spatial sequences, hence powerful smoothness assumptions are not needed. No compromises are made with respect to the spatial resolution and the immunity to noise. This is achieved by utiliz-

* Please address all correspondence to the following address: Dr. Nahum Kiryati, Department of Electrical Engineering - Systems, Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: nk@eng.tau.ac.il

ing the color characteristics of the projector and the camera, by projecting more than one pattern and by adapting the patterns to the lightness and color of the scene. A significant reduction in the number of required projection patterns with respect to the Gray code method is obtained. Further theoretical and experimental details, that are beyond the scope of this paper, can be found in [3].

2 Principle of Operation

The suggested structured light system relies on a computer controlled color video projector and a color camera. Suppose that a point in the scene is illuminated by a light plane. Let $\vec{c}$ denote the RGB instruction triplet given to the graphics adapter for setting the color of that plane. Let the vector $\vec{C}$ denote the triplet of RGB camera readings at the corresponding pixel in the image. We model the relation between $\vec{C}$ and $\vec{c}$ as follows:

$$\vec{C} = \mathbf{A} \cdot \mathbf{K} \cdot \vec{P}\{\vec{c}\} + \vec{C}_o \ , \tag{1}$$

where $\mathbf{A}$ is a matrix that depends on the response of the color filters in the camera and the projector, $\mathbf{K}$ is a diagonal matrix whose components k_R, k_G and k_B are proportional to the reflectivity of the scene point within the three color bands, $\vec{P}$ is a triplet of nonlinear monotonic transformations from graphics adapter instructions to projected intensities and $\vec{C}_o$ represents the triplet of RGB camera reading under the ambient light only. The matrix $\mathbf{A}$, the transformation $\vec{P}$ and its inverse $\vec{P}^{-1}$ can be obtained in advance by colorimetric calibration.

In our system, color coding of light planes (with time multiplexing of projection patterns) consists of the following layers. **Code alphabet:** Each code letter corresponds to one of the primary colors. Let n_R, n_G, n_B denote the number of brightness levels used in the red, green and blue channels respectively. Then each code letter is an integer in the range $0...n_R-1$ or $0...n_G-1$ or $0...n_B-1$ according to its color. Three different alphabets are thus used, one for each primary color. **Symbol:** A triplet of letters, one from each primary color. A symbol corresponds to the color of a specific light plane in a certain illumination pattern. **Code word:** Each light plane is generally represented by a code word that is a combination of symbols. When using M color illumination patterns, the code word is M symbols long ($3 \cdot M$ letters). **Projection patterns:** A single projection pattern carries a symbol for each light plane. The set of projection patterns carries the complete code word for each light plane.

In the beginning of a range imaging session, the user sets two of the following three parameters:

- The number of light planes L
- The number of projection patterns M
- A noise immunity parameter α

From this point on, the following procedure is automatically carried out:

1. A reference image is taken under the ambient light only. The camera readings $\vec{C}_o$ in all pixels are recorded.
2. A second reference image is taken under uniform white illumination, i.e., with $\vec{c} = \vec{c}_w = [255, 255, 255]^T$ for all light planes. The camera readings $\vec{C}_w$ in all pixels are recorded.
3. Black and shadow regions are identified by thresholding and dilation. They are marked and excluded from further consideration.
4. For each pixel, the matrix $\mathbf{K}$ is computed by solving Eq. 1 with $\vec{c} = \vec{c}_w$ and $\vec{C} = \vec{C}_w$.
5. The lowest combinations of $R_w - R_0$, $G_w - G_0$ and $B_w - B_0$ outside the black and shadow regions are found and the respective values of k_R, k_G and k_B are stored. Among the parameters L, M and α, the one that was *not* set by the user is now automatically adapted. The camera color space is sampled on a rectangular lattice with intervals that depend on k_R, k_G, k_B and α. The sampling points are transformed to projection instruction triplets. Code words are built according to the generalized Gray code [4].
6. M patterns are projected and M images are acquired.
7. For each pixel in each of the M images, $\vec{c}$ is obtained by solving Eq. 1. (At this stage the diagonal matrix $\mathbf{K}$ is known for every pixel).
8. For each pixel, the code word nearest to the M $\vec{c}$ vectors is determined. It corresponds to a light plane.
9. The 3-D location of the scene point that corresponds to each pixel is obtained by triangulation between the light plane and the line of sight.

Currently, the whole procedure takes about 30 seconds that are spent mostly on hard disk access. Once this is eliminated, computing time will be reduced to a few seconds. With hardware acceleration, acquisition time approaching the limiting value of M/R can be achieved, where R denotes the frame rate.

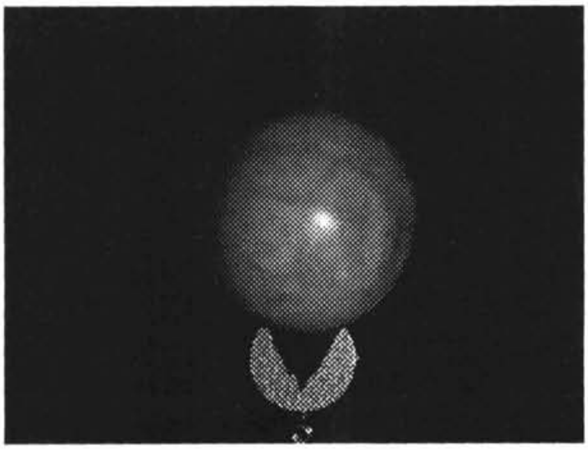

Figure 1: An image of a ball, taken under uniform white illumination.

3 Experiment

The projector in our experimental system is an Epson EMP-3000 video & data color VGA projector designed for presentations, capable of projecting up to 640 distinct light planes. In the projector, a source of white light is split by dichroic mirrors into three beams, red green and blue. Each beam is modulated by a liquid crystal spatial light modulator. The three beams are then combined by means of a beam combiner and projected through a common lens. The camera is a Sony XC-003P 3-CCD 768 × 576 analog color camera.

The geometric calibration of the system follows [9]. The typical distances of the camera and the projector from the scene are 1.5m and 1.0m respectively. The optical axes of the camera and the projector roughly intersect in the workspace, creating an angle of about 20^o. Due to pixelization in the projector and in the camera, in our setup the standard deviation of the depth error cannot be smaller than about 0.5mm. Better accuracy could be achieved by using higher resolution projector and camera.

Fig. 1 is an image of a green, textured plastic ball, taken under uniform white illumination. The diameter of the ball is 18cm. It is positioned at a distance of about 1.5m from the camera.

Range imaging was carried out in the following conditions:

1. Adaptive color coding with $L = 640$ light planes and a noise immunity factor of $\alpha = 6$. The number of projection patterns was automatically set to $M = 5$, with two levels in green and blue and none in red.

2. Black & white Gray coding with $M = 10$ illumination patterns.

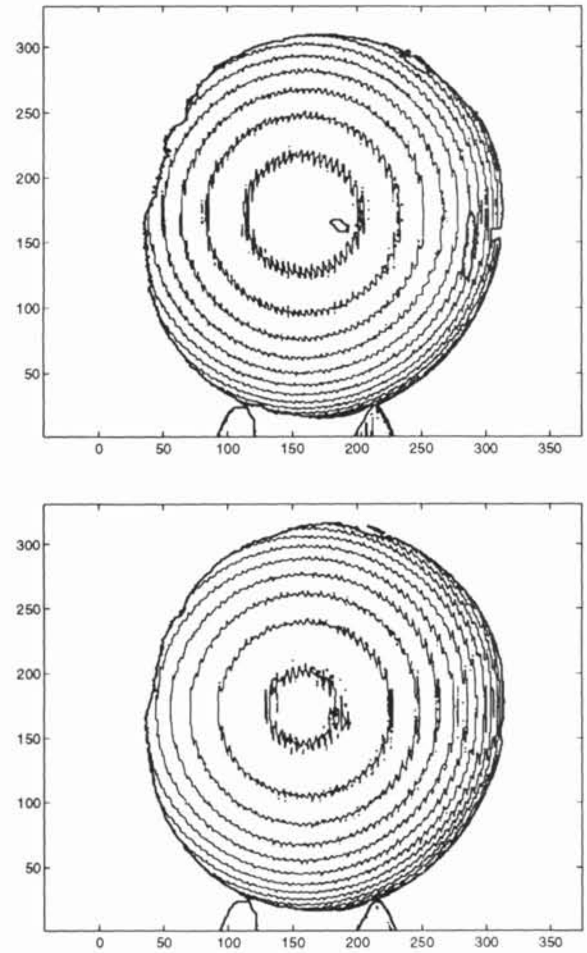

Figure 2: *Top:* Equal height contours representing the reconstruction of the ball using 5 color projection patterns. *Bottom:* Equal height contours representing the reconstruction of the ball using 10 black & white Gray code projection patterns.

Fig. 2 shows equal height contours representing the range data obtained with color coding (top) and with black & white Gray coding (bottom). In the reconstruction based on the five color patterns, there are two small patches on the surface of the ball in which the system provides an error indication and no depth data is available. One is a very dark spot on the right side of the scene that is considered to be black hence excluded from processing. The other is the point of specular reflection in which the camera is saturated. Both problems can be alleviated by employing a camera with a higher dynamic range. In the ten pattern black & white Gray code reconstruction, valid depth data is obtained in the dark spot, but there are some difficulties in the specular patch. Spherical surfaces were fitted to the two reconstructions of the ball. The standard deviation in both cases was about 0.9mm.

4 Conclusions

We presented a structured light method that reaches the accuracy and robustness of the Gray code technique [7] while using fewer projection patterns. The time interval during which the scene must remain static can therefore be reduced, thus widening the range of potential applications. This is accomplished by two new aspects of the suggested approach.

One is a new approach for the use of color in structured light. It has been previously believed that color neutrality of the scene is a prerequisite for successful color labeling of light planes. We have shown that even with colorful scenes, by using reference images, the color of an impinging light plane can be identified from the image of the illuminated scene. Identification is carried out in each pixel separately and does not rely on information from neighbors, hence powerful smoothness assumptions are not needed and the spatial resolution is not compromised.

The other new aspect is the online adaptation of the number and form of the projection patterns to the scene, as part of the acquisition process. Since the signal to noise ratio in each color channel eventually depends on the reflection from the scene in that band, adaptation eliminates the need for wasteful worst case design of the set of projection patterns. Instead, noise margins are matched to the actual noise levels, thus keeping the number of projection patterns at the necessary minimum. With darker scenes, the adapted projected patterns converge to the standard Gray code method.

The fundamental assumption in structured light methods that are based on intensity ratios is linear dependence of the intensity of light reflected from a point in the scene on the intensity of the light plane projected onto that point. Violations of this assumption, e.g., due to mutual illumination or fluorescence, can lead to errors. In our method, as long as these effects are small, they can be absorbed in the noise margins. Beyond a certain level, performance will deteriorate.

Acknowledgments

We are grateful to the Image Science Laboratory of the ETH in Zurich, Switzerland, and in particular to Dr. Marjan Trobina, for the system calibration software. We thank Dr. Boris Spektor, Technion, for his help and advice.

References

[1] P. Besl, Active optical range imaging sensors, *Machine Vision and Applications*, Vol. 1, pp. 127-152, 1988.

[2] K.L. Boyer and A.C. Kak, Color-coded structured light for rapid image ranging, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 9, pp. 14-28, 1987.

[3] D. Caspi, Adaptive color structured light, M.Sc. Thesis, Dept. of Electrical Eng., Technion, Haifa, Israel, 1997. See also D. Caspi, N. Kiryati and J. Shamir, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 20, pp. 470-480, 1998.

[4] M.C. Er, On generating the N-ary reflected gray codes, *IEEE Trans. Computers*, Vol. 33, pp. 739-741, 1984.

[5] E. Horn and N. Kiryati, Towards optimal structured light patterns, in *Proc. Int. Conf. on Recent Advances in 3-D Digital Imaging and Modeling*, Ottawa, Canada, pp. 28-37, 1997. To appear in *Image and Vision Computing.*

[6] H. Hügli G. Maître, Generation and use of color pseudo random sequences for coding structured light in active ranging, in *Proc. SPIE Industrial Inspection*, Vol. 1010, pp. 75-82, 1988.

[7] K. Sato and S. Inokuchi, Three-dimensional surface measurement by space encoding range imaging, *J. of Robotic Systems*, Vol. 2, pp. 27-39, 1985.

[8] J. Tajima and M. Iwakawa, 3-D data acquisition by rainbow range finder, in *Proc. 10th ICPR*, Atlantic City, pp. 309-313, 1990.

[9] M. Trobina, Error model of a coded-light range sensor, Technical Report BIWI-TR-164, ETH, Zurich, Switzerland, 1995.