Second International Symposium on Computer Vision and the Internet (VisionNet’15)

Census Filtering Based Stereomatching Under Varying Radiometric Conditions

Sarika.S*, Deepambika V.A, M.Abdul Rahman

*Department of ECE, LBS Institute of Technology for Women, Trivandrum-695012, India
bDepartment of ECE, LBS Institute of Technology for Women, Trivandrum-695012, India
cKerala Technical University, Trivandrum, India

Abstract

Census transform is a non-parametric local transform which is based on the relative ordering of pixels instead of the pixel intensity values within a window; therefore it is insensitive to radiometric variations like exposure changes and illumination variations. This work focuses on an efficient census filtering based local stereomatching for images taken under varying radiometric conditions. For matching cost value computation hamming distance is taken for the dissimilarity measure. Finally for cost aggregation guided filtering is used. Guided filter has edge preserving property and runs faster than the bilateral filtering based cost aggregation. The experimental results shows that the proposed algorithm provides better disparities for images taken under varying radiometric conditions than the conventional stereomatching algorithm.

Keywords: Stereomatching; Radiometric variations; Census transform; Hamming distance; Guided filter

1. Introduction

Stereovision is the process of extraction of depth information from 2D images. Dense Stereomatching has many applications like object tracking, depth extraction, image reconstruction, unmanned vehicle navigation etc.

* Corresponding author. Tel.:+919895043054.
E-mail address: sarikaece01@gmail.com
The most challenging task of this field is the stereo correspondence which aims at finding the corresponding matching pixels between 2 images. Many stereocorrespondence algorithms have been developed in recent years for getting an accurate disparity map and satisfactory results have been obtained. Stereocorrespondence algorithm can be basically classified into two types, global and local methods\(^1\). Global methods can produce accurate disparity map comparable to local methods, but computationally it is very expensive and time consuming. On the other hand local methods achieve satisfactory results quickly and is very appropriate for real time applications. However this method does not produce accurate results in the textureless areas and near depth discontinuities and researches are still going on to solve these issues in local methods. Stereocorrespondence method consists of four steps\(^1\): (1) matching cost computation (2) cost aggregation (3) disparity computation/optimization (4) disparity refinement. Global methods do not involve step 2. Matching cost can be pixel based like absolute difference (AD), squared intensity difference (SD) etc. In local methods cost aggregation step involves the aggregation of matching cost over a window which depends on the intensity values of images and usually makes some implicit smoothness assumptions. Most of the local stereo algorithm works well only when the input image pairs are under similar radiometric conditions. It is based on the assumption that pixel intensity value in the left image and right image is the same. But in real world scenario this may not hold, there may be variations in the intensity values between the two images due to radiometric conditions. The radiometric variations include illumination variations and exposure changes. Due to these variations stereocorrespondence search becomes very difficult. Hirschmuller and Scharstein\(^2\) evaluate the performance of 15 different matching cost methods on images taken under different radiometric conditions and they have found that census filtering\(^3\) performs better in all conditions and is found to be more robust. Yufu Qu et al\(^4\) also have used census based approach to compensate for radiometric variations and have obtained satisfactory result. For cost aggregation to avoid edge fattening Veksler\(^5\) adopted an adaptive window based approach and Kang et al\(^6\) adopted a multiple window based method. Both these approaches improve the matching result to a certain extent but the size and shape of the window are restricted in these approaches. Yoon and Keon\(^7\) proposed an adaptive support weight aggregation which is actually bilateral filtering where the window size is fixed, but the pixels in the window have different weights depending on the colour proximity and similarity compared to the centre pixel and Gestalt law. Even though it provides good result near depth discontinuities the execution time is very long and hence not apt for real time applications. Guided filtering\(^8,\,9\) based cost aggregation proved to be the better performing local correspondence method. It is very fast compared to bilateral filtering and hence is suitable for real time applications.

In this paper, to compensate for radiometric variations, census filtering is performed as a pre-processing step. The matching cost computation is based on hamming distance measurement. For cost aggregation Guided filter weight is adopted. The disparity value is then computed using winner-take all optimization (WTA) and left-right consistency check. Finally median filtering is performed as a post processing step to get the final disparity map.

The rest of the paper is organized as follows, the proposed algorithm is described in section 2. In section 3 experimental results of the proposed algorithm is compared with conventional SAD based stereomatching algorithm. Section 4 provides the conclusion.

2. Proposed Method

In order to compensate for radiometric variations like exposure and illumination changes, first census filtering is carried out as a pre-processing step. Hamming distance is taken as the matching cost which is used to find out the pixel correspondence. For getting an accurate disparity map the cost value is aggregated over a window using guided filter. Finally the disparity value is selected by choosing the minimum aggregated cost value corresponding to every pixel using winner take all optimization (WTA).

2.1. Census transform

Census transform is a non-parametric local transform\(^3\) which depends on relative ordering of the pixel other than the intensity values. Therefore it preserves the local structure of an image even when there are radiometric variations and noises. Since the variations between the images do not affect the ordering of the pixels, the corresponding pixel value can be found out easily and an accurate disparity map can be obtained. Census transform depends only on the comparisons between the pixels and therefore they are invariant to changes in input bias and gain.
It maps the neighbouring pixels surrounding a pixel into a bit string depending on whether these pixels intensity value is greater or smaller than the pixel under consideration.

Consider the reference image, let \( I(q) \) be the intensity of the pixel point \( q(u,v) \) and \( W \) represent a 3x3 window surrounding \( q(u,v) \),if the intensity of the surrounding pixel \( q'(u,v) \) is greater than \( q(u,v) \) i.e. if \( I(q') \geq I(q) \) then \( n(q_i, q_i) = 1 \), otherwise \( n(q_i, q_i) = 0 \). Then the 8 bit string is given by:

\[
R(q) = \bigoplus_{w} (n(q_{i}, q_{i} + [u, v]))
\]  

(1)

To illustrate the way in which this transform works, consider a 3x3 window around a pixel

\[
\begin{array}{ccc}
32 & 112 & 34 \\
156 & 98 & 56 \\
32 & 72 & 98 \\
\end{array}
\]

\[
\begin{array}{ccc}
0 & 1 & 0 \\
1 & x & 0 \\
0 & 0 & 1 \\
\end{array}
\]

\[
\begin{array}{cccccccc}
0 & 1 & 0 & 0 & 1 & 0 & 0 & 1
\end{array}
\]

Fig1. Census transform

2.2. Matching cost computation

For census transformed images corresponding point is matched using hamming distance to extract the disparity value. Census filtering transforms the intensity difference between the pixels into 0 or 1 in a 1D vector form. Hamming distance measures the number of positions this vector differs. The cost value using hamming distance is given by:

\[
C_{\text{census}}(p, d) = H(R(q_l), R(q_r))
\]  

(2)

where \( H \) is the hamming distance between the Census transformed left image and right image

2.3. Cost aggregation using guided filter

In this step the matching cost value computed is aggregated over a window. In order to get satisfactory result near depth discontinuities and textureless region, matching cost is aggregated using guided filter. The advantage of using guided filter is that it has better edge preserving property and is very fast compared to the bilateral filtering. The aggregated cost value with guided filter is given by:

\[
C_{\text{agg}}(p, d) = \sum_{q} W_{p,q}(I) C_{\text{census}}(p, d)
\]  

(3)

where \( C_{\text{agg}}(p, d) \) is the filtered cost value of the pixel \( p \) for each slices of \( d \). \( W_{p,q}(I) \) is the weight of the guided filter which depends on \( I \), which is the reference image. Here the guidance image is the left image which is used to filter the guided image which is the \((x, y)\) slices of the cost volume. For a gray scale guidance image \( I \) the filter weight can be simply given by:

\[
W_{p,q}(I) = 1/|w|^{2} \sum_{w} (1 + \frac{(I_{p} - \mu_{I})(I_{q} - \mu_{I})}{\sigma_{I}^{2} + \epsilon})
\]  

(4)

where \( \mu_{I} \) and \( \sigma_{I}^{2} \) are mean and variance of the guidance image over the window \( w \) centred at pixel \( k \) with dimension \( \gamma \times \gamma \). \( \epsilon \) as the smoothness parameter.
For colour images filter weights can be given by:

\[
W_{p,q}(l) = \frac{1}{|w|^2} \sum_{k \in \mathcal{F}} (1 + \frac{(I_p - \mu_k)(I_q - \mu_k)^{-1}}{\sigma_k^2})^{-(1 + \frac{1}{2})}
\]  

(5)

where \(I_p, I_q\) are the 1x3 vectors, \(U\) is a 3x3 identity matrix and \(\varepsilon_k\) is the covariance matrix.

2.4. Disparity optimization

After the cost volume filtering, final disparity value for each pixel is fixed by selecting the minimum value from the aggregated cost value using winner take all optimization (WTA).

\[
d_p = \arg\min_{d \in D} (C_{agg}(p, d))
\]

(6)

D is the disparity search range.

2.5. Disparity refinement

In disparity refinement stage the disparity of the right image is computed in the same way as the left image (reference image). Then a left right consistency check is performed to detect the occluded areas and then the holes due to occlusion are filled by the disparities of the neighbouring pixels. Finally a median filter is applied as a post processing step to get the final dense disparity map.

3. Experimental Results

The performance of the algorithm is evaluated on Middlebury dataset. The experiments are conducted on both radiometrically clean and radiometric invariant images inorder to find out the robustness of the method to radiometric variations. The experiment is performed on the windows 8 operating system with Intel core i5 processor. The parameters used are \(\{\varepsilon, \gamma\} = \{0.0001, 9\}\). To quantitatively evaluate the performance of algorithm on radiometrically clean images, percentage of bad pixels is calculated in Non-occluded regions with error threshold greater than 1.

The test is conducted on standard stereo pair Cone, Teddy and Venus. The results are shown in Table 1. The data shows that proposed method is much better than the most commonly used conventional stereomatching algorithm SAD.

Table 1.Percentage of bad pixels for the Non-occluded pixels with error threshold>1

| Algorithm          | Venus | Teddy | Cone |
|--------------------|-------|-------|------|
| SAD                | 4.2   | 20    | 12   |
| Proposed algorithm | 1.69  | 9.58  | 6.04 |

Robustness of the method to radiometric variations is tested on image set (Aloe, Baby1). The image set used for the experiments is captured under 3 different exposure conditions (exp0, exp1, exp2) and three different illumination conditions (Illum1, Illum2, Illum3). The disparity maps obtained using SAD and proposed algorithm under varying exposure conditions with illumination condition fixed at Illum1 for Aloe and Baby1 are shown in fig(2) and fig(3) respectively. The result shows that the SAD algorithm is very sensitive to radiometric variations while the proposed method yields satisfactory result under varying radiometric conditions and the computational time for the proposed algorithm is 38s.
To quantitatively compare the performances of the algorithms, the RMS error of the computed disparity map against the ground truth disparity is calculated. Fig (4) and Fig (5) represent the RMS errors of the algorithms under varying exposure condition with illumination condition fixed at Illum1 and varying illumination condition with exposure condition fixed at exp2 respectively. From the result we can see that the proposed algorithm gives much smaller RMS error than the SAD algorithm.

Fig4.RMS Error (non-occluded pixels) under different exposure conditions. (a) Aloe image; (b) Baby1
4. Conclusion

A local stereomatching algorithm for images taken under varying radiometric conditions using census filtering is done. Since Census transform is based on the relative ordering of pixels in an image, it is insensitive to radiometric variations like exposure changes and illumination variations. For the census transformed image hamming distance is considered for the similarity measurement. For cost aggregation guided filtering is adopted. Experimental result gives good result for images under varying radiometric conditions and performs much better than the conventional SAD based stereomatching algorithm.

References

1. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereocorrespondence algorithm,” Int. J.Computer.Vision,vol. 47, pp. 7–42, Apr. 2002.
2. Heiko Hirschmuller And Daniel Scharstein“Evaluation Of Stereo Matching Costs On Images With Radiometric Differences” IEEE transactions on Pattern Analysis And Machine Intelligence August 2008
3. Ramin Zabih and John Woodfill,”Non–parametric Local Transforms for Computing visual Correspondence”, proceedings of European conference on Computer Vision, May 1994,pages 151-158
4. Yufu Qu1, Jixiang Jiang1, Xiangjin Deng2, Yanhong Zheng2” Robust local stereo matching under varying radiometric conditions” IET comput. vis., 2014, Vol. 8, Iss. 4, pp. 263–276
5. Veksler, O.; 'Fast variable window for stereo correspondence using integral images’. IEEE Conf. Computer Vision and Pattern Recognition, 2003, pp. 556–561
6. Kang, S.B., Szeliski, R., Chai, J.: ‘Handling occlusions in dense multi-view stereo’. IEEE Conf. Computer Vision and Pattern Recognition, 2001, pp. 103–110
7. Yoon, K.-J., Kweon, I.S.: ‘Adaptive support-weight approach for correspondence search’, IEEE Trans. Pattern Anal. Mach. Intell.,2006, 28, (4), pp. 650–656
8. K. He, J. Sun, and X. Tang, “Guided image filtering,” in ECCV, 2010.
9. C. Rhemann et al., “Fast Cost-Volume Filtering for Visual Correspondence and Beyond,” Proc. IEEE Int’l Conf. Computer Vision and Pattern Recognition, 2011, pp. 3017–3024.
10. Http://vision.middlebury.edu/stereo/