Wavelet packet transform-based robust video watermarking technique

GAURAV BHATNAGAR[1,*] and BALASUBRMANIAN RAMAN[2]

[1]Department of Electrical and Computer Engineering, University of Windsor, Windsor-N9C 1M2, ON, Canada
[2]Department of Mathematics, Indian Institute of Technology Roorkee, Roorkee 247 667, India
e-mail: goravb@uwindsor.ca; balaiitr@ieee.org

Abstract. In this paper, a wavelet packet transform (WPT)-based robust video watermarking algorithm is proposed. A visible meaningful binary image is used as the watermark. First, sequent frames are extracted from the video clip. Then, WPT is applied on each frame and from each orientation one sub-band is selected based on block mean intensity value called robust sub-band. Watermark is embedded in the robust sub-bands based on the relationship between wavelet packet coefficient and its 8-neighbour (D_8) coefficients considering the robustness and invisibility. Experimental results and comparison with existing algorithms show the robustness and the better performance of the proposed algorithm.

Keywords. Video watermarking; multiresolution analysis; wavelet packet transform; block mean intensity value.

1. Introduction

Due to the phenomenal development of the computer and network technologies, the reproduction and distribution of digital media become much simpler and faster than before. Therefore, copyright protection has become a very vital issue for the digital media. In the recent years, digital watermarking has drawn much attention of researchers as a potential solution of these issues. The basic idea behind watermarking is to insert an information (the watermark) into a digital media, which can be later extracted or detected for variety of purposes including identification and authentication purposes. The embedding is done in such a way that it must not cause serious degradation to the original digital media. The digital media is either an image or an audio or a video. If any conflict happens to the copyright identification and authentication, the embedded watermark can be extracted to verify the authority. In the recent years, a lot of work has been done for securing images but a few work has been done for the videos. Hence, the present work is concentrated on the robust digital video watermarking. Since, the video can

---

*For correspondence

be regarded as the sequence of the correlated images. Therefore, one can do video watermarking either frame-by-frame or block-wise (block of consecutive frames).

In this paper, a robust video watermarking algorithm is proposed which employs frame-by-frame. To achieve high robustness, the watermark is embedded into wavelet packet coefficients. First, each frame is transformed into various sub-bands via wavelet packet transform (WPT) and then one robust sub-band is selected from each orientation based on block mean intensity value. Watermark embedding is done in wavelet packet coefficients using proposed D_8 neighbouring coefficients mean value. A reliable watermark extraction is developed for extracting watermark from distorted images. Comparison of proposed method is made with the methods described by Hsu & Wu (1998), Hong *et al* (2001), Joumaa & Davoine (2005), Sun & Liu (2004), Alattar *et al* (2003), Hartung & Girod (1998) and Kong *et al* (2006), respectively. Experimental results and comparison with existing algorithms show the robustness and the better performance of the proposed algorithm. Before introducing proposed video watermarking scheme, the requirements of an effective watermarking scheme and the related work are depicted first.

1.1 *Requirements of watermarking schemes*

Generally, a practical watermarking system embeds some copyright information into the host data as a proof of rightful ownership and must meet requirements. Obviously, different applications have different requirements for watermarking system. Therefore, it is quite difficult to have a unique set of requirements that all watermarking system must satisfy. The requirements with respect to copyright protection and rightful ownership are as follows (Cox *et al* 2007).

 (i) *Robustness*: Robustness refers to the ability of the watermark to be preserved even after distortions introduced by standard or malicious data processing, which may be either intentionally or un-intentionally. These distortions are also known as watermarking attacks.
 (ii) *Imperceptibility*: The imperceptibility of the watermark refers to its perceptual transparency. In other words, the human eye should not be able to detect differences between the watermarked and original video.
(iii) *Capacity*: Capacity refers to the maximum amount of information that can be hidden in the media. This directly affects the robustness and perceptual transparency.
(iv) *Security*: Security refers to the fact that un-authorized persons should neither detect nor read the watermark; however, it must be retrieved correctly by the authorized user.

1.2 *Related work*

A lot of video watermarking algorithms have been proposed in the literature employed either in spatial (Mobasseri 2000) or frequency domain (Tsai & Chang 2004; Ge *et al* 2003; Hsu & Wu 1998; Hong *et al* 2001; Liu *et al* 2002; Niu *et al* 2000). Recently, some video watermarking algorithms have been proposed with ICA (Joumaa & Davoine 2005; Sun & Liu 2004), PCA (Mirza *et al* 2007) and SVD (Kong *et al* 2006). The detailed description of these schemes are as follows.

Mobasseri (2000) has proposed a spatial domain watermarking scheme for compressed videos. Authors have showed that it is possible to embed a watermark in raw video and still recover it from MPEG decoder, by exploiting the inherent processing gain of direct sequence spread spectrum. Tsai & Chang (2004) have proposed a novel watermarking scheme for a compressed video sequence via VLC decoding and VLC code substitution. To have better imperceptibility, they used Watson's DCT-based visual model for video watermarking. Ge *et al* (2003) have presented a novel adaptive approach to video watermarking. It takes full advantage

of both intra-frame and inter-frame information of video content to guarantee the perceptual invisibility and robustness of the watermark. A major advantage of this scheme is that the watermark can be extracted without referring to the original video while embedded adaptively in accordance with the human visual system and signal characteristics.

Hsu & Wu (1998) have proposed a MPEG-based technique for digital video watermarking. They embedded watermarks into both intraframe and non-intraframe with different residual masks. For embedding, first original watermark image is degraded by pixel-based permutation and block-based permutation in succession and then embedding is done in DCT domain in middle frequency coefficients that are collected in zig-zag order. Hong *et al* (2001) have proposed DWT based algorithm in which middle frequencies are modified and a flag is generated for extraction process. In the extraction process, another flag is generated from watermarked image and compared with the original flag. In this algorithm, instead of taking watermark image, authors have used generated flag as watermark. Doerr & Dugelay (2003) and Zhong & Huang (2006) have presented video watermarking based on spread-spectrum techniques to improve watermarking robustness. Each watermark bit is spread over a large number of chip rate (Cr) and then modulated by a pseudo-random binary sequence. The robustness of these algorithm increases with the increase of the variance of the pseudo-noise sequence. However, the increase of Cr will reduce the embedding rate of watermark information, where as the increase of variance may result in the perceptibility of the watermark.

Liu *et al* (2002) have proposed a wavelet transform-based video watermarking scheme where multiple information bits are embedded into uncompressed video sequences. Embedding is done in LL sub-band, reducing error probabilities of detection via BHC code. Niu *et al* (2000) have presented a new type of watermarking scheme via two-dimensional and three-dimensional multiresolution signal decomposing. The decomposed watermark image with different resolution is embedded in the corresponding resolution of the decomposed video. To enhance the robustness of watermarking, the watermark information is coded by error correction coding of Hamming code in our algorithm. This technique is robust enough against the attacks of frame dropping, averaging and lossy compression. Mirza *et al* (2007) have proposed a video watermarking scheme based on Principal Component Analysis. They embedded watermark in the three colour channels RGB of an input video file. The main advantage of this scheme is that the same or multiwatermark can be embedded into the three colour channels of the image in order to increase the robustness of the watermark. Furthermore, use of PCA transform allows to choose the suitable significant components to embed the watermark.

Joumaa & Davoine (2005) have proposed a video watermarking scheme, using an informed trellis, applied in two transformed domains obtained by using the DCT transform and an ICA coding technique, respectively. They considered data hiding in digital TV channels where data are compressed using MPEG-2. Sun & Liu (2004) have proposed a blind video watermarking scheme, which embeds a watermark into video independent components by using the four-neighbouring-mean-based method. Alattar *et al* (2003) have proposed an MPEG-4 compressed-domain video watermarking method and tested its performance at low bit-rates. This approach is similar to the approach presented by Hartung & Girod (1998). The only difference is the use of a synchronization template to combat geometric attacks. Both these methods also feature the control algorithm that adjusts the embedding strength of the watermark, depending on local image characteristics. Kong *et al* (2006) have proposed a novel SVD based blind video watermarking algorithm where the watermarks are embedded in specially selected singular values to achieve considerable visible quality and robustness.

This paper is organized as follows: Section 2 gives the brief description of Wavelet Packet Transform. Our proposed watermark embedding and extraction algorithm are described in detail

in section 3. In section 4, experimental results using our watermarking scheme are presented and finally in section 5 the concluding remarks are given.

2. Mathematical preliminaries: Wavelet packet transform

The wavelet packet transform (WPT) (Daubechies 1990; Posch 1992) generalizes the wavelet transform (Mallat 1989; Daubechies 1992) and provides a more flexible tool for the time-scale analysis of the data. All advantages of the wavelet transform are retained because the wavelet bases are in the repertoire of bases available with the wavelet packet transform. Given this, the WPT may eventually become a standard tool in signal and image processing. Using a pair of low and high-pass filters to split a space corresponds to splitting the frequency content of a signal into roughly a low and a high-frequency components. In wavelet decomposition, we leave the high-frequency part alone and keep splitting the low-frequency part. Also in wavelet packet decomposition, we split the high-frequency part into low and high-frequency parts. So, in general, wavelet packet decomposition divides the frequency space into various parts and allows better frequency localization of signals.

Basically, wavelet decomposition re-expresses a function in the terms of the wavelet basis. This amounts to decomposing the function space $L^2(\mathbb{R})$ into a direct sum of orthogonal subspaces and choosing the combination of the orthonormal bases as the basis for $L^2(\mathbb{R})$. In the case of finite data with information up to a resolution level J, a wavelet transform performs a decomposition of the space V_J into a direct sum of orthogonal subspaces

$$V_J = W_{J-1} \oplus V_{J-1} = W_{J-1} \oplus V_{J-2} \oplus V_{J-2} = ... = \oplus_{j=0}^{J-1} W_j \oplus V_0 \tag{1}$$

and the union of the bases of these subspaces forms a basis for the wavelet decomposition. This, of course, is by no means the only way to decompose the space $L^2(\mathbb{R})$ or V_J. After generalizing the wavelet decomposition, resulting in a new family of orthonormal bases for function space, which is nothing but wavelet packet decomposition of the space.

It is evident from multiresolution analysis that if the basis functions $\{\phi_{1,k}(t)\}$ of V_1 then $\{\phi(t-k)\}$ and $\{\psi(t-k)\}$ constitute an orthonormal basis for V_0 and W_0 such that $V_1 = W_0 \oplus V_0$, where

$$\phi(t) = \sqrt{2} \sum_k h_k \, \phi(2t-k) \text{ and } \psi(t) = \sqrt{2} \sum_k g_k \, \phi(2t-k). \tag{2}$$

So the V space can be decomposed into a direct sum of the two orthogonal subspaces defined by their basis functions given by the equation 2. Hence, W space is also decomposed using the same splitting algorithm. Mathematically,

$$w_2(t) = \sqrt{2} \sum_k h_k \, \psi(2t-k) \text{ and } w_3(t) = \sqrt{2} \sum_k g_k \, \psi(2t-k), \tag{3}$$

where $\{w_2(t-k)\}$ and $\{w_3(t-k)\}$ are orthonormal basis functions for the two subspaces whose direct sum is W_1. Further,

$$w_4(t) = \sqrt{2} \sum_k h_k \, w_2(2t-k) \text{ and } w_4(t) = \sqrt{2} \sum_k g_k \, w_2(2t-k), \tag{4}$$

where $\{w_4(t-k)\}$ and $\{w_5(t-k)\}$ are orthonormal basis functions for the two subspaces whose direct sum is W_2. In general, a sequence of functions for $n = 0, 1, 2, 3, ...$ is defined as follows:

$$w_{2n}(t) = \sqrt{2} \sum_k h_k \, w_n(2t-k) \text{ and } w_{2n+1}(t) = \sqrt{2} \sum_k g_k \, w_n(2t-k). \qquad (5)$$

If we take $n = 0$ then we get scaling function i.e., $w_0(t) = \phi(t)$ and $n = 1$ yields the mother wavelet i.e., $w_1(t) = \psi(t)$. So far we have been using the combination of $\{\phi(2^j t - k)\}$ and $\{\psi(2^j t - k)\}$ to form a basis for V_J but now we have a sequence of functions $\{w_n(t)\}$ and the combinations of these with their dilations and translations can represent the orthonormal bases for the function space. This collection is called wavelet packet bases and the function of the form $\{w_{n,j,k} = 2^{j/2} \, w_n(2^j t - k)\}$ is called a wavelet packet. The following diagram (figure 1)
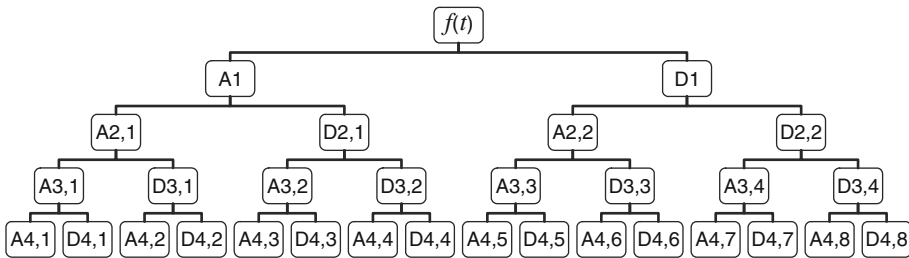

Figure 1. Decomposition of the signal $f(t)$ via WPT.
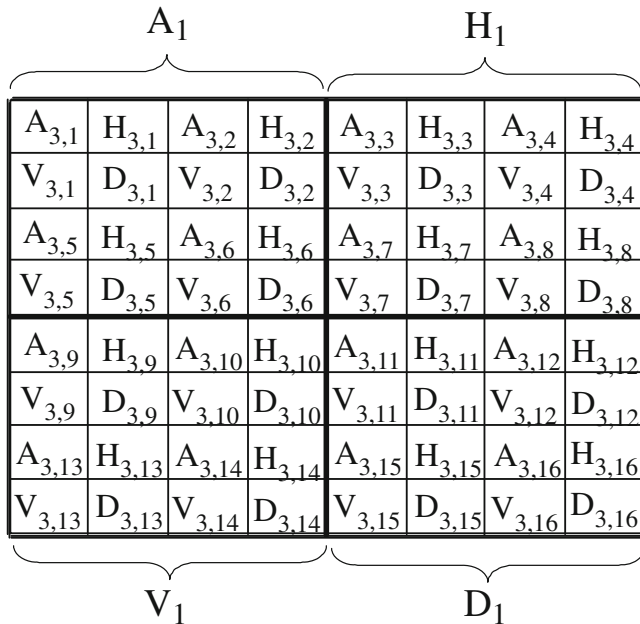

Figure 2. Orientation of WPT in proposed algorithm.

illustrates the decomposition of the signal $f(t)$ using wavelet packets. The complete proof of the above can be found in Coifman *et al* 1992 and Mallat 1998.

In the case of images, the orientation is first defined in WPT (as used in the proposed algorithm). From the literature, it is evident that 1-level wavelet transform and 1-level wavelet packet transform are same (Percival & Walden 2000). Therefore, the orientation for any sub-band from A_1, H_1, V_1, D_1 are A, H, V, D, respectively (see figure 2).

3. Proposed video watermarking algorithm

In this section, we discuss some motivating factors in the design of our approach to video watermarking scheme. WPT is used for developing the algorithm. Let us consider $\mathcal{V}$ is the host video and $\mathcal{W}$ is the binary watermark image. The host video is a colour video of size $M \times N \times 3 \times n$. We save the positions of robust sub-bands for extraction process. The watermark $\mathcal{W}$ is a binary watermark of size $M_w \times N_w$. The block diagram of proposed algorithm is shown in figure 3.

3.1 *Embedding process*

Embedding Process is given as follows:

1. Sequent video frames are extracted from the video $\mathcal{V}$, denoted by F^i, where i is the total number of frames and $i \in [1, n]$.
2. Transform the RGB colour space of the extracted video frames to the YCbCr colour space. The luminance (Y) component has been selected from YCbCr to apply the WPT, since the luminance has the same range as the gray image between 0 and 255. Let us denote luminance component for each video frame by F_Y^i.
3. Perform l-level *discrete wavelet packet transform* on F_Y^i, which is denoted by $(f_Y^i)_{l, p}^{\theta}$, where $\theta \in \{A, H, V, D\}$ and $p = 1, 2, 3,..., 2^{l-1} \times 2^{l-1}$.
4. Calculate block mean intensity value for all sub-bands

$$M_{(f_Y^i)_{l, p}^{\theta}} = \frac{1}{m_1 \times n_1} \sum_{j=1}^{m_1} \sum_{k=1}^{n_1} C(j, k), \tag{6}$$
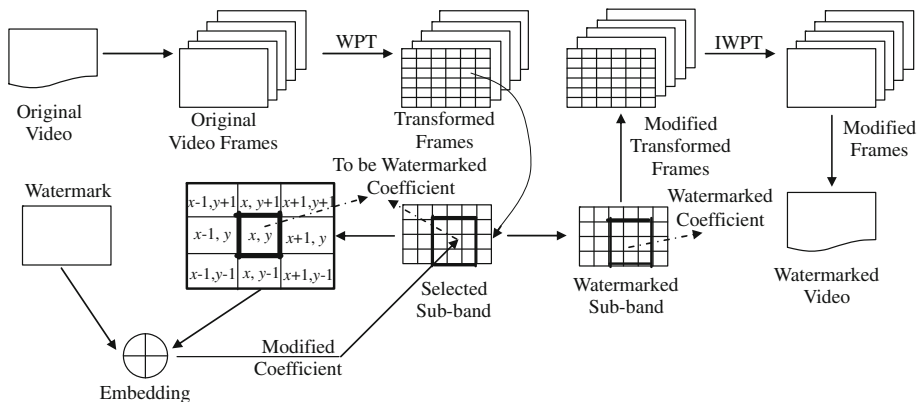

Figure 3. Block diagram of proposed video watermarking algorithm.

where $C(j, k)$ represents the magnitude of the coefficients and $m_1 \times n_1$ is the total number of coefficients in the sub-band $(f_Y^i)_{l,\ p}^\theta$.

5. Select one robust sub-band from each orientation, based on block mean intensity value, denoted by $(f_Y^i)_{sel}^\theta$. Select that sub-band as robust sub-band whose block mean intensity value is maximum.

6. Embed watermark $\mathcal{W}$ in the robust sub-bands as follows:

   (a) Computation of D_8 mean value: If the coefficient value of the watermark embedding position is $c_{x,y}$, then D_8 mean value $\mathcal{D}_{x,y}$ is computed as follows:

$$\mathcal{D}_{x,y} = \frac{1}{8} \left(\sum_{j=x-1}^{j=x+1} \sum_{k=y-1}^{k=y+1} c_{j,k} - c_{x,y} \right) \tag{7}$$

   (b) Modify coefficient value as follows:

   • if $c_{x,y} > \mathcal{D}_{x,y}$, $W(x, y) = 0$ or $c_{x,y} < \mathcal{D}_{x,y}$, $W(x, y) = 1$ then,

$$\begin{aligned} c'_{x,y} &= c_{x,y} + \alpha(c_{x,y} - \mathcal{D}_{x,y}), \\ \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) &= 1, \end{aligned} \tag{8}$$

   • if $c_{x,y} > \mathcal{D}_{x,y}$, $W(x, y) = 1$ or $c_{x,y} < \mathcal{D}_{x,y}$, $W(x, y) = 0$ then,

$$\begin{aligned} c'_{x,y} &= c_{x,y} - \alpha(\mathcal{D}_{x,y} - c_{x,y}), \\ \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) &= 0. \end{aligned} \tag{9}$$

   • If $c_{x,y} = \mathcal{D}_{x,y}$, $W(x, y) = 1$ then,

$$\begin{aligned} c'_{x,y} &= c_{x,y} + \alpha(c_{x,y} - \mathcal{D}_{x,y}), \\ \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) &= 1. \end{aligned} \tag{10}$$

   • If $c_{x,y} = \mathcal{D}_{x,y}$, $W(x, y) = 0$ then,

$$\begin{aligned} c'_{x,y} &= c_{x,y} - \alpha(\mathcal{D}_{x,y} - c_{x,y}), \\ \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) &= 0, \end{aligned} \tag{11}$$

   where α is the watermark strength. $\mathcal{F}_{(f_Y^i)_{sel}^\theta}$ is a matrix to determine the existence of watermark.

7. Map modified robust sub-bands to their original positions and perform l-level *inverse discrete wavelet packet transform* to construct watermarked video frames.

8. Constitute all watermarked video frames to get watermark video.

3.2 *Extraction process*

In the extraction procedure, the estimate of the original watermark is obtained. For watermark extraction from watermarked image, original image is not required. Hence this extraction is called *blind*. The extraction process is formulated as follows:

1. Sequent video frames are extracted from the watermarked video V_w, denoted by $(F_w)^i$, where i is the total number of frames and $i \in [1, n]$.

2. Transform the RGB colour space of the extracted watermarked video frames to the YCbCr colour space. Let us denote luminance component for each watermarked video frame by $(F_w)_Y^i$.

3. Perform l-level *discrete wavelet packet transform* on $(F_w)_Y^i$, which is denoted by $(fw_Y^i)_{l,p}^\theta$, where $\theta \in \{A, H, V, D\}$ and $p = 1, 2, 3,..., 2^{l-1} \times 2^{l-1}$.

4. Select watermarked robust sub-band from each orientation, based on saved positions, which are denoted by $(fw_Y^i)_{sel}^\theta$.

5. Extract watermark W' from the watermarked robust sub-bands as follows:

    (a) Compute D_8 mean value for each coefficients of watermarked robust sub-bands by equation 7.

    (b) Extract watermark as follows:

$$W_i'(x, y) = 0 \begin{cases} c_{x,y}' > \mathcal{D}_{x,y}', \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) = 0; \\ c_{x,y}' < \mathcal{D}_{x,y}', \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) = 1; \\ c_{x,y}' = \mathcal{D}_{x,y}', \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) = 0. \end{cases}$$

$$W_i'(x, y) = 1 \begin{cases} c_{x,y}' > \mathcal{D}_{x,y}', \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) = 1; \\ c_{x,y}' < \mathcal{D}_{x,y}', \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) = 0; \\ c_{x,y}' = \mathcal{D}_{x,y}', \mathcal{F}_{(f_Y^i)_{sel}^\theta}(x, y) = 1. \end{cases}$$

6. Extracted watermark from each watermarked video frame and finally taking exclusive OR operation on all extracted patterns to get the resultant watermark i.e.,

$$W^{ext} = \mathbf{XOR}(W_i'). \tag{12}$$

4. Results and discussions

4.1 *Experimental set-up*

In order to explore the performance of proposed video watermarking algorithm, two videos of size $352 \times 288 \times 3 \times 20$ are considered, named mobile and template. Two different binary mean-ingful logos are used as watermarks. As the watermark logo, IEEE (figure 3c) and CVGIP LAB logo are used (figure 4c). The size of watermark logo is 88×72. Logo IEEE is embedded into mobile video, while CVGIP LAB logo is embedded into template video. The watermarked video quality is measured using PSNR (Peak Signal to Noise Ratio). For a video, PSNR is calculated by taking average of PSNR values of all frames and called Average PSNR. Watermarked mobile and template videos are having average PSNR values of 61.4308 dB and 67.1915 dB, respectively. For watermark embedding, strength factor α is set to 0.01. Daubechies filter coefficients are used and 2-level of decomposition is performed, for wavelet packet decomposition. For the better analysis of a signal, we should select our 'mother wavelet' carefully. Mother wavelet will not only determine how the original signal is well-estimated in terms of translation and dilation, but also, it will affect the frequency spectrum of the signal. The choice of mother wavelet can be based either on the cumulative energy over some interval of interest or based on similarity between original and reconstructed signals. We choose to select the mother wavelet based on the similarity. For this purpose, *PSNR* is calculated between original and reconstructed image and that mother wavelet is chosen which gives the maximum *PSNR*.

To verify the presence of watermark, different measures can be used to show the similarity between the original and the extracted watermarks. In the proposed algorithm, the Normalized Hamming Similarity (NHS) is used to evaluate the effectiveness of the proposed algorithm. The NHS between the embedded binary watermark W and that extracted $\widetilde{W}$ is defined as:

$$NHS = 1 - \frac{HamDist(W, \widetilde{W})}{M_w \times N_w},$$

(13)

where $HamDist(\,\cdot\,,\,\cdot\,)$ denotes the Hamming Distance between two binary strings, i.e., the number of different bits in the two binary strings. We can easily show that $NHS \in [0, 1]$. If higher NHS values are acquired, then the extracted watermark is more similar to the original one. If $NHS > T_s$, where T_s is the threshold for watermark detection, then the watermark is detected. In our experiments, the threshold for watermark detection is $T_s = 0.6$.

Original video frame, watermarked video frame, original logo and extracted logo images are shown in figures 4 and 5, for Mobile and Template video, respectively. If we observe original and watermarked video frame, we cannot find any perceptual degradation. In table 1, average PSNR and NHS are given for both the videos and frame-wise PSNR and NHS values are shown in figures 6 and 7, respectively. The complete videos for visual assessment can be found on our website (http://gaurav.iitr.21.googlepages.com/video_water).

4.2 *Selection of threshold*

In watermarking algorithm, the threshold value plays an important role for the extraction of watermark as well as to maintain the robustness of the algorithm. In this context, the threshold value should be selected to make a balance between these two requirements. Here, we performed
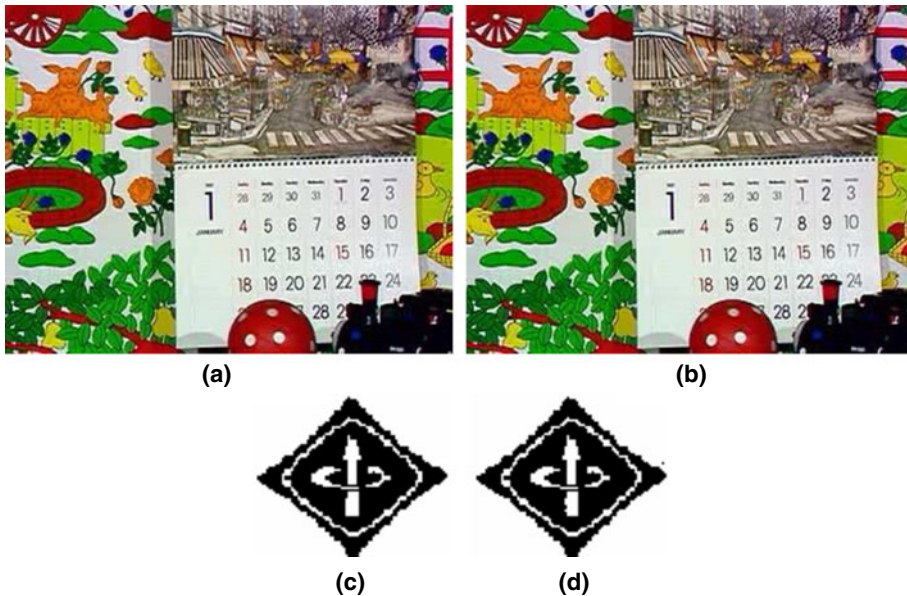

Figure 4. (**a**) First mobile video frame. (**b**) Watermarked first frame. (**c**) Original logo. (**d**) Extracted logo.
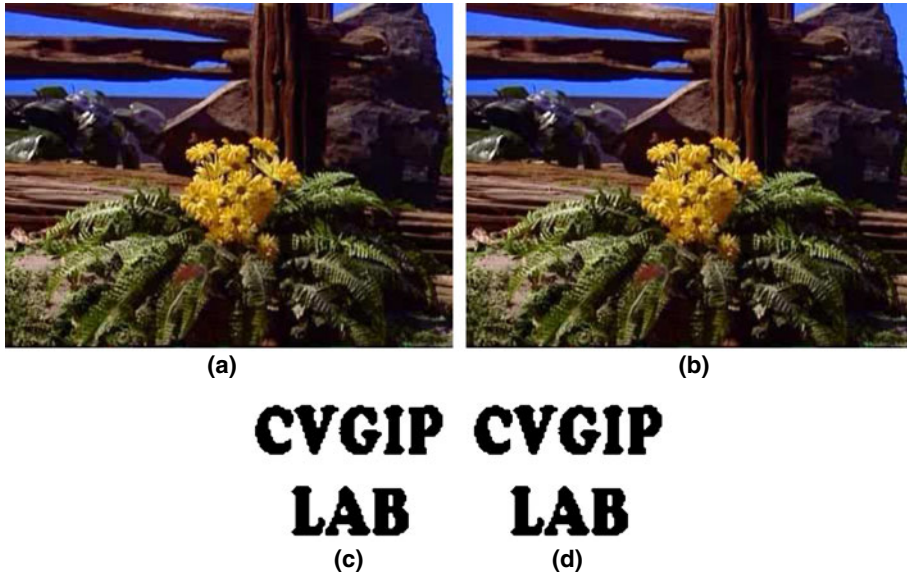
Figure 5. (**a**) First template video frame. (**b**) Watermarked first frame. (**c**) Original logo. (**d**) Extracted logo.

a number of experiments to compute the optimal value of T_s. We generated 25 binary matrices randomly and embedded in the host video by the proposed algorithm. In each case we extract the watermark and calculate NHS between original watermark and extracted binary matrix and taking the maximum value of NHS as desired threshold. A graph is plotted (figure 8) by taking the number of binary matrices on x-axis and corresponding NHS on y-axis. The corresponding highest value (i.e., the highest peak on plot) is taken as the optimal value for threshold.

Without loss of generality and for our convenience, we compute threshold via generating 25 binary matrices randomly but for the good security one can generate large number of matrices and then threshold must be computed. Since, threshold is determined using similarity between original watermark and extracted binary random matrix. Therefore, the threshold value will vary for each video and the used watermark. As a result, the change in the threshold value with respect to the number of generated random matrices and videos are discussed here. For experiments, we have used mobile and template videos with IEEE and CVGIP LAB logos as watermark. First threshold is determined with 25 randomly generated binary matrices. Then, we increase the number to 100 and then by multiple of 100. The final number which we have used is 10000. For each set of random matrices, we calculate the maximum NHS between original watermark and extracted random matrix and the corresponding results are depicted in figure 9. From figure 9, it is clear that if we increase the number of random matrix then the threshold will also

Table 1. PSNR and NHS for mobile and template videos.

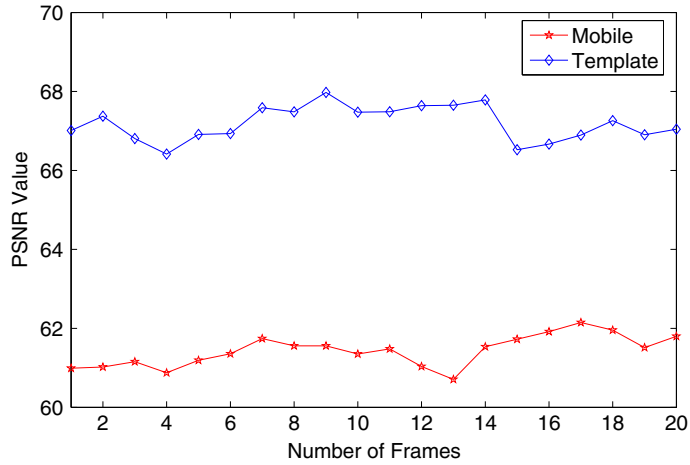| Image	Average PSNR (in DB)	NHS
Mobile	61.4308	0.9983
Template	67.1915	0.9981

Figure 6. PSNR (in DB) values for all video frames.

increase and after one stage it become constant. Finally, maximum of all thresholds corresponding to each set of random matrices is set to be optimal threshold value. In the case of Mobile video, threshold comes out to be 0.6195, whereas it is 0.6054 for template video and the optimal threshold is comes out to be 0.6195(=max(0.6195,0.6054)) which is close to the threshold used in the proposed method i.e., 0.6.

4.3 *Histogram analysis*

A histogram is basically a graphical representation of the tonal distribution in the image. In other words, a histogram is a graph which shows the size of the area of the image that is captured for each tonal variation that the camera is capable of recording. By looking at the histogram for a
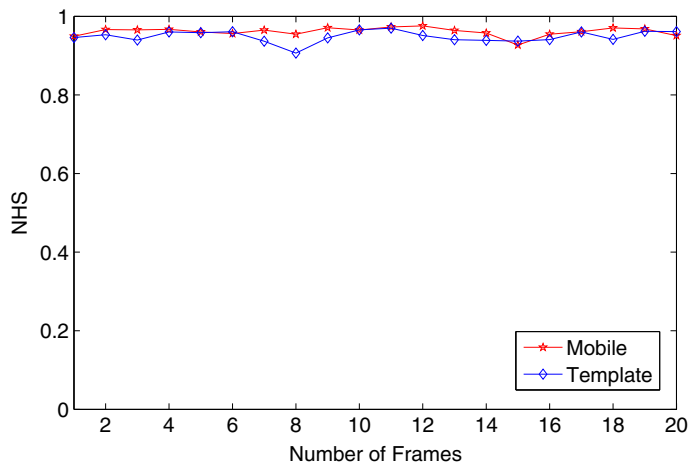

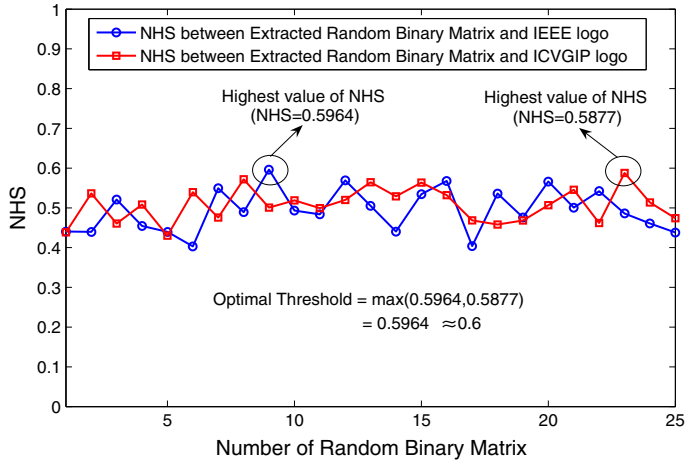**Figure 7.** NHS for extracted watermark from each video frame.

Figure 8. Selection procedure for threshold.

specific image you will be able to judge the entire tonal distribution at a glance. It tells you how much detail is at the light end of the spectrum and how much at the dark.

According to human visual system, a watermarking scheme is good and effective if we cannot differentiate the original and watermarked video (or images) histogram (Shannon 1949; Bhatnagar & Raman 2010). But it is possible only for loss-less watermarking schemes. Since almost every watermarking scheme is lossy hence there are some changes in some tonal values. Hence, for lossy watermarking schemes the changes can be viewed at the peaks. In this case, scheme is said to be efficient if the variation is very less and overall shape of histogram is maintained. In figures 10 and 11, the histogram analysis for first frame of both the videos are given. It is clear from the figures that the overall shape of the histogram is maintained.
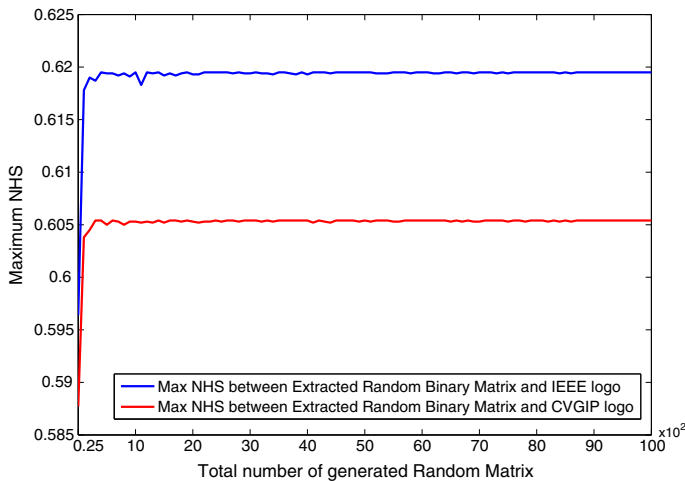

Figure 9. Variation of threshold with the number of generated random matrices.
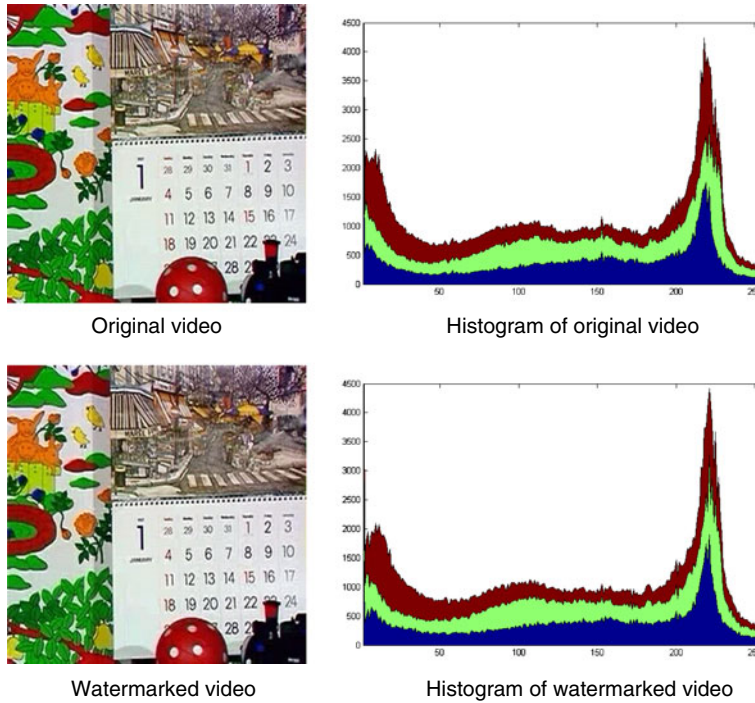
Original video

Histogram of original video

Watermarked video

Histogram of watermarked video

Figure 10. Histogram analysis of first mobile video frame.

Original video

Histogram of original video

Watermarked video

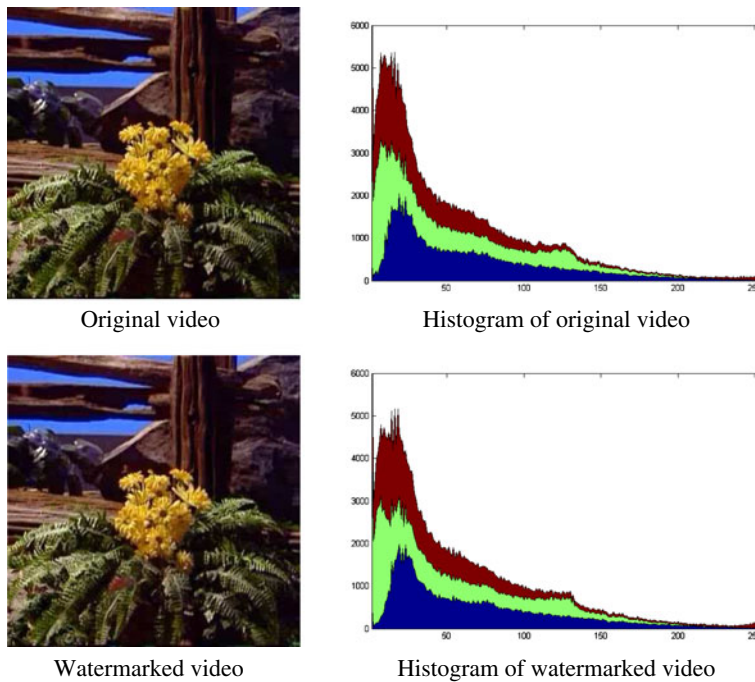Histogram of watermarked video

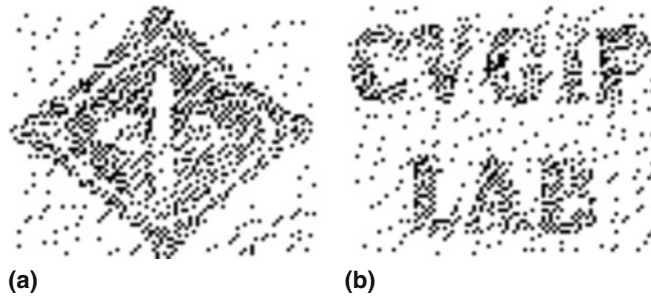**Figure 11.** Histogram analysis of first template video frame.

Figure 12. Watermark extracted after average filtering attack from (**a**) mobile. (**b**) Template video.

4.4 *Attack analysis*

To measure the robustness of the proposed algorithm, various attacks are used like average and median filtering, frame dropping, MPEG coding, resizing and rotation. The most common manipulation in digital media is filtering. In figures 12 and 13, extracted watermark after applying 3×3 averaging and median filtering have been shown. To verify the robustness of the proposed algorithm, another measure is frame dropping. In our experiments, we have dropped P frames randomly. The missing frame is replaced with the average of the two neighbouring frames. In figure 14, the extracted watermark from 15 frame dropped video has been shown. In real life applications, storage and transmission of digital data, a lossy coding operation is often performed on the data to reduce the memory and increase transmission efficiency. Hence, we have also tested our algorithm for the MPEG coding and the result is shown in figure 15. To test the MPEG coding, watermark video is coded at 1.5 Mbps. We have also tested our algorithm for resizing and rotation attack. For resizing, first the size of each video frame is reduced to 128×128 and again carried back to the original size (figure 16). Even though the wavelet packet transform is not rotational invariant, proposed method can extract the watermark for small rotations. Extracted logos from $0.5°$ are shown in figure 17.

4.5 *Comparative analysis*

Proposed algorithm is compared with the different existing algorithms. Among these two algorithms (Hsu & Wu 1998; Hong *et al* 2001) are proposed for images. Due to the similar
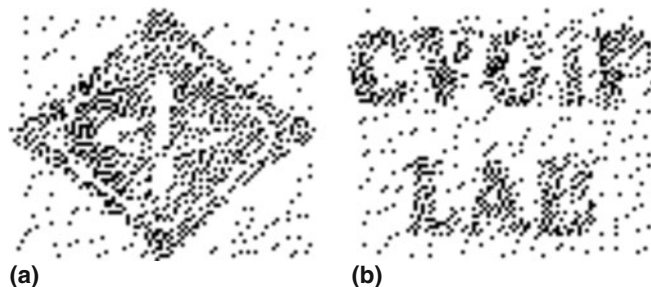

Figure 13. Watermark extracted after median filtering attack from (**a**) mobile. (**b**) Template video.

Figure 14. Watermark extracted after frame dropping attack from (**a**) mobile. (**b**) Template video.

Figure 15. Watermark extracted after MPEG coding attack from (**a**) mobile. (**b**) Template video.

Figure 16. Watermark extracted after resize attack from (**a**) mobile. (**b**) Template video.

Figure 17. Watermark extracted after rotation attack from (**a**) mobile. (**b**) Template video.

Table 2. NHS for extracted watermarks from mobile video after different attacks.

| Attacks | Hsu & Wu | Existing algorithms by | | | | | | Proposed algorithm |
		Hong et al	Joumaa & Davoine	Sun & Liu	Alattar et al	Hartung & Girod	Kong et al	
Average filtering	0.5109	0.7356	0.7021	0.7160	0.7427	0.6567	0.7598	0.7614
Median filtering	0.4829	0.7134	0.6978	0.6948	0.7568	0.6142	0.7564	0.7577
Frame dropping	0.6143	0.9418	0.9143	0.9342	0.9495	0.6498	0.9536	0.9508
MPEG coding	0.5190	0.7896	0.7673	0.7655	0.8186	0.5667	0.8131	0.8138
Rotate	0.5978	0.7834	0.7909	0.7952	0.8165	0.6022	0.7437	0.8247
Resizing	0.6085	0.8141	0.8151	0.8135	0.8403	0.6678	0.7717	0.8619

Table 3. NHS for extracted watermarks from template video after different attacks.

| Attacks | Hsu & Wu | Existing algorithms by | | | | | | Proposed algorithm |
		Hong et al	Joumaa & Davoine	Sun & Liu	Alattar et al	Hartung & Girod	Kong et al	
Average filtering	0.5242	0.7531	0.7248	0.7256	0.7887	0.5456	0.7931	0.8106
Median filtering	0.4997	0.7688	0.7232	0.7277	0.7957	0.5245	0.7957	0.8098
Frame dropping	0.6279	0.9483	0.8946	0.9134	0.9380	0.6243	0.9609	0.9612
MPEG coding	0.5353	0.8073	0.8063	0.8147	0.7961	0.5548	0.8160	0.8299
Rotate	0.6011	0.8116	0.8407	0.8491	0.6225	0.6207	0.8056	0.9398
Resizing	0.6114	0.8191	0.7838	0.7922	0.6441	0.6329	0.8032	0.8884

mechanics, these two algorithms are extended and implemented for videos. Apart form these, all other algorithms are proposed specifically for videos and these includes Joumaa & Davoine (2005), Sun & Liu (2004), Alattar *et al* (2003), Hartung & Girod (1998) and Kong *et al* (2006), respectively. In tables 2 and 3, comparisons are given in detail for mobile and template video, respectively. From tables, it is clear that NHS is greater for proposed algorithm, which shows the better performance of the proposed algorithm. The algorithms presented in Hsu & Wu (1998) and Hartung & Girod (1998) give baseline results. These methods give poor results relative to other algorithms. The algorithms in Joumaa & Davoine (2005) and Sun & Liu (2004) give better results than baseline results but are not comparable with other algorithms. This was expected because these algorithms have no scale selectivity. This drawback is rectified in multiresolution based algorithms (Hong *et al* 2001; Alattar *et al* 2003). Though the performance of these algorithms are close to the proposed algorithm but only in specific situations. The algorithm in Alattar *et al* (2003) works in the compressed domain and especially designed for MPEG4 compressed videos where as algorithm in Kong *et al* (2006) works in the uncompressed domain like proposed algorithm but is not robust against resizing and rotation attacks.

5. Conclusions

Present research on watermarking mainly concentrates on the still images. In this paper, a new robust video watermarking scheme is presented and employed in wavelet packet transform domain. The watermark is a meaningful binary logo instead of randomly generated Gaussian sequence. Watermark is embedded not only in higher, but also in lower frequency without any degradation in the video. Experiments show better robustness against different attacks, specially against frame dropping and MPEG coding. The proposed algorithm is simple, efficient and with less complexity. In the present work, we already considered small rotations and the problem with large rotations can be addressed in the future.

Acknowledgements

References

Alattar A M, Lin E T and Celik M U 2003 Digital watermarking of low bit rate advanced simple profile MPEG-4 compressed video. *IEEE Transactions on Circuits and Systems for Video Technology* 13(8): 787–800

Bhatnagar G and Raman B 2010 Distributed multiresolution discrete Fourier transform and its application to watermarking. *Int. J. Wavelets, Multiresolution and Information Processing* 8(2): 225–241

Coifman R R, Meyer Y and Wickerhauser M V 1992 *Wavelet analysis and signal processing*. In: M B Ruskai *et al* (eds.), *Wavelets and their applications*. Boston: Jones and Bartlett, pp. 153–178

Cox I, Miller M, Bloom J, Fridrich J and Kalker T 2007 *Digital watermarking and steganography*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Daubechies I 1090 The wavelet transform timeŰfrequency localization and signal analysis. *IEEE Transactions on Information Theory* 36(5): 961–1005

Daubechies I 1992 *Ten lectures notes on wavelet*. Philadelphia USA: SIAM

Doerr G and Dugelay J L 2003 A guide tour of video watermarking. *Signal Processing: Image Communication* 18(4): 263–282

Ge Q, Lu Z and Niu X 2003 Oblivious video watermarking scheme with adaptive embedding mechanism. *Proc. Int. Conf. Machine Learning and Cybernetics*, Xian, China 5: 2876–2881

Hartung F and Girod B 1998 Watermarking of uncompressed and compressed video. *Signal Processing* 66(3): 283–301

Hong I, Kim I and Han S S 2001 A blind watermarking technique using wavelet transform. *Proc. IEEE Int. Sym. Industrial Electronics*, Pusan, Korea 3: 1946–1950

Hsu C T and Wu J L 1998 A DCT-based watermarking for videos. *IEEE Transactions on Consumer Electronics* 44(1): 206–216

Joumaa H and Davoine F 2005 Performance of an ICA video watermarking scheme using informed techniques. *Proc. IEEE Int. Conf. Image Processing*, Genoa, Italy 1: 261–264

Kong W, Yang B, Wu D and Niu X 2006 SVD based blind video watermarking algorithm. *Proc. Int. Conf. Innovative Computing, Information and Control*, Beijing, China 1: 265–268

Liu H, Chen N, Huang J, Huang X and Shi Y Q 2002 A robust DWT-based video watermarking algorithm. *Proc. IEEE Int. Sym. Circuits and Systems*, Scottsdale, Arizona 3: 631–634

Mallat S G 1989 A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transcations on Pattern Analysis and Machine Intelligence* 11(7): 674–693

Mallat S G 1998 *A wavelet tour of signal processing*. San-diego, CA, USA: Academic Press

Mirza H H, Thai H D, Nagata Y and Nakao Z 2007 Digital video watermarking based on principal component analysis. *Proc. Int. Conf. Innovative Computing, Information and Control*, Kumamoto, Japan 290–294

Mobasseri B G 2000 A spatial digital video watermark that survives MPEG. *Proc. Int. Conf. Information Technology: Coding and Computing*, Las Vegas, USA 68–73

Niu X, Sun S and Xiang W 2000 Multiresolution watermarking for video based on gray-level digital watermark. *IEEE Transcations on Consumer Electronics* 46(2): 375–384

Percival D B and Walden A T 2000 *Wavelet methods for time series analysis*. New York, USA: Cambridge University Press

Posch T 1992 The wavelet packet transform (WPT) as applied to signal processing. *Proc. IEEE-SP Int. Sym. TimeFrequency TimeScale Anal*, Victoria, BC, Canada 143–146

Shannon C E 1949 Communication theory of secrecy system. *Bell System Technical Journal* 28(4): 656–715

Sun J and Liu J 2004 Data hiding with video independent components. *IEEE Electronics Letters* 40(14): 858–859

Tsai H M and Chang L W 2004 Highly imperceptible video watermarking with the Watson's DCT-based visual model. *Proc. IEEE Int. Conf. on Multimedia and Expo*, Taipei, Taiwan 3: 1927–1930

Zhong J D and Huang S T 2006 An enhanced multiplicative spread spectrum watermarking scheme. *IEEE Transactions on Circuits and Systems for Video Technology* 16(12): 1491–1506