2002

Spread spectrum-based video watermarking algorithms for copyright protection

Serdean, Cristian Vasile

http://hdl.handle.net/10026.1/563

http://dx.doi.org/10.24382/3920
University of Plymouth

Chapter

6

"All animals are equal but some animals are more equal than others."

George Orwell, 1903-1950, "Animal Farm"

Wavelet Domain Watermarking

This chapter begins with an introduction to the wavelet transform, with the accent on the 2-D DWT (Discrete Wavelet Transform) case. The multiple advantages of the DWT transform are discussed and compared with the traditional FFT/DCT transforms [Jain, 1989], [Misiti et al, 2001], taking into account the specific framework of digital watermarking. Choosing a proper basis constitutes an important step which will be also discussed.

Due to major advantages of the DWT, the wavelet coefficients are one of the most suitable places to insert a watermark. The proposed watermarking system is described in detail during this chapter, including the HVS aspects of the scheme and error correction. The performance of the system will be then analysed for both image watermarking (in order to compare the results with the existing image watermarking schemes described in the literature) and video watermarking.

6.1 Short introduction to the Wavelet transform

Historically speaking the wavelet analysis is a relatively new method, although some of the mathematical background dates back to the theory of Fourier in the nineteenth century. Fourier set the basis of the frequency analysis which for a long time was the best and the only approach existent in signal analysis.

The research gradually moved from frequency-based analysis to scale-based analysis when the researchers realised that an approach measuring average fluctuations at different scales might prove less sensitive to noise. And so, the wavelet transform was born. The first recorded mention of what we call now a "wavelet" dates back to 1909 in Alfred Haar's thesis. The concept of wavelets in its present theoretical form was proposed later by Jean Morlet. The methods of wavelet analysis have been developed mainly by Yves Meyer and his colleagues, and the main algorithm was provided by Stephane Mallat in 1988. Since then, the research has become international [Burke-Hubbard, 1998], [Misiti et al, 2001].

6.1.1 Wavelet versus Fourier

Fourier analysis

The Fourier transform [Jain, 1989], [Misiti et al, 2001] is perhaps the most well-known way of analysing a signal. The signal is break down into its constituent sinusoids of different frequency e.g. transforming the signal from the time domain to frequency domain.

This analysis is very useful since in most of the cases the frequency content of a signal is very important. One important drawback of the Fourier analysis is that during this transformation time information is lost. In other words looking at the Fourier transform of a signal is impossible to say when a particular event took place. For stationary signals this is not a problem, but for real life signals which usually are non-stationary the transitory events are of capital importance.

Short-time Fourier analysis

In a bid to overcome this deficiency, the signal can be windowed e.g. the Fourier transform analyses only a small section of the signal at a time, given by the size of the window. Introduced by Gabor in 1946, this technique is called Short-Time Fourier Transform (STFT) [Misiti et al, 2001].

In this case the signal is mapped into a two-dimensional function of time and frequency. The STFT offers some information about when and at what frequencies a signal event occurs, but only with limited precision, given by the size of the window.

This compromise between time and frequency information is very useful but has a drawback too: once the size of the window is chosen, that window size is used for all frequencies. The STFT is a good start, but many signals require an even more flexible

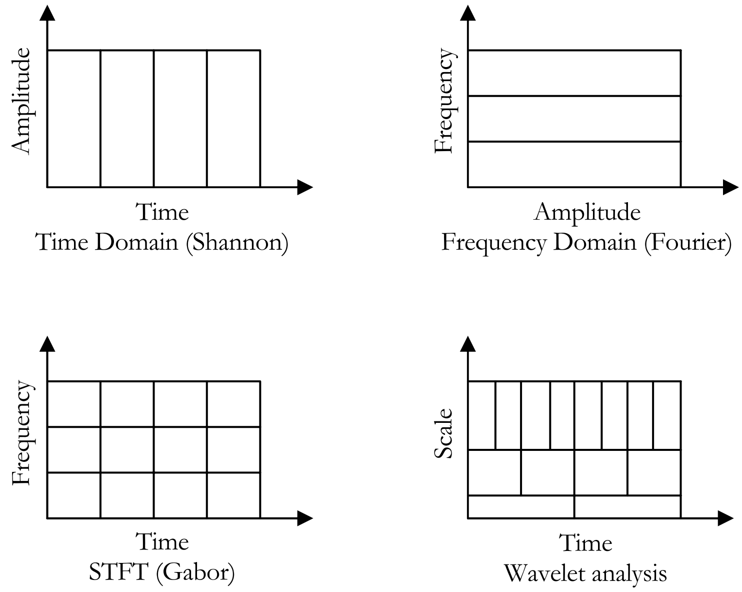


Figure 6-1 Methods of signal analysis: a comparison between time domain, Fourier, STFT and wavelet analysis

approach, where the window size can be varied in order to determine more accurately either time or frequency.

Wavelet analysis

The later requirement can be achieved by the wavelet analysis, which offers a windowing technique with variable sized regions. The wavelet transform allows the use of long time intervals where we want more precise low-frequency information and shorter intervals where we want high frequency information.

The concept is illustrated in **Figure 6-1** compared with the other 3 traditional approaches: time domain analysis, Fourier analysis and STFT analysis.

One major advantage of the wavelet transform is the ability to analyse a localised area of a larger signal. In this way, wavelet analysis is capable to reveal aspects of the data that other signal analysis techniques miss, for example breakdown points, trends, discontinuities in higher derivatives and self-similarity.

6.1.2 Wavelet transform

Unlike the FFT and the DCT, the DWT is a hierarchical transform which offers the possibility of analysing a signal at λ different resolutions or levels (with λ integer). Such multiresolution analysis gives a frequency domain representation as a function of time (or space in the 2-D case) i.e. both time/space and frequency localisation as shown in **Figure 6-1**.

One cannot achieve infinite resolution simultaneously in both time and frequency (Heisenberg uncertainty principle); high time resolution forces poor frequency resolution and vice-versa. This trade-off is used by the wavelet transform to provide multiresolution analysis.

In order to analyse the signal in terms of both frequency and time, the analysing functions must have different frequencies and they also have to be localised in time. Formally we refer to scale and resolution, where, for the dyadic case, scale is defined as $a = 2^{\lambda}$ and resolution as $r = \frac{1}{a} = 2^{-\lambda}$. Scale in this case means simply stretching (or compressing) the wavelet. The smaller the scale factor, the more compressed is the wavelet, or in other words, the greater the resolution, the smaller and finer are the details that can be analysed. There is a correspondence between wavelet scales and frequency as revealed by wavelet analysis:

- Low scale → Compressed wavelet → Rapidly changing details → High frequency
- High scale → Stretched wavelet → Slowly changing, coarse features → Low frequency

A wavelet can be defined as a waveform of effectively limited duration that has an average value of zero as opposed of the infinite duration sine waves used in the Fourier analysis. Also the wavelets tend to be irregular and asymmetric rather than smooth and predictable as the sine waves. A representation of the Antonini 7.9 wavelet for different scales is given in **Figure 6-2**.

For the 1-D case, a certain wavelet is defined by the mother wavelet function $\Psi(x)$ and a scaling function (or father wavelet) $\Phi(x)$, and the analysing wavelets are scaled

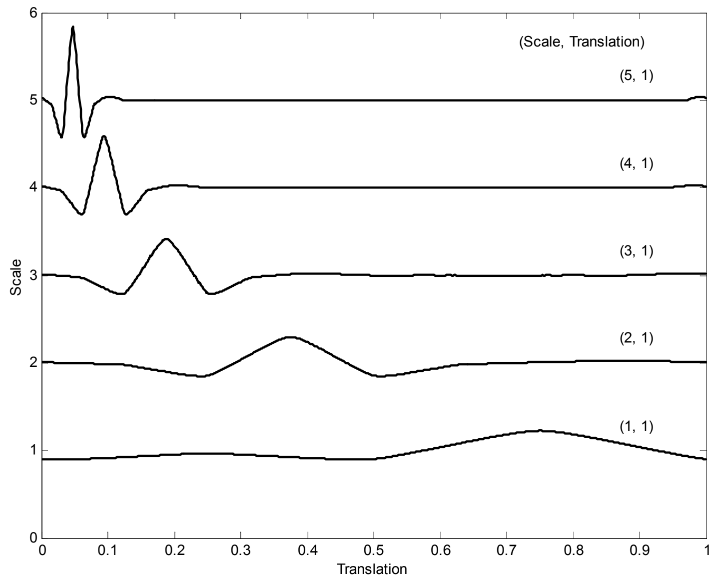


Figure 6-2 The Antonini 7.9 wavelets at various scales (same translation).

and translated versions of the mother wavelet:

$$\frac{1}{\sqrt{a}}\Psi\left(\frac{x-b}{a}\right) \tag{6.1}$$

The wavelet analysis breaks up the signal into shifted (translated) and scaled versions of the original (mother) wavelet. Defining translation $b = ka$, (where k, λ are integers) the dyadic case becomes:

$$\begin{aligned} \Psi_{\lambda,k}(x) &= 2^{-\frac{\lambda}{2}}\Psi(2^{-\lambda}x - k) \\ \Phi_{\lambda,k}(x) &= 2^{-\frac{\lambda}{2}}\Phi(2^{-\lambda}x - k) \end{aligned} \tag{6.2}$$

Given the input signal $f(x)$, a wavelet coefficient can be defined as:

$$C(\lambda,k) = \int_{-\infty}^{\infty} f(x)\Psi_{\lambda,k}(x)\,dx \tag{6.3}$$

For the 2-D case, the mother wavelet can be described as:

$$\frac{1}{\sqrt{a_1 a_2}}\Psi\left(\frac{x-b_1}{a_1}, \frac{y-b_2}{a_2}\right) \tag{6.4}$$

In this case (b_1, b_2) represents the translation vector and (a_1, a_2) is the scaling parameter. Furthermore there are one scaling function $\Phi(x, y)$ and three wavelet functions $\Psi_\theta(x, y)$, where θ denotes orientation:

$$\Phi(x, y) = \Phi(x)\Phi(y) \tag{6.5}$$

and

$$\begin{aligned} \Psi_H(x, y) &= \Phi(x)\Psi(y) \\ \Psi_V(x, y) &= \Psi(x)\Phi(y) \\ \Psi_D(x, y) &= \Psi(x)\Psi(y) \end{aligned} \tag{6.6}$$

Different orientations extract different features of the frame, such as vertical, horizontal, and diagonal information. This fact is very well illustrated in **Figure 6-3**. Generally speaking, edges and textures will be represented by large coefficients in the high frequency sub-bands, and are well localised within the sub-band.

In practice wavelet analysis is performed using multilevel filter banks. Essentially this comprises a succession of filtering (lpf and hpf) and sub-sampling operations and has been widely described in the literature [Kingsbury, 1997], [Antonini et al, 1992], [Villasenor et al, 1995], [Watson et al, 1996 and 1997] and [Xia et al, 1998].

6.1.3 Main applications of the Wavelet transform

Several thousand papers have been written within the last 15 years about the wavelets and their applications. This proves once more the success of wavelet analysis.

Some of the most popular applications are briefly presented below. Many applications were developed for signal/image processing, including de-noising and compression. Watermarking tends to become another successful application area. Probably one of the most popular applications of the wavelets is the compression of the FBI fingerprints and the recent JPEG 2000 standard.

Medicine is a very prolific application field for wavelets, especially in heart diagnosis (EKG/ECG – Electrocardiography), EEG (Electroencephalography), mammography and MRS (Magnetic Resonance Spectra).

Many papers were published in oceanography and earth studies.

6.2 Choosing the right basis

For watermarking one needs to select an appropriate wavelet or basis. Most of the basis development has taken place in the context of image compression [Villasenor et al, 1995], and fortunately watermarking and compression have many things in common. On the other hand, it is very important to choose a basis that offers compact support. The smaller the support of

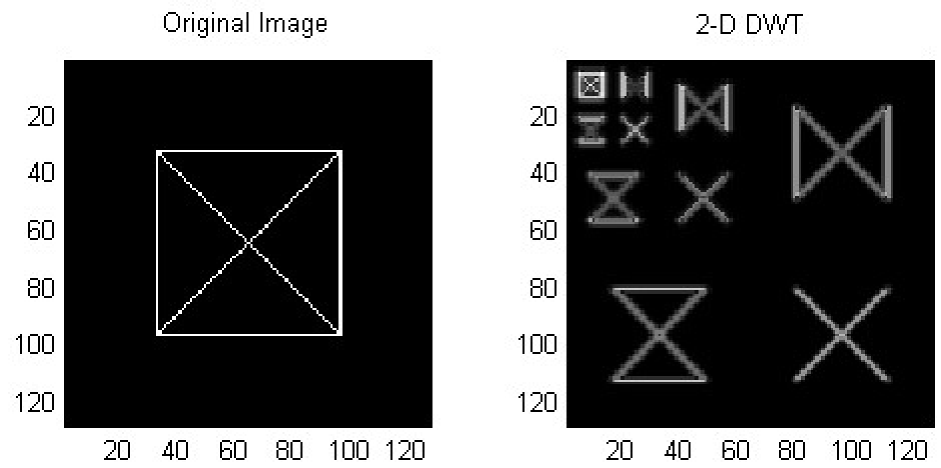


Figure 6-3 The 2-dimensional DWT: the original image and the wavelet decomposition for $\lambda = 3$.

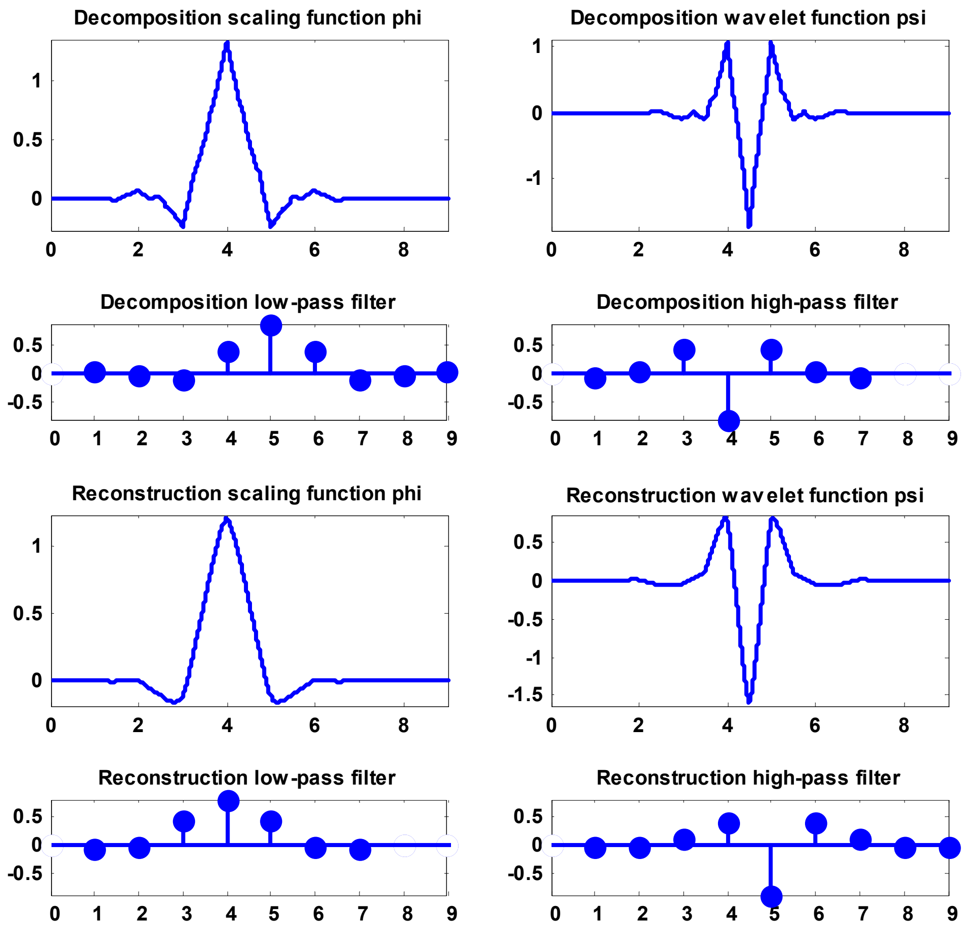


Figure 6-4 Decomposition and reconstruction filters for the Antonini 7.9 wavelet and the corresponding wave shapes.

the wavelet, the less nonzero wavelet coefficients will correspond to an edge for example, so basically the transform compacts more energy in the high frequency sub-bands [Lewis et al, 1992]. Also we are restricted to a class of either orthogonal or bi-orthogonal wavelets.

To narrow the choice even more, filter regularity, symmetry and a smooth wavelet function are important aspects for the reconstructed image quality. In addition, from the watermarking perspective this time, the selected basis must have a reasonably good HVS model designed for it.

The wavelet selected for this work is the Antonini 7.9 wavelet, presented in **Figure 6-4**, which is one of the best wavelets available for image compression [Kingsbury, 1997], [Antonini et al, 1992] and [Villasenor et al, 1995]. Its main properties are highlighted below:

- Bi-orthogonal wavelet
- Compact support, symmetric
- Good regularity (each filter has 2 factors)
- The lpf and hpf are quite similar

- Simple filters (only 7 and respectively 9 taps)
- Linear (zero) phase
- HVS model available [Watson et al, 1996 and 1997]
- Smooth wavelet function

This wavelet is widely used in image compression algorithms like EZW and SPIHT, but perhaps its most important application to date is the FBI fingerprint compression standard which uses this particular basis.

6.3 Advantages of the wavelet transform

The basis function for the DFT ($f(x) = \exp(i\omega x)$) or DCT (infinite cosine) has perfect localisation in frequency but is not time/space localised. In contrast, as already mentioned, wavelets offer a trade-off between time/space and frequency/scale, and so a watermarking scheme based on the DWT will produce a watermark with both spatially local and spatially global support (see **Figure 6-3**). This localisation makes a wavelet based scheme more robust than the DCT scheme, given geometric attacks such as cropping and scaling.

For instance, in the case of cropping, the lower frequency levels will be affected more than the high frequency ones, because of the fact that the watermark from the higher levels corresponds to a smaller spatial support. Looked at in the frequency domain, cropping corresponds to convolving the frequency components with a *sinc* function, where the width of the main lobe is inversely proportional to the width of the cropped window size [Podilchuck et al, 1998]. This will affect all the frequency components of any scheme based on a global transform, but since the wavelet scheme has a watermark with local spatial support, the watermark will be unaffected by the cropping.

For scaling, because the DWT coefficients are localised both in space and frequency,

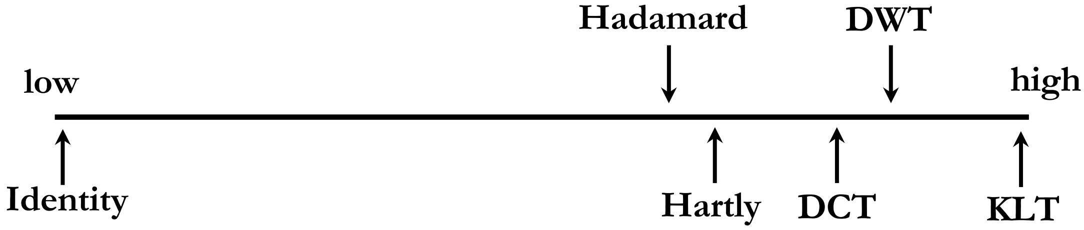


Figure 6-5 The energy compaction scale for several transforms.

whilst the DCT coefficients are only localised in frequency, it is likely that this kind of attack will be less serious for a DWT scheme. Simulation confirms this to be the case. Finally, the global spatial support of a DWT scheme will tend to be robust to operations such as low pass filtering/compression (which attenuate high frequency levels).

Another fundamental advantage of the DWT lies in the fact that it performs an analysis similar to that of the HVS. In fact the wavelet transform can be regarded as a rough HVS model by itself. The HVS splits an image into several frequency bands and processes them independently. In a similar way, the DWT permits the independent processing of different sub-bands without significant perceptible interaction between them. Again, this is because the analysing functions Ψ are localised in space, being zero outside a space domain U i.e. the signal values located outside of domain U are not influencing the values of the coefficients within U. Similarly, if Ψ is translated to position b, the wavelet coefficient will analyse the signal around b. This local analysis is specific to the compact support wavelets. Basically, for a small scale a local analysis is performed, whilst a large scale corresponds to a global analysis. **Figure 6-2** illustrates how the wavelet functions change for different scales.

Finally, more general advantages of the DWT are:

- It is not a block based transform, and so, the annoying blocking artefacts associated with the DCT are absent.
- Its hierarchical, multiresolution property offers more degrees of freedom compared with the DCT (for example separate or hierarchical cross-correlation).
- Higher watermark capacity and better robustness to attacks compared with the more traditional transforms (FFT, DCT).
- Lower computational cost than the FFT or DCT: $O(n)$ instead of $O(n\log(n))$, where n is the order of the transform input vector (lower hardware requirements) [Lumini et al, 2000], [Jain, 1989].
- Better energy compaction than both the FFT and DCT, in the sense that it is closer to the optimal Karhunen-Love transform as **Figure 6-5** depicts [Ramkumar et al, 1998-2]. More details about these transforms can be found in [Jain, 1989] and [Ramkumar et al, 1998-2].
- Some wavelets have useful invariance properties (for example complex wavelets are shift invariant [Kingsbury, 1997, 1998 and 1999]).

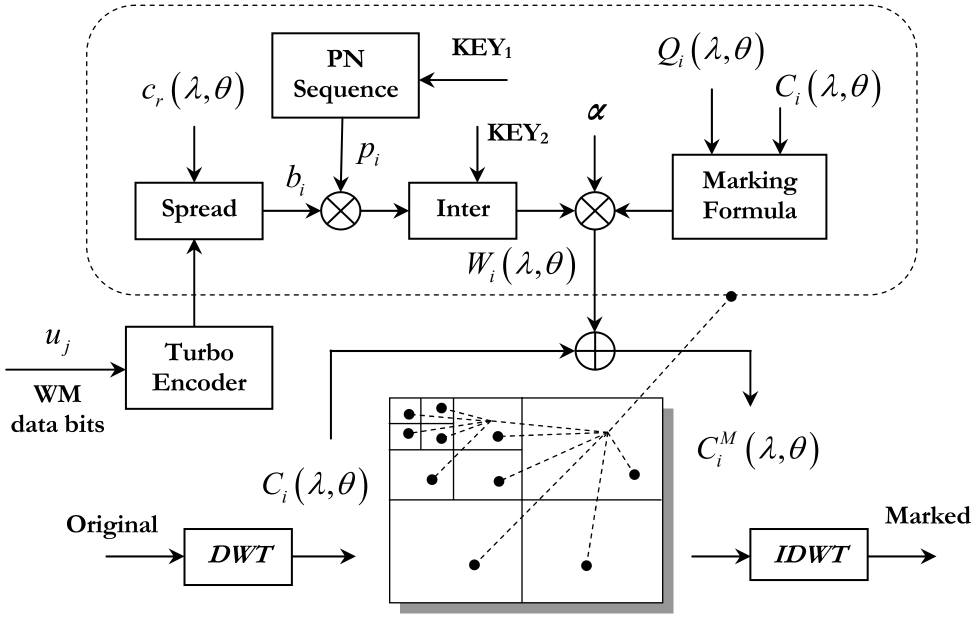


Figure 6-6 Wavelet-based watermark embedding.

6.4 The DWT-based watermarking scheme

6.4.1 Wavelet-based watermark embedding

Watermark embedding is illustrated in **Figure 6-6**. The preferred choice of the parameter λ is 3 (three levels of decomposition) [Misiti et al, 2001].

As for DCT systems, embedding uses the spread-spectrum approach and retrieval is via cross-correlation (matched filtering). The interleaver uses a separate key to that of the PN sequence in order to enhance system security and provide a random distribution of the data bits within each sub-band.

The hierarchical nature of the DWT is exploited here, by choosing to insert a self-contained watermark in each sub-band. This means that all of the data bits are inserted in each sub-band, the chip rate reducing as λ increases. Although reducing chip rate may appear to be a disadvantage, the advantage of this type of marking comes at the retrieval.

The watermark is embedded using amplitude modulation as follows:

$$C_i^M = \begin{cases} C_i + \underbrace{\alpha \frac{Q(\lambda,\theta)}{Q_{\min}} \cdot \frac{|C_i|}{mean(|C_i|)}}_{S} \cdot W_i, & \text{(details)} \\ if \quad S > 24, \quad then \quad S = 24. & \\ C_i + \alpha \frac{Q(\lambda,\theta)}{2} \cdot \frac{|C_i|}{mean(|C_i|)} \cdot W_i, & \text{(approximation)} \end{cases} \tag{6.7}$$

where $Q_{\min}$ is the minimum value within the quantisation matrix Q, W_i is the watermark, C_i is the original wavelet coefficient and C_i^M is the marked coefficient. The value of the factor S was experimentally derived from subjective visibility tests.

6.4.2 The HVS model

The HVS is incorporated in the quantisation matrix $Q(\lambda,\theta)$, where θ represents the orientation. Although this is a much simpler HVS model compared with the one used in the DCT scheme, the overall performance of the scheme is better. This illustrates once more the superiority of the wavelet transform over conventional transforms like FFT and DCT.

The matrix $Q(\lambda,\theta)$ offers only one quantisation factor for an entire sub-band, and incorporates only limited information about the HVS (essentially only the frequency sensitivity of the eye). In other words, the model is HVS dependent since it incorporates some aspects of the human vision (MTF of the eye), but unfortunately it is not media dependent, which constitutes an important drawback.

The quantisation matrix is computed according to the visual model developed by Watson for the Antonini 7.9 DWT [Watson et al, 1996 and 1997]:

$$\log Y = \log a + k\left(\log f - \log g_\theta f_0\right)^2 \tag{6.8}$$

where:

$$a = 0.495$$
$$k = 0.466$$
$$f_0 = 0.401$$
$$g_\theta = \{1.501,\ 1,\ 0.534,\ 1\},\ for\ \ \theta \in \overline{1\ldots4}$$

The rest of the parameters can be defined as:

$$f = r \cdot 2^{-\lambda} \quad \text{(cycles/degree)} \tag{6.9}$$

and:

$$r = dv\tan\left(\frac{\pi}{180}\right) \simeq dv\frac{\pi}{180} \tag{6.10}$$

where d is the display resolution in pixels/cm and v is the viewing distance in cm. Finally, the quantisation factor for each sub-band is derived as:

$$Q(\lambda,\theta) = \frac{2Y_{\lambda,\theta}}{A_{\lambda,\theta}} = \frac{2}{A_{\lambda,\theta}} \cdot a \cdot 10^{k\left(\log\frac{2^{\lambda} f_o\, g_\theta}{r}\right)^2} \tag{6.11}$$

where $A_{\lambda,\theta}$ represents the basis function amplitude for the Antonini 7.9 wavelet [Watson et al, 1996 and 1997].

The quantisation matrix $Q(\lambda,\theta)$ is a rough measure of the visibility for each sub-band, and, as stated, it is not media dependent. This dependence is required for a robust watermark and is provided by the embedding algorithm in equation (6.7). This marks more heavily the high frequency sub-bands and the largest coefficients, since modification of these coefficients is less likely to incur visible artefacts.

6.4.3 Wavelet-based watermark recovery

Watermark retrieval is shown in **Figure 6-7**. As mentioned, it is advantageous to have a self-contained watermark (all data bits) in each sub-band, since a SNR can be determined for

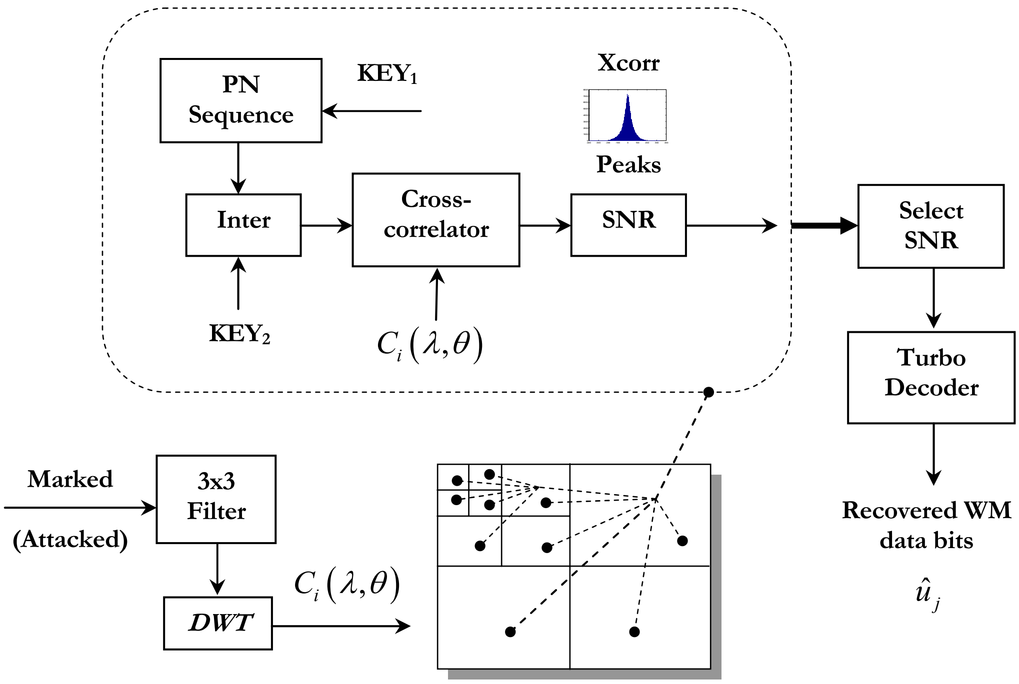


Figure 6-7 Wavelet-based watermark recovery.

each sub-band as an indicator of sub-channel quality. Different types of attack affect different levels and orientations in different ways, so it is always possible to select an optimal sub-band via SNR. In this way we can take advantage of any strong structure associated with the original image/video.

Correlation is therefore performed separately for each sub-band, obtaining a set of cross-correlation peaks (one peak for each embedded data bit) for each sub-band. A SNR is then computed for each set of cross-correlation peaks (section 4.3), and retrieval is carried out (only) for the sub-band with the highest SNR.

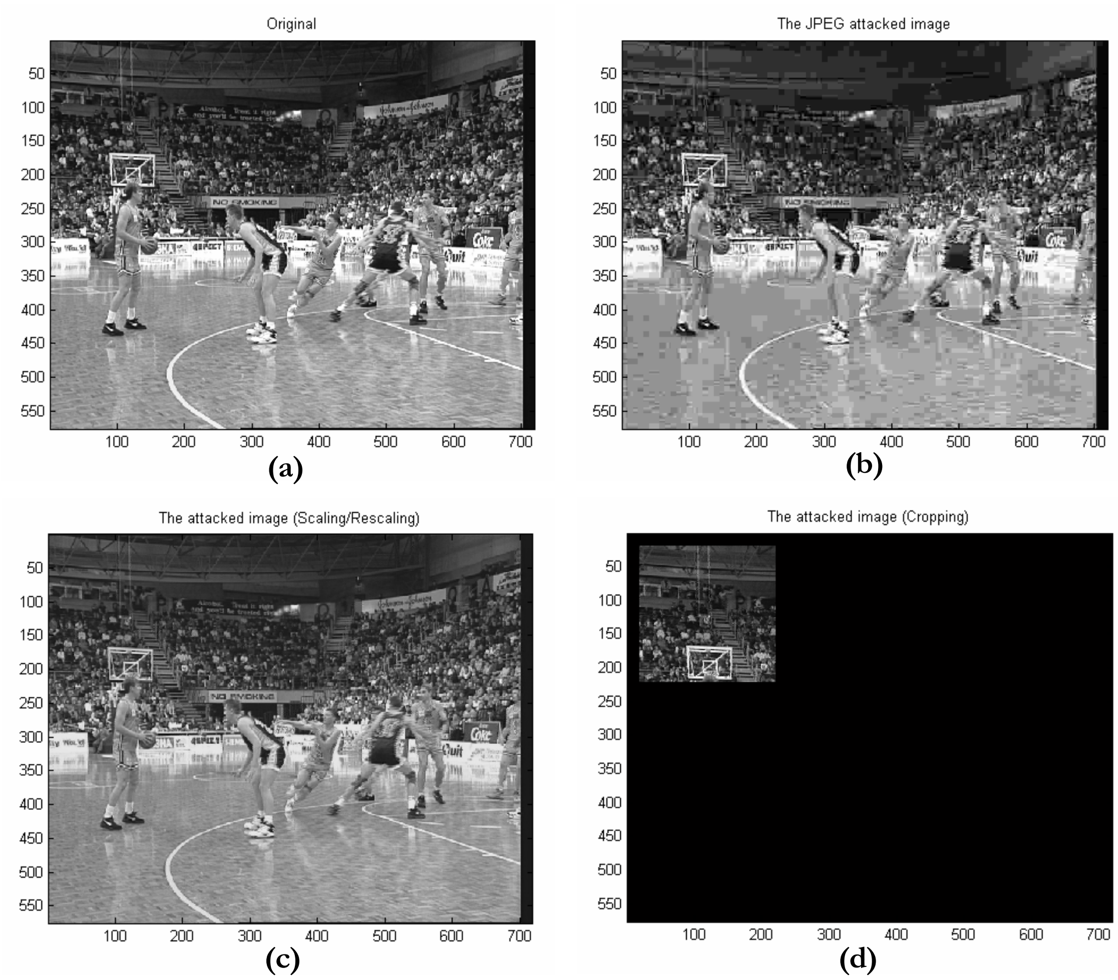
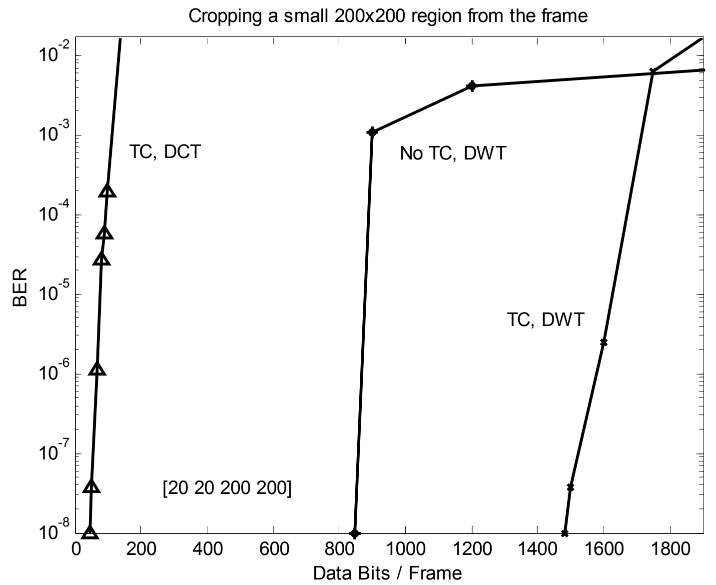


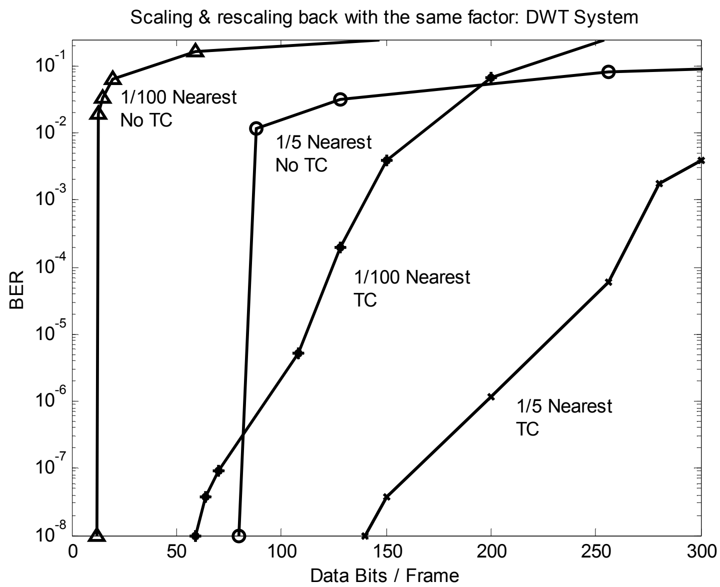
Figure 6-8 A frame from the original "Basketball" sequence **(a)** and the effects of different attacks: **(b)** JPEG compression (5% quality factor, 30:1 compression ratio), **(c)** scaling/rescaling (1/5 and back using the 'nearest' method) and **(d)** cropping a small area from the original (200x200 pixels rectangle with the upper left corner at the location [20,20]).

6.5 Performance of the Wavelet-based system

The visual effect of attacks like cropping, scaling/rescaling and JPEG compression is illustrated in **Figure 6-8**. The magnitude of these attacks is quite extreme, leading to unacceptable visual artefacts. The scaling for example was performed with a very bad quality



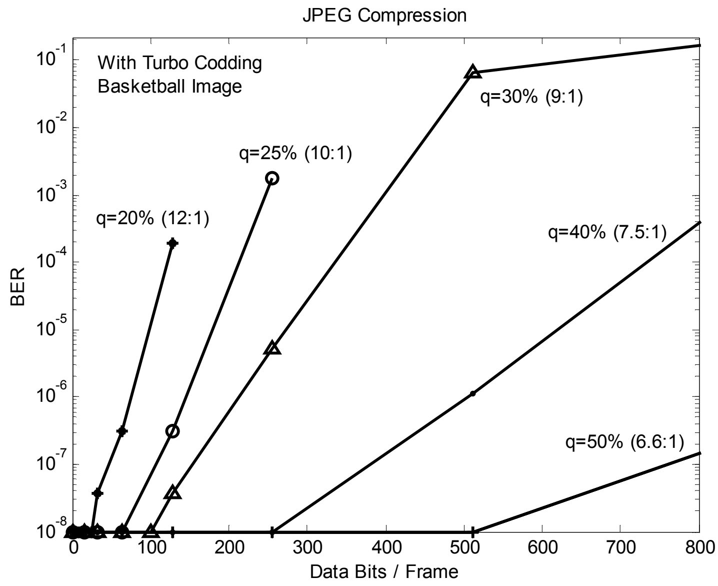
(a)



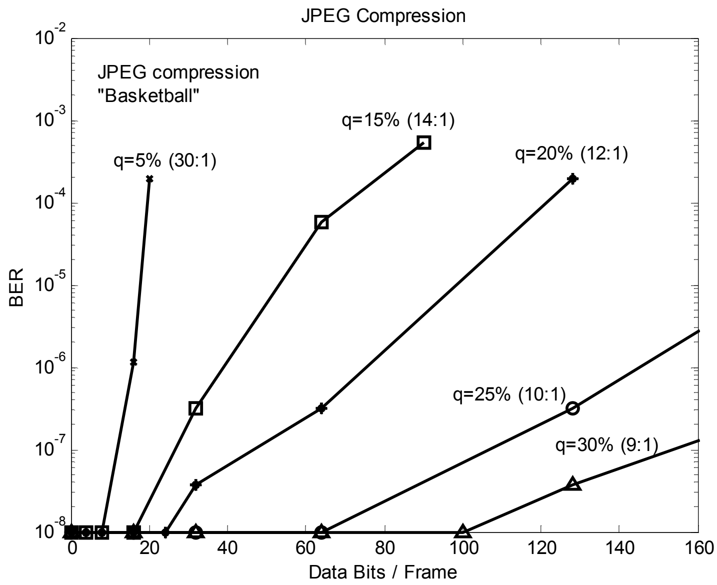
(b)

Figure 6-9 The performance of the DWT system for: **(a)** cropping and **(b)** scaling-rescaling.

interpolation filter, just to see how well the system performs. For JPEG compression, important artefacts become visible for a quality factor lower than about 25% (10:1 compression). **Figure 6-9 ↔ Figure 6-14** presents the performance of the DWT system for various attacks and some comparisons between the DCT scheme and the DWT scheme in order to outline the superiority of the DWT-based approach. The DCT scheme used for comparison is the one described in **Chapter 5**.



(a)



(b)

Figure 6-10 The performance of the DWT system for: **(a)** medium quality JPEG compression and **(b)** low quality JPEG compression.

As might be expected from the compact support, the most significant advantage of wavelets occurs under cropping and scaling. For cropping, a rectangle of 200x200 pixels was selected from the upper left corner of the frame, as shown in **Figure 6-8(d)**. This location was selected since it has average detail. Clearly, cropping to this degree is an extreme case and is unlikely to occur in practice. It is apparent from **Figure 6-9(a)** that the DCT scheme has poor performance even with FEC (TC, DCT curve), whereas the DWT scheme performs very well

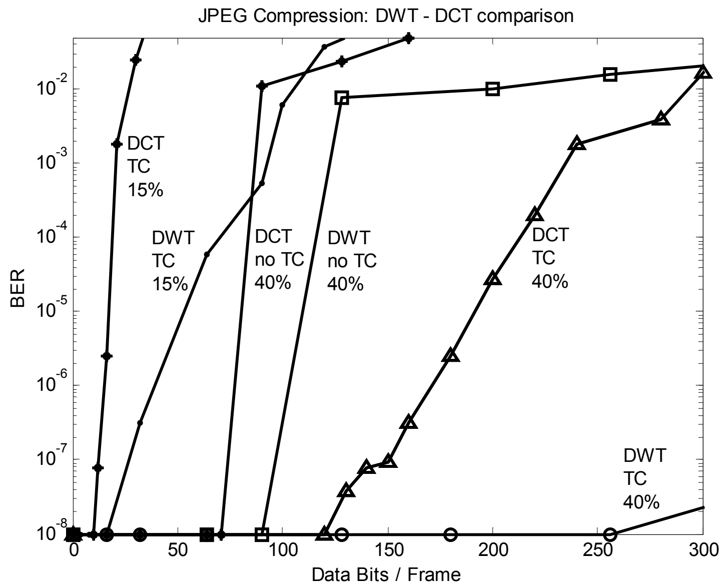


Figure 6-11 Comparison between the DCT and DWT systems for medium quality JPEG attack.

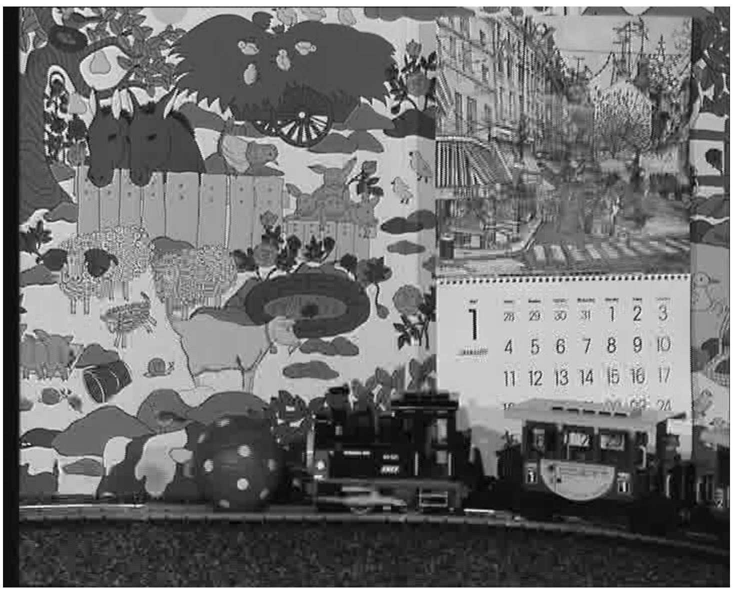
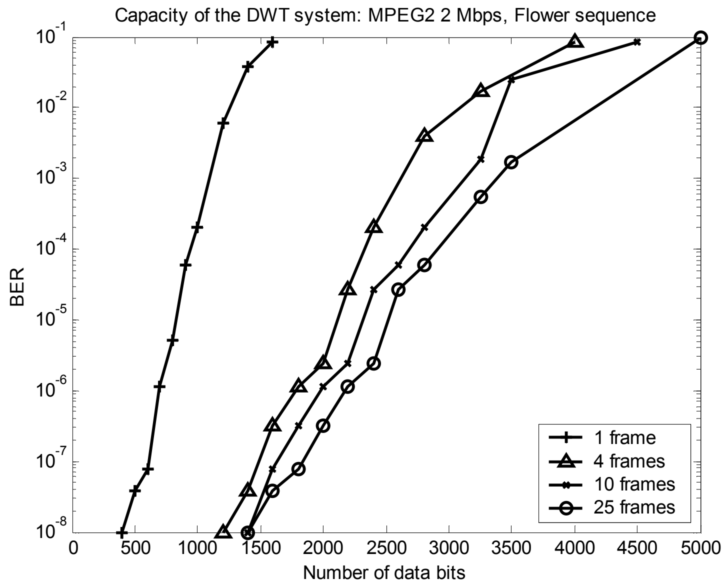


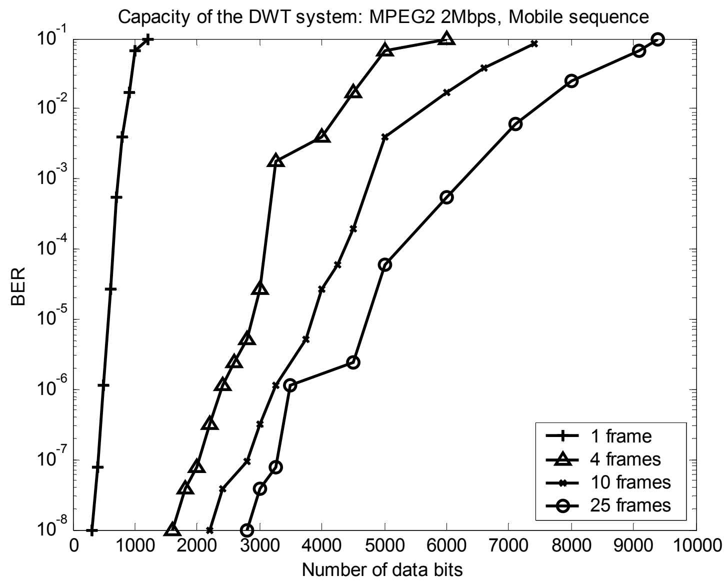
Figure 6-12 The "mobile" sequence MPEG2 compressed to 2Mbps: it is easy to spot the blocking artefacts even on a still frame.

without FEC (No TC, DWT curve) leading to over 20 kbps at $BER = 10^{-8}$. With FEC the capacity increases to 37 kbps, but will reduce markedly under a combined attack.

Figure 6-9(b) shows the results for scaling. The frame is scaled up or down and then brought back to the original size (720x576). Even so, with the worst kind of scaling, the DWT system performs quite well. The effect of this kind of attack results in luminosity changes and geometric distortion, **Figure 6-8(c)**. A DCT system can't cope with this attack. In contrast, the



(a)



(b)

Figure 6-13 Performance of the DWT system under 2Mbps MPEG2 compression attack for: **(a)** "flower" video sequence and **(b)** "mobile" video sequence.

DWT gives very acceptable performance, especially when using FEC. For example, for 1/5 "nearest" scaling, the capacity is about 80 bpf (bits per frame) which translates to 2 kbps (25 frames per second) and increases to about 140 bpf (3.5 kbps) with FEC.

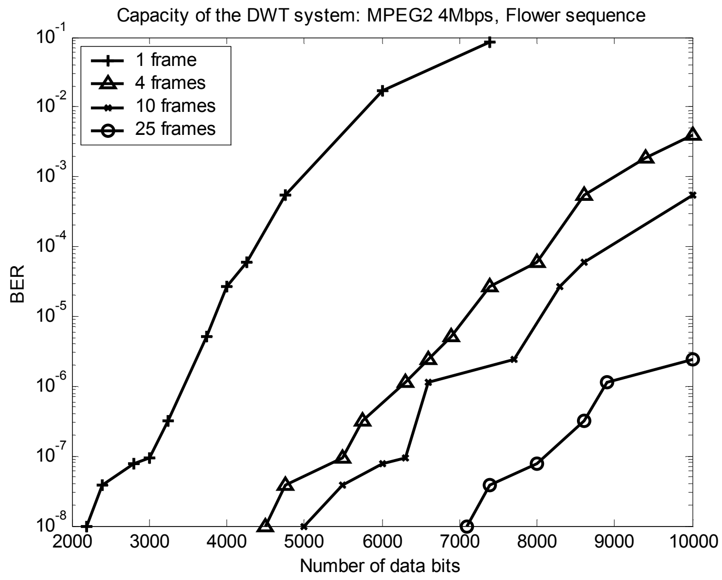
The results for JPEG compression with several different quality factors are presented in **Figure 6-10(a)** and **Figure 6-10(b)**. As **Figure 6-10(b)** indicates, for a relatively high compression factor of 10:1 (25% quality, slight visual artefacts) and with Turbo coding, the wavelet scheme can achieve a capacity of 64 bpf. Even under extreme JPEG compression (30:1 compression, 5% quality, with heavy blocking artefacts) the wavelet scheme still has a capacity of 8 bpf. This attack is illustrated in **Figure 6-8(b)**.

A comparison of the DCT and DWT schemes under JPEG compression attack is shown in **Figure 6-11**. For a quality factor of 40%, the DWT more than doubles the capacity when Turbo coding is used, achieving a capacity over 6 kbps at $BER = 10^{-8}$. This result clearly shows the advantage of FEC. The wavelet scheme is net superior to the DCT scheme, especially for higher quality factors and when using FEC. Again, with FEC, for a quality factor of 40% (7.5:1 compression) the capacity of the DWT scheme is double compared with the DCT scheme.

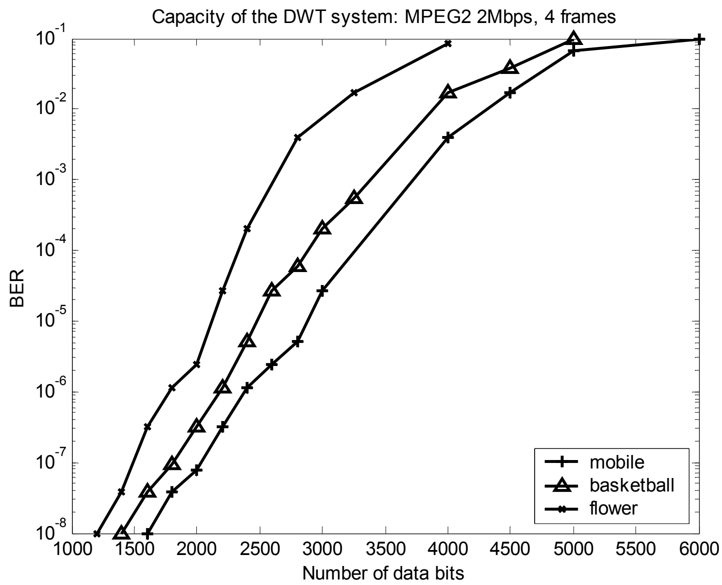
Since the capacity per frame is quite high, we can afford to increase the robustness in the expense of the capacity (trading capacity for robustness) and still achieve a higher capacity compared with the single frame case, by inserting the same watermark in a number of $n\left(n \le 25\right)$ successive frames. In this way the recovery is much simplified since takes place only once, and is easier to combat frame dropping.

This case is illustrated in **Figure 6-13(a)** for MPEG2 compression attack, which gives an impressive capacity of about 1Kbps, when at least 4 frames are averaged together. The improvement between the 4, 10 and respectively 25 frames averaging seems to be quite small, however this is due to the high compression applied in this case (2Mbps); for a medium level of compression the difference between these cases are much more obvious **Figure 6-14(a)**.

It must be said that MPEG2 compression at 2Mbps is a drastic attack, which leads to important visual artefacts **Figure 6-12**, which can be best seen in a moving sequence. The results mentioned above are quite remarkable considering that were obtained for the "Flower garden" sequence, which is notorious for its difficulty and usually gives worse results compared with the other test sequences. This can be seen very well comparing the results obtained for the "Flower garden" sequence with those of the "Mobile" sequence presented in **Figure 6-13(b)**. A direct comparison between three typical sequences is provided in **Figure 6-14(b)**.



(a)



(b)

Figure 6-14 Performance of the DWT system under MPEG2 attack for: **(a)** "flower garden" video sequence, compressed at 4Mbps and **(b)** different video sequences, compressed at 2Mbps, 4 frames averaging.

6.6 Conclusions

The use of the wavelet transform in digital watermarking has many advantages compared with the traditional FFT/DCT transform, fact very well illustrated by the performance of the DWT-based scheme.

The properties of the wavelet transform itself lead to a significant increase in capacity compared with the DCT based systems, in spite of the much simpler HVS used in wavelet's case. From the robustness perspective, wavelets also offer much better results. The results suggest that the DWT has significant advantages under attacks which are likely to be encountered in studios: compression, scaling and cropping.

Under JPEG compression attack, the DWT can more than double the capacity of a DCT system. Subjected to a MPEG2 compression attack, the DWT system can achieve capacities four times as much as the DCT system. For a typical scaling/re-scaling attack, a Turbo coded DWT scheme can yield capacities in excess of 1 kbps, whilst under the same conditions a DCT scheme fails. The DWT scheme has been found to be particularly robust to cropping: for example, the Turbo coded DWT scheme had a capacity of some 37 kbps, compared to 1 kbps for the DCT scheme.

The improved robustness of the DWT scheme is mainly attributed to the spatially local and spatially global support of wavelets. For example, wavelets with local support are less likely to be affected by cropping, compared to the theoretically infinitely long basis functions used in Fourier analysis. The multiresolution feature can also be exploited to optimize retrieval, by embedding all data bits in each sub-band and measuring sub-band SNR. The hierarchical, multiresolution nature of the DWT has also a fundamental advantage for embedding, by performing an analysis similar to that of the HVS. Since the DWT is a HVS model by its nature, even a relatively simple HVS model may suffice. Finally, the DWT has a computational advantage compared with the DCT and it does not suffer from the blocking artefacts so common to the DCT.