Improved Stereo Matching With Boosting Method

Shiny B, Dr. Deepa J.

Abstract: This paper presents an approach based on classification for improving the accuracy of stereo matching methods. We propose this method for occlusion handling. This work employs classification of pixels for finding the erroneous disparity values. Due to the wide applications of disparity map in 3D television, medical imaging, etc, the accuracy of disparity map has high significance. An initial disparity map is obtained using local or global stereo matching methods from the input stereo image pair. The various features for classification are computed from the input stereo image pair and the obtained disparity map. Then the computed feature vector is used for classification of pixels by using GentleBoost as the classification method. The erroneous disparity values in the disparity map found by classification are corrected through a completion stage or filling stage. A performance evaluation of stereo matching using AdaBoostM1, RUSBoost, Neural networks and GentleBoost is performed.

Index Terms: Stereo matching, occlusion handling, initial processing, classification, completion stage, median filtering, performance evaluation

1 Introduction

STEREO MATCHING is an important research topic in computer vision and image processing. This finds the correspondences of pixels between two stereo images i.e., left image and right image that are captured by using stereo camera which are set at a distance usually equal to the distance between the human eyes. So, stereo matching simulates human binocular vision and it estimates the 3D positions. The lenses used for capturing images may be parallel or converging lenses. Two or more lenses can be used for capturing images. For obtaining the disparity map, the stereo correspondence can be used. That is the significance of stereo matching problem in 3D reconstruction. The disparity map finds application in medical imaging, 3D television machine vision, robot navigation and control etc. The two images captured will be slightly different since they are taken from different viewpoints. Most of the pixels in the left image have a corresponding pixel in the right image. So the disparity is found by finding the difference between the coordinate positions of the corresponding pixels. Situations may occur such as some pixels in the left image may not be present in the right image due to varying depths or occlusion. There are also problems with color inconsistency, discrete nature of images etc., which affects exact matching of images. This work handles occlusion by using classification. The stereo matching methods used are Normalized Cross Correlation (NCC) and Sum of Absolute Distances (SAD). Even though they are local methods, they are faster which makes them suitable for real time applications. Supervised classification is used in which classes are known. Fig. 1 shows the Middlebury [1] evaluation images which are used to check the accuracy of the stereo matching algorithms.

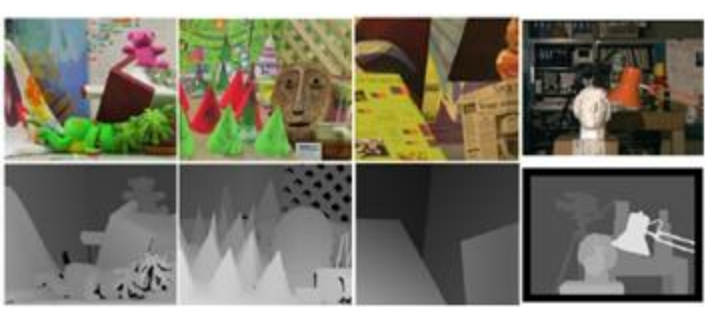


***Fig. 1** Middlebury evaluation images (Top row), from left to right, Teddy, Cones, Venus, Tsukuba [1], Corresponding disparity map (bottom row).*

In essence, classification can and is proven to combine a certain set of measurements for the purpose of improving stereo matching proficiently [1]

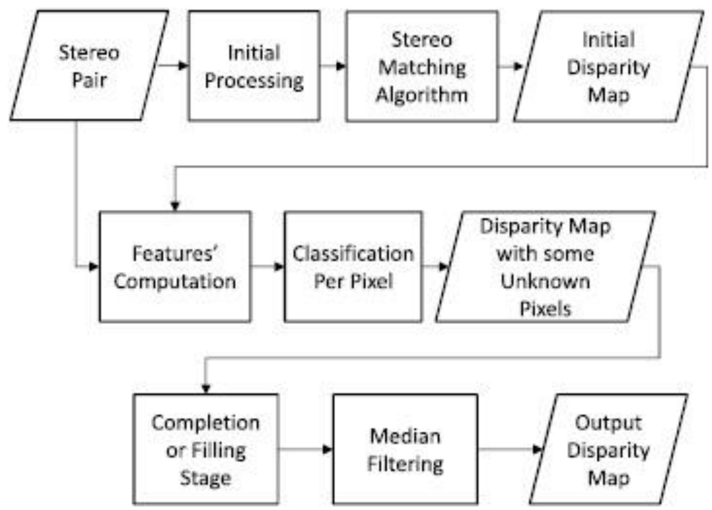


***Fig. 2** Flowchart showing the stages of the proposed system*

The suggested system follows the steps shown in the flowchart of Fig. 2 [1]. The method starts by performing initial processing to the stereo pair. Then, a stereo matching algorithm is performed to obtain the required disparity map. After that, the computed map along with the input images are used to compute the chosen features for every pixel. The pixels are then classified as acceptable or not which leads to a disparity map with unknown values. The unidentified values are then filled using a completion stage followed by median filtering to produce the final disparity map [1].

2 Related Works

2.1 Review Stage

There are many methods to tackle the stereo matching problem. In general, the methods are divided into three types; local, global, and semi-global ones that combine the other two. The local methods usually implement a measure of the difference between the pixels and are generally the fastest. The three basic local methods are the Sum of Absolute Differences (SAD), the Sum of Squared Differences (SSD), and the Normalized Cross Correlation (NCC) [1]. [1] used classification with these local stereo matching methods to enhance the accuracy of stereo matching method. The classification methods include AdaBoostM1, RUSBoost and Neural Networks. In addition to this, a completion or filling stage is also used which further improves the output disparity map. There are other methods in addition to these local methods such as Birchfield Tomasi (BT) measure in [2].Also, adaptive windows are used instead of fixed windows such as SAD [3] and adaptive support weights (ASW) [4]. The global stereo matching methods aren't suitable for realtime applications since it demands more time. These methods consider the pixel's disparity with the neighboring ones [1] . These methods rely on graph theory [5], cooperative optimization [6], [7], nonlinear optimization techniques [8], [9], Markov Chain based methods [10] and Belief Propagation (BP) as in [11] and [12]. The semi-global approaches use mutual information as in [13] and in [14] implemented a new distinctive similarity measure by combining local and global methods. The two key properties of a stereo matching method are accuracy and speed with constantly updated research dealing with these aspects [1] .In [15], the accuracy was improved by accounting for the edge points which can be problematic due to disparity alterations [1]. The main problems that reduce accuracy in stereo matching are occlusion and illumination differences. In regards to occlusion, there are many works that concentrated on it like [13] and [16]. Also, [17] dealt with large occlusion stereo, and [18] used disparity alterations and the uniqueness properties of different pixels for occlusion handling [1]. All of the previous methods relied on LRC while in this work; we handle occlusion through classification using GentleBoost as the classification method. Concerning the illumination issues, these are addressed here through histogram information according to [19]. The work in [19] further summarizes some of the illumination related work in stereo matching [1].

3 Proposed Improved Stereo Matching with Boosting Method

3.1 Initial Processing

At first, images are preprocessed to reduce the illumination variances. Then stereo matching is applied on the processed images [1]. The preprocessing method is based on the work in [19]. This basically involves finding the histograms of each stereo image. Then, a set of peaks of the two histograms are found. After that, the absolute differences between the corresponding peaks of each histogram are found. If the median difference is relatively high (≥ 2 in a 0-255 quantized image), histogram matching, or specification, is applied to match the image having the higher peak with the other. This leads to a slightly modified pair that are better matched than the original ones [19],[1].

3.2 Stereo Matching

In this work, local stereo matching methods are used. The methods include Normalized Cross Correlation (NCC), Sum of Absolute Differences (SAD) and Sum of Squared Differences (SSD). These can be defined by the following equations:

$$\mathrm{SADx,\ y} = \sum_{i=x-Rad}^{x+Rad} \sum_{j=y-Rad}^{y+Rad} \left| L(i,j) - R(i,j) \right| \tag{1}$$

$$\mathrm{SSDx,\ y} = \sum_{i=x-Rad}^{x+Rad} \sum_{j=y-Rad}^{y+Rad} \left(L(i,j) - R(i,j) \right)^2 \tag{2}$$

$$\mathrm{NCCx,\ y} = \frac{1}{n-1} \sum_{i=x-Rad}^{x+Rad} \sum_{j=y-Rad}^{y+Rad} \frac{\left(L(i,j)-\bar{L}\right)}{\sigma_L} \frac{\left(R(i,j)-\bar{R}\right)^2}{\sigma_R} \tag{3}$$

3.3 Feature Computation

Various features are computed from the initially obtained disparity map and the input stereo image pairs. The different features used are mean intensity difference, mean edge disparity value, number of similar disparities, mean disparity difference, disparity value, candidate accurate disparity, histogram disparity value difference, regional histogram disparity value differences and local histogram disparity value differences [1]. They are defined by the following equations:

$$f1 = \frac{1}{n} \sum_{i=x-Rad}^{x+Rad} \sum_{j=y-Rad}^{y+Rad} \left| I(i,j) - I(x,y) \right| \tag{4}$$

$$f2 = \frac{1}{n} \sum_{i=x-Rad}^{x+Rad} \sum_{j=y-Rad}^{y+Rad} E_D(i,j) \tag{5}$$

$$f3 = length\left(\sum_{i=x-Rad*2}^{x+Rad*2} \sum_{j=y-Rad*2}^{y+Rad*2} \left| D(i,j) - D(x,y) \right| \rangle T_D \right)$$

$$Condition: \left| I(i,j) - I(x,y) \right| \le i_D \tag{6}$$

$$f4 = \frac{1}{n} \sum_{i=x-Rad*2}^{x+Rad*2} \sum_{j=y-Rad*2}^{y+Rad*2} \left| D(i,j) - D(x,y) \right| \tag{7}$$

$$f5 = D(x,y) \tag{8}$$

$$f6 = 1;if\ D_A(x,y) exists;0 otherwise \tag{9}$$

$$f7 = \left| H_{DA}(I(x,y)) - D(x,y) \right| \tag{10}$$

$$f8 = \left| Hi_{DA}(I(x,y)) - D(x,y) \right| \tag{11}$$

$$f9 = \left| hi_{DA}(I(x,y)) - D(x,y) \right| \tag{12}$$

3.4 Classification method

As this work focuses on improving stereo matching with classification, GentleBoost has been chosen as the classification method. GentleBoost or Gentle Adaboost is a variant of real Adaboost algorithm. Since this work uses binary classification, GentleBoost is chosen as the classification method. The difference between Adaboost and GentleBoost is in the updation of weights The learning starts with a set of N samples with equal weights, w_i = 1/N and a single weak learner. Then the learner is trained and the error rate is calculated with respect to the weights, w_i. This is repeated for a number of iterations, t. The weights of true pattern is updated and renormalized using the equation

$$w_i = w_i \cdot e^{-h_t(x_i)y_i} \quad (13)$$

where $h_t(x_i)$ is the weighted return value of the weak classifier for x_i. After 't' iterations, the weak classifiers are combined to obtain a strong classifier to perform the classification.

3.5 Completion Or Filling Stage

This stage is used to fill the unknown disparity values in the disparity map obtained after classification. The unknown values are filled by combining two techniques [1]. The first technique quantizes the image to obtain a new image (Iq). Then, each connected empty disparity region is treated as a single region. An empty region is a region where the disparities are currently unknown since they were found erroneous. Then in the empty region, assign the unknown disparities with the median disparity of similar known values in the vicinity of the region. The vicinity of the region is every known disparity pixel that neighbors the region. The obtained map is termed (D1). The second technique (Cb) is to fill in the unknown disparity with the nearest spatially known disparity along the same row to yield the map (D2) [20],[21],[22],[1] These two techniques are then combined meaning that if the two disparity values have an absolute difference ≤TD, they are averaged, while otherwise the value is set to unknown again. This validation approach leaves a set of unknown pixels which are now filled by the first process unless no similar known value exists. These unknown values are filled with (Cb) to yield the map (D3) [1]. An idea that enhanced the results is to special treat the pixels that lie before 60% of the disparity range. These disparities are set to the maximum value between that of (D3) and (D2) leading to the final map [1].

4 Results

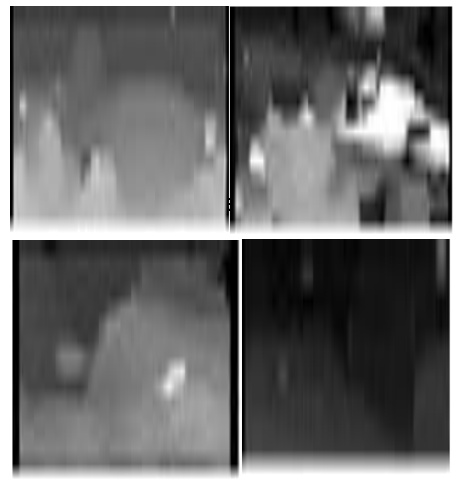
The results are shown for the Middlebury evaluation set [1]. Since the work relies on classification for enhancing the output disparity map of the stereo matching methods, a training dataset is created with images from the Middlebury evaluation set. The classification method used is GentleBoost. The results tabulated here are the average of different runs for different image sizes and window sizes. Also, the figures shown are of a random case. For calculating errors and comparing the obtained disparity values with the true disparities, two measures are used [1]. The first measure used is percentage error threshold which finds the percentage of pixels whose disparity differs from the true disparity by a certain difference threshold according to equation (14):

$$\%error = \frac{\#\,pixelswith\left(\left|disp - truedisp\right\rangle thresh\right)}{\#\,allpixels} * 100 \quad (14)$$

The other measure used is mean absolute error (MAE) which is not governed by a threshold.

$$MAE = \frac{1}{\#\,allpixels} \sum_{allpixels} \left|disp - truedisp\right| \quad (15)$$

Fig. 3 shows the obtained disparity map using NCC and GentleBoost and Fig. 4 shows the obtained disparity map using SAD and GentleBoost. Table 1 shows the results on Middlebury test images measured by percentage error threshold and Table 2 shows the results measured by mean absolute error. The values shown for AdaboostM1, RUSBoost and Neural networks are from [1].



(a)(c)
(b)(d)

Fig. 3 *Disparity maps obtained by using NCC+proposed approach (a)Cones(b) Tsukuba(c)Teddy(d)Venus*

(a)(c)
(b)(d)

Fig. 4 *Disparity maps obtained by using SAD+proposed approach (a)Cones(b) Tsukuba(c)Teddy(d)Venus*

TABLE 1
RESULTS ON TEST IMAGES MEASURED BY PERCENTAGE ERROR THRESHOLD AFTER COMPLETION STAGE

Error / Method	Image	Teddy	Cones	Tsukuba	Venus
NCC+GentleBoost		5.02	5.5	3.5	1.2
SAD+GentleBoost		4.6	4.6	3.7	1.6
NCC+AdaboostM1		11.6	12.5	5.9	2.7
SAD+AdaboostM1		11.5	12.5	5.2	2.6
NCC+RUSBoost		10.8	12.7	5.5	2.3
SAD+RUSBoost		10.9	12.2	5.4	2.4
NCC+NN		11.3	11.8	4.9	2.6
SAD+NN		11.9	11.9	4.9	2.7

All the images were initially processed and the threshold is equal to 2 or 1 according to the disparity range of image. The classification was performed using GentleBoost algorithm to find the erroneous disparity values.

TABLE 2
RESULTS ON TEST IMAGES MEASURED BY MEAN ABSOLUTE ERROR AFTER COMPLETION STAGE

MAE / Method	Image	Teddy	Cones	Tsukuba	Venus
NCC+GentleBoost		0.81	0.88	0.31	0.24
SAD+GentleBoost		0.74	0.87	0.2	0.28
NCC+AdaboostM1		1.13	1.22	0.39	0.38
SAD+AdaboostM1		1.09	1.19	0.41	0.36
NCC+RUSBoost		1.06	1.19	0.41	0.36
SAD+RUSBoost		1.02	1.18	0.42	0.35
NCC+NN		1.05	1.17	0.38	0.37
SAD+NN		11.9	11.9	4.9	2.7

5 Conclusion

This work used classification based approach to enhance the output disparity map obtained from the local, global or semi-global stereo matching methods. One advantage of this work is that classification can be combined with any of the stereo matching methods. An initial processing stage is used to reduce the illumination variances in the input stereo images so that they are further enhanced. The completion or filling stage after classification further enhances the output disparity map by filling in the erroneous disparity values.

Acknowledgment

The authors wish to thank M.Baydoun, M.A.Al-Alaoui .

References

[1] M.Baydoun, M.A.Al-Alaoui,"Enhancing stereo matching with classification", IEEE Access, Digital Object Identifier 10.1109/ACCESS.2014.2322101.

[2] S. Birchfield and C. Tomasi, ``A pixel dissimilarity measure that is insensitive to image sampling,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 20,no. 4, pp. 401-406, Apr. 1998.

[3] T. Kanade and M. Okutomi, ``A stereo matching algorithm with an adaptive window: Theory and experiment,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 16, no. 9, pp. 920-932, Sep. 1994.

[4] K.-J. Yoon and I.-S. Kweon, ``Locally adaptive support-weight approach for visual correspondence search,'' in Proc. IEEE Comput. Soc. Conf. CVPR, vol. 2. Jun. 2005, pp. 924-931.

[5] V. Vineet and P. J. Narayanan, ``CUDA cuts: Fast graph cuts on the GPU,''in Proc. IEEE Comput. Soc. Conf. CVPRW Jun. 2008, pp. 1-8.

[6] M.-H. Ju and H.-B. Kang, ``A new method for stereo matching using pixel cooperative optimization,'' in Proc. 16th IEEE ICIP, Nov. 2009, pp. 2105-2108.

[7] Z.-F. Wang and Z.-G. Zheng, ``A region based stereo matching algorithm using cooperative optimization,'' in Proc. IEEE Conf. CVPR, Jun. 2008,pp. 1-8.

[8] J. Y. Goulermas, P. Liatsis, and T. Fernando, ``A constrained nonlinear energy minimization framework for the regularization of the stereo correspondence problem,'' IEEE Trans. Circuits Syst. Video Technol., vol. 15,no. 4, pp. 550-565, Apr. 2005.

[9] E. Park and K. Wohn, ``Stereo and motion correspondences using nonlinear optimization method,'' Comput. Vis. Image Understand., vol. 101, no. 3, pp. 194-203, 2006.

[10] W. Kim, J. Park, and K. Lee, ``Stereo matching using population-based MCMC,'' Int. J. Comput. Vis., vol. 83, no. 2, pp. 195-209, 2009.

[11] P. F. Felzenszwalb and D. P. Huttenlocher, ``Ef cient belief propagation for early vision,'' Int. J. Comput. Vis., vol. 70, no. 1, pp. 41-54, 2006.

[12] Q. Yang, L. Wang, and N. Ahuja, ``A constant-space belief propagation algorithm for stereo matching,'' in Proc. IEEE Conf. CVPR, Jun. 2010, pp.1458-1465.

[13] I. Ernst and H. Hirschmüller, ``Mutual information based semi-global stereo matching on the GPU,'' in Proc. 4th Int. Symp. Adv. Vis. Comput., 2008, pp. 228-239.

[14] K.-J. Yoon and I. S. Kweon, ``Stereo matching with the distinctive similarity measure,'' in Proc. IEEE ICCV, Oct. 2007, pp. 1-7.

[15] J. Gibson and O. Marques, ``Stereo depth with a uni ed architecture GPU,'' in Proc. IEEE Comput. Soc. Conf. CVPRW, Jun. 2008, pp. 1-6.

[16] J. Sun, Y. Li, S. B. Kang, and H.-Y. Shum, ``Symmetric stereo matching for occlusion handling,'' in Proc. IEEE

Comput. Soc. Conf. CVPR, vol. 2. Jun. 2005, pp. 399-406.

[17] A. F. Bobick and S. S. Intille, ``Large occlusion stereo,'' Int. J. Comput. Vis., vol. 33, no. 3, pp. 181-200, 1999.

[18] C. L. Zitnick and T. Kanade, ``A cooperative algorithm for stereo matching and occlusion detection,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 7, pp. 675-684, Jul. 2000.

[19] M. Baydoun and M. A. Al-Alaoui, ``Enhancing stereo matching with varying illumination through histogram information and normalized cross correlation,'' in Proc. 20th IWSSIP, Jul. 2013, pp. 5-9.

[20] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, and M. Gelautz, ``Fast cost-volume filtering for visual correspondence and beyond,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 2, pp. 504-511, Feb. 2013.

[21] A. Hosni, M. Bleyer, C. Rhemann, M. Gelautz, and C. Rother,``Real-time local stereo matching using guided image filtering,'' in Proc. IEEE Int. Conf. ICME, Jul. 2011, pp. 1-6.

[22] W.-H. Lee, Y. Kim, and J. B. Ra, ``Efficient stereo matching based on a new confidence metric,'' in Proc. 20th EUSIPCO, 2012, pp. 1139-1143.

[23] D. Scharstein and R. Szeliski, ``High-accuracy stereo depth maps using structured light,'' in Proc. IEEE Comput. Soc. Conf. CVPR, vol. 1. Jun. 2003, pp. I-195-I-202.