Proc. 13th Int. Conf. on Pattern Recognition ICPR´96, Vienna, Austria, August 25-29, 1996, Vol. I, pp. 835-839.

Color Stereo Vision Using Hierarchical Block Matching and Active Color Illumination

Andreas Koschan [1], Volker Rodehorst [2], Kathrin Spiller [2]

[1] Dept. of Computer Science, FR 3-11
Technical University of Berlin
Franklinstr. 28-29
10587 Berlin, Germany
koschan@cs.tu-berlin.de

[2] Photogrammetry and Cartography, EB 9
Technical University of Berlin
Strasse des 17. Juni 135
10623 Berlin, Germany
vr, kathrin@fpk.tu-berlin.de

Abstract

Stereo is a well-known technique for obtaining depth information from digital images. Nevertheless, this technique still suffers from a lack in accuracy and/or long computation time needed to match stereo images. A new hierarchical algorithm using an image pyramid for obtaining dense depth maps from color stereo images is presented. We show that matching results of high quality are obtained when using the new hierarchical chromatic Block Matching algorithm. Most stereo matching algorithms can not compute correct dense depth maps in homogenous image regions. This paper shows that using an active color illumination will considerably improve the quality of the matching results. We present results for synthetic and for real images.

1. Introduction

During the past few years the development of stereo algorithms has been a subject of considerable research activity. The key problem in stereo is how to find the corresponding points in the left and in the right image, referred to as the correspondence problem. Whenever the corresponding points are determined, the depth can be computed by triangulation. Although, more than 300 papers have been published dealing with stereo vision this technique still suffers from a lack in accuracy and/or long computation time needed to match stereo images. There is still a need for more precise and faster algorithms.

Stereo techniques can be distinguished by either matching edges and producing sparse depth maps or matching all pixels in the images and producing dense depth maps. The objective of the application always effects the decision whether the preference is given to dense stereo correspondence or to edge-based correspondence. For a successful reconstruction of complex surfaces it is essential to compute dense disparity maps defined for every pixel in the entire image.

Unfortunately, most of the existing dense stereo techniques are very time consuming (see, e.g., [1, 2]). In an earlier investigation [3], we found the Block Matching technique using color information to be very suitable for dense stereo. The precision of the matching results always improved by 20 to 25 % when using color information instead of gray value information. In this paper, we present a new hierarchical algorithm using an image pyramid for obtaining dense depth maps from color stereo images. We show that matching results of higher quality are obtained when using the hierarchical algorithm instead of the non-hierarchical algorithm we developed earlier. Additionally, the algorithm becomes faster than the standard one at the same time.

Most stereo matching algorithms can not compute correct dense depth maps in homogenous image regions. From other experiments we realize that the discriminability of objects can be enhanced by controlling the illumination color [4]. Therefore, finding the optimal illumination color for recognizing objects in structured environments [5] is a promising but rather difficult task. Fortunately, we do not have to find an optimal illumination color to improve stereo matching results. Kanade and his colleagues [6] projected a sinusoidal varying intensity onto the scene. They found an improvement in the results but still got some false matches due to the limited dynamic range of their camera, particularly with dark surfaces. From this they conclude that projecting a colored code may serve to reduce these effects. In this paper, we propose to project a color code onto the scene that represents a rainbow like spectrum. The quality of the matching results always improves, especially in homogenous regions, when this active color illumination is used.

2. Stereo analysis using Block Matching

The main idea of Block Matching is a similarity check between two equal sized blocks (*n* x *m*-matrices) in the left and the right image (area-based stereo). The mean square error *MSE* between the pixel values inside the respective blocks defines a measure for the similarity of two blocks. We propose to employ an approximation of the Euclidean distance to measure color differences.

The left color image F_L and the right color image F_R may be represented in the *RGB* color space as

$$F_L(i,j) = \left(R_L(i,j), G_L(i,j), B_L(i,j)\right) \text{ and}$$
$$F_R(i,j) = \left(R_R(i,j), G_R(i,j), B_R(i,j)\right).$$

The *MSE* is defined with $n = m = 2k+1$ as

$$MSE_{color}(x,y,\Delta) = \frac{1}{n \cdot m} \sum_{i=-k}^{k} \sum_{j=-k}^{k} \Big(\left| R_R(x+i, y+j) - R_L(x+i+\Delta, y+j) \right|^2 + \left| G_R(x+i, y+j) - G_L(x+i+\Delta, y+j) \right|^2 + \left| B_R(x+i, y+j) - B_L(x+i+\Delta, y+j) \right|^2 \Big),$$

where Δ is an offset describing the difference $(x_L - x_R)$ between the column positions in the left and in the right image. The block (of size *n* x *m*) is shifted pixel by pixel inside the search area. Using standard stereo geometry the epipolar lines match the image line. Therefore, the matching process can be simplified as corresponding pixels can only be found in the same row in both images. The disparity *D* of two blocks in both images is defined by the horizontal distance, that shows the minimum mean square error. Furthermore, the search area in the right image is limited in the horizontal direction by a predefined maximum disparity d_{max}.

$$D = \min_{|\Delta| \le d_{max}} \left\{ MSE_{color}(x,y,\Delta) \right\}.$$

Block disparities are median filtered to avoid outliers. A dense disparity map is generated when applying a pixel selection technique to every pixel in the image (see [3] for further details).

We applied the proposed algorithm to a stereo image pair named "ANDREAS" (see Fig. 1). It can be easily seen that the left eye in the left image is falsely matched to the right eye in the right image. This artifact occurs due to the repetitive pattern in the scene and the property of the chromatic Block Matching method of matching all single blocks independently to each other.

Reducing the search space for the disparities could be a solution to the problem. This could be obtained by a very restrictive use of the continuity constraint proposed in [7]. It produces a smoothed depth map where fine structures can not be represented. Discontinuities in depth that are typical for object edges get smoothed. Thus, any segmentation may fail. This disadvantage can be solved using a pyramid model.

Fig. 1: Gray value reproduction of the left color stereo image "ANDREAS" and the depth map (right) obtained when applying standard Block Matching to the images.

3. Hierarchical Block Matching using image pyramids

The idea of using pyramid models in image analysis was introduced by Tanimoto and Pavlidis [8] as a solution to edge detection. Since the 1970´s, considerable research activities have been taken place in this area and image pyramids are now used in several areas of image analysis (see, e.g., [9], [10]). One important property of the pyramid model is that it is computationally extremely efficient [11].

We enhanced the chromatic Block Matching algorithm by using a quad pyramid (see Fig. 2). The disparities $D(s+1)$ at level $(s+1)$ can be derived from the disparities $D(s)$ of the preceding level (s) by applying a modified block matching algorithm to the image of level $(s+1)$. The search space for the disparity of each block at level $(s+1)$ is derived from the disparity of the corresponding block at level (s) by a tolerance factor D_T. This parameter defines the width D_Δ of the reduced search space $[D_{MIN}, D_{MAX}]$ and controls the smoothness of the disparity map.

$$D_\Delta(s) = 2^{(s-1)} \cdot D_T\ ,$$

$$D_{MIN}(s) = \begin{cases} D(0) - D_\Delta(s) & \text{for} \quad s = 1 \\ D_{MIN}(s-1) - D_\Delta(s-1) & \text{for} \quad s > 1 \end{cases},$$

$$D_{MAX}(s) = \begin{cases} D(0) + D_\Delta(s) & \text{for} \quad s = 1 \\ D_{MAX}(s-1) + D_\Delta(s-1) & \text{for} \quad s > 1 \end{cases}.$$

When choosing a small value for the tolerance factor D_T, the difference between the final disparities and the average disparity found at level 0 will be very small. This is equivalent to a small variation of disparities over the whole image. A larger tolerance factor will cause a bigger

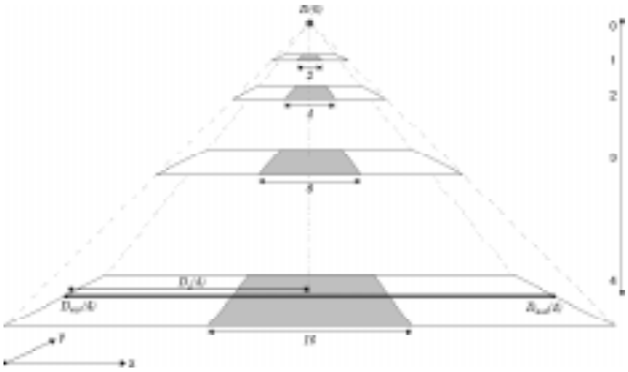

Fig 2: Definition of the search space with a tolerance factor D_T = 3.0.

search space and the influence of the computed disparities in the preceding levels will decline. Then the artifacts described above may occur.

Nevertheless, this hierarchical method is more robust than the non-hierarchical one. Comparing the results of the non-hierarchical Block Matching with the results of the hierarchical Block Matching the enhanced algorithm shows better results (see Fig. 3). Not only the artifact with the eyes can be solved, even the depth of small structures as the ears are estimated more correctly. Furthermore, the hierarchical approach can be implemented very efficiently in parallel to achieve high speed execution.

4. Active color illumination for enhancing stereo matching

Most stereo matching algorithms can not compute correct dense depth maps for homogenous image regions. This is due to the ambiguity of image values inside these regions. The ambiguity can be eliminated by adding a synthetic texture to the scene. The results improve when intensity coded light is projected onto the scene [6].

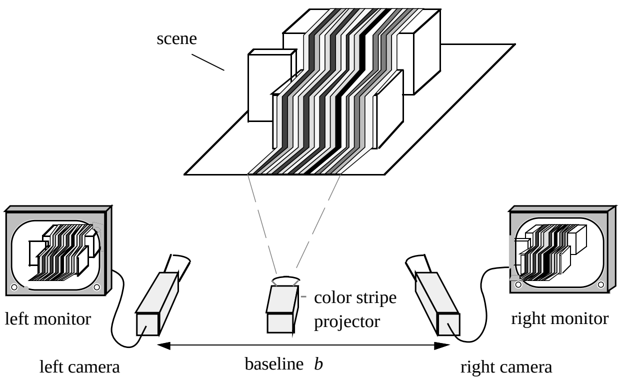

Fig. 4: Principle of the stereo system using active color illumination.

Fig. 3: Enhanced depth map using hierarchical Block Matching and the 3-D reconstruction with texture mapping.

False matches can considerably be reduced when projecting a color code onto the objects instead of an intensity code since color provides much more distinguishable information than intensity [12]. Since the projected colored light mixes with the unknown object colors, blended colors are reflected onto the image planes of the two cameras. Fortunately, we must not determine either object colors or projected colors, when using stereo vision. The colors obtained by superimposing object colors with the color coded illumination are identical in both images. Therefore, the pixels in the images can be matched without any additional knowledge about the object colors or the light colors, respectively. The principle of the stereo arrangement using active color illumination is outlined in Fig. 4.

We projected a rainbow like color spectrum onto the scene. Every row of length n in the color spectrum S_{RGB} was generated using the equations

$$S_R = sin\left(\frac{i}{n}\cdot\pi\right)\cdot\left(\frac{G_{MAX}}{2}-1\right)+\frac{G_{MAX}}{2}\ ,$$

$$S_G = sin\left(\left(\frac{2}{3}+\frac{i}{n}\right)\cdot\pi\right)\cdot\left(\frac{G_{MAX}}{2}-1\right)+\frac{G_{MAX}}{2},\ i=\{0,..,n\},$$

$$S_B = sin\left(\left(\frac{4}{3}+\frac{i}{n}\right)\cdot\pi\right)\cdot\left(\frac{G_{MAX}}{2}-1\right)+\frac{G_{MAX}}{2}\ ,$$

where i denotes the column position in the spectrum image and G_{MAX} denotes the maximum intensity value in every color channel.

The color spectrum was synthetically generated using the equations mentioned above. Afterwards, a slide representing this synthetic color stripe code was exposed by using a scanner for color reversal films. This slide is projected onto the objects using a standard slide projector. The two stereo images are acquired with color CCD-cameras. To take advantage of the structured illumination we chose the chromatic Block Matching algorithm because it considers the color coding best.

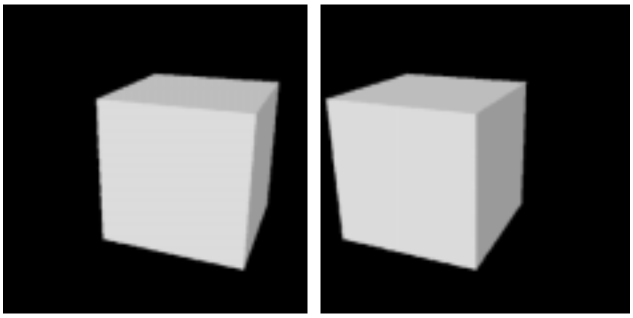

Fig 5: The synthetic gray value stereo images CUBE.

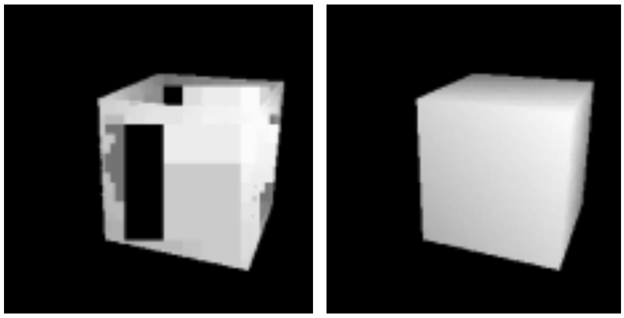

Fig. 6: Disparity map computed for CUBE (left) and results obtained with the color spectrum (right).

5. Experimental results with synthetic images

The quality of a stereo algorithm can hardly be evaluated if exclusively real images and the computed disparity maps are investigated. Often there does not exist enough information to predict the correct disparity values, particularly in dense disparity maps. Opposed to this, the correct depth map and an integer map representing the disparity values can be generated when using synthetic images and a stereo simulator.

Now we present some results using a synthetic image pair, we name "CUBE", of size 512 x 512 pixels (see Fig. 5). The matching of surface patches with uniform shading and without texture is ambiguous. Fig. 6 shows the computed disparity map obtained when applying the Block Matching algorithm to the original images. Several mismatches occurred due the ambiguity of the intensity values. Furthermore, we superimposed the synthetic images with the color spectrum introduced in the previous section. The resulting disparity map is shown in Fig. 6.

We compared both disparity maps to the ideal disparity map generated by the stereo simulator. The computed results improved significantly when using structured light (see Fig. 6 and Tab. 1).

Difference (in pixel)	Intensity images (%)	Color images (%)
0	7.9	62.2
1	12.2	25.8
≥ 2	79.8	12.0

Tab. 1: Distribution of the matching errors (in percentage) for the gray value images and for the images superimposed with the color spectrum.

The image "CUBE" was chosen to clarify the effect of colored light to the matching problem. Of course the results can not be used to predict similar improvements for real scenes. Several additional investigations were carried out for synthetic images. Unfortunately, we can not present further details due to limited space. In summary, we found that the matching results always improved when a color spectrum is projected into the scene. Furthermore, the results are rather robust concerning noise, contrast variations, and intensity variations.

6. Experimental results with real images

Since the projection of color codes onto the scene considerably improves the quality of the matching results for synthetic images we applied the same technique to real images. We applied the Block Matching algorithm to a color stereo image named "RELAXING JACK" (see Fig. 7, upper row). The computed depth map, that was obtained by triangulation from the disparity map, shows a defect beneath the head (magnified section). Another artifact occurs at the bottom part of the arm-chair, although this uniform shaded section is quite small. In another experiment the color spectrum was projected onto the same scene. The artifacts that occurred when using the

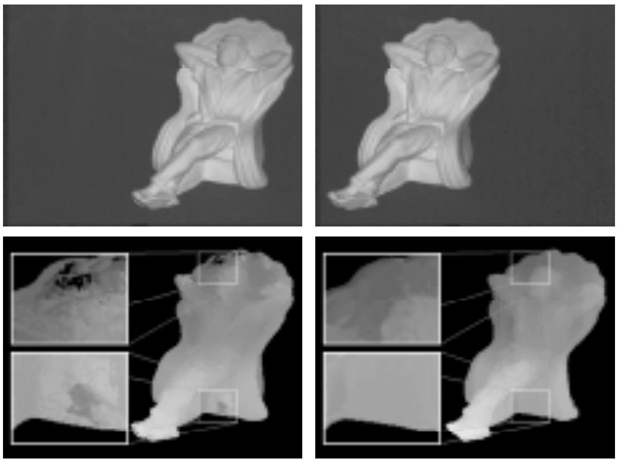

Fig. 7: Upper row: gray value representation of the real color stereo image "RELAXING JACK". Lower row: computed depth maps obtained with (right) and without (left) active illumination.

Fig. 8: Left: Difference (scaled) between both depth maps shown in Fig. 7; right: shaded representation of the reconstructed scene "RELAXING JACK".

original color images, do not show up (see magnified sections in Fig. 7, lower row, right). More distinctions are pointed out by visualizing the difference of the two depth maps (see Fig. 8, left). Additionally, the right image in Fig. 8 shows the reconstructed scene mapped with the texture of the right image.

The experiments mentioned above show a considerable improvement in the matching results for synthetic and for real images when active color illumination is utilized. Regarding other investigations with structured illumination the main advantage of this method is, that it deals with only one pair of images. Thus, the method can also be applied to moving or non-rigid objects. Additionally, we studied different parallel implementations for stereo matching [13]. Using 10 processing units on a SGI Power Challenge, the hierarchical Block Matching takes between 0.03 and 0.65 sec depending on the image size (256 x 256 or 768 x 566 pixel, respectively).

7. Conclusion

A combination of two new approaches for dense stereo matching has been presented. The first approach uses an image pyramid model and a hierarchical implementation of Block Matching for color stereo images. It has been shown that the quality of the matching results can be improved with this hierarchical approach.

The second approach we presented uses active color illumination for stereo matching. We showed the benefit of the approach for synthetic and for real images. The quality of the matching results always considerably improved when employing active colored illumination, particularly in homogenous regions. This holds for every dense stereo technique. We used a combination of the hierarchical approach with active color illumination to produce high quality results. Additional tests and investigations are necessary for a more detailed evaluation of the techniques. Currently, this is under construction and further results will be presented soon.

In summary, we should like to emphasize that active colored illumination always serves to improve stereo matching results. Therefore, we believe that more precise results can be efficiently obtained in dense stereo matching when combining hierarchical chromatic Block Matching with the active color illumination approach.

Acknowledgment

This work was funded by the Deutsche Forschungsgemeinschaft (DFG).

References

[1] J.R. Jordan III, A.C. Bovik, "Using Chromatic Information in Dense Stereo Correspondence," *Pattern Recognition,* vol. 25, no. 4, pp. 367-383, 1992.

[2] M. Okutomi, O. Yoshizaki, G. Tomita, "Color Stereo Matching and its Application to 3-D Measurement of Optic Nerve Head," *Proc. 11th IAPR Int. Conf. on Pattern Recognition*, the Hague, the Netherlands, vol. I, pp. 509-513, 1992.

[3] A. Koschan, "Chromatic block matching for dense stereo correspondence," *Proc. 7th Int. Conf. on Image Analysis and Processing 7ICIAP*, Capitolo, Monopoly, Italy, pp. 641-648, 1993.

[4] M. Vriesenga, G. Healey, K. Peleg, J. Sklansky, "Colored Illumination for Enhancing Discriminability in Machine Vision," *Journal of Visual Communication and Image Representation,* vol. 6, no. 3, pp. 244-255, 1995.

[5] H. Murase, S.K. Nayar, "Illumination Planning for Object Recognition in Structured Environments," *Proc. CVPR´94*, Seattle, Wa., pp. 31-38, 1994.

[6] S.B. Kang, J.A. Webb, C.L. Zitnick, T. Kanade, "A Multibaseline Stereo System with Active Illumination and Real-time Image Acquisition," *Proc. 5th Int. Conf. on Computer Vision ICCV´95*, Cambridge, Ma., pp. 88-93, 1995.

[7] D. Marr, "Vision - A Computational Investigation into the Human Representation and Processing of Visual Information," Freeman & Co, New York, 1982.

[8] S. Tanimoto, T. Pavlidis, "A Hierarchical Data Structure for Picture Processing," *Computer Graphics and Image Processing*, vol. 4, pp. 104-119, 1975.

[9] J.-M. Jolion, A. Rosenfeld, "A Pyramid Framework for Early Vision," Kluwer, Dordrecht, the Netherlands, 1994.

[10] I. Pitas, "Digital Image Processing Algorithms," Prentice Hall, Hempstead, UK, 1993.

[11] W.G. Kropatsch, "Properties of Pyramidal Representations," *Computing Suppl. 11*, pp. 99-111, 1996.

[12] R. Sasse, "Bestimmung von Entfernungsbildern durch aktive stereoskopische Verfahren," Vieweg, Braun–schweig / Wiesbaden, 1994 (in German language).

[13] A. Koschan, V. Rodehorst, "Towards Real-Time Stereo Employing Parallel Algorithms For Edge-Based And Dense Stereo Matching," *Proc. of the IEEE Workshop on Computer Architectures for Machine Perception CAMP´95*, Como, Italy, pp. 234-241, 1995.