Robust Digital Video Watermarking in the Spatial and Wavelet Domain

Radu Ovidiu Preda, Cristina Oprea, Ionuţ Pirnog, Lucian Andrei Perişoară
Faculty of Electronics, Telecommunications and Information Technology
Politehnica University of Bucharest
Bucharest, Romania
radu@comm.pub.ro, cristina@comm.pub.ro, ionut@comm.pub.ro, lperisoara@yahoo.com

Abstract — **This paper presents two blind video watermarking techniques in the spatial and wavelet domain proposed by the authors and compares the two approaches. The original watermark and the original, unwatermarked videos are not required for the watermark extraction process. The two methods are combinations of spread-spectrum and quantization based techniques. The watermarks used are binary images, containing the copyright information. The watermark is protected against singular bit errors with a Hamming error correction code. The spatial domain technique embeds a watermark bit by spreading it in a luminance block. The actual embedding is done using a quantization based approach. The wavelet based technique embeds the same watermark bit into a number of chosen detail wavelet coefficients of the middle wavelet sub-band. The resilience of the schemes is improved by redundantly embedding the same watermark in a number of video frames. We have tested the perceptual quality of the watermarked videos and the resilience of our schemes to eight different attacks in the spatial, temporal and compressed domain, for different quantization step sizes and different number of redundant frames. The test results show that our wavelet domain technique achieves better video quality and robustness to attacks than the spatial domain method.**

Keywords - Digital Video Watermarking; Copyright Protection; Spatial Domain; Wavelet Domain; Comparison; Perceptual Quality; Robustness to Attacks.

I. INTRODUCTION

Video watermarking techniques are characterized by the domain that the watermark is being embedded or detected, their capacity, the perceptual quality of the watermarked videos and their robustness to particular types of attacks. They can be divided into three main groups according to the domain, in which the watermark is embedded and extracted: spatial domain, frequency domain and compressed domain watermarking. We will focus here on spatial and frequency domain watermarking.

Spatial domain algorithms embed the watermark into the pixel values and no transforms are applied to the host signal during the embedding process. The most common techniques to insert the watermark into the host data in the spatial domain is via Least Significant Bit (LSB) modification, Spread Spectrum Modulation (SS) and Quantization Index Modulation (QIM). The LSB techniques are not robust to attacks because the LSB plane can be easily replaced by random bits, removing the watermark.

Spread spectrum methods view watermarking as a problem of communication through a noisy channel. As a means to combating this noise or interference, spread-spectrum techniques are employed to allow reliable communication in such noisy environments. In this case, the watermark data is coded with a pseudorandom code sequence to spread its power spectrum in the image or video, thus increasing its robustness to attacks. One of the first methods was the one-dimensional spread spectrum approach [1]. Here, the watermark is a pseudo-random sequence spread over the video frames by direct spatial domain addition. The watermark is repeatedly embedded throughout the video in a sequential manner. Other more complicated spread-spectrum methods were proposed in [2][3].

Quantization Index Modulation (QIM) refers to a class of data hiding schemes that exploit Costa's [4] findings by embedding information in the choice of quantizers. Over the past few years, QIM-based data hiding has received increasing attention from the data hiding community because it is more robust than techniques such as spread spectrum and LSB modification. State of the art proposed QIM schemes include Chen and Wornell's QIM and dither modulation [5], Eggers et al's scalar Costa scheme (SCS) [6], Jie and Zhiqiang's color image QIM scheme [7] and Kalantari and Ahadi's logarithmic QIM scheme [8].

In frequency domain watermarking, the most common transforms being used are the Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT) and Discrete Wavelet Transform. The main advantage offered by transform domain techniques is that they can take advantage of special properties of the alternate domains to address the limitations of pixel-based methods or to support additional features. Also, they have better resistance to compression based attacks. Generally, the main drawback of transform domain methods is their higher computational requirements.

Lately, algorithms in the Wavelet domain have gained more popularity due to their excellent spatial localization, frequency spread, and multi-resolution characteristics [9-14].

A lot of research has been done lately in developing new and improved watermarking techniques, but there is a difficulty in comparing the research results, because independent researchers use very different watermarks, watermark capacity, test videos, parameters for watermark

 ISBN: 978-1-61208-193-9

embedding and extraction and attacks with different parameters to test the robustness of their schemes. There is a need to compare the watermarking methods in different domains. Our paper addresses this issue by proposing two approaches in the spatial and Wavelet domain that have similar specifications, like watermark, watermark capacity, test videos, attacks with the same parameters. Both approaches embed the same watermark (binary image) with spatial and temporal redundancy and use a blind method for watermark extraction.

The rest of this paper is organized as follows: Sections II and III describe the proposed video watermarking techniques in the spatial and DWT domain, respectively, providing detailed diagrams and description of the watermark embedding and extraction strategies. Section IV contains the experimental results and a detailed comparison of the proposed methods in terms of perceptual quality and robustness to different attacks. Finally, Section V presents the conclusions of our work.

II. THE PROPOSED VIDEO WATERMARKING SCHEME IN THE SPATIAL DOMAIN

The watermark embedding process, illustrated in Fig. 1, is described in the following:

First, the original video is partitioned into groups of k frames. Every frame of the group is converted to the YC_bC_r color space.

The binary image matrix is transformed into a binary row vector w of size $P = h \times v$. To protect the watermark against bit errors, a Hamming error correction code (m,n) with codeword length of m bits and data-word length of n bits is applied to the vector w. The size of the resulting watermark vector w_c is:

$$P' = P\frac{m}{n} \tag{1}$$

The binary sequence w_c is partitioned into a number of $\frac{F}{k}$ sequences $w_c(j)$ of size $P'\frac{k}{F}$, where $j = \overline{1, \frac{F}{k}}$, F is the number of frames of the video and k is the number of redundant frames. The dimensions h and v of the watermark are chosen so that $P'\frac{k}{F}$ is an integer. The same sequence $w_c(j)$ will be inserted into every frame of a group j of k frames.

The size l of a square bloc of $l \times l$ luminance values is calculated in (2) to embed a bit of the watermark:

$$l = \left[\sqrt{\frac{MNC}{P'k}}\right], \tag{2}$$

where [.] is the integer part operator.

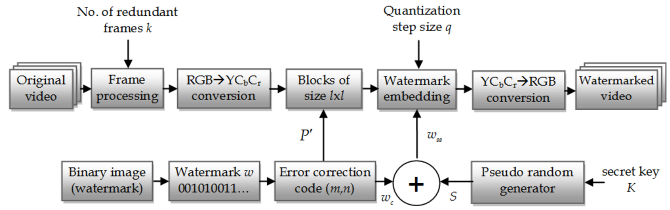

Figure 1. Block diagram of the spatial watermark encoder

A spread-spectrum technique is used to spread the power spectrum of the watermark data, thus, increasing its robustness against attacks. First a binary pseudo-random sequence $S = \{s_r | s_r \in \{0,1\}, r = 1,...,l^2\}$ of size l^2 with equal number of zeros and ones is generated using the Mersenne-Twister algorithm proposed in [15] with the use of the last 64 bits of the secret key K as seed for the generator. This method generates numbers with a period of $(2^{19937} - 1)/2$. For every bit of the watermark $w_c(j)$, the corresponding spread spectrum sequence is:

$$w_{ss} = \begin{cases} [s_1, s_2, ..., s_{l^2}], & \text{if } w_c = 0 \\ [\overline{s}_1, \overline{s}_2, ..., \overline{s}_{l^2}], & \text{if } w_c = 1 \end{cases} \tag{3}$$

A sequence S (representing one bit of the original watermark) is embedded in every bloc of $l \times l$ luminance values. A bit of S is embedded into the luminance value of the pixel of the same index by rounding its value to an even or odd quantization level. Rounding to an even quantization level embeds a "0", while rounding to an odd quantization level embeds a "1", as shown in (4):

$$L_w(i,j) = \left[\frac{L}{2q}\right] \cdot 2q + q \cdot w \cdot sign\left(L(i,j) - \left[\frac{L(i,j)}{2q}\right] \cdot 2q\right), \tag{4}$$

where $L(i,j)$ is the original luminance value, $L_w(i,j)$ is the watermarked luminance value, q is the quantization step size and $sign()$ is defined as:

$$sign(x) = \begin{cases} -1, & if\ x \le 0 \\ 1, & if\ x > 0 \end{cases} \tag{5}$$

The video is converted back to the RGB format, obtaining the watermark video.

The choice of the quantization step size q is a tradeoff between the perceptual quality of the watermarked video (q must have a small value) and the resilience of the watermarking scheme to attacks (q must have a big value).

The watermark extraction process, shown in Fig. 2, implies the following steps:

The watermarked video is partitioned into groups of k frames. Every frame of the group is converted to the YC_bC_r format. Every luminance frame is partitioned into square blocks of $l \times l$ luminance values. A bit of the spread spectrum sequence w_{ss}' of size l^2 is extracted from every luminance value of a block of size $l \times l$ using (6):

$$w' = \mathrm{mod}\,2\left(round\left(\frac{L_w(i,j)}{q} \right) \right), \tag{6}$$

where w' is the extracted watermark bit, $L_w(i,j)$ is the luminance value of the pixel at position (i,j), q is the quantization step size and *mod2* is the modulo2 function.

Using the 64 bit seed from the secret key K the binary sequence S is generated locally. The extracted watermark bit for the corresponding block is:

$$w_b' = \begin{cases} 0, & \text{if } \sum_{r=1}^{l^2} \left| w_{ss,r}' - s_r \right| \le \frac{l^2}{2} \\ 1, & \text{if } \sum_{r=1}^{l^2} \left| w_{ss,r}' - s_r \right| > \frac{l^2}{2} \end{cases} \tag{7}$$

A binary sequence $w'_{c,i}(j)$ is extracted from every frame of a group of k frames, where $i = \overline{1,k}$. The sequence $w'_c(j)$ is computed from $w'_{c,i}(j)$ using (8):

$$w'_c(j) = \begin{cases} 0, & \text{if } \sum_{i=1}^{k} w'_{c,i}(j) \le \frac{k}{2} \\ 1, & \text{if } \sum_{i=1}^{k} w'_{c,i}(j) > \frac{k}{2} \end{cases}, \quad j \in \{1,2,\ldots,P'\} \tag{8}$$

The resulting watermark bit stream w'_c of size *P'* is error corrected and the watermark w' of size P is obtained. The extracted binary image is obtained by reshaping the vector w' to a matrix of size $h \times v$.

The choice of the quantization step size q is a tradeoff between the perceptual quality of the watermarked video (q should have a small value) and the resilience of the watermarking scheme to attacks (q should have a big value).

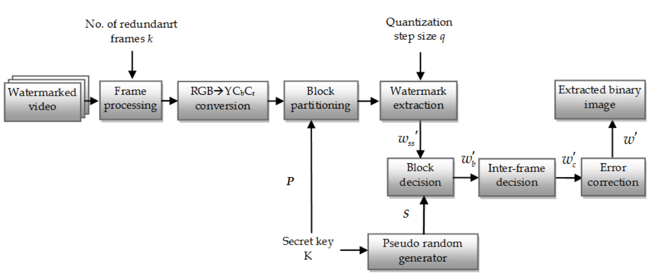

Figure 2. Block diagram of the spatial watermark decoder

III. THE PROPOSED VIDEO WATERMARKING SCHEME IN THE WAVELET DOMAIN

The watermark is embedded in the selected wavelet coefficients of the luminance Y of every frame of the video. The wavelet decomposition of the luminance is done using the 2D Discrete Wavelet Transform. We have chosen a Wavelet decomposition on L=3 resolution levels. The watermark is embedded in the wavelet coefficients of the LH, HL and HH sub-bands of the second Wavelet decomposition level. The choice of the second decomposition level is a tradeoff between the invisibility of the watermark and the resilience to attacks. A watermark embedded in the wavelet coefficients of the LH_1, HL_1 and HH_1 sub-bands is very sensitive to attacks, because these sub-bands contain the finest details of the frame. On the other hand, if we embed the watermark in the LH_3, HL_3 and HH_3 sub-bands, the perceptual quality of the video will be significantly altered. For these reasons, the best choice for watermark embedding is the second wavelet decomposition level.

For videos of resolution $M \times N$, the number of selected wavelet coefficients for a frame is:

$$C = 3\frac{MN}{2^{2(L-1)}} \tag{9}$$

The maximum capacity of the watermarking scheme is $C' = FC$ where F is the number of video frames and can be achieved by embedding a watermark bit in every selected wavelet coefficient. For example, for CIF videos of resolution 352x288 and 30 frames/s, the maximum capacity is 556kb/s. This maximum capacity is not needed in most applications, thus we will reduce it to improve the robustness of the scheme. Fig. 3 shows the block diagram of our Wavelet based watermark embedding scheme and is described in the following steps:

The binary image matrix is transformed into a binary row vector w of size $P = h \times v$. To protect the watermark against bit errors, a Hamming error correction code with codeword length of m bits and data word length of n bits is applied to vector w, resulting in a watermark vector w' of size P'.

A same spread-spectrum technique is used to spread the power spectrum of the watermark data. First the binary sequence $S = \left\{ s_j \,\middle|\, s_j \in \{0,1\}, j = 0,1,\ldots,G \right\}$ with equal number of zeros and ones is generated using the Mersenne-Twister algorithm with the use of 64 bits of the secret key K as seed for the generator. For every bit of the watermark *w'*, the corresponding spread spectrum sequence is:

$$w_{ss}(i) = \begin{cases} [s_1, s_2, \ldots, s_G], & if\ w'(i) = 0 \\ [\overline{s}_1, \overline{s}_2, \ldots, \overline{s}_G], & if\ w'(i) = 1 \end{cases}, \quad i = 1,\ldots,P' \tag{10}$$

Every sequence $w_{ss}(i)$ (representing one bit of the original watermark) is embedded into a number G of wavelet coefficients, every bit of $w_{ss}(i)$ in a wavelet coefficient.

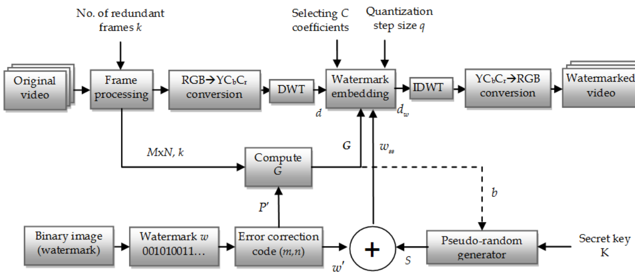

Figure 3. Block diagram of the wavelet watermark encoder

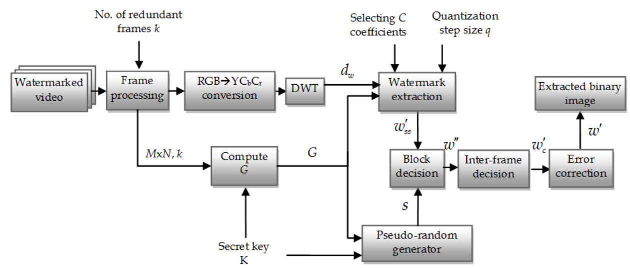

Figure 4. Block diagram of the wavelet watermark decoder

The number G depends on the number C of the selected wavelet coefficients, the number of frames F of the original video and the size P' of the watermark:

$$G = \left\lfloor \frac{C \cdot F}{P'} \right\rfloor \tag{11}$$

A bit of the binary sequence S is embedded in the selected wavelet coefficient by rounding its value to an even or odd quantization level. Rounding to an even quantization level embeds a "0", while rounding to an odd quantization level embeds a "1", as shown in (12):

$$d_w = \left\lfloor \frac{d}{2q} \right\rfloor \cdot 2q + q \cdot w \cdot sign\left(d - \left\lfloor \frac{d}{2q} \right\rfloor \cdot 2q \right), \tag{12}$$

where d is the original wavelet coefficient, d_w is the watermarked wavelet coefficient and q is the quantization step size.

After the entire watermark has been embedded, the 2D Inverse Discrete Wavelet Transform is computed for every frame to obtain the watermarked video.

The watermark extraction process, shown in Fig. 4, is explained in the following:

First, the wavelet decomposition of the watermarked, possibly attacked video is performed, then the wavelet coefficients used for embedding are selected. Parameter G is computed using the information about the size of the watermark provided by the secret key K. From every selected coefficient a bit is extracted according to (13), resulting in a sequence $w'_{ss}(j)$ of G bits from every group.

$$w' = \mathrm{mod}_2\left(round\left(\frac{d_w}{q} \right) \right), \tag{13}$$

where d_w is the watermarked wavelet coefficient.

Using the 64 bit seed from the secret key K the binary sequence S of size G is generated. The extracted watermark bit $w''(i)$ corresponding to a group of G wavelet coefficients is computed in (14).

$$w''(i) = \begin{cases} 0, if \ \sum_{j=1}^{G} [w'_j(i) - s_j] \le \frac{G}{2} \\ 1, if \ \sum_{j=1}^{G} [w'_j(i) - s_j] > \frac{G}{2} \end{cases}, i = 1, ..., P' \tag{14}$$

The resulting watermark bit stream of size P' is error corrected and the watermark w' of size P is obtained. The extracted binary image is obtained by reshaping the vector w' to a matrix of size $h \times v$.

To improve the resilience of the algorithm against temporal attacks we embedded the same watermark redundantly in every k frames. Thus, the number of wavelet coefficients used for embedding a watermark bit is decreased from G to G/k.

IV. Comparison of the Proposed Techniques

Our algorithms were tested using the first 27 frames of the videos "stefan", "forman" and "bus" in RGB uncompressed avi format, of resolution 352x288, 24 bits/pixel and frame rate of 30 frames/s. The binary image used as watermark is a copyright logo containing the name of one of the authors. The resolution of the image depends on the error correction code used, the number of redundant frames and the resolution of the initial video. The size of the watermark used is rather big, in order to better compare the two approaches. Using watermarks with smaller payload would improve the robustness, with BER values very close to zero for both methods, making it harder to compare them.

We have conducted the experiments for both methods using the quantization step sizes $q = 2$, $q = 4$, embedding of the same watermark in $k = 3$ and $k = 9$ frames, without using an error correction code and using a Hamming (7,4) error correction code.

To compare the perceptual quality of the watermarked video with the original one, we have computed the mean Peak Signal to Noise Ration (PSNR) of all frames of the video. The PSNR results are shown in Fig. 5. We can see that the best quality is obtained using the Wavelet approach. The PSNR results for the spatial watermarking scheme are quite low for quantization with bigger quantization step sizes (for $q = 4$ and $q = 8$ below the accepted value of 40 dB). For $q = 8$ only the wavelet based technique achieves a PSNR value higher than 40 dB.

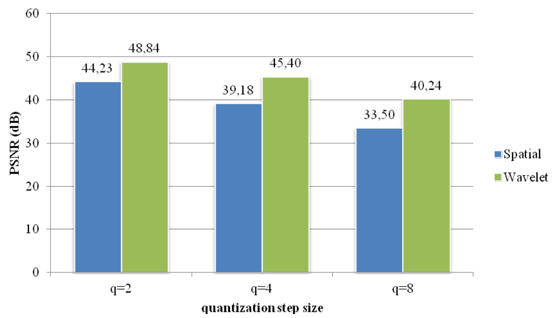

Figure 5. PSNR values for the proposed methods for different quantization step sizes

Next, we wanted to test the robustness of the proposed watermarking schemes. For this purpose we have carried out a range of eight attacks on the watermarked videos: (a) blurring of 2x2 pixel blocks, (b) brightening, adding Y=6 to the luminance of every pixel, (c) addition of Gaussian noise with mean 0 and variance 0,0003, (d) median filtering using a 3x3 pixel neighborhood, (e) addition of "salt and pepper" noise with density 0,3%, (f) frame averaging of 20% of the frames, where the current frame is the mean of the previous, current and next frame of the video, (g) JPEG compression of every frame using a quality factor Q=60 and (h) MPEG-2 compression at 4 and 2 Mbps. The parameters of the attacks were chosen in such a manner, that the visual degradation of the attacked videos is acceptable, because, by attacking a watermarked video, an attacker wants to destroy the watermark, but not the video quality.

To evaluate the robustness objectively, we have calculated the mean values of the decoding BER for the watermarks extracted from all test videos after they were attacked and plotted 6 different graphs (Fig. 6 - 8), where we represent the mean decoding BER for every method and every attack. The variables are the quantization step size q (chosen 2, 4 and 8) and the number of frames k used for embedding the same watermark (chosen 3 and 9). For $q = 2$ no error correction code was used, because the corresponding BER values are quite high and the Hamming (7,4) error correction would not work for such high bit error rates. For $q = 4$ and $q = 8$, with lower BER values, we used the Hamming (7,4) code, which can correct single bit errors.

The method working in the spatial domain is vulnerable to the brightening attack. For example by adding Y=6 to every luminance value, the decoding BER is 100% for every combination of parameters. We didn't represent this value on the graphs, because we didn't want to scale all BER values to 100%. On the other hand, the spatial embedding method has the best resilience to median filtering attacks. The weakness of the wavelet-based method to 3x3 median filtering can be improved by embedding the watermark in the third level wavelet subbands instead of the second. Because of the lower computational complexity, the spatial method could be used for real time processing.

The best overall resilience is achieved by the method working in the wavelet domain, with perfect decoding of the watermark for $q = 8$, $k = 9$ and Hamming (7,4) error correction.

V. Conclusion

In this paper we have compared our two proposed, blind video watermarking techniques in the spatial and wavelet domain. The original watermark and the original, unwatermarked videos are not required for the watermark extraction process. The methods are combinations of spread-spectrum and quantization based techniques. The watermarks used are binary images, containing the copyright information. The watermark is protected against singular bit errors using a Hamming error correction code.

The spatial domain technique embeds a watermark bit by spreading it in a luminance block. The actual embedding into a luminance value is done using a quantization based approach. The wavelet based technique embeds the same watermark bit into a number of detail wavelet coefficients of the middle wavelet sub-bands.

The resilience of the schemes is improved by redundantly embedding the same watermark in a number of k video frames. We have tested the perceptual quality of the watermarked videos and the resilience of the schemes to eight different attacks in the spatial, temporal and compressed domain, for different quantization step sizes and different number of redundant frames.

The experimental results show, that the wavelet domain technique achieves better video quality and better robustness to most attacks. The spatial domain method is most vulnerable to the brightening attack. The wavelet based technique achieves very good overall scores, being the better candidate for robust video watermarking.

References

[1] F. Hartung and B. Girod, "Watermarking of uncompressed and compressed video", Signal Processing, vol. 66, no. 3, pp. 283–301, May 1998.

[2] H. O. Altun, A. Orsdemir, G. Sharma, and M. F. Bocko, "Optimal Spread Spectrum Watermark Embedding via a Multistep Feasibility Formulation", IEEE Transactions on Image Processing, vol. 18, no. 2, pp. 371-387, Feb. 2009.

[3] S. P. Maity and S. Maity, "Multistage Spread Spectrum Watermark Detection Technique Using Fuzzy Logic", IEEE Signal Processing Letters, vol. 16, no. 4, pp. 245-248, April 2009.

[4] M. H. M. Costa, "Writing on dirty paper", IEEE Transactions on Information Theory, vol. IT-29, no. 3, pp. 439–441, May 1983.

[5] B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding", IEEE Transactions on Information Theory, vol. 47, pp. 1423–1443, May 2001.

[6] J. J. Eggers, R. Bauml, R. Tzschoppe, and B. Girod, "Scalar Costa scheme for information embedding", IEEE Transactions On Signal Processing, vol. 51, no. 4, pp. 1003–1019, 2003.

[7] N. Jie and W. Zhiqiang, "A new public watermarking algorithm for RGB color image based on Quantization Index

Modulation", International Conference on Information and Automation, ICIA '09, pp. 837-841, June 2009.

[8] N.K. Kalantari and S.M. Ahadi, "A Logarithmic Quantization Index Modulation for Perceptually Better Data Hiding", IEEE Transactions on Image Processing, vol. 19, no. 6, pp. 1504-1517, June 2010.

[9] M. Barni, F. Bartolini, and A. Piva, "Improved wavelet-based watermarking through pixel-wise masking", IEEE Transactions on Image Processing, vol. 10, no. 5, pp. 783-791, 2001.

[10] D. Zou, Y.Q. Shi, Z. Ni, and W. Su, "A Semi-Fragile Lossless Digital Watermarking Scheme Based on Integer Wavelet Transform", IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, no. 10, pp. 1294-1300, Oct. 2006.

[11] G. S. El-Taweel, H. M. Onsi, M.Samy, and M.G. Darwish, "Secure and Non-Blind Watermarking Scheme for Color Images Based on DWT", GVIP Special Issue on Watermarking, 2007.

[12] L.E. Coria, M. R. Pickering, P. Nasiopoulos, and R. K. Ward, "A Video Watermarking Scheme Based on the Dual-Tree Complex Wavelet Transform", IEEE Transactions on Information Forensics and Security, vol. 3, no. 3, pp. 466-474, Sept. 2008.

[13] R. O. Preda and N. Vizireanu, "Robust wavelet-based video watermarking scheme for copyright protection using the human visual system", Journal of Electronic Imaging 20, 013022, 2011.

[14] R. O. Preda and N. Vizireanu, "Quantization-based video watermarking in the wavelet domain with spatial and temporal redundancy", International Journal of Electronics, Vol. 98, Issue 3, pp. 393-405, 2011.

[15] M. Matsumoto and T. Nishimura, "Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudorandom Number Generator", ACM Transactions on Modeling and Computer Simulation, vol. 8, no. 1, pp. 3-30, 1998.

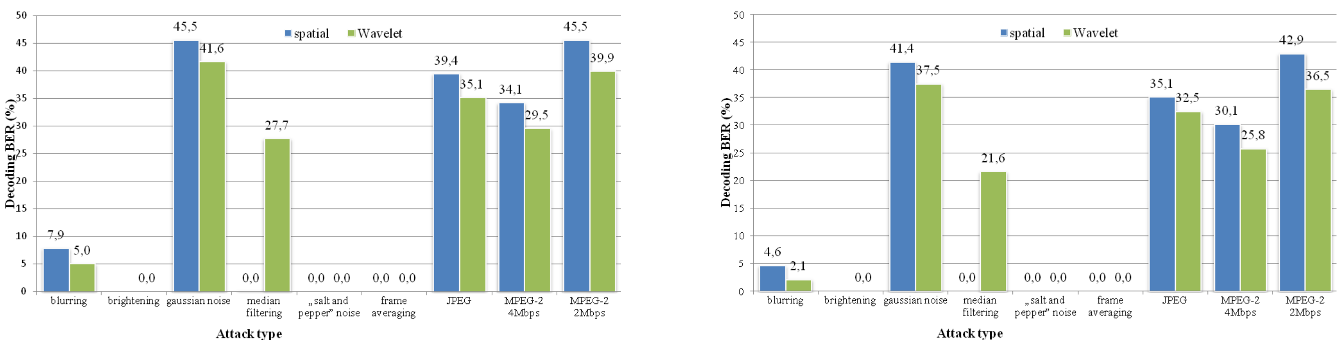

Figure 6. Decoding BER (%) for the proposed methods using $q = 2$, no error correction code and (a) k=3 and (b) k=9 redundant frames

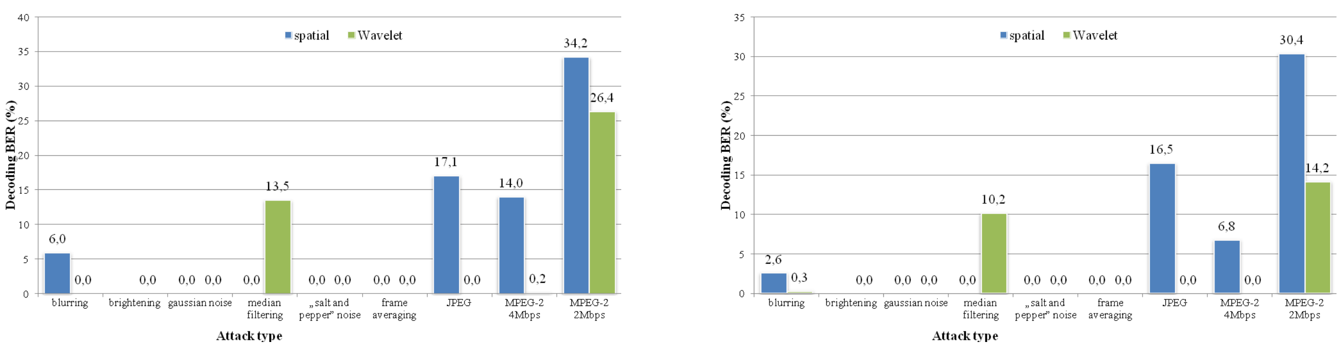

Figure 7. Decoding BER (%) for the proposed methods using $q = 4$, no error correction code and (a) k=3 and (b) k=9 redundant frames

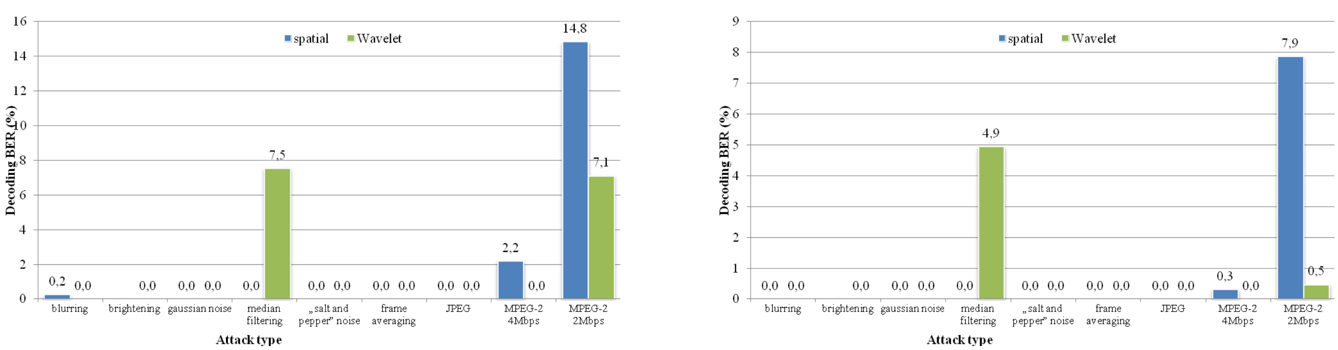

Figure 8. Decoding BER (%) for the proposed methods using $q = 8$, no error correction code and (a) k=3 and (b) k=9 redundant frames

 ISBN: 978-1-61208-193-9