2

Impact of Wavelets and Multiwavelets Bases on Stereo Correspondence Estimation Problem

Asim Bhatti and Saeid Nahavandi
Centre for Intelligent Systems Research, Deakin University
Australia

1. Introduction

Finding correct corresponding points from more than one perspective views in stereo vision is subject to number of potential problems, such as occlusion, ambiguity, illuminative variations and radial distortions. A number of algorithms has been proposed to address the problems as well as the solutions, in the context of stereo correspondence estimation. The majority of them can be categorized into three broad classes i.e. local search algorithms (LA) L. Di Stefano (2004); T. S. Huang (1994); Wang et al. (2006), global search algorithms (GA) Y. Boykov & Zabih (2001); Scharstein & Szeliski (1998) and hierarchical iterative search algorithms (HA) A. Bhatti (2008); C. L. Zitnick (2000). The algorithms belonging to the LA class try to establish correspondences over locally defined regions within the image space. Correlations techniques are commonly employed to estimate the similarities between the stereo image pair using pixel intensities, sensitive to illuminative variations. LA perform well in the presence of rich textured areas but have tendency of relatively lower performance in the featureless regions. Furthermore, local search using correlation windows usually lead to poor performance across the boundaries of image regions. On the other hand, algorithms belonging to GA group deals with the stereo correspondence estimation as a global cost-function optimization problem. These algorithms usually do not perform local search but rather try to find a correspondence assignment that minimizes a global objective function. GA group algorithms are generally considered to possess better performance over the rest of the algorithms. Despite of the fact of their overall better performance, these algorithms are not free of shortcomings and are dependent on how well the cost function represents the relationship between the disparity and some of its properties like smoothness, regularity. Moreover, how close that cost function representation is to the real world scenarios. Furthermore, the smoothness parameters makes disparity map smooth everywhere which may lead to poor performance at image discontinuities. Another disadvantage of these algorithms is their computational complexity, which makes them unsuitable for real-time and close-to-realtime applications. Third group of algorithms uses the concept of multi-resolution analysis Mallat (1999) in addressing the problem of stereo correspondence. In multi-resolution analysis, as is obvious from the name, the input signal (*image*) is divided into different resolutions, i.e. *scales and spaces* Mallat (1999); A. Witkin & Kass (1987), before estimation of the correspondence. This group of algorithms do not explicitly state a global function that is to be minimized, but rather try to establishes correspondences in a hierarchical manner J. R. Bergen & Hingorani (1992); Q'ingxiong Yang & Nister (2006), similar to iterative optimization algorithms Daubechies (1992). Generally, stereo correspondences established in lower resolutions are propagated to higher resolutions in an

iterative manner with mechanisms to estimate and correct errors along the way. This iterative error correction minimizes the requirements for explicit post processing of the estimated outcomes. In this work, the goal is to provide a brief overview of the techniques reported within the context of stereo correspondence estimation and wavelets/multiwavelets theory and highlight the deficiencies inherited in those techniques. Using this knowledge of inherited shortcomings, we propose a comprehensive algorithm addressing the aforementioned issues in detailed manner. The presented work also focuses on the use of multiwavelets basis that simultaneously posses properties of orthogonality, symmetry, high approximation order and short support, which is not possible in the wavelets case A. Bhatti (2002); Ozkaramanli et al. (2002). The presentation of this work is organized by providing some background knowledge and techniques using multiresolution analysis enforced by wavelets and multiwavelets theories. Introduction of wavelets/ multiwavelets transformation modulus maxima will be presented in section 3. A simple, however, comprehensive algorithm is presented next, followed by the presentation of some results using different wavelets and multiwavelets bases.

2. Wavelets / multiwavelets analysis in stereo vision: background

The multi-resolution analysis is generally performed by either Wavelets or Fourier analysis Mallat (1999; 1989; 1991). Wavelets analysis is relatively newer way of scale space representation of the signals and considered to be as fundamental as Fourier and a better alternative A. Mehmood (2001). One of the reasons that makes wavelet analysis more attractive to researchers is the availability and simultaneous involvement of a number of compactly supported bases for scale-space representation of signals, rather than infinitely long sine and cosine bases as in Fourier analysis David Capel (2003). Approximation order of the scaling and wavelet filters provide better approximation capabilities and can be adjusted according to input signal and image by selecting the appropriate bases. Other features of wavelet bases that play an important role in signal/ image processing application are their shape parameters, such as symmetric and asymmetric, and orthogonality (i.e. $\langle f_i, f_j \rangle = 0$ if $i \neq j$) and orthonormality (i.e. $\langle f_i, f_j \rangle = 1$ if $i = j$). All these parameters can be enforced at the same time in multiwavelets bases however is not possible in scaler wavelets case A. Bhatti (2002). Wavelet theory has been explored very little up to now in the context of stereo vision. To the best of author's knowledge, Mallat Mallat (1991); S. Mallat & Zhang (1993) was the first who used wavelet theory concept for image matching by using the zero-crossings of the wavelet transform coefficients to seek correspondence in image pairs. In S. Mallat & Zhang (1993) he also explored the the signal decomposition into linear waveforms and signal energy distribution in time-frequency plane. Afterwards, Unser M. Unser & Aldroubi (1993) used the concept of multi-resolution (coarse to fine) for image pattern registration using orthogonal wavelet pyramids with spline bases. Olive-Deubler-Boulin J. C. Olive & Boulin (1994) introduced a block matching method using orthogonal wavelet transform coefficients whereas X. Zhou & Dorrer (1994) performed image matching using orthogonal Haar wavelet bases. Haar wavelet bases are one of the first and simplest wavelet basis and posses very basic properties in terms of smoothness, approximation order Haar (1910), therefore are not well adapted for correspondence problem. In aforementioned algorithms, the common methodology adopted for stereo correspondence cost aggregation was based on the difference between the wavelet coefficients in the perspective views. This correspondence estimation suffers due to inherent problem of translation variance with the discrete wavelet transform. This means that wavelet transform coefficients of two shifted versions of the same image

may not exhibit exactly similar pattern Cohen et al. (1998); Coifman & Donoho (1995). A more comprehensive use of wavelet theory based multi-resolution analysis for image matching was done by He-Pan in 1996 Pan (1996a;b). He took the application of wavelet theory bit further by introducing a complete stereo image matching algorithm using complex wavelet basis. In Pan (1996a) He-Pan explored many different properties of wavelet basis that can be well suited and adaptive to the stereo matching problem. One of the major weaknesses of his approach was the use of point to point similarity distance as a measure of stereo correspondences between wavelet coefficients as

$$SB_j((x,y),(\acute{x},\acute{y})) = |B_j(x,y) - \acute{B}_j(\acute{x},\acute{y})| \tag{1}$$

Similarity measure using point to point difference is very sensitive to noise that could be introduced due to many factors such as difference in gain, illumination, lens distortion, etc. A number of real and complex wavelet bases were used in both Pan (1996a;b) and transformation is performed using wavelet pyramid, commonly known by the name Mallat's dyadic wavelet filter tree (DWFT) Mallat (1999). Common problem with DWFT is the lack of translation and rotation invariance Cohen et al. (1998); Coifman & Donoho (1995) inherited due to the involvement of factor 2 down-sampling as is obvious from expressions 2 and 3.

$$S_A[n] = \sum_{-\infty}^{\infty} x[k]L[2n-k] \tag{2}$$

$$S_W[n] = \sum_{-\infty}^{\infty} x[k]H[2n+1-k] \tag{3}$$

Where L and H represent filters based on scaling function and wavelet coefficients Mallat (1999); Bhatti (2009). Furthermore similarity measures were applied on individual wavelet coefficients which is very sensitive to noise. In Esteban (2004), conjugate pairs of complex wavelet basis were used to address the issue of translation variance. Conjugate pairs of complex wavelet coefficients are claimed to provide translation invariant outcome, however increases the search space by twofold. Similarly, Magarey J. Magarey & Kingsbury (1998); J. Margary & dick (1998) introduced algorithms for motion estimation and image matching, respectively, using complex discrete *Gabor-like* quadrature mirror filters. Afterwards, Shi J. Margary & dick (1998) applied *sum of squared difference* technique on wavelet coefficients to estimate stereo correspondences. Shi uses translation invariant wavelet transformation for matching purposes, which is a step forward in the context of stereo vision and applications of wavelet. More to the wavelet theory, multi-wavelet theory evolved Shi et al. (2001) in early 1990s from wavelet theory and enhanced for more than a decade. Success of multiwavelets bases over scalar ones, 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 Mallat (1999); A. Bhatti (2002); Ozkaramanli et al. (2002). Being a new theoretical evolution, multi-wavelets are still new and are not yet applied in many applications. In this work we will devise a new and generalized correspondence estimation technique based wavelets and multiwavelets analysis to provide a framework for further research in this particular context.

3. Wavelet and multiwavelets fundamentals

Classical wavelet theory is based on the dilation equations as given below

$$\phi(t) = \sum_h c_h \phi(Mt - h) \tag{4}$$

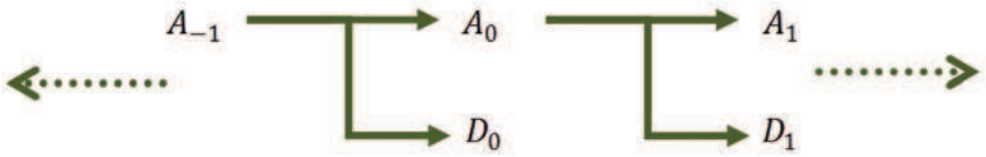


Fig. 1. wavelet theory based Multiresolution analysis

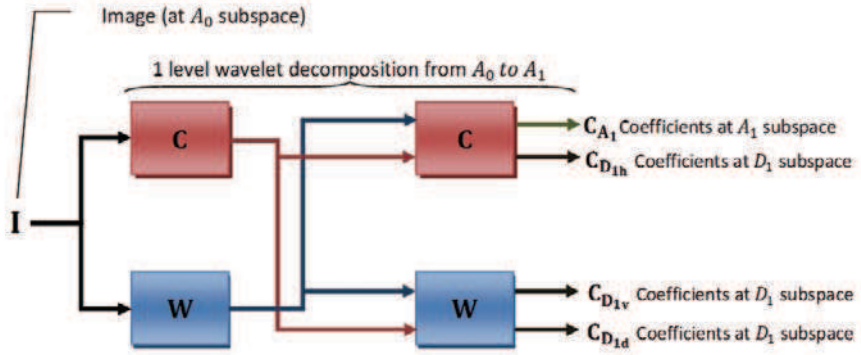


Fig. 2. Mallat's dyadic wavelet filter bank

$$\psi(t) = \sum_h w_h \phi(Mt - h) \tag{5}$$

Expressions (4) and (5) define that scaling and wavelet functions can be represented by the combination of scaled and translated version of the scaling function. Where c_h and w_h represents the scaling and wavelet coefficients which are used to perform discrete wavelet transforms using wavelet filter banks. Similar to scalar wavelet, multi-scaling functions satisfy the matrix dilation equation as

$$\Phi(t) = \sum_h C_h \Phi(Mt - h) \tag{6}$$

Similarly, for the multi-wavelets the matrix dilation equation can be expressed as

$$\Psi(t) = \sum_h W_h \Phi(Mt - h) \tag{7}$$

In equations 6 and 7, C_h and W_h are real and matrices of multi-filter coefficients. Generally only two band multiwavelets, i.e. $M = 2$, defining equal number of multi-wavelets as multi-scaling functions are used for simplicity. For more information, about the generation and applications of multi-wavelets with, desired approximation order and orthogonality, interested readers are referred to Mallat (1999); A. Bhatti (2002).

3.1 Multiresolution analysis

Wavelet transformation produces scale-space representation of the input signal by generating scaled version of the approximation space and the detail space possessing the properties as

$$\cdots A_{-1} \supset A_0 \supset A_1 \cdots \tag{8}$$

$$\overline{\bigcup_{-\infty}^{\infty} A_s} = L^2(R) \tag{9}$$

$$\bigcap_{-\infty}^{\infty} A_s = 0 \tag{10}$$

$$A_0 = A_1 \bigoplus D_1 \tag{11}$$

In expression (8) subspaces A_s are generated by the dilates of $\phi(Mt-h)$, whereas translates of $\phi(t-h)$ produces basis of the subspace A_0 that are linearly independent. A_s and D_s represents approximation and detail subspaces at lower scales and by direct sum constitutes the higher scale space A_{s-1}. In other words A_s and D_s are the sub-spaces of A_{s-1}. Expression (11) can be better visualize by the Figure 1. 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 are involved. A multi-wavelet basis is characterized by r scaling and r wavelet functions. Here r denotes the multiplicity of the scaling functions and wavelets in the vector setting with $r > 1$. In case of multiwavelets, the notion of multiresolution changes as the basis for A_0 is now generated by the translates of r scaling functions as

$$\Phi(t) = \begin{bmatrix} \phi_0(t) \\ \phi_1(t) \\ \vdots \\ \phi_{r-1}(t) \end{bmatrix} \tag{12}$$

and

$$\Psi(t) = \begin{bmatrix} \psi_0(t) \\ \psi_1(t) \\ \vdots \\ \psi_{r-1}(t) \end{bmatrix} \tag{13}$$

The use of Mallat's dyadic filter-bank Abhir Bhalerao & Wilson (2001) results in three different detail space components, which are the horizontal, vertical and diagonal. Figure 2 can best visualize the graphical representation of the used filter-bank, where C and W represents the coefficients of the scaling functions and wavelets, respectively, as in 6 and 7. Figure 3 shows transformation of Lena image using filter bank of Figure 2 and Daubechies-4 B. Chebaro & Castan (1993) wavelet coefficients.

3.2 Translation invariance

Discrete wavelets and multiwavelets transformations inherently suffer from lack of translation invariance. In the context of stereo vision, translation invariant representation of the signal is of extreme importance. The translation of the signal should only translates the numerical descriptors of the signal but should not modify it, otherwise recognition of the similar features within the translated representation of the signal could be extremely difficult. The problem of translation variance arises, in discrete dyadic wavelet transform, due to the factor-2 decimation which stands for the disposal of every other coefficient without considering its significance. To address this inherent shortcoming of translation invariance we have adopted the approach of utilizing wavelet transformation modulus maxima coefficients instead of simple transformation coefficients. The filter bank proposed by Mallat Mallat (1999) is modified in this work by removing the decimation of factor 2, which discards every

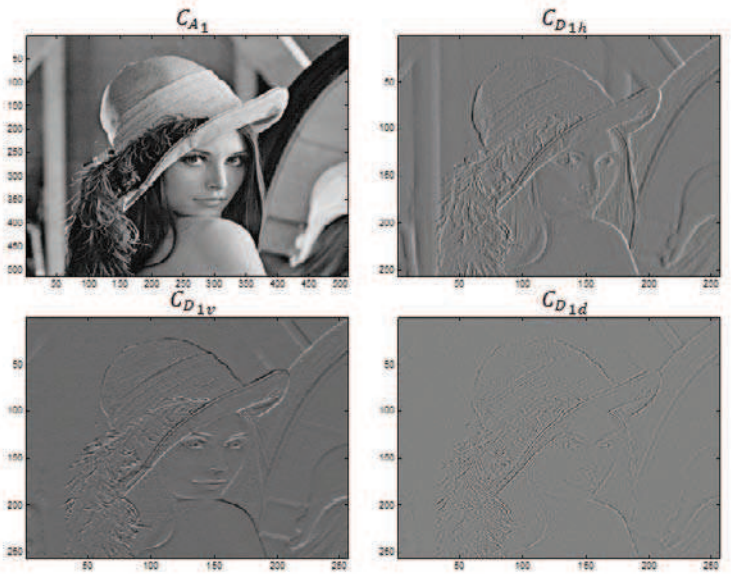


Fig. 3. 1-level discrete wavelet transform of Lena image using figure 2 filter bank

second coefficient, consequently creating an over complete representation of coefficients at subspaces (D_j). Instead, zero padding is performed for coefficients that are not transform modulus maxima. For correspondence estimation between stereo pair of images wavelet transform modulus maxima coefficients are employed to provide translation invariance representation. The proposed approach in achieving translation invariance is motivated by Mallat's approach of introducing critical down sampling Mallat (1999; 1991) into the filter bank instead of factor-2. Before proceeding to translation invariant representation of wavelets and multiwavelets transformation, concept of scale normalization is adopted (Figure 2) as

$$\zeta_s = \left| \frac{C_{D_{s,j}}}{C_{A_s}} \right| \qquad \forall\, s \text{ and } j \in \{h, v, d\} \tag{14}$$

$|.|$ defines the absolute values of the coefficients' magnitudes at scale s. The benefit of wavelets and multiwavelets scale normalization is two fold. Firstly, it normalizes the variations in coefficients, at each transformation level, either introduced due to illuminative variations or by filters gain. Secondly, if the wavelets and multiwavelets filters are perfectly orthogonal, the features in the detail space become more prominent. Let wavelet transform modulus (WTM) coefficients in polar representation be expressed as

$$\Xi_s = \zeta_s \angle \Theta_{\zeta_s} \tag{15}$$

Where ζ_s defines the magnitude of (WTM) coefficients and can be further expanded by referring to (2) as

$$\zeta_s = \frac{1}{3}\left(\sqrt{C_{D_{sh}}^2 + C_{D_{sv}}^2 + C_{D_{sd}}^2}\right) \tag{16}$$

Where D_{jh}, D_{jv} and D_{jd} represents D_1 subspace coefficients, which in visual terms represent discontinuities of the input image **I** along horizontal, vertical and diagonal dimensions. The

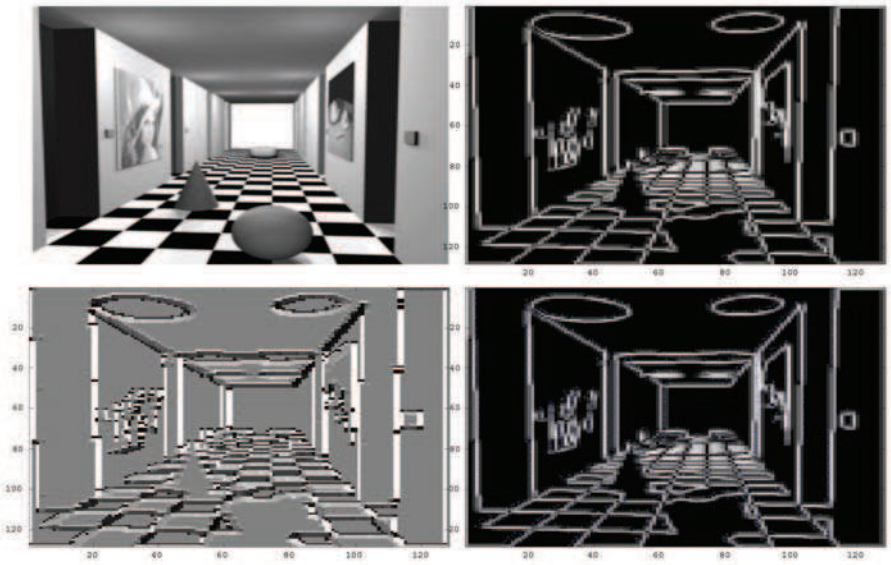


Fig. 4. Top Left: Original image, Top Right: Wavelet Transform Modulus, Bottom Left: wavelet transform modulus phase, Bottom Right: Wavelet Transform Modulus Maxima with Phase vectors

phase of (WTM) coefficients (Θ_{ζ_s}), which in fact is the phase of the discontinuities (edges) pointing to the normal of the plan that edge lies in, can be expressed as

$$\Theta_{\zeta_s} = \begin{cases} \alpha & \text{if} \quad C_{D_{sh}} > 0 \\ \pi - \alpha & \text{if} \quad C_{D_{sh}} < 0 \end{cases} \tag{17}$$

where

$$\alpha = tan^{-1}\left(\frac{C_{D_{jv}}}{C_{D_{jh}}}\right) \tag{18}$$

These discontinuities are referred by Mallat as multi-scale edges Mallat (1999) (section 6.3, page 189).The vector $\vec{n}(k)$ points to the direction, normal to the plan where the discontinuity lies in, as

$$n(k) = [cos(\Theta_{\zeta_s}), sin(\Theta_{\zeta_s})] \tag{19}$$

A discontinuity is the point p at scale s such that Ξ_s is locally maximum at $k = p$ and $k = p + \varepsilon n(k)$ for $|\varepsilon|$ small enough. These points are known as wavelet transform modulus maxima Ξ_n, and are translation invariant through the wavelet transformation and can be expressed by reorganizing expression 15 as

$$\Xi_{ns} = \zeta_{ns} \angle \Theta_{\zeta_{ns}} \tag{20}$$

Through out the rest of presentation, **coefficients** term will be used for wavelet transform modulus maxima coefficients instead of wavelets and multiwavelets coefficients, as in 20. An example of wavelet transform modulus maxima coefficients can be visualized by Figure 4. For further details in reference to wavelet modulus maxima and its translation invariance, reader is kindly referred to Abhir Bhalerao & Wilson (2001) (section 6).

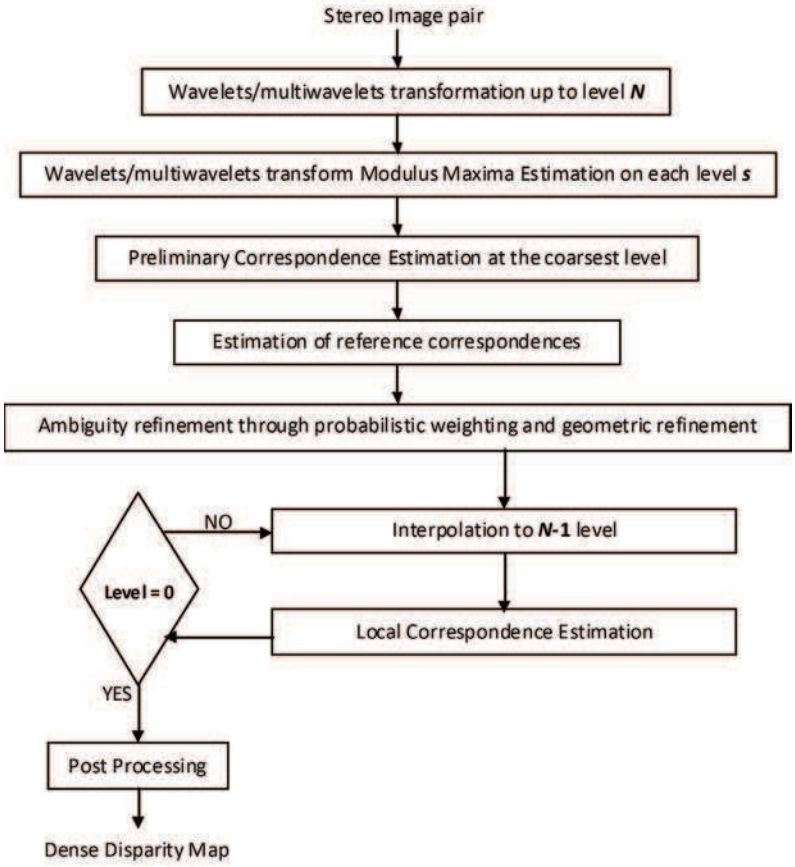


Fig. 5. A simple block representation of the proposed algorithm

4. Correspondence estimation

In the light of multiresolution techniques, presented in section 2 and their inherited shortcomings, we propose a novel wavelets and multiwavelets analysis based stereo correspondence estimation algorithm. The algorithm is developed to serve two distinct purposes; 1) to exploit the potential of wavelet and multiwavelets scale-space representation in solving correspondence estimation problem; and 2) providing a test-bed to explore the correlations of embedded properties of wavelets and multiwavelets basis, such as approximation order, shape and orthogonality/orthonormality with the quality of stereo correspondence estimation. The correspondence estimation process of the proposed algorithm is categorized into two distinct steps. First part of the algorithm defines the correspondence estimation at the coarsest transformation level, i.e. at signal decomposition level N. Figure 2 can facilitate visualization of signal decomposition considering the presented filter bank decomposes the signal up to level 1. Second phase of the algorithm defines the iterative matching process from finer $(N-1)$ to finest (0) transformation level, which according to Figure 1 refers to subspace A_0. Correspondence estimation at the coarsest

level is the most important part of the proposed algorithm due to its hierarchical nature and dependance of finer correspondences on the outcomes of coarser level establishments. Estimation of correspondences at finer levels use local search methodology searching only at locations where correspondences have already been established in the coarser level search. A block diagram representing the process of the proposed algorithm is shown in Figure 5.

4.1 Similarity measure

To establish initial correspondences, similarity measure is performed on modulus maxima coefficients (Ξ_s) using correlation measure Medioni & Nevatia (1985) enforced by multi-window approach Alejandro Gallegos-Hernandez (2002) (Figure 6) as

$$C_{\Xi} = C_{\Xi,W_0} + \sum_{i=1}^{n_W/2} C_{\Xi,W_i} \tag{21}$$

Where C_{Ξ} represents the correlation score of wavelets transform modulus maxima, under investigation and n_W represents the number of surrounding windows, usually taken as 9, without considering W_0. The second summation term in (21) represent the summation of best $n_W/2$ windows out of n_W. An average of the correlation scores from these windows is taken to keep the score normalized i.e. within the range of $[0\ 1]$.

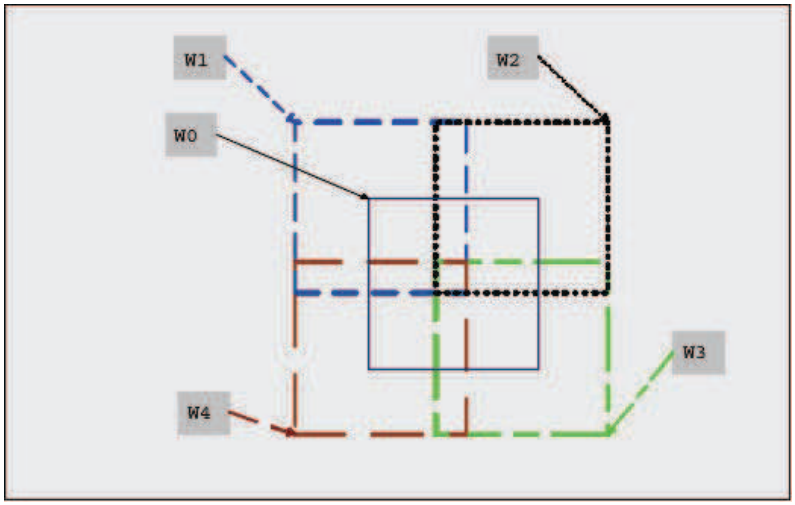


Fig. 6. Multi-window approach for correlation estimation

4.2 Probabilistic weighting

Wavelets and multiwavelets transformations, using filter-bank (Figure 2), produce r^2 sub-spaces for each bank at each scale. r defines the multiplicity of scaling functions and wavelets, which is one (i.e. $r = 1$) for wavelets, whereas $r > 1$ in case of multiwavelets, as illustrated in (12 and 13). Figure 7 represents one level multiwavelets transformation using GHM basis C. Baillard & Fitzgibbon (1999) with $r = 2$, therefore each subspace $(C_{A_1}, C_{D_{sh}}, C_{D_{sv}}, C_{D_{sd}})$ has produced 4 subspaces in contrast to one subspace as shown in Figure 3. Consequently, multiwavelets transform modulus maxima representation will consists of r^2 subspaces (16) for correspondence estimation process at each scale s. To ensure

the contribution of all coefficients from r^2 subspaces, probabilistic weighting is introduced to strengthen correlation measure of (21). In case of wavelets with $r = 1$, this step is bypassed. Probabilistic weighting defines the probability of optimality for any corresponding pair of coefficients. To define this probability; let Ξ_{c1} be the reference coefficient that belongs to one image of the stereo pair and Ξ_{c2_j} be the corresponding coefficients from the other image. The term j in Ξ_{c2_j} is due to the fact that sometimes different coefficients from r^2 subspace of the other image appear to be the potential correspondences for Ξ_{c1} coefficient. This phenomena is generally referred to as ambiguity Baker & Binford (1981).

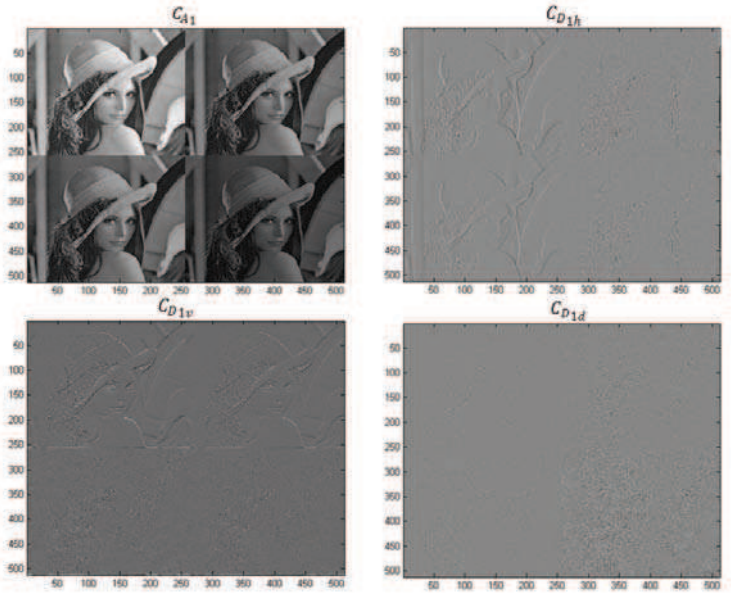


Fig. 7. 1-level discrete multiwavelets transform of Lena image using figure 2 filter bank and GHM multiwavelets C. Baillard & Fitzgibbon (1999)

The probability expression for corresponding pair (Ξ_{c1}, Ξ_{c2_j}) is defined as

$$P_{\Xi_{c2_j}} = n_{\Xi_{c2j}}/r^2 \quad \text{where} \quad 1 \leq n_{\Xi_{c2_j}} \leq r^2, \ \forall j \tag{22}$$

where $n_{\Xi_{c2_j}}$ is the number of times coefficient Ξ_{c2_j} is appeared as potential correspondence for Ξ_{c1}. In case of no ambiguity, Ξ_{c2} will appear as corresponding coefficient for Ξ_{c1} throughout r^2 subspaces, producing the $P_{\Xi_{c2}} = \frac{r^2}{r^2} = 1$. It is obvious from expression (22) that the P_{Ξ_c} lies between the range of $[1/r^2 \ \ 1]$. The correlation score in expression (21) is then weighted with P_{Ξ_c} as

$$\aleph_{\Xi_{c2_j}} = \frac{P_{\Xi_{c2_j}}}{r^2} \sum_{n_{\Xi_{c2j}}} C_{\Xi_{c2_j}} \tag{23}$$

r^2 term in expression (23) is for normalization of the correlation scores which will be accumulated over r^2 subspaces. In case of no ambiguity between the correspondence of Ξ_{c1}

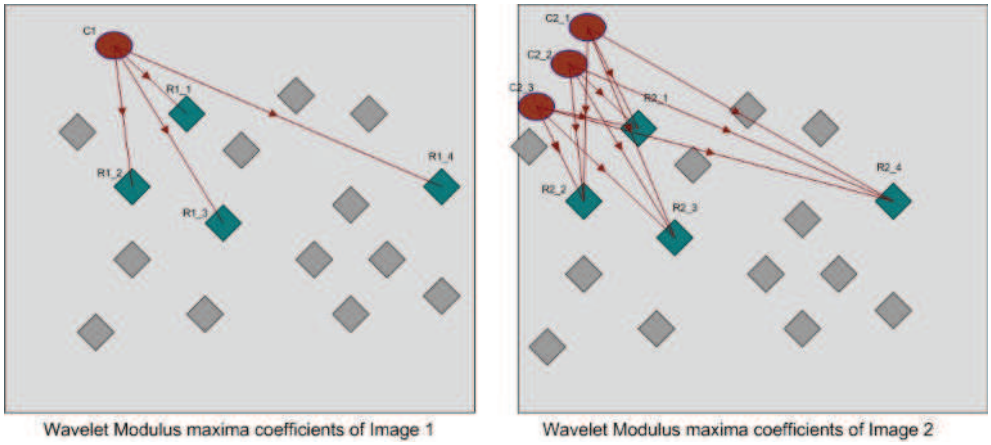


Fig. 8. Geometric refinement procedure

and Ξ_{c2} throughout r^2 subspaces, expression 23 will be simplified to C_Ξ as in expression 21 as

$$\aleph_{\Xi_{c2}} = \frac{1}{r^2}\left(r^2 \times C_{\Xi_{c2}}\right) = C_{\Xi_{c2}} \tag{24}$$

Simplification of expressions from 23 to 24 is of course under the assumption that C_Ξ is constant for the corresponding pair trough out the r^2 subspace, which is found to be true majority of the times. Corresponding pairs with $P_\Xi = 1$ and C_Ξ above predefined threshold, usually within the range of $[0.6 \quad 0.7]$, are used as references in addressing the ambiguity problem for rest of the correspondences. These reference coefficients provide a test ground to measure the credibility of rest of the correspondences by employing geometric refinement technique, presented in the following section.

4.3 Geometric refinement

Geometric refinement is employed to filter credible coefficients' correspondences, out of the ambiguous ones, using established reference correspondences from previous section-4.2. Three geometric features, relative distance difference (*RDD*), absolute distance difference (*ADD*) and relative slope difference (*RSD*), are employed to perform geometric refinement. The selection of these geometric features is influenced by their invariant nature through geometric transformations, such as Projective, Affine, Matric and Euclidean Siebert (1998). Geometric refinement procedure can be best visualized by Figure 8 where red circles represent candidate coefficient correspondences and squares represent reference coefficients. In Figure 8, $C1$ represents the coefficients from first image with potential corresponding coefficients $C2\_i$ from second image. Similarly, $R1$ and $R2$ represents reference corresponding coefficients with respect to first and second images, respectively. Small number of randomly chosen reference correspondences are employed in this phase to keep the process computationally less expansive. Let n_r be the number of randomly chosen reference correspondences out of N_r total reference correspondences and n_c be the number of candidate corresponding coefficients represented by $C2_j$ in Figure 8. With trial and error it has been found that n_r within the range $[3 \quad 5]$ produces desired outcome. Let Ξ_{n_r} and $\acute{\Xi}_{n_r}$ be the reference corresponding coefficients and Ξ_{n_c} and $\acute{\Xi}_{n_c}$ be the corresponding candidate

coefficients for left and right images, respectively. According to Figure 8 aforementioned coefficients can be mapped as $(\Xi_{n_r} : R1)$, $(\acute{\Xi}_{n_r} : R2)$, $(\Xi_{n_c} : C1)$ and $(\acute{\Xi}_{n_c} : C2)$. To calculate *ADD*, we can define the expression as

$$D_{A_{\Xi_{c_j}}} = \left[\left|\frac{d_{\Xi_{c_j}} - d_{\Xi_r}}{d_{\Xi_{c_j}} + d_{\Xi_r}}\right|\right]_m \tag{25}$$

Where $|.|$ represents the absolute values and $D_{A_{\Xi_{c_j}}}$ defines *ADD* for jth candidate coefficient of second image corresponding to d_{Ξ_r} from first image. Process of 25 is averaged over m times repetitions to minimize any bias introduced by the coefficients belonging to any particular area of image as well as involvement of any wrong candidate pair that could have been assigned the tag of reference coefficients as $\Xi_{n_{r_j}}$ and $\acute{\Xi}_{n_{r_j}}$. Similarly, *RDD* can be defined by the following expression

$$D_{R_{\Xi_{c_j}}} = \left[\left|\frac{d_{\Xi_c r} - d_{\acute{\Xi}_{cr_j}}}{d_{\Xi_{cr_j}} + d_{\acute{\Xi}_{cr_j}}}\right|\right]_n \tag{26}$$

Similar to *ADD* (25), (26) is repeated n times. Finally, *RSD* is calculated by defining relative slopes between candidate and reference coefficients as

$$S_{\Xi_{c_j}} = \left[\left|\frac{s_{\Xi_c r} - s_{\acute{\Xi}_{cr_j}}}{s_{\Xi_c r} + s_{\acute{\Xi}_{cr_j}}}\right|\right]_n \tag{27}$$

The term $(.)_n$ defines the average over n repetitions for each jth candidate coefficient. Employing expressions (25), (26) and (27), a generalized expression of geometric refinement is defined for each jth candidate correspondence by weighting the established correlation score from (23) as

$$F_{\Xi_{c_j}} = \frac{\aleph_{\Xi_{c_j}}}{3}\left(e^{-D_{A_{\Xi_{c_j}}}} + e^{-D_{R_{\Xi_{c_j}}}} + e^{-S_{\Xi_{c_j}}}\right) \tag{28}$$

It is obvious from expression (28), $\acute{\Xi}_{n_{c_j}}$ with highest score will be the one having closest geometrical topology with respect to the reference coefficients Ξ_{n_r} and $\acute{\Xi}_{n_r}$. For instance, for an optimal correspondence between Ξ_{n_c} and $\acute{\Xi}_{n_c}$, expression $(F_{\Xi_{c_j}})$ will boil down to simple correlation score $\aleph_{\Xi_{c_j}}$ from (23) as the term $\left[\frac{1}{3}\left(e^{-D_{A_{\Xi_{c_j}}}} + e^{-D_{R_{\Xi_{c_j}}}} + e^{-S_{\Xi_{c_j}}}\right)\right]$ will become 1.

5. Finer levels correspondence estimation

Correspondence estimation process (section-4) at coarsest wavelet transformation level, i.e. level N, produces set of optimal correspondences between coefficients belonging to first and second images. These correspondences are then projected to finer level, i.e. level $N-1$, where a local search is performed to authenticate correspondences, established at coarsest level N, as well as to estimate new ones. Referring back to section (3.1) and (3.2), transform modulus maxima that belongs to lower frequency components disappear at higher transformation levels. Authentication of correspondences at $N-1$ level, using the information of N level correspondences, provides a structured ground to constraint the search of new coefficient correspondences to local search regions. This local search eliminates the need of computationally expansive geometric refinements leaving the processes

Basis	r	$[C_s, C_w]$	A_p	Orth	Shape
Haar Haar (1910)	1	[2 , 2]	1	o	s
D-4 Daubechies (1988)	1	[4 , 4]	2	o	as
D-8 I. Daubechies & Lagarias (1992)	1	[8 , 8]	4	o	as
BI9 Strela (1998)	1	[9 , 7]	4	bo	s
BI7 Strela (1998)	1	[7 , 9]	4	bo	s
BI5 Strela (1998)	1	[5 , 3]	2	bo	s
BI3 Strela (1998)	1	[3 , 5]	2	bo	s
GHM Gernimo et al. (1994)	2	[4 , 4]	2	o	s
CL Chui & Lian (1996); Chui (1992)	2	[3 , 3]	2	o	s
SA4 Strela (1996)	2	[4 , 4]	1	o	s
BIH52S Strela (1998)	2	[5 , 3]	2	bo	s
BIH32S Strela (1998)	2	[3 , 5]	4	bo	s
BIH54N Strela (1998)	2	[5 , 3]	4	bo	s
MW1 A. Bhatti (2002); Ozkaramanli et al. (2002)	3	[6 , 6]	2	o	s
MW2 A. Bhatti (2002); Ozkaramanli et al. (2002)	3	[6 , 6]	3	o	as
MW3 A. Bhatti (2002); Ozkaramanli et al. (2002)	3	[8 , 8]	4	o	s

Table 1. Employed wavelets and multiwavelets bases with embedded attributes

of sections (4.1) and (4.2) necessary and sufficient to achieve desired quality of correspondence estimation. This procedure can be considered as iterative optimization process Daubechies (1992).

6. Analysis of the effect of different wavelet and multi-wavelet bases

To address the influence of wavelets and multiwavelets bases on the quality of correspondence estimation, 16 wavelets and multiwavelets bases are employed. These bases are carefully chosen to cover range of properties such as orthogonality, bi-orthogonality, symmetry, asymmetry, multiplicity and approximation order Asim Bhatti & Zheng (2003) as presented in Table 1.

Referring to Table 1, parameters *o* and *bo*, in the Orth column represents orthogonality and bi-orthogonality ,respectively, of the bases. *s* and *as* are the shape parameters, representing symmetric and asymmetric, whereas *r* defines the multiplicity. It is obvious from the Table 1 that scalar wavelets possess unit multiplicity, i.e. one scaling function and one wavelet. The coefficients related to wavelets and multiwavelets bases presented, in Table 1, can be found in Bhatti (2009). Statistical analysis is performed using *root mean squared error* (RMS) and *percentage of bad discrete pixel disparities* (BPD), employed from D. Scharstein & Szeliski (n.d.), for qualitative measure of the correspondences estimation. Disparity maps generated using

Basis	r	$[C_s, C_w]$	A_p	Orth	Shape
CL Chui & Lian (1996); Chui (1992)	2	[3 , 3]	2	o	s
MW2 A. Bhatti (2002); Ozkaramanli et al. (2002)	3	[6 , 6]	3	o	as
MW3 A. Bhatti (2002); Ozkaramanli et al. (2002)	3	[8 , 8]	4	o	s

Table 2. Selected multiwavelets basis

estimated correspondences are compared with the ground truth disparity maps.

$$RMS = \sqrt{\frac{1}{N}\sum_{x,\,y} |d_E(x,y) - d_G(x,y)|^2} \tag{29}$$

and

$$PBD = \frac{1}{N}\sum_{(x,y)} |d_E(x,y) - d_G(x,y)| > \xi \tag{30}$$

where d_E and d_G are the estimated and ground truth disparity maps, N is the total number of discrete disparity values in the disparity map whereas ξ represents the disparity error tolerance, taken as 1. In other words any difference greater than 1 between ground truth disparity maps and the estimated disparity is considered as bad discrete disparity. These statistics are related to the images *Map*, *Bull*, *Teddy*, *Cones* and *Venus*, taken from D. Scharstein & Szeliski (n.d.). Referring to visual representation in Figures 9 and 10, a distinguished higher performance of multi-wavelets bases can be observed throughout the set of employed images. This statistical behavior of the estimated data strengthens earlier established understanding about the superior performance of multiwavelets bases over the scalar ones Strela (1996). Their success stems from the fact that they can simultaneously posses the good properties of orthogonality, symmetry, high approximation order and short support G. Strang & Strela (1995; 1994), which is not possible in the scalar wavelets case G. Strang & Strela (1994); Daubechies (1992). Out of 9 multiwavelets bases, CL, MW2 and MW3 has outperformed rest of the bases with major contribution from *MW2*. Analyzing embedded attributes of these multiwavelets bases, separated in Table 2, we see a clear pattern of commonality in terms of multiplicity and orthogonality contributing into the higher performance of these multiwavelets bases. Although it is hard to visualize a clear correlation pattern, explicitly, between the attributes of presented wavelets and multiwavelets bases and the quality of correspondences, however we would initiate a short discussion to address some possible effects of these attributes to correspondence estimation problem as:

Orthogonality dictates that coefficients in subspaces of D_{sj} and A_s and linearly independent, as in Figures 1 and 2, and their direct sum produces the subspace A_{s-1} (11). In classical signal processing terms, subspace A_s contains lower frequency components of the input signal whereas D_{sj} contains higher frequency components depending on the approximation order of the scaling functions. Perfect separation, due to orthogonality, between lower and higher frequency components and into scales and subspaces provides a sparse representation of high value features that are easier to track.

Multiplicity influences the size of search space by producing r^2 subspaces of coefficients (sections (3.1) and (4.2)). Consequently producing expanded search space to establish and authenticate coefficient correspondences. In general, multiplicity and orthogonality

together, influences the separation of signal components into distinct subspaces making it easier to establish robust correspondences. This leads to the notion of scale-space representation Chui & Lian (1996). In this particular study employing wavelet transform modulus maxima coefficients with higher multiplicity and orthogonality ensured the involvement of high profile features at different scales and spaces making the algorithm robust and resistant to errors.

Approximation order defines the approximation capabilities of the scaling functions. Multiwavelets bases are set to have approximation order p if a linear combination of the scaling functions can produce polynomials up to degree $p-1$. In other words, polynomials up degree $p-1$ are in linear span of scaling space spanned by the shifts of scaling functions $\phi_0(t), \phi_1(t), \cdots \phi_{r-1}(t)$. This means polynomials up to degree 1, i.e. $f = t$ are in the linear span of multiscaling functions of D4, BI5, BI3, GHM, BIH52S and MW1 (Table 1). Similarly, $f = t^2$ and $f = t^3$ polynomials are in the linear span of MW2 and MW3 bases, respectively. In the context of image processing, polynomials can be represented by the gradient intensity change. Single color without any intensity variations can be represented as polynomial of degree zero ($f = t^0 = 1$), that is a constant function. A constant intensity variation would refer to polynomial of degree 1, i.e. $f = t$. Based on this understanding of approximation order, we can say, higher approximation order leads to higher order modulus maxima coefficients in $D_s j$ subspaces (Figure 1 and 2). In other words higher approximation order ensures the separation of higher order features or modulus maxima coefficients from lower order features, consequently allowing the algorithm to focus on global aspects rather than getting stuck into local minima introduced by low value coefficients. Considering, very high approximation order could also result in filtering the important coefficients into the approximation space rather than detail space, which is used for correspondence estimation, it can be argued; what is the optimal approximation order? It is very hard to conclude at this stage however our future work involves the extension of statistical analysis utilizing bigger data base of images and multiwavelets bases.

7. Conclusion

In this presentation we have tried to initiate a discussion about the potential of multiwavelets bases into the domain of robust correspondence estimation. We have addressed some embedded attributes of wavelets and multiwavelets bases that could play a key role in establishing highly robust correspondences between two and more views. Seven wavelets and nine multiwavelets bases were employed covering a range of well known attributes including orthogonality, approximations order, support and shape. For statistical performance analysis, five well known images with diverse range of intensity complexities were employed. In addition, a novel and robust correspondence estimation algorithm is presented to provide a test bed to exploit the potential of wavelets and multiwavelets bases. The proposed algorithm uses multi-resolution analysis to estimate correspondences. The translation invariant multiwavelets transform modulus maxima (WTMM) are used as matching features. To keep the whole matching process consistent and resistant to errors an optimized selection criterion is introduced involving the contribution of probabilistic weighted normalized correlation and geometric refinement. Probabilistic weighting involves the contribution of more than one search spaces, whereas geometric refinement addresses the problem of geometric distortion between the perspective views. Moreover, beside that comprehensive selection criterion

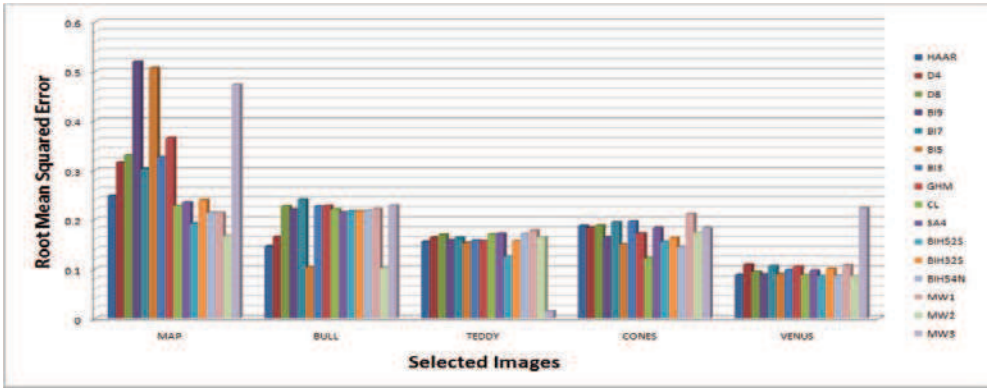


Fig. 9. Root Mean Square Error (RMS) for number of images

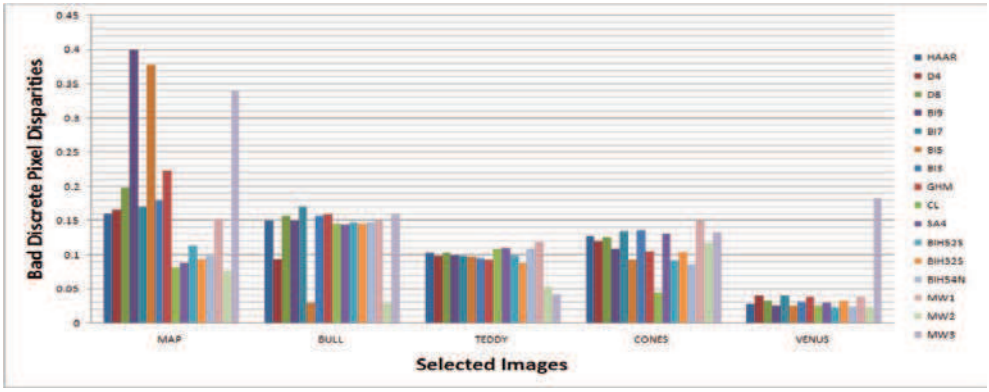


Fig. 10. Percentage of Bad Pixel Disparity (BPD) for number of images

the whole matching process is constrained to uniqueness, continuity and smoothness. We are currently in the process of expanding the experimental envelope and would hope to present clearer picture of correlations between the embedded attributes of the bases and correspondence problem in future presentations.

8. References

A. Bhatti, ., H. . (2002). M-band multiwavelets from spline super functions with approximation order, *in* IEEE (ed.), *International Conference on Acoustics Speech and Signal Processing, (ICASSP 2002)*, Vol. 4, IEEE, pp. 4169–4172.

A. Bhatti, ., S. N. (2008). Stereo correspondence estimation using multiwavelets scale-space representation based multiresolution analysis, *Cybernetic and Systems* 39(6): 641–665.

A. Mehmood, ., A. S. (2001). Digital reconstruction of buddhist historical sites (6th b.c-2nd a.d) at taxila, pakistan (unesco, world heritage site), *in* IEEE (ed.), *Proceedings of the Seventh International Conference on Virtual Systems and Multimedia (VSMM01)*.

A. Witkin, D. T. & Kass, M. (1987). Signal matching through scale space, *Int. J. of Computer Vision* 1: 133–144.

Abhir Bhalerao, . & Wilson, R. (2001). A fourier approach to 3d local feature estimation from volume data, *Proc. of British Machine Vision Conference*, Vol. 2, pp. 461–470.

Alejandro Gallegos-Hernandez, Francisco J. Ruiz-Sanchez, J. R. V.-C. (2002). 2d automated visual inspection system for the remote quality control of smd assembly, *in* IEEE (ed.), *28th Annual Conference of Industrial Electronics Society (IECON 02)*, Vol. 3, IEEE, pp. 2219–2224.

Asim Bhatti, S. N. & Zheng, H. (2003). Disparity estimation using ti multi-wavelet transform, *Fourth International Conference on Intelligent Technologies (Intech'03)*.

B. Chebaro, A. Crouzil, L. M.-P. & Castan, S. (1993). Fusion of the stereoscopic and temporal matching results by an algorithm of coherence control and conflicts management, *Int. Conf. on Computer Analysis of Images and Patterns*, pp. 486–493.

Baker, H. & Binford, T. (1981). Depth from edge and intensity based stereo, *Int. Joint Conf. on Artificial Intelligence*, Vancouver, Canada, pp. 631–636.

Bhatti, A. (2009). *Stereo Vision and Wavelets/Multiwavelets Analysis*, Lambert Academic Publishing.

C. Baillard, C. Schmid, A. Z. & Fitzgibbon, A. (1999). Automatic line matching and 3d reconstruction of buildings from multiple views, p. 12.

C. L. Zitnick, T. K. (2000). A cooperative algorithm for stereo matching and occlusion detection, *IEEE PAMI* 22(7): 675–684.

Chui, C. K. (1992). *Wavelets: A tutorial in theory and applications*, Acadmic press.

Chui, C. K. & Lian, J. (1996). A study of orthonormal multi-wavelets, *J. Applied Numerical Math* 20(3): 273–298.

Cohen, I., Raz, S. & Malah, D. (1998). Adaptive time-frequency distributions via the shiftinvariant wavelet packet decomposition, *Proc. of the 4th IEEE-SP Int. Symposium on Time-Frequency and Time-Scale Analysis*, Pittsburgh, Pennsylvania.

Coifman, R. R. & Donoho, D. L. (1995). Translation-invariant de-noising, *Wavelet and Statistics, Lecture Notes in Statistics*, a. antoniadis and g. oppenheim, ed. edn, Springer-Verlag, pp. 125–150.

D. Scharstein, . & Szeliski, R. (n.d.). www.middlebury.edu/stereo.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.* 41: 309–996.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, Philadelphia.

David Capel, A. Z. (2003). Computer vision applied to super-resolution, p. 10.

Esteban, C. H. (2004). *Stereo and Silhouette Fusion for 3D Object Modeling from Uncalibrated Images Under Circular Motion*, Ph.d. dissertation.

G. Strang, . & Strela, V. (1994). Orthogonal multiwavelets with vanishing moments, *J. Optical Eng.* 33: 2104–2107.

G. Strang, . & Strela, V. (1995). Short wavelets and matrix dilation equations, *IEEE Trans. on SP* 43: 108–115.

Gernimo, J., Hardin, D. & Massopust, P. R. (1994). Fractal functions and wavelet expansions based on several functions, *J. Approx. Theory* 78: 373–401.

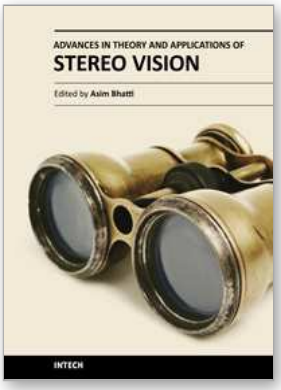
Haar, A. (1910). Zur theorie der orthogonalen funktionen-systeme, *Math* 69: 331–371.

I. Daubechies, . & Lagarias, J. (1992). Two-scale difference equations i. existence and global regularity of solutions, *SIAM J. Math. Anal.* 22: 1388–1410.

J. C. Olive, ., J. D. & Boulin, C. (1994). Automatic registration of images by a wavelet-based multiresolution approach, *SPIE*, Vol. 2569, pp. 234–244.

J. Magarey, . & Kingsbury, N. (1998). Motion estimation using a complex-valued wavelet

transform, *IEEE Transections on signal proceessings* 46(4): 1069–1084.
J. Margary, . & dick, A. (1998). Multiresolution stereo image matching using complex wavelets, *Proc. 14th Int. Conf. on Pattern Recognition (ICPR)*, Vol. 1, pp. 4–7.
J. R. Bergen, P. Anandan, K. J. H. & Hingorani, R. (1992). Hierarchical model-based motion estimation, *ECCV*, pp. 237–252.
L. Di Stefano, ., M. M. S. M. G. N. (2004). A fast area-based stereo matching algorithm, *Image and Vision Computing* 22(12): 938–1005.
M. Unser, . & Aldroubi, A. (1993). A multiresolution image registration procedure using spline pyramids, *SPIE Mathematical Imaging* 2034: 160–170.
Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. PAMI* (11): 674–693.
Mallat, S. (1991). Zero-crossings of a wavelet transform,, *IEEE Transactions on Information Theory* 37: 1019–1033.
Mallat, S. (1999). *A Wavelet Tour of Signal Processing*, Vol. 2nd edition, Academic Press.
Medioni, G. & Nevatia, R. (1985). Segment based stereo matching, *Computer Vision, Graphics and Image Processing* 31(1): 2–18.
Ozkaramanli, H., Bhatti, A. & Bilgehan, B. (2002). Multi wavelets from b-spline super functions with approximation order, *International Journal of Signal Processing, Elsevier Science* 82(8): 1029–1046.
Pan, H.-P. (1996a). General stereo matching using symmetric complex wavelets, *Wavelet Applications in Signal and Image Processing* 2825.
Pan, H.-P. (1996b). Uniform full-information image matching using complex conjugate wavelet pyramids, *XVIII ISPRS Congress, International Archives of Photogrammetry and Remote Sensing*, Vol. 31.
Q'ingxiong Yang, ., L. W. R. Y. H. S. & Nister, D. (2006). Stereo matching with color-weighted correlation, hierarchical belief propagation and occlusion handling, *CVPR*, Vol. 2, pp. 2347–2354.
S. Mallat, . & Zhang, S. (1993). Matching pursuits with time-frequency dictionaries, *IEEE Transactions on Signal Processing* 41(12): 3397–3415.
Scharstein, D. & Szeliski, R. (1998). Stereo matching with nonlinear diffusion, *Int. J. of Computer Vision* 28(2): 155–174.
Shi, F., Hughes, N. R. & Robert, G. (2001). Ssd matching using shift-invariant wavelet transform, *British Machine Vision Conference*, pp. 113–122.
Siebert, A. (1998). A linear shift invariant multiscale transform, *in* IEEE (ed.), *International Conference on Image Processing*.
Strela, V. (1996). *Multiwavelets: Theory and Applications*, Phd.
Strela, V. (1998). *A note on construction of Biorthogonal Multi-scaling functions*, Contemporary Mathematics, A. Aaldoubi and E. B. Lin (eds.),.
T. S. Huang, ., A. N. N. (1994). Motion and structure from feature correspondences: A review, *in* IEEE (ed.), *Proc. of the IEEE*, Vol. 82, pp. 252–268.
Wang, L., Liao, M., Gong, M., Yang, R. & Nister, D. (2006). High-quality real-time stereo using adaptive cost aggregation and dynamic programming, *International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06)*, pp. 798–805.
X. Zhou, . & Dorrer, E. (1994). Automatic image-matching algorithm based on wavelet decomposition, *IAPRS* 30(1): 951–960.
Y. Boykov, ., O. V. & Zabih, R. (2001). Fast approximate energy minimization via graph cuts, *IEEE Transections of Pattern Analysis and Machine Intelligence* 23(11): 1222–1239.



Advances in Theory and Applications of Stereo Vision
Edited by Dr Asim Bhatti

ISBN 978-953-307-516-7
Hard cover, 352 pages
Publisher InTech
Published online 08, January, 2011
Published in print edition January, 2011

The book presents a wide range of innovative research ideas and current trends in stereo vision. The topics covered in this book encapsulate research trends from fundamental theoretical aspects of robust stereo correspondence estimation to the establishment of novel and robust algorithms as well as applications in a wide range of disciplines. Particularly interesting theoretical trends presented in this book involve the exploitation of the evolutionary approach, wavelets and multiwavelet theories, Markov random fields and fuzzy sets in addressing the correspondence estimation problem. Novel algorithms utilizing inspiration from biological systems (such as the silicon retina imager and fish eye) and nature (through the exploitation of the refractive index of liquids) make this book an interesting compilation of current research ideas.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Asim Bhatti and Saeid Nahavandi (2011). Impact of Wavelets and Multiwavelets Bases on Stereo Correspondence Estimation Problem, Advances in Theory and Applications of Stereo Vision, Dr Asim Bhatti (Ed.), ISBN: 978-953-307-516-7, InTech, Available from: http://www.intechopen.com/books/advances-in-theory-and-applications-of-stereo-vision/impact-of-wavelets-and-multiwavelets-bases-on-stereo-correspondence-estimation-problem

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821