An Application of Multidimensional Time-Frequency Analysis as a base for the Unified Watermarking Approach

Srdjan Stanković, Irena Orović, and Nikola Žarić

Abstract— **A watermarking approach based on multidimensional time-frequency analysis is proposed. It represents a unified concept that can be used for different types of data such as audio, speech signals, images or video. Time-frequency analysis is employed for speech signals, while space/spatial-frequency analysis is used for images. Their combination is applied for video signals. Particularly, we focus on the 2-D case: space/spatial-frequency based image watermarking procedure that will be subsequently extended to video signal. A method that selects coefficients for watermarking by estimating the local frequency content is proposed. In order to provide watermark imperceptibility, the non-stationary filtering is used to model the watermark which corresponds to the host signal components. Furthermore, the watermark detection within the multidimensional time-frequency domain is proposed. The efficiency and robustness of the procedure in the presence of various attacks is proven experimentally.**

I. INTRODUCTION

Digital watermarking has become an active research area focused on digital data protection, such as ownership and copyright protection [1]-[4]. It is based on embedding a secret signal called watermark into the digital content that should be protected. In most cases, the quality of original data has to be preserved i.e. the watermark should be imperceptible. In digital communications, the data are often exposed to various malicious or non-malicious attacks. Hence, the watermark should be robust and detectable within the host data even in the presence of attacks. However, the watermark imperceptibility and robustness are mutually opposite requirements and the main challenge is to find the best trade-off. Various approaches are developed depending on the type of host signal, since it is difficult to provide a general solution. Also, different domains have been used: time (or space domain), spectral domains such as DFT, DWT, DCT domain and joint time/space-frequency domain [5]-[16].

IEEE Transactions on Image Processing, Vol. 1, No. 3, March 2010

A watermarking procedure based on the multidimensional time-frequency analysis is proposed in this paper. It represents a unified watermarking approach that can be used for different types of signals: 1-D (speech/audio), 2-D (image) and 3-D signals (video). One-dimensional time-frequency analysis is employed for audio signals, 2-D for images, while their combination is used for video signals. Here, the 2-D and 3-D cases are considered, while the 1-D case has already been studied in [9].

The time-frequency representation is used to select the dynamic regions of the host signal where the presence of watermark is difficult to perceive. Moreover, the watermark is modeled to follow specific signal components in the selected time-frequency region. For this purpose, the concept of non-stationary filtering is adapted to produce the appropriate time-frequency mask [17], [18]. Also, this concept provides the inverse mapping from the time-frequency domain to signal domain. The watermark embedding and detection are performed within the same domain.

In particular, we focus on the 2-D image signal and the corresponding watermarking procedure in the joint space/spatial-frequency domain. Additionally, the extension to video watermarking approach is considered. The initial idea has been introduced in [10]. It is also interesting to mention a few related pro-

cedures from the literature. In [11], a watermark is created as a 2-D chirp signal embedded in the spatial domain. The Radon-Wigner distribution is used to detect the peaks that appear at certain frequencies. Further, the use of the Wigner distribution in image watermarking has also been reported in [12], [13]. The 1-D Wigner distribution is calculated for image rows or columns and then certain time-frequency cells are used for watermarking. A difficulty appears in the inversion of the Wigner distribution, which is formulated as signal synthesis problem.

Unlike other approaches, the proposed procedure is based on the space/spatial-frequency representation that provides the estimation of local frequency content for each image pixel. The efficient space/spatial-frequency representation is provided by using the 2-D S-method [19], [20]. It provides information about distribution of local frequency content within the considered space/spatial-frequency region. In order to ensure watermark imperceptibility, the local frequency content is used to determine whether the image region is suitable for watermarking. Also, the watermark characteristics are modeled according to the space/spatial-frequency components of each selected image region. The modeling procedure and inverse mapping from the space/spatial-frequency domain are realized by using the concept of space-varying filtering. The successful correlation based watermark detection is provided within the space/spatial-frequency domain. The efficiency of watermark detection in the presence of attacks is demonstrated by various examples. This approach has also been used for digital video signals, where the combination of 1-D and 2-D time-frequency analysis is applied.

II. Theoretical Background – 2-D Time-Frequency Analysis

Time-frequency analysis has been efficiently used in many practical applications with 1-D signals, such as speech, radar and sonar signals [18]. Subsequently, it has been extended to space/spatial-frequency analysis for the applications of image filtering and texture segmentation [20]. In this work, the time-frequency analysis and the concept of non-stationary filtering is essential for the proposed unified watermarking approach. Hence, a review of these concepts for 2-D signals is provided in the sequel.

The simplest time-frequency representation is obtained by using the short-time Fourier transform (STFT). The discrete form of the STFT for 2-D signal $I(x,y)$ can be written as:

$$STFT(x,y,\omega_x,\omega_y) = \\ = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{I(x+\xi,y+\nu)\; w(\xi,\nu) \\ \times e^{-j(\omega_x\xi+\omega_y\nu)}\}\, d\xi d\nu \;, \qquad (1)$$

where $w(\xi,\nu)$ is a 2-D window function. The spectrogram is defined as a squared module of the STFT: $SPEC(x,y,\omega_x,\omega_y) = |STFT(x,y,\omega_x,\omega_y)|^2$.

Similarly as in the 1-D case, if the second and higher order derivatives of signal phase function are not negligible, the spectrogram cannot provide a satisfying concentration at the local frequency. In order to improve concentration, the Wigner distribution has been used. Its pseudo form is defined as [20]:

$$WD(x,y,\omega_x,\omega_y) = \\ = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(\xi,\nu)w^*(-\xi,-\nu)I(x+\tfrac{\xi}{2},y+\tfrac{\nu}{2}) \qquad (2)$$

Note that the 2-D pseudo Wigner distribution can be written in the form:

$$WD(x,y,\omega_x,\omega_y) = \\ = \frac{1}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{STFT(x,y,\omega_x+\theta_x,\omega_y+\theta_y) \\ \times STFT^*(x,y,\omega_x-\theta_x,\omega_y-\theta_y)\}\, d\theta_x d\theta_y. \qquad (3)$$

It provides a good concentration in the time-frequency domain, but in the case of multi-component signals, beside the auto-terms, the Wigner distribution also contains the cross-terms. Starting from the idea to remove the cross-terms, but to preserve the concentration as in the Wigner distribution, the S-method

has been introduced in [19]. The 2-D form of the S-method can be defined as [20], [24]:

$$SM(x,y,\omega_x,\omega_y) = \\ = \frac{1}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(\theta_x,\theta_y) \\ \times STFT(x,y,\omega_x+\theta_x,\omega_y+\theta_y) \\ \times STFT^*(x,y,\omega_x-\theta_x,\omega_y-\theta_y)d\theta_x d\theta_y. \tag{4}$$

Note that for $P(\theta_x,\theta_y) = 2\pi\delta(\theta_x,\theta_y)$ the spectrogram is obtained, while for $P(\theta_x,\theta_y) = 1$, the pseudo Wigner distribution follows. Thus, one may conclude that the S-method is "between" the spectrogram and the Wigner distribution. Moreover, it combines good properties of both. The S-method improves the concentration of spectrogram towards the concentration of the Wigner distribution. In the discrete form, the S-method can be written as:

$$SM(n_1,n_2,\omega_1,\omega_2) = \\ = \sum_{i_1=-L}^{L}\sum_{i_2=-L}^{L} P(i_1,i_2) \\ \times STFT(n_1,n_2,\omega_1+i_1,\omega_2+i_2) \\ \times STFT^*(n_1,n_2,\omega_1-i_1,\omega_2-i_2) \\ = SPEC(n_1,n_2,\omega_1,\omega_2)+ \\ +2Re\{\sum_{i_1=0}^{L}\sum_{i_2=0}^{L} P(i_1,i_2) \\ \times STFT(n_1,n_2,\omega_1+i_1,\omega_2+i_2) \\ \times STFT^*(n_1,n_2,\omega_1-i_1,\omega_2-i_2)\}, \tag{5}$$

where $P(i_1,i_2)$is a rectangular separable window of width $2L+1$ in both direction, while the discrete STFT is given by:

$$STFT(n_1,n_2,\omega_1,\omega_2) = \\ = \sum_{k_1=-N/2}^{N/2-1}\sum_{k_2=-N/2}^{N/2-1} I(n_1+k_1,n_2+k_2) \\ \times w(k_1,k_2)e^{-j2\pi/N(\omega_1 k_1+\omega_2 k_2)}. \tag{6}$$

The S-method is a numerically more efficient approach than the Wigner distribution itself [20],[24]. The convergence within the $P(i_1,i_2)$ is fast, so that the satisfactory concentration could be achieved even for a small value of L in (5) (e.g. L=3).

The time-frequency analysis combined with the concept of non-stationary filtering can be used to generate a signal with a specific time-frequency characteristics. The concept of non-stationary filtering has been introduced in [17]. In order to make it suitable for practical implementations, the 2-D pseudo form of non-stationary filtering is defined as follows:

$$(Hs)(x,y) = \frac{1}{4\pi^2}\int_{\omega_x}\int_{\omega_y} L_H(x,y,\omega_x,\omega_y). \\ \times STFT(x,y,\omega_x,\omega_y)d\omega_x d\omega_y \tag{7}$$

The support function $L_H(x,y,\omega_x,\omega_y)$is defined as Weyl symbol mapping to the space/spatial-frequency domain [17]. Assuming that the signal components lie inside the 2-D region R_f, while the noise is outside this region, the support function $L_H(x,y,\omega_x,\omega_y)$can be defined as [17]:

$$L_H(x,y,\omega_x,\omega_y) = \\ = \begin{cases} 1 & \text{for} \quad (x,y,\omega_x,\omega_y) \in R_f \\ 0 & \text{for} \quad (x,y,\omega_x,\omega_y) \notin R_f \end{cases} \tag{8}$$

The support region R_f is usually determined by using the Wigner distribution, or equivalently, by using the S-method. The determination of the support region based on the spectrogram would be appropriate only when local frequency does not vary in space or varies very slowly, since in this case the spectrogram is sufficient to provide good concentration [17].

III. Image watermarking approach based on the space/spatial-frequency representation

In this section the space/spatial-frequency representation and the concept of non-stationary filtering are used for the image watermarking purposes. This approach is realized through the following steps:

- A procedure based on the space/spatial-frequency representation is used to select the

image pixels suitable for watermark embedding. A pixel $I(n_1,n_2)$ is suitable for watermarking only if the surrounding region of pixels, captured by the NxN window of STFT, is a busy region.

• The local space/spatial-frequency characteristics of each selected pixel are used to model the space-spatial frequency content of watermark. Here, the middle frequency coefficients of 2-D STFT are considered. As a result, each pixel will have a corresponding NxN matrix that is the STFT of watermark.

• The watermark is additively embedded in the space/spatial-frequency domain. Hence, the watermarked pixel $I_w(n_1, n_2)$is obtained as:

$$I_w(n_1, n_2) = \\ = \frac{1}{4\pi^2} \sum_{\omega_1} \sum_{\omega_2} [STFT_I(n_1, n_2, \omega_1, \omega_2) \\ + STFT_{w_{key}}(n_1, n_2, \omega_1, \omega_2)] , \quad (9)$$

where $STFT_I(n_1, n_2, \omega_1, \omega_2)$ and $STFT_{w_{key}}(n_1, n_2, \omega_1, \omega_2)$ are the STFTs of image pixel and watermark on a position (n_1,n_2), respectively. Frequencies ω_1 and $\omega_2$belongs to the middle range, which is commonly used for watermarking in other frequency domains, as well (e.g. DCT or DFT). The same procedure is performed for all the selected pixels. It is important to mention that unlike, for example, the DCT blocks, the space/spatial-frequency regions (of size NxN) could overlap, since any modification of the region affects only the central pixel.

• The detection is performed by using the form of correlation detector in the space/spatial-frequency domain. Namely, the STFT is calculated for each watermarked pixel and then correlated with the corresponding STFT of the watermark. In this case, the information about watermark will be spread over a large number of coefficients in the space/spatial-frequency domain, even when a small number of pixels is used. In other words, the proposed detection is flexible regarding the number of considered pixels, because for a small number of pixels there is still large number of elements in correlation.

In the sequel, the procedure for pixel selection and watermark modeling procedure will be discussed. Also, the form of correlation based detector in the space/spatial-frequency domain will be defined.

A. Procedure for selection of pixels suitable for watermarking

Watermark embedding in the pixels of busy image regions, i.e. regions with significant dynamics, can make it difficult to perceive. At the same time, the watermark can be more robust (a stronger watermark can be embedded). In order to determine busy image regions that will be used for watermarking, we need a space/spatial-frequency representation that is able to emphasize its dynamics. Thus, instead of the simplest choice i.e. the spectrogram, the S-method is used. In comparison with spectrogram, it will increase the concentration of components within busy regions, especially at the middle frequencies. In the case of flat regions, middle frequencies spectrogram components are too weak or do not exist, so they will remain insignificant within the S-method, as well.

As an illustration of the S-method advantages, let us observe two grayscale regions in Fig 1.a and Fig 1.d. The first one consists of two flat parts with different luminance, while the second one contains more dynamic changes and would be more appropriate for watermark embedding. Note that the spectrogram is similar for both regions. However, the S-method significantly highlights the components in the middle frequency range of the second region that correspond to the fast varying details (Fig 1.f).

The procedure for deciding whether a certain region is suitable for watermarking may be summarized as follows:

1. The S-method is calculated for the region of pixels within the window of size NxN and the middle frequency range is considered: $D_m^n = \{(\omega_1, \omega_2) : \rho_1 < \omega_1, \omega_2 < \rho_2\}$. The S-method within$D_m^n$ is denoted asSMD_m^n.
2. If SMD_m^n contains sufficient number of components whose energy is above a floor value:

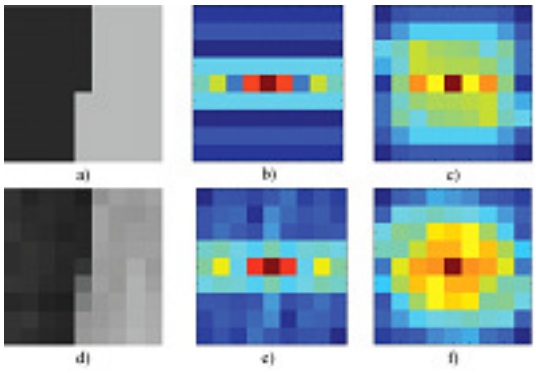


Fig. 1. Comparison between the S-method and the spectrogram: a) first region with two flat parts, b) spectrogram of the first region, c) the S-method of the first region, d) second region with more luminance changes, e) spectrogram of the second region, f) the S-method of the second region

$$No\{|SMD_m^n(n_1,n_2,\omega_1,\omega_2)| > S\} > No_{Ref}, \quad (10)$$

the region contains significant dynamics and will be used for watermarking. The function $No\{\}$ returns the number of components that satisfy the condition within the parenthesis, while$No_{Ref}$is a reference number of points used to make a decision about region's dynamics. The parameter S is an energy floor. For different image regions, the energy of middle frequency content changes consistently with the DC frequency component that brings the information about average energy of the region. It is a maximal component of the S-method within the whole region: $\max(SM(n_1,n_2,\omega_1,\omega_2))$. Therefore, the energy floor S can be determined as a portion of the S-method maximum:

$$S = \lambda \cdot 10^{\lambda \log_{10}(\max(SM(n_1,n_2,\omega_1,\omega_2)))}, \quad (11)$$

where λ takes a value between 0 and 1. This procedure provides efficient distinguishing between the dynamic regions on one side, and flat or slow varying regions on the other side. If the spectrogram was used instead of the S-method, it would be difficult to apply and set up the energy floor defined in terms of its maximal component. Namely, in the case of spectrogram, the energy of the middle frequency content is for several orders of magnitude lower than its maximal component. On the other side, by using the S-method, the energy of middle frequency content is increased. Consequently, it allows more efficient and precise use of the proposed energy floor S.

The energy floor S is obtained by using λ=0.7. The reference number of points is obtained as: No_{Ref}= $\lambda \cdot n$, where n is the total number of points in D_m^n. The value of λ is obtained experimentally by performing a large number of tests with various regions and images. Note that the same λ can be used for different window sizes in the STFT calculation (from 8x8 to 16x16). Some experimental results for various 9x9 image regions are given in Table I (the total number of points in the middle frequency range is n=20 and No_{Ref}=14). In addition, an example with other values of λ (close to 0.7) is given in Table II. Note that by using λ=0.65 even the flat region will be selected for watermarking (Region 3). On the other side λ=0.75 is a too strong criterion: it does not consider Region 1 and Region 2 although they are dynamic and suitable for watermarking. Finally, λ=0.7 suits all the cases. The pixels of images Lena and Peppers, that are selected by using the proposed procedure (with λ=0.7), are shown in Fig 2. (non-suitable pixels are marked with black points).

B. *Watermark modeling procedure*

The space/spatial-frequency content of watermark should be modeled according to the host components that are used for watermark embedding. The concept of non-stationary filtering is adapted and used for this purpose.

For each selected pixel, the watermark will be embedded in the middle frequency content D_m^n of 2-D STFT. The positions of components within D_m^n are determined by the support function:

$$L_{H_1}(n_1,n_2,\omega_1,\omega_2) = \\ = \begin{cases} 1 & \text{for} \quad \omega_1,\omega_2 \in D_m^n \\ 0 & \text{otherwise} \end{cases}. \quad (12)$$

In order to discard weak and insignificant components within D_m^n, the resulting support

TABLE I

SOME EXPERIMENTAL RESULTS OBTAINED BY USING THE PROPOSED PROCEDURE FOR DISCRIMINATION BETWEEN BUSY AND FLAT REGIONS

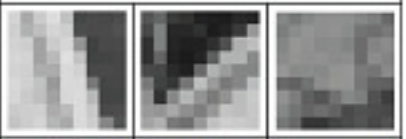
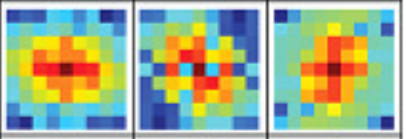
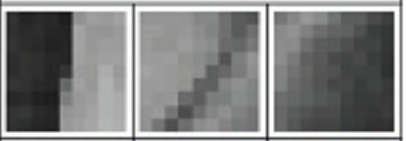
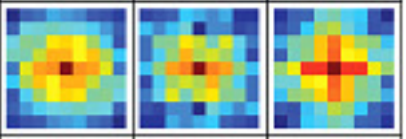
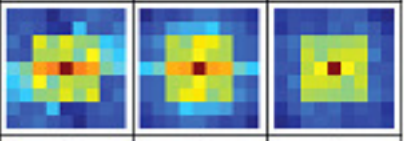
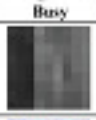
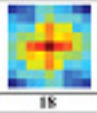
No_{Ref}=14	BUSY REGIONS		
Image region			
SM region			
No from 20	20	20	18
Decision	Suitable	Suitable	Suitable
No_{Ref}=14	BUSY REGIONS		
Image region			
SM region			
No from 20	18	16	14
Decision	Suitable	Suitable	Suitable
No_{Ref}=14	FLAT REGIONS		
Image region			
SM region			
No from 20	8	4	0
Decision	Non-suitable	Non-suitable	Non-suitable

TABLE II

RESULTS OF AN EXPERIMENT WITH DIFFERENT VALUES OF PARAMETER λ

	Region 1 Busy	Region 2 Busy	Region 3 Flat
Image region			
SM region			
No from 20, λ=0.65	18	18	16
No from 20, λ=0.7	16	14	8
No from 20, λ=0.75	10	8	4

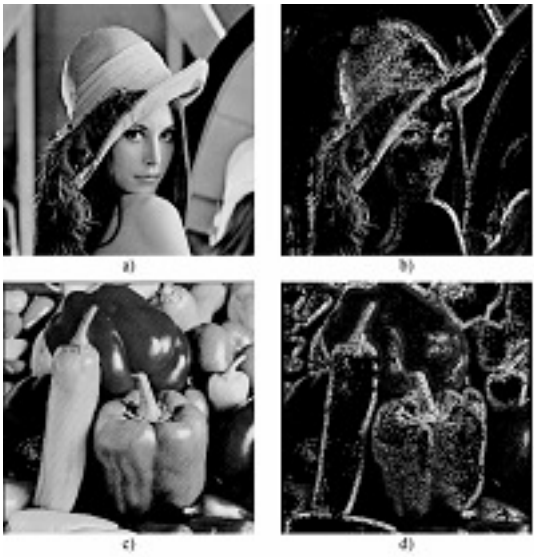


Fig. 2. a) Original image Lena, b) pixels of image Lena that are suitable for watermarking, c) original image Peppers, d) pixels of image Peppers that are suitable for watermarking

function is obtained as: $L_M(n_1, n_2, \omega_1, \omega_2) = L_{H_1}(n_1, n_2, \omega_1, \omega_2) \cap L_{H_2}(n_1, n_2, \omega_1, \omega_2)$, where:

$$L_{H_2}(n_1, n_2, \omega_1, \omega_2) = \\ = \begin{cases} 1 & \text{for } (\omega_1, \omega_2) \in \\ & \in |SMD_m^n(n_1, n_2, \omega_1, \omega_2)| > S \\ 0 & \text{for } (\omega_1, \omega_2) \in \\ & \in |SMD_m^n(n_1, n_2, \omega_1, \omega_2)| \leq S \end{cases}. \quad (13)$$

Thus, only the middle frequency components whose energy is above floor values S are used for watermark embedding. Their

locations within the space/spatial-frequency plane are determined by using the support function$L_M(n_1, n_2, \omega_1, \omega_2)$. Consequently, the local frequency content of watermark should be distributed over the same space/spatial-frequency locations. Thus, we may start with the STFT of an arbitrary random 2-D sequence $p(n_1, n_2)$and model its space/spatial-frequency characteristics by using L_M. The resulting watermark is obtained at the output of non-stationary filter as follows:

$$w_{_{key}}(n_1, n_2) = \\ = \sum_{\omega_1} \sum_{\omega_2} L_M(n_1, n_2, \omega_1, \omega_2) \\ \times STFT_p(n_1, n_2, \omega_1, \omega_2) \ , \qquad (14)$$

or equivalently, the STFT of watermark is:

$$STFT_{w_{key}}(n_1, n_2, \omega_1, \omega_2) = \\ = L_M(n_1, n_2, \omega_1, \omega_2) STFT_p(n_1, n_2, \omega_1, \omega_2).$$

In this way, the watermark with specific space/spatial-frequency characteristics is obtained. Moreover, the STFT of watermark contains only a certain middle frequency content of $STFT_p$ which is determined by L_M. This approach improves the watermark imperceptibility and provides high value of PSNR.

C. Watermark detection

Instead of a single pixel in the spatial domain, there is a corresponding NxN matrix of coefficients in the space/spatial-frequency domain. Consequently, in this domain the watermark will be distributed over large number of components which is very important for its detection. Thus, the suitable form of correlation detector is obtained by using the middle frequency coefficients of 2-D STFT, as follows:

$$Det = \sum_{i=1}^{N_w} \left[\sum_{\omega_1} \sum_{\omega_2} STFT_{I_w}(n_1, n_2, \omega_1, \omega_2) \right. \\ \left. \times \ STFT_{w_{key}}(n_1, n_2, \omega_1, \omega_2) \right] \ , \qquad (15)$$

where $STFT_{Iw}$ is STFT of watermarked pixel $I_w(n_1,n_2)$, while N_w is the total number of watermarked pixels. In the space/spatial-frequency domain, the number of components that brings the information about watermark could be more than ten times higher than in the spatial domain. This fact significantly contributes to the efficient correlation based detection, as it will be proved in the sequel.

IV. Examples

The efficiency of the proposed procedure will be demonstrated by the examples. The procedure is repeated for 500 trials (500 right keys - watermarks), in each example. For each right key, 100 wrong keys are generated (50000 in total). The detection is performed for all right and wrong trials and the mean values of detector responses are calculated:$\overline{D}(w_{key})$for right keys (watermarks) and $\overline{D}(w_{wr})$for wrong keys. The standard deviations of detector responses ($\sigma^2_{w_{key}}$ and $\sigma^2_{w_{wr}}$for right and wrong keys) are calculated, as well. Considering these parameters, the measure of detection quality is obtained as [21]:

$$R = \frac{\overline{D}(w_{key}) - \overline{D}(w_{wr})}{\sqrt{\sigma^2_{w_{key}} + \sigma^2_{w_{wr}}}}. \qquad (16)$$

The measure R corresponds to the detectability index used in signal detection theory to evaluate detection performance [21], [22]. It is used to calculate the probability of detection error:

$$P_{err}(R) = \frac{1}{2} erfc(\frac{R}{\sqrt{2}}). \qquad (17)$$

Example 1: The aim of this example is to illustrate the advantages of the proposed watermark detection in the space/spatial-frequency domain. The watermark embedding is done with a relatively high PSNR=50dB. The STFT is calculated by using the window of size 9x9. The number of watermarked pixels (second column in Table III), selected by using the proposed approach, is significantly less than the total number of image pixels. However, the proposed procedure provides a very low probability of error (Table III, **S/S-F detection**). It is the result of the significantly higher number of coefficients, within

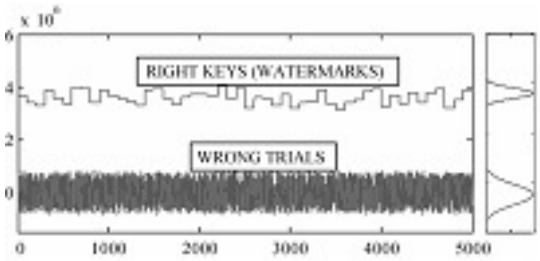


Fig. 3. The detector responses for 50 right keys and 5000 wrong trials

TABLE III
MEASURES OF DETECTION EFFICIENCY AND PROBABILITIES OF ERROR

Image	**No of pixels**	**S/S-F detection**		**Spatial detection**	
		R	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$
Lena	**4330**	8.25	$\tilde{}10^{-17}$	0.75	0.22
Peppers	**4830**	9.15	$\tilde{}10^{-20}$	0.9	0.18
F16	**3304**	7.95	$\tilde{}10^{-14}$	0.55	0.29
Boat	**6015**	9.2	$\tilde{}10^{-20}$	0.95	0.17
Barbara	**7833**	9.45	$\tilde{}10^{-21}$	1.1	0.13

the space/spatial-frequency domain, containing information about watermark. Experimentally obtained detector responses for right and wrong keys, for the image Lena, are given in Fig 3 (in order to provide better readability of their variations only 10% of the considered trials are presented).

On the other side, consider the case of additive watermark embedding in the spatial domain by using the same set of pixels. The watermark is a random sequence whose strength is adapted to provide the same value of PSNR=50dB as in the case of space/spatial-frequency based approach. Observe that the detection in spatial domain based on the correlation between watermarked pixels and watermark results in significantly higher probabilities of error (Table III, **Spatial detection**). Note that detection in spatial domain is tested with the same number of trials as in the case of the proposed procedure.

Example 2: The proposed procedure is tested in the presence of various attacks. The tests are performed by using different images.

Fig. 4. a) Original image Lena, b) Watermarked image Lena

The results for several of them are reported. The number of pixels selected for watermarking is: 4330 for the image Lena, 4830 for the image Peppers, 6015 for the image Boat, 3304 for the image F16, and 7833 for the image Barbara. The watermark is embedded in the middle frequency coefficients of STFT of size 9x9, providing PSNR=50dB. The original and the watermarked image Lena are shown in Fig 4. The measures of detection quality R and the corresponding probabilities of error are given in Table IV and Table V (**S/S-F procedure**). In order to compare the results, the DCT based image watermarking procedure is considered (with the same PSNR=50dB). Namely, according to [7], [8] all middle frequency DCT coefficients are used for watermarking (22050 coefficients for images of size 256x256, which is significantly greater than in the case of the proposed procedure). The detection is performed by using the standard correlation detector (**DCT procedure - Corr. Detector** in Table IV and Table V), detector based on the Generalized Gaussian function (GGF) [7] (**DCT procedure – GGF detector** in Table IV and Table V) and Cauchy detector proposed in [8] (**DCT procedure – Cauchy detector** in Table IV and Table V). The measures of detection quality and corresponding probabilities of error are calculated in the same way as for the proposed procedure.

Both procedures are tested in the presence of the following attacks (Table IV and Table V): JPEG compression with quality factors QF=80, 60, 50 and 40, median filter 3x3 and 5x5, average filter 3x3, impulse noise with variance 0.01, Gaussian noise with variance 0.003,

image lightening and image darkening for 20% of the original pixels values. Note that the proposed procedure outperforms the DCT based procedures by providing lower probabilities of detection error in the presence of all the considered attacks.

V. Digital Video Watermarking

Observe that the proposed procedure can be extended and used for video watermarking, as well. In the sequel, we propose an idea for time-frequency based video watermarking that may be combined with the existing approaches.

As a 3-D signal, video sequence should be a subject of 3-D time-frequency analysis that results in a 6-D time-frequency representation. At first sight it appears as an insurmountable obstacle. However, this can be solved by using the space/spatial-frequency analysis within a frame and, after that, by applying the time-frequency analysis for selected positions (x,y) within the sequence of video frames. A simple watermarking procedure can be based on the pixels whose space/spatial-frequency representation will not vary in time for a certain number of frames. Thus, the first step is selection of regions/pixels suitable for watermarking within one (initial) frame, following the procedure proposed for the images. The second step uses time-frequency analysis to determine which of the selected pixels produces space/spatial-frequency representation that is stationary in time (for a certain number of consecutive frames). Note that the space/spatial-frequency representation is stationary if the whole region of pixels used for its calculation is unaltered for a certain time interval. In most cases it is sufficient to observe a pixel on the central position and four pixels at the corners of the rectangular region associated with the size of 2-D windows for the STFT and the S-method calculation. Since the noise in video sequences can cause small variations of pixel values, time-frequency analysis is an appropriate tool for the pixel stationarity analysis.

The sequence of pixel values on the position (x,y) along the K consecutive frames is considered:

$$I_t(x,y) = [I_1(x,y), I_2(x,y), ..., I_K(x,y)]. \tag{18}$$

Based on $I_t(x,y)$, a frequency modulated signal $z(t)$ is created as:

$$z(t) = e^{j\mu(I_t(x,y)-\overline{I_t}(x,y))\cdot t}, \tag{19}$$

where: $\overline{I_t}(x,y) = mean(I_t(x,y))$, while μ is a constant and t is a time vector. Time-frequency distribution of $z(t)$ is obtained by calculating the 1-D S-method as follows:

$$SM_z(t,\omega) = \\ = \sum_{i=-L}^{L} P(i) STFT_z(t,\omega+i) STFT_z^*(t,\omega-i). \tag{20}$$

Note that, $\overline{\omega} = \arg\max\{SM_z(t,\omega)\} = I_t(x,y) - \overline{I_t}(x,y)$.
Therefore, if $\overline{\omega} = const$, the value of pixel on the position (x,y) will be unaltered within K consecutive frames. The illustrations of the $SM_z(t,\omega)$for pixels on different positions (x,y) are shown in Fig 5. Fig 5.a is obtained for a pixel that is unaltered within K frames, while Fig 5.b is obtained for a pixel that is changed due to the object motion within the considered sequence of frames. Similarly, in Fig 5.c and Fig 5.d five pixels are traced: at the corners and the center of the selected region. The value μ=0.0016 was used.

Example: The test video sequence was 60 frames long and then repeated in order to provide better illustrations. An illustration of video sequence with the object position in different frames is shown in Fig 6. Furthermore, two types of pixels are illustrated within a few frames in Fig 7. The first type is marked with white rectangle and represents a pixel that will remain unaltered within the sequence. The second type is marked with black ellipse and will change due to the object motion.

The selected pixels from busy frame regions will carry the watermark within K frames and then the selection procedure should be repeated for the next K frames. Note that the watermarking of one frame is done in the same manner described in previous sections.

TABLE IV

MEASURES OF DETECTION EFFICIENCY AND PROBABILITIES OF ERROR

Lena	S/S-F procedure		DCT procedure - Corr. detector		DCT procedure - GGF detector		DCT procedure - Cauchy detector	
	R	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$
No attack	8.27	$6.73{\cdot}10^{-17}$	6.53	$3.16{\cdot}10^{-11}$	7.1	$6{\cdot}10^{-13}$	7.25	$2{\cdot}10^{-13}$
JPEG 80	7.31	$1.25{\cdot}10^{-13}$	5.15	$1.27{\cdot}10^{-7}$	5.92	$1.6{\cdot}10^{-9}$	5.95	$3.31{\cdot}10^{-9}$
JPEG 60	6.75	$7.39{\cdot}10^{-12}$	3.72	$9.9{\cdot}10^{-5}$	4.76	$9.68{\cdot}10^{-7}$	4.9	$4.8{\cdot}10^{-7}$
JPEG 50	5.67	$7.08{\cdot}10^{-9}$	3.22	$6.4{\cdot}10^{-4}$	4.3	$8.53{\cdot}10^{-6}$	4.41	$5.17{\cdot}10^{-6}$
JPEG 40	5.15	$1.27{\cdot}10^{-7}$	2.93	$1.7{\cdot}10^{-3}$	3.48	$2.5{\cdot}10^{-4}$	3.6	$1.6{\cdot}10^{-4}$
Median 3x3	4.45	$4.17{\cdot}10^{-6}$	3.61	$1.53{\cdot}10^{-4}$	3.72	$9.9{\cdot}10^{-5}$	4.15	$1.66{\cdot}10^{-5}$
Median 5x5	2.86	$2.1{\cdot}10^{-3}$	2	$2.2{\cdot}10^{-2}$	2.03	$2.1{\cdot}10^{-2}$	2.22	$1.32{\cdot}10^{-2}$
Impulse noise	7.1	$6{\cdot}10^{-13}$	5.08	$1.88{\cdot}10^{-7}$	5.57	$1.25{\cdot}10^{-8}$	5.63	$9{\cdot}10^{-9}$
Gaussian noise	7.24	$2.2{\cdot}10^{-15}$	5.34	$4.64{\cdot}10^{-8}$	5.54	$1.5{\cdot}10^{-8}$	5.6	$1{\cdot}10^{-8}$
Lightening 20%	8.27	$6.73{\cdot}10^{-17}$	6.31	$1.4{\cdot}10^{-10}$	6.75	$7.39{\cdot}10^{-12}$	6.8	$5.2{\cdot}10^{-12}$
Darkening 20%	8.27	$6.73{\cdot}10^{-17}$	6.3	$1.48{\cdot}10^{-10}$	6.75	$7.39{\cdot}10^{-12}$	6.8	$5.2{\cdot}10^{-12}$
Average 3x3	5.11	$1.54{\cdot}10^{-7}$	3.11	$9.35{\cdot}10^{-4}$	3.4	$3.37{\cdot}10^{-4}$	3.68	$1.16{\cdot}10^{-4}$
Peppers	S/S-F procedure		DCT procedure - Corr. detector		DCT procedure - GGF detector		DCT procedure - Cauchy detector	
	R	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$
No attack	9.14	$3.1{\cdot}10^{-20}$	6.31	$1.4{\cdot}10^{-10}$	7.57	$1.86{\cdot}10^{-14}$	7.7	$6.8{\cdot}10^{-15}$
JPEG 80	8.14	$1.91{\cdot}10^{-16}$	5.15	$1.27{\cdot}10^{-7}$	6.8	$5.2{\cdot}10^{-12}$	6.86	$3.44{\cdot}10^{-12}$
JPEG 60	7.48	$3.68{\cdot}10^{-14}$	3.85	$5.9{\cdot}10^{-5}$	5.66	$7.58{\cdot}10^{-9}$	5.7	$6{\cdot}10^{-9}$
JPEG 50	6.53	$3.16{\cdot}10^{-11}$	3.57	$1.78{\cdot}10^{-4}$	4.93	$4.11{\cdot}10^{-7}$	5.08	$1.88{\cdot}10^{-7}$
JPEG 40	5.7	$5.8{\cdot}10^{-9}$	2.85	$2.2{\cdot}10^{-3}$	4	$3.16{\cdot}10^{-5}$	4.21	$1.27{\cdot}10^{-5}$
Median 3x3	5.08	$1.88{\cdot}10^{-7}$	3.1	$9.6{\cdot}10^{-4}$	4.25	$1.06{\cdot}10^{-5}$	5.34	$4.64{\cdot}10^{-8}$
Median 5x5	3.58	$1.66{\cdot}10^{-4}$	2	$2.2{\cdot}10^{-2}$	2.05	$1.8{\cdot}10^{-2}$	2.24	$1.25{\cdot}10^{-2}$
Impulse noise	8.23	$8.63{\cdot}10^{-17}$	5.05	$2.2{\cdot}10^{-7}$	5.4	$3.3{\cdot}10^{-8}$	5.46	$2.38{\cdot}10^{-8}$
Gaussian noise	8.29	$5.48{\cdot}10^{-17}$	5.02	$2.58{\cdot}10^{-7}$	5.15	$1.3{\cdot}10^{-7}$	5.37	$3.93{\cdot}10^{-8}$
Lightening 20%	9.14	$3.1{\cdot}10^{-20}$	5.7	$6{\cdot}10^{-9}$	7.19	$3.24{\cdot}10^{-13}$	7.35	$9.9{\cdot}10^{-14}$
Darkening 20%	9.14	$3.1{\cdot}10^{-20}$	5.67	$7.13{\cdot}10^{-9}$	7.19	$3.24{\cdot}10^{-13}$	7.3	$1.4{\cdot}10^{-13}$
Average 3x3	4.65	$1.6{\cdot}10^{-6}$	2.45	$7.1{\cdot}10^{-3}$	2.56	$5.2{\cdot}10^{-3}$	2.9	$1.9{\cdot}10^{-3}$
F16	S/S-F procedure		DCT procedure - Corr. detector		DCT procedure - GGF detector		DCT procedure - Cauchy detector	
	R	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$
No attack	7.95	$8.92{\cdot}10^{-16}$	5.57	$1.25{\cdot}10^{-8}$	6.66	$1.37{\cdot}10^{-1}$	6.4	$7.76{\cdot}10^{-11}$
JPEG 80	6.85	$3.59{\cdot}10^{-12}$	4.02	$2.9{\cdot}10^{-5}$	5.11	$1.61{\cdot}10^{-7}$	5.33	$4.9{\cdot}10^{-8}$
JPEG 60	5.95	$3.31{\cdot}10^{-9}$	3.3	$4.8{\cdot}10^{-4}$	4.54	$2.81{\cdot}10^{-6}$	4.6	$2.1{\cdot}10^{-6}$
JPEG 50	5.57	$1.25{\cdot}10^{-8}$	2.68	$3.7{\cdot}10^{-3}$	3.64	$1.36{\cdot}10^{-4}$	3.7	$1{\cdot}10^{-4}$
JPEG 40	5	$2.86{\cdot}10^{-7}$	2.3	$1{\cdot}10^{-2}$	2.75	$3{\cdot}10^{-3}$	2.95	$1.6{\cdot}10^{-3}$
Median 3x3	4.76	$9.68{\cdot}10^{-7}$	3.25	$5.77{\cdot}10^{-4}$	3.76	$8.5{\cdot}10^{-5}$	4.02	$2.9{\cdot}10^{-5}$
Median 5x5	2.65	$4{\cdot}10^{-3}$	2	$2.2{\cdot}10^{-2}$	2.05	$2{\cdot}10^{-2}$	2.15	$1.4{\cdot}10^{-2}$
Impulse noise	6.25	$2{\cdot}10^{-10}$	4.56	$2.55{\cdot}10^{-6}$	5.25	$7.6{\cdot}10^{-8}$	5.4	$3.3{\cdot}10^{-8}$
Gaussian noise	6.27	$1.8{\cdot}10^{-10}$	4.53	$2.95{\cdot}10^{-6}$	5.25	$7.6{\cdot}10^{-8}$	5.35	$4.4{\cdot}10^{-8}$
Lightening 20%	7.95	$8.92{\cdot}10^{-16}$	4.84	$6.5{\cdot}10^{-7}$	5.7	$6{\cdot}10^{-9}$	5.7	$6{\cdot}10^{-9}$
Darkening 20%	7.95	$8.92{\cdot}10^{-16}$	4.85	$6.17{\cdot}10^{-7}$	5.67	$7.14{\cdot}10^{-9}$	5.67	$6.73{\cdot}10^{-9}$
Average 3x3	5.15	$1.3{\cdot}10^{-7}$	2.25	$1.2{\cdot}10^{-2}$	2.66	$3.9{\cdot}10^{-3}$	3.03	$1.2{\cdot}10^{-3}$
Boat	S/S-F procedure		DCT procedure - Corr. detector		DCT procedure - GGF detector		DCT procedure - Cauchy detector	
	R	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$
No attack	9.21	$1.54{\cdot}10^{-20}$	5.82	$2.94{\cdot}10^{-9}$	6.85	$3.69{\cdot}10^{-12}$	7.25	$2.08{\cdot}10^{-13}$
JPEG 80	8.14	$1.91{\cdot}10^{-16}$	4.8	$7.9{\cdot}10^{-7}$	5.75	$4.46{\cdot}10^{-9}$	5.96	$1.26{\cdot}10^{-9}$
JPEG 60	7.6	$1.48{\cdot}10^{-14}$	4.19	$1.39{\cdot}10^{-5}$	5.03	$2.45{\cdot}10^{-7}$	5.15	$1.3{\cdot}10^{-7}$
JPEG 50	6.75	$7.39{\cdot}10^{-12}$	3.81	$6.94{\cdot}10^{-5}$	4.45	$4.29\cdot10^{-6}$	4.45	$4.29\cdot10^{-6}$
JPEG 40	5.95	$3.31{\cdot}10^{-9}$	3.14	$8.44{\cdot}10^{-4}$	3.56	$1.85{\cdot}10^{-4}$	3.62	$1.47{\cdot}10^{-4}$
Median 3x3	4.75	$1{\cdot}10^{-6}$	3.1	$9.6{\cdot}10^{-4}$	3.95	$3.9{\cdot}10^{-5}$	4.03	$2.78{\cdot}10^{-5}$
Median 5x5	3.15	$8.16{\cdot}10^{-4}$	1.98	$2.39{\cdot}10^{-2}$	2.2	$1.4{\cdot}10^{-2}$	2.2	$1.4{\cdot}10^{-2}$
Impulse noise	8.15	$1.89{\cdot}10^{-16}$	4.82	$7.17{\cdot}10^{-7}$	5.67	$7.14{\cdot}10^{-9}$	5.7	$6{\cdot}10^{-9}$
Gaussian noise	8.1	$2.74{\cdot}10^{-16}$	4.75	$1{\cdot}10^{-6}$	5.65	$8{\cdot}10^{-9}$	5.65	$8{\cdot}10^{-9}$
Lightening 20%	9.21	$1.54{\cdot}10^{-20}$	5.25	$7.6{\cdot}10^{-8}$	5.76	$4.2{\cdot}10^{-9}$	6.14	$4.12{\cdot}10^{-10}$
Darkening 20%	9.21	$1.54{\cdot}10^{-20}$	5.25	$7.6{\cdot}10^{-8}$	5.76	$4.2{\cdot}10^{-9}$	6.15	$3.87{\cdot}10^{-10}$
Average 3x3	5.15	$1.3{\cdot}10^{-7}$	4.53	$2.95{\cdot}10^{-6}$	4.53	$2.95{\cdot}10^{-6}$	4.8	$7.9{\cdot}10^{-7}$

TABLE V

CONTINUED - MEASURES OF DETECTION EFFICIENCY AND PROBABILITIES OF ERROR

Barbara	S/S-F procedure		DCT procedure - Corr. detector		DCT procedure - GGF detector		DCT procedure - Cauchy detector	
	R	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$	**R**	$\mathbf{P}_{err}$
No attack	9.46	$1.49{\cdot}10^{-21}$	6.52	$3.51{\cdot}10^{-11}$	7.75	$4.59{\cdot}10^{-15}$	8.13	$2.14{\cdot}10^{-16}$
JPEG 80	8.55	$6.15{\cdot}10^{-18}$	5.41	$3.15{\cdot}10^{-8}$	6.56	$2.69{\cdot}10^{-11}$	6.85	$3.69{\cdot}10^{-12}$
JPEG 60	7.65	$1{\cdot}10^{-14}$	4.6	$2.11{\cdot}10^{-6}$	5.51	$1.79{\cdot}10^{-8}$	5.82	$2.94{\cdot}10^{-9}$
JPEG 50	6.66	$1.37{\cdot}10^{-1}$	4.13	$1.81{\cdot}10^{-5}$	4.79	$8.34{\cdot}10^{-7}$	4.9	$4.8{\cdot}10^{-7}$
JPEG 40	5.95	$3.31{\cdot}10^{-9}$	3.45	$2.8{\cdot}10^{-4}$	3.86	$5.67{\cdot}10^{-5}$	4.05	$2.56{\cdot}10^{-5}$
Median 3x3	4.65	$1.6{\cdot}10^{-6}$	3	$1.3{\cdot}10^{-3}$	3.6	$1.6{\cdot}10^{-4}$	3.9	$4.8{\cdot}10^{-5}$
Median 5x5	2.66	$3.9{\cdot}10^{-3}$	1.87	$3{\cdot}10^{-2}$	2.06	$1.97{\cdot}10^{-2}$	2.12	$1.7{\cdot}10^{-2}$
Impulse noise	8.27	$6.73{\cdot}10^{-17}$	5.04	$2.37{\cdot}10^{-7}$	5.6	$1.1{\cdot}10^{-8}$	5.95	$3.31{\cdot}10^{-9}$
Gaussian noise	8.15	$1.89{\cdot}10^{-16}$	5.2	$9.96{\cdot}10^{-8}$	5.55	$1.42{\cdot}10^{-8}$	5.72	$5.32{\cdot}10^{-9}$
Lightening 20%	9.6	$4{\cdot}10^{-22}$	5.61	$1{\cdot}10^{-8}$	7	$1.28{\cdot}10^{-12}$	7.02	$1.1{\cdot}10^{-12}$
Darkening 20%	9.6	$4{\cdot}10^{-22}$	5.61	$1{\cdot}10^{-8}$	6.96	$1.7{\cdot}10^{-12}$	7.02	$1.2{\cdot}10^{-12}$
Average 3x3	5.55	$1.42{\cdot}10^{-8}$	3.22	$4.8{\cdot}10^{-4}$	3.34	$4.19{\cdot}10^{-4}$	3.43	$3{\cdot}10^{-4}$

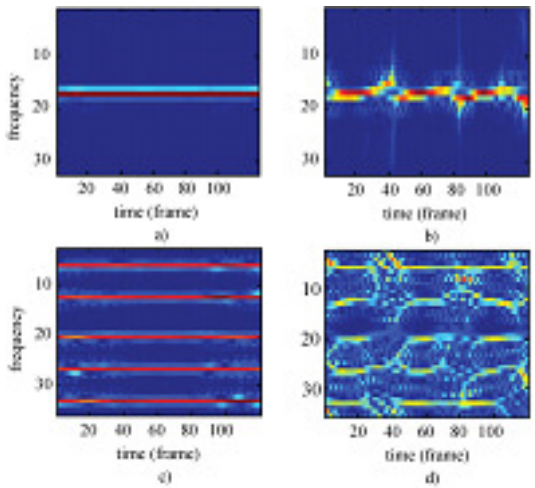


Fig. 5. $SM_z(t,\omega)$ obtained within K=120 frames for: a) a stationary pixel, b) a pixel that changes in time, c) five stationary pixels, d) five pixels that vary in time

Fig. 6. Object position in frames 10, 20, 30, 40, 50

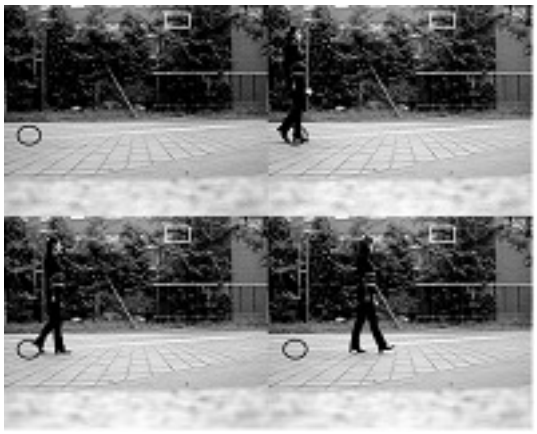


Fig. 7. Stationary pixel (white rectangle) and pixel that changes due to the object motion (black ellipse) in frame 1, frame 23, frame 33 and frame 50

Note that the number of pixels that remain unaltered, i.e. the number of watermarked pixels, will depend on the number of frames K. Thus, K should be chosen to provide a sufficient number of watermarked pixels for watermark detection. The experimental results show that the minimal number of pixels that still provide a satisfactory detection within one frame/image is approximately 600 (with P_{err}~10^{-6}). This number is obtained by performing tests with various images/frames. Also, it is experimentally obtained that even within a sequence of K=50 frames (2s of video signal, PAL system), the number of unaltered

TABLE VI
MEASURES OF DETECTION EFFICIENCY AND PROBABILITIES OF ERROR

	Uncompressed video		Compressed video (MPEG4)	
	R	**Perr**	**R**	**Perr**
Sequence 1	8.5	$9.5{\cdot}10^{-18}$	3.4	$3.3{\cdot}10^{-4}$
Sequence 2	8.9	$2.8{\cdot}10^{-19}$	3.6	$1.5{\cdot}10^{-4}$
Sequence 3	9.1	$4.5{\cdot}10^{-20}$	3.75	$8.8{\cdot}10^{-5}$

pixels from busy frame regions is higher than 2500.

The watermark robustness is tested in the presence of MPEG4 compression. The watermarking procedure is implemented in Matlab 7. The video sequence in AVI format is converted to MPEG 4 (1500kbps) by using the OJOsoft Total Video Converter. The results for three test sequences are given in Table VI.

The proposed idea for video watermarking could be interesting and efficient in video surveillance applications, especially when fixed camera is used.

Finally, it is important to mention that simple and efficient hardware solutions for real-time implementations of 1-D and 2-D S-method (for time-frequency and space/spatial-frequency analysis) have already been done [23], [24].

VI. CONCLUSION

A generalized watermarking approach for multidimensional data protection has been proposed. The multidimensional time-frequency analysis is employed in different stages of the procedure: selection of signal regions, watermark modeling, watermark embedding, as well as in watermark detection. The theoretical considerations are elaborated for 2-D case – space/spatial frequency based image watermarking. The middle frequency components of 2-D S-method are used to define a criterion that selects pixels for watermarking. Additionally, the imperceptibility and robustness is improved by modeling the watermark characteristics according to these components. The successful watermark detection is provided in the space/spatial-frequency domain. The proposed approach provides low probabilities of error for relatively high PSNR and for a small number of watermarked pixels. The procedure exhibits a significant robustness to various attacks which is verified in the experiments.

ACKNOWLEDGEMENT

The authors are very thankful to the anonymous reviewers for their comments that helped to improve the paper. We also thank to Prof. Cleo Baras and Prof. Zdravko Uskoković for very helpful discussions and suggestions during the work on this paper.

REFERENCES

[1] I. J. Cox, M. L. Miller, J. A. Bloom, *Digital Watermarking,* Academic Press, 2002.

[2] M. Barni, F. Bartolini, *Watermarking Systems Engineering.* New York: Marcel Dekker, Inc. 2004.

[3] Proceedings of the IEEE: Special Issue on Identification and Protection of Multimedia Information, vol. 87, July 1999.

[4] E. Muharemagić, B. Furht, "Survey of Watermarking Techniques and Applications," *CH. 3 in Multimedia Watermarking Techniques and Applications*, B. Furht and D. Kirovski, editor, Auerbach Publication, pp. 91-130, 2006.

[5] A. Nikolaidis and I. Pitas, "Asymptotically optimal detection for additive watermarking in the DCT and DWT domains," *IEEE Transaction on Image Processing*, vol. 12, no. 5, pp. 563-571, 2003.

[6] J. R. Hernandez, M. Amado, and F. Perez Gonzales, "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE Transaction on Image Processing*, vol.9, pp. 55-68, Jan. 2000.

[7] A. Briassouli and M. G. Strintzis, "Locally Optimum Nonlinearities for DCT Watermark Detection," *IEEE Transaction on Image Processing*, vol. 13, No. 12, pp. 1604-1618, Dec. 2004.

[8] Q. Cheng and T. S. Huang, "An Additive Approach to Transform-Domain Information Hiding and Optimum Detection Structure," *IEEE Transaction on Multimedia*, vol. 3, No. 3, pp. 273-284, Sept 2001.

[9] S. Stanković, I. Orović, N. Žarić, "Robust speech watermarking in the time-frequency domain," *EURASIP Journal on Advances in Signal Processing*, Vol. 2008, Issue ID 519206, May 2008.

[10] N. Žarić, I. Orović, S. Stanković, C. Ioana, "Space/Spatial-Frequency Based Image Watermarking," *in Proc. 50^{th} International Symposium* ELMAR 2008., pp.101-104.

[11] S. Stanković, I. Djurović, I. Pitas, "Watermarking in the space/spatial-frequency domain using 2-D Radon-Wigner distribution," *IEEE Transaction on Image Processing*, vol. 10, pp. 650-658, Apr. 2001.

[12] B. G. Mobaseri, "Digital watermarking in the joint time-frequency domain," *IEEE International Conference on Image Processing*, vol. 3, pp. 481-484, Sept. 2002, New York

[13] M. Al-khassaweneh and S. Aviyente, "A time-frequency inspired robust image watermarking," *in IEEE Conference Record of the Thirty-Eighth Asilomar Conference*, 2004, vol. 1, pp. 392-396.

[14] I. Djurović, S. Stanković, I. Pitas, "Digital Watermarking in the Fractional Fourier Transformation Domain," Journal of Network and Computer Applications, Academic Press, vol. 24, No. 2, Apr. 2001.

[15] M. Mitrea, F. Preteux, A. Vlad, "Spread spectrum Colour Video Watermarking in the DCT domain," Journal of Optoelectronics and Advanced Materials, vol. 7, No. 2, pp. 1065-1071, Apr. 2005.

[16] S. Thiemert, T. Vogel, J. Dittmann, M. Steinebach, "A Hogh-Capacity Block Based Video Watermarking," *Proc. of* 30^{th} *EUROMICRO Conference* 2004, pp. 457-460, Sept. 2004.

[17] LJ. Stanković, S. Stanković, I. Djurović, "Space/Spatial-Frequency Based Filtering," IEEE *Transaction on Signal Processing*, vol. 48, no. 8, Aug. 2000.

[18] B. Boashash, "Time-Frequency Signal Analysis and Processing," *Elsevier*, 2003.

[19] LJ. Stanković, "A method for Time-Frequency Signal Analysis," IEEE *Trans. on Signal Processing*, vol. 42, No. 1, Jan. 1994.

[20] S. Stanković, LJ. Stanković, Z. Uskoković, "On the local Frequency, Group Shift and Cross Terms in Some Multidimensional Time-Frequency Distributions; A Method for Multidimensional Time-Frequency Analysis," IEEE *Transactions on Signal Processing,* vol. 43, no. 7, July 1995.

[21] D. Heeger, "Signal Detection Theory," Dept. Psych., Stanford Univ., Stanford, CA, Teaching Handout, Nov. 12, 1997.

[22] T. D. Wickens, *Elementary Signal Detection Theory.* Oxford, U.K., Oxford Univ. Press, 2002.

[23] S. Stanković, LJ. Stanković, V. Ivanović, R. Stojanović, "An Architecture for the VLSI Design of Systems for Time-Frequency Analysis and Time-Varying Filtering," *Annales des Telecommunications*, vol. 57, No.9/10, pp. 974-995, Sept./Oct. 2002.

[24] S. Stanković, I. Djurović, V. Vuković, "System Architecture for Space-Frequency Image Analysis," *Electronic Letters*, vol. 34, No. 23, Nov. 1998.