# Analysis of Video Compression Algorithms Using DCT, DWT and FFT

Thoon Sandy, Khant Kyawt Kyawt Theint
*University of Computer Studies, Yangon*
*thonsandi123@gmail.com, khant2theint@gmail.com*

## Abstract

*Compression creates a brand new document that shops records in a format that require much less area. Compression strategies are used in video conferencing, video telephony, video on mobile telephone, video on internet. The compression of an picture or video record is of splendid concern in virtual communication machine due to the storing capacities and constraints in the transmission rate. And a video compression format or a video compression specification is a specification for digitally representation of video as a file or a piece circulate. Amongst them, this system used the avi format for video file compression. In this system, the avi format video is compressed by using* ***discrete cosine transform (****DCT), d****iscrete wavelet transform (****DWT) and f****ast fourier transform (****FFT). And then, this system analyze the performance of each compression techniques according to the compression ratio.*

***Key words:*** *Video Compression, avi, DCT, DWT, FFT*

## 1. Introduction

Video is a sequence of nonetheless pictures which can be referred to as frames [5]. The consumers using digital video growing each day, so video compression is necessary to reduce the size. Video compression has two important benefits. First, it makes it possible to apply virtual video in transmission and garage environments that could now not help uncompressed video for example cutting-edge net throughput costs are inadequate to handle uncompressed video in real time. Video compression enables greater green use of transmission and storage sources. If the excessive bit rate transmission channel is to be had, then it is greater attractive proposition to ship a high resolution compressed video or a couple of compressed video channels than ship a single, low resolution, uncompressed circulate.

In spite of regular advances in storage and transmission capability, compression is probably to be an critical component of multimedia services for decades to come back.

Uncompressed digital video can require big bandwidth mainly as spatial and temporal decision growth. One of the most important downside of an uncompressed video is storage and transmission necessities. Video compression is wanted in internet video streaming and television broadcast where multimedia alerts are transmitted over a hard and fast quantity of limited bandwidth channels. Video compression techniques are used to reduce redundancy in video records without affecting visual pleasant. In video compression, the needless facts and repetitive facts are discarded.

## 2. Related Work

Anamitra Bardhan proposed the compression technique that plays an critical function in prognosis, prognosis and analysis of ischemic coronary heart diseases. It's also foremost for its fast facts sending capability within the area of telemedicine. Discrete consine transformation (DCT), discrete wavelet transformation (DWT), fast fourier transformation (FFT) and Walsh Hadamard transformation (WHT) are the main used in this system. On this paper a comparative take a look at of FFT, DCT, DWT, and WHT is proposed the use of ECG and PPG signal [1]. This paper promises a comparative study based totally on compression ratio and Peak-Signal-to-Noise-Ratio (PSNR) values of image features for corresponding techniques. [6].

Rosa A Asmara proposed the virtual picture watermarking that is used lately to comfy the photo via embedding every other digital picture. Frequency area transformation techniques are used extensively in digital image compression and virtual photograph watermarking. Popular transformation strategies used are 2D DCT, 2D DFT, and 2D DWT. This paper will display the evaluation result of these three transformation method. The experiments are assessment evaluation of picture watermark excellent the use of top signal to noise ratio (PSNR), color changing, photograph resizing, picture optical

scanning and the noise-tolerant of the photograph watermarked with the aid of giving Gaussian Noise.

## 3. Background Theory

**Video Compression:** Video compression is a set of techniques for reducing and removing redundancy in video data. The compressed video must have a much smaller size compared to the uncompressed video. This allows the video to be saved in a smaller file or sent over a network more quickly.

A video comprises of a sequence of frames. A video, of the duration of 1 second, generated by a TV camera usually contains from 24 frames to 30 frames/ second uncompressed video scanning.


**Figure 1. Frames of Un-compressed Video**

MPEG4 is a lossy video compression technique based on video compression algorithms, quantization, and entropy encoding [2]. Lossy way that records is lost, or thrown away, for the duration of compression, so great after decoding is much less than the original picture. Strategies of mpeg-four video compression standards are as follows [1]:

**Video compression algorithms:**

**Discrete cosine transform (DCT)** is used to transform spatial facts into frequency statistics. This lets in the encoder to discard information, corresponding to better video frequencies, which might be much less visible to the human eye.

**Quantization** is carried out to the DCT coefficients of either authentic frames or the DCT of the residual to limit the set of viable values transmitted through putting them into companies of values which can be nearly the identical.

**Entropy encoding:** entropy encoding is a piece assigning and lossless module for any transmission multimedia coding. Given that entropy encoding is a lossless module, compression ratio is the best constraint. This sort of encoding consists of run-period encoding and Huffman encoding.

## 4. Overview of the Proposed System

In the proposed system, the avi layout video is used as an input video and each body of video is converted by way of (discrete cosine transform / discrete wavelet transform / fast fourier transform) as first step as shown in figure 2. Second step is quantization to discard the redundant frame to restriction the organization of viable values transmitted. Third step is the technique of encoding by means of using run length encoding that is a chunk assigning and lossless module for any transmission. To get original avi layout video report, interpreting system should be completed by using inverse transform of DCT, DWT and FFT. In the interpreting method, the inverse techniques of encoding are performed as shown in above figure 2.

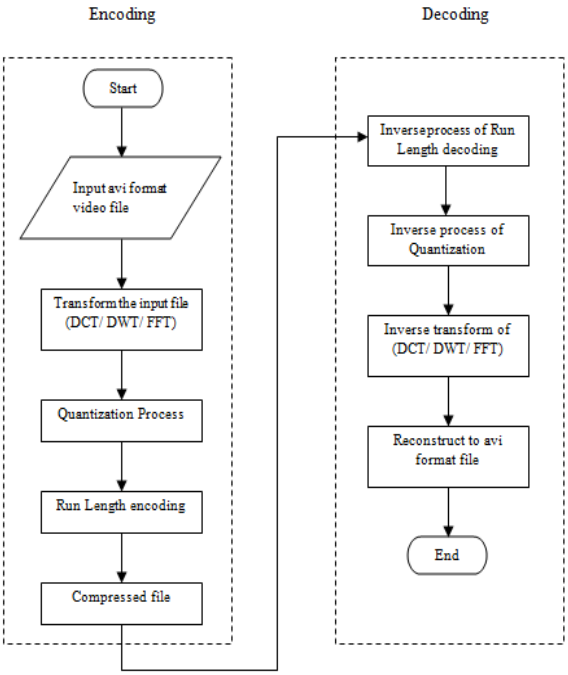



**Figure 2. The System Overview**

### 4.1. Discrete Cosine Transform (DCT)

DCT converts the spatial variations in the macro-block into frequency versions without changing the records. The DCT picture is damaged into 8x8 blocks of pixels.

Operating from left to right, top to bottom, the DCT is implemented to each block. Every block's elements are compressed then quantization way dividing by a few particular fee. The array of compressed blocks that represent the photo is stored in a considerably decreased amount of area.

The output of a DCT function is a DCT coefficient matrix containing the information in frequency area. After that for decreasing the garage space DCT coefficients are quantized through

dividing via quantization matrix. Block length fee additionally influences pleasant and compression ratio. In reality the better the block length, better the compression ratio it influences lack of more records and pleasant. So that big price is needed to lessen small size of area, so choice of quantization matrix, entropy encoding. If we take small fee for quantization then we get the better best or less compress ratio.

A two-dimension DCT function can be represented as:

$$F(u,v) = \frac{2}{MxN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\cos\frac{(2x+1)u\pi}{2M}\cos\frac{(2y+1)v\pi}{2N}$$

Eq. (1)

Where u = Horizontal spatial frequency for the integers: 0 <= u <= M-1, N-1

v = Vertical spatial frequency for the integers: 0 <= v <= M-1, N-1

f(x , y) = Pixel value at coordinates (x ,y )

F(u ,v ) = DCT coefficient at coordinates (u ,v )

f(x, y) is the (x , y)element of the image represented by the matrix F. M,N is the size of the block that the DCT is done on. The equation calculates one entry (u , v) of the transformed image from the pixel values of the original image matrix. For the standard 8x8 block that MPEG compression uses, N equals 8 and x and y range from 0 to 7.

## 4.2. Quantization

Quantization reduces the amount of information in better frequency DCT coefficient components using a default quantization matrix defined through mpeg-4 general. Default quantization matrix incorporates regular values as shown in figure 3. Every aspect in DCT coefficient matrix is divided through its corresponding regular cost in default quantization matrix and a quantized DCT coefficient matrix is computed. A quantization function may be represented as:

$$F(u,v) = Integer\ Round\left(\frac{F_{(u,v)}}{Q_{(u,v)}}\right)$$

Eq. (2)

The human eye isn't sensitive to the excessive frequency content in an picture. Therefore elimination of these spatial frequencies does not perceptible loss in photograph best. The spatial frequency content material of the picture is received through the use of the DCT operation, that's accompanied by way of an elimination of the high frequency content that is the quantization system.

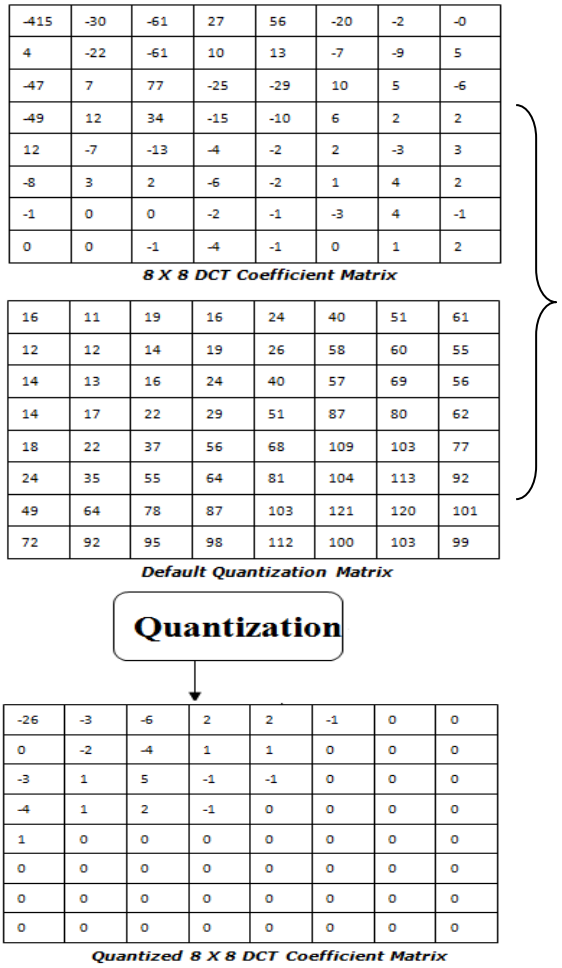



**Figure 3. Example Operation of Quantization**

## 4.3. Run Length Encoding

Run length encoding: additives of quantized DCT coefficient matrix is read. 2D matrix of quantized DCT coefficients are represented inside the form of a single-dimensional vector. After quantization, most of the excessive frequency coefficients (decrease right nook) are zero. To make the most the range of zeros, run-length encoder is used depend allows all the coefficients and decrease frequency coefficients are encoded using run-period encoding. Run period encoding is a lossless information compression technique of a series wherein same information value takes place in lots of consecutive facts factors.

*For example-*

WWWWWWWWWWWWBWWWWWWWW
WWWWBBB *(Sequence)* .

Here, the runs of data are stored as single data value and count. That is, the above sequence can be represented as *12W1B12W3B.*

Run length encoding is a very easy form of facts compression wherein runs of digital information are saved as a single records fee and be counted as opposed to because the unique run.

### 4.4. Discrete Wavelet Transform

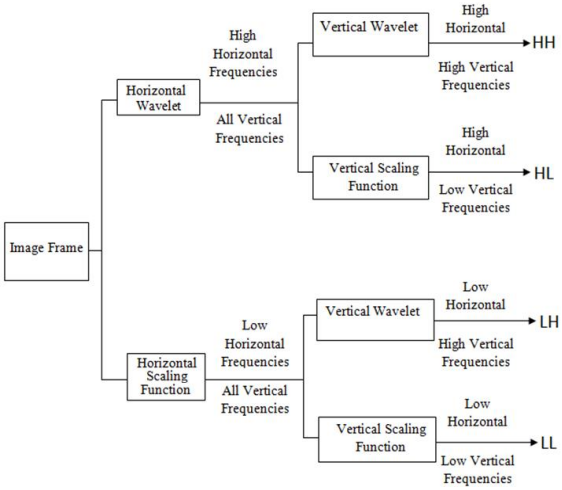



**Figure 4. Frame Decomposition**

The discrete wavelet transform (DWT) is easy to enforce and decreases the computation time and cast off pointless sources facts. DWT is split with the aid of successive low-skip and excessive-skip filtering of second photographs are created by using horizontally and vertically, then second wavelet should be decomposed for 4 sub- images as shown in figure 4. To construct 2D wavelets, the image is break up in parts, excessive frequency and occasional frequency running. The two ensuing sub pix include each high- and low-frequency vertical statistics. Every of the sub pix is now convolved with the wavelet and the scaling characteristic vertically, every generating two new separations. The method is shown in above figure6.

So, a single-stage wavelet transformation consists of a filtering operation that can decompose a two-dimensional signal into four frequency bands.

### 4.5. Fast Fourier Transform (FFT)

FFT derivatives from the discrete fourier transform (DFT) and it work with digital photos and finite range of discrete samples. These discrete samples are the pixels that compose an image. FFT and DFT in reality that they both perform fourier transform on discrete signals. The simplest difference is that the implementation of set of rules is optimized to dispose of redundant calculations. The output of the transformation represents the picture inside the fourier or frequency area, while the enter image is the spatial area equal.

The fourier transform is used in a extensive range of packages, consisting of photo analysis, photograph filtering, photo reconstruction and photograph compression and digital transmissions. The variety of frequencies corresponds to the number of pixels in the spatial area photo, i.e. the photo inside the spatial and fourier domain names are of the same size.

$$F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(\frac{ux}{M}+\frac{vy}{N})}$$

Eq. (3)

Where, u = Horizontal spatial frequency for the integers

v =Vertical spatial frequency for the integers

$$0 \le u \le M-1, N-1$$
$$0 \le v \le M-1, N-1$$

f (x , y) = Pixel value at coordinates (x ,y )

F (u, v) = FFT coefficient at coordinates (u, v)

The spatial domain and the exponential term is the basis function corresponding to each point F(u , v) in the Fourier space. The equation can be interpreted as: the value of each point F(u , v) is obtained by multiplying the spatial image with the corresponding base function and summing the result. The Fourier image can be re-transformed to the spatial domain. Input sequence of real numbers results in Fourier Transform output of complex numbers and efficiently computed using Fast Fourier Transform.

## 5. Performance Evaluation Measure

To measure the performance of each video compression techniques, this system uses the compression ratio, running time and peak signal-to-noise ratio (PSNR).

**Compression ratio:** Data compression ratio is defined as the ratio between the uncompressed size and compressed size.

**Peak Signal-to-noise ratio (PSNR):** PSNR measures the ratio between an arbitrary signal level and noise. Peak Signal to Noise ratio used to be a measure of image quality. The PSNR between two

images having 8 bits per pixel or sample in terms of decibels (dB) is given by:

$$PSNR = 10 \log_{10}\left(\frac{255^2}{MSE}\right)$$

Eq. (4)

Generally, when PSNR is 40 dB or greater, then the original and the reconstructed images are virtually indistinguishable by human observers [5].

**Mean Square Error (MSE):**

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[L(i,j) - K(i,j)]^2$$

Eq. (5)

Where m x n is the image size and L (i, j) is the input image and K (i, j) is the retrieved image. The mean squared error (MSE) is defined as the mean of the square of the difference between the original and reconstructed pixels, x and x 0. Mean square error is a criterion for an estimator: the choice is the one that minimizes the sum of squared errors due to bias and due to variance.

## 5.1. Experimental Results

Compression ratio presents in figure 5 is calculated by the dividing original video and compressed video size. The compression ratio CR can be defined as:

$$CR = \frac{\text{Size of Original Image}}{\text{Size of Compressed Codestream}} = \frac{\text{Size in byte}}{\text{Size in byte}}$$

Eq. (6)

Then, the experimental result is analyzed based on two nature of video: fixed_nature and motion_nature video. In fixed_nature video, the focus is fixed and background scene of the video frames are almost same. In motion nature video, the focus of the frame are moving left to right or up to down and the background scene of the video frames are quite different or motion quickly.

Based on the experimental result of figure 5, for the aspect of processing time comparison between DCT, DWT and FFT; DCT takes the longest processing time than DWT and FFT for both nature of video (Fix and Motion nature). But DWT and FFT are not distinct to determine for the processing time.

High PSNR value points the quality of compressed video file for various compression techniques. In this system experiments, DWT has the high PSNR values on Fix nature video and FFT has high PSNR value which means the compression has high quality and lower noise. For quick review, high quality compres sed values arc circled as a high light.

For compressed ratio, viptraffic.avi video compression ratio in DCT is 2.512:1 means DCT technique can compress the video file 2.512 times of original file size. DWT is 1.985:1 means DWT technique can compress the video file 1.985 times of original file size. FFT is 1.846:1 means FFT technique can compress the video file 1.846 times of original file size. So, the experimental result pointed that DCT has best compression rate. Therefore, the compressed size of DCT on each video files are high light as red color.

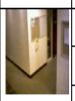
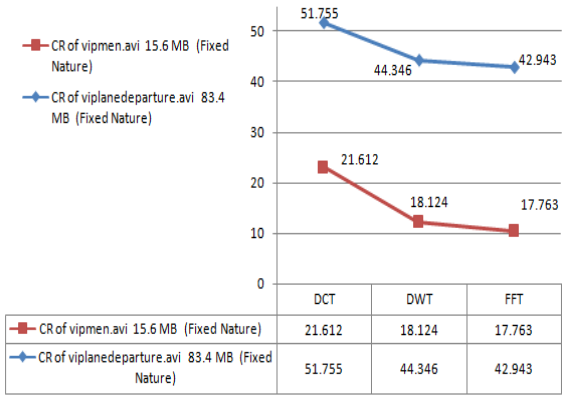
| No | Video Name | Movie Sample | Compression Techniques | Origima l Size | Compressed Size | Processing Time | MSE | PSNR | Compressed Ratio |
|---|---|---|---|---|---|---|---|---|---|
| 1 | viptraffic.avi(Fixed Nature) |  | DCT | 813KB | 324KB | 19.3764 seconds | 0.033042 | 62.9401 | 2.512:1 |
| | | | DWT | | 410KB | 7.2015 seconds | 0.012385 | 63.1274 | 1.985:1 |
| | | | FFT | | 441KB | 6.5893 seconds | 0.032031 | 63.0751 | 1.846:1 |
| 2 | vipmen.avi(Fixed Nature) |  | DCT | 15.6MB | 743KB | 47.0773 seconds | 0.028521 | 63.5791 | 21.612:1 |
| | | | DWT | | 686KB | 16.5687 seconds | 0.028485 | 63.5846 | 18.124:1 |
| | | | FFT | | 904KB | 15.2992 seconds | 0.030461 | 63.2934 | 17.763:1 |
| 3 | viptrain.avi(Fixed Nature) |  | DCT | 10.0MB | 5.29MB | 400.9117 seconds | 0.028313 | 63.6109 | 1.905:1 |
| | | | DWT | | 7.55MB | 177.1119 seconds | 0.027598 | 63.7221 | 1.334:1 |
| | | | FFT | | 7.97MB | 178.3032 seconds | 0.032212 | 63.0507 | 1.264:1 |
| 4 | rhinos.avi (Motion Nature) |  | DCT | 25.0MB | 729KB | 66.7308 seconds | 0.03263 | 62.9947 | 35.192:1 |
| | | | DWT | | 886KB | 28.8115 seconds | 0.03227 | 63.0428 | 28.956:1 |
| | | | FFT | | 876KB | 28.9269 seconds | 0.037408 | 62.4011 | 29.287:1 |
| 5 | NewH.avi (Motion Nature) |  | DCT | 1.82MB | 1.54MB | 202.3189 seconds | 0.014726 | 66.45 | 1.180:1 |
| | | | DWT | | 1.60MB | 114.5617 seconds | 0.014724 | 66.4506 | 1.136:1 |
| | | | FFT | | 1.58MB | 115.736 seconds | 0.016033 | 66.7106 | 1.171:1 |
| 6 | vipmosaicking.avi(Motion Nature) |  | DCT | 15.6MB | 540KB | 42.2902 seconds | 0.038709 | 62.2527 | 29.770:1 |
| | | | DWT | | 649KB | 18.3403 seconds | 0.03849 | 62.2773 | 24.770:1 |
| | | | FFT | | 610KB | 18.2748 seconds | 0.041772 | 61.9219 | 26.354:1 |

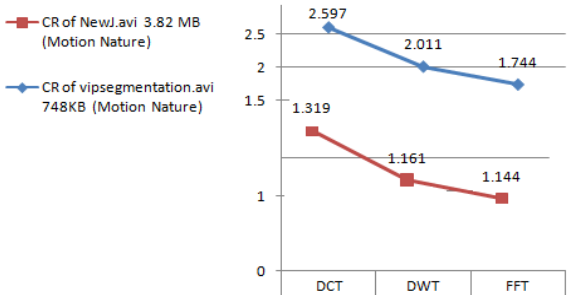
**Figure 5. Compression Analysis Based on DCT/DWT/FFT**

In this gadget, video compression evaluations are made on two varieties of video nature (fixed and motion) via the use of DCT, DWT and FFT as shown in figure 5, figure 6 and figure 7. However, DCT's compression charge on both kinds of nature is more than DWT and FFT due to the fact DCT is suitable for frequent statistics fee due to DCT's coefficient base calculation. So, the compression prices of DCT are maximum most appropriate over DWT and FFT.




| | DCT | DWT | FFT |
|---|---|---|---|
| CR of vipmen.avi 15.6 MB (Fixed Nature) | 21.612 | 18.124 | 17.763 |
| CR of viplanedeparture.avi 83.4 MB (Fixed Nature) | 51.755 | 44.346 | 42.943 |

**Figure 6. Analysis on Fixed_nature Video**




| | DCT | DWT | FFT |
|---|---|---|---|
| CR of NewJ.avi 3.82 MB (Motion Nature) | 1.319 | 1.161 | 1.144 |
| CR of vipsegmentation.avi 748KB (Motion Nature) | 2.597 | 2.011 | 1.744 |

**Figure 7. Analysis on Motion_nature Video**

In fixed_nature video, DWT's compression rate is greater than FFT because of the videos within the repair nature have greater comparable pixel price and no need to divide the input coding into non-overlapping 2-d blocks, it has better compression ratios keep away from blocking artifacts.

In motion_nature video, FFT's compression fee is more than dwt due to the videos in the speedy motion_nature have maximum special pixel value than fixed_ nature; and FFT is a quicker, computational, green algorithm to compute the discrete fourier transform (DFT) and its inverse. FFT is suitable for maximum specific cost of bit streams enter.

## 6. Conclusion

Video compression is essential role for transmission and storage of multimedia signals over the internet. So, this system compressed avi format video file by using DCT, DWT and FFT. According to the analysis of DCT, DWT and FFT, DCT provides the higher compression ratio on any nature video but it can take more processing time than other techniques and degrade the result in quality and performance than DWT. DWT is suitable for fixed_ nature video to get better quality, compression ratio over FFT and increase the result in performance achieved good PSNR, MSE values. FFT is supported for motion_nature video to get better compression ratio than DWT and better performance. Based on the experimental results; DCT provides the optimal compression rates and DWT has the optimal PSNR and MSE values upon the all of any nature videos.

## References

**[1]** A. B. Roy , D. Dey, B. Mohanty. D. Banerjee, "Comparison of FFT, DCT, DWT, WHT Compression Techniques on Electrocardiogram & Photoplethysmography Signals", Special Issue of International Journal of Computer Applications (0975 – 8887) International Conference on Computing, Communication and Sensor Network (CCSN) 2012

**[2]** I. E. G. Richardson, "H.264 and MPEG-4 Video Compression", Video Coding for Next – generation Multimedia, The Robert Gordon University, Aberdeen, UK, 2003.

**[3]** T. Sikora, "MPEG-1 and MPEG-2 Digital Video Coding Standards", McGRAW-HILL Book Company.

**[4]** M. Igarta, "A Study of MPEG-2 and H.264 Video Coding", Master of Science in Electrical and Computer Engineering", Purdue University, December, 2004.

**[5]** H. Balasko, "Comparison of Compression Algorithms for High Definition and Super High Definition Video Signals", Audio Video Consulting Ltd, Croatia.

**[6]** Sharon Shen UMBC "Discrete Wavelet Transform (DWT)"

**[7]** Walsh and M. Bourges-Se´venier (eds), MPEG-4 Jump Start, Prentice-Hall, 2002.