Adaptive Multiscale Stereo Images Matching Based on Wavelet Transform Modulus Maxima

Abdelhak EZZINE *aezzine@uae.ac.ma*
ENSAT/ SIC/LABTIC
Abdelmalek ESSAADI University
Tangier, 99000, Morocco

Youssef YazidiAlaoui *yazidi17@gmail.com*
ENSAT/ SIC/LABTIC
Abdelmalek ESSAADI University
Tangier, 99000, Morocco

Achraf Cohen *cohenachraf@gmail.com*
ENSAT/ SIC/LABTIC
Abdelmalek ESSAADI University
Tangier, 99000, Morocco

Abstract

In this paper we propose a multiscale stereo correspondence matching method based on wavelets transform modulus maxima. Exploitation of maxima modulus chains has given us the opportunity to refine the search for corresponding.Based on the wavelet transform we construct maps of modules and phases for different scales, then extracted the maxima and then we build chains of maxima. Points constituents maxima modulus chains will be considered as points of interest in matching processes. The availability of all its multiscale information, allows searching under geometric constraints, for each point of interest in the left image corresponding one of the best points of constituent chains of the right image. The experiment results demonstrate that the number of corresponding has a very clear decrease when the scale increases. In several tests we obtained the uniqueness of the corresponding by browsing through the fine to coarse scales and calculations remain very reasonable.

Keywords: Maxima modulus, stereo matching, wavelet transform, maxima modulus chains.

1 INTRODUCTION

Matching the extracted visual indexes of stereoscopic images is a very significant step within the process of 3D version. It is made up of bringing together the primitive extracts of right and left images so as to realize a 3D reconstruction of the scene. This is based on estimating the offset between the positions of matched primitives and building a disparity map showing the scene by highlighting the depth of scene objects. The disparity map generated from the correspondence matching process, along with a stereo camera are used to estimate the depth map and produce the 3D view of the scene.

This problematic is largely treated by a variety of researchers yet still there is no general method [6,8,10,26], many approaches exist in the field and exploit generally the primitives such as reference points, contours or areas but generally via the same characteristics which are either photometric, geometric or either morphologic.

This technique remains difficult due to the multiplicity of essential parameters, mainly the pertinence of considerable attributes, in the process of matching and the difficulty to find the good corresponding primitive is subject to a number of potential problems in vision process like occlusion, ambiguity, illuminative variations and radial distortion and all algorithms proposed are adapted to control at least some of these problems. We find many class of stereo vision algorithms, the first class is named global algorithms (GA) who deal with the correspondence estimation process as global cost function optimization problem and the second one is the local algorithms (LA) and are categorized into either area based or feature based algorithms [23]. The area-based algorithm used the correlation of the two image functions over locally defined regions and the feature-based algorithms establish correspondence between the selected primitives of the image.

Our method is one of the new class of algorithms [5,6], it is based on the concept of wavelets because the matching is the result of a multi-scale operation of the modulus and the phase of wavelet coefficients. This class is considered in the middle way between the local and global algorithm, these combine the best features of the LA and GA techniques and uses the multiresolution concept to involves the matching of two images at different scales.

This method is based on the wavelets modulus maxima chain (CMI) and involves matching of the different points forming the chain (CMI) to find the best corresponding in the right image for each extremity of one chain in left image. The search of the best one is done by estimating a similarity test between the points of each left chain and points of the right chains satisfying the geometric constraint like epipolar and orientation. The disparity map generated has a multiscale character as it is built based on the simultaneous treatment of space and scale. The multiscale character will permit the refinement of the disparity map and possibility to refine the selection of the potential candidates using a multiscale criterion based on normalized correlation and geometric refinement. As proved in Mallat[7,22,12], the extremity of the chain coincides with a singularity of the image characterized by its coefficient of Lipchitz, thus with a good choice for constructing maxima chains, the proposed matching algorithm not only provides the precise results but it also converges rapidly.

Firstly, we are going to recapitulate the main principles of wavelet transform modulus maxima (MMW). Then, we move to present the method and its particularities. Finally, we synthesis the major results of the method we intend to compare with other attributes.

2. WAVELETS TRANSFORM MODULUS MAXIMA

Based on the idea that in the image processing, visual data is put in a hierarchic way via the scale, wavelets transform decompose image into elementary blocs that are well localized both on scale and space. As a result of this decomposition, the image can be represented by the approximation to the coarse resolution and detail to all intermediate scales.

Among the main interests of representation in the form of wavelets we state the coefficient wavelet length exam of an image can inform us about it analysis. Indeed, in [2, 9], it is proved that the regularity accentuates the decay of wavelet coefficients; we are talking about the local regularity related to the analyzing wavelet support [17, 14].

Thereby, if the image is continuously regular everywhere except at a few isolated points the estimated wavelet coefficients will be affected only if the support of the analyzing wavelet contains these points or edges. In this sense, multiscale representations are better suited to focus the information in an image whose regularity is not homogeneous because the wavelet coefficients above a certain threshold focus only near the singularities (edges, single points, etc.). One of the multiscale decompositions that exploits well the wavelet properties which permit making a study, in a different levels, about the discontinuity points present in an image is wavelet transform modulus maximas proposed by Mallat and Zhong[1,7].

The principle of wavelet modulus maxima can give the image of characteristics to identify it, they transmit obvious features of the image depending on well defined models and well determined directions. For more details about wavelets modulus maxima we refer to [2, 13,14, 21] and present bellow the essential terms.

Suppose two wavelets ψ^x and ψ^y suh that $\psi^x = -\frac{\partial G}{\partial x}$ and $\psi^y = -\frac{\partial G}{\partial y}$ and $\int_{-\infty}^{+\infty} G(t)dt \neq 0$. For this work we choose G as a Gaussian wavelet to ensure both a valorization of wavelet maxima and continuity of the wavelet maxima chains [4, 7]. Thus wavelet transform of an image I can be written as a multiscale differential operator

$$\begin{pmatrix} W^x I(u,v,j) \\ W^y I(u,v,j) \end{pmatrix} = \begin{pmatrix} I * \psi_j^x(-u,-v) \\ I * \psi_j^x(-u,-v) \end{pmatrix} = 2^j \vec{\nabla} (I * G_j(-u,-v)) \qquad (1)$$

The polarized representation of this gradient vector offers the wavelet transform modulus

$$MI(u,v,j) = \sqrt{|W^x I(u,v,j)|^2 + |W^y I(u,v,j)|^2} \qquad (2)$$

and it is direction is defined by the angle $\theta I(u,v,j) = \tan^{-1}\left(\frac{W^y I(u,v,j)}{W^x I(u,v,j)}\right)$. (3)

Then the local maximum of the wavelet transform modulus $MMI(u,v,j)$ can be found by solving $\partial MI(u,v,j) = 0$.

Otherwise on some scale j, the points (u_0, v_0) such that $MMI(u,v,j)$ is a local maximum, according the L_∞ norm, on a neighborhood of (u_0, v_0) in the direction $\theta MI(u,v,j)$. We note that for image color we consider the maxima according DiZenzo approach[11].

3. MULTILEVEL MATCHING METHOD

In works of Mallat and Zhong[7] is prove that a wavelet coefficient can be influenced by a singularity resides within support of the analyzing wavelet; we talk about the cone of influence to express the set of points where the wavelet coefficients may be influenced by a singularity at a point (u_0, v_0). This singularity will generate local maxima of wavelet coefficient upon the cone and till a particular scale. In order to characterize this singularity, we will observe the behavior of local maxima in terms of scale; this will be much easier to follow if the maxima, between two successive scales, are connected together to form a curve in the plane scale space so-called line or chain maxima.

In practice, we start with an image of size $(N, N), N = 2^J$. We obtain the wavelet transform according dyadic approach [3,11,18] in which scale parameter is expressed by the power of $2, 2^j, j = J \ldots L, \text{L is the corsest scale.}$

This method generates on each scale, a wavelet transform with the same number of pixels as the image I; in spite of that the support of the wavelet is increasingly large when scales grow. After that we calculate the wavelets transform modulus and directions by equations (2) and (3). The results are two matrix, the maxima modulus matrix $MMI(.,.,j)$ and the angle matrix $\theta MI(.,.,j)$ for each scale $j = J-1 \ldots L$. The image I is then represented by $(MMI(.,.,j))_{j=J-1\ldots L}$, positions and values of the wavelet coefficients at each scale j when the module is a local maximum in the direction θ.

We construct the line or chain of maxima by a chaining (x_j, y_j), the locationof each element of$MMI(.,.,j)$, with (x_{j-1}, y_{j-1}),the location of its successor in $MMI(.,.,j-1)$[21,23]. The search for these successors is limited to the neighborhood defined by the cone of influence and every chain can be indexed by$(x_j, y_j)_{j=J\ldots m}$and noted$CMI(x_j^J, y_j^J, m)$[14,11]. The integer m indicate the length of the chain, and is defined such that the last point of the chain is ones of maxima of the wavelet transform at scale J-m. We note that each chain is composed of maxima and is characterized by its extremity $MMI(\mathrm{x_J}, \mathrm{y_J}, \mathrm{J}$ at the level J that coincides with a singular point of the image.

3.1 The matching approach

Given two rectified color images, from a vision system calibrated. we associate to each one its maxima modulus chains representation$(x_j^i, y_j^i)_{j=J\ldots m}$, $i = g$for the left image and$i = d$for the right one. With different representations defined above, we are faced with three possible methods to mapping. The first is to match directly the maxima modulus transforms$MMI_i(.,.,\mathrm{j})$of the two images at each scale$j = J - 1 \ldots L$. The second performs the matching of the extremity of each chain$CMG(x_J^g, y_J^g, m)$with an extremity of an chain$CMD(x_J^d, y_J^d, m')$by an non-linear algorithm. The third method proposes to consider the maxima chains as primitives.In this paper we develop the second method and we will give an idea about the treatment of the third, we note that the first one looks approximately like the work of Bhatti and we invite the reader to discover in this work [5].
To chain $CMG(x_J^g, y_J^g, m)$of the left image , we will search for all Chain of the right image such that their extremities respect the epipolar constraint with the pixel (x_J^g, y_J^g); all maxima of the left that have no match in the right image verifying the epipolar constraint are dropped. We then perform a test of similarity between $MMI(x_J^g, y_J^g, \mathrm{J})$and each one of the maxima modulus on extremities of the rights chains selected before. We note $\left\{(x_{J,k}^d, y_{J,k}^d)_{k=1..c}\right\}$the set of extremity of the right chains validating this test. If $c = 1$, we conclude that $(x_J^g, y_J^g)$is matched with $(x_{J,1}^d, y_{J,1}^d)$.If $c > 1$, we move to the next scale and we redo the estimating test between the successors, i.e. $(x_{J-1}^g, y_{J-1}^g)$and $\left\{(x_{J-1,k}^d, y_{J-1,k}^d)_{k=1..c}\right\}$. This step is repeated, recursively on the scale j, until the end of the chain and it will result necessarily matching of $(x_J^g, y_J^g)$and ones of $\left\{(x_{J,k}^d, y_{J,k}^d)_{k=1..c}\right\}$points.

The similarity is evaluated, under the orientation constraint, by using ones off different similarity measure available in the literature. Major similarity measures used in matching methods are measures of distortion (SAD) and (SSD) and the normalized cross correlation (NCC)[15]. It is known that SAD and SSD are computationally fast then NCC who is more accurate and it allowed us to refine the similarity test.

At fine scales, there are many edge points created by the image noise in the wavelet transform maxima modulus representation. By reason of the presence of these wrong maxima and because it is well known that the SSD and SAD are justified when the additive noise distribution is Gaussian or exponential, we are led to combine the NCC and SAD to the similarity test for the fine scales. We note that at larges scales the smoothing process removes most of the wrong maxima and the NCC is sufficient.

To find the correspondence between pixels of$\mathrm{M}MG(x_\mathrm{j}^g, y_\mathrm{j}^g, \mathrm{j})$ and $\mathrm{M}MD(x_\mathrm{j}^d, y_\mathrm{j}^d, \mathrm{j})$, one can maximize a correlation or minimize the distortion measure. For each point(x_j^g, y_j^g), we search for a point(x_j^d, y_j^d) in $CMD(x_J^d, y_J^d, m')$, whose a neighborhood of a size s has a maximum correlation or maximum distortion with the neighborhood of the point(x_j^g, y_j^g) in $CMD(x_J^g, y_J^g, m)$.Due to the epipolar constraint we have$y_j^d = y_j^g$and the search is reduced to points(x_j^d, y_j^g)on the epipolar line.

4. EXPERIMENT RESULTS

The algorithm presented in this paper is tested on the images from web site [10,10b]www.middlebury.edu/stereo/. These images are created for the stereo matching algorithms we present the results obtained for four of them with the same illumination. In order to give visual results, the performance of the proposed nonlinear algorithm is presented for different scales j with the dyadic approach wavelet transform so we can follow the behavior of a maximum on different scales.

The experimental results were generated using the Gaussian wavelets. In the first column of thefigure 1, we present the left images cones, teddy and tsukuba. In the fifth column we present the right view of the same images. In the second and fourth column we present the maxima modulus on scale J=4.

Into the third columnare presented the number (Z axis) of maxima of the right image validating the similarity test with each one of the maxima (X axis) of the left image. We present these results of the simulationonly for 50 points (on the X axis) represents the great maxima of each one of the tree images. The third axis (Y axis) represents the scale and for each maxima of the left image we represent the behavior of the number of corresponding maxima of the right image on each scale j.

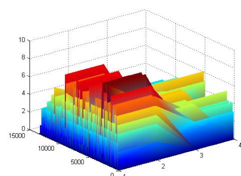
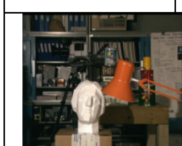
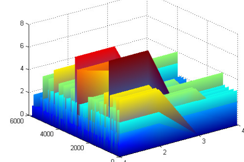
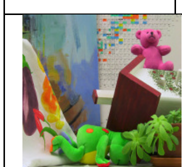
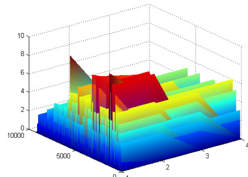
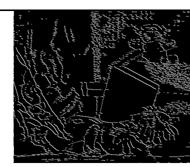
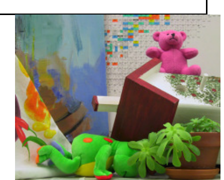
Left Image	Maxima modulus for scale j:4	Behavior of corresponding maxima	Maxima modulus for scale j:4	Right image
				
				
				

Figure 1: in third column we observe the decreasing behavior of number of corresponding maxima in right image at scale j:4 for everyone of maxima in left image (x axis :scale; y axis: maxima and z axis: number of corresponding)

We observe that the number of corresponding decrease clearly when the scale increases. There are more than 6 corresponding in the first scale, our method allows the reduction of this number to 1 for several maximas. We observe also that maxima corresponding to the strong singularity require further exploration scales.

This observation is more important and we are working to prove the relationship between the category of singularities and the speed of convergence of this method. To compare the results of this method with other work and especially by platform Middleburry[10,10b], we need to validate the disparity map based on the points of interest of our method and this is one of the future direction of our research.

5. CONCLUSION

This paper introduced a nonlinear algorithm matching based on wavelet transform maxima modulus. The maxima modulus chains are used to find the best corresponding maxima in the right image of each maximum in left image. This process exploits intelligently the multiscale character and offers an accurate and fast algorithm. We note also that the first results of the construction of disparity maps are very encouraging.

Following this work is, firstly, affine the disparity map and compare it to techniques benchmarked and published in the Middleburry database; and secondly integer the Lipschitz regularity of pixels [2,16,17] in the similarity process as an additive information, because it was proven that there is a strong relationship between maxima and this regularity.

This will allow us to exploit the nature of the maxima as edge points in the image and propose a non-linear feature-based algorithm. The next future direction on this work is to explore different wavelets basis and observe the behavior of the nonlinear algorithm of this paper according the choice of wavelets basis.

6. REFERENCES

[1] S. Zhong, "Edge Representation from Wavelet Transform Maxima". PhD thesis, New York University, September 1990.

[2] Y. Meyer, " Ondelettesetopérateurs ", vol. I. Paris: Hermann, 1990.

[3] M. Shensa, " Wedding the à trous and Mallat algorithms, " IEEE Trans. Signal Processing, vol. 40, no. 10, pp. 2464-2482, 1992.

[4] J. Canny, " computational approach to edge detection ", IEEE Trans. on PAMI, vol. 8, no. 6, pp. 679-698, 1986.

[5] A. Bhatti,S. Nahavandi "Stereo Image matching Using Wavelet Scale-Space Representation", Proceeding of CGIV'06, pp 267-272, 2006

[6] P. B. Zadeh, C. V. Serdean " A Hierarchical multiwavelet based stereo correspondence matching techniques ", EURASIP'2011, pp 393-397

[7] S. Mallat and S. Zhong, " Characterization of signals from mul- tiscale edges ", IEEE Trans. on PAMI, vol. 14, pp. 710-732, July 1992.

[8] G. Qu, D. Zhang and P. Yan " Medical image fusion by wavelet transform modulus maxima", Optics Express Vol 9, No 4 pp 184-190, 2001.

[9] A. Cohen and A. Ezzine "Quadratures singulières et fonctions d'echelle ", CRAS, t.323, Serie I, 829-834, 1996.

[10] D. Scharstein and R. Szeliski, www. Middlebury.edu/stereo/,

[10b] D. Scharstein and R. Szeliski, Stereo matching with nonlinear diffusion. Int. J. of Computer vision, 28(2), 1998.

[11] S. Dizenzo, " A note on the gradient of a multi-image ", CVGIP, vol. 33, pp. 116-125, 1986.

[12] S. Mallat and W. Hwang, " Singularity detection and processing with wavelets ", IEEE Trans. Information Theory, vol. 38, no. 2, pp.617-643, 1992.

[13] P. Carré, “ Maxima d’ondelettes couleur, colorwavelet maxima “, traitement du signal 2004 vol 21 special L’image numérique couleur 623-632.

[14] S. Mallat, “ Wavelets for a vision “, Processing of IEEE, vol. 84, NO 4, pp. 604-614, April 1996.

[15] L. Distefano, Stefano Mattocia, “ Fast template matching using bounded partial correlation “, Machine vision and application, NO 13, pp. 213-221, 2003.

[16] S. Jaffard, “ Exposants de Holder en des points données et coefficients d’ondelettes”, CRAS France,308ser I ,pp79-81,1989.

[17] I. Daubechies, Ten lectures on Wavelets, SIAM, Philadelphia, PA, 1992.

[18] S. Mallat, a wavelet tour of signal processing, Academic Press, second edition, 1998.

[19] J. Canny A computational approach to edge detection, IEEE Trans. Patt. Anal. and Mach. Intell.,vol PAMI-8 no. 6, pp. 679-698, Novembre 1986.

[20] M. Holschneider, R. Kronland-Martinet, J. Morlet, P. Tchamitchian Wavelets, Time-Frequency Methods, and Phase Space, Springer-Verlag, Berlin, 1989.

[21] O. Lecadet Méthodes d’ondelettes pour la segmentation d’images. Applications à l’imagerie médicale et au tatouage d’images, these de doctorat, 2003.

[22] S. Meignen, S. Achard, Time Localization of Transients with Wavelet Maxima Lines, IEEE. Trans. sign. Proc., accepted, 2004.

[23] Y. Ruichek, Perception de l’environnement par stereovision , Application à la sécurité dans les systèmes de transports terrestres, Phdthesis, USTL, France 2006.