Color-compensated Stereo Matching Algorithm for Color Inconsistent Stereo Pair

Jae-Il Jung and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712, Korea
Telephone: +82-62-970-2258 Fax: +82-62-970-3164
E-mail: {jijung, hoyo}@gist.ac.kr

Abstract

In conventional algorithms for a stereo matching, corresponding color information between a pair of images is mainly used to find matching points. However, due to various reasons, color information of both images cannot be matched accurately in most practical cases. In this paper, we propose a color-compensated stereo matching algorithm for severely distorted stereo images. Initial corresponding points are extracted by a feature-based matching method from histogram-equalized stereo images. From this information, we estimate relation of color formation of stereo images with a sigmoid function, and compensate for differences. Then, we apply the block-based stereo matching algorithm using the graph cut method. Our method shows good performance when input images are severely distorted.

Keywords: Stereo Matching, Color Compensation

1. Introduction

In recent times, the three-dimensional (3-D) video service has attracted much attention due to its various applications. The 3-D video service provides a dramatic enhancement in the viewing experience and 3-D feeling to users. It is based on 3-D information, depth information, which represents distance between a camera and objects in an image. Acquisition of depth information, therefore, is an essential part for 3-D video service. Various methods for acquisition have been researched and can be classified into two types: an active sensor-based method and a passive sensor-based method. In the active sensor-based methods, depth information is obtained from the natural scenes directly using physical sensors, such as laser or infrared ray (IR) sensors [1]. Although these sensors can extract accurate depth information, these equipments are expensive and have several critical problems. Passive sensor-based methods estimate depth information indirectly from 2-D images captured by two or more cameras at different position. Among these methods, one of the well-known indirect depth methods is the stereo matching. The stereo matching is one of the most extensively researched topics in computer vision and tries to identify corresponding points in both images that refer to the same scene point. Various approaches have been developed for the steo matching, and most algorithms have been researched under a common assumption that corresponding pixels in stereo images have a similar color value. This property is called the intensity conservation assumption. Since this assumption cannot be valid in a practical environment, pixels that correspond to the same scene point can have the different values due to the different property of each camera, jitter of shutter speed and aperture, or illumination condition [2,3]. The human visual system contains a color compensating process which is able to compute colors irrespective of radiometric variations and estimate the reflectance of the objects [4]. However, most stereo matching algorithms do not contain this process. Although there are some alternative cost functions to handle radiometric variation in stereo pair, they cannot cover regions which do not contain feature points, such as edges and corners or should compute matching cost with their own cost function. In this paper, a stereo matching algorithm containing a color compensation part without changing of the cost function is proposed to overcome radiometric variation.

2. Related Works

In this section, we review related stereo matching algorithms. The performances of conventional stereo matching algorithms depend on the radiometric variation between input stereo pair Hence, alternative methods are necessary for handling this problem. Wang *et al.* presented a new invariant for stereo reconstruction called light transport constancy, which allows completely arbitrary scene reflection. This invariant can be used to formulate a rank constraint on the stereo matching [5]. Their method requires at least two stereo image pairs with different illumination conditions to be available for making use of rank constraint. Normalized cross correlation is a very popular and powerful cost function for matching contrast varying images [6]. It is particularly useful since it is insensitive to both signal strength and level. However it is only suitable for matching affine-transformed values and cannot cover regions which do not contain feature points. Therefore, a stereo matching method which is robust to radiometric variation and independent to the cost function is needed. In this paper, a stereo matching algorithm containing a color compensation part without changing of the cost function is proposed.

3. Proposed Algorithm

In this paper, we introduce the color-compensated stereo matching algorithm for severely distorted stereo images. We compensate the different color distribution between stereo images before the stereo matching process. The feature-based matching algorithm and the non-linear regression are used for compensation. Then, the stereo matching with the cost function, sum of absolute difference (SAD), is used. Figure 1 shows the flowchart of the proposed algorithm.

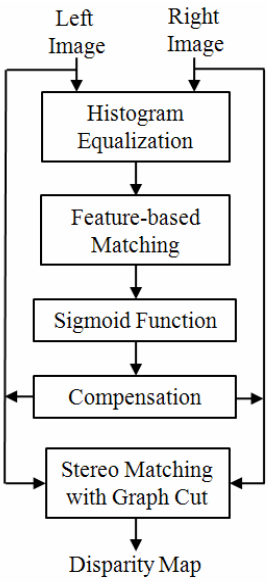


Figure 1. Flowchart of the proposed algorithm

3.1 Histogram Equalization

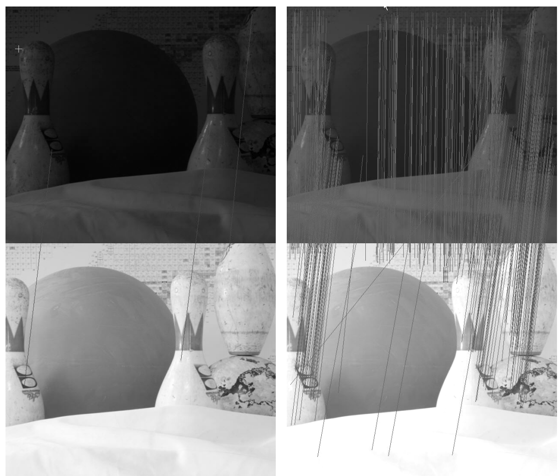
In the second step, we find corresponding points using the feature-based matching algorithm. However, this process is not straightforward due to severe difference of color values between input stereo pair. In order to solve this problem, we equalize the histograms of both input images by the following formula. In this formula, s_k and r_k are pixel values in the histogram equalized and original images, respectively.

$$\begin{aligned} s_k &= T(r_k) = \sum_{j=0}^{k} p_r(r_j) \\ &= \sum_{j=0}^{k} \frac{n_j}{n} \quad , k = 0,1,\ldots,L-1 \end{aligned} \tag{1}$$

where n is the total number of pixels in the image, n_k is the number of pixels that have value r_k and L is the total number of possible levels in the image. $p_r(r_k)$ denotes the probability density functions of random variables r_k.

3.2 Feature-based Matching

After applying the histogram equalization, we extract correspondence by using the sift-invariant feature transform (SIFT) [7] which is a method for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. Figure 2 (a) shows the result of the SIFT process without the histogram equalization process, few correspondences are found due to the severely distorted color distribution. The lines in the figure stand for correspondence between two images. Contrastively, when we extract correspondence based on the histogram equalized images, plenty of correspondences are extracted as shown in Fig. 2 (b).



(a) (b)

Figure 2. Results of the feature-based matching process (a) without and (b) with the histogram equalization

Although finding correspondence on the histogram equalized images, we extract pixel values of correspondence from original images. It means the histogram equalization only helps to find exact positions of correspondences without any influence to pixel values of correspondences.

3.3 Color Relation

Then, we define the color relation of stereo images, and it will be used to compensate color distribution. Because the property of the color relation is not a linear form and can contain a saturation part, as you can see sample data in Fig. 3, we use the sigmoid function in Eq. (2).

$$V_{right} = y_0 + a(1 + e^{-\left(\frac{V_{left} - x_0}{b}\right)})^{-c} \quad (2)$$

where V_{left} and V_{right} stand for pixel values of the left image to be compensated and pixel values of the right image as a reference, respectively. In order to define the relation between stereo images, five coefficients, y_0, x_0, a, b, and c, in the equation should be estimated. By using the non-linear regression with corresponding pixel values, we estimate these coefficients.

Figure 3 shows the example relation of color distribution between stereo pair. Blue points are sample pixel values extracted by the SIFT algorithm and the red line represents the estimated data using the sigmoid function. As you can see, the sigmoid function reflects the color saturation region, which can appear in severely distorted stereo images, and non-linear property. This task that estimates coefficients should be processed for each color channel, individually.

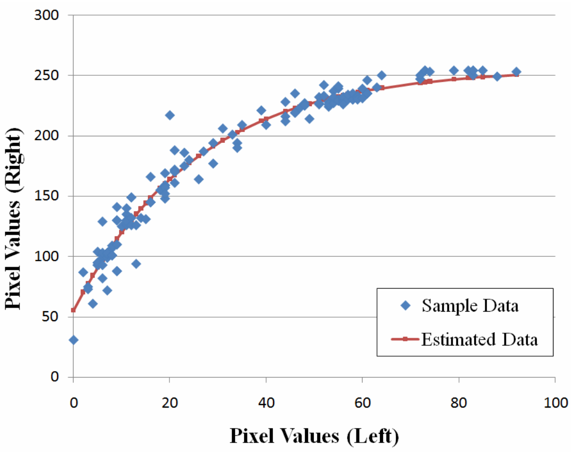


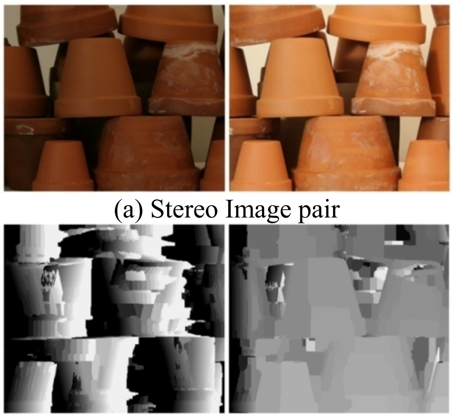
Figure 3. The relation of color distribution between left and right images

3.4 Block-based Stereo Matching

With a reference to the sigmoid function and estimated coefficients, the pixel values of left image are compensated. Then we find an initial depth map using the block-based stereo matching algorithm with the SAD cost function. Because SAD does not consider color different problem of input image, we can maximize the effect of out color compensation part. After finding the initial depth map, it is refined with graph cut algorithm [8, 9].

4. Experimental Results

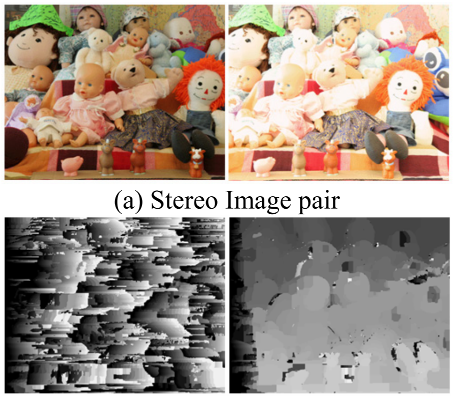
In order to evaluate the performance of the proposed algorithm, we utilized a data set of Middlebury [10] as a test image. Each data set has different exposure time.



(a) Stereo Image pair

(b) SAD +GC (c) CC+SAD+GC

Figure 4. Result of *flowerpot* Images



(a) Stereo Image pair

(b) SAD +GC (c) CC+SAD+GC

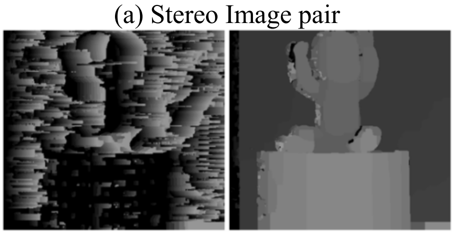
Figure 5. Result of *doll* Images

Figure 4,5 and 6 show the results of proposed algorithm for different input images. In each figure, (a) is input stereo pair, and (b) and (c) represent the depth map without and with color compensation (CC) part.

Test set	Original Images SAD + GC					Distorted Images SAD + GC					Distorted Images CC + SAD + GC				
	all	non occ	occ	textur ed	textur eless	all	non occ	occ	textur ed	textur eless	all	non occ	occ	textur ed	textur eless
baby	5.3	2.2	17.0	2.1	2.3	25.9	26.5	17.8	20.9	30.6	4.9	1.9	16.3	1.9	1.9
bowling	14.5	9.7	30.0	4.2	10.9	38.7	37.3	46.3	32.9	38.6	24.0	21.6	34.7	5.9	24.7
doll	11.3	3.1	30.3	3.1	2.9	30.1	29.1	36.0	27.7	31.7	12.4	5.8	30.7	5.6	6.1
flowerpot	16.9	8.7	33.6	13.7	7.7	38.8	36.7	46.0	32.5	37.3	19.1	10.9	36.8	10.6	11.0

Table : Middlebury stereo evaluation (root mean square error)

(a) Stereo Image pair



(b) SAD +GC (c) CC+SAD+GC

Figure 6. Result of *baby* Images

As you can see, the depth maps without color compensation part have coarse quality because the SAD cost function cannot cover the color difference. However, after applying proposed algorithm, the quality of depth map is remarkably increased. Table 1 shows the results of quantitative evaluation. The results of original and distorted images are from input images captured at same and different exposure time, respectively. It shows that our proposed algorithm can effectively handle radiometric variations without changing of the cost function.

4. Conclusions

In this paper, we have proposed the color-compensated stereo matching algorithm for severely distorted stereo images. We extracted initial corresponding points using the feature-based matching method from histogram-equalized input images. Then, the relation of color formation of both images was estimated with a sigmoid function. After compensating, the stereo matching algorithm using graph cut was applied. From the experimental results, we can confirm that our proposed method shows good performance under severely distorted input images.

Acknowledgements

This work was supported in part by ITRC through RBRC at GIST (IITA-2009-C1090-0902-0017).

References

[1] C. S. Swiss Ranger SR-2, The swiss center for electronics and microtechnology. http://www.csem.ch/fs/imaging.htm.

[2] G. D. Finlayson, B. Schiele, and J. L. Crowley, “Comprehensive colour image normalization,” In *Proc. of European Conference on Computer Vision*, 1998.

[3] H. Hirschmuller and D. Scharstein, “Evaluation of cost functions for stereo matching,” In *Proc. of Computer Vision and Pattern Recognition*, 2007.

[4] D. J. Jobson, Z. Rahman, and G. A. Woodell, “A multiscale retinex for bridging the gap between color images and the human observation of scenes,” *IEEE Trans. Image Processing*, vol. 6, no. 7, pp. 965-976, 1997.

[5] L.Wang, R. Yang, and J. E. Davis. “BRDF invariant stereo using light transport constancy,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 29, no. 9, pp. 1616-1626, 2007.

[6] O. Faugeras, B. Hotz, H. Mathieu, T. Vieville, Z. Zhang, P. Fua, E. Theron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy, “Real time correlation based stereo: algorithm implementations and applications,” INRIA Technical Report, 1993.

[7] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[8] D. Scharstein, R. Szeliski, “A Taxonomu and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms,” International journal of Computer Vision, vol. 47, no. 1, pp.7-42, 2002.

[9] V. Kolmogorov and R. Zabih, “Multi-camera Scene Reconstruction via Graph Cuts,” Proceedings of the 7th European Conference on Computer Vision, pp. 82-96, 2002.

[10] http://vision .middlebury.edu/stereo/