Color Invariant Census Transform for Stereo Matching Algorithm

Soo-Chang Pei, *Fellow*, *IEEE*, and Yu-Ying Wang

Abstract – **Illumination and color invariance are important problems in computer vision (CV). The Census transform (CT) can resist change of illumination intensity and is widely used for many applications in CV and consumer electronics. However, only grayscale images can be processed in the CT algorithm. In this paper, a new color Census transform (CCT) based on a color invariance model for stereo matching is proposed because color images are with more significant features and most source images are color. To evaluate the proposed method, its resulting disparity maps and computation time are compared with grayscale modified CT (MCT). Experimental results show that the computation time of the proposed method is little more than the grayscale MCT. But the proposed method is able to significantly improve structure features of disparity maps, compared to the grayscale MCT. Further, the CCT is not to be affected by shadows, highlights, and variations in illumination.**

I. INTRODUCTION

Stereo correspondence matching is an important research topic in image processing and computer vision (CV) for consumer applications. It is developed for finding correspond-ding points in stereo pair images and the result is a disparity map.

Census transform (CT) is a non-parametric local transform which was introduced by Zabih and Woodfill [1]. It has been used widely for many consumer applications such as face detection [2], motion segmentation [3], and auditory vision [4]. However, CT only processes grayscale images. Because color images are with more significant features and most source images are color, this paper presents a new color CT on the basis of color invariance. The Gaussian color model (GCM) [5] is used for the proposed algorithm. J. Geusebroek et al. showed the shadows and highlights are successfully discounted using GCM and perfect measurements of surface reflectance properties are also proved in [5]. In this paper, by using the color invariance model, variations in shading or illumination have weak influence on the disparity maps.

Experimental results on the Middlebury database show that the proposed GCM color Census transform (GCM CCT) achieves much better performance than the grayscale modified CT (MCT) with a small increase of computation time.

II. COLOR CENSUS TRANSFORM

A. Census Transform using Color Invariance Model

Color invariance is a fundamental problem in computer vision, because the color is sensitive to the illumination condition. Many algorithms have been developed to let the features in color images more invariant but a computational and complicated algorithm is not suitable for stereo vision as a pre-processing step. The color invariance model proposed in [5] is used for this paper to improve the CT algorithm due to its robustness and simplicity. Three channels E^1, E^2, and E^3 can be obtained from RGB components using the following simple linear transform [5]:

$$\begin{bmatrix} E^1 \\ E^2 \\ E^3 \end{bmatrix} = \begin{bmatrix} 0.06 & 0.63 & 0.27 \\ 0.30 & 0.04 & -0.35 \\ 0.34 & -0.60 & 0.17 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \tag{1}$$

Color images can be converted from RGB color space to GCM by using (1) and the color differences between two pixels can be calculated by the formula of Euclidean distance. ΔE_G is the Euclidean distance between two pixels in GCM. The color differences between the center pixel and its neighbors in the mask are calculated. The average of these distances is used as the reference distance denoted as ΔE_m. The Census transform is defined as an ordered set of comparisons of pixel intensities in a certain mask. Therefore, it consists of a comparison function ξ, which is defined as

$$\xi(p_1, p_2) = \begin{cases} 1, & p_2 > p_1 \\ 0, & p_2 \le p_1 \end{cases} \tag{2}$$

Therefore, the Census transform creates a bit string for each pixel in an image I. The color Census transform of a pixel with coordinate (u,v) can be expressed as

$$CCT(u,v) = \bigotimes_{i=-m}^{m} \bigotimes_{j=-n}^{n} (\xi(\Delta E_m, I(u+i, v+j))), \tag{3}$$

where $\otimes$ denotes the concatenation operator and the Census mask size is $(2m+1)\times(2n+1)$.

B. Color Census Transform for Stereo Matching

For the calculation of the matching costs of each disparity level d, the cost function is defined as the Sum of Hamming distance (SHD). The calculated costs are stored in a 3D disparity space image (DSI). And the disparity level with the lowest cost is used as an integer disparity value.

$$\begin{aligned} &\underset{d\in[1,L]}{DSI}(u,v,d) \\ &= \underset{d\in[1,L]}{\mathrm{SHD}}(R_{CCT}(u,v), L_{CCT}(u+d,v)) \end{aligned} \tag{4}$$

where R_{CCT} and L_{CCT} are the CCT results of the right and left images respectively. L is the maximum disparity search range.

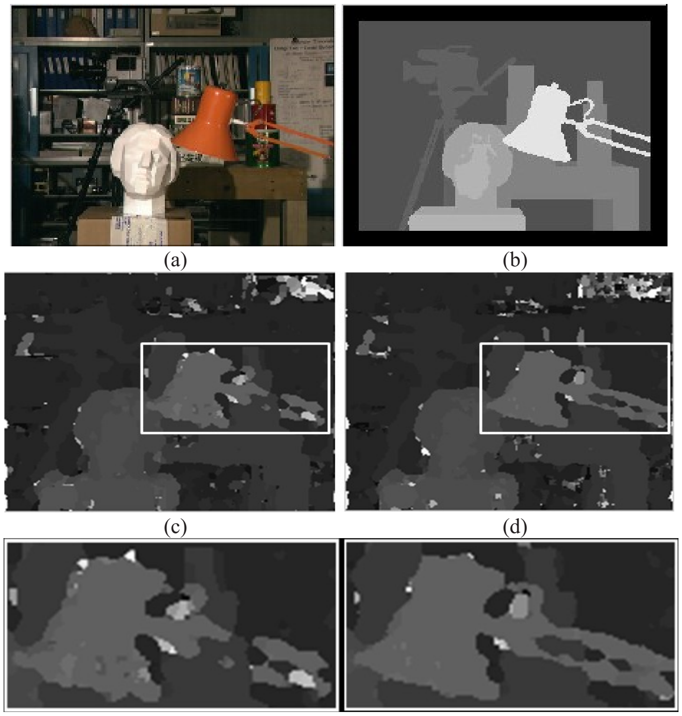

Fig. 1. (a) Original left image, (b) ground truth image. From (c) to (d): the disparity maps of 9×9 MCT and GCM CCT respectively. From (e) to (f): zoom in on areas of interest. The lamp stand in (e) is broken, but is entire in (f).

III. EXPERIMENTAL RESULTS

A. Stereo Matching Results

The experiments show the results of the grayscale MCT and the GCM CCT by using the Middlebury database. The CT mask with size of 9×9 was used for evaluation. Additionally, a 3×3 mode filter is applied as a post-processing step for stereo matching in this paper. Figs. 1(a)-(b) show the results of Tsukuba using MCT and GCM CCT. Figs. 1(c)-(d) are the magnified images of the interesting areas. The lamp stand in Fig. 1 (c) is broken, but is entire in Fig. 1 (d). It also can be seen that there are more noise in the disparity map of the MCT, compared to the disparity map of the GCM CCT. The GCM CCT presents the structure features of the image more clearly than the MCT. The GCM CCT not only maintains the structure features of the image but also represents a clear disparity map. From Figs. 1 (a), it can be noticed that there is a shadow in the original color image. And the GCM CCT performs well in the shadow (see Figs.1 (d)). In this case, the GCM CCT looks possible to get significant improvement in color stereo matching.

The disparity error (DE) of GCM CCT, a percentage of pixels with an incorrect disparity measure, is taken into considerations in performance evaluations and compared to those obtained with MCT. Table I summarizes the DE of GCM CCT and MCT for each pair of images. It shows that the GCM CCT have lower DE percentage than the MCT.

TABLE I
DISPARITY ERRORS FOR DIFFERENT METHODS

	Tsukuba	Teddy	Laundry	Art
Image resolution	384×288	450×375	447×370	463×370
Max disparity (pixel)	40	60	50	50
DE of GCM CCT (%)	28	32	28	28
DE of MCT (%)	33	34	31	29

TABLE II
COMPUTATION TIME RATIO OF DIFFERENT METHODS

	Tsukuba	Teddy	Laundry	Art
MCT	1	1.254	1.187	1.224
GCM CCT	1.102	1.345	1.275	1.287

B. Computation time comparison

Table II displays the computation time of whole procedures for all methods. In Table I, the computation time for Tsukuba using MCT is set to 1 as the reference. The computation time of the proposed method is little more than the MCT.

IV. CONCLUSIONS

In this paper, a Gaussian color model color Census transform (GCM CCT) is proposed as a new color invariant Census Transform for stereo matching. This algorithm is based on color invariance approach [5]. The experimental results show that the performance of the GCM CCT is better than the MCT because the structure features of the color image are presented more clearly by using the GCM CCT. Moreover, the GCM CCT provides better disparity maps which are not to be affected by variations in shadows. The proposed method can be naturally extended to any grayscale-based matching algorithms and improve the performance with a small increase of computation time. Future research will concentrate on finding a more suitable stereo matching algorithms based on GCM to achieve better performance in color stereo matching, such as an appropriate matching cost evaluation, aggregation and a new method to compute the color CT bit string.

ACKNOWLEDGMENT

This work was supported by the National Science Council, R.O.C. under Contract 101-2221-E-002-021-MY3.

REFERENCES

[1] R. Zabih and J. Woodfill, "Non-Parametric Local Transforms for Computing Visual Correspondance," *Proc. Third European Conf. Computer Vision*, pp. 151–158, 1994.

[2] B. Froba and A. Ernst, "Face Detection with the Modified Census Transform," *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (AFGR)*, 2004.

[3] K. Yamada, K. Mochizuki, K. Aizawa, and T. Saito, "Motion Segmentation with Census Transform," *IEEE Pacific Rim Conference on Multimedia*, pp. 903–908, 2001.

[4] S. C. Pei and Y. Y. Wang, "Census-based Vision for Auditory Depth Images and Speech Navigation of Visually Impaired Users," *IEEE Transactions on Consumer Electronics*, vol. 57, No. 4, pp. 1883–1890, November 2011.

[5] J. Geusebroek, R. van Den Boomgaard, A. Smeulders, and H. Geerts, "Color Invariance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.23, no. 12, pp. 1338–1350, 2001.