Color Stereo Matching Cost Applied to CFA Images

Hachem Halawana, Ludovic Macaire, and François Cabestaing

LAGIS, USTL, Bt. P2, Cit Scientifique, 59655 Villeneuve dAscq. France
hachem.halawana@ed.univ-lille1.fr,
{ludovic.macaire,francois.cabestaing}@univ-lille1.fr
http://lagis-vi.univ-lille1.fr/

Abstract. Most color stereovision setups include 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: Color stereovision, CFA image, Demosaicing.

1 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 minimizing 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 [6]. 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 [8].

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 image size while reducing 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 CFA image, where each pixel is characterized by a single color component. Figure 1 shows the Bayer CFA which is the most widely used one. To estimate the color vector $(R\ G\ B)^T$ at each pixel, one has to determine the levels of the two missing components in figure 1. 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 [1] [4] [7] .
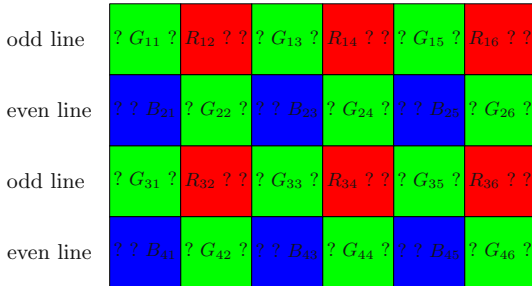

Fig. 1. Bayer Color Filter Array image (? corresponds to the missing color components)

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 [10]. So, 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 CFA images. In section 2, we briefly introduce dense color stereovision. In the third section of this paper, we present the problems inherent in CFA stereovision. In the fourth section, we detail our proposed matching cost function which is specifically designed to compare levels of pixels in CFA images. Experimental results on synthetic and real color stereo images are provided in the last section, in order to show the effectiveness of our proposed matching cost applied to pairs of CFA images.

2   Dense Color 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 [2]. 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 [3]. On the other hand, dense stereovision matching techniques search the right homologous pixel of every left pixel.

When the geometry of the stereovision setup is precisely adjusted as Bumble-bee system does (available at `http://www.ptgrey.com/products/bumblebee2/index.asp`), 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 $(x_L \ y_L)^T$. The spatial coordinates of its right homologous pixel P_R in the line at the same vertical position of the right image are $(x_R \ y_R)^T$ (see figure 2). The disparity d, estimated at the left pixel P_L, is expressed as:
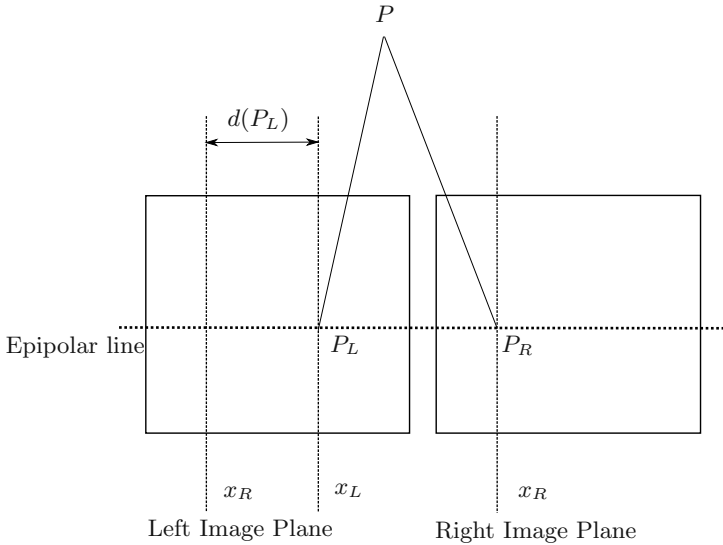
$$d(P_L) = x_L - x_R \ . \tag{1}$$

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

$$SSD(x_L, y_L, s) = \frac{1}{(2w+1)^2} \sum_{i=-w}^{w} \sum_{j=-w}^{w} \tag{2}$$

$$\|C(P_L(x_L + i, y_L + j)) - C(P_R(x_L + i - s, y_L + j))\|^2 \; ,$$

where $C(P)$ is the color vector $(R\ G\ B)^T$ of a pixel P, s is the spatial shift along the horizontal epipolar line, and w the half-width of a $(2w+1) \times (2w+1)$ correlation window.

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

$$\widehat{d}(P_L) = \arg\min_s (SSD(x_L, y_L, s)) \; . \tag{3}$$

In order to show the limits reached by applying SSD-based matching to a pair of color demosaiced images, we propose to consider the benchmark Tsukuba color stereo images [9], whose ground truth disparity map is available (figure 3(a) shows the left image of the Tsukuba pair).

For comparing SSD-based performance on full and demosaiced color images, we compute artificial left and right CFA images by removing two color components of each pixel according to the Bayer CFA (see figure 1). Then, the
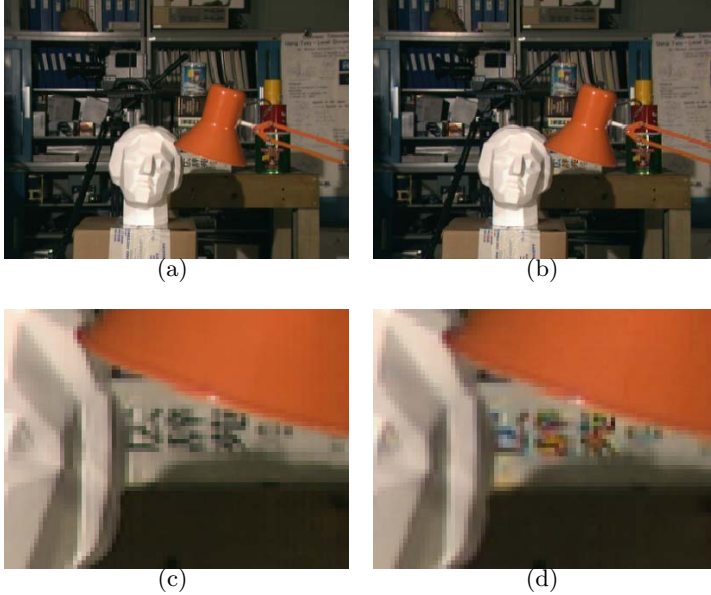

(a)   (b)

(c)   (d)

Fig. 3. Tsukuba left full color image (a), left demosaiced color image (b), zoom on full color image (c), and zoom on demosaiced color image (d)
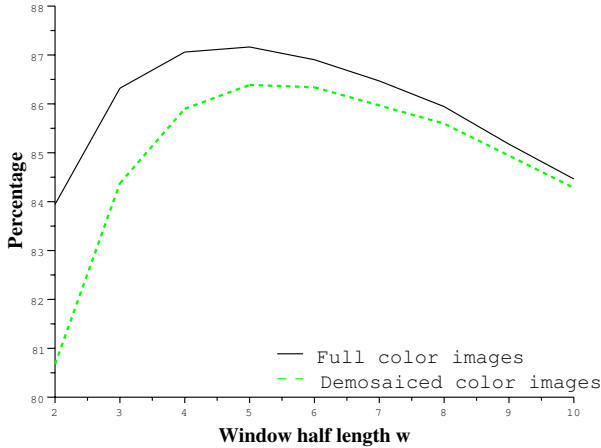
Fig. 4. Percentage of pixels correctly matched by SSD applied to Tsukuba color images of figure 3

two missing color components are estimated by Hamilton's method [5] to produce demosaiced images. In those, each pixel P is characterized by an estimated three-dimensional color vector denoted $\tilde{C}(P)$. Hamilton's method has been selected since it reaches a good compromise between demoisaicing quality and processing time [10]. The original left color image (fig. 3(a)) and the demosaiced one (fig. 3(b)) look very similar. However, zooming on textured areas (fig. 3(c) and fig. 3(d)) shows that these two images are locally very different: here, false colors have been appeared in the demosaiced image.

We match all the pixels by computing color SSD scores on these two pairs of color stereo images. By comparing the estimated disparity $\widehat{d}(P_L)$ 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.

Figure 4 shows this percentage which is obtained by SSD applied to the full and demosaiced color image pairs ,respectively, with respect to different correlation window half-widths w. It arises that, whatever the window width, the percentage of correctly matched pixels with demosaiced images is lower than with full color images. We notice that the difference of correctly matched pixel percentages ranges between 5% and 2% although the demosaiced color image of Tsukuba seems to be visually identical to the full color one.

When w ranges between 2 and 5, the assumption about constant disparity inside the correlation window is verified. The difference between the rates obtained with full and demosaiced images is mainly due to the error of missing color components estimation. That explains why the percentage of correctly matched pixels increases with respect to w. However, when w is higher than 5,

the assumption about constant disparity is no longer verified and the percentage of correctly matched pixels decreases. This is the reason why the difference between the rates obtained with full and demosaiced images decreases when w is higher than 5. These experimental results demonstrate that the demosaicing step degrades the quality of stereo matching. That leads us to propose a SSD score specifically designed for CFA images.

3   CFA Stereovision

From a pair of stereo CFA images acquired by two single-sensor cameras, we aim at calculating the disparity map. We have found no algorithm available in the literature that computes the disparity map by processing directly two CFA images. We propose to adapt the matching cost function designed for color images (eq. 3) in order to take into consideration the specific properties of CFA images.

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

This example shows that the SSD score cannot be computed directly from CFA values, since the assumption that the CFA levels of homologous pixels are similar is not met for odd disparities. More precisely, the integer disparity value for each left pixel can be:
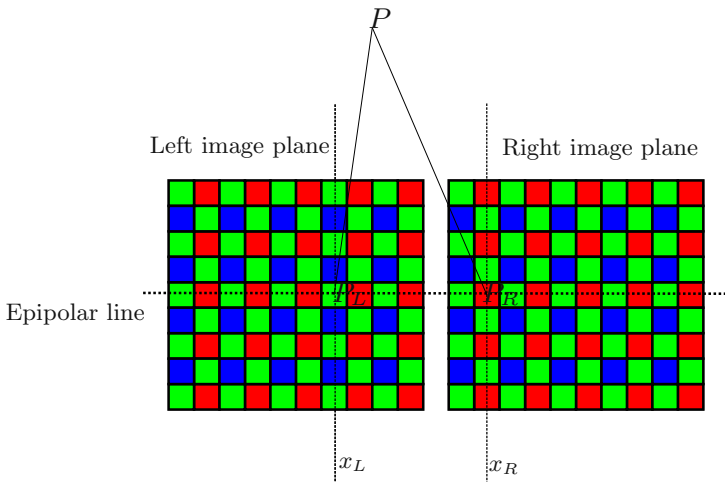

Fig. 5. Problem of matching CFA images

- Even: the right homologous pixel is characterized by the same available color component in the right CFA image. The assumption about similarity of levels is met.
- Odd: the right homologous pixel is characterized by another available color component in the right CFA image. The assumption about similarity of levels is not met.

Whatever the parity of the actual disparity $d(P_L)$, the adapted SSD score must reach a minimum when s is equal to the actual disparity d so that the estimated disparity $\widehat{d}(P_L)$ is equal to the actual one $d(P_L)$.

4   Partial Demosaicing for CFA Stereovision

We assume that the matching errors of standard SSD applied to demosaiced color images are mainly caused by the error of estimation of the two missing color components at each pixel. We propose to reduce the matching error by estimating only one missing color component at each pixel.

The Bayer pattern is designed so that the pixels in one line of the CFA image are characterized by one among two possible color components. Figure 1 shows that pixels of odd (resp. even) lines are characterized by red or green levels (resp. blue or green levels). So, the modified SSD score can consider only two color components, the available one and an estimated one, rather than three. For matching purposes, we propose to estimate only the missing color component at each pixel (called hereafter second component) that exists in the same line as shown in figure 6.

We estimate 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 so-called *partially* demosaiced image. Each pixel in the partially demosaiced image is characterized by a two-dimensional partial color vector denoted $\tilde{C}_P(P)$. The single missing color component is also estimated by Hamilton's approach [5].
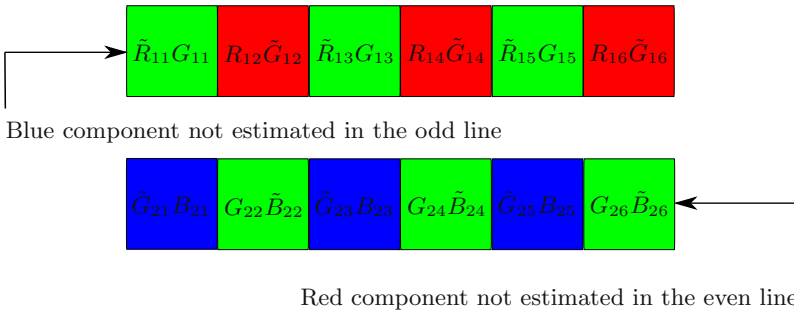

Blue component not estimated in the odd line

Red component not estimated in the even line

Fig. 6. Partially demosaiced image

The SSD score described by eq. (3) is modified so that it can be applied to partially demosaiced images:

$$SSD_P(x_L, y_L, s) = \frac{1}{(2w+1)^2} \sum_{i=-w}^{w} \sum_{j=-w}^{w} \tag{4}$$

$$\left\| \tilde{C}_P(P_L(x_L + i, y_L + j)) - \tilde{C}_P(P_R(x_L + i - s, y_L + j)) \right\|^2.$$

Since the pixels of horizontal lines with the same parity in the left and right partially demosaiced images are characterized by the same two color components, we can reasonably assume that the partial color vectors of two homologous pixels are similar. Since our partial SSD_P score compares the partial color vectors of left and right pixels located on the same horizontal lines, we assume that SSD_P reaches a minimum when the shift s is equal to the actual disparity.

5   Experimental Results

In order to compare the quality of pixel matching, we first apply the partial SSD_P score on the Tsukuba partially demosaiced images. Figure 7 shows the rates of correctly matched pixels obtained by analyzing the original full color images, the demosaiced images and the partially demosaiced images. We remark that our partial SSD score outperforms the standard SSD 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 partial demosaicing and for computing the SSD_P score is lower
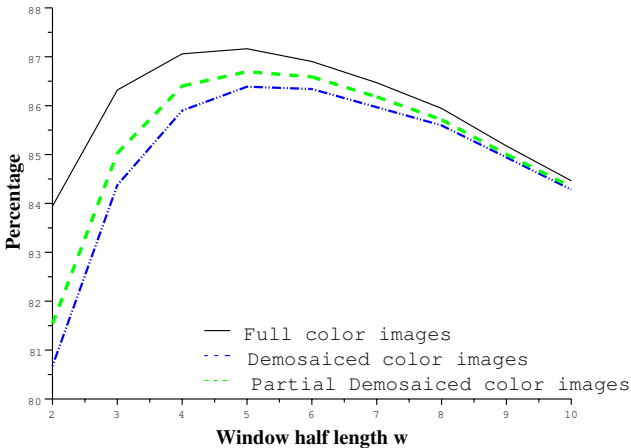

Fig. 7. Percentage of correctly matched pixels by SSD applied to Tsukuba color images(full color and demosaiced) and SSD_P applied to partially demosaiced images
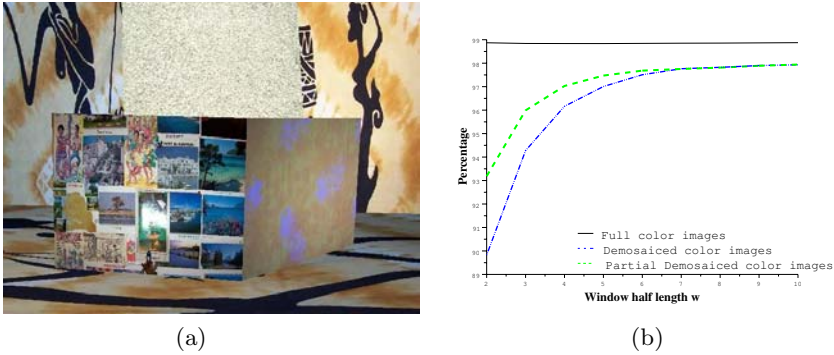
(a)   (b)

Fig. 8. Percentage of exactly matched pixels by SSD applied to mur color images (full color and demosaiced) and SSD_P applied to partial demosaiced images

than that required by total demosaicing and SSD computation since only one color component is estimated for each pixel and since the partial SSD_P takes into consideration two color components instead of three.

We have explained in section 3 that the main problem of CFA stereovision is caused by odd disparities. In order to examine the behavior of our approach in this case, we use the photo-realistic image 'murs' designed by Bocquillon and available at `http://www.irit.fr/~Benoit.Bocquillon/MYCVR/download.php`. All the pixel locations of the left image are shifted by an odd value (nine pixels in our case) to produce the right image. The actual disparity between all pairs of homologous pixels of these two images is therefore equal to nine. Then, we measure the percentage of exactly matched pixels, i.e. pixels whose difference between estimated and ground truth disparity is zero, for the full color, demosaiced and partially demosaiced images respectively. Figure 8 shows that the exactly matched pixel rate obtained by considering the full color images is close to 100%. We also notice that the difference between exactly matched percentages obtained by analyzing demosaiced and partially demosaiced images ranges from 0% to 4%. These results demonstrate that the partial demosaicing improves the matching process, and is robust against the odd disparity problem.

6   Conclusion

In this paper, we have outlined that the demosaicing step can decrease the quality of pixel matching by considering images acquired by single-sensor color cameras. We have proposed a modified SSD score, specifically designed to match pixels of stereo CFA images. We have experimentally shown that using partial demosaicing instead of total demosaicing improves the disparity estimation results. Moreover, the proposed method is faster than the classical one.

Future works will show the reason why the partial demosaicing improves the quality of matching by analyzing not only the stereo matching results but also the color estimation errors. The second color component of partially demosaiced images is estimated thanks to Hamilton's method which uses a two-dimensional window for color estimation. We will try in future work to simplify this step by using a simple one-dimensional interpolation and study its influence on the quality of disparity estimation. It would be also interesting to study the performance of our matching applied only to the luminance information extracted from CFA images.

References

1. Battiato, S., Guarnera, M., Messina, G., Tomaselli, V.: Recent patents on color demosaicing. Recent Patents on Computer Science 1(3), 194–207 (2008)
2. Chambon, S., Crouzil, A.: Color stereo matching using correlation measures. In: SEE (ed.) Proceedings of the First International Conference on Complex Systems Intelligence and Modern Technological Applications, Cherbourg, France, September 2004, pp. 520–525 (2004)
3. Gouet, V., Montesinos, P., Pelé, D.: Stereo matching of color images using differential invariants. In: Proceedings of the 5th International Conference on Image Processing, ICIP 1998, Chicago, IL, USA, October 1998, vol. 2, pp. 152–156 (1998)
4. Gunturk, B.K., Glotzbach, J., Altunbasak, Y., Schafer, R.W., Mersereau, R.M.: Demosaicking: Color filter array interpolation. IEEE Signal Processing Magazine 22(1), 44–54 (2005)
5. Hamilton, J.F., Adams, J.E.: Adaptive color plan interpolation in single sensor color electronic camera. US patent 5, 629, 734, to Eastman Kodak Co., Patent and Trademark Office, Washington, DC (May 1997)
6. Hirschmüller, H., Scharstein, D.: Evaluation of cost functions for stereo matching. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007), Minneapolis, June 2007, pp. 1–8 (2007)
7. Lukac, R.: Single-Sensor Imaging: Methods and Applications for Digital Cameras. CRC Press, Boca Raton (2008)
8. Pinhasov, E., Shimkin, N., Zeevi, Y.: Optimal usage of color for disparity estimation in stereo vision. In: 13th European Signal Processing Conference (EUSIPCO 2005), Antalya, Turkey (September 2005)
9. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision 47, 7–42 (2002)
10. Yang, Y., Losson, O., Duvieubourg, L.: Quality evaluation of color demosaicing according to image resolution. In: Proceedings of the 3rd International Conference on Signal-Image Technology & Internet-based Systems (SITIS 2007), Shanghai Jiaotong University, China, December 2007, pp. 640–646 (2007)