Journal of Information Hiding and Multimedia Signal Processing ©2016 ISSN 2073-4212
Ubiquitous International Volume **7**, Number **6**, November **2016**

A Blind Video Watermarking Algorithm for Copyright Protection based on Dual Tree Complex Wavelet Transform

P. Senatore

Department of Information, Electrical and Telecommunication Engineering
Sapienza University of Rome
via Eudossiana 18, 00184 Roma, Italy
pietro.senatore@gmail.com

A. Piva

Department of Information Engineering
University of Florence
via S.Marta 3, 50139 Firenze, Italy
alessandro.piva@unifi.it

F. Garzia, R. Cusani

Department of Information, Electrical and Telecommunication Engineering
Sapienza University of Rome
via Eudossiana 18, 00184 Roma, Italy
fabio.garzia@uniroma1.it, roberto.cusani@uniroma1.it

Received March, 2016; revised June, 2016

Abstract. *DVDs and Blu-rays are among the most frequent victims of video content counterfeiting. Primarily, illegal distribution of movies on Internet is a growing menace to film industry. For this reason, authentication techniques are required to ensure the integrity and security of a video: the digital video watermarking can play a leading role in this field. In this paper, a blind digital video watermarking algorithm for copyright protection is proposed, in which the watermark is embedded into the third level of the Dual Tree Complex Wavelet Transform of both chrominance channels, to provide a robust watermarking algorithm. With the purpose of copyright protection, the information inserted in the video will be composed of two contributions: a digital watermark and a binary signature, generated from a desired name. This second contribution will ensure the copyright protection, in order to avoid the illicit distribution of the content. In addition, the inserted bits will be scattered along all the video's frames, with the aim of preventing their interception and destruction, protecting the binary signature. The specific watermark of a frame is extracted from it without the key used for the watermark generation: this will provide robustness to temporal synchronization attacks, such as the frame rate conversion. This scheme is also robust to downscaling in resolution, noise addition, geometric attacks (such as upscaling and cropping), compression, watermark estimation, temporal frame averaging, multiple watermark embedding.*

Keywords: Video Watermarking, Dual Tree Complex Wavelet Transform, Copyright protection, Binary signature, Robustness, Data hiding

1. Introduction.

In recent years, innovations in electronics and information technology, combined with the sudden growth of powerful techniques of digital and multimedia signal processing, have made the diffusion of movies on Internet easier: from streaming sites an user can freely watch a movie released a few days before, or proceed to download the file itself [1]. The question that arises concerns the copyright protection. For an effective content protection, it could be employed the usual encryption scheme "Alice and Bob", commonly used in many fields related to security and encryption [2]. In particular, Alice, in this case owner of the content, wants to send the information to Bob, potentially unreliable, without at the same time revealing to Bob the detailed analysis of the transmitted content: a similar problem does not show an obvious solution of cryptographic nature. However, there have been efforts at international level to allow a redefinition of the problem as follows: Alice sends the digital data to Bob, but he can only treat the information obtained with a device designed for that purpose. In this way, the protection relies on the inability of Bob to directly access to data [3].

Consequently, new solutions are required to ensure the security of a digital video, underlining the importance that acquires the digital video watermarking. With watermarking, a digital code, i.e. the watermark, is unperceivably embedded into the data in such a way that a given piece of information, is indissolubly tied to the data itself. Later on, such an information can be extracted. To cope with these increasingly critical scenarios regarding the illegal distribution of multimedia contents, the context in which the proposed watermarking scheme is inserted is the copyright protection. The watermarking algorithm will hide then into the video a code informing users about the identity of the rights-holder; such an embedded information will have to be inserted in such a way that if an attacker removes the watermark, he will also destroy the video he was interested to.

For DVDs and Blu-rays, of course, the detection of the watermark should be fast and economic, requiring little additional hardware in the readers/recorders to be able to perform efficiently the extraction. An additional requirement is related to the probability of false alarm, i.e. the circumstance in which a frame is erroneously considered watermarked. Obviously, the probability of false alarm should be as low as possible.

Moreover, with the technological evolution of the cameras, the easiness with which the movies can be captured from cinemas has greatly increased, allowing to pirate the film from a large screen. However, the obtained video quality is not comparable to the original one, being the resolution considerably lower than that of the original content. In order to have an acceptable quality captured film, video upscaling, cropping and other operations on the recorded file are performed: similar operations also have the effect of damaging the inserted watermark, therefore the proposed scheme has to be resistant to such attacks, as shown in the experimental section of the paper.

2. Algorithm's Purpose and Innovations.

The proposed scheme, inspired by the algorithm [4], is based on the Dual Tree Complex Wavelet Transform (DT CWT), developed by Kingsbury [5]. The innovations introduced in this algorithm have the aim of ensuring that the newly developed scheme is able to robustly protect the copyright of a video content, without any perceptible loss in visual quality. In particular, both the chrominance channels are watermarked to increase the robustness, while at the same time guaranteeing almost the same imperceptibility that would have been with one chrominance channel watermarking. A new approach for the computation of the detection threshold, based on the Neyman Pearson criterion has also been designed. Moreover, a binary signature is also introduced which allows to link the digital video to a person or a group of users.

As an additional novelty, besides the PSNR (Peak Signal-to-Noise Ratio) to evaluate the imperceptibility of the watermark, it was used in the Section VI the SSIM (Structural

Similarity Index) index. The SSIM considers the structural distortion of the video as an estimate of the distortion perceived by the user: it is based on the assumption that the human visual system is particularly suitable for extracting the structural information from the visual field. Then, the level of degradation perceived is proportional to the loss of structural information perceived, instead of the errors perceived [6].

Thereby, the presented algorithm reaches the purposes of both embedding a watermark, whose presence will be ascertained in the detection, and linking a digital content, in the specific case a video, to a person or a group of users. In fact, the insertion of a signature is dictated by the need to protect the content from its illicit distribution: if the video is given to unauthorized people, through the extraction of the watermark and then of the inserted signature, the user that has allowed the illicit distribution will be identifiable [7].

The watermark is embedded in each frame of the test videos, using the chrominance channels U and V of the YUV color space. In parallel, the bits composing the binary signature are scattered on a group of chosen frames constituting a video, alternating the use of U or V chrominance channels, as exposed in Section III: this way, their interception by a potential attacker is prevented. Then, in the decoding phase, it is checked if the received bit matches with the complete binary signature.

Since the considered context is the copyright protection, the algorithm is blind: the decoder does not have access to the original content, therefore it must be able to detect the presence of the watermark, and its binary signature, only by the analysis of the watermarked video. Besides, the watermark will be changed after △ consecutive frames, to prevent estimation and temporal frame averaging attacks, as detailed in Section IV.

The remaining part of the paper is organized as follows: Section III describes how the to be embedded information is generated starting from a set of pseudorandom keys, specifying the design choices and how to implement the proposed algorithm; Section IV is dedicated to the experimental analysis with a large spectrum of attacks on test videos, both in standard resolution and HD; finally, Section V draws the conclusions of the paper, suggesting some promising ways to further develop the algorithm and projecting it to the modern technologies of digital videos. In Fig.1 a summary scheme of the main features of the proposed algorithm is exposed.

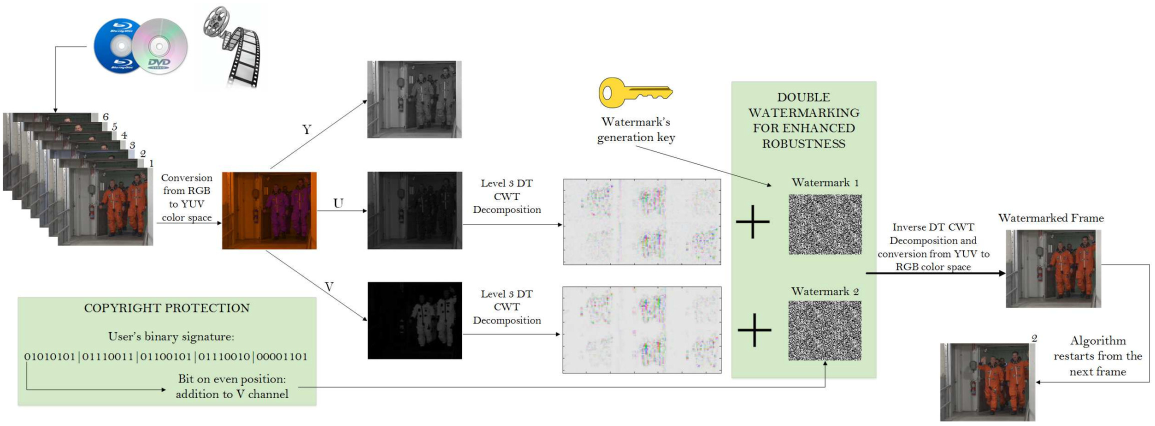

6
5
4
3
2
1
Conversion from RGB to YUV color space
Y
U
V
Level 3 DT CWT Decomposition
Level 3 DT CWT Decomposition
Watermark's generation key
DOUBLE WATERMARKING FOR ENHANCED ROBUSTNESS
Watermark 1
Watermark 2
Inverse DT CWT Decomposition and conversion from YUV to RGB color space
Watermarked Frame
2
Algorithm restarts from the next frame
COPYRIGHT PROTECTION
User's binary signature:
01010101|01110011|01100101|01110010|00001101
Bit on even position: addition to V channel

Figure 1. Brief outline of the proposed algorithm. In the parts highlighted in green the introduced innovations are visible: the double chrominance channel watermarking, to enhance the robustness of the scheme, and the copyright protection of the multimedia contents.

3. **Binary Signature and Watermarking Generation.** The information embedded in the videos is composed of two components:

- A pseudonoise watermark, composed of values -1 and +1, generated starting from a pseudorandom key K_W,
- A binary signature, generated from a character string like e.g. the sequence ”FirstnameLastname”, identifying the copyright owner of the video content.

Let's see in more details how these information are generated. The embedded watermark W is a two-dimensional array composed of +1 or -1: given a set of integer numbers generated from a pseudorandom number generator, a +1 is chosen if the generated number is even, a -1 if it is odd. The generator is driven by a key K_W, defined as:

$$K_W = K_{time} + K_\Delta \tag{1}$$

where the key K_{time} is initialized with the execution time of the algorithm, to ensure the greatest possible pseudorandomness, whereas K_Δ is also a positive integer constant, but it will change its value (without assuming it for more than once) every Δ consecutive frames. The choice of Δ should be made as a compromise between two possible situations: a high value could make the algorithm susceptible to estimation attacks, while for a low value the scheme could be vulnerable to the temporal frame averaging attack [8].

The horizontal and vertical dimensions of the 2-D watermark are chosen to be one-sixteen of the dimensions of the original frame: this factor ensures that the watermark will be, after some steps of DT CWT, the quarter part of the resolution of the frame, as described later.

The watermark just created does not constitute the final one that will be embedded in the U and V channels. In fact, there is the need to ensure an effective defense against attacks based on frame dropping or frame rate manipulation. As suggested in [4], the final watermark is designed in order to allow the decoder to determine whether the examined sequence is watermarked or not without the need of the original key in the decoding phase. For this purpose, a watermark W_Z is generated as:

$$W_Z = \begin{bmatrix} W_1 \cdot Z_1 & W_2 \cdot Z_2 \\ W_3 \cdot Z_3 & W_4 \cdot Z_4 \end{bmatrix} \tag{2}$$

where W is the previously generated basic watermark; Z is a vector of length 4, consisting of values that can assume -1 or +1, depending on the binary signature. In particular, if a bit of the signature is a 1, the first and the fourth value of Z will be set with the same sign (both +1 or both -1). Alternatively, if the bit is a 0, their sign will be set discordant. In this way, in the decoding step, analyzing the cross correlation between the first and the fourth section of the final watermark W_F, it is possible to decipher the inserted bit: if the cross correlation is positive, the transmitted bit is a 1; otherwise, it is a 0.

Finally, the watermark W_Z is subjected to an upscaling by a factor of 2 in both dimensions, obtaining the definitive watermark W_F that will be hidden into the video frames. This upscaling of a factor 2 is required to increase the robustness against geometrical modifications of the watermarked frames. The choice of the factor 2 is a compromise between visual quality and robustness of the algorithm.

The final watermark W_F is not embedded directly, but only after being subjected to level 1 DT CWT decomposition. This is done since the DT CWT is a redundant transform and some components of a pseudorandom sequence, like W_F, in the DT CWT domain may be lost during the inverse transform. The loss of information corresponds to the sequence that lies in the null space of the inverse DT CWT [9]. One way to reduce this information loss is by embedding in the chrominance channels the DT CWT coefficients of the watermark, instead of the original watermark.

In Fig.2 it is shown a diagram illustrating in detail the steps performed for the final watermark generation, the insertion of the binary signature and the effects of the DT CWT decomposition: the creation of the six high-frequency sub-bands HW1,1, HW1,2,..., HW1,6 and of a low-frequency coefficients band LW. About the notation of Fig.2, the apexes indicate the object of the decomposition, while the subscripts, respectively, the level and the sub-band considered.

After the generation of the watermark, the DT CWT is performed on W_F.

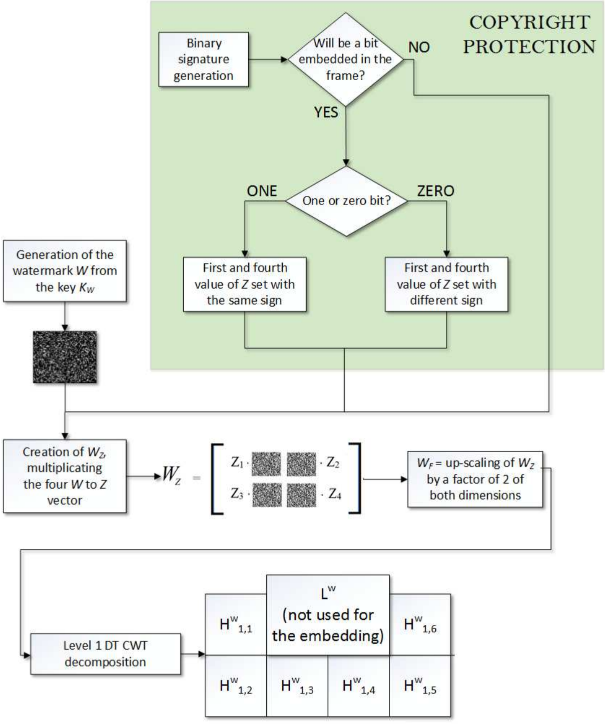

COPYRIGHT PROTECTION
Binary signature generation
Will be a bit embedded in the frame?
NO
YES
ONE
One or zero bit?
ZERO
Generation of the watermark W from the key K_W
First and fourth value of Z set with the same sign
First and fourth value of Z set with different sign
Creation of W_Z, multiplicating the four W to Z vector
W_Z =
Z_1 ·
· Z_2
Z_3 ·
· Z_4
W_F = up-scaling of W_Z by a factor of 2 of both dimensions
Level 1 DT CWT decomposition
$H^W_{1,1}$
L^W (not used for the embedding)
$H^W_{1,6}$
$H^W_{1,2}$
$H^W_{1,3}$
$H^W_{1,4}$
$H^W_{1,5}$

FIGURE 2. Watermark's generation and insertion of the binary signature; as final step the level 1 DT CWT is applied on WF, with the creation of the six high frequency sub-bands and one low frequency component. In the parts highlighted in green it is possible to notice the introduction of the binary signature, which makes the copyright protection possible.

4. **Watermark Embedding.** First of all, the algorithm converts the input video sequence frame from RGB to YUV color space, using the (3). After this, the U and V channels are used to make a double watermarking. The choice of the chrominance channels depends on the fact that the human visual system is less sensitive to changes in chrominance than in luminance (Y channel) [10], [11].

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} Y \\ U \\ V \end{pmatrix} \tag{3}$$

Before proceeding with the watermarking, it is necessary to define a visual mask that ensures a greater imperceptibility to the embedded watermark [12]. This mask is composed of coefficients generated from the six high frequency sub-bands of the level 3 DT CWT decomposition of the U and V channels. The visual masks are generated exactly from those channels in which the watermark is inserted. This choice depends on the consideration that when a watermarked frame is affected by the action of one or more attacks, the extraction of the watermark is performed with greater accuracy if the masks are generated by those same distorted channels: watermark extraction and mask estimation are both performed on the same source (U and V), compensating the distortion of both.

For the visual masks calculation, the magnitudes of the high frequency sub-bands coefficients of the level 3 DT CWT decomposition are considered; these values are subjected to the action of a low pass filter F_{low} and finally divided by a positive integer Ψ, used to control the amplitude of the mask coefficients. Higher values of Ψ lead to smaller mask coefficients, which make the watermark more imperceptible, but at the same time less robust; smaller values of Ψ cause the opposite effect. The final expressions for the coefficients masks of the two channels are:

$$M^U_{3,bw} = \frac{\left|H^U_{3,bw}\right| * F_{low}}{\Psi} \quad M^V_{3,bw} = \frac{\left|H^V_{3,bw}\right| * F_{low}}{\Psi} \tag{4}$$

where $\left|H^U_{3,bw}\right|$ and $\left|H^V_{3,bw}\right|$ are the magnitudes of the high-frequency coefficients of DT CWT decomposition of U and V channels, for $bw = 1, 2, ..., 6$ sub-bands of level 3; F_{low} is the low-pass filter defined as follows [13]:

$$F_{low} = \begin{bmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{bmatrix} \tag{5}$$

Therefore, the coefficients of level 1 DT CWT decomposition of the watermark W_F, indicated with $H^W_{1,bw}$, are added to the high-frequency coefficients of U and V channels, according to the following expressions [14]:

$$\widetilde{H^U_{3,bw}} = H^U_{3,bw} + \rho \times \left(M^U_{3,bw} \cdot H^W_{1,bw}\right) \tag{6}$$

$$\widetilde{H^V_{3,bw}} = H^V_{3,bw} + \rho \times \left(M^V_{3,bw} \cdot H^W_{1,bw}\right) \tag{7}$$

where the symbol $\cdot$ denotes the element-wise matrix multiplication (Hadamard product); ρ is a scalar factor that controls the strength of the watermark: its increase leads to a stronger, but less imperceptible watermark; its decrease produces the opposite circumstance. For these reasons, it is necessary to find an optimal value for ρ, i.e. an acceptable compromise between the two requirements: this will be exposed in Section IV.

After watermarking, the inverse DT CWT of the $\tilde{U}$ and $\tilde{V}$ channels, and a conversion from YUV to RGB color space using the (8) are performed.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1.13983 \\ 1 & -0.39465 & -0.58060 \\ 1 & 2.03211 & 0 \end{pmatrix} \begin{pmatrix} Y \\ U \\ V \end{pmatrix} \tag{8}$$

In Fig.3 is shown a detailed diagram of the necessary steps for the watermarking, with the part highlighted in green about the double watermarking on U and V, used to enhance the robustness of the scheme.

Note that to increase the robustness to the attacks, this procedure here applied to each frame of the video sequence, can be modified as follows: the scheme will check if the embedded bit has, in the complete signature, even or odd position: if it is even, the DT CWT watermark coefficients with the embedded bit will be inserted only in the V chrominance channel of the frame, otherwise only in the U chrominance channel.

In this way, a potential attacker will have only one channel for the bit estimation, furthermore ignoring the specific frame in which the bit is embedded, since only encoder and decoder will know in advance the bits position scattered along the frames: with these countermeasures a potential attacker will not decipher the signature inserted.

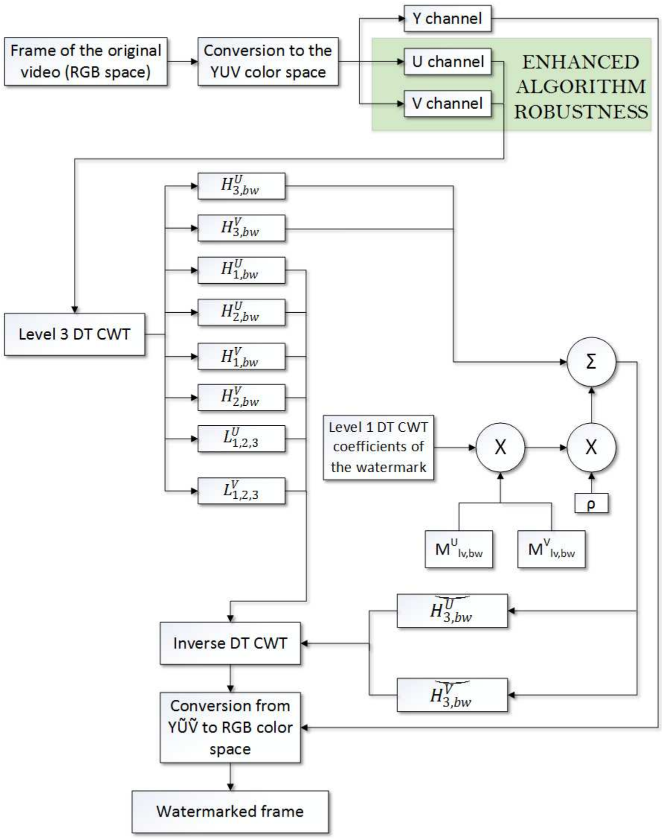


FIGURE 3. Diagram of the encoder, with the embedding of DT CWT coefficients of the watermark in the U and V channel. In the part highlighted in green it is visible the double watermarking of U and V channels, that increases the robustness of the algorithm.

5. **Watermark Extraction.** The decoder performs the inverse operation of the encoder. At first, the frame submitted to the decoder is converted from the RGB to YUV color space. Then, the level 3 DT CWT decomposition is applied on U and V channels, obtaining, respectively, $\widetilde{H^U_{lv,bw}}$ and $\widetilde{H^V_{lv,bw}}$. After choosing one or more levels from which the watermark is extracted, it is performed an estimate of the visual masks coefficients utilized by the encoder, which is different depending on the considered channels and level: $iM^{\tilde{U}}_{lv,bw}$ and $iM^{\tilde{V}}_{lv,bw}$. In general, the inverse coefficients are calculated as follows:

$$iM^{\tilde{C}}_{lv,bw} = \begin{bmatrix} \dfrac{1}{M^{\tilde{C}}_{lv,bw}(0,0)} & \cdots & \dfrac{1}{M^{\tilde{C}}_{lv,bw}(0,D)} \\ \vdots & \ddots & \vdots \\ \dfrac{1}{M^{\tilde{C}}_{lv,bw}(E,0)} & \cdots & \dfrac{1}{M^{\tilde{C}}_{lv,bw}(E,D)} \end{bmatrix} \tag{9}$$

where D and E are, respectively, the horizontal and vertical dimensions of the high frequency sub-bands of the U and V channels, that change depending on the decomposition level; $\tilde{C}$ indicates the watermarked channel from which the mask coefficients are estimated.

Finally, the high-frequency sub-bands coefficients of U and V channels are multiplied by the inverse of the estimated visual masks coefficients, leading to an estimate of the high frequency coefficients of the embedded watermark, indicated with $\widetilde{H^{W_C}_{lv,bw}}$ and defined as follows:

$$\widetilde{H^{W_U}_{lv,bw}} = \widetilde{H^{U}_{lv,bw}} \cdot iM^{\tilde{U}}_{lv,bw} \tag{10}$$

$$\widetilde{H^{W_V}_{lv,bw}} = \widetilde{H^{V}_{lv,bw}} \cdot iM^{\tilde{V}}_{lv,bw} \tag{11}$$

The complete decoding scheme is shown in Fig.4.

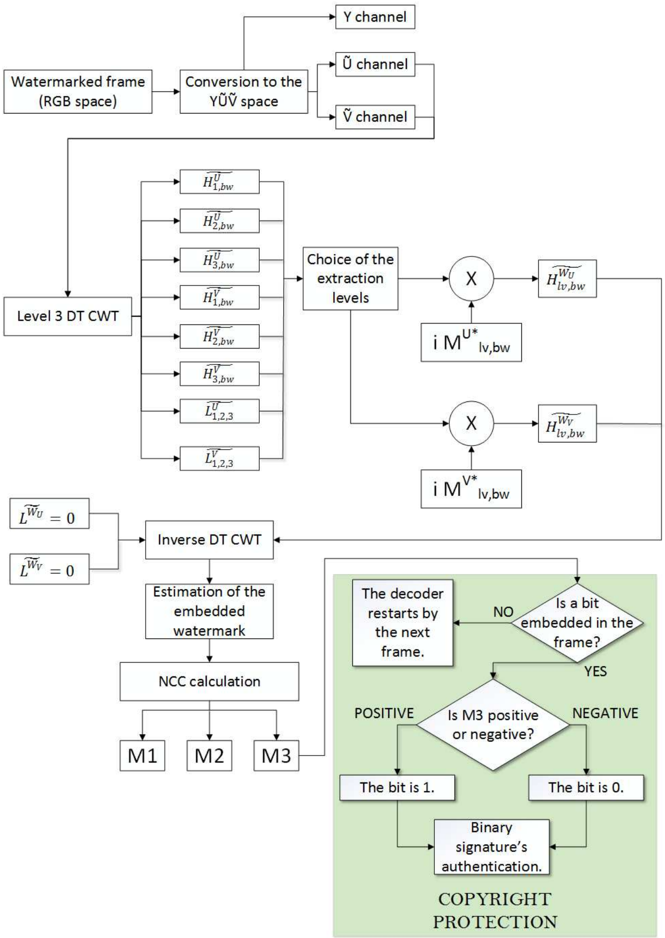

Y channel
Ũ channel
Ṽ channel
Watermarked frame (RGB space)
Conversion to the YŨṼ space
$\widetilde{H^{U}_{1,bw}}$
$\widetilde{H^{U}_{2,bw}}$
$\widetilde{H^{U}_{3,bw}}$
$\widetilde{H^{V}_{1,bw}}$
$\widetilde{H^{V}_{2,bw}}$
$\widetilde{H^{V}_{3,bw}}$
$\widetilde{L^{U}_{1,2,3}}$
$\widetilde{L^{V}_{1,2,3}}$
Level 3 DT CWT
Choice of the extraction levels
X
$\widetilde{H^{W_U}_{lv,bw}}$
i M$^{U*}_{lv,bw}$
X
$\widetilde{H^{W_V}_{lv,bw}}$
i M$^{V*}_{lv,bw}$
$\widetilde{L^{W_U}} = 0$
$\widetilde{L^{W_V}} = 0$
Inverse DT CWT
Estimation of the embedded watermark
NCC calculation
M1
M2
M3
The decoder restarts by the next frame.
NO
Is a bit embedded in the frame?
YES
POSITIVE
Is M3 positive or negative?
NEGATIVE
The bit is 1.
The bit is 0.
Binary signature's authentication.
COPYRIGHT PROTECTION

Figure 4. Diagram of the decoder, with the watermark estimation and the NCC calculation. In the part highlighted in green it is visible the copyright protection with the binary signature's authentication, which decodes the received bits and compares them, at the end of the video, with the correct binary signature.

By inversion of the previously extracted coefficients, an estimate of W_Z: a perfect reconstruction is not possible because, on the encoding phase, only the high-frequency coefficients of level 1 DT CWT of the watermark were considered (as pointed out in Fig.2). In fact, on the decoding phase, the low frequency coefficients LW of the estimated watermark are considered as an array of zeros.

The extracted watermark follows this four-sections structure:

$$\widetilde{W_F} = \begin{bmatrix} \widetilde{W^{1,lv}_{ex}} & \widetilde{W^{2,lv}_{ex}} \\ \widetilde{W^{3,lv}_{ex}} & \widetilde{W^{4,lv}_{ex}} \end{bmatrix} \tag{12}$$

where $lv = 1, 2, 3$ indicates one or more levels chosen for the watermark extraction and the tilde sign indicates the performed estimation. Note that $\widetilde{W_{ex}^{n,lv}} = W_n \cdot Z_n \cdot 2$, for $n = 1, 2, 3, 4$ watermark's sections.

At this point, the decoder, without knowing the original key K, must determine whether the watermark is present or not. For this purpose, it is calculated the normalized cross-correlation (NCC) between the four sections constituting the estimated watermark: for example between W_1 and W_2, indicating the calculation with C_{12}.

In the circumstance in which it's used more than one level to embed the watermark, the NCC is calculated between sections of the watermark extracted from the used levels, for example the first and the third: $C_{12,1}$ and $C_{12,3}$. Then, it is chosen the maximum peak and it is marked as M_1. The identical operation is applied between $W_{1,l}$ and $W_{3,l}$, obtaining M_2, and between $W_{1,l}$ and $W_{4,l}$, obtaining M_3. This process is applied to each frame of the video sequence and finally, after collecting all the values for each frame, the average of all the results obtained is calculated, in order to produce three representative NCC values of the entire video:

$$\overline{M_1} = \frac{1}{\nu}\sum_{n=1}^{\nu} W_{1,l} * W_{2,l} \quad \cdots \quad \overline{M_3} = \frac{1}{\nu}\sum_{n=1}^{\nu} W_{1,l} * W_{4,l} \tag{13}$$

where ν indicates the number of total frames of the entire video and the symbol $*$ indicates the normalized cross correlation.

Therefore, the NCC oscillates between two values: one positive (and as close as possible to +1), if the two sections of the watermark are identical; one negative (and as close as possible to -1), if the two sections of the watermark are one the opposite of the other. In both cases, considering the values of the NCC in modulus, the presence of the watermark is confirmed. If the NCC is close to zero, it will mean that the watermark is absent.

Therefore, the designed algorithm makes each frame independent to the others: if a frame is dropped or the frame rate has been altered, the other frames of the video will not suffer any effect, since they have their own watermark, distinguished from the others. In addition, an attacker could also make the calculation of the NCC within a frame, but this will not allow anyway to discover the exact pattern of the watermark, that remains hidden to potential attackers.

5.1. **Threshold detection.** Whenever a video is submitted to the decoder, there's the need to establish with certainty whether the watermark is present or not. As already claimed, the correlation between a fake watermark and the real one is very low, while between the embedded watermark and the authentic one is very high. Therefore, it's necessary to define a watermark detection threshold that, if exceeded, implies the watermark's presence. By formulate the detection problem as a classical hypothesis testing problem, the Neyman Pearson criterion can be applied, according to which the probability of correctly detecting the watermark is maximized subject to a prescribed limit on the probability of false alarm PFA and then observing the resulting threshold[15]. Assuming a P_{FA} equal to 10^{-6}, under gaussianity assumption the following equation can be derived:

$$T = \sqrt{2}\,\sigma_z \,\mathrm{erfc}^{-1}(2P_{FA}) + \mu_{z|H_0} \tag{14}$$

where T is the detection threshold; erfc^{-1} is the inverse complementary error function; σ_z is the variance of the correlation; $\mu_{z|H_0}$ the mean of the correlation under the assumption that the examined video does not contain the watermark. Secondly, it is possible to invert the equation in order to obtain the consequent value of T:

$$T = \sqrt{2\sigma_I^2/n}\,\mathrm{erfc}^{-1}(2P_{FA}) + \mu_I\overline{w} \tag{15}$$

where σ_I^2 is the variance of the not watermarked image; n the number of samples on which the calculation is performed; μ_I the mean of the image (frame); $\overline{w}$ the sample average of the watermark. This last contribution in our context is zero, since the watermark inserted in the video has zero mean. In Table 1 they are indicated the results of the calculated thresholds for all videos, both standard definition and HD.

Table 1. Calculation results of the watermark threshold T for the considered videos

T of SD videos	Akiyo	Coast	Bus	Football
	0.1574	0.1423	0.1512	0.1488
T of HD videos	Sky	Riverbed	Mobcal	Trees
	0.1289	0.1310	0.1498	0.1452

5.2. **Extraction levels.** The main effect of the downscaling attack is to truncate the high frequency components of a video frame. Since these components are the ones used in the watermarking, it is vital to conceive an effective defense against this technique.

Primarily, it is necessary to understand the downscaling effects, in our case in the circumstance in which it has been adopted the watermarking by level 3 DT CWT decomposition. After this scheme is applied on an image, only the third level should be altered by the presence of the watermark, but what emerges instead is that the watermarking has also direct effects on the lower levels. Besides, the application of downscaling has the effect of shifting the DT CWT coefficients of the watermark towards the lower levels: this effect is proportional to the rate of the applied downscaling.

Therefore, during the downscaling attack, the extraction of the watermark is not performed exclusively on level 3: other levels are taken in consideration, on the decoder side, in order to make a correct estimate of the embedded watermark. A solution could be to consider all possible levels and to perform the extraction steps designed above, but it would be too heavy in terms of computational complexity. Therefore, only certain level(s) for the extraction will be chosen in presence of downscaling of the original video [4]:

- From level 2 and 3 for horizontal resolution higher than 750 pixels;
- From level 1 and 2 for horizontal resolution between 400 and 750 pixels;
- From level 1 for horizontal resolution lower than 400 pixels.

6. Experimental Analyss and Effects of V Channel Watermarking.

6.1. **Examined videos and chosen parameters.** A watermarking system can be considered well designed if after a successful attack, i.e. the watermark has been removed, the commercial value of the watermarked content is destroyed with it [16]. In our context, it means that the watermarked video is too compromised to be correctly watched. To test the robustness of the algorithm to the attacks, they have been considered, as test videos, sequences in uncompressed format (extension .yuv).

The characteristics of the considered videos are the following: four CIF (Common Intermediate Format) videos of 352x288 pixels, therefore in standard definition (SD): ”Akiyo”, ”Coastguard”, ”Bus”, ”Football”; four HD videos: two of 1920x1080 pixels (”Blue Sky” and ”Riverbed”) and two of 1280x720 pixels (”Mobcal” and ”Trees”). All the sequences under examination, illustrated in Fig.5, are in video format YUV 4:2:0, in which it is presented a downsampling by a factor of 2, both vertically and horizontally, of the chrominance channels U and V with respect to the luminance channel Y; besides, the U and V channels are centered vertically halfway between scan lines [17].

The pattern of the watermark has been changed every $\Delta = 20$ consecutive frames; therefore, taking the example of a video with 460 frame, in total the pattern has been changed $460/20 = 23$ times. After some tests evaluating the best compromise between the mask strength and the watermark strength ρ, on the encoding phase the Ψ value was set to 11, while on the decoding phase it was set to 3.

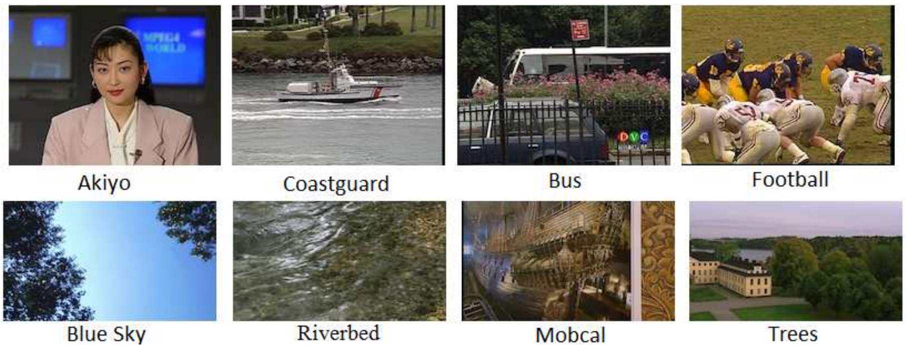


FIGURE 5. Considered test videos: in the top are shown the four videos in Standard Definition (CIF format, 352x288 pixels); in the bottom, there are the HD videos: Blue Sky and Riverbed of 1920x1080 pixels, Mobcal and Trees of 1280x720 pixels.

6.2. **Effects of V channel watermarking.** In this scheme, to increase the robustness of the algorithm, it has been decided to embed the watermark also in the V channel. This improvement makes it more difficult to remove the watermark by an attacker, who then should perform the attack on both chrominance channels, and at the same time destroy those frames in which the bits of the binary signature are embedded.

At this point it must be verified that the addition of V channel watermarking has not compromised the visual quality of the watermarked videos, because, as already described, there is an inverse proportionality relationship between the strength of the watermark and its imperceptibility. Therefore, to evaluate the imperceptibility of the watermark (a particular example is shown in Fig.6), the PSNR and the SSIM were calculated on the frames of the watermarked video.

The choice of these two indices to measure the degradation of the video quality is due to their different nature: the PSNR stands its bases mainly on mathematical computations, therefore sometimes it does not reflect the real video quality perceived by a human being. On the other hand, it is highly sensitive to changes in contrast and brightness. The considered PSNR expression is the following:

$$\mathrm{PNSR} = 10 \log_{10} \left(\frac{\mathrm{peakval}}{\mathrm{MSE}} \right) \tag{16}$$

where peakval in our case is assumed to be 255, while MSE is the Mean Square Error between the original image and the watermarked one.

Instead, the SSIM has the advantage of being more close to the human visual system, but it is almost insensitive to changes in brightness, contrast and hue (and if such changes become substantial, the index can be affected significantly) [18]. Manipulating the value of ρ, it is therefore possible to increase or decrease the strength of the watermark: a stronger watermark increases the NCC values, but it degrades the watermarked video quality. Conversely, with a lowest ρ, the video quality will benefit at the expense of the NCC. Thus, it is necessary to find a compromise value between the two requirements.

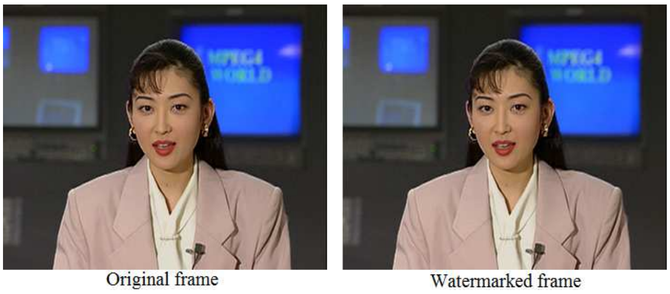


Figure 6. Watermarking of the video Akiyo; in this case the watermarking is difficult due to the presence of the blue screen in the background, because one of the channel used for the watermarking is the chrominance blue channel U. Nevertheless, the watermarked frame on the right doesn't have noticeable artifacts.

Experimentally, as shown in Fig.7, it was observed that values of ρ less than 30 lead to an optimum video quality (PSNR higher than 40 dB and SSIM approximately equal to 0.9), but the NCC falls below 0.5. Therefore, the basic ρ has been set to 30 (the effects of higher values are visible in Fig.7). Examining the PSNR and SSIM values, it is evident that the visual quality is not compromised. Focusing on the results of the "Galeon" (also known as Mobcal) and "Trees" videos, high values of PSNR can be noticed, but SSIM values are lower if compared with the other high definition videos. In this context, the values of PSNR index correctly reflect, better than the SSIM values, the good video quality: in the two videos, in fact, there aren't noticeable artifacts or aberrations.

6.3. **Considered attacks.** Regarding the resistance to the attacks, the following types of attacks have been considered: Geometric attacks (cropping, upscaling); Lossy compression with M-JPEG encoding; Noise addition; Additive attack; Downscaling; Watermark estimation; Temporal frame averaging.

6.3.1. *Cropping.* Cropping the edges or more extended portions of a video sequence is one of the most common attack techniques [19], especially when a watermark is a simple logo inserted at the top or bottom of a multimedia content. In addition, the cropping is a simple task to accomplish, being included in most of common image editing tools. The attacks conducted by cropping were made removing in all videos the 10% of the size of the frames.

For similar levels of cropping, in absence of other attacks, the results have led to NCC values that exceed the detection threshold of the watermark, and to very low BER (Bit Error Rate) values: the results are shown in Table 2.

Table 2. Average correlation peaks after cropping

SD Videos	Akiyo	Coast	Bus	Football
NCC peak	0.6875	0.6959	0.6517	0.6069
Bit Error Rate	2%	1%	3%	1%
HD Videos	Sky	Riverbed	Mobcal	Trees
NCC peak	0.7295	0.7374	0.7126	0.7207
Bit Error Rate	0%	2%	2%	3%

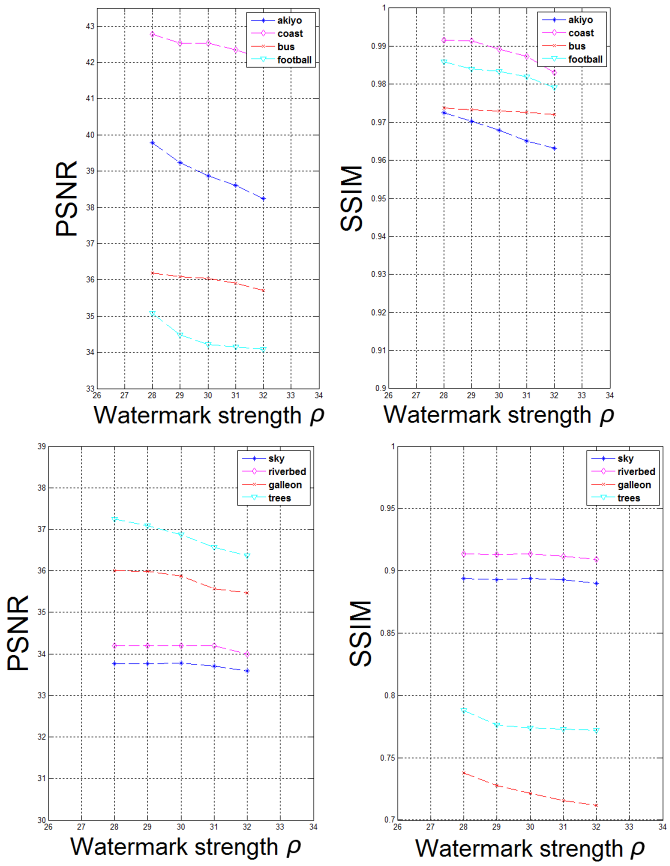
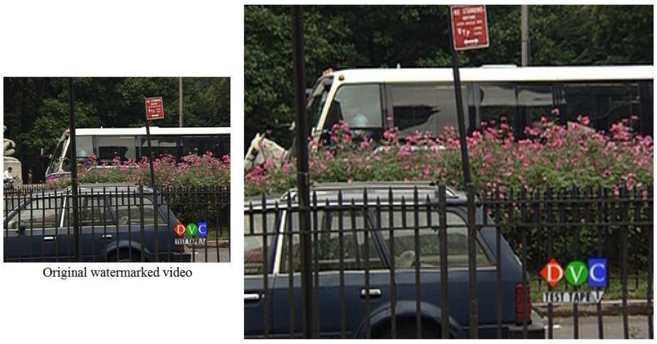


Figure 7. Variation of PSNR and SSIM with the increase of the watermark strength ρ; in the top results with SD videos, in the bottom with HD videos.

6.3.2. *Cropping and upscaling.* A common technique consists in joining the cropping to the upscaling attack. Therefore, the video sequences have been subjected to this combination of attack, in particular the videos are primarily watermarked and then upscaled by a factor of 2, subsequently they are cut in a uniform manner along the edges (removing the 10% of the video size). An example is illustrated in Fig.8. As exposed in Table 3, even with the presence of both attacks, the average correlation peaks and the BER are not subjected to a significant worsening.



Original watermarked video

Upscaled and cropped video

Figure 8. Example of upscaling by a factor of 2 with an uniform cropping of 10% along the edges.

Table 3. Average correlation peaks after upscaling and cropping

SD Videos	Akiyo	Coastguard	Bus	Football
NCC Peak	0.5992	0.6375	0.6303	0.5864
Bit Error Rate	5%	2%	7%	3%
HD Videos	Sky	Riverbed	Mobcal	Trees
NCC Peak	0.7013	0.7159	0.6849	0.6961
Bit Error Rate	3%	5%	8%	8%

6.3.3. *Lossy compression.* For what concerns the lossy compression attacks, it has been considered the Motion JPEG format (also known as M-JPEG or MJPEG). In this compression format each frame of a digital video sequence is separately compressed as a JPEG image. This format is currently subject of interest since it is often used in video surveillance, both by privates and institutions, as in the metro railway [20], [21].

The compression ratio was 4:1, as example it is possible to consider the compression performed on the video ”Bus”: after MJPEG encoding, the original video of 43.5 MB became a file of 10.85 MB. The results have led to low values of BER and low false negative rate, as shown in Table 4. It can be noticed that although the false negative rate is present, even if it's small, the Bit Error Rate is absent or negligible, since in the designed algorithm the information related to the transmitted bit is intimately linked to the presence or absence of peak's sign of the NCC. To considerably compromise the embedded binary signature, it would be necessary to abate the NCC peaks, which means to completely destroy the inserted watermark.

Table 4. Results of the average correlation peaks after the MJPEG lossy compression

SD Videos	Akiyo	Coastguard	Bus	Football
False negative rate	10%	11%	7%	9%
Bit Error Rate	0%	1%	0%	0%
HD Videos	Sky	Riverbed	Mobcal	Trees
False negative rate	8%	7%	9%	11%
Bit Error Rate	1%	0%	0%	1%

6.3.4. *Noise addition.* The noise represents an unwanted modification to the video. In particular, corrupting the digital video with ”Salt & Pepper” noise is a common circumstance, for example for hardware malfunctions, problems in the decoding channel, signal reducing in communications, transmission on a disturbed channel [22]. Salt & Pepper noise alters the pixels values, leading them to assume the minimum (0) or maximum (255) value. In this context, regarding the density of the noise, the choice of the pixels that will vary will be random.

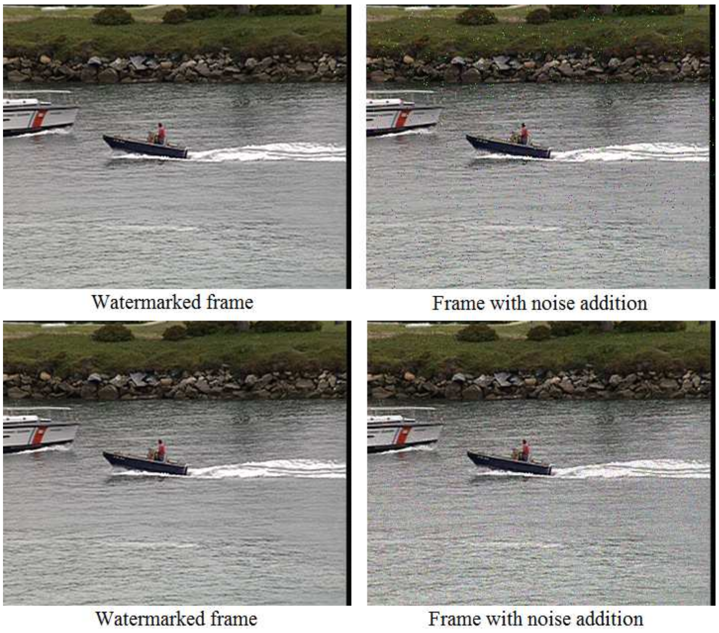
The second considered noise is the Speckle: it is not a source of noise, but a variation of the noise itself. In particular, it is often defined as a multiplicative noise and it is characterized by a granular pattern [23]. An example of the effects of both noises on Coastguard video is shown in Fig.9, while the results of an attack with Salt & Pepper and Speckle noise are reported, respectively, in Table 5 and Table 6. The false negative rate grows when the Speckle noise is present, but for both sources of noise the BER is low: although the noises reduce the correlation peaks, they do not erase the sign of the normalized cross correlation between W_1 and W_4, used for the detection of the correct bit.

Table 5. False negative rate and BER after the salt & pepper noise addition

SD Videos	Akiyo	Coastguard	Bus	Football
False negative rate	7%	7%	8%	9%
Bit Error Rate	2%	3%	3%	0%
HD Videos	Sky	Riverbed	Mobcal	Trees
False negative rate	5%	3%	6%	4%
Bit Error Rate	2%	0%	3%	1%

Table 6. False negative rate and BER after the speckle noise addition

SD Videos	Akiyo	Coastguard	Bus	Football
False negative rate	9%	11%	10%	8%
Bit Error Rate	2%	3%	4%	1%
HD Videos	Sky	Riverbed	Mobcal	Trees
False negative rate	7%	5%	6%	9%
Bit Error Rate	1%	0%	0%	3%



Watermarked frame

Frame with noise addition

Watermarked frame

Frame with noise addition

Figure 9. Attack with noise addition; in the top are shown the effects of salt & pepper noise, in the bottom the effects of speckle noise.

6.3.5. *Additive attack.* Another attack that can be used to remove the correlation peaks is to embed a second watermark to the video [24]. The limit of such an attack is the degradation that typically occurs in the resulting video, since multiple watermarks are present at the same time. In order to test the resistance of the algorithm to this type of technique, a second watermark ρ^* has been applied to the video (already watermarked with the desired W). The generation of this second watermark has been conducted exactly as for the authentic one, in order to objectively verify the effects of a second insertion. As with the original watermark, the pattern of the second watermark was changed every $\Delta = 20$ frames. Therefore, different values of ρ^* have been considered in order to verify, in correspondence, how much the PSNR and SSIM of the resulting video decrease. In

Fig.10 the effects of the attack on standard resolution and HD videos, respectively, are shown.

It can be noticed that, in the first case, as soon as the value of ρ^* exceeds the unit, the PSNR decrease drastically: in particular, it is possible to identify the presence of a second watermark by observing the lower limit of PSNR (for the chosen value of $\rho = 30$) of the SD and HD videos. If the PSNR drops below 34 dB for SD videos and 35 dB for the HD videos, it will be supposed the presence of a second watermark.

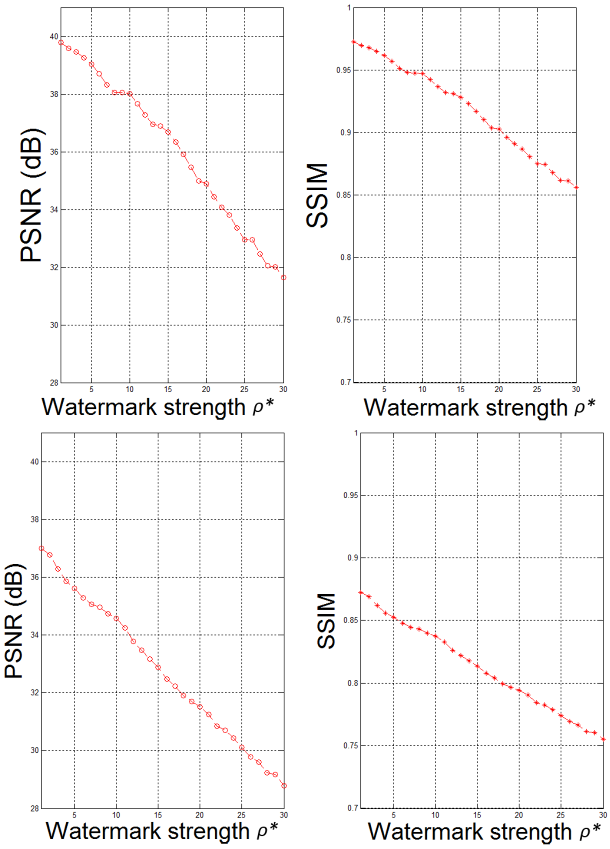


Figure 10. Effects of the embedding of a second watermark on the SD (top) and HD (bottom) videos, based on average values of PSNR and SSIM.

Finally, in Fig.11 it is exposed the average correlation peaks of the considered videos, for increasing values of ρ^*: as it can be noticed, the attack was not successful to cancel the existing correlation. Moreover, for all the tested videos, the BER is equal to 0%.

6.3.6. *Downscaling.* Due to the different resolutions of the considered videos, the standard resolution sequences were downscaled by a factor of 2: the original size of 352x288 pixels has changed to 176x144 pixels. Despite the resulting sequences have been clearly small, the results in terms of false negatives and BER were at optimum levels, as reported in Table 7. About HD videos, due to the greater resolution, it was decided for two levels of downscaling: one by a factor of 2 and the other by a factor of 4:

1. Blue Sky and Riverbed (1920x1080 pixels) were downscaled firstly to 960x540, then to 480x270 pixels.
2. Mobcal and Trees (1280x720 pixels) were downscaled firstly to 640x360 and then to 320x180 pixels.

The results, shown in Table 8, were good even in this case, especially regarding the BER values.

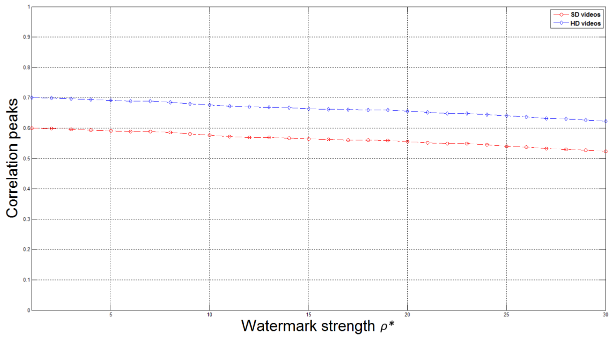


Figure 11. Average correlation peaks of the watermarked videos, after the insertion of a second watermark ρ^*.

Table 7. Downscaling of SD videos to 176X144 pixels resolution

SD Videos downscaled to 176x144 pixels	Akiyo	Coastguard	Bus	Football
False negative rate	0%	0%	1%	1%
BER	0%	1%	1%	0%

Table 8. Downscaling of HD videos to various resolutions

HD Videos	Sky *(to 960x540)*	Riverbed *(to 960x540)*	Mobcal *(to 640x360)*	Trees *(to 640x360)*
False negative rate	5%	4%	5%	6%
BER	1%	2%	2%	1%
HD Videos	Sky *(to 480x270)*	Riverbed *(to 480x270)*	Mobcal *(to 320x180)*	Trees *(to 320x180)*
False negative rate	8%	8%	7%	9%
BER	3%	4%	4%	5%

6.3.7. *Estimation attack and temporal frame averaging.* Since in the decoding phase it wasn't used a key to decrypt the watermark, the security of the proposed scheme has been tested with two attacks: watermark estimation [25] and temporal frame averaging.

The first attack can be performed by collecting a certain number of frames, then estimating their watermarks and finally performing the average on the estimations to obtain a unique estimated watermark. The simulation of this attack has been conducted by calculating the difference between a frame and its version subjected to a high pass filter: the watermark, in fact, has been added to the high frequency coefficients. Therefore, the difference is performed to estimate the watermark and, finally, the estimated watermark is subtracted from the original frame to remove the watermark.

The average correlation peaks of examined videos, with increasing number of frames used to estimate the watermark, are shown in Fig.12 (left). It can be noticed that in the first $\Delta = 20$ frames the correlation peaks are reduced, since the pattern of the watermark, in the proposed algorithm, varies only after 20 frames: in this first interval the watermark will remain the same, however the correlation peaks are sufficiently above the threshold T, thus useful for the detection of the watermark. Then, the correlation peaks increase again, reaching the usual correlation values in the intermediate frames.

The simulation of the temporal frame averaging attack has been performed with a time window that captures multiple frames of a video. In particular, a sequence is subjected to a low pass filter: if the pattern of the watermark is varied in a small number of frames, the attack could be successful without excessively compromising the visual quality of the content. Even for this reason, in the designed algorithm the pattern of the watermark is changed every $\Delta = 20$ frames. In Fig.12 (right) it is illustrated the average correlation peaks after having applied this technique, using an increasing number of frames (up to 20). As a result, it is possible to note that only for high values of the window frames number (above 20), the correlation is considerably reduced (still above the threshold). However, for increasing frames numbers, the visual quality typically suffers considerably: if the adjacent frames of a video change quickly and they are not very similar, a low pass filtered window, composed even just by 3 frames, will lead to a poor video quality due to the considerable variability of the content.

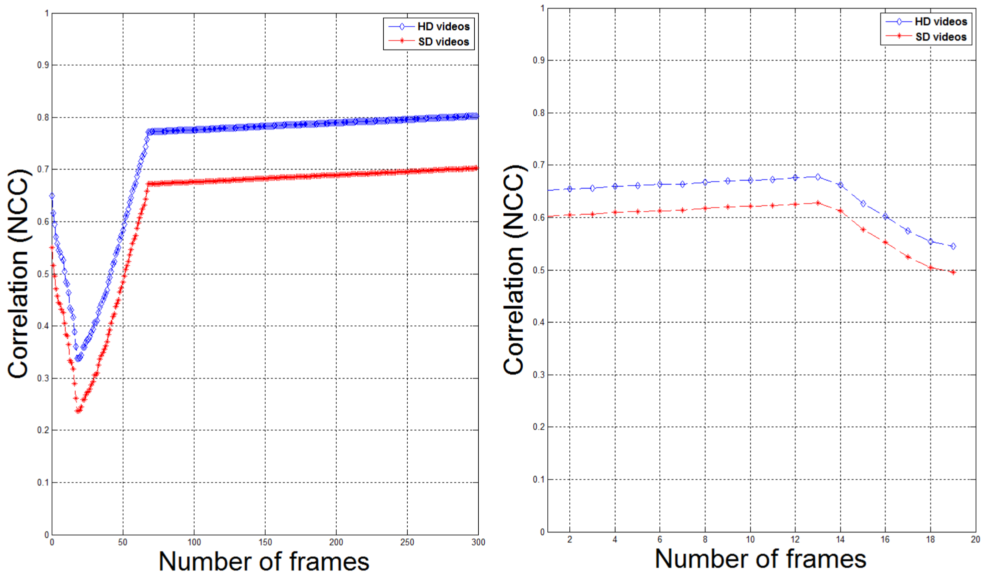


FIGURE 12. Effects of the estimation attack (left) and of the average time attack (right) on the average correlation peaks of the watermarked videos, with increasing number of collected frames used for the watermark's estimation.

6.3.8. *Execution time.* Encoding and decoding operations require an amount of time to be performed, depending on the type of the watermarking scheme, test videos resolution and length [26]. Therefore, the execution times of the two operations have been measured and then compared to similar algorithms. The experimental tests have been performed with MATLAB R2011b on a Personal Computer with i7-720QM 1.6 GHz Processor, 8 GB RAM and Windows 7 Professional OS; the execution times have been measured with the MATLAB functions tic and toc. The results are shown in Table 9. On the encoding phase, the execution time has been taken from the start of the video loading to the end of all the watermarking operations, including the creation of the watermarked file.

On the decoding phase, the execution time has been taken from the start of the watermarked video reading to the end of all the decoding operations, including NCC peaks calculation and binary signature authentication. We have compared the obtained results with those of four watermarking schemes in the state-of-the-art based on the Discrete Wavelet Transform (DWT). Since the analyzed watermarking tools and the corresponding datasets used in the experiments are not available, we were only able to carry out a rough comparison; at this aim, note also that the DWT has less computational costs than the DT CWT here adopted [27]. The collected computational costs are the following:

1. In [26] a level 3 DWT based data hiding algorithm is proposed; experiments on this paper have been carried out on different SD videos. These videos are longer than those used in this paper, but their resolution is on average lower than that of the SD test videos used in this paper. The embedding time is on average 6 minutes and 38 seconds, the extraction time is on average 3 minutes and 1 second.
2. Two approaches are described in [28]: one based on the level 3 DWT and another based on the level 5 DWT. The results of the second scheme (the fastest one) are 6 minutes and 19 seconds for the encoding, 2 minutes and 50 seconds for the decoding. Again, the considered videos are longer than those used in this paper, while they have on average a lower resolution than that of the SD videos used in this paper.
3. A mixed DWT-DCT approach is described in [29]. It isn't applied to the videos, but to the single image Apple: the watermarking process, lasts on average 68 seconds.
4. Finally, in [30] three watermarking schemes based on level 0 DWT are described: the scheme that models the Human Visual System takes on average 61 seconds to watermark a single image.

If compared with the previous methods, it appears that the computational cost of the proposed method is very competitive with respect the state of the art, such that an efficient implementation of it could allow its use in a practical application scenario.

Table 9. Algorithm's execution time

Quality of test videos	Average encoding time	Average decoding time	Average videos size
SD (Akiyo, Coastguard, Bus, Football)	41 seconds	1 minute and 46 seconds	60.45 MB
HD (Blue Sky, Riverbed, Mobcal, Trees)	23 minutes and 15 seconds	19 minutes and 14 seconds	1.05 GB

7. Conclusions.

The embedding of a binary signature and the channel V watermarking have proved efficacy to ensure a robust watermarking, with the possibility of linking a video to a user or a group of users. Besides, the enhanced robustness introduced by the V channel watermarking did not affect the visual quality of the tested videos, as evidenced by the reported PSNR and SSIM values.

Some attacks have determined effects in lowering NCC peaks and creating false negatives, but they have failed to compromise the proper binary signature extraction. Other attacks have led to a wrong interpretation of some bits of the embedded binary signature, but not to the lowing of the cross correlation peaks. In general, the algorithm has avoided the presence of both problems at the same time. Therefore, in the future it will be necessary to create a methodology that can effectively resist to the exposed attack techniques, without being affected either on the NCC or on the correct interpretation of the binary signature: both information must survive almost intact to the attacks.

Finally, it is possible to identify some possible paths for future developments:

1. Hardware implementation. A hardware implementation of the exposed algorithm can bring substantial benefits: faster execution, application in more fields and better overall performance.

2. Research in the Human Visual System (HVS). The DT CWT decomposition accurately models the human visual system, making it a suitable instrument for the realization of imperceptible watermarking algorithms [31]. It will be useful to continue the research in the HVS, in order to ensure more sophisticated solutions for a better hiding of the watermark.
3. Development of watermarking algorithm for 3D videos. The emerging market of 3D will inevitably produce new problems concerning the copyright of such contents, which will be protected depending on the technology used [32].

REFERENCES

[1] A. Garboan, M. Mitrea, and F. Preteux, Camcorder recording robust video fingerprinting. *Consumer Electronics (ISCE), 2012 IEEE 16th International Symposium on*, 2012.

[2] B. Schneier, Applied cryptography (2nd ed.): protocols, algorithms, and source code in C. 1995: John Wiley & Sons, Inc. 758.

[3] M. Maes, et al., Digital watermarking for DVD video copy protection, *Signal Processing Magazine, IEEE*, 2000. 17(5): p. 47-57.

[4] M. Asikuzzaman, et al., Imperceptible and Robust Blind Video Watermarking Using Chrominance Embedding: A Set of Approaches in the DT CWT Domain., *Information Forensics and Security, IEEE Trans. on*, 2014. 9(9): p. 1502-1517

[5] N. Kingsbury, Shift invariant properties of the dual-tree complex wavelet transform, *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 3. Mar. 1999, pp. 12211224.

[6] M. Vranjes, S. Rimac-Drlje, D. Zagar, Objective video quality metrics, *ELMAR, 2007*, 2007.

[7] T. Chen, W. Jingchun, and Z. Yonglei, Combined digital signature and digital watermark scheme for image authentication. *Info-tech and Info-net, 2001. Proceedings. ICII 2001 - Beijing. 2001 International Conferences on*, 2001.

[8] L. Chun-Shien, C. Jan-Ru, F. Kuo-Chin., Resistance of content-dependent video watermarking to watermark-estimation attacks, *Communications, 2004 IEEE International Conference on*, 2004.

[9] P. Loo, N. Kingsbury, Digital watermarking using complex wavelets, *Image Processing, 2000. Proceedings. 2000 International Conference on.* 2000.

[10] R. Hoffman, Data Compression in Digital Systems, 1997: p.98 Chapman & Hall, Ltd. 415.

[11] C. Jen-Shiun, et al., Saturation adjustment method based on human vision with YCbCr color model characteristics and luminance changes, *Intelligent Signal Processing and Communications Systems (ISPACS), 2012 International Symposium on*, 2012.

[12] R. Kwitt, P. Meerwald, and A. Uhl, Blind DT-CWT domain additive spread-spectrum watermark detection, *Digital Signal Processing, 2009 16th International Conference on*, 2009.

[13] G. Yan, and Z. Lin-lin, Simulation study of FIR filter based on Matlab, *in Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on*, 2010.

[14] M. Asikuzzaman, et al., A blind digital video watermarking scheme with enhanced robustness to geometric distortion, *Digital Image Computing Techniques and Applications (DICTA), 2012 International Conference on*, 2012.

[15] M. Barni and F. Bartolini, Watermarking Systems Engineering, *Signal Processing and Communications, 21*, CRC Press, Inc., Boca Raton, FL, USA, 2004.

[16] F. Garzia, Handbook of Communications Security, WIT Press, Southampton UK and Boston USA, 2013.

[17] V.G. Devereux, Limiting of YUV digital video signals, *BBC Research Department Report*, Dec. 1987.

[18] Z. Kotevski, and P. Mitrevski, Experimental comparison of PSNR and SSIM metrics for video quality estimation, *ICT Innovations 2009*, D. Davcev and J. Gmez, Editors, Springer Berlin Heidelberg, pp. 357-366 2010.

[19] M. Tong, T. Yan, and J. Hongbing, Watermarking technique resisting to strong cropping, *Intelligent Information Technology Application, 2008. IITA '08. Second International Symposium on*, 2008.

[20] A. Cozzolino, et al., Evaluating the effects of MJPEG compression on motion tracking in metro railway surveillance, *Advanced Concepts for Intelligent Vision Systems*, J. Blanc-Talon, et al., Editors, Springer Berlin Heidelberg, pp. 142-154, 2012

[21] W. Chuen-Ching, C. Ming-Jun, and C. Yao-Tang, New watermarking algorithm with coding efficiency improvement and authentication in video surveillance, em Intelligent Information Hiding and Multimedia Signal Processing, IIH-MSP '09. Fifth International Conference on, 2009.

[22] S.S. Al-Amri, N.V. Kalyankar, and S.D. Khamitkar, A comparative study of removal noise from remote sensing image, *CoRR, abs/1002.1148*, 2010.

[23] J.W. Goodman, Some fundamental properties of speckle, *Journal of the Optical Society of America*, 66(11): p. 1145-1150, 1976.

[24] R. Manjula and N. Settipalli, A new relational watermarking scheme resilient to additive attacks, *International Journal of Computer Applications*, 10(5): pp. 1-7, November 2010.

[25] T. Kalker, J.P. Linnartz, and M. van Dijk, Watermark estimation through detector analysis, *Image Processing, ICIP 98. Proceedings International Conference on*, 1998.

[26] T. Tabassum, and S.M.M. Islam, A digital video watermarking technique based on identical frame extraction in 3-Level DWT, *Computer and Information Technology (ICCIT), 15th International Conference on*, 2012.

[27] E. Lefebvre, Advances and Challenges in Multisensor Data and Information Processing, *Volume 8 NATO Security through Science Series: Information and Communication Security)*, IOS Press, 2007.

[28] V. C. Neha Goel, ,Krishna Mohan Rai, Study of 5 level DWT and comparative performance analysis of digital video watermarking techniques using 3-L-DWT and 5-L-DWT, *International Journal of Engineering and Technical Research (IJETR) 3(7)*, July 2015.

[29] R.B. Keta Raval, Unified approach to secure and robust digital watermarking scheme for image communication, *International Journal of Recent Technology and Engineering (IJRTE) 2(1)*, March 2013.

[30] C.S. Woo, et al., Performance factors analysis of a wavelet-based watermarking method, em Proceedings of the 2005 Australasian Workshop on Grid Computing and E-Research - Volume 44, Newcastle, New South Wales, Australia, Australian Computer Society, Inc.: 89-97, 2005.

[31] F.K.A Bouridane, M. Byrne, S. Boussakta, Host adaptive colour image watermarking using complex wavelets, *Q.s.U.B.S.o.E.a.E.E. School of Computer Science*, University of Leeds, Editor.

[32] K. Hee-Dong, et al., Robust DT-CWT watermarking for DIBR 3D images. *Broadcasting, IEEE Transactions on*, 58(4): p. 533-543, 2012.