The Fifteenth International Symposium on Artificial Life and Robotics 2010 (AROB 15th '10),
B-Con Plaza, Beppu,Oita, Japan, February 4-6, 2010

Dense stereovision using mono-CCD color cameras

Hachem Halawana[1] and Hani Hamdan[2]

[1]*LAGIS, USTL, Bât. P2, Cité Scientifique, 59655 Villeneuve d'Ascq, France*
[2]*Dept. of Signal Processing and Electronic Systems, SUPELEC, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette cedex, France*
Hachem.Halawana@ed.univ-lille1.fr, Hani.Hamdan@supelec.fr

Abstract: Most of the stereo algorithms were based only on the analysis of the luminance information. However with the advances in camera technology, in addition to the fact that color information can robustly improve matching, color stereovision gained more and more interest. Color stereovision setups are usually based on single-sensor cameras which provide Color Filter Array (CFA) images. In those, a single color component is sampled at each pixel rather than the three required ones (R, G, B). We show that standard demosaicing techniques, used to interpolate missing components, are not well adapted when the resulting color pixels are matched for estimating image disparities. In order to avoid this problem while exploiting color information, we propose a new matching cost designed for dense stereovision based on pairs of CFA images.

Keywords: Stereovision, demosaicing, matching cost.

I. INTRODUCTION

Dense stereo correspondence algorithms are based on measures of the similarity between image locations in a pair of stereo images. Typically, a matching cost is computed at each pixel of the left image for all the shifts in a predefined range, *i.e.* for a limited set of candidate pixels in the right image. Then, the candidate pixel which minimizes the cost is retained and its position yields the disparity. Matching costs assume that homologous pixels have almost the same component values, but they cope with limited radiometric changes and/or with noise. Common window-based matching costs include the sum of absolute or squared differences (SAD/SSD), normalized cross-correlation (NCC), and census transform [1]. Chambon et al have compared widely used stereo matching costs applied to gray level and color images [2]. They have shown that taking into account color information generally improves the performance of matching costs [3].

Color images can be acquired by two types of cameras: those including three sensors associated with beam splitters and color filters providing the so-called full-color images, and those including a single image sensor. Many recent digital cameras include a single-chip CCD (Charge Coupled Device) or CMOS (Complementary Metal Oxide Semiconductor) sensor, to increase the image size while reducing the device cost. The surface of such a sensor is covered with an array of small spectrally selective filters, arranged in an alternating pattern, so that each photo-sensitive element samples only one of the three color components Red (R), Green (G) or Blue (B). These single-sensor cameras actually provide a color filter array (CFA) image, where each pixel is characterized by a single color component. Fig. 1 shows the Bayer CFA image acquired by most of these cameras. To estimate the color vector $(R\ G\ B)^T$ at each pixel, one has to determine the levels of the two missing components. This process is commonly referred to as CFA demosaicing, and yields a color demosaiced image where each pixel is characterized by an estimated color vector [4] [5] [6].

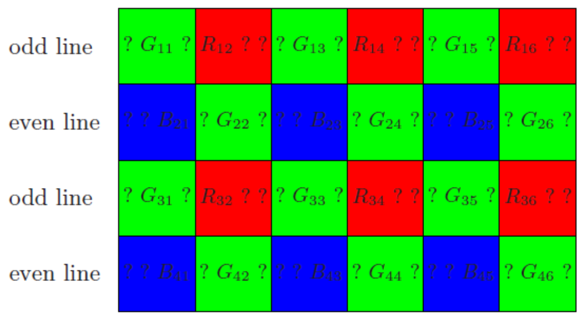

Fig. 1. Bayer CFA Filter.

Since the demosaicing methods intend to produce "perceptually satisfying" demosaiced images, they attempt to reduce the presence of color artifacts, such as false colors or zipper effects, by filtering the images [7]. So, some useful color texture information may be erased in the color demosaiced images. However, to match homologous pixels, window-based stereo matching costs need as much local texture information as possible. Thus, the quality of stereo matching on color demosaiced image pairs may suffer either from color artifacts or from the removal of color texture caused by demosaicing schemes.

In order to avoid this problem while exploiting color, we propose a new matching cost designed for stereovision based on color mono CCD cameras. In Section II, we briefly introduce dense color stereovision. The problem of demosaicing is discussed in Section III. Then, we propose in Section IV a new cost to deal with CFA image. Finally, experimental results are presented in Section V.

II. DENSE STEREOVISION

Stereovision schemes aim at computing a three dimensional representation of a scene observed by two cameras. Stereo correspondence of homologous pixels, *i.e.* pixels in the left and right images onto which the same physical point of the scene is projected, allows for 3D reconstruction. One of the key points of stereovision is to find these homologous pixels through stereo matching [4]. Sparse stereovision matching techniques match only the pixels marked on salient image features, such as lines or corners. Their performance depends on the quality of the primitive detection stage [2]. On the other hand, dense stereovision matching techniques search the homologous of every pixel.

When the geometry of the stereovision setup is precisely known, the images can be rectified. After rectification, epipolar lines correspond to horizontal lines in the images, and homologous pixels have the same vertical coordinate. Let us consider a pixel in a left image, called left pixel and denoted P_L with spatial coordinates (x_L, y_L). The spatial coordinates of its homologous pixel P_R in the line at the same vertical position of the right image are (x_R, y_L) (see Fig. 2). The disparity d, estimated at the left pixel P_L, is expressed as:

$$d(P_L) = x_L - x_R \tag{1}$$

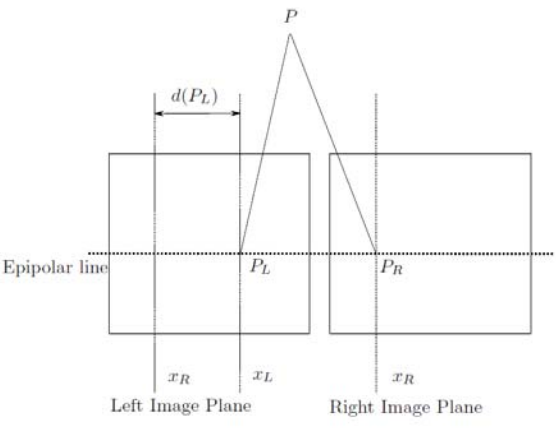

Fig. 2. Disparity between two homologous pixels.

The objective of the dense stereovision scheme is to estimate the disparity at each left pixel in order to produce the disparity map from which it is possible to reconstruct the 3D scene. For this purpose, it measures local similarity between the levels of the neighbors of the considered left pixel and the levels of the neighbors of each candidate right pixel thanks to correlation scores.

The sum of the absolute differences (SAD) between colors of neighboring pixels is one of the most widely used matching cost functions. The SAD score between the left pixel P_L with spatial coordinates (x_L, y_L) and a candidate pixel in the right image, with the s-shifted spatial coordinates $(x_L - s, y_L)$, is expressed as:

$$\begin{aligned} C(x_L, y_L, s) = \sum_{i=-w}^{w}\sum_{j=-w}^{w} & \left|R^L(x_L+i, y_L+j) - R^R(x_L+i-s, y_L+j-s)\right| \\ & + \left|G^L(x_L+i, y_L+j) - G^R(x_L+i-s, y_L+j-s)\right| \\ & + \left|B^L(x_L+i, y_L+j) - B^R(x_L+i-s, y_L+j-s)\right| \end{aligned} \tag{2}$$

where *R*, *G* and *B* are the color components of a pixel *P*, *s* is the spatial shift along the horizontal epipolar line, and *w* the half-width of a (*2w* +1)×(*2w* +1) correlation window.

SAD scores computed for different candidates, *i.e.* for different shifts *s*, are then compared. With respect to the *winner takes all (WTA)* method, the candidate pixel yielding the lowest SAD score is matched to the considered left pixel and the estimated disparity $\hat{d}(P_L)$ is given by:

$$\hat{d}(P_L) = \arg\min_s C(x_L, y_L, s) \tag{3}$$

III. DEMOSAICING AND COLOR STEREOVISION

In order to show the limits reached by applying SAD-based matching to a pair of color demosaiced images, we propose to consider Murs image pair available at http://www.irit.fr/~Benoit.Bocquillon and shown in Fig. 3.

For comparing SAD-based performance on full and demosaiced color images, we computed artificial left and right CFA images by removing two color components of each pixel according to the Bayer CFA (see Fig. 1). Then, the two missing color components were estimated by the Hamilton method [5] to produce demosaiced images. In those, each pixel *P* is characterized by an estimated three-dimensional color vector denoted $(R\ G\ B)^T$.

Hamilton method was selected since it reaches a good compromise between demosaicing quality and processing time [8].

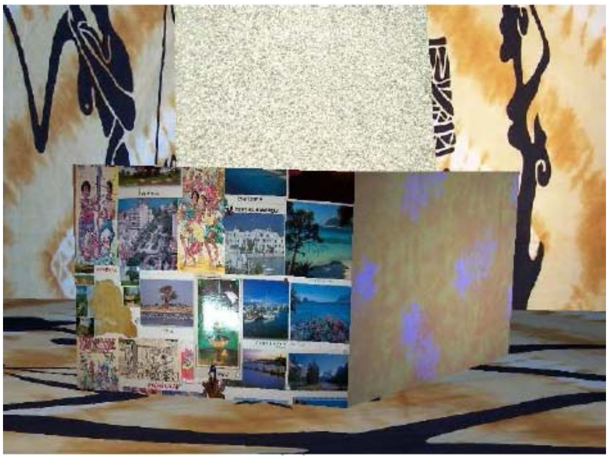

Fig. 3. Murs left image.

We matched all the pixels by computing color SAD scores on these two pairs of color stereo images. By comparing the estimated disparity $\hat{d}(P_L)$ and the ground truth disparity $d(P_L)$, we can estimate the percentage of correctly matched pixels, i.e. pixels for which the difference between the estimated and the ground truth disparities is lower than or equal to one pixel. Fig. 4 shows this percentage for full and demosaiced color image pairs as function of correlation window half-width w. It arises that, whatever window width, the percentage of correctly matched pixels for demosaiced color images is lower than for full color images. We notice that the difference of the correctly matched pixel percentages ranges between 2% and 4% although the demosaiced color Murs image seems to be visually identical to the full color one. This experimental result demonstrates that the demosaicing step degrades the quality of stereo matching. That led us to propose a SAD score specifically designed for CFA images.

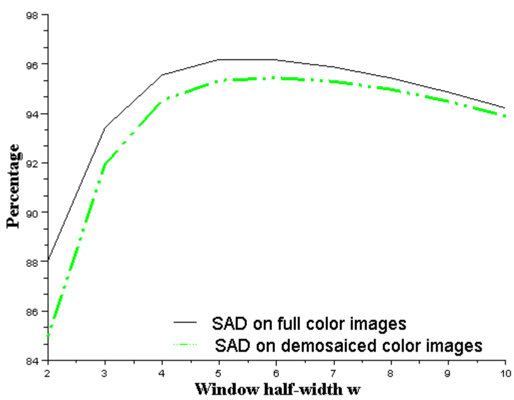

Fig. 4. Matching percentages on full and demosaiced color images.

IV. PARTIAL SAD COST

The main problem with CFA stereovision is that the available color components of homologous pixels in the left and the right images may be different. For example, let us examine Fig. 5 that shows a situation in which a physical space point P is projected onto a green pixel P_L in the left CFA image and onto a red pixel P_R in the right CFA image. A green (resp. red) pixel in a CFA image is characterized by only the green (resp. red) color component.

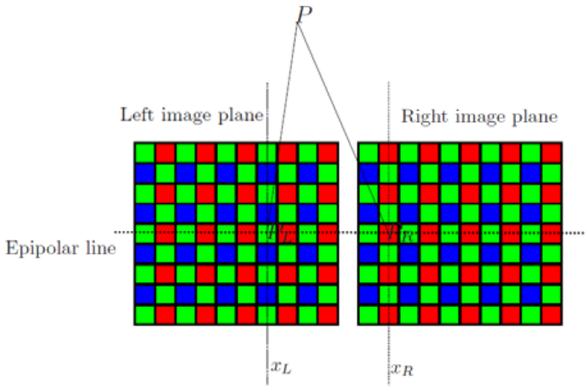

Fig. 5. CFA stereovision problem.

Therefore, one cannot assume that the green level, only available value for the left pixel, is equal to the red level of its homologous pixel in the right CFA image.

We use only the missing red or green level of each pixel located on an odd line, and only the missing blue or green level of each pixel located on an even line. Therefore, each pixel of an odd (resp. even) line is characterized only by its red (resp. blue) and green levels in the cost function. The missing color component is estimated by Hamilton's approach [5].

We propose to modify the SAD score of equation (2) to:

$$C(x_L,y_L,s)=\sum_{i=-w}^{w}\sum_{j=-w}^{w}\left|G^L(x_L+i,y_L+j)-G^R(x_L+i-s,y_L+j-s)\right| + \left|RB^L(x_L+i,y_L+j)-RB^R(x_L+i-s,y_L+j-s)\right| \quad (4)$$

where *RB* is the red color component in the odd lines and the blue color component in the even ones. We call this modified SAD score *partial SAD cost*.

V. RESULTS

In order to compare the quality of pixels matching, we apply the partial SAD coston the Murs demosaiced images. Fig. 6 shows the rates of correctly matched pixels obtained by applying SAD cost on the original full color images and the demosaiced color images, and partial SAD cost on the demosaiced images. We remark that our partial SAD costoutperforms the standard SAD cost applied to demosaiced color images. Obviously, our method does not reach the matching quality obtained on full color images even if the difference between these two rates decreases when the size of the correlation window increases. Furthermore, the processing time needed for demosaicing and for computing the partial SAD score is lower than that required by total demosaicing and SAD score computation since only one color component is estimated for each pixel and since the partial SAD cost takes into consideration two color components instead of three.

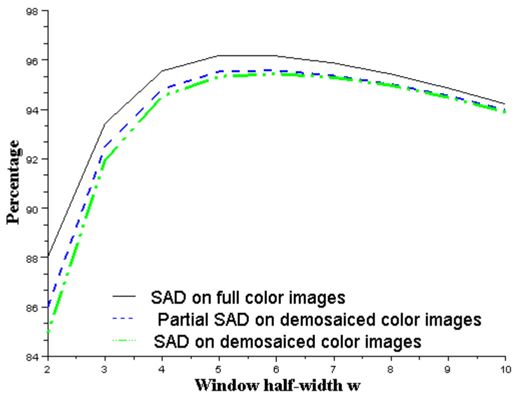

Fig. 6. Matching percentages of SAD applied on full and demosaiced color images and partial SAD applied on demosaiced color images.

VI. CONCLUSION

In this paper, we outlined that the demosaicing step can decrease the quality of pixels matching by considering images acquired by single-sensor color cameras.

We proposed a modified SAD cost, specifically designed to match pixels of stereo CFA demosaiced color images. We have experimentally shown that using partial cost instead of the standard one total demosaicing improves the disparity estimation results. Moreover, the proposed method is faster than the classical one.

We have assumed that the pair of CFA demosaiced images had been previously rectified so that assumptions on epipolar lines are satisfied. However, the rectification of CFA demosaiced images is an open problem.

REFERENCES

H. Hirschmuller and D. Scharstein. Evaluation of cost functions for stereo matching. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07)*, Minneapolis, USA, June 2007, pp. 1–8.

S. Chambon and A. Crouzil. Color stereo matching using correlation measures. In *SEE, editor, Proceedings of the First International Conference on Complex Systems Intelligence and Modern Technological Applications*, Cherbourg, France, september 2004, pp. 520–525.

E. Pinhasov, N. Shimkin, and Y. Zeevi. Optimal usage of color for disparity estimation in stereo vision. In *13th European Signal Processing Conference (EUSIPCO'05)*, Antalya, Turkey, september 2005.

S. Battiato, M. Guarnera, G. Messina, and V. Tomaselli. Recent patents on color demosaicing. *Recent Patents on Computer Science*, 1(3): 194–207, november 2008.

B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau. Demosaicking: color filter array interpolation. *IEEE Signal Processing Magazine*, 22(1): 44–54, january 2005.

R. Lukac. *Single-Sensor Imaging: Methods and Applications for Digital Cameras*. CRC Press, 2008.

Y. Yang, O. Losson, and L. Duvieubourg. Quality evaluation of color demosaicing according to image resolution. In *Proceedings of the 3rd International Conference on Signal-Image Technology & Internet-based Systems (SITIS'07)*, Shanghai Jiaotong University, China, December 2007, pp. 640–646.

E. Pinhasov, N. Shimkin, and Y. Zeevi. Optimal usage of color for disparity estimation in stereo vision. In *13th European Signal Processing Conference (EUSIPCO'05)*, Antalya, Turkey, september 2005.