Rapid 3D Face Data Acquisition Using a Color-Coded Pattern and a Stereo Camera System

Byoungwoo Kim, Sunjin Yu, Sangyoun Lee, and Jaihie Kim

Biometrics Engineering Research Center,
Dept. of Electrical and Electronics Engineering, Yonsei University,
134 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Korea
{bwkim, biomerics, syleee, jhkim}@yonsei.ac.kr

Abstract. This paper presents a rapid 3D face data acquisition method that uses a color-coded pattern and a stereo camera system. The technique works by projecting a color coded pattern on an object and capturing two images with two cameras. The proposed color encoding strategy not only increased the speed of feature matching but also increased the accuracy of the process. We then solved the correspondence problem between the two images by using epipolar constraint, disparity compensation based searching range reduction, and hue correlation. The proposed method was applied to 3D data acquisition and time efficiency was compared with previous methods. The time efficiency of the suggested method was improved by about 40% and reasonable accuracy was achieved.

1 Introduction

Although current 2D face recognition systems have reached a certain level of maturity, their performance has been limited by external conditions such as head pose and lighting. To alleviate these conditions, 3D face recognition methods have recently received significant attention, and the appropriate 3D sensing techniques have also been highlighted [1][2].

Previous approaches in the field of 3D shape reconstruction in computer vision can be broadly classified into two categories; active and passive sensing. Although the stereo camera, a kind of passive sensing technique, infers 3D information from multiple images, the human face has an unlimited number of features. Because of this, it is difficult to use dense reconstruction with human faces. Therefore, passive sensing is not an adequate choice for 3D face data acquisition.

On the other hand, active sensing projects a special pattern onto the subject and reconstructs shapes from reflected pattern imaging with a CCD camera. Because active sensing is better at matching ambiguity and also provides dense feature points, it can act as an appropriate 3D face-sensing device.

The most simple approach in structured lighting is to use a single-line stripe pattern, which greatly simplifies the matching process, although only a single line of 3D data points can be obtained with each image shot. To speed up the acquisition of 3D range data, it is necessary to adopt a multiple-line stripe pattern instead. However, the matching process then becomes much more difficult. One possibility is to use color information to simplify this difficulty [2][3].

D. Zhang and A.K. Jain (Eds.): ICB 2006, LNCS 3832, pp. 26 – 32, 2005.

Furthermore, in the single-camera approach, it is necessary to find the correspondence between the color stripes projected by the light source and the color stripes observed in the image. In general, due to the different reflection properties (or surface albedos) of object surfaces, the color of the stripes recorded by the camera is usually different from that of the stripes projected by the light source (even when the objects are perfectly Lambertian.) It is difficult to solve these problems in many practical applications [4].

On the other hand, this does not affect our color-lighting/stereo system if the object is Lambertian, because the color observed by the two cameras will be the same, even though this observed color may not be exactly the same as the color projected by the light source. Therefore, by adding one more camera, the more difficult problem of lighting-to-image correspondence is replaced by an easier problem of image-to-image stereo correspondence. Here, the stereo correspondence problem is also easier to solve than traditional stereo correspondence problems because an effective color pattern has been projected onto the object [4].

In this paper, we show how we have developed and implemented a new method for 3D range data acquisition that combines color structured lighting and stereo vision. In the proposed system, we developed a new coded color pattern and a corresponding point matching algorithm. Once the correspondence problem was solved, the 3D range data was computed by the triangulation technique. Triangulation is a well-established technique for acquiring range data with corresponding point information [5].

This paper is organized as follows; in section 2, we address system calibration, and section 3 discusses generating a new color-coded pattern. Stereo matching methods are dealt with in section 4. In section 5, experimental results are presented. Finally, section 6 concludes the paper.

2 Camera Calibration

Calibration is the process of estimating the parameters that determine a projective transformation from the 3D space of the world onto the 2D space of image planes. A set of 3D-2D point pairs for calibration was obtained with a calibration rig. If we know 6 point pairs, calibration matrix is uniquely determined. However, in many cases, since there exists errors, more than 6 point pairs are recommended, and it results in over-determined problem. Then the stereo camera system was calibrated with the DLT (Direct Linear Transform) algorithm [5][6].

3 Color-Coded Pattern Generation

The color-coded pattern generates an effective color sequence that can solve the corresponding problem and provide strong line edge segments. For pattern design, line segments have been effectively used in many 3D data acquisition systems, so we have exploited these line segments in our pattern design [7].

Previous research has shown that the HSI model is an effective color model for stereo matching [3][8]. Using the line features and the HSI color model, a set of unique color encoded vertical stripes was generated.

Each color-coded stripe was obtained as follows. Stripe color was denoted as $stripe(\rho,\theta) = \rho e^{j\theta}$, where ρ is the saturation value and θ is the hue value in the HS polar coordinate system shown in Fig. 1. To obtain a distinctive color sequence, we defined four sets of color. Each set contained three colors whose hue was separated by 120° within the set. We used only one saturation value (saturation=1) because hue information was enough to distinguish each stripe for matching process. Finally, the stripe color equation was denoted as (1).

$$color(m,n) = e^{j(mH_{jmp}+\varepsilon n)} \tag{1}$$

(a) Hue-Saturation polar coordinates (b) Generated color-coded pattern

Fig. 1. Generation of color-coded pattern

Next, the color-coded sequence was obtained as follows. First, we chose one out of the four, and 3 elements from this set were used. The next set elements were then used sequentially. After a 12-color sequence was generated, the next 12 color stripes were generated in the same manner. Fig. 1(b) shows the generated color-coded pattern.

4 Stereo Matching

In this section, rectification, and the corresponding points matching method are introduced. The color stripes to be projected onto the face were captured by both the left and right cameras. The captured images were then processed and were represented by thinned color lines. Then, the preprocessed image pairs were rectified using calibration information. Finally we found the corresponding point pairs quickly using the proposed method.

4.1 Epipolar Rectification

After thinning, the obtained image pairs were rectified using the camera calibration information. This step transforms the images so that the epipolar lines are aligned horizontally. In this case, the stereo matching was able to take advantage of the epipolar constraint and the search space was reduced to one dimension. Rectification is important when finding the corresponding points of the left image (i_l, j_l). We only needed to look along the scanline $j_r = j_l$ in the right image [5][9].

4.2 Disparity Compensation

To minimize computational complexity, we needed to restrict the searching ranges. After rectification, the difference between the pair of stereo images was small and was caused by horizontal shifts, it was necessary to compensate for the disparity of the stereo images. We used the SAD (Sum of Absolute Difference) to get the disparity value. Because it would take too much time to compensate for every image row, we only did so at multiples of 100 rows. We compensated at the K_{th} row using the following equation:

$$SAD_K = \sum_{i}^{N_x} \sum_{j}^{N_y} \left| Hue_L(i,j) - Hue_R(i+k,j) \right| \tag{2}$$

where N_x and N_y are 3 by 3 block size, and $Hue_L(i,j)$ is the hue value of the i, j pixel positions in the left image. At the equal row line, we found the minimum SAD:

$$SAD^p{}_{MIN} = MIN(\sum SAD_K) \tag{3}$$

Finally we found the background disparity of the whole image by maximizing equation (4):

$$SAD_{MAX} = MAX(\sum SAD^p{}_{MIN}) \tag{4}$$

By this process, we found K, which is the background disparity of the stereo images:

$$Right^{compensated} = Right_{t-K} \tag{5}$$

$$Left^{compensated} = Left_{t+K} \tag{6}$$

4.3 Stereo Matching

At the stereo matching step, we obtained the corresponding pairs of the two captured images. We found the hue distribution of two images very similar. However, the hue distribution of the captured left image and that of the captured right image are more similar than the hue distribution of the pattern image. Matching between the two captured images is more robust and accurate than between one of the captured images and the pattern image. This result confirms one of the major benefits of our new proposed system.

Up to the thinning step, we obtained two images that contained thinned color lines. With the epipolar constraint, the corresponding point pair fell on the epipolar line. With this constraint, the searching range was reduced to a line. Furthermore, we needed to limit the searching range of the epipolar line. Because the same color stripes were used twice in the designed color sequence, one point of the left image was matched twice on the epipolar line. To solve this problem, we used the disparity compensation method to restrict the searching range. So we never considered matching pixels with a disparity of more than (K+40), or less than –(K+40). We only compared the hue values of about 4 points on the epipolar line. In this case, there was no chance of getting two corresponding pairs. Three constraints including the epipolar constraint, disparity compensation-based searching range reduction, and hue information allowed us to find the corresponding points very rapidly. This is another major benefit of the proposed method. Fig. 2. shows the matching process.

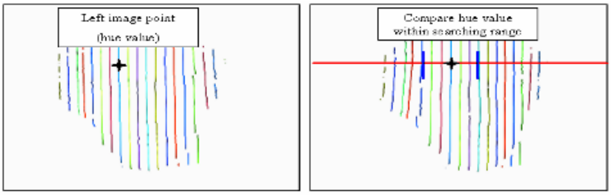

Fig. 2. Matching process

4.4 3D Reconstruction

Triangulation is the process of obtaining a real 3D position from two intersecting lines [5]. These lines are defined by the corresponding pairs and information from each calibration. After camera calibration, the triangulation method was used to obtain the 3D depth information. The triangulation method was solved with the SVD(Singular Value Decomposition) algorithm and 3D points were reconstructed [6].

5 Experiments

The system underwent an initializing step prior to inferring the 3D coordinates. After initializing the step, the color-coded pattern illuminated the subject. Corresponding point matching was then followed.

5.1 Accuracy Test

To test accuracy, we used a skin-colored box. We estimated the width, height and degree of the box. The metric RMS error between the real value and the reconstructed value was used as the accuracy measure. Table 1 shows the obtained results. From table 1, we can see that our system produced the maximum 2.39% RMS(Root Mean Squared) error when compared to the real values.

Table 1. The Accuracy test results

	Width	Length	Height	Degree A	Degree B	Degree C
Real value	14.5	12.5	9.5	90	88	92
Reconstruction result	13.89	11.32	9.28	88.32	86.12	89.48
RMS error	0.6211	1.2135	0.2641	1.86	2.35	2.39

unit: cm, degrees

Table 2. Time efficiency test results

	Dataset1			Dataset2		
Process	Previous Method1	Previous Method2	Proposed Method	Previous Method1	Previous Method2	Proposed Method
Preprocessing	3904	2942	1206	3889	3124	1284
Matching	736	720	946	814	749	856
Triangulation	242	237	244	287	264	255
Total Time	4888	3899	2396	4990	4137	2395
Total Points	5620	5644	5723	6920	6425	6324
Time / Point	0.8690	0.6908	0.4187	0.7210	0.6439	0.3787

unit: ms

Table 3. Computation time of the proposed matching method versus the DP matching method

	Proposed method	DP matching
Corresponding pairs	6947	7012
Time	945	1344

unit: ms

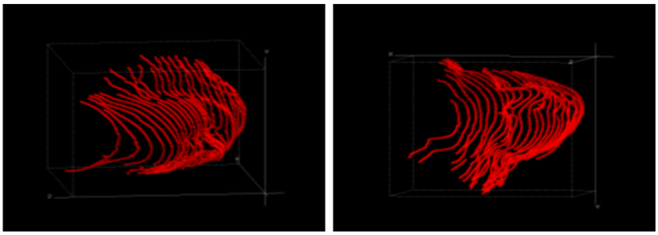

Fig. 3. 3D reconstruction results: Facial range data from two different viewing points

5.2 Time Efficiency

To test time efficiency, we estimated one 3D point reconstruction time. This is because, even for the same object, the number of reconstructed data points were different for each acquisition system. This made it impossible to estimate time efficiency by reconstruction time per total data points. We compared our system with a previous method [10][11]. The results are shown in table 2. We found that the time efficiency of our system improved by about 40% compared to a previous method. Table 3 also shows the comparison results between the proposed matching algorithm and the DP matching algorithm [7][12]. Performance of the proposed matching algorithm improved by about 30% compared to the DP matching algorithm. Fig. 3 shows the results of 3D face data reconstruction.

6 Conclusions

One significant advantage of our approach is that there is no need to find the correspondence between the color stripes projected by the light source and the color stripes observed in the image. In general, it is quite difficult to solve the above matching problem because the surface albedos are usually unknown. By not having to deal with this, we were able to focus on the easier image-to-image stereo correspondence problem. This process was also easier than traditional stereo correspondence because a good color pattern was projected onto the object.

Experimental results show a value of about 2% of depth error for the polyhedral object, but its performance decreased a little around the curved object. Also, the time efficiency of the proposed system is better than previous color structured lighting methods and the DP matching method. A drawback of this system is that color-coded stripes are usually sensitive to ambient light effects. Also, for dense reconstruction, the number of lines needs to be increased. Therefore, future works will include developing a more robust color pattern for ambient illumination and dense reconstruction.

Acknowledgement. This work was supported by the Korea Science and Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center (BERC) at Yonsei University.

References

1. H.S. Yang, K.L. Boyer and A.C. Kak.: Range data extraction and interpretation by structured light. Proc. 1st IEEE Conference on Artificial Intelligence Applications, Denver, CO, (1984) 199-205.
2. K.L. Boyer and AC. Kak.: Color-encoded structured light for rapid active ranging. IEEE Trans. Pattern Analysis and Machine Intelligence, (1987) 14-28.
3. C.H. Hsieh, C.J. Tsai, Y.P. Hung and SC. Hsu.: Use of chromatic information in region-based stereo. Proc. IPPR Conference on Computer Vision, Graphics, and Image Processing, Nantou, Taiwan, (1993) 236-243
4. C. Chen, Y. Hung, C. Chiang, and J. Wu.: Range data acquisition using color structured lighting and stereo vision. Image and Vision Computing, Mar. (1997) 445-456
5. Emanuele Trucco and Alessandro Verri.: Introductory Techniques for 3-D Computer Vision, Prentice Hall (1998)
6. R. Hartley and A. Zisserman.: Multiple view Geometry in computer vision, Cambridge University Press (2000)
7. Y. Ohta, and T. Kanade.: Stereo by intra and inter scan line search using dynamic programming. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 7, No. 2, Mar. (1985) 139-154
8. R.C. Gonzales and R.E. Woods.: Digital Image Processing, Addison-Wesley, Reading, MA, (1992).
9. H. Jahn.: Parallel Epipolar Stereo Matching. IEEE int. Conf. on Pattern Recognition, ICPR2000, (2000) 402-405
10. Dongjoe Shin.: The hard calibration of structured light for the Euclidian reconstruction of face data. Master's Thesis. Dept. of Electrical and Electronic Engineering. Yonsei University. (2004)
11. Sungwoo Yang, Sangyoun Lee and Jaihie Kim.: Rapid Shape Acquisition for Recognition Using Absolutely Coded Pattern. Comm. Int. Symp. Intell. Signal Process., Comm. Systems (ISPACS). Seoul, Korea. Nov. (2004) 620-624
12. L. Zhang, and B. Curless, and S. M. Seitz.: Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programming," Proc. of First International Symposium on 3D Data Processing Visualization and Transmission, Jun. (2002) 24-36