Illumination Invariant Data Cost using Modified Census Transform

Raghavendra U.
Dept. of Computer Applications
MIT, Manipal, India

Krishnamoorthi Makkithaya
Dept. of CSE
MIT, Manipal, India

Karunakar A. K.
Dept. of Computer Applications
MIT, Manipal, India

ABSTRACT
Stereo matching in non-ideal illumination is a challenging area of research. It assumes identical corresponding color values and this assumption is not guaranteed for real-time environment. As a result most of the stereo algorithms fail to generate good disparity. This paper proposes a Modified Census Correlation (MCC) data cost for stereo matching. The proposed data cost will be derived from modified census transformed indexed image and it is robust to change in lighting direction, exposure and illumination color. The obtained total energy is optimized for disparity estimation using Graph-Cut. An exhaustive evaluation using Middlebury stereo image proves the robustness of the proposed technique for variety of illumination and exposure conditions.

Keywords
Stereo Matching, Radiometric Difference and Computer Vision

1. INTRODUCTION
Stereo matching aims to extract 3D information from the scene taken from two different viewpoints. Corresponding point estimation is the basic requirement for extracting this 3D information. Radiometric difference is one of the important factor which prevents corresponding pixel from having similar color value in real-time. This might be due to variation in camera settings, exposure variations, image noise, etc,. Under this variation, majority of the existing stereo algorithms fail to identify corresponding point that leads to poor disparity estimation [1]. Hence, there is a need for robust and computationally efficient data cost for stereo matching.This paper proposes a Modified Census Correlation (MCC) data cost which is robust in handling local and global radiometric variations. Though Modified census transform is not new, extracting its correlation values, generating data cost from it and optimizing the computed energy using Graph-Cut is the actual contribution and novelty of the paper.
The paper is organized as follows. Related work is given in Section 2. Stereo energy formulation is given in Section 3. MCC data cost generation is given in Section 4. Section 5 gives Global energy model. Results and discussions are given in Section 6. Section 7 gives conclusion of the paper.

2. RELATED WORK
An efficient data cost is the basic need for every stereo matching algorithm. This cost takes an important role in deciding the quality of the output disparity [2, 3]. In the meantime, it is even more difficult to decide the type of data cost for the stereo image captured under radiometric variations. Hirschmuller et al. [1] have carried out detailedevaluation of a set of data cost with radiometrically different images caused by various factors such as change in light source, varying exposure, gamma correction variations, noise etc. They have concluded that none of them can compensate both local and global variations. Normalized Cross Correlation (NCC) [4] is a popular similarity measure for stereo matching, which is suitable solely for matching affine-transformed color values. It measures the cosine of angles between the matching vectors. It suffers from fattening effect across object boundaries as in the case of Sum of Absolute Difference (SAD) and Sum of Squared Difference (SSD). Heo et al. [5] have proposed a technique using log chromaticity normalization with adaptive normalized cross correlation. Though the technique is efficient in terms of disparity error, computational complexity is high. Ogale et al. [6] and Xu et al. [7] have presented two techniques for variation in contrast for local matching, which can compensate only global variations. The major drawback of existing methods is the usage of raw color as a data cost. This paper shows the effectiveness of MCC data cost derived from modified census transformed indexed image for compensating local and global radiometric variations.

3. STEREO ENERGY FORMULATION
The stereo matching problem can be treated as finding an optimal disparity which minimizes the global energy function. This work defines following energy function in the MAP-MRF framework [8].

$$E(f) = E_{data}(f) + E_{smooth}(f) \qquad (1)$$

$$E_{data}(f) = \sum_{p} D_p(f_p) \qquad (2)$$

$$E_{smooth}(f) = \sum_{p} \sum_{q \in N(p)} V_{pq}(f_p, f_q) \qquad (3)$$

where*N(p)* is the neighbourhood pixels of *p* and $D_p(f_p)$is the data cost, which is a measure of the dissimilarity between pixel *p* in the left image and pixel $p+f_p$ in the right image. $V_{pq}(f_p, f_q)$ is the smoothness cost that favors the piecewise smooth object. By taking these costs, optimal disparity can be derived by minimizing the total energy in Equation 1.

4. MCC DATA COST FOR STEREO MATCHING
The data cost plays a vital role in every stereo matching algorithm and it is one of the effective parameter that decides the quality of the disparity. The selection of appropriate data cost for radiometrically different image is still a challenging area of research. It is observed in the literature that raw intensity or color data cost (parametric) is not advantageous for real-time environment. There are some nonparametric methods, which considers relative ordering of local neighborhood as a data cost [9]. This data cost is independent of original intensity and produces a local structure known as kernel structure.

The Kernel indexed image can be obtained from the original image by applying the kernel structure generation function to each and every pixel. Hence, every pixel of original image will be replaced with its kernel indexed value. To compensate the local variation, an effective correlation function has to be carried out over the kernel indexed image rather than applying it on the raw image. This correlation function can handle local radiometric variation as the function operates on the local neighborhood of indexed image. By observing this, an efficient data cost is proposed for stereo matching application. Initially, modified census transformed index image will be generated using input stereo image. NCC based correlation computation will be carried out in the next stage. Stereo energy formulation will be done in the third stage followed by disparity estimation using Graph-Cut. The complete functional block diagram of the proposed method is given in Figure 1.

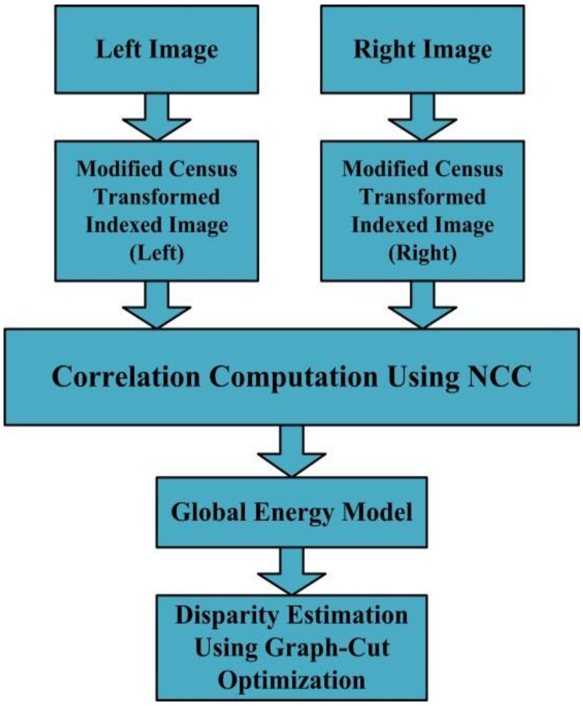

Fig 1.Functional Block Diagram of the Proposed Method

4.1 Modified Census Transform Based Indexed Image Construction

This work constructs an illumination invariant indexed image for stereo matching using a nonparametric method. Nonparametric methods are those, which considers relative ordering of the neighboring intensities instead of its raw intensities, namely Rank transform (RT) and the Census transform (CT). The Rank transform replaces the intensity of a pixel with its rank among all the pixels within a certain neighborhood [9]. Due to the ordering of pixel intensities, it increases the robustness of local stereo matching against outliers. The effective data cost for handling both outliers and radiometric condition is the crucial requirement for stereo matching application.

CT is a type of nonparametric local transform [9], which is defined as an ordered set of comparison of pixel intensities in a local neighborhood. For example, a local neighborhood$\eta(x)$ centered at x i.e. $x \in \eta(x)$. CT generates a bit string vector V, which represents a set of neighborhood having lesser intensities than the center pixel x i.e.,

$$V(x) = \begin{cases} 1 & if I(x) < I(y) \\ 0 & Otherwise \end{cases} \qquad (4)$$

where*I(x)* is the neighborhood pixels. The pixel intensities in a local neighborhood are always positive or zero. With this assumption the CT can be written as,

$$C(x) = \otimes_{y\in\eta}\ \zeta\big(I(x), I(y)\big) \qquad (5)$$

where*C(x)* is the census transformed bit string and ⊗ denotes concatenation operation. Every bit of *C(x)* will have same significance level and *C(x)* can be interpreted as the index of kernel structure defined on $\eta(x)$ with center pixel set to zero. Based on this interpretation, pixels in the kernel structure represents the outcome of a single census comparison ζ(.) at the corresponding location in the neighborhood.

In the original formulation of the CT given in Equation 5, it can compute only a subset of 2^8= 256 (out of 512) kernel structure defined on a 3×3 neighbourhood due to the centre pixel value set to zero. Hence, CT cannot produce exact kernel structure for local radiometric variation, since local radiometric variation effects mainly on a small area of an image [5]. It needs accurate representation of the local structure. In order to obtain all possible kernel structure, Kublbeck et al.[10] have proposed an extended version of the CT known as Modified Census Transform (MCT) which considers the mean intensity of the local neighborhood for comparison rather than the center pixel intensity itself.

Let $\eta'(x)$ be a local neighbourhood with center pixel at x. The mean intensity of the neighbourhood is given by *I'(x)*. Now, the original CT equation can be reformulated as

$$\Gamma(x) = \otimes_{y\in\eta'}\ \zeta\big(I'(x), I(y)\big) \qquad (6)$$

where*Γ(x)* is the MCT bit string. Equation 6 determines all possible kernel structure defined by the local neighborhood. Hence, it effectively captures any different local image structure caused by the radiometric variations as shown in Figure 4.

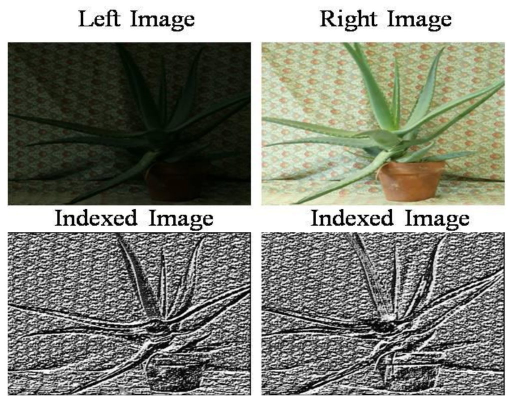

Fig. 4. Example for radiometric invariance property of the indexed image using MCT.

4.2 Correlation Computation Using NCC

A common approach for finding similarity in nonparametric method is finding its absolute or Hamming distance between the transformed bit strings. To improve the robustness to local radiometric variation, this work considers NCC [4] based correlation measure on the modified census transformed indexed image. NCC is a well-known traditional similarity measure that can compensate for camera gain and bias.

Let $I_L(p)$ and$I_R(p+d)$ are the corresponding pixels, where $I_L(p)$ is a value in the left image at a pixel p and $I_R(p+d)$

is a value in the right image at a pixel *(p+d)*. The similarity measure for the MCC data cost is given by,

$$MCC = \frac{\sum_{i=1}^{M}\left(I_L(p)\right).\left(I_R(p+d)\right)}{\sqrt[2]{\sum_{i=1}^{M}(I_L(p))^2 . \sum_{i=1}^{M}(I_R(p+d))^2}} \qquad (7)$$

where$M$ is an empirical value, which contains number of pixels in the neighborhood and d is the disparity range. The output of MCC generates correlation data cost.

5. GLOBAL ENERGY MODEL

MCC is a correlation measure that ranges from -1 to 1. In order to arrive at a non-negative cost between pixel p and $p+f_p$ in the left and right image respectively, *MCC* is subtracted from +1 and it is given by,

$$D_p(f_p) = 1 - (\exp(MCC))(7)$$

where*MCC* is the correlation data cost generated by modified transformed indexed image. Exponential function will increase the discriminability among the correlation values. Truncated quadratic cost based smoothness prior is used for pair wise smoothness and it is given by,

$$V_{pq}(f_p, f_q) = \lambda . \min\left(\left|f_p - f_q\right|^2, V_{max}\right)(8)$$

whereλ is the per pixel pair weight and $V_{max}$is thetruncation constant. By combining (7) and (8), total energy is defined by,

$$E(f) = \sum_{p} D_p(f_p) + \sum_{p} \sum_{q \in N(p)} V_{pq}(f_p, f_q) \qquad (9)$$

This total energy is optimized using Graph-cut (αexpansion) algorithm [8, 12] in order to find the disparity map.

6. RESULTS AND DISCUSSIONS

6.1 Experimental Settings

Aloe stereo image from Middlebury dataset [2] with different illumination and the exposure conditions are used. The experiment is conducted by considering $\lambda = 1.09$, $V_{max} = 0.19$, modified census window η' (9×7) and MCC correlation window M (9×9). A set of most popular stereo algorithms such as SAD with window size of 9×9, SSD with window size of 9×9, BT with default settings [11] , NCC with window size of 9×9 are optimized in the MRF framework using Graph-Cut(GC) with truncated quadratic smoothness cost and compared with proposed method.

6.2 Change in Light Source

To test the effect of light source (illumination), exposure index of the stereo image is fixed to 1 (constant) and the illumination index is varied from 1 to 3. Figure 5 shows the Aloe stereo image taken under extremely different illumination condition (left and right image have been taken at an index of illumination 1 and 3 respectively). Figure 5 also shows disparity maps of test stereo algorithms for Aloe stereo image. Change in light source will result in various local radiometric variations and it is one of the most difficult factors among radiometric effects. SAD, SSD and BT are very sensitive to this local variation and produces poor disparity. BT can produce very good results for ideal condition as well as for small variations in the contrast. NCC algorithm with raw matching cost is less dependent on the lighting condition than the SAD, SSD and BT algorithm but still more dependent than the proposed method. NCC algorithm suffers from fattening effect and produces large error near object boundaries and it assumes only the global affine transformed difference. Proposed method shows better disparity than the others. Change in scaling and the rotation will not effect as the stereo camera is calibrated in prior.

6.3 Change in Exposure

To test the effect of change in exposure, illumination index is fixed to 2 (constant) and the exposure index is varied from 0 to 2 (left and right images have been captured with an exposure index of 0 and 2 respectively). Figure 6 shows the left Aloe image, right Aloe image and the disparity maps of test stereo algorithms.

Change in exposure will create a global intensity variation between stereo images. SAD, SSD and BT algorithms are seriously affected due to these global variation and produces very dark or bright regions in the output disparity. It is observed that NCC is also not strong at global variation but it has maintained very less error as compared to SAD, SSD and BT algorithms. It is also observed that the proposed method remains stable and performs better with respect to exposure change. The change in exposure highlights the robustness of the developed algorithm.

7. CONCLUSION

This paper proposed an illumination invariant data cost known as Modified Census Correlation for stereo matching application. The evaluation section proves the robustness of the proposed technique for a variety of illumination, exposure and real-time conditions.

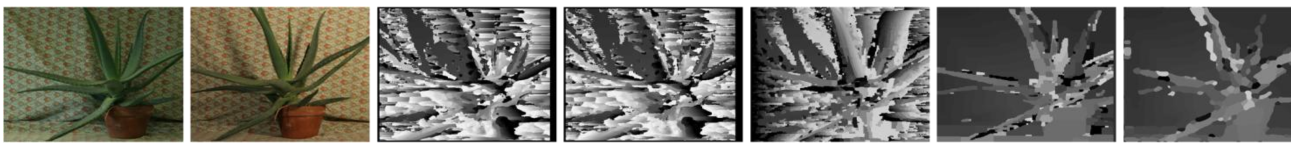

Fig. 5. Results of Aloe stereo image with varying illumination. From left to right: Left image with illumination (1)-exposure (1), Right image with illumination (3)-exposure (1), disparity of SAD, SSD, BT, NCC and proposed method.

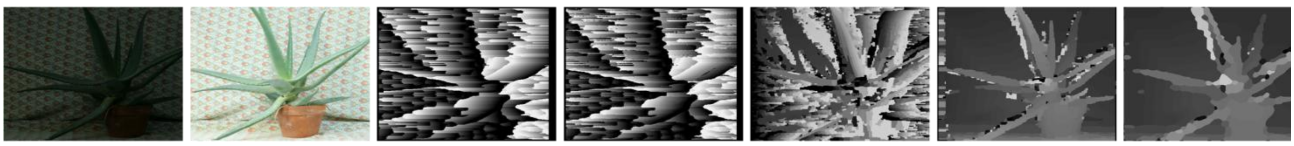

Fig. 6. Results of Aloe stereo image with varying exposure. From left to right: Left image with illumination (2)-exposure (0), Right image with illumination (2)-exposure (2), disparity of SAD, SSD, BT, NCC and proposed method.

ACKNOWLEDGEMENT
This research work is carried out under the scheme Structured Ph.D. of Manipal University

8. REFERENCES
[1] H. Hirschmuller and D. Scharstein, "Evaluation of stereo matching costs on images with radiometric differences," IEEETransactions on Pattern Analysis and Machine Intelligence,vol. 31, no. 9, pp. 1582–1599, 2009.

[2] D. Scharstein and R. Szeliski, "A taxonomy and evaluation ofdense two-frame stereo correspondence algorithms," Journal of Computer Vision, vol. 47, no. 1-3, pp. 7–42, 2001.

[3] Raghavendra U, KrishnamoorthiMakkithaya, Karunakar A. K,Qualitative and Quantitative Evaluation of Correlation Based StereoMatching Algorithms, Lecture Notes in Computer Science, SpringerBerlin/Heidelberg, International Conference on Advanced Computing,Networking and Security, pp. 242- 252, 16th Dec 2011, NITK,Surathkal.

[4] O. Faugeras, B. Hotz, H. Mathieu, T. Viville, Z. Zhang, P. Fua, E. Thron, and P. Robotvis, "Real time correlationbased stereo: Algorithm, implementations and applications," Technical Report RR-2013 INRIA, 1996.

[5] Y. S. Heo, K. M. Lee, and S. U. Lee, "Robust stereo matchingusing adaptive normalized cross-correlation", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 4, pp. 807–822, 2011.

[6] A. S. Ogale and Y. Aloimonos, "Robust contrast invariant stereo correspondence," in ICRA USA, 2004, pp. 819–824.

[7] Z. Xu, L. Ma, M. Kimachi, and M. Suwa, "Efficient contrastinvariant stereo correspondence using dynamic programmingwith vertical constraint," Vis. Comput., vol. 24, no. 1, pp. 45–55, Nov. 2007.

[8] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," IEEETransactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[9] R. Zabih and J. Woodfill, "Non-parametric local transforms for computing visual correspondence," in ECCV (2) Sweden,1994, pp. 151–158.

[10] C. Kublbeck and A. Ernst, "Face detection and tracking in video sequences using the modifiedcensus transformation," Image and Vision Computing, vol. 24, no. 6, pp. 564–572,2006.

[11] S. Birchfield and C. Tomasi, "A pixel dissimilarity measurethat is insensitive to image sampling," IEEE Transactions onPattern Analysis and Machine Intelligence, vol. 20, no. 4, pp.401–406, 1998.

[12] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A comparative study of energy minimization methods for markov random fields with smoothness-based priors,"IEEETransactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 6, pp. 1068-1080, 2008.