STEREO MATCHING ALGORITHM BASED ON CURVELET DECOMPOSITION AND MODIFIED SUPPORT WEIGHTS

Dibyendu Mukherjee, Guanghui Wang, Q.M.Jonathan Wu

Department of Electrical and Computer Engineering,
University of Windsor,
401 Sunset Avenue, Windsor, Ontario, Canada N9B 3P4

ABSTRACT

We present a novel multiresolution analysis based stereo matching method using curvelets and modified adaptive support weight. Multiresolution analysis has long been applied to stereo correspondence. However, previous methods suffer from false matches arising from textureless region or repetitive textures and fattening effect due to area based matching. In the proposed approach, we have reduced false matches by using curvelet coefficients in different scales and orientations. Curvelet coefficients can uniquely represent different image points and increase matching accuracy. The fattening effect is reduced using support weights modified for curvelets. The proposed method is verified and compared with state-of-the art methods by extensive tests, and good results are obtained.

Index Terms— Stereo correspondence, multiresolution analysis, curvelet decomposition, support weights

1. INTRODUCTION

Stereo reconstruction is one of the most popular techniques for estimation of depth from digital images. It has two steps - Correspondence matching and reconstruction. Stereo correspondence is a process of matching the similar points in the stereo image pairs. The apparent shift of position between two corresponding image points in two images is called disparity. There are different approaches for correspondence search. Local area based approaches rely on some statistical correlation of color or intensity values. They estimate the disparity of a pixel by correlating a support window around the pixel with a similar support window in the other image. Area based methods assume all pixels in the support window have similar depth. This assumption is violated in depth discontinuities and results in fattening effect near these regions. The approaches select the best correlation matches for each pixel. This sometimes results in wrong disparity estimation for image points having ambiguity of depth. As a consequence, textureless regions, repetitive textured regions and regions with depth discontinuity fail to match correctly. The popular correlation functions like, Sum-Absolute-Difference, Sum-Squared-Difference and Normalized-Cross-Correlation all suffer from the problems. Several solutions have been proposed out of which support weights [1] is one of the most popular approach.

For improving correlation search, multiresolution methods also been used before. Wavelets have been proposed [2, 3] for multiresolution analysis of images. Unfortunately, the wavelet coefficients have strong dependencies across scales and subbands. Also, wavelets are shift-variant and the coefficients get highly affected by shifts. Hence, their application in stereo correspondences is limited. Contourlets and Curvelets are found to be more effective as local feature descriptors. They can clearly represent the intrinsic geometrical structures like curvatures from images. Contourlet based correspondence has been implemented in [4]. Curvelets are relatively new and have gained recognition from researchers for image denoising, face recognition etc. Their application in correspondence search is novel till now. They can be used for local area based matching along with any correlation function. Curvelets decompose the image into a number of scales and orientations. A correspondence match can be found in each of these scales and orientations for each pixel of the stereo images. The best match is one of these matches and can be estimated by simple comparison of the correlation values or left-right disparity consistency check. However, to reduce the fattening effect, support weights can be combined with curvelets.

In the paper, we propose a novel method for stereo matching based on curvelets with Modified-Adaptive-Support-Weights, which is termed as Curv-MASW. In the proposed method, color information is limited to original scale because the curvelet coefficients are treated as normal gray scale values. Thus the support weights are modified to fit to gray scale images. The proposed method is verified by extensive tests and evaluations.

The remaining part of the paper is organized as follows. We first present a brief description of curvelets in Section 2, followed by a description of the MASW and the implementation details are described in Section 3. The experimental results and comparison with current state-of-the-art methods in Middlebury database are presented in Section 4. Finally, Section 5 concludes our paper.

2. CURVELET TRANSFORM

Curvelets were developed by Candes and Donoho in 1999 mainly for image analysis. They have strong directional characteristics and due to their variable width and length with a parabolic scaling of length^2 $\sim$ width, the coefficients are highly anisotropic at fine scales.

As described by Candes in [5], curvelets can be thought of as obtained by applying parabolic dilations, rotations and translations to a specifically shaped function ψ ; they are indexed using scale $a(0 < a < 1)$, location b and orientation θ as

$$\psi_{a,b,\theta}(x) = a^{-\frac{3}{4}}\psi(D_a R_\theta(x-b)) \qquad (1)$$

with

$$D_a = \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{\sqrt{a}} \end{pmatrix}$$

where, D_a is a parabolic scaling matrix, R_θ is a rotation by θ radians. Second generation curvelet transforms are simpler and faster than their previous versions. There are two types of implementation - via Unequally Spaced Fast Fourier Transform (USFFT) or via wrapping [6]. Curvelets via wrapping has been used in the paper because it is faster and can be applied to any size of images effectively.

The curvelets measure the information of a signal at specified scale and locations but only along specified orientations. So, they effectively represent objects with *"curve-punctuated smoothness"* - smoothness except discontinuity along a general curve with bounded curvature [5]. Real images have lots of curves i.e. edges along different orientations. Using curvelets, these features can be represented with much greater accuracy than wavelets. For the current paper, the images have been decomposed into curvelet coefficients of scale=3 and orientation=8. Higher number of coefficients would yield better accuracy in expense of less speed. The current choice is selected as a trade-off between speed and accuracy and has been experimentally verified. The main advantage of using a multiresolution method for matching lies in initial disparity estimation. This will be explained later on.

3. ALGORITHM

A brief description of adaptive support weight is necessary to understand the modifications we made. Adaptive support weight has been proposed in [1]. The support weight of a pixel's neighborhood is written as

$$w(p,q) = k.\exp\left(-\left(\frac{\Delta c_{pq}}{\gamma_c} + \frac{\Delta g_{pq}}{\gamma_p}\right)\right) \tag{2}$$

where, Δc_{pq} and Δg_{pq} represent the color difference and spatial distance between the pixels p and q. k is a proportionality constant. γ_c and γ_p are constants. The details behind the theory can be found from [1]. The color space is chosen to be CIELab because of the ease of three dimensional representation of color. In the color space, Δc_{pq} is expressed as

$$\Delta c_{pq} = \sqrt{(L_p - L_q)^2 + (a_p - a_q)^2 + (b_p - b_q)^2} \tag{3}$$

where, (L_p, a_p, b_p) are the three color components of pixel p. In our implementation, the curvelets basically provide a group of two dimensional matrices that are treated as gray level images. Thus, we cannot use the color information to compute Δc_{pq} as in equation (3). As a replacement, we modify the support weight by including the gray level difference value

$$\Delta c_{pq} = \mid m_p - m_q \mid \tag{4}$$

where, m_p represents the gray level value of pixel p. Using only gray level information, the accuracy decreases but it is superseded by the accuracy improvement due to the search operation in multiple resolution and using curvelet coefficients. Having introduced the modified adaptive support weights, we come to the implementation details listed in Table 1.

Table 1: The implementation details

1. Perform curvelet transform on the gray scale stereo pairs; 2. Compute the correspondence match in the lowest approximate scale of curvelets by eq. (4) and (2); 3. Shift to the next scale, first orientation; 4. Compute scale_factor = (size of current scale image)/(size of previous scale image); 5. For each pixel, divide the coordinate of the pixel by scale_factor to get the coordinates in previous scale. Then, compute the initial_curr_disparity = scale_factor disparity in previous scale; 6. Take a range of search around the initial_curr_disparity; 7. Find the best match using MASW from eq. (2); 8. If this is the last orientation, go to step 10; 9. Shift to the next orientation and go to step 5; 10. For each pixel, there is a match in each orientation at current scale. Best match can be found by normal correlation value check or left-to-right and right-to-left disparity consistency check; 11. If this is the last curvelet scale, go to step 12, else go to step 3; 12. Compute final disparity map using MASW for the original image;

The implementation does not depend on any initial disparity estimation provided externally. The initial estimation is already performed in the lowermost scale i.e. the approximation. The approximated disparity map is then improved in each scale using different orientations of curvelets. Thus, the method completes a total disparity computation rather than a pre or post processing. The initial disparity map defines the range of search and the limit of maximum search. With the curvelet coefficients obtained at scale=3, orientation=8, the approximated image for Tsukuba pair has a dimension of 97x129. In this dimension, using the window size of 21x21, the maximum disparity obtained is 5. Multiplying this by a scale factor equal to the ratio of original image size and approximated image size, the disparity for normal image has maximum range of 15. This limit of range reduces search area and the uncertainty of wrong disparity choices. Finally, the use of support weights reduce the fattening effect that arises in local correspondence matching.

4. EXPERIMENTAL RESULTS

The experiments have been carried out on the Middlebury dataset [14, 15] and the results are compared in the Middlebury website with other methods. We also show the comparisons with wavelets by combining wavelets with MASW. We term it as Wave-MASW. This helps us compare the two multiresolution methods on the same benchmark.

The parameters in the tests are specified as follows - Size of support window = 33×33, γ_p=36, γ_c=7 and k=1.5. For the smallest approximation, window size has been taken as 21×21. The search area for higher scales have ±10 pixels deviation from original disparity position of the initial disparity map and reduces to ±5 pixels for the original size image pairs.

The experimental results are shown in Figure 1. The comparison results with recentmost methods are tabulated in Table 2. The results are evaluated with the percentage of bad pixels in the non-occluded regions, all regions and regions with discontinuity respectively. The formula for percentage of bad pixels is defined as

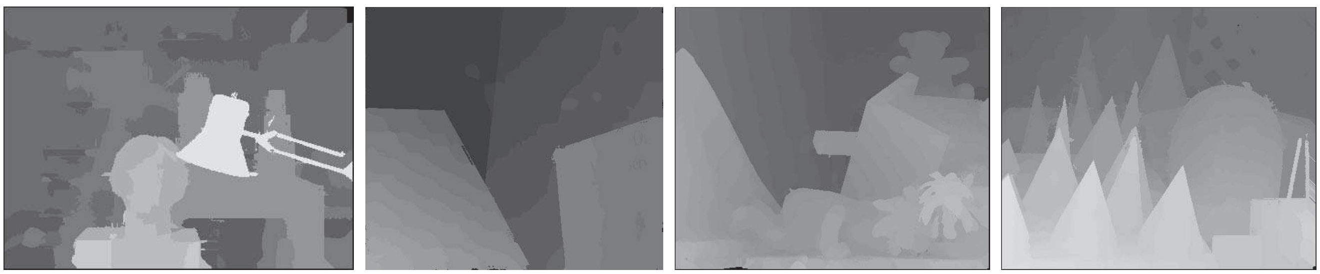
$$\frac{1}{N}\sum_{x,y}(\mid d_c(x,y) - d_t(x,y) \mid) > \delta_{thresh}$$

where, N is the total number of pixels in the region of interest, d_c and d_t represent the computed and ground truth disparity maps respectively, and δ_{thresh} is the threshold for bad pixels (usually equal to 1.0). Some results for Middlebury 2001 datasets [16] are also presented in Figure 2 for normal support weights and wave-MASW. Percentages of bad pixels for all regions are tabulated in Table 3.

As seen from Table 2, Curv-MASW has better performance than the original support weights, region Tree DP, realtime BP etc. methods. In comparison to support weights, the Tsukuba and Venus pairs have not shown considerable improvements but Cones and Teddy

Table 2: Comparison with state-of-the-art methods, the proposed method is denoted as Curv-MASW

Images / Methods	Tsukuba			Venus			Teddy			Cones		
	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc	nonocc	all	disc
CoopRegion [7]	0.87	1.16	4.61	0.11	0.21	1.54	5.16	8.31	13.0	2.79	7.18	8.01
DoubleBP [8]	0.88	1.29	4.76	0.13	0.45	1.87	3.53	8.30	9.63	2.90	8.78	7.79
DistinctSM [9]	1.21	1.75	6.39	0.35	0.69	2.63	7.45	13.0	18.1	3.91	9.91	8.32
Curv-MASW	**1.40**	**1.84**	**7.42**	**1.00**	**1.11**	**4.42**	**7.85**	**8.84**	**16.8**	**3.82**	**6.22**	**8.24**
RegionTreeDP [10]	1.39	1.64	6.85	0.22	0.57	1.93	7.42	11.9	16.8	6.31	11.9	11.8
AdaptWeight [1]	1.38	1.85	6.90	0.71	1.19	6.13	7.88	13.3	18.6	3.97	9.79	8.26
SegTreeDP [11]	2.21	2.76	10.3	0.46	0.60	2.44	9.58	15.2	18.4	3.23	7.86	8.83
RealtimeBP [12]	1.49	3.40	7.87	0.77	1.90	9.00	8.72	13.2	17.2	4.61	11.6	12.4
GC+occ [13]	1.19	2.01	6.24	1.64	2.19	6.75	11.2	17.4	19.8	5.36	12.4	13.0



(a) Curv-MASW (b) Curv-MASW (c) Curv-MASW (d) Curv-MASW

(e) Ground truth (f) Ground truth (g) Ground truth (h) Ground truth

Fig. 1: Dense disparity maps for the Middlebury images using our method and their corresponding ground truths. The first row shows the results from our method and second row shows the corresponding groundtruth.

pair are significantly improved. The reason for this lies in the initial disparity map creation. For Tsukuba and Venus, the maximum disparity is much smaller compared to Cones and Teddy. For smaller disparities, the initial disparity map creation process does not produce effectively better results than other methods. But, with increasing amount of disparity, the method works better than others due to the uniqueness of the curvelet coefficients in lower resolutions. This difference is prominent in Cones and Teddy. Cones pair has been listed in the top two in Middlebury images for all regions matching. This result can be further improved by using higher number of scales and angles or using better outlier removal methods in combination.

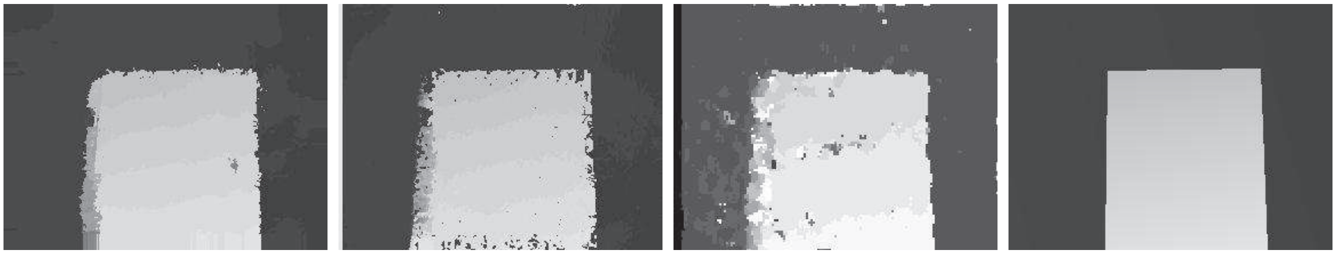
5. CONCLUSION

In this paper, we have proposed a new multiresolution based approach for stereo correspondence search. The approach adopts curvelets as the multiresolution method and support weights as cost minimization method. The method doesn't depend on the initial disparity estimation and does not suffer from the foreground fattening effect. The outputs are comparable to the state-of-the-art algorithms.

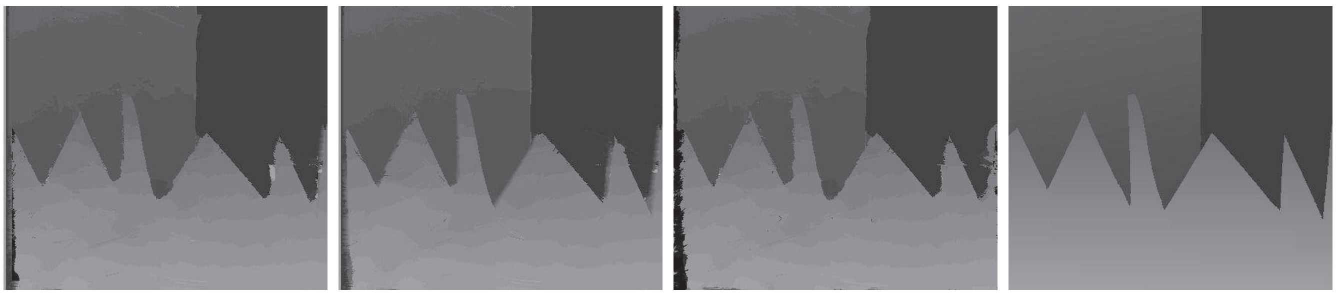
In the future, we plan to take more scales and orientations for better matching and combine the method with better outlier removal methods.

Table 3: Comparison for Middlebury 2001 dataset

Images / Methods	Map	Sawtooth
	% of bad pixels	
Curv-MASW	0.8212	0.5382
AdaptWeight	0.8554	0.5472
Wave-MASW	0.9711	0.5442



(a) Curv-MASW (b) Support weight (c) Wave-MASW (d) Ground truth



(e) Curv-MASW (f) Support weight (g) Wave-MASW (h) Ground truth

Fig. 2: Disparity maps for Sawtooth and Map pairs produced by Curv-MASW, Support weights and Wave-MASW

6. ACKNOWLEDGMENTS

The work is supported in part by the Natural Sciences and Engineering Research Council of Canada.

7. REFERENCES

[1] K. Yoon and I. Kweon, "Locally adaptive support-weight approach for visual correspondence search," *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 924–931, 2005.

[2] G. Caspary and Y. Y. Zeevi, "Wavelet-based multiresolution stereo vision," in *In Proceedings of the 16th International Conference on Pattern Recognition*, 2002, vol. 3, pp. 680–683.

[3] S.G. Mallat, "Wavelets for a vision," *In Proceedings of IEEE*, vol. 84, no. 4, pp. 604–614, 1996.

[4] H. Ding, M. Fu, and M. Wang, "Shift-invariant contourlet transform and its application to stereo matching," in *In Proceedings of the First International Conference on Innovative Computing, Information and Control*, 2006, pp. 87–90.

[5] E. J. Candes, "What is... a curvelet?," *In Notices of American Mathematical Society*, vol. 50, 2003.

[6] E. J. Candes, L. Demanet, D. L. Donoho, and L. Ying, "Fast discrete curvelet transforms," *In Multiscale Modelling and Simulation*, vol. 5, pp. 861–899, 2006.

[7] Z.F. Wang and Z.G. Zheng, "A region based stereo matching algorithm using cooperative optimization," in *In IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.

[8] Q. Yang, L. Wang, R. Yang, H. Stewénius, and D. Nistér, "Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling," *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 31, no. 3, pp. 492–504, 2009.

[9] K.J. Yoon and I.S. Kweon, "Stereo matching with the distinctive similarity measure," in *In IEEE 11th International Conference on Computer Vision*, 2007, pp. 1–7.

[10] C. Lei, J. Selzer, and Y.-H. Yang, "Region-tree based stereo using dynamic programming optimization," in *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, pp. 2378–2385.

[11] Y. Deng and X.Y. Lin, "A fast line segment based dense stereo algorithm using tree dynamic programming," in *In European Conference on Computer Vision*, 2006, pp. 201–212.

[12] Q.X. Yang, L. Wang, and R.G. Yang, "Real-time global stereo matching using hierarchical belief propagation," in *In British Machine Vision Conference*, 2006, pp. 989–998.

[13] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions via graph cuts," in *In International Conference on Computer Vision*, 2001, pp. 508–515.

[14] H. Hirschmuller and D. Scharstein, "Evaluation of cost functions for stereo matching," in *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.

[15] D. Scharstein and R.S. Szeliski, "High-accuracy stereo depth maps using structured light," in *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003, vol. 1, pp. 195–202.

[16] D. Scharstein and R.S. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *In International Journal of Computer Vision*, vol. 47, no. 1-3, pp. 7–42, 2002.