

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Wavelet Transform Based Motion Estimation and Compensation for Video Coding

Najib Ben Aoun, Maher El'arbi and Chokri Ben Amar  
*REsearch Groups on Intelligent Machines (REGIM)*  
*University of Sfax, National Engineering School of Sfax (ENIS)*  
*Tunisia*

## 1. Introduction

With the big evolution in the quantity of video data issued from an increased number of video applications over networks such as the videophone, the videoconferencing, and multimedia devices such as the personal digital assistants and the high-definition cameras, it has become crucial to reduce the quantity of video data which will be stored or transmitted. In fact, since the capacity of the storage Medias has become high and sufficient, the data storage problem was resolved but the transmission of the data remains an important problem especially with the limited channel bandwidth.

Actually, the necessity of the development of an efficient video coding method has made video compression a fundamental task for video-based digital communications. Video compression reduces the quantity of video data by eliminating the spatial and the temporal redundancy. Spatial compression is done by transforming video frames and representing them otherwise using the spatial correlation between frames pixels. In the other side, motion estimation and compensation are employed in video coding systems to remove temporal redundancy while keeping a high visual quality. They are the most important parts of the video coding process since they require the most computational power and the biggest consumption in resources and bandwidth. Therefore, many techniques have been developed to estimate motion between successive frames.

Motion estimation and compensation (ME/MC) was conducted in many domains such as spatial domain by applying it directly on images pixels without any transformation, the frequency domain by driving it on the Discrete Cosine Transform (DCT) or the Discrete Fourier Transform (DFT) coefficients. It can be also done in the multiresolution domain by running it on the Discrete Wavelet Transform (DWT) coefficients. However, giving the promising performances of the multiresolution analysis especially the DWT which provides a multiresolution expression of the signal with localization in both space and frequency, many methods have been developed to construct a wavelet based video coding system (Shenolikar, 2009) and the DWT was integrated in new coding standards such as JPEG2000, MPEG-4, and H.264. Furthermore, recently, many motion estimation and compensation systems (BEN AOUN, 2010) have also confirmed that the DWT is the most suitable and the most efficient domain that gives efficient and precise motion estimation.

For this, we have developed a block based ME/MC method in the wavelet domain. Our method exploits the benefits of DWT and the hierarchical relationship between its subbands

(Quadtree) to drive ME/MC on wavelet coefficients, especially in the low frequency subband where we find the most significant visual information. This method is consolidated by several techniques to ameliorate the results. With this method, we have achieved good results in terms of prediction quality, compression performance and computational complexity.

The goal of this chapter is to introduce new motion estimation and compensation system based on the DWT which has given better and superior results compared with others systems conducted in spatial or frequency domains. Our system is also based on the Block Matching Algorithm (BMA) which is the simplest, the most efficient and the most popular technique for motion estimation and compensation. Additional techniques are introduced to accelerate the estimation process and improve the prediction quality. In Section 2, we introduce the multiresolution domains and especially the DWT as a multiresolution description for the image which has proved its efficiency for ME/MC. Section 3 presents the motion estimation principle and methods focusing on the DWT based systems. Section 4 describes our DWT and BMA based proposed method. In Section 5, we will introduce some supplementary techniques which have been developed to improve our method and give the main causes which have made of them crucial parts for an efficient motion estimation system. In Section 6, we evaluate our method and compare it to others conventional methods conducted in different domains. This will prove that our method outperforms conventional method in many terms. Finally, Section 7 summarizes the key findings and suggests future research possibilities. We should mention that, along this chapter, when we say motion estimation, we imply implicitly the motion compensation.

## 2. Wavelet transform domain

The wavelet transform, as a multiresolution domain that hybrid the frequency and the spatial domain, has proved that it is a very appropriate and reliable domain for a powerful motion estimation and compensation. For this, we have been encouraged to study and exploit it, and more precisely the DWT, in our motion estimation system.

The DWT consists on applying hierarchically low-pass (L) and high-pass (H) filters after decimation (sub-sampling the image on two parts). This procedure is repeated until reaching a prefixed level. Figure.1 shows the decomposition of an image with DWT. In this example there are two levels of DWT decomposition.

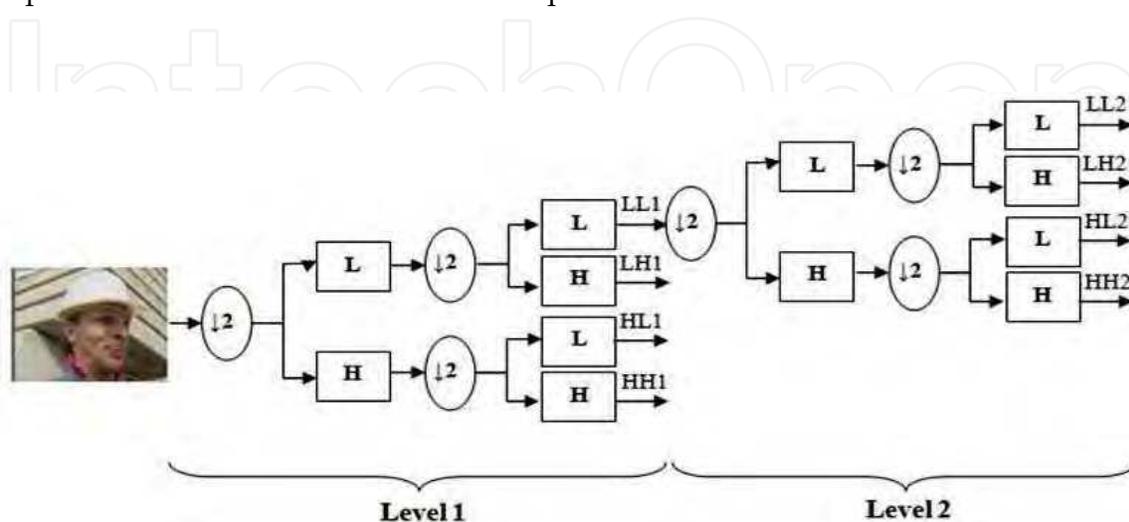


Fig. 1. DWT decomposition (2 levels)

The DWT decomposes the image into different subbands, as shown in Figure.2, aiming to isolate the high frequencies that are not interesting to the human eyes. So, we will have the most important information concentrated in the subband LL of the highest level called also DWT approximation (LL3 in the Fig.2).

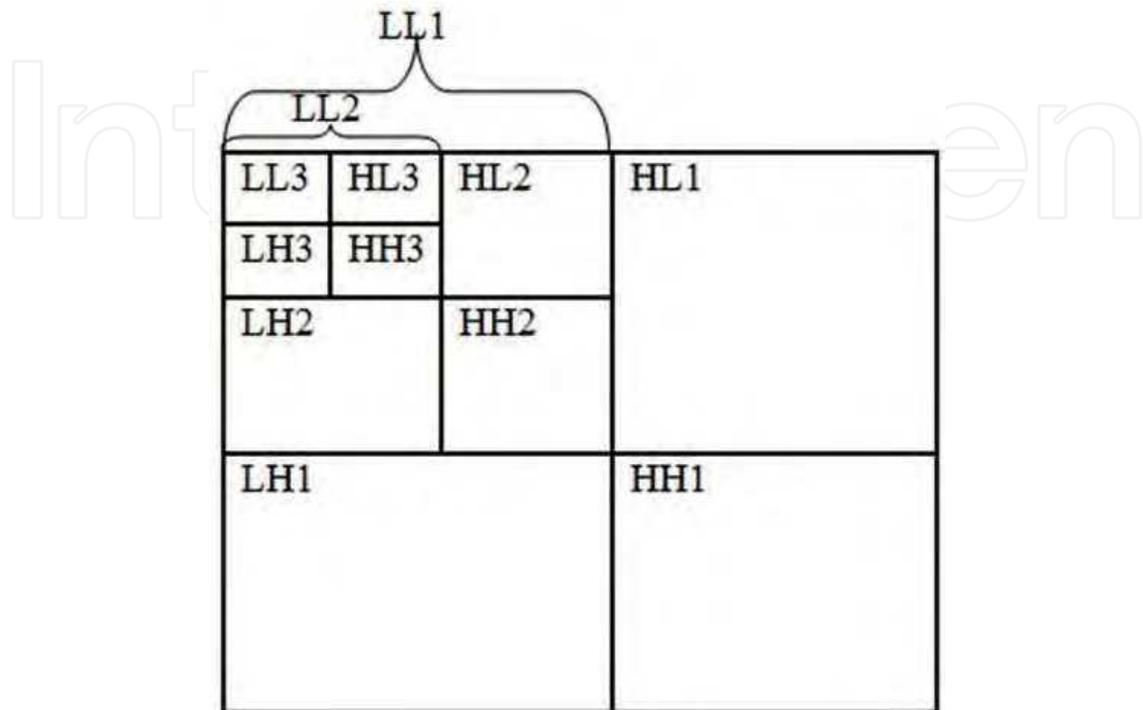


Fig. 2. Different DWT subbands (3 levels)

The Figure.3 below shows the decomposition of the Foreman image into three level of DWT. This example illustrates clearly that the DWT approximation presents the most significant information that the human eyes are sensible to. The others subbands (DWT details) give the high frequencies existing in the image along different orientations.



(a) Original image

(b) DWT decomposition

Fig. 3. Three levels DWT decomposition applied to Foreman

The fact that the DWT approximation contains the most of the information issued from the original image was encouraging to benefit of this DWT propriety. For this, the motion estimation was conducted principally in this subband which accelerates the motion estimation process.

The discrete wavelet transform (DWT) as a powerful tool for signal processing has found its application in many areas of research. Image compression is still one of the most successful applications in which the DWT has been applied. So, it is natural that researchers are interested in creating a DWT based new technologies for video compression and motion estimation (Kutil, 2003).

### 3. Motion estimation and compensation

With the continuous growth in the volume of video data in the multimedia databases, it has become crucial to reduce the quantity of the data to be transmitted and stored by video compression and coding. That is why, motion estimation is introduced as a solution to reduce the quantity of data by eliminating the temporal redundancy between adjacent frames in an image sequence. ME/MC are the fundamental parts of video coding systems and form the core of many video processing applications. Motion estimation eliminates temporal redundancy from video by exploiting the temporal correlation between successive frames, so that it reduces the amount of data to be transmitted or stored while maintaining sufficient data quality. However, ME extracts temporal motion information from video sequences, while MC uses this motion information for efficient interframe coding.

Motion estimation process serves to predict motion between two successive frames and produce the motion vectors (MVs) which represent the displacements between these two frames. Consequently, instead of transmitting two frames, we will send only one frame which is the reference frame, the motion vectors and the residue which is the difference between the current frame and the reconstructed frame by motion compensation. So, the MVs and the prediction error are transmitted instead of the frame itself. With this process, the encoder will have sufficient information to faithfully reproduce the frame sequence. The combination of the motion estimation and motion compensation is a key part of the video coding.

There are many methods to achieve ME/MC. In fact, They can be divided into two classes: the statistical methods, the differentials methods as indirect methods (applied to image features) and the optical flow, and the block based method as direct ones (applied to image pixels). Block matching algorithm (Gharavi, 1990) is an effective and popular technique for block based motion estimation. It has been widely adopted in various video coding standards and highly desirable since it maintains an acceptable prediction errors.

Block-based motion estimation is most used method because of its simplicity and performances, which made it the standard approach in the video coding systems. The procedure of BMA is to divide the frames into a block of  $N \times N$  pixels, to match every block of the current frame (CF) with his most similar block inside a research window in the reference frame (RF) and to generate the motion vector. Consequently, for this method, the most important parameters here are the size of the block  $N$  and the size of the search

window P. However, the block matching is based on minimizing a criterion like the Mean Absolute Error (MAD) or the Mean Square Error (MSE) which is the most common block distortion measure for matching two blocks and it provides more accurate block matching. The MV will be applicable to every pixels of the same block which reduces the computational requirement.

To identify the best corresponding block, the simplest way is to evaluate every block in the reference frame (exhaustive search, ES). But, although this method finds generally the appropriate block, it consumes a high computation time. Hence, others fast searching strategies (Barjatya, 2004) have been developed where search is done in a particular order. There are the Three Step Search (TSS), the Simple and Efficient Search (SES), the Four Step Search (4SS), the Adaptive Rood Pattern Search (ARPS) and the Diamond Search (DS) which has proved to be the best searching strategies coming close to the ES results. So, the DS was improved in many variants such as the Cross DS (CDS), the Small CDS (SCDS) and the New CDS (NCDS).

In conventional coding systems such as H.261 and MPEG-1/2, BMA is conducted directly on frame which needs a large computing power. That is why many studies have been made and proved that it is better to transform the frame before executing the ME techniques. However, with the development of new video coding standards, wavelets have received an important interest since it has shown good and effective results. The main idea behind wavelet is to generate a space-frequency representation focusing only on the spatial frequencies that are most significant to the human eye. This wavelet decomposition is a reversible procedure which is performed by successive approximations of the initial information (original frame). This process, will improve the coding efficiency since the wavelet coefficients are much correlated and this representation reduces the blocking effects especially in the edges.

Initially, the DWT was used to encode the MVs and the estimation errors after conducting the motion estimation in the spatial or the frequency domains (Figure.4.a). Thereafter, given that the DWT is a spatial-frequency representation for the image that concentrates the most important information in one subband (DWT approximation subband) and since the different DWT subbands are hierarchically correlated, the DWT was used as a domain to conduce the motion estimation and it has shown a great success.



(a) Conventional ME + DWT based MVs and ME errors encoding

(b) Motion estimation in the wavelet domain

Fig. 4. Video coders based on DWT

Exploiting the hierarchical relationship between the wavelet coefficients of the different subbands in different levels, different hierarchical ME methods were developed which are adapted to the wavelet transformation. The hierarchical relationship gives that every wavelet coefficients has four descendants in the lower level of the DWT. The motion estimation is conducted hierarchically so that it is calculated firstly in one of the DWT level and it is corrected with the estimation obtained, thereafter, at the others levels.

In fact, there are two main ME categories of approaches for DWT based: forward and backward approaches. The forward approach consists on conducting the ME in the DWT details subbands of the low level and using it to determine the motion in the higher level subbands (coarse-to-fine). Researchers like Meyer and al (Meyer, 1997) have followed the forward approach to propose a ME method with a new pyramid structure. They have taken the aliasing effect, caused by the BMA used, into consideration and build a ME system given a good perceptual quality after MC. Also, P.Y Cheng and al (Cheng, 1995) has proposed a multiscale forward ME working on the DWT coefficients. They have built a new pyramidal structure overcoming the shift variant problem of the DWT.

Nosratinia and Orchard (Nosratinia, 1995) were the first researchers who developed a ME system based on DWT following a backward approach (coarse-to-fine) where they estimated the motion in the finest DWT resolution (higher level) and then progressively refined the ME by incorporating the finer level. Furthermore, Conklin and Hemami (Conklin, 1997) have proved the superiority of the backward ME approach over the forward one in terms of compression rate and visual quality after compensation. This is what encourages more recent researchers (Lundmark, 2000; Yuan, 2002) to follow this approach in their ME systems.

The effectiveness of the BMA and the suitability of the DWT in the video coding, have led us to develop a block matching based motion estimation method in the wavelet domain.

#### **4. Our proposed method**

The motion estimation and compensation are the most important parts in the video coding process. For this, many works have focused on these video coding parts aiming to improve them. But, the results reached still insufficient especially for the real time applications. That is what encourages us to work on these parts and improve them.

The Block Matching Algorithm still one of the most efficient and the most used method for motion estimation since it works directly on image pixels and it accelerates the estimation process by working on pixels blocks. This method suffers like all others methods from some problems such as the Blocking effect (discontinuity across block boundary) in the predicted image. But, we have overcome this problem in our system with several motion estimation improvement techniques.

Thanks to its proprieties and its suitability as a domain to apply motion estimation and compensation, the multiresolution domain has been adopted in our system to conduct the motion estimation directly on its coefficients. Among the method to obtain a multiresolution representation for the image, we have the DWT that has proved its efficiency not only for data compression but also for motion estimation.

The proposed method makes use of the wavelet properties to apply the motion estimation directly in the wavelet coefficients. We have adopted the fine-to-coarse motion estimation strategy which has shown its success by many previous works. After applying the DWT on both CF and RF, the motion is estimated firstly between the DWT approximations of the two images. So, we have provided a better estimation since the approximation contains the most visual information. The motion vectors of the approximation are directly calculated. We have exploited that every DWT coefficient has four descendants in the lower DWT level (Quadtree structure). So, the motion vectors of the details subbands are deducted using the hierarchical relationship that exists between the DWT subbands as shown in Figure 5. We compute the motion vectors of the details subbands following this formula:

$$V_{i,j} = 2^{L-i} V_{L,1}(x,y) + \delta_{i,j} \quad (1)$$

Working on a three level DWT ( $L=3$ ), we will have  $i=\{1, 2, 3\}$  which is the level,  $j=\{1, 2, 3, 4\}$  representing the subband number,  $V_{i,j}(x, y)$  is the motion vector for the subband “j” at the level “i” and  $\delta_{i,j}$  is the refinement factor (equal to 0 if “i” is equal to L). The displacement of every subband block is the double of the displacement of the same subband block in the lower DWT level where we add to it a refinement factor  $\delta_{i,j}$  which correct the estimation error as given in the equation and presented in the Figure above.

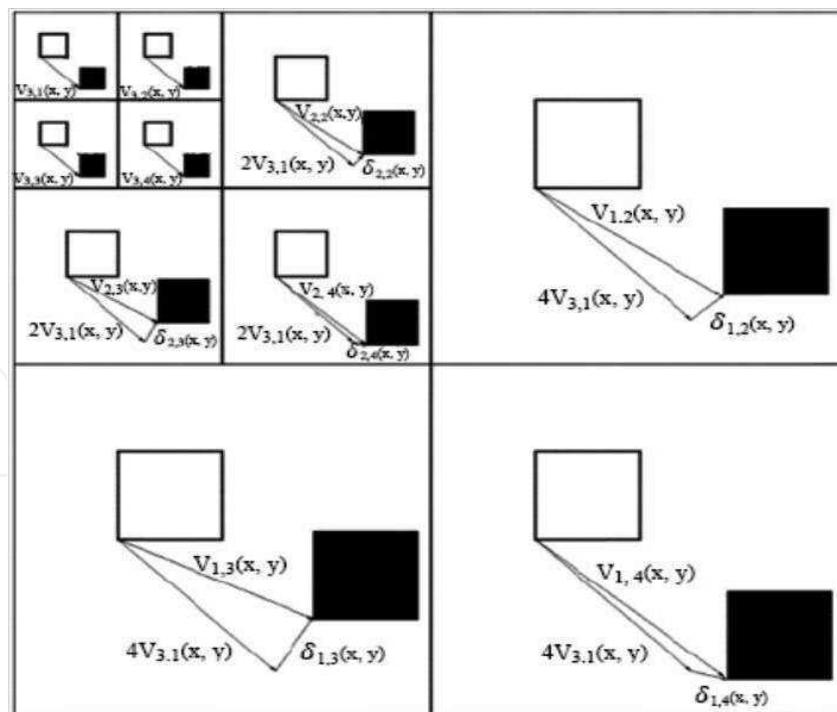


Fig. 5. DWT subbands motion vectors representation ( $L=3$ )

Moreover, by predicting the motion only in the approximation which has a small size compared to the original frame and contains the most significant information, not only the

computation requirement is highly reduced and the compression ratio is increasing, but also our method maintains a good prediction quality.

The BMA is an efficient method for motion estimation which encourages us to use it in our multiresolution based method. Unfortunately, despite their encouraging proprieties and their promising results, the BMA and DWT suffer from some problems. For this, a several improvement techniques have been implemented to surmount these problems and make our method more robust giving best results.

## 5. Additional improvement techniques

Despite that it outperforms the conventional motion estimation methods, our proposed DWT based method still having some problems. As we have mentioned before, the DWT representation suffers from the problem of aliasing and the fact that it is a shift variant transformation. Moreover, the block based motion estimation causes the blocking effect which gives a discontinuity in the block boundaries of the predicted image. That is what drives us to develop some additional techniques to overcome these problems.

These techniques make the motion estimation process more precise and more rapid by detecting the moving zones and limiting the estimation operation to it, adding a sub-pixel precision to the motion vector computing, applying the motion estimation to a shifting variants of the original image aiming to make the estimation a shift invariant operation, overlapping the frame blocks to correct the motion vector by their neighbouring vectors and finally, refining the prediction by changing the block size and re-predicting the blocks which are falsely predicted. In this section we will describe these techniques as well as the causes that conduct us to implement them.

### 5.1 Moving zones detection

To accelerate the ME process, we have focused on the image zones where there are movements so that we will conduct the motion estimation only in them. Many techniques have been developed to detect the moving zones in an image. The simplest method is to subtract the background by comparing every image pixels displacement to a prefixed threshold and assuming that it belongs to the foreground if it is superior to this threshold and it is declared as a background's pixel otherwise (Spagnolo, 2006). Hence, the foreground is considered as moving zones. This method is not very efficient since it depends essentially on the prefixed threshold. For this, recently, more sophisticated methods have been built to overcome this limit. Criminisi and al (Criminisi, 2006) have developed a bilayer segmentation method based on the calculation of a complex energy function.

In our system, we have used the background subtraction technique develop by Zivkovic and van der Heijden (Zivkovic, 2006) which models every image pixel's colour values distribution with a mixture of Gaussians (GMM). The mean and the covariance of each component in the mixture are updated for each new video frame (image) to reflect the change of the pixel values. In the case when the new pixel value is far enough the mixture, the pixel is considered as a foreground. This method has shown its rapidity and its good segmentation results in a big variety of videos (as shown on Figure.5).



Fig. 6. Background subtraction results with the method of Zivkovic

This temporal segmentation based moving zones detection has allowed us to estimate the motion only on limited zones. Thereby, this technique will reduce the computational time of the ME process and gives a more precise estimation with the assumption that the motion vectors of the blocks which are out of the detected zones will have a null value. This gain is increased if the movement is concentrated in very limited zones.

## 5.2 Sub-pixel precision

Block based motion estimation assumes that every block have an integer pixel displacement which is, in reality, not true. Therefore, to improve the motion estimation and to increase the accuracy of the prediction, we have moved to sub-pixel precision by developing a sub-pixel technique with a bilinear interpolation process. This is done by interposing a line between each two lines of the image  $I$  (see Figure.7) and a column between each two columns of the image. Then, ME is applied to the new image  $O$ .

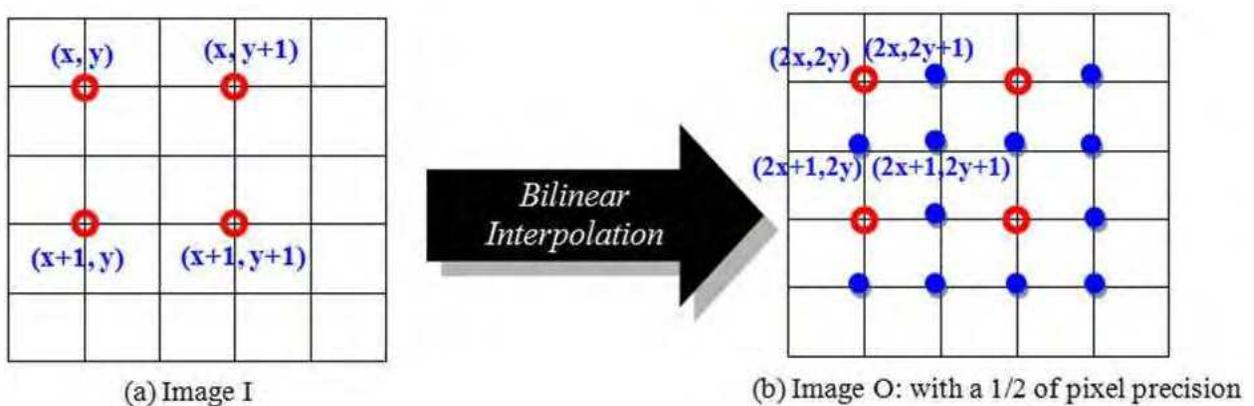


Fig. 7. Bilinear Interpolation for 1/2 pixel precision

The values of the pixels that are in the 1/2 pixel positions are determined relatively to their neighbouring pixels in the integer positions as follows:

$$O(2x, 2y) = I(x, y) \quad (2)$$

$$O(2x+1, 2y) = (I(x, y) + I(x+1, y)) / 2 \quad (3)$$

$$O(2x, 2y+1) = (I(x, y) + I(x, y+1)) / 2 \quad (4)$$

$$O(2x+1, 2y+1) = (I(x, y) + I(x+1, y) + I(x, y+1) + I(x+1, y+1)) / 4 \quad (5)$$

With this technique, a motion vector can point in a half or quarter of pixel position or even more. In this case, a block which has a real location at a fraction of pixels will be better predicted. The sub-pixel precision can not only increase the accuracy of motion vectors and reduce errors, but also filter the image to eliminate noise and rapid changes. The results of conducting the ME on some standard video sequences shown on the table below prove the efficiency of the sub-pixel precision technique.

| Sequence \ Precision | Tennis  | Susie   | Foreman |
|----------------------|---------|---------|---------|
| Integer pixel        | 31.7586 | 33.1613 | 31.2889 |
| 1/2 of pixel         | 34.2206 | 37.8811 | 33.6719 |
| 1/4 of pixel         | 34.7099 | 40.0285 | 36.6072 |
| 1/8 of pixel         | 31.5650 | 37.4465 | 37.7870 |

Table 1. PSNR of the reconstructed image with different sub-pixel precision

Using the sub-pixel technique as a pre-treatment step for the motion estimation process will improve it. Taken the Tennis sequence results in Table.1, the Peak Signal to Noise Ratio (PSNR), which is a criterion to compare the original frame to the reconstructed frame after motion compensation, is augmented from 31.7586 dB without using the sub-pixel technique to 34.2206 dB with a 1/2 of pixel precision and to 34.7099 dB with a 1/4 of pixel precision. This confirms the need to this technique for motion estimation. It should be noticed here that augmenting the sub-pixel precision level (to 1/8 of pixel precision or more) is not always beneficial since it can, in the most times, perturb the estimation.

That is true that this technique causes a doubling of image size, but is not a big problem since we conduct the motion estimation on the DWT approximation which has a reduced size. Furthermore, this technique saves time since it allows a quick search for the BMA by minimizing the path to find the corresponding block. For all this, in block based ME methods, sub-pixel technique is becoming crucial.

### 5.3 Shifting technique

The DWT has many advantages of multiresolution domain, which has made this spatial-frequency transformation very useful for the ME. However, the shift-variant property of the DWT caused by the decimation process has made the ME/MC less efficient in the wavelet domain. Otherwise, there is a big difference between the DWT of an image and the DWT of the same image shifted by even one pixel as shown in the Figure.8. This property touches especially the high frequencies in the image's edges, but it has less effect on the low frequencies.

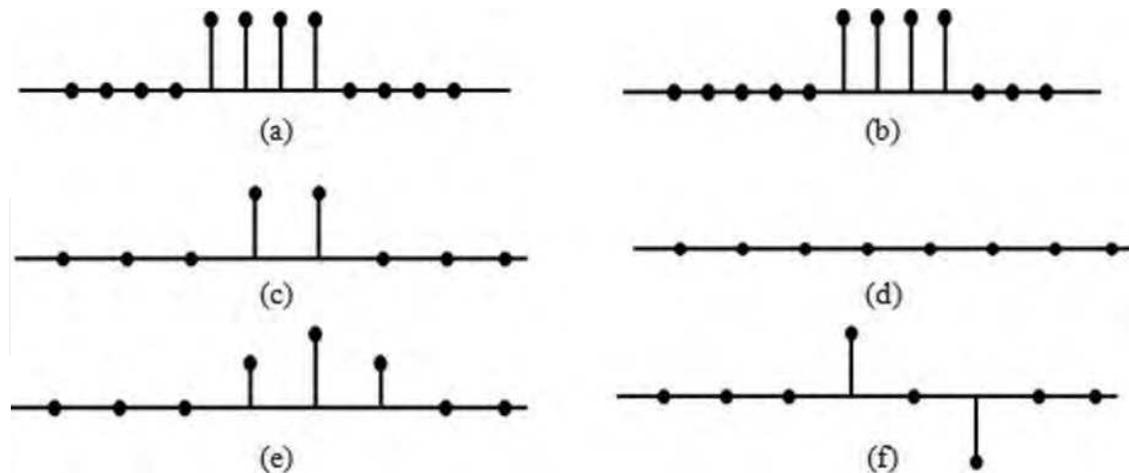


Fig. 8. Example of a DWT coefficients (*Haar* wavelet) for a 1-D signal  $s(n)$  and a shifted signal  $s(n+1)$  by one pixel. (a) original signal  $s(n)$ , (b) shifted signal, (c) low-pass frequencies subband  $s(n)$ , (d) high-pass frequencies subband of  $s(n)$ , (e) low-pass frequencies subband of  $s(n+1)$ , (f) high-pass frequencies subband of  $s(n+1)$ .

In the Figure.8,  $s(n+1)$  is a shifted variant by one pixel (shifting to the right) of the 1-D signal  $s(n)$ . As illustrated in this figure, the difference between the high-pass frequencies subband before and after shifting is much important than the low-pass frequencies subband before and after shifting. This is a simple and a 1-D signal example but it is also the case of the 2-D signal. Hence, this is reinforced more our choice to conduct ME in the approximation (low-pass frequencies subband) of the DWT.

To overcome the shift-variant property of the DWT, a shifting technique is used which increase the prediction quality (Yuan, 2002). Before applying ME, we shift the frame in spatial domain by one pixel in all directions. Then, the shifted frames are transformed to the wavelet domain for motion estimation more precise and more real. After calculating a motion vector for the block in every direction, we generate the final motion vector which is the mean of all calculated vectors.

This technique has increased the estimation results by smoothing the predicted vectors and reducing the aliasing effect. By adding this technique to the ME process, the estimation was remarkably ameliorated as shown in the Table 2. However, this technique has improved the PSNR of the reconstructed image after MC for the Tennis sequence from 31.7586 dB to 32.3164 dB.

| Precision \ Sequence | Tennis  | Susie   | Foreman |
|----------------------|---------|---------|---------|
| ME without shifting  | 31.7586 | 33.1613 | 31.2889 |
| ME with shifting     | 32.3164 | 35.5236 | 32.6301 |

Table 2. PSNR of the reconstructed image without/with the Shifting technique

### 5.4 Blocks overlapping technique

Supplementary technique for improving the motion estimation is to overlap the neighbouring block to smooth the motions vectors in a way to have a more real prediction (as shown in Figure.9). So, each motion vector will be the average of itself and the direct neighbouring motion vectors with a certain weighting (every MV will have a weight stronger than the weight of the MVs of the neighbouring blocks).

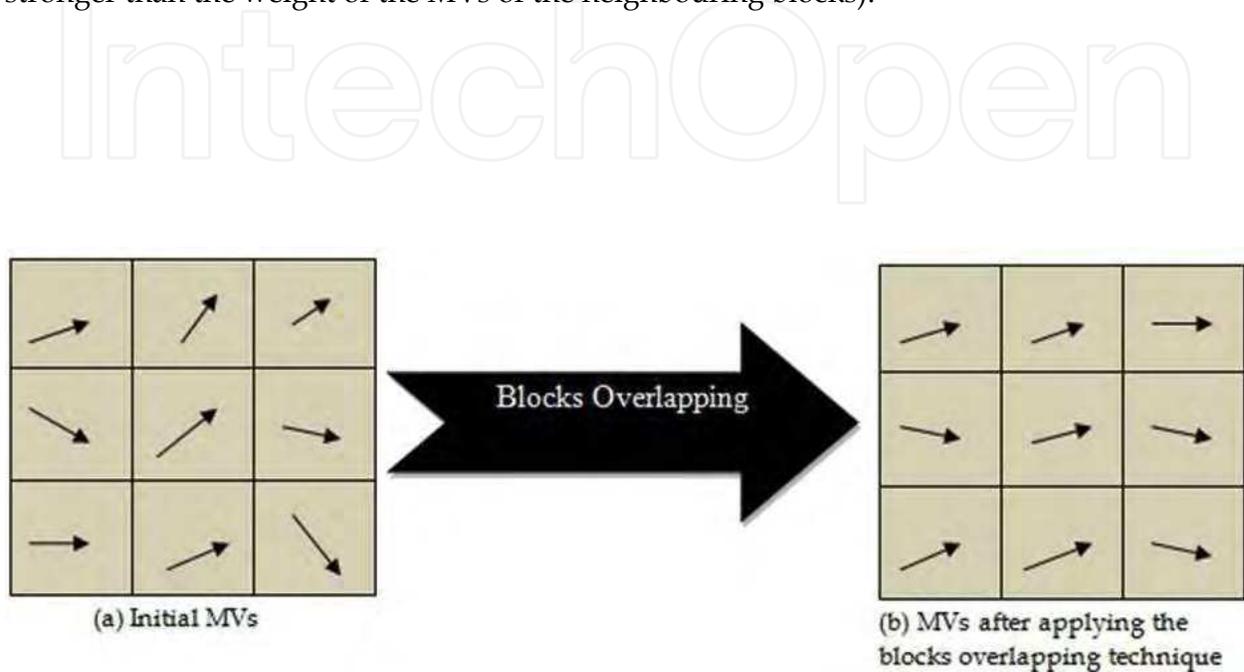


Fig. 9. Correcting the MVs with blocks overlapping technique

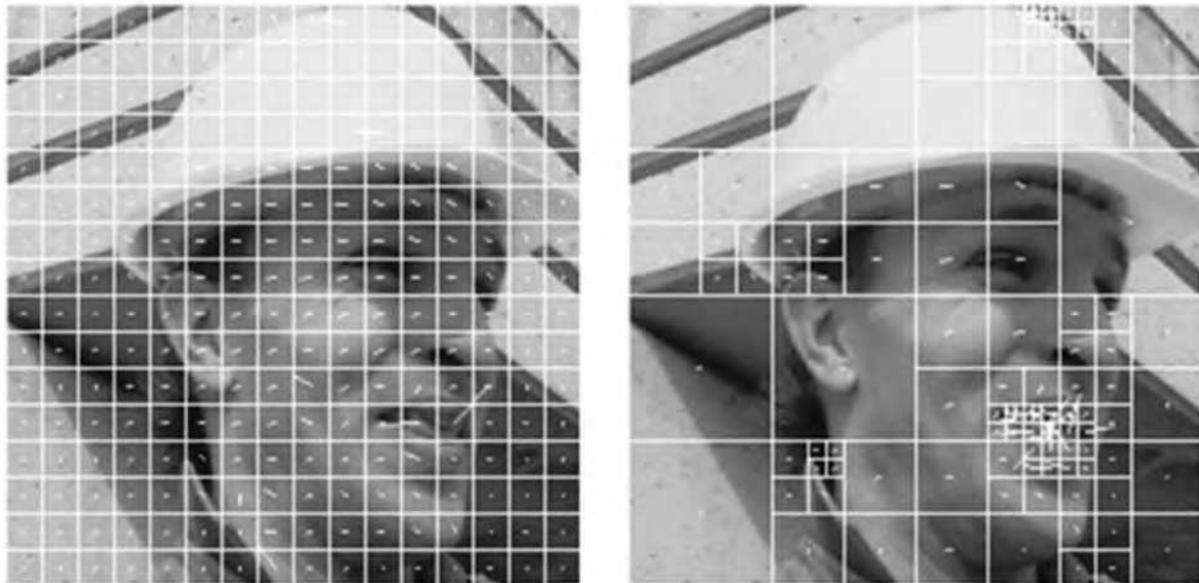
This blocks overlapping technique will surmount the false prediction especially the discontinuity at the edges which gives the high frequencies in the estimated image. This is done since the technique is somewhat averaging the possible candidates for each pixel and correcting then a probable false estimation. Hence, this technique will make the visual quality more clear and net.

### 5.5 Refinement techniques

The basic idea in the BMA is to divide the frame into blocks of a fixed size  $N \times N$ . This means that all the pixels of the same block have the same displacement. But, this is not true in most cases, since there may be different movements in the same block (movements intra-block).

For this, we have developed two techniques which aim to take into consideration this problem and give each image pixel a MV representing its real movement.

The first technique consists on dividing the blocks which are poorly predicted and conducting a re-estimation on the new sub-blocks. This will fix the blocks size relatively to the movements and we will build then a variable block size ME system (see Figure.10) as develop by Arvanitidou et al (Arvanitidou, 2009).



(a) ME with a pre-fixed block size

(b) ME with variable block size

Fig. 10. Motion estimation with different block size strategies

This technique is very powerful since it corrects the motion vectors by a hierarchical procedure based on modifying the block sizes. It provides a good estimation and tries to minimize the error by taking into account the intra-block movements.

Another refinement technique is also carried out for our method, which consists on moving the estimation to a lower level (larger resolution) of the DWT. This process is not performed for all blocks, but it runs only on poorly predicted blocks. The refinement will re-estimate the motion of the blocks that has an error greater than certain threshold. This technique has given a more accurate estimation prediction quality.

| Precision \ Sequence                           | Tennis  | Susie   | Foreman |
|--|---------|---------|---------|
| ME without refinement                          | 31.7586 | 32.5908 | 17.7091 |
| ME + refinement with changing the block size   | 32.0609 | 33.0652 | 17.6722 |
| ME + refinement by moving to a lower DWT level | 32.6278 | 34.3762 | 17.9133 |

Table 3. PSNR of the reconstructed image with different refinement techniques

As presented in the Table above, the second refinement technique has better results, which have encouraged us to use it in our method.

All these techniques have united to improve our methods which make it fast, efficient and accurate. In addition, we can even exploit the human visual system and remove the small variations not recognized by the human eye between the two frames. The motion vectors and the prediction error can be encoded after transformed by DWT using the Embedded Zerotree wavelet algorithm (EZW) developed by Shapiro (Shapiro, 1993) or by the Set Partitioning in Hierarchical Trees Algorithm (SPIHT) developed by Said and Pearlman (Said, 1996) which are algorithms that exploit the wavelet structure for an efficient coding.

## 6. Experimental results

In our block based method, we have fixed the Diamond Search as a block searching strategy and the MSE as a block matching criterion since it gives better compression performance while not sacrificing image quality. We have also fixed the size of the window to 7 and the size of the block to 2 since we work in the approximation in the third level of the DWT. Furthermore, we have integrated all the techniques mentioned previously with a quarter of pixel precision and a refinement technique by moving to lower DWT level to re-estimate the poorly predicted blocks.

Our method has proved its performance and robustness for several video benchmarks used to test the ME/MC methods such as the "Tennis", "Foreman", "Susie", "Claire" sequences and even the "Football" sequence which contains large movements.

| Sequence \ Methods | Tennis         | Foreman        | Susie          | Claire         |
|--------------------|----------------|----------------|----------------|----------------|
| Spatial domain     | 34.3983        | 33.5550        | 36.6450        | 37.7992        |
| DCT domain         | 28.2568        | 31.3646        | 31.2833        | 33.0233        |
| Conventional DWT   | 31.7586        | 31.2889        | 33.1613        | 32.5908        |
| Proposed method    | <b>35.6263</b> | <b>34.6025</b> | <b>38.3417</b> | <b>38.5418</b> |

Table 4. PSNR of the reconstructed image

The reached results showed large performance in terms of quality of reconstructed frame as shown in Table.4 and also in terms of compression ratio. All this, is due to the accuracy of the estimation and the corrections made for the motion vectors.

Our experiments verify the superiority of the proposed ME system, not only versus several other well-known ME systems in the frequency and the multiresolution domains, but also versus the ME systems in the spatial domain. Moreover, it is faster than other methods and the compression ratio is highly increased because it works on the approximation level of the DWT, which is 8 times smaller than the original image.

Furthermore, it is clear from the Figure.11 that there is a big difference between the visual qualities of the reconstructed frames using these different ME/MC systems. We can

observe that when the motion estimation is applied on the DCT domain, block effects appeared. On the other hand, using the classical DWT domain, there are also blocks effects, despite its superiority to the DCT domain. Our method gives a better visual quality that resembles to the quality of the reconstructed frame by the spatial domain based ME/MC system.

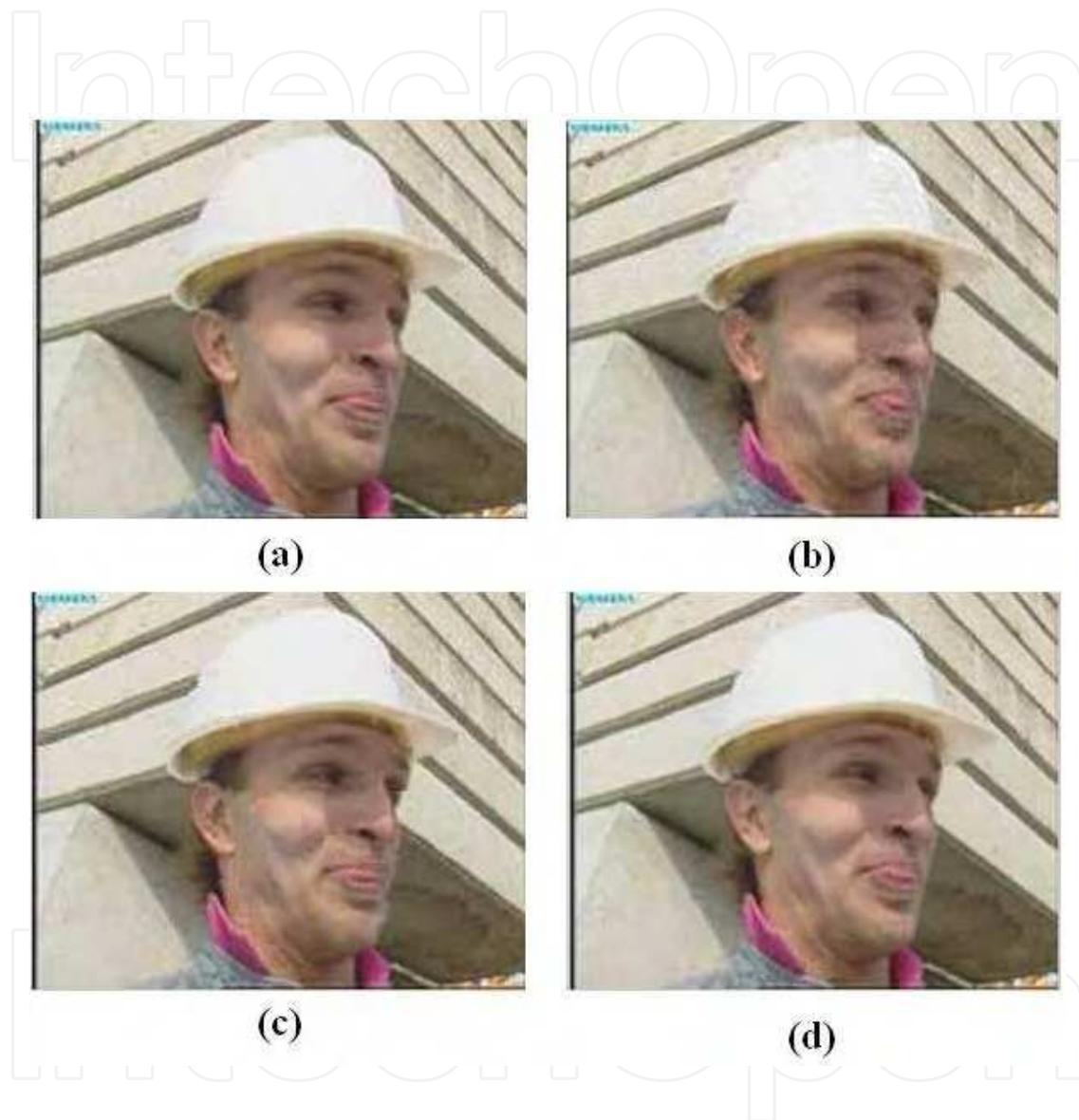


Fig. 11. The ME/MC results on the 129<sup>th</sup> frame of the “foreman” sequence. (a) The original image. The estimated frame: (b) ME/MC in the DCT domain, (c) ME/MC in the DWT domain, (d) with our method.

The efficiency of our motion estimation method is well confirmed by the results, in the visual qualities of the reconstructed frames, reached by applying the ME/MC on the Tennis sequence conducted in several domains. The results mentored in Figure.12 consolidate the fact that our motion estimation method outperforms other motion estimations methods conducted in different domains.

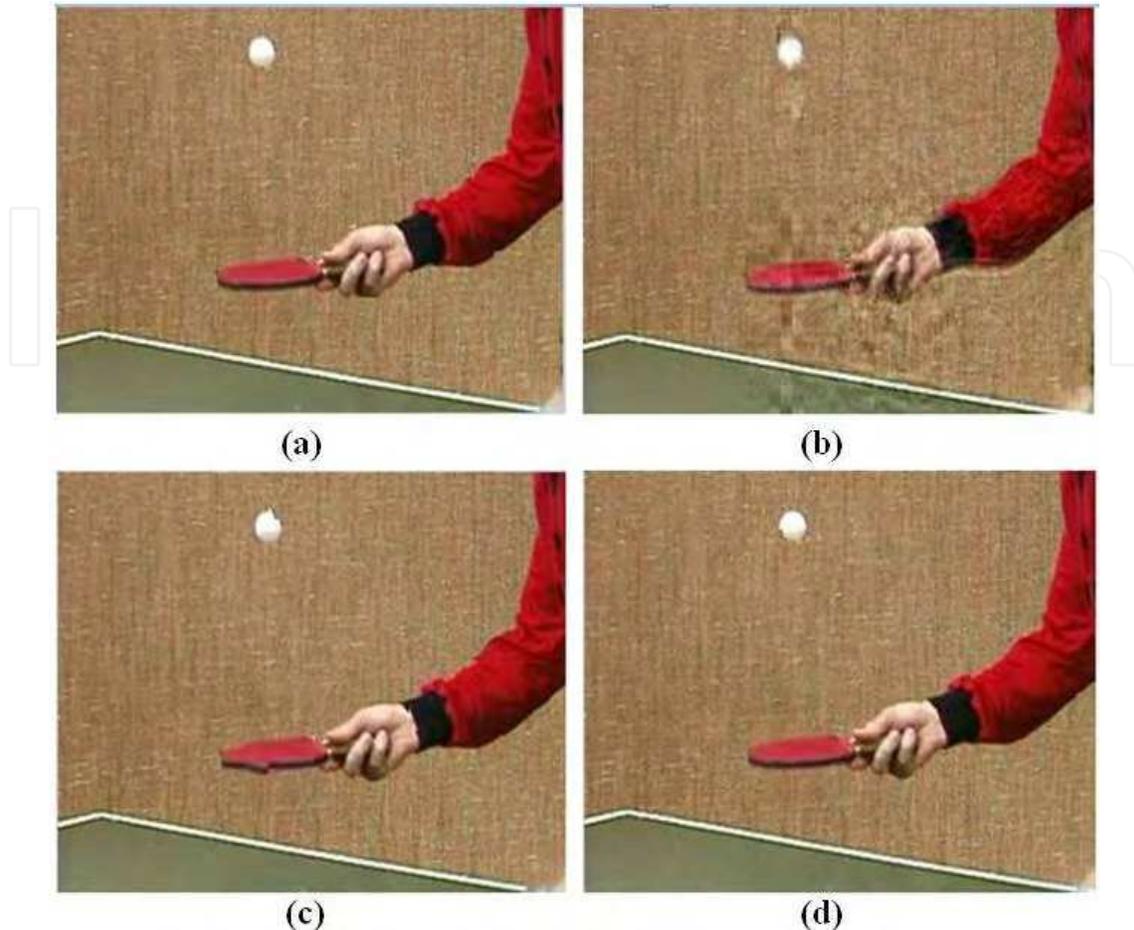


Fig. 12. The ME/MC results on the 17<sup>th</sup> frame of the “Tennis” sequence. (a) The original image. The estimated image: (b) ME/MC in the DCT domain, (c) ME/MC in the DWT domain, (d) with our method.

## 7. Conclusion

Video coding has received an increased interest because of the big growth in the quantity of the video data. That is why a big interest has been made for developing an efficient video coding system and improving the motion estimation part which represents the most important part since it consumes most computation time and most resources used for video coding. Making the motion estimation a fast and efficient process was the goal of many researchers. But, unfortunately, that was not reached in the spatial domain. That’s why, new ME systems have been conducted in other domain such as the frequency and the multiresolution domain. That is why many studies have been made to improve and simplify the ME methods. In this chapter, we have studied the wavelet as a domain for ME and we have proposed a multiresolution motion estimation and compensation method based on block matching applying in the wavelet coefficients. Because of some problems presented in this chapter, we have integrated some improvements techniques to ameliorate our ME system. As a future works, we will reinforce our method with others techniques such as the spatial segmentation which makes the estimation more accurate by trying to identify real objects in the predicted moving zones.

## 8. Acknowledgment

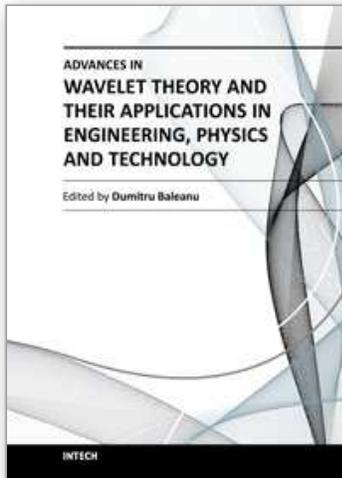
The authors would like to acknowledge the financial support of this work by grants from the General Direction of Scientific Research (DGRST), Tunisia, under the ARUB program.

## 9. References

- Arvanitidou, M.G. et al. (2009). Global motion estimation using variable block sizes and its application to object segmentation, *10<sup>th</sup> Workshop on Image Analysis for Multimedia Interactive Services*, pp. 173-176, ISBN 978-1-4244-3609-5, London, UK, May 6-8, 2009
- Barjatya, A. (2004). Block Matching Algorithms For Motion Estimation, DIP 6620 Spring 2004 Final Project Paper, Available from:  
[http://read.pudn.com/downloads153/sourcecode/graph/texture\\_mapping/675918/BlockMatchingAlgorithmsForMotionEstimation.PDF](http://read.pudn.com/downloads153/sourcecode/graph/texture_mapping/675918/BlockMatchingAlgorithmsForMotionEstimation.PDF)
- Ben Aoun, N., El'Arbi, M. & Ben Amar, C (2010). Multiresolution motion estimation and compensation for video coding, *Proceedings of the 10th International Conference on Signal Processing, Part II*, pp.1121-1124, ISBN 978-1-4244-5897-4, Beijing, China, October 24-28, 2010
- Cheng, P.Y, Li, J. & Jay Kuo, C.-C. (1995). Multiscale video compression using wavelet transform and motion compensation, *International Conference on Image Processing*, pp. 606-609, Vol.1, ISBN 0-8186-7310-9, Washington D.C., USA, October 23-26, 1995.
- Conklin, G.J., Hemami, S.S. (1997). Multi-resolution motion estimation, *International Conference on Acoustics, Speech, and Signal Processing*, pp. 2873-2876, Vol.4, ISBN 0-8186-7919-0, Munich, Bavaria, Germany, April 21-24, 1997
- A. Criminisi, A., Cross, G., Blake, A. & Kolmogorov, V. (2006). Bilayer Segmentation of Live Video, *International Conference on Computer Vision and Pattern Recognition*, pp. 53-60, Vol.1, ISBN 0-7695-2597-0, New York, NY, USA, June 17-22, 2006.
- Gharavi, H. & Mills, M. (1990). Block Matching Motion Estimation Algorithms: New Results. *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.37, No.5, (May 1990), pp. 649-651, ISSN 0098-4094
- Kutil, R. (2003). Evaluation of wavelet domain block motion compensation (WBMC), *Proceedings of the International Picture Coding Symposium (PCS'03)*, pp. 513-518, Saint Malo, France, April 23-25, 2003
- Lundmark, A., Li, H. & Forchheimer, R. (2000). Motion vector certainty reduces bit rate in backward motion estimation video coding, *SPIE Proceedings of Visual Communications and Image Processing*, Vol.4067, pp. 95-104, ISBN 0-8194-3703-4, Perth, AUSTRALIE, June 20-23, 2000
- Meyer, F.G, Averbuch A. & Coifman R.R. (1997). Motion compensation of wavelet coefficients for very low bit rate video coding, *Proceedings of IEEE International Conference on Image Processing*, Vol.3, pp. 638-641, ISBN 0-8186-8183-7, Washington, DC, USA, October 26-29, 1997
- Nosratinia, A. & Orchard, M.T. (1995). Multi-resolution backward video coding, *International Conference on Image Processing*, pp. 563-566, Vol.2, ISBN 0-8186-7310-9, Washington D.C., USA, October 23-26, 1995.

- Park, H. & Kim, H. (2000). Motion estimation using lowband-shift method for wavelet-based moving-picture coding. *IEEE Transactions on Image Processing*, Vol.9, No.4, (April 2000), pp. 577-587, ISSN 1057-7149.
- Said, A. & Pearlman, W.A. (1996). A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees. *IEEE transactions on circuits and systems for video technology*, Vol. 6, No. 3, (June 1996), pp. 243-250, ISSN 1051-8215.
- J. M. Shapiro, J.M. (1993). Embedded image coding using zerotrees of wavelet coefficients, *IEEE Transactions on Signal Processing*, Vol.41, No.12, (December 1993), pp. 3445-3463, ISSN 1053-587X.
- Shenolikar, P.C. & Narote, S.P. (2009). Motion estimation on DWT based image sequence. *International Journal of Recent Trends in Engineering*, Vol.2, No.4, (November 2009), pp. 168-170, ISSN 1797-9617
- Spagnolo, P., Orazio, T.D., Leo, M. & Distante, D. (2006). Moving object segmentation by background subtraction and temporal analysis. *Image and Vision Computing*, Vol.24, No.5, (May 2006), pp. 411-423, ISSN 0262-8856.
- Yuan, Y. & Mandal, K.M. (2002). Low-Band-Shifted Hierarchical Backward Motion Estimation and Compensation for Wavelet-Based Video Coding, *Proceedings of the 3<sup>rd</sup> Indian Conference on Computer Vision, Graphics and Image Processing*, pp. 185-190, Ahmedabad, India, December 16-18, 2002
- Zivkovic, Z. & van der Heijden, F. (2006). Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, Vol.27, No.7, (May 2006), pp. 773-780, ISSN 0167-8655

IntechOpen



## **Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology**

Edited by Dr. Dumitru Baleanu

ISBN 978-953-51-0494-0

Hard cover, 634 pages

**Publisher** InTech

**Published online** 04, April, 2012

**Published in print edition** April, 2012

The use of the wavelet transform to analyze the behaviour of the complex systems from various fields started to be widely recognized and applied successfully during the last few decades. In this book some advances in wavelet theory and their applications in engineering, physics and technology are presented. The applications were carefully selected and grouped in five main sections - Signal Processing, Electrical Systems, Fault Diagnosis and Monitoring, Image Processing and Applications in Engineering. One of the key features of this book is that the wavelet concepts have been described from a point of view that is familiar to researchers from various branches of science and engineering. The content of the book is accessible to a large number of readers.

### **How to reference**

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

Najib Ben Aoun, Maher El'arbi and Chokri Ben Amar (2012). Wavelet Transform Based Motion Estimation and Compensation for Video Coding, *Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology*, Dr. Dumitru Baleanu (Ed.), ISBN: 978-953-51-0494-0, InTech, Available from: <http://www.intechopen.com/books/advances-in-wavelet-theory-and-their-applications-in-engineering-physics-and-technology/wavelet-transform-based-motion-estimation-and-compensation-for-video-coding>

**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](http://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

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen