Stereo image matching using multi-wavelets

AUTHOR(S)

Asim Bhatti, Saeid Nahavandi, H Zheng

PUBLICATION DATE

01-01-2003

HANDLE

10536/DRO/DU:30005258

Proceedings of the 1st International Light Metals Technology Conference 2003 <Edited by Arne Dahle>

STEREO IMAGE MATCHING USING MULTI-WAVELETS

Asim Bhatti, Saeid Nahavandi, Hong Zheng

CRC for Cast Metals Manufacturing (CAST),
[2]School of Engineering and Technology, Deakin University, Geelong 3217, Australia

Abstract

A multi-resolution image matching technique; based on multi-wavelets, followed by a *coarse to fine* strategy is presented. The technique addresses the estimation of corresponding points in the presence of occlusion, ambiguity and illuminative variations in the two images. In order to tackle the problem of occlusion and ambiguity a geometric topological approach is used along with the normalized correlation which performs well in the presence of illuminative variations. A comparison between scalar and multi-wavelets is also performed, in the context of finding the optimal corresponding points.

1 Introduction

The main focus of this work is to generate 3D models of the die casting surfaces using more than one 2D perspective view. The 3D reconstruction [1, 2] process by itself can be categorized into three main categories, calibration (calculating the intrinsic and extrinsic parameters of the camera) [3,4] finding the corresponding pairs of points projected from the same 3D point on to the two perspective views [2,3,4], and triangulation to project the 2D information back to the 3D space in order to create a 3D model [3,4,5,13]. The calibration and triangulation strategies are quite mature in both theoretical and applicative perspective but finding corresponding points from two views still suffers from many problems like occlusion, ambiguity, illuminative, radial distortions, etc.

For that purpose a multi-resolution technique was established, based on multi-wavelets. Ever since their discovery, multi-wavelets (wavelets with more than one scaling and wavelet functions) have been the focus of a lot of research in signal processing and pure mathematics. The interest in multi-wavelets is mainly due to the fact that they produce promising results in many applications such as speech, image, video compression, denoising, machine vision [8,9,11,12], etc. Their success stems from the fact that they can simultaneously posses the good properties of orthogonality, symmetry, high approximation order and short support which is not possible in the scalar case. Due to these properties multi-wavelets drew the attention of many researchers to apply them in different applications.

Multi-wavelets already have been proven to perform better, than scalar wavelets, in applications like image compression and denoising [8, 9] etc. In the application of correspondence matching some work has already been done using scalar wavelets [11, 12] and convincing results have been achieved. As multi-wavelets have proven to perform better than scalar ones, due to their extra properties, there is a great deal of moti apply them in the application of corre matching.

The organization of the rest of the work is as the next section a brief introduction of th multi-wavelets and their related filter banks section 3, a complete image matching algorit which can be subcategorized into r decomposition, correlation, geometric topol refinement and interpolation. In section 4, some results are given; obtain from the simple implementation of the algorithm given in section 3. Conclusions are presented in section 5.

2 Multi-Wavelets and Filter Banks

Multi-resolution can be generated not just in the scalar context i.e., with just one scaling function and one wavelet but also in the vector case where there is more than one scaling functions and wavelets. The latter case leads to the notion of multi-wavelets. A multi-wavelet basis is characterized by r scaling and r wavelet functions. Here r denotes the multiplicity in the vector setting with $r > 1$. Multi-scaling functions satisfy the matrix dilation equation or the refinement equation

$$\mathbf{\Phi}(t) = \sum_k \mathbf{C}_k \mathbf{\Phi}(2t - k) \tag{1}$$

Similarly for the multi-wavelets the matrix dilation equation is given by

$$\mathbf{\Psi}(t) = \sum_k \mathbf{D}_k \mathbf{\Phi}(2t - k) \tag{2}$$

where $\mathbf{\Phi}(t) = [\phi_0(t)\ \phi_1(t) \ldots \phi_{r-1}(t)]^T$,
$\mathbf{\Psi}(t) = [\psi_0(t)\ \psi_1(t) \ldots \psi_{r-1}(t)]^T$ and $\mathbf{C}_k$ and $\mathbf{D}_k$ are real r by r matrices or multi-filter coefficients that iteratively and with rescaling define the scaling and wavelet functions respectively i.e., the link between functions and filters is iteration with rescaling. For more information, about the generation and applications of multi-wavelets, with desired approximation order and orthogonality, the interested readers are referred to [6,7,8,10].

The filter bank corresponding to multi-wavelet systems with r scaling functions and r wavelets is shown in Figure 1.

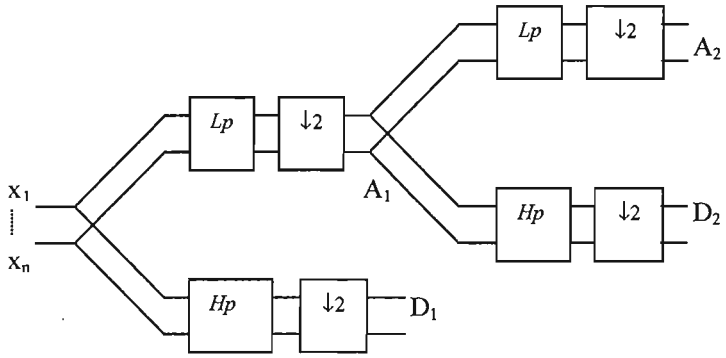


Figure 1. 2 level multiwavelet filter bank

In Figure 1, $\mathbf{A}_i$ and $\mathbf{D}_i$ represents the coefficient matrices of approximation and detail spaces, respectively, possessing the properties as follows

$$\cdots \subset A_1 \subset A_2 \cdots \subset A_j \subset \cdots \tag{3}$$

$$A_1 = A_2 \oplus D_2 \tag{4}$$

Similar to the scalar case, low pass filter Lp and high pass filter Hp consist of coefficients, corresponding to the matrix dilation equations (3) and (4).

$$Lp = [C_0 \quad \cdots \quad C_{N-1}] \tag{5}$$

$$Hp = [D_0 \quad \cdots \quad D_{N-1}] \tag{6}$$

where $[C_0 \quad \cdots \quad C_{N-1}]$ and $[D_0 \quad \cdots \quad D_{N-1}]$ are r by $r \times N$-1 coefficient matrices of scaling functions and wavelets respectively. As it is shown in Figure 1, filter banks need multiple streams instead of one, so input signals need to be preprocessed in order to be compatible with multiple stream criteria. For that purpose repeated row preprocessing is performed [8, 9]. An example of multi-wavelet decomposition is shown in Figure 2

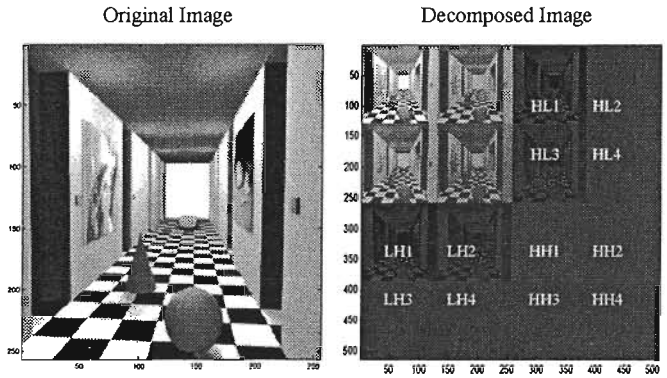


Figure 2. Decomposed image after 1-level

LHi, HLi and HHi, in figure 2, represents the horizontal, vertical and diagonal details of the image, which are actually the search spaces for the pixel matching process. During the rest of the process these image details LHi, HLi and HHi, will be called detail windows.

3 Image Matching Algorithm

The first step is the decomposition of images using a wavelet filter bank up to level N this ends up with 3×2^r number of wavelet coefficient matrices for each level. Where r is the multiplicity of the multi-wavelets where as scalar wavelets had unit multiplicity. These wavelet coefficients are also known as detail space coefficients of the images. For example using "Chui-Lian" multi-wavelet [10], after decomposition the result is 12 detail space coefficients matrices for each level and similarly using "mw112_r3_p2" [7], it will be 27, due to the multiplicity 2 and 3 respectively. For simplicity, detail space coefficient matrices will be denoted by detail windows, through the rest of the work.

3.1 Correlation

Initially, at the coarsest level, area based search is performed for each pixel in the left image, through all pixels in the right. As these are 3×2^r details windows, we will perform a comprehensive search for each pixel is performed in each detail window of the left camera throughout all detail windows of the right image. For that purpose a normalized correlation expression is used as a criterion to pick the optimal corresponding match. The correlation expression that have been used here is

$$CS(x,y,d) = \frac{\Sigma_{i,j}\left[(p_l - \overline{p_l}) \times (p_{rd} - \overline{p_{rd}})\right]^2}{\sqrt{\Sigma_{i,j}[p_l - \overline{p_l}]^2} \times \sqrt{\Sigma_{i,j}[p_{rd} - \overline{p_{rd}}]^2}} \tag{7}$$

where $p_l = p_l(x+i, y+j)$,

$p_{rd} = p_{rd}(x+i+d, y+j+d)$ where x and y represents the coordinates of the pixels and d represents the disparity (shift of the point in x and y direction) related to the pixel location at x and y in the left image, and

$$\overline{p_l} = \sum_n p_l(x+i, y+i)/n^2 \tag{8}$$

$$\overline{p_{rd}} = \sum_n p_{rd}(x+i+d, y+i+d)/n^2 \tag{9}$$

are the averages of the pixel intensities within the correlation windows in the left and right images. The main purpose of subtracting the averages is to minimize the effect of illuminative variations between the two images. The second good feature of that correlation expression is that it is invariant to change from P_l and P_{rc} to $a_1 p_l + b_1$ and $a_2 p_{rd} + b_2$, respectively [13]. During that correlation process a window of pixels, surrounding

the pixel under consideration, is used instead of the single pixel. In the proposed work, windows of 7×7 and 9×9 are used. That windowed correlation process can be better explained graphically as

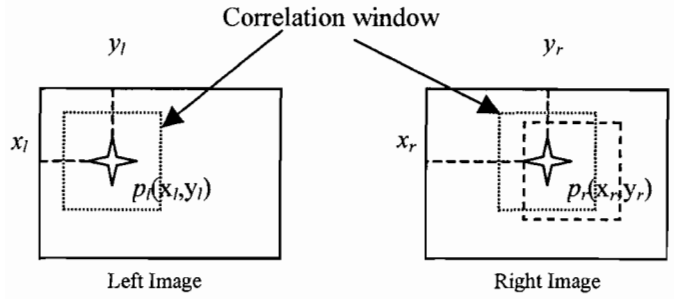


Figure 3. Windowed Correlation Process

As it is quite obvious that the score $CS(x,y,d)$ lies within [-1 1], with -1 for the correlation windows not similar at all and 1 for the identical ones. For each pixel in the left image there will be a set of 3×2^r matches, as mention before, with the same number of $CS(x,y,d)$ scores. That is in fact the upper limit of the number of corresponding matches for each pixel in the reference image. A constraint is then applied to select the most consistent matches. A pixel with correlation score higher than a predefined threshold will be selected as a candidate match, as in (10)

$$\max_{(x,y)}(CS(x,y,d)) >= \text{threshold} \qquad (10)$$

The threshold is usually taken within the range [0.5 0.8].

3.2 Geometric topological refining

After the correlation technique, as described above, for each pixel there will be a set of candidate matches with the upper limit of 3×2^r. These candidate matches can either be pointing to the same pixel location or different. Here an assumption for further processing is used.

Assumption: *If all the candidate matches are pointing to the same location and the size of the set of candidate matches is more than or equal to (3 times 2^r)/2, then that pair will be considered as true match.*

In order to find an optimal match from the set of candidate matches, containing different locations, a geometric topological approach is used. This method is quite robust against the occlusions and ambiguity. It involves the calculation of some geometric features and along with their correlations scores will be called, it the strength of the candidate match. The geometric feature that have been taken into account are *relative distances* and *angles* (slopes of lines instead of angels for simplicity), which are the invariant features through many geometric transforms like affine, metric, etc. These transforms are very common in the applications of stereo vision and 3D modeling, which is the main goal. The match strength is defined as

$$MS_k(x,y) = CS_k\left(\frac{e^{-rdd_k} + e^{-spd_k}}{2}\right) \qquad (11)$$

where k is the number of candidate matches for a specific pixel in the left image and CS_k is the average correlation score of k^{th} candidate match. rdd_k is the average of the relative distance differences and spd_k is the average of the slope differences between a pair of points, reference to candidate k. In order to minimize the effect of a wrong true match, chosen after the correlation step, rdd_k and spd_k are calculated more than once by picking a random point every time from the bin of true matches and then averaging the values. The optimal match will be the one with maximum match strength.

3.3 Interpolation

After the completion of matching process at the coarsest level there are a number of matches. The constellation relation between the pixels at coarser and finer levels can be understood by looking at factor 2 decimation involved in each level of wavelet filter bank, and is shown below

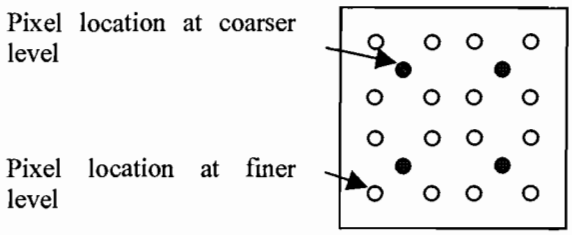


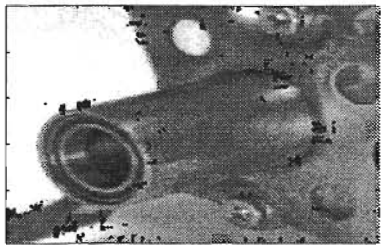
Figure 4. Coarser and finer level pixels' constellation relation

From this constellation, it is clear that each pixel at coarser level represents 4 pixels at finer scale. After the correlation step there will be a number of matching pairs. In the interpolation step these matches will be projected or interpolated to the next finer level. A windowed correlation process, as described in section 3.1, is performed again to check the consistency of these matches at the finer level. All the matches, at that level, with correlation scores satisfying the criterion of threshold, will be interpolated to the next finer level. That process will be continued until the finest resolution level is reached.

4 Results

The main purpose of this work was to produce 3D models of die casting surfaces. Due to copyright reasons, the die casting parts' images are not shown here. In order to show the performance of the technique presented above, an image of a mechanical part is used. For comparative reasoning, both wavelets and multi-wavelets are used for multi-resolution analysis. Due to the space shortage only the results obtained from Mallat's scalar wavelet [8] and mw112_r3_p2 [7] multi-wavelet are shown in Figures 5 and 6.

Left Image



Right Image

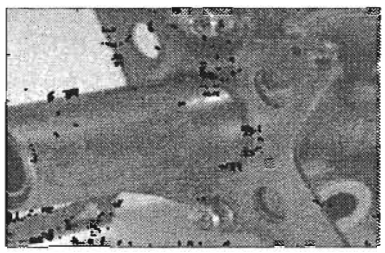
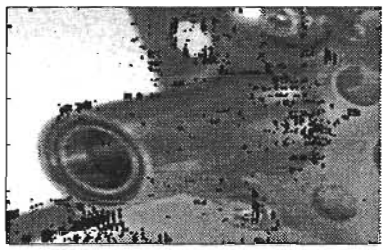


Figure 5. Corresponding point matches (695) found using Mallat's wavelets

Left Image



Right Image

Figure 6. Corresponding point matches (2038) found using Mw112_r3_p2 Multi-wavelets

For ease of demonstration, only small portions of the images are shown. There is no performance check criterion available to check the comparative performance of the scalar and multi-wavelets. The only thing that can be discussed, in terms of the performance here, is the number of corresponding matched points that have been obtained by using, scalar 695 and multi-wavelets 2038 and it was quite obvious that multi-wavelets have performed better than scalar one.

5 Conclusion

A multi-resolution image matching technique based on multi-wavelets is presented. Multi-wavelets have performed well and proved to have the potential, as a good tool, of solving the problems of image matching. It is also found and shown that multi-wavelets perfo better than scalar ones, as has been proven in many oth applications.

References

[1] E. Trucco and A. Verri, *Introductory Techniques for D Computer Vision*, Prentice Hall, 1998.

[2] A. Faugeras, "*Three dimentional Computer Vision - geometric Approach Viewpoin.* MIT Press, 1993.

[3] Z. Zhang. "A flexible new technique for came calibration". *IEEE Transactions on Pattern Analys and Machine Intelligence,* 22(11):1330-1334, 2000.

[4] W. H. Liao and J. K. Aggarwal, "Cooperati matching paradigm for the analysis of stereo ima sequences," *Int. J. of Imaging Systems a Technology*, 9(3):192–200, 1998.

[5] F. Kahl, R. Hartley, and K. Astrom, "Critic configurations for n-view projective reconstruction In *Proc. Conf. Computer Vision and Patte Recognition, Hawaii, USA*, 2001.

[6] A. Bhatti, H. Özkaramanli, "M- Band multiwavele from spline super functions with approximation order *International Conferance on Acoustics Speech an Signal Processing ICASSP, Orlando, FL, USA,* Ma 2002.

[7] Özkaramanli H, Bhatti A., Bilgehan B., "Mul wavelets from B-Spline Super Functions wit Approximation Order", *Signal Processing, Elsevie Science,* pp. 1029-1046, Aug. 2002.

[8] S. Mallat, "A thery for multiresolution signa decomposition: the wavelet representation", *IEE Trans. PAMI* 11 (1989) 674-693.

[9] V. Strela, P.N. Heller, G. Strang, P. Topiwala and C Heil, "The application of multi-wavelet filter banks t signal and image processing," *IEEE Trans. on Imag Processing*, 8(4):548-563, April 1999.

[10] C. K. Chui and J. Lian, "A Study of Orthonorma Multi-wavelets," *J. Applied Numerical Math.* Vol.20 no.3, pp273-298, 1996

[11] Magarey J. and Dick A., "Multiresolution stereo imag matching using complex wavelets", *Proc. 14th Int Conf. on Pattern Recognition (ICPR '98)*, Brisbane Australia, August 1998, Vol. I, pp. 4-7.

[12] Pan, H. P., " Uniform Full-Information Imag Matching Using Complex Conjugate Wavele Pyramids", XIII ISPRS Congree, Vienna, 1996 Internation Archieves of Photogrammetry and Remote Sensing, Vol. XXXI.

[13] Olivier Faugeras, Bernard Hotz, Hervé Mathieu, Thierry Viéville, Zhengyou Zhang, Pascal Fua, Eric Théron, Laurent Moll, Gérard Berry, Jean Vuillemin, Patrice Bertin, Catherine Proy, Real time correlation-based stereo: algorithm, implementations and application, *INRIA Reseach Report 2013* , August 1993.