Stereo Matching with Cost Combination and Disparity Refinement

Deepa

Abstract: This paper presents an approach that use combined cost method and disparity refinement for stereo matching. Here disparity is obtained using the intensity, gradient, histogram of gradients for cost computation. Later disparity refinement is performed using Weiner filter. The correctness of matching is enhanced in the proposed method. The combined use of cost measure ensures the effective evaluation. The proposed approach produces good quality disparity map.

Keywords : Disparity, Gradient, Weiner..

I. INTRODUCTION

Stereo matching finds its application in 3D reconstruction. It is one of major topic that is researched in computer vision [1]. Stereo matching uses camera which contains two or sometimes more lenses. It generates images that are similar to human vision and find 3D position of objects. The gap between the lenses is same as the gap that exists between the human eyes. Vision lets us to perform certain basic tasks such as inferring shape and depth of the objects by finding the differences of the projection on our two eyes. Apart from these simple capabilities our eyes can also perform complex tasks such as identifying and also interpreting the context of our natural surroundings. Stereo matching tries to obtain 3D information from 2D images. The stereo matching aims at computing the disparity at each individual pixel. The change in positions of pixels between matching pixels is known as disparity. Using camera parameters depth of the pixels can be calculated. Apart from stereo matching, several other methods are available for finding such as laser scanning [2]. These methods need a large setup. Stereo matching is challenging because of poor or repetitive textures, illumination differences. Further, quick calculations are needed for real-time applications. Stereo methods are classified into local approaches and global approaches [3]. Local methods tend to be simple but less accurate. Global method use energy function to find disparity map.

The proposed method uses a novel technique by combining the measure of intensity, gradient, histogram of gradients (HOG) for the disparity computation. Disparity refinement is performed using wiener filter. The resultant disparity map shows an improvement over the conventional method.

The rest part of the paper is summarized below. Section II includes literature survey. Methodology is given in III section. The results obtained are demonstrated in the section IV. Section V shows conclusion.

Revised Manuscript Received on January 5, 2020

Deepa, Department of IS&E, N.M.A.M. Institute of Technology, Nitte, Karkala, Karnataka, India. deepashetty17@nitte.edu.in

II. LITERATURE SURVEY

Several stereo matching algorithms are proposed in the past. The local methods generally use certain measures such Sum of Squared Differences (SSD). Apart from this measure the measure such as Birch field Tomasi is proposed. The correctness of local methods depends on correct widow size selection. [4] Used variable sized windows. [5] Proposed multiple windows to improve accuracy at depth discontinuities. Adaptive weight [6] is used recently to obtain good disparity map. These are computationally expensive. Segmentation based stereo matching is proposed by [7, 8].Their performance depends on accuracy of segmentation. Fast stereo matching method based on area is proposed by L. D. Stefano [9]. Here searching is performed in single direction. Hence it is called as Single Matching Phase (SMP). It discards previous matches when better result is found. It uses SAD technique for error calculation. This method generates a dense disparity map. The method used in [10] proposes zero mean normalized cross correlation. It uses neural model .Correct window size and shapes are selected by the neural network for each considering region. The results obtained by the network are better but they are expensive in terms of computational power. So it is not applicable for real time applications . Matching based on shape is proposed in [11]. It tries to show the relevance of the horizontal and slanted surfaces. Here the researchers replaced the uniqueness constraint which refers to pixels by uniqueness constraint which refers to line segments along the given scan line. It calculates interval matching rather than pixel matching. It uses absolute intensity difference for the matching factor. Several global algorithms are proposed in the fast like belief propagation [12], dynamic programming, graph cut[13]. Intensity is the common measure used for cost computation, but they fail to perform at the object boundaries. To overcome this problem authors in [14] used gradient information. Weighted sum is used in [15].Here the gradient information may be susceptible to noise. The proposed method uses combined cost measure

III. METHODOLOGY

The steps followed in the paper are given in Fig. 1.

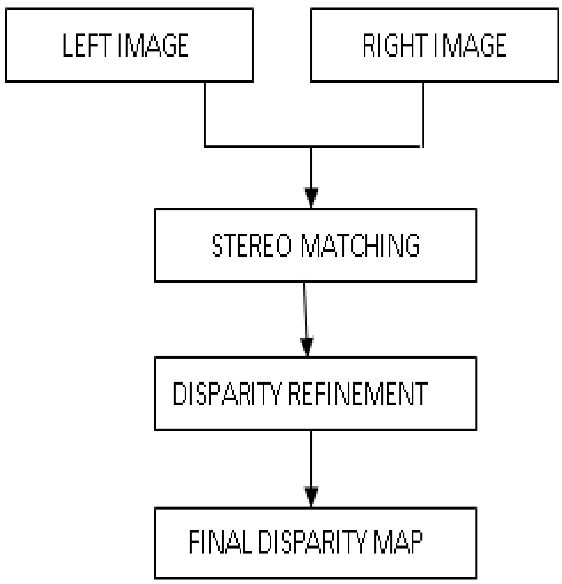

Fig.1. Steps of proposed approach

A. Stereo Matching

Stereo matching that are local based uses cost computation for matching. The matching cost is found for all pixels. Typically matching cost between corresponding pixels are found based on intensity differences. Color intensity is very sensitive to illumination differences .Image gradient preserves structural details present in image. Hence it is less sensitive to change in illumination. Gradient of images are obtained using convolving filter. Sobel filter is used in the proposed approach. Gradient images in x and y direction is calculated using the following equation.

$$\Delta f = \begin{bmatrix} Gx \\ Gy \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (1)$$

Here $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ represent gradient in x, y direction respectively.

Gradient information is sensitive to noise. To overcome this issue HOG descriptor is considered in calculating disparity map. HOG is invariant to photometric transformations. This calculates the total count of occurrences of orientation of gradient in a part of image. The image is arranged as tiny connected regions which are known as cells. For each and every connected region histogram of unsigned gradient orientation is calculated. All the histograms of connected regions are combined to form the HOG descriptor.

The cost based on intensity, gradient, HOG is calculated as follows

$$CI(x,y) = \sum_{(x,y)\in W}^{N} \left| Il(x,y) - Ir(x-d,y) \right|^2 \quad (2)$$

$$CG(x,y) = \sum_{(x,y)\in W}^{N} \left| Gl(x,y) - Gr(x-d,y) \right|^2 \quad (3)$$

$$CH(x,y) = \sum_{(x,y)\in W}^{N} \left| Hl(x,y) - Hr(x-d,y) \right|^2 \quad (4)$$

Here $Il(x,y)$, $Ir(x-d,y)$ is the intensity values, $Gl(x,y)$, $Gr(x-d,y)$ is gradient information . $Hl(x,y)$, $Hr(x-d,y)$ are the HOG values of pixel (x, y), d denotes value of disparity. W is the window .

The minimum difference over the window gives the best matching pixel. This pixel location gives the optimal disparity. The disparity map using intensity, gradient and HOG feature is found. The disparity map is obtained by combining the weighted sum of these three disparity map.

$$d = (\propto * d_1) + (\beta * d_2) + (\gamma * d_3) \quad (5)$$

Here d_1, d_2, d_3 are the disparity map obtained based on intensity, gradient and HOG respectively. $\propto + \beta + \gamma = 1$. $\propto = 0.4$, $\beta = 0.4$ and $\gamma = 0.2$.

B. Disparity refinement

Disparity refinement is performed with the help of wiener filter. This filter is applied in frequency domain. It can filter the noise from the corrupted image. It tries to minimize the mean square error by keeping the variation between the original and reconstructed image as low as possible. It helps in deblurring the image. The commonly used filter to restore corrupted image is inverse filter. But the inverse filters are sensitive to additional noise. Wiener filter removes noise and helps in deblurring at the same time.

IV. RESULTS AND DISCUSSIONS

The proposed system uses images taken from Middlebury dataset as the input .Here rectified images are considered to evaluate the performance. The images like Cloth, Baby, and Flowerpot were used. The performance is evaluated using qualitative and quantitative measure. The Qualitative results are depicted as below

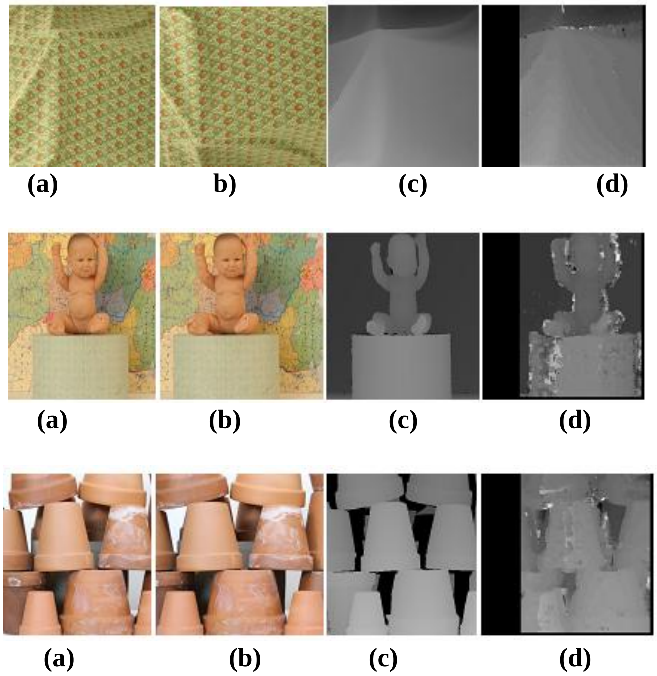

Fig. 2 Results for images Cloth, Baby, Flowerpot . (a) Left image, (b) Right image (c) Ground truth image (d) Disparity Map

Root Mean Square Error (RMSE) is used for quantitative measure. RMSE is computed between the ground truth and disparity map. Lesser the resultant value of RMSE better is the quality of the image. The following equation is used to calculate RMSE.

$$RMSE = \left(\frac{1}{M * N} \sum \sum \left(dc\,(x,y) - dg\,(x,y) \right)^2 \right)^{1/2} \quad (6)$$

Here dc is disparity map generated and dg is the ground truth image. M represents rows and N represents columns present in the image.

Table I gives the value of RMSE of disparity map obtained by the proposed system and disparity map using gradient. The RMSE value of the disparity map of proposed approach is lesser than that of the disparity map using gradient method which ensures better performance.

TABLE I RMSE values

Input Image	Gradient method	Proposed method
Cloth	9.95	8.91
Baby	11.09	10.51
Flowerpot	13.56	13.40

V. CONCLUSION

In this paper, disparity map is generated using the combined cost measure of intensity, gradient and HOG features. Disparity refinement is performed using wiener filter. The proposed system is tested using the images from Middlebury dataset. The results obtained shows that this method produces better quality disparity map.

REFERENCES

1. D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light", *Proceedings IEEE Computing Society,* Conf. CVPR, vol. 1, pp. I-195-202, 2003.
2. Z. Lei, G. Xianwei, Z. Cheng, and D. Xiuze, "Research on the technology of extracting 3D face feature points on basis of binocular vision", in *Proceedings. 2nd International CISP*, pp. 1-4, Oct. 2009.
3. D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence", *International Journal Comput.vis,* vol. 47, nos 1-3, pp. 7-42, 2002.
4. Y. Boykov, O. Veksler, and R. I. Zabi, "A variable window approach to early vision", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1283–1294, 1998.
5. H. Hirschmuller, P. R. Innocent, and J. Garibaldi, "Real-time correlation-based stereo vision with reduced border errors," *International Journal of Computer Vision*, vol. 47, no. 1–3, pp.229–246, 2002.
6. K.-J. Yoon and I. S. Kweon, "Adaptive support-weight approach for correspondence search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 650–656, 2006.
7. M. Bleyer, C. Rother, P. Kohli, D. Scharstein, and S. Sinha, "Object stereo Joint stereo matching and object segmentation", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* , pp. 3081–3088, Washington, DC, USA, 2011.
8. A. Klaus, M. Sormann, and K. Karner, "Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure," *Proceedings of the 18th International Conference on Pattern Recognition* , pp. 15–18, Hong Kong, 2006.
9. Lugi Di Stefano, Massimiliano Marchionni, Stefano Mattoccia, "A fast Area Based Stereo Matching Algorithm", *Image and vision Computing* 22, pp. 983-1005, 2004.
10. Elisabetta Binaghi, Ignazio Gallo, Giuseppe Marino, Mario Raspanti, "Neural adaptive stereo matching", *Pattern Recognition Letters 25*, pp. 1743-1758, 2004.
11. Abijit S. Ogale and Yiannis Aloimonos, "Shape and the Stereo Correspondence Problem", *IJCV* 65, 3, pp. 147-167, 2005.
12. J. Sun, N. N. Zheng, H. Y. Shum, "Stereo matching using belief propagation", *In Proc. IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 25, pp 787-800, July 2003.
13. Y. Boykov, O. Veksler, R. Zabih, "Fast approximate energy minimization via graph cuts", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222-1239, Nov 2001.
14. L. De-Maeztu, A. Villanueva, R. Cabeza , "Stereo matching using gradient similarity and locally adaptive support-weight", *Pattern Recogniton* ,pp1643–1651,2011.
15. Cheng-Tao Zhu, Yau-Zen Chang, Huai-Ming Wang, Kai He, Shih-Tseng Lee, and Chung-Fu Lee, "Efficient Stereo Matching with Decoupled Dissimilarity Measure Using Successive Weighted Summation," *Mathematical Problems in Engineering*, vol. 2014, Article ID 127284, 9 pages, 2014.

AUTHORS PROFILE

Ms. Deepa, received her B.E degree in Computer Science and Engineering from V.T.U. Belgaum in the year 2005 and M. Tech degree in Computer Science & Engineering in the year 2012. She is currently a research scholar at V.T.U. Belagum. Her research interests are image processing, depth estimation and computer vision.