

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

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

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

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

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



---

# Watermarking Technique for Multimedia Documents in the Frequency Domain

---

Maha Bellaaj and Kaïs Ouni

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.79370>

---

## Abstract

In order to secure and maintain the authenticity and integrity of multimedia documents, we use digital watermarking. This discipline can be applied to images, audios, and videos. For this reason, and to be independent of the nature of the signal composing the document to be watermarked, we will propose in this chapter two watermarking techniques, one for the audio and another for the image to watermark a video containing the two components audio and image. MDCT is combined with Watson model and a motion detection algorithm in the image watermarking technique and is combined with a psychoacoustic model to elaborate the audio watermarking technique. For the two techniques, the bits of the mark will be duplicated to increase the capacity of insertion and then inserted into the least significant bit (LSB). We will use an error correction code (Hamming) on the mark for more reliability in the detection phase. To highlight our experimental results point of view robustness and imperceptibility, we will compare the proposed techniques with some other existing techniques.

**Keywords:** multimedia documents, watermarking, MDCT, Watson model, motion detection, psychoacoustic model, hamming

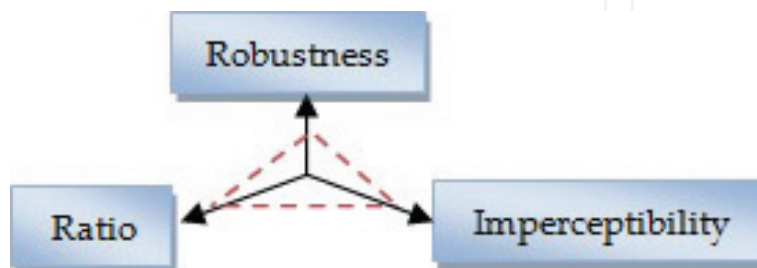
---

## 1. Introduction

The spread of multimedia documents and by virtue of the development of technologies in connection with the computer directs the world toward an era where the digital takes a primordial place. In addition, the development of the Internet and, more generally, the new means of communication authorized the large-scale dissemination of digital data. Despite the mentioned advantages, we are facing serious problems: multimedia documents become unprotected, digital data are distributed in an illegal manner, and copyrights are unprotected.

Where does the digital watermarking come from as a security mechanism complementary to encryption? Its basic idea is to insert the information in a robust and imperceptible way in multimedia documents [1]. On after the literature, digital watermarking has received substantial interest as a research topic in the 90s [2, 3]. For the past 28 years, the work on digital watermarking continue to multiply in order to find watermarking techniques for multimedia documents that must meet the following criteria: robustness against a maximum number of attacks and manipulations, high capacity insertion, and imperceptibility of the mark. An appropriate watermarking system must provide the best compromise between these three main features (**Figure 1**).

A watermarking system is formed mainly by two processes: insertion and detection. A mark  $W$  is inserted in a multimedia document  $M$  to obtain the watermarked document  $M'$  by applying the insertion process. In some watermarking systems, we can use a secret key  $C$  to perform the insertion. The marked document  $M'$  can undergo transformations, and we obtain the resulting document  $M''$ . Subsequently, we move to the detection of the mark. There are several detection schemes which we quote: the private scheme where the original digital document is given to the detector, the mark is detected by comparing the original with the watermarked, and the semi-private scheme which gives an answer in the presence or absence of the mark (true or false) without using the original document and the blind scheme, in which only the secret key is needed to extract the mark. To design a watermarking system, the choice of the insertion area is considered as a very important step [4, 5]. We can distinguish three major fields of insertion: the domain without transformation (spatial domain and time domain), the frequency domain, and the multi-resolution domain. The domain without transformation can be the spatial domain for the image and the video and the time domain for the audio. One of the advantages of the methods operating in this field is that they are very fast, since no initial treatment is necessary. However, such a domain does not offer much resistance against existing attacks. The frequency domain is obtained after the application of a transformation such as fast Fourier transform (FFT), discrete cosine transformation (DCT) [6], etc. The most important benefit of using the transformed domain is that it is already used to prepare multimedia information in communication standards such as JPEG for still images [7], MPEG2 for video sequences [8], and MPEG1 for audio [9]. Techniques operating in the frequency domain have the advantage of being robust against the compression operation, since they use the same space that is used for coding. The development of new compression standards such as JPEG2000 [7] and MPEG4 [8] has led researchers to use other insertion domains as the multiresolution domain [10]. The information represented in this area is well localized in



**Figure 1.** Compromise between robustness, ratio, and imperceptibility.

frequency and time. The sub-band decomposition allows isolating the low frequency components. The middle and high components constitute a less sensitive insertion space.

In the following, we will present some watermarking techniques for video existing in the literature.

- Shaveta and Daljit [11]: in this technique, the authors apply the SWT to the images of the video. Subsequently, they apply the SVD to each subband of the red layer. Then, they change the singular values of the HH band with the singular values of the HH band of the brand. For the other two layers, they select the block with the highest S values and then apply the DCT to the selected band. Finally, they insert the mark on each of the selected bands. The detection scheme is the inverse of that of insertion.
- Shital et al. [12]: In this article, the author used a watermarking technique to detect tampering in a video. The technique operates in the frequency domain using DCT as a transformation. After generating the mark (hash value of the frame, the micro-block numbers, and the frame number), the latter is inserted into the frames in the frequency domain. The insertion is done by replacing the LSB of the highest non-zero DCT coefficient by the bit of the corresponding mark.
- Supriya and Navin [13]: in this chapter, the author proposes a hybrid technique for video based on the discrete wavelet transform and singular value decomposition. In this technique, the mark is inserted into the original video images by first converting it into the YCbCr color space. Next, the luminance portion (Y component) is broken down into four subbands using a discrete wavelet transform. Finally, the singular values of the sub-band LL are perceptually shaped by singular values of the image of the watermark. The detection scheme is the inverse of that of insertion.

In this chapter, we will propose a watermarking system for multimedia documents based on the following ideas:

- The frequency space is a good space points of view robustness and imperceptibility, hence the choice of the modified discrete cosine transformation (MDCT) to switch to the frequency domain.
- The temporal methods based on the least significant bit (LSB) provide good results in terms of imperceptibility, insertion capacity, and robustness. For these reasons, came the idea of using the concept of LSB not in the time domain but in the frequency domain to take advantage of the latter.
- To have a blind detection and to reduce the error rate, we had the idea to use a substitute method with an error correction code.
- To select the places of insertion, we exploited the properties of the psychoacoustic models 2 of MPEG 1 for audio component, the properties of the human visual system, the Watson model for image component, and a motion detection algorithm to watermarking video.
- Finally, to improve the robustness against attacks, we thought to duplicate the bits of the mark several times.

This chapter is organized as follows: in Section 2, we will detail some related works and the process of insertion and detection for the proposed techniques. Section 3 will present the experimental results and compare the results obtained by the proposed watermarking system with other existing in the literature. In the last section, we give a conclusion for this work.

## 2. The proposed algorithm

### 2.1. Related works

#### 2.1.1. MDCT

According to the literature, watermarking techniques for still images and videos in the frequency domain use DCT. And since the latter is a block transformation, it can introduce block effects causing noticeable distortions. Then, MDCT has emerged as a very effective and dominant tool in the coding of high quality signals because of its particular properties. The MDCT simultaneously performs critical sampling, reduction of block effects, and flexible windows switching [14]. The coefficients obtained after the application of the MDCT are separated into two bands: high frequencies band and low frequencies band. In our work, we will use a modified version of the MDCT.

The direct and inverse MDCT defined for the audio signal are given by:

$$X(k) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x(n) \cos\left(\frac{\pi}{N} \left(n + \frac{1}{2}\right) \left(k + \frac{1}{2}\right)\right), \quad (1)$$

where:

- $x(n)$  is the sample number  $n$ ,
- $k$  is the number of the frequency line ( $k \in [0, N - 1]$ ).

$$y(n) = \sum_{k=0}^{N-1} \sqrt{\frac{2}{N}} X(k) \cos\left(\frac{\pi}{N} \left(n + \frac{1}{2}\right) \left(k + \frac{1}{2}\right)\right), \quad (2)$$

where:

- $n$  is the number of the temporal sample,  $n \in [0, N]$ ,
- $k$  is the number of the frequency line  $k \in [0, N]$ .

For the image, and as we are going to work on blocks of two dimensions, we will use the MDCT for two-dimensional arrays.

The direct and inverse MDCT defined for the image signal are given by:

$$I'(k, l) = \sum_{i=0}^{N1-1} \sum_{j=0}^{N1-1} I(i, j) \cos\left(\frac{\pi}{N1} \left(k + \frac{1}{2}\right) \left(i + \frac{1}{2}\right)\right) \cos\left(\frac{\pi}{N1} \left(l + \frac{1}{2}\right) \left(j + \frac{1}{2}\right)\right), \quad (3)$$

where:

- $N1 \times N1$  size of the image  $I$ ,
- $I(i, j)$  value of the pixel at position  $i, j$  of the image  $I$ .

$$J(i, j) = \sum_{k=0}^{N1-1} \sum_{l=0}^{N1-1} I'(k, l) \cos\left(\frac{\pi}{N1} \left(k + \frac{1}{2}\right) \left(i + \frac{1}{2}\right)\right) \cos\left(\frac{\pi}{N1} \left(l + \frac{1}{2}\right) \left(j + \frac{1}{2}\right)\right). \quad (4)$$

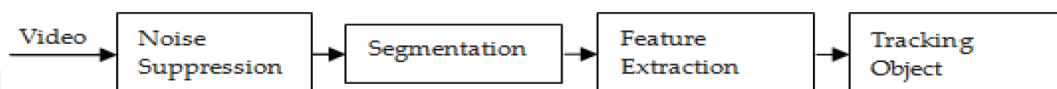
### 2.1.2. Motion detection

To improve the robustness of the video watermarking technique, it is preferable to insert the mark in moving objects [15, 16]. For this reason, we have chosen to use a motion detection algorithm, the one proposed by Peddireddi [17], to identify the objects in motion in the video where we will insert the bits of the mark. The algorithm is composed of four main blocks presented in the following figures (**Figures 2 and 3**).

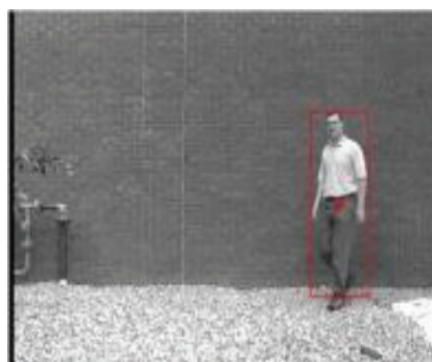
### 2.1.3. JND

JND (or just noticeable difference), also known as just perceptible difference or differential threshold, is the minimum amount by which the intensity of the stimulus must be modified to produce a noticeable variation in a sensory experience [18]. This measure is used in the Watson model which consists of the following steps:

- Change the domain of study by calculating the DCT.
- Definition of the quantization matrix. This model uses the  $Q_m$  quantization matrix of the JPEG standard [19].
- Calculate the frequencies sensitivity coefficients.



**Figure 2.** Blocks of the motion detection algorithm.



**Figure 3.** Detecting the moving object in the video: samplevideo.avi.



- Calculate the sensitivity to the luminance.
- Calculate the contrast masking threshold,  $M$ .
- Finally, calculate the quantization error  $E$  divided by  $M$  to obtain the  $JND$  threshold.

$$JND = \frac{E}{M}. \quad (5)$$

In our work, we will change this model. To achieve the change of the domain study, we will use the MDCT instead of the DCT to exploit its advantages. This choice is also due to the fact that the MDCT has better coding performance than the DCT and also due to the calculation complexity of MDCT which has been reduced in recent years.

#### 2.1.4. Psychoacoustic model

In our work, we will use the psychoacoustic model 2 of the MPEG1 standard. We chose to incorporate this model into our proposed watermarking technique for the audio component of the video, if it exists, in the search for insertion positions. In this model, we do not distinguish between tonal and non-tonal components, but we calculate tonal indices that determine whether the components appear to be tonal or nontonal (noise) [9]. This model is applied on time frames and calculates a masking curve that we will note,  $thr\omega$ .

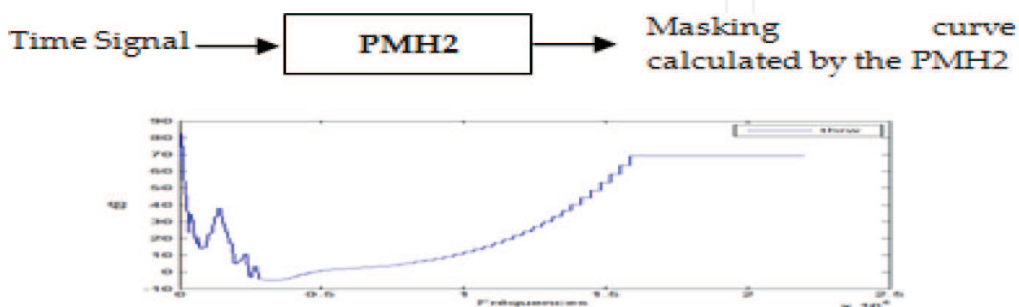
**Figure 4** shows the masking curve  $thr\omega$  for a test signal that has been selected.

## 2.2. Insertion scheme

The diagram we will adopt can be summarized in **Figure 5**.

In this section, we will give the general principle of the process of inserting the brand for the video watermarking technique. For the realization of this technique, we will adopt a proposed watermarking technique for the still image and another proposed technique for the audio. The insertion is performed at moving objects and in non-successive images. This choice is inspired by the fact that:

- Successive images are strongly correlated, and a mark can be detected and deleted easily by a hacker.



**Figure 4.** Psychoacoustic model 2,  $thr\omega$ .

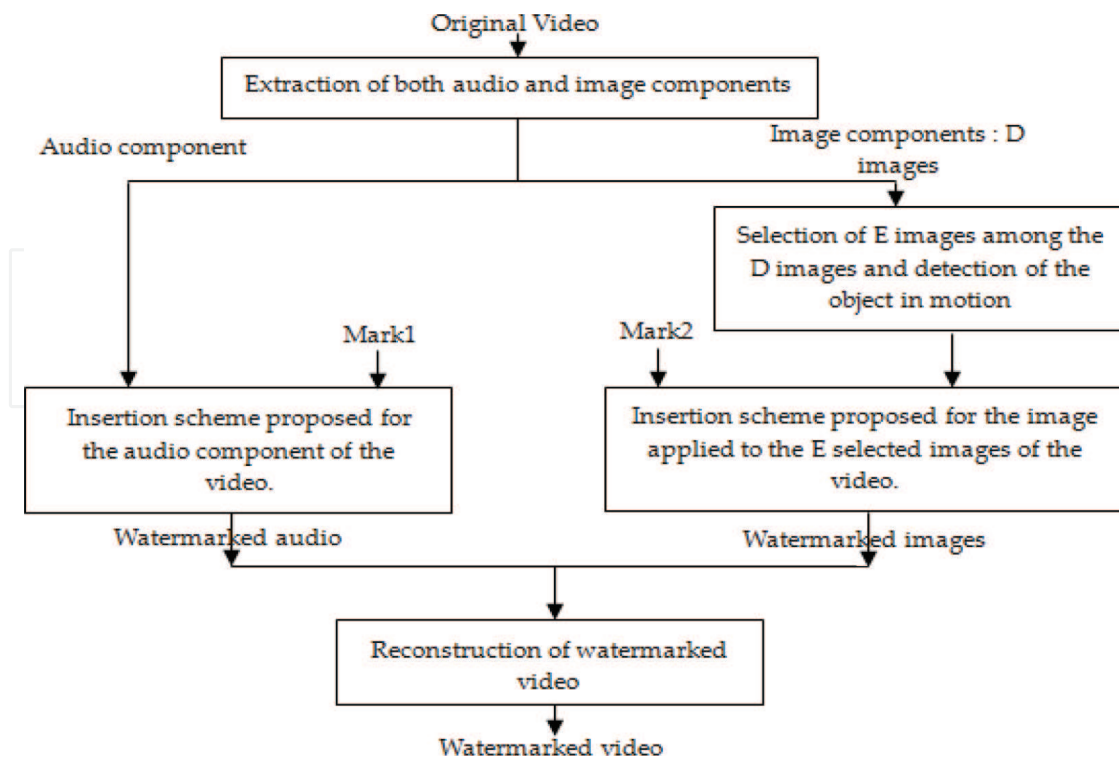


Figure 5. General scheme of insertion of the mark for the video.

- Moving objects are considered a very important factor as, for example, in MPEG4 compression. So, to guarantee a good robustness criterion especially against the compression, we inserted the bits of the mark in the moving objects of the video. We can also improve the invisibility criterion as the mark moves with the objects.
1. The initial input signal is an uncompressed video file. The latter may include or not an audio component.
  2. After reading the original video, we proceed to the separation of the two audio and image components. For this reason, the first step is to check if the video has an audio component or not. If the video does not have an audio component, then we extract only the different images constituting the video.
  3. In this technique, we will insert the mark “Mark1” in the audio component and the mark “Mark2” in the image component. Before proceeding with the insertion of the two marks, we must binarize them. The insertion process of the proposed technique can integrate any type of mark (text, image, and beep sound). The length of the marks is chosen to be multiple of 8. After binarization of the two marks, we obtain two binary vectors of length multiple of 8. This choice will then be useful for performing a Hamming coding (12,8) [20] on each byte of the binary vectors. The use of the Hamming error correction code makes it possible to improve the detection rate of the two marks, as the inserted bits can be modified (inversion from 0 to 1 or from 1 to 0). It will ensure the correction of errors if necessary. Hamming (12,8) is a linear code whose principle is to add 4 control bits to



encode an 8-bit word. At the end, we obtain two coded bit vectors which represent the two coded marks, of length multiple of 12.

4. To obtain a robust watermarking technique against the different manipulations, we will insert the bits of the mark "Mark2" in no-successive images. Hence, the interest of the module allows to select  $E$  images among the  $D$  images of the video. Subsequently, we proceed to the detection of the moving object in these images while using a motion detection algorithm. As an output, this algorithm gives the images of the object in motion.
5. Insertion scheme proposed for the image: we will tattoo the different images of the object in motion detected.
  1. Set the block size to  $(8 \times 8)$  pixels.
  2. Replicate the edges of the image to make its dimensions a multiple of 8.
  3. Decompose the image into blocks of  $8 \times 8$  pixels in the spatial domain.

$$Bloc\_image = \sum_{i=1}^{N1} \sum_{j=1}^{M1} image\_re(i : i + bloksize - 1, j : j + bloksize - 1), \quad (6)$$

where:

$i = 1 \dots N1$  and  $j = 1 \dots M1$  with a step equal to block size = 8.

4. Move to the frequency domain by applying the MDCT, (Eq. (3)). To obtain the frequency coefficients for each block, we must apply the MDCT for each block of  $8 \times 8$  pixels.
5. Separate the frequencies and extract the low frequencies band. We chose to insert the mark bits in the low frequencies band as it is much less sensitive to attacks than the high frequencies band. At the end of this step, we obtain for each block all the low frequencies.
6. Since the human eye is more sensitive to the noise introduced into the low frequency band, we will introduce the Watson model to look for the least perceptible insertion places in the frequencies band. This model calculates the just perceptible difference "JND" for each frequency coefficient of each block.
7. Substitute the insertion of the mark bits: we will look for insertion positions that belong to the band of low frequencies and allow keeping the mark imperceptible (Figure 6).

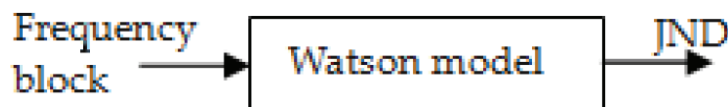


Figure 6. Watson model.

- Select a coefficient of the low frequencies band.
- Binarize the selected coefficient.
- Select the least significant bit (LSB) of the binary representation of the coefficient.
- Substitute the least significant bit by bit stream of watermark to insert.
- Calculate the decimal value of the watermarked coefficient.
- Calculate the difference between the coefficient before the insertion of the mark bit and after the insertion:  $Var\_coef$ .
- Compare this value obtained with that which corresponds to the matrix containing the  $JND$  values generated by the Watson model.
  - If  $Var\_coef < JND$ , so we can insert watermarking bit in this position and we can change the coefficient value without noticing the difference.
  - Else, the insertion in this position will be visible to the eye.

The insertion is performed on all the blocks of the image to improve the robustness. Therefore, we will proceed with the duplication of bits of the brand  $F$  times.  $F$  is calculated according to the number of components where insertion is invisible to the eye, " $NBCom\_INV$ ," and brand size  $Lmark$ :

$$F = \frac{NBCom\_INV}{Lmark}. \quad (7)$$

At the end of this step, we get a watermarked block in the frequency domain.

8. Go back to the space domain by applying the IMDCT (Eq. (4)) to reconstruct the watermarked image.

All previous steps are applied to all blocks in the image and for all selected images in the video.

6. Insertion scheme proposed for the audio: we will integrate this model in the insertion process to exploit its properties in the search for insertion positions. Similarly, and as for the image, this technique operates in the frequency domain using the MDCT (Eq. (1)). The various steps constituting the insertion process are:
  1. Decompose the original audio signal into blocks of 1024 samples each (23 ms duration).
  2. Integrate the psychoacoustic model 2 on each time frame of 1024 samples obtained from the previous step. This model will generate a masking curve  $thr\omega$ .
  3. In parallel with the previous step, apply the MDCT (Eq. (1)) transformation on blocks of 1024 samples to pass to the frequency domain. We obtain blocks of 1024 frequency coefficients in the frequency domain.

4. Extraction of low frequencies: the coefficients obtained are separated at low frequencies and high frequencies. We take each block of frequencies components and set the low frequencies band to half, at the occurrence of  $N/2$  ( $N = 1024$ ).
5. Substitute insertion: we will inject the watermarking bits into the frequency components of the low frequency band under the masking curve *thrw* (Figure 7).

We will look for the insertion positions  $P_0$  belonging to the low frequency band and lying under the curve. After the binarization and the hamming coding of the Mark1, we will obtain a binary sequence  $bi \{0, 1\}$  of length  $L_{mark1}$ . In order to improve the robustness criterion of the proposed technique, we duplicated each bit of the sequence  $bi$ ,  $F1$  times.  $F1$  is calculated as the integer part of the ratio between the number of components at positions  $P_0$ ,  $NB\_TH$  and the length of the mark  $L_{mark1}$ .

$$F1 = \text{Integerpart}\left(\frac{NB\_TH}{L_{mark1}}\right). \quad (8)$$

We will have a binary sequence  $b'i \{0, 1\}$  of length  $L'_{mark1}$ .

$$L'_{mark1} = L_{mark1} * F1 \quad (9)$$

After the search for the different frequency components located at the  $P_0$  positions, we proceed to the binarization of the values of these components. Next, we substitute the least significant bit (LSB) of each component with the current bit of the watermarked message. At the end, we get watermarked block in the frequency domain.

6. Go back to the time domain by applying the IMDCT (Eq. (2)) to reconstruct the watermarked audio. All previous steps are applied to all blocks in the audio.
7. After getting the watermarked audio signal and different watermarked images, we join these two components (audio and image) to form the final watermarked video signal

### 2.3. Detection scheme

The detection is blind (we do not have the original document; only the secret key is needed to extract the mark) and the reverse of the insertion. For the detection of the two marks Mark1 and Mark2 inserted, we will need as keys “Key1,” “Key2”:

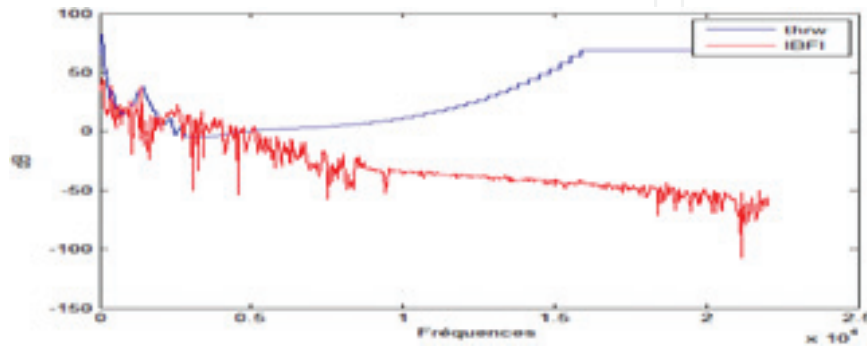


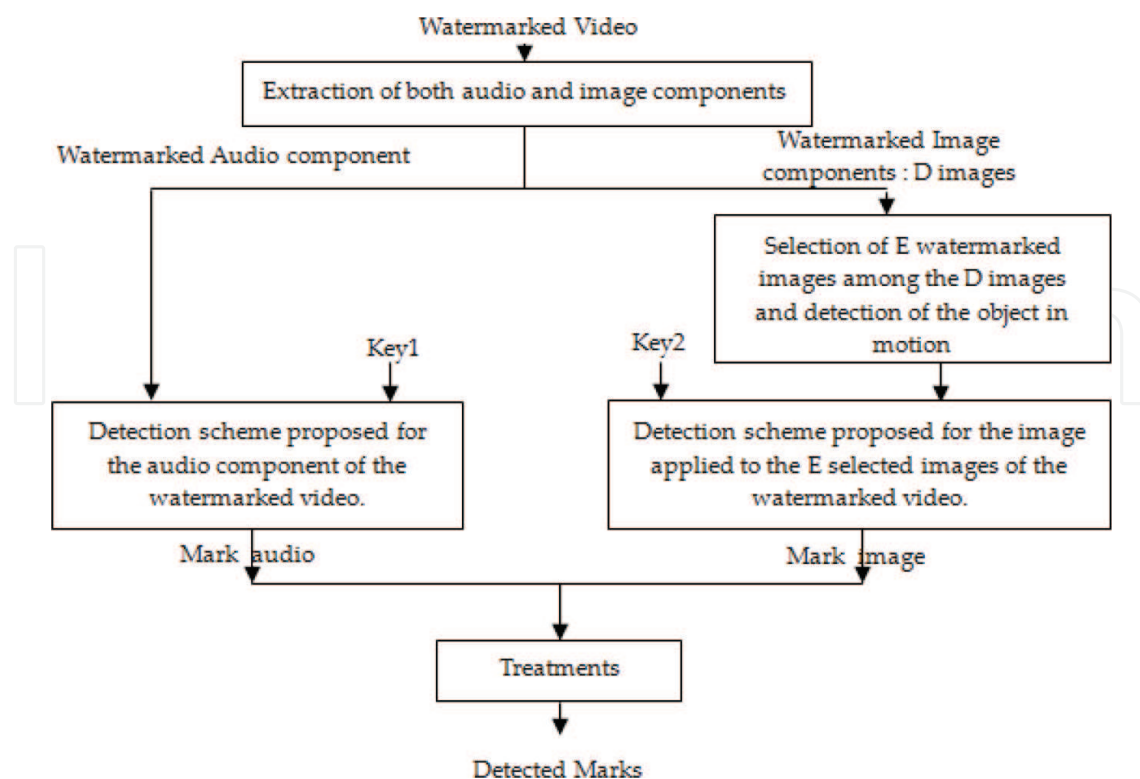
Figure 7. Curve in red “low frequencies” and curve in blue *thrw* for a chosen test signal.

**Figure 7** shows the masking curve  $thr_w$  in blue and the curve of low frequencies samples in red for a signal that has been chosen.

- Duplication numbers  $F$  and  $F1$  that we can insert a bit.
- List of the positions of the components under the masking curve that are sought by the psychoacoustic model 2 in the insertion phase.
- Positions of the components sought by the Watson model in the insertion phase.

The entry of the detection process is the watermarked video resulting from the insertion process. After separating the two audio components, if it exists, and image and using the two keys ( $Key1$  and  $Key2$ ), we extract the two marks inserted into each component.

1. Detection scheme proposed for the image: we begin by replicating the edges of the watermarked image, breaking down into blocks size  $8 \times 8$  pixels in the spatial domain, and applying the MDCT to switch to the frequency domain. Our detection scheme is blind. For this reason, we only use the duplication numbers  $F$  and the positions of the invisible components generated using the Watson model in the insertion phase.
1. From these, we can detect the bits of the message inserted in the components corresponding to these positions. We will then have as a result a binary vector containing the watermark bits corresponding to the coded signature but with duplication  $F$  times for each bit. Finally, to detect the bits of the mark without duplication, we use the parameter  $F$  to eliminate the duplication. We will have as a result the extracted encoded binary brand, of size multiple of 12.



**Figure 8.** General scheme of detection of the mark for the video.

2. Hamming decoding to finally find useful binary brand, corrected multiple of 8.
  3. Reconstruction of the final mark (**Figure 8**).
1. Detection scheme proposed for the audio: after decomposing the watermarked audio signal into blocks of 1024 samples and applying the MDCT on each block to pass to the frequency domain, we proceed to the detection of the bits of the mark.
    - From the positions of the watermarked components under the masking curve, sought by the psychoacoustic model 2 in the insertion phase, we determine the values of these components. Subsequently, we proceed, as we did in the insertion process, to the binarization of these values. Then, we extract from the least significant bit of the inserted message. We obtain then a binary sequence with duplication of length  $L'_{mark1}$ . Finally, to detect the bits of the mark without duplication, we use the parameter  $F1$  to eliminate the duplication. We will have as a result the extracted encoded binary brand, of size multiple of 12.
    - Hamming decoding to finally find useful binary brand, corrected multiple of 8.
    - Reconstruction of the final mark.

### 3. Experimental results and comparative analysis

In this section, we will present, in detail, all the experimental results obtained. The algorithm is tested on MATLAB R2013a with an Intel (R) core (TM) i7-6500U CPU 2.59 GHz, 8 GB memory computer. The experimental corpus is formed by six videos of .avi format (**Table 1**).

#### 3.1. Performance evaluation indexes

##### 3.1.1. PSNR


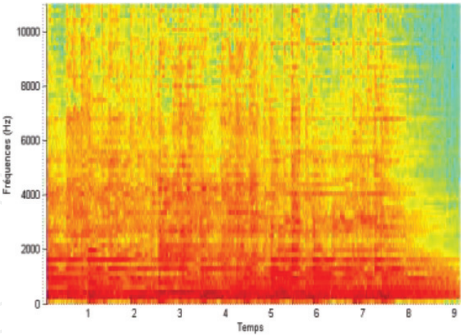


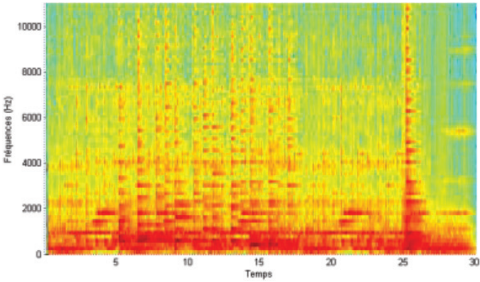

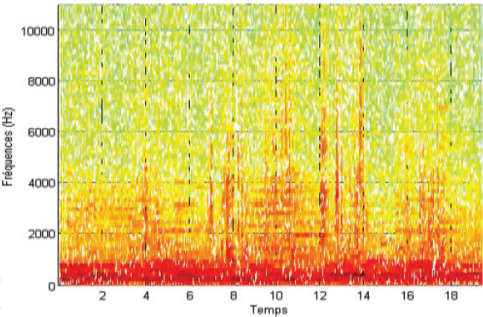

Peak signal-to-noise ratio (PSNR) is an objective quality evaluation measure whose unit is (dB). It measures the quality of the altered (watermarked) image compared to the original image. In particular, we used the PSNR to evaluate the invisibility of our watermarking system. PSNR is defined as:

$$PSNR_{seq\_video} = 10\log_{10} \left( \frac{255^2}{\frac{1}{RN1M1} \sum_{r=1}^R \sum_{n1=1}^{N1} \sum_{m1=1}^{M1} (I_{r,n1,m1} - I'_{r,n1,m1})^2} \right), \quad (10)$$

where:

- $I_{r,i,j}$  and  $I'_{r,i,j}$ : values of pixels (i, j) in the  $r^{th}$  image of the original and watermarked video.
- $(M1 \times N1)$ : size of the video image.
- $R$  : total number of video frames.



Video (.avi)	Image component	Spectrogram audio component
Windows1		
WildLife11		No audio component
horses		
TV		
Sample		No audio component



Video (.avi)	Image component	Spectrogram audio component
Foreman		No audio component

**Table 1.** Tested videos.

### 3.1.2. SNR

Signal-to-noise ratio (SNR) is a measure that will allow us to calculate the similarity between watermarked audio and original audio. It is usually expressed in decibels (dB). SNR is defined as:

$$SNR_{dB} = 10 \log_{10} \left[ \frac{\left( \sum_n x(n)^2 \right)}{\sum_n [x(n) - x'(n)]^2} \right], \quad (11)$$

where:

- $x(n)$ : sample number  $n$  of the original signal.
- $x'(n)$ : sample number  $n$  of the watermarked signal.

### 3.1.3. Objective difference grade

Objective difference grade (ODG) is a score calculated by the PEAQ algorithm [21]. This algorithm compares the original signal and the watermarked signal and assigns a comparative score between 0 and  $-4$ . If  $ODG = 0$ , there is no degradation. If we get a GDO rating that varies between  $-0.1$  and  $-1$ , the deterioration is noticeable but not annoying. For an ODG rating that ranges between  $-1.1$  and  $-2$ , the degradation is slightly annoying. If the ODG value obtained varies between  $-2.1$  and  $-3$ , the degradation is annoying. Finally, if the ODG score obtained is in the range  $[-3, 1; -4]$  so the distortion is very boring.

### 3.1.4. Universal quality index

The universal quality index (UQI) is proposed by [22]. It is an objective evaluation of the visual quality of images and whose range of values varies between  $[0, 1]$ . Higher UQI values represent a better criterion of imperceptibility. The UQI is defined by:

$$UQI = \frac{4\sigma_{\bar{I}\bar{I}'}}{(\sigma_I^2 + \sigma_{I'}^2) \left[ (\bar{I})^2 + (\bar{I}')^2 \right]}, \quad (12)$$

where:

- $I$  and  $I'$  are, respectively, the average values of the original image  $I$  and the processed image  $I'$ .
- $\sigma^2_I$  and  $\sigma^2_{I'}$  are, respectively, the variances of  $I$  and  $I'$ .
- $\sigma_{II'}$  is the covariance of  $I$  and  $I'$ .

### 3.1.5. NC

To test the robustness of the technique against attacks, we will calculate the correlation  $NC$  between the original brand inserted and the mark detected after the exposure of watermarked files to different attacks. For the image, the formula for normalized intercorrelation is given by:

$$NC = \frac{\sum_{i,j=1}^{L_{mark2}} bin(i,j) * bin'(i,j)}{\sqrt{\sum_{i,j=1}^{L_{mark2}} bin'(i,j)^2 * \sum_{i,j=1}^{L_{mark2}} bin(i,j)^2}}, \quad (13)$$

where:

- $bin$  is the binary vector of the inserted mark.
- $bin'$  is the binary vector of the mark detected after application of the attacks.
- $L_{mark2}$  is the length of the inserted mark.

For audio, the formula for normalized intercorrelation is given by:

$$NC = \frac{\sum_{j=1}^{L_{mark1}} b_i(j) * bin(j)}{\sqrt{\sum_{j=1}^{L_{mark1}} bin(j)^2 * \sum_{j=1}^{L_{mark1}} b_i(j)^2}}, \quad (14)$$

where:

- $b_i$  is the binary vector of the inserted mark.
- $bin$  is the binary vector of the mark detected after application of the attacks.
- $L_{mark1}$  is the length of the inserted mark.

## 3.2. Marks

- **Mark1:** in the audio component of the video, we will insert the text mark “audiowatermarking,” of length 136 bits and after the hamming coding, its length reaches 204 bits (after that, each bit will be duplicated  $F1$  times).
- **Mark2:** in the image component of the video, we will insert the image “logo.bmp,” of size  $32 \times 32$  pixels and after the hamming coding, its length reaches 1536 bits (after that, each bit will be duplicated  $F$  times) (**Figure 9**).



Figure 9. logo.bmp binarised.

Video	PSNR_video	UQI	SNR_audio	ODG
Windows1	62,01	1	63,35	0
WildLife11	59,62	0,99	—	—
Horses	62,28	1	68,52	0
TV	60,78	0,99	62,94	−0.1
Sample	60,4	0,99	—	—
Foreman	53,9	0,99	—	—

Table 2. Experimental results.

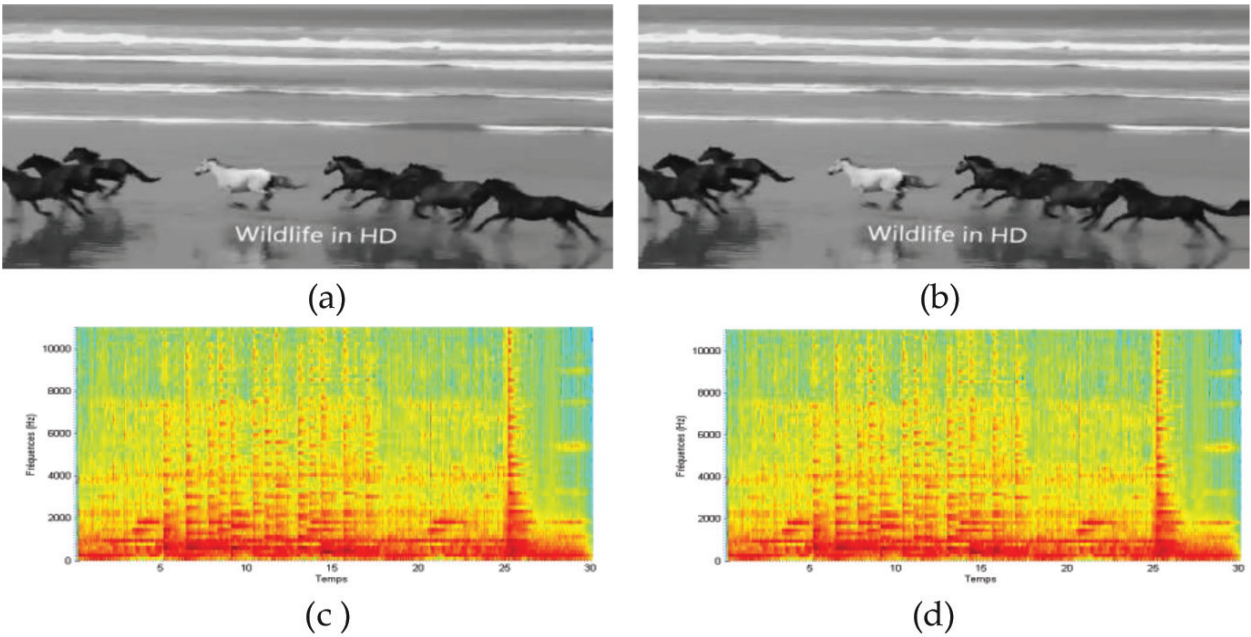


Figure 10. Video test: “horses.avi”; (a) frame (i) original; (b) frame (i) watermarked; (c) spectrogram of the original audio component; and (d) spectrogram of the watermarked audio component.

3.3. Imperceptibility

Table 2 gives PSNR, UQI, SNR, and ODG values for the imperceptibility tests.

By analyzing and comparing the original image (a) with its watermarked equivalent (b) of the video horses.avi, we notice that they do not present remarkable differences and they are even

identical. So the proposed watermarking technique does not affect the quality of images and the inserted brand remains invisible to the human eye. We also note that the spectrogram of figure (c) faithful to that of figure (d). This shows the imperceptible criterion of the technique (Figure 10).

From the results in Table 2, we note that the PSNR values for video sequences vary between 53.90 and 62.28 dB. These values show that the proposed technique gives very adequate results point of view of imperceptibility of the mark despite the high insertion capacity. These values are highlighted by the values of UQI which vary between 1 and  $0.99 \approx 1$ . Similarly for the audio component, we note that the SNR values vary between 62.94 and 68.52 dB, which is very interesting. ODG values further enhance the imperceptible criterion of the technique and they vary between 0 and  $-0.1$ . All these values prove the good criterion of imperceptibility guaranteed by the proposed watermarking technique.

### 3.4. Robustness

To evaluate the robustness, we will apply different types of attacks on the audio and video component of the video: MP3 compression/decompression with the MPEG1 coder "lame.exe" at different compression rates: 128, 96, and 64 kbit/s, the attack by impulsive and Gaussian noise, cropping, frame swapping, frame dropping, frame averaging, and change the coding rate. We will calculate the NC values between the mark before and after the attacks for both components (Table 3).

According to the results, we note that the NC values for watermarking system vary between 1 and 0.85 that is very interesting. For values of  $NC = 1$ , it means that the mark detected after the attacks is perfectly identical to the initial mark. We also notice that the watermarking system is robust against MPEG1 and MPEG2 compression.

### 3.5. Comparative analysis

According to the study of the existing, the watermarking techniques for the video watermark only the image component. It is among the contributions of our watermarking system.

In Table 3, the notation "—" indicates that data are not available.

Video.avi	128 Kbps	96 Kbps	64 Kbps	MPEG-2	Cropping	Dropping	Swapping	Impulse noise	Gaussian
	NC: audio component			NC: image component					
Windows1	1	1	0.9	0.92	0.97	0.94	0.98	0.93	0.92
WildLife11	—	—	—	0.92	0.91	0.96	0.95	0.85	0.88
Horses	1	0.99	0.95	0.93	0.95	0.90	0.97	0.92	0.91
TV	1	1	0.92	0.85	0.98	0.96	0.96	0.91	0.89
Sample	—	—	—	0.89	0.95	0.90	0.97	0.87	0.9
Foreman	—	—	—	0.9	0.97	0.95	0.97	0.86	0.87

Table 3. NC values.

Techniques	PSNR	MPEG-2	Cropping	Dropping	Swapping	Impulse noise	Gaussian
		NC					
Li and Wang [23]	39.08	0.98	1	—	0.5	—	0.98
Dolley and Manisha [24]	—	0.92	—	0.93	1	0.90	0.41
Supriya and Navin [13]	42.12	—	1	0.98	0.98	0.98	0.98
Shaveta and Daljit [11]	44.88	—	0.91	—	—	0.62	—
Himanshu et al. [25]	—	0.86	0.92	—	0.85	0.69	—
Chitrasen and Tanuja [26]	40.04	—	—	0.59	—	0.50	0.50
Proposed method	60	0.90	0.95	0.93	0.97	0.89	0.89

Table 4. Comparative analysis.

On after PSNR values shown in **Table 4**, we note that the proposed watermarking system guarantees the best criteria of imperceptibility PSNR = 60 dB. In addition, the proposed technique shows good performance against attacks. The NC values vary between 0.89 and 0.97. Comparing the results obtained by the proposed watermarking system with those obtained by Dolley and Manisha in [24], we note that the proposed technique is more robust against the Gaussian attack and for the other attacks, the results are close but from the description of this technique, we found that the detection scheme requires the presence of the original video, while our proposed method requires only the keys which makes the detection faster. In addition, we note that the results obtained by the proposed method are better than those obtained by Chitrasen and Tanuja in [26] which shows the contribution of our watermarking system.

### 4. Conclusion

In this chapter, we proposed a watermarking system for multimedia documents operating in the frequency domain using MDCT. This watermarking system watermarks both image and audio components, if it exists. For the image component, we injected the message bits into the detected moving object using a motion detection algorithm and in the LSB of the components searched using the Watson model. If the video has an audio component, we have injected the mark in the LSB of the components under the masking curve sought by the psychoacoustic model 2 of the MPEG 1 standard. The imperceptibility of the watermarking system is tested by calculating four measures: PSNR, UQI, ODG, and the SNR. For the image component, we obtained a total PSNR value equal to 60 dB, and an average UQI index equal to  $0.99 \cong 1$ . For the audio component, we obtained a total SNR value equal to 64.93 dB, and an average ODG value equal to  $-0.03 \cong 0$ . These values show that the watermarking system ensures a good criterion of imperceptibility. The robustness of the watermarking system is tested by applying several attacks known in the literature as: Gaussian noise, impulse noise, compression MPEG II, MP3 compression,... The results show the interest of the technique point of view robustness. We obtained NC values between 0.85 and 1.



As a conclusion, we can draw from these obtained results that:

- The frequency domain, in particularly the MDCT, has shown its proof point of view imperceptibility and robustness. There is still a very interesting area.
- The integration of the psychoacoustic model 2 of the MPEG I standard, the use of the Watson model and the motion detection algorithm, the insertion in the LSB, and the Hamming coding improves the performance of the proposed watermarking system.
- We have been able to increase the insertion capacity while maintaining a good criterion of imperceptibility and robustness.

## Author details

Maha Bellaaj\* and Kaïs Ouni

\*Address all correspondence to: [maha\\_bellaaj@yahoo.fr](mailto:maha_bellaaj@yahoo.fr)

Research Laboratory Smart Electricity & ICT, SEICT, LR18ES44, National Engineering School of Carthage, ENICarthage, University of Carthage, Tunis, Tunisia

## References

- [1] Gwenaël D, Jean-Luc D. Problématique de la Collusion en Tatouage Vidéo. 2005;22(6): 563-574
- [2] Hembrooke EF. Identification of sound and like signals. United States Patent: 3,004,104; 1961
- [3] Cox I, Miller LM. The first 50 years of electronic watermarking. EURASIP. Journal on Advances in Signal Processing. 2002:126-132. DOI: <https://doi.org/10.1155/S1110865702000525>
- [4] Cléo B. Tatouage informé de signaux audio numériques [thesis]. National School of Telecommunications; 2005
- [5] Fadwa D. Tatouage d'images Par Techniques Multidirectionnelles et Multirésolution. [DEA Memory]. National Institute of Applied Sciences of Lyon; 2003
- [6] Martin T, Burg J. Digital Representation, Comparison of DCT and DFT. Science of Digital Media; 2007. Supplement to Chapter 4
- [7] Marcellin MW, Gormish MJ, Bilgin A, Boliek MP. An overview of JPEG-2000. In: Proceedings of IEEE Data Compression Conference; 2000. pp. 523-541
- [8] Sikora T. MPEG digital video coding standards. In: Jurgens R, editor. Digital Consumer Electronics Handbook. New York: McGraw-Hill Book Company; 1997. Chapter 9



- [9] Norme Internationale, ISO/CEI 11172-3. Technologies de l'information codage de l'image animée et du son associé pour les supports de stockage numérique jusqu'à environ 1,5 Mbit/s, partie 3: Audio
- [10] Kundur D, Hatzinakos D. Digital watermarking using multiresolution wavelet decomposition. In: Proceedings of IEEE ICASSP '98, Vol. 5. Seattle, WA, USA; 1998. pp. 2969-2972
- [11] Shaveta, Daljit K. Scaled wavelet transform video watermarking method using hybrid technique: SWT-SVD-DCT. International Journal of Advanced Technology in Engineering and Science. 2015;3(01):78-86. ISSN: 2348-7550
- [12] Shital D, Nitin D, Suresh R. Tampering detection and localization in video using fragile watermarking. International Journal of Innovative Research in Computer and Communication Engineering. 2016;4(7):13106-13113
- [13] Supriya A, Navin S. Digital video watermarking using Dwt and Pca. IOSR Journal of Engineering. 2013;3(11):45-49. e-ISSN: 2250-3021, p-ISSN: 2278-8719
- [14] Mu-Huo H, Yu-Hsin C. Fast IMDCT and MDCT algorithms—A matrix approach. In: Proceedings of IEEE Transactions on Signal Processing; 2003. pp. 221-229
- [15] Delp EJ. Scene Adaptive Video Watermarking. Purdue University School of Electrical and Computer Engineering Purdue Multimedia Testbed. Video and Image Processing Laboratory; 2000
- [16] Zlomek M. Video watermarking [thesis]. Charles University in Prague Faculty of Mathematics and Physics; 2007
- [17] Peddireddi L. Object Tracking and Velocity Determination Using TMS320C6416T DSK. Klagenfurt: Institute of Networked and Embedded Systems Pervasive Computing; 2008
- [18] Weber's Law of Just Noticeable Differences [Internet]. Available from: <http://www.usd.edu/psyc301/WebersLaw.htm> [Accessed: August 15, 2016]
- [19] Marcellin W, Michael J, Gormish A, Bilgin A, Boliek MP. An overview of JPEG-2000. In: Proceedings of IEEE Data Compression Conference; 2000. pp. 523-541
- [20] Hamming RW. Error detecting and error correcting codes. Bell System Technical Journal. 1950;26(2):147-160
- [21] Union Internationale des Télécommunications (UIT). Recommandation B.S. 1387: Méthode de mesure objective de la qualité du son perçu; 2001
- [22] Wang Z, Lu L, Bovik C. Video quality assessment based on structural distortion measurement. Signal Processing: Image Communication. 2004;19:121-132
- [23] Li X, Wang R. A video watermarking scheme based on 3D-DWT and neural network. In: Proceedings of Ninth IEEE International Symposium on Multimedia; 2007. pp. 110-115

- [24] Dolley S, Manisha S. A new approach for scene-based digital video watermarking using discrete wavelet transforms. *International Journal of Advanced and Applied Sciences*. 2018:148-160
- [25] Himanshu A, Rakesh A, Bedi S. Highly robust and imperceptible luminance-based hybrid digital video watermarking scheme for ownership protection. *International Journal of Image, Graphics and Signal Processing*. 2012;(11):47-52. DOI: 10.5815/ijigsp.2012.11.07
- [26] Chitrasen TK. Robust video watermarking using discrete cosine transform and third level discrete wavelet transform. *International Journal of Engineering Research and Applications*. 2017;7(10.Part-1):87-92. ISSN: 2248-9622

IntechOpen

