A TIME-FREQUENCY BASED PERCEPTUAL AND ROBUST WATERMARKING SCHEME

Mahmood Al-khassaweneh and Selin Aviyente

Department of Electrical and Computer Engineering,
Michigan State University, East Lansing, MI 48824
e-mail:{alkhassa, aviyente}@egr.msu.edu

ABSTRACT

In the past decade, watermarking has become an important tool in copyright protection applications. In this paper, we introduce a new watermarking scheme in the joint time-frequency domain. Wigner distribution is used to transform an image into the spatial-spectral domain. The proposed method selects the time-frequency cells to be watermarked based on the particular image's energy distribution in the joint domain. This approach ensures the imperceptibility of the embedded watermark. It is shown that embedding in the time-frequency domain is equivalent to a nonlinear embedding function in the spatial domain. A corresponding watermark detection algorithm is also introduced. The performance of the proposed watermarking algorithm under possible attacks, such as noise, re-sampling, rotation, filtering, and JPEG compression is illustrated.

1. INTRODUCTION

With the advance of internet and multimedia tools, watermarking became a powerful tool for data hiding and copyright protection. The past research on image watermarking focused on the spatial and the spectral domains. In the spatial domain, the regions to be watermarked are selected based on the texture of the given image [1, 2]. While in the spectral domain, the watermark is embedded in the transform domain using methods such as DCT and DWT, in the mid-frequency range to ensure transparency and robustness of the watermark, simultaneously [3, 4, 5]. The transform domain algorithms have more advantages in terms of robustness, but it is sometimes difficult to satisfy imperceptibility constraints in the spatial domain simultaneously with the spectral domain constraints. In order to take full advantage of both the spatial and the transform domains, researchers started looking at the joint spatial-spectral representation of the image, which gives a more comprehensive representation of the image [6, 7]. The joint spatial-spectral domain provides flexibility in how much data can be hidden and where it should be hidden inside an image.

In this paper, a new time-frequency based watermarking method will be introduced using Wigner distribution. A complete mathematical derivation for both embedding and detection stages will be given. An equivalent simplification of this Wigner domain method in the time domain will also be shown.

2. BACKGROUND

Time-frequency distributions are bilinear transforms of the signal that represent the energy distribution over both time and frequency. In this paper, Wigner distribution, which is a well-known member of the Cohen's class of distributions, is used for watermarking. For a 1-D discrete signal, $s(n)$, the discrete Wigner distribution is given by [8],

$$WD(n,\omega) = 2\sum_{m} s(n+m)s^*(n-m)e^{-j2\omega m}. \quad (1)$$

where n and $\omega = 2\pi k/N$ are the time and the frequency variables respectively.

The resulting Wigner distribution has some important properties that make it a good choice in watermarking applications; it is always real, satisfies the marginals, invertible and symmetric. These properties make Wigner distribution a good choice in watermarking applications. Invertibility is especially important in watermarking applications where detecting the watermark is one of the primary goals. In the case of images, where we have positive and real pixel values, the Wigner distribution is symmetric and the original signal, $s(n)$, can be retrieved completely from its Wigner distribution as,

$$s(n) = \sqrt{\sum_{\omega} WD(n,\omega)}. \quad (2)$$

Equation (2) implies that for a positive real valued signal, the original signal can be retrieved from its Wigner distribution by taking the inverse Fourier transform of the Wigner distribution evaluated at $m = 0$ and taking the square root of each element in this row. This result will simplify the embedding and the detection of the watermark.

3. WATERMARK EMBEDDING

Wigner distribution for 2-D signals such as images are defined in [8]. However, due to the computational complexity and the difficulty of analyzing the resulting representation, in this paper we transform each row of the image to the time-frequency domain individually using (1).

The watermark embedding algorithm can be summarized as follows:

1. Transform each row of the image, $I(x,y)$, to the Wigner domain:

$$WD_x(y,\omega_y) = 2\sum_{m} I(x,y+m)I(x,y-m)e^{-j2\omega_y m}. \quad (3)$$

2. Transform the watermark sequence w to the Wigner domain:

$$WD_w(y,\omega_y) = 2\sum_{m} w(y+m)w^*(y-m)e^{-j2\omega_y m}. \quad (4)$$

3. Embed the Wigner distribution of the watermark sequence into the Wigner distribution of each row:

$$\hat{WD}_x(y,\omega_y) = WD_x(y,\omega_y) + A_x(y,\omega_y)WD_w(y,\omega_y), \quad (5)$$

where $A_x(y,\omega_y)$ is the time-frequency dependent weighting matrix for row x.

4. Take the inverse transform to obtain the watermarked image:

$$\hat{I}(x,y) = \sqrt{\sum_{\omega_y} \hat{WD}_x(y,\omega_y)}. \tag{6}$$

The fact that not all time-frequency distributions are valid Wigner distributions puts restrictions on the choice of the weighting matrix $A_x(y,\omega_y)$. This implies that $A_x(y,\omega_y)$ should be symmetric with respect to the center and the additive component in (5) should correspond to the Fourier transform of a local autocorrelation function. These are extremely restrictive conditions and hard to satisfy. Assuming that the distribution in (5) is still a valid Wigner distribution, the inversion in (6) is still true and the embedding algorithm can be simplified as follows,

$$\hat{I}(x,y) = \sqrt{\sum_{\omega_y} \hat{WD}_x(y,\omega_y)}, \tag{7}$$

$$= \sqrt{\sum_{\omega_y} (WD_x(y,\omega_y) + A_x(y,\omega_y) WD_w(y,\omega_y))}, \tag{8}$$

$$= \sqrt{2\sum_m I(x,n+m)I(x,n-m)\delta(2m) + \sum_{\omega_y} A_x(y,\omega_y) * w^2(y)}, \tag{9}$$

$$= \sqrt{I^2(x,y) + \left(\sum_{\omega_y} A_x(y,\omega_y)\right) * w^2(y)}, \tag{10}$$

$$\hat{I}(x,y) = \sqrt{I^2(x,y) + \left(\sum_{\omega_y} A_x(y,\omega_y)\right) * w^2(y)}, \tag{11}$$

where $*$ corresponds to convolution. The simplification in (11) reduces $\hat{I}(x,y)$ to a nonlinear function of the image and the watermark sequence in the spatial domain. The time-frequency dependence of the embedding function is through the time-frequency dependent weighting matrix $A_x(y,\omega_y)$. Although the result in (11) is only true when $\hat{WD}_x(y,\omega_y)$ is a valid Wigner distribution, we will use this result in embedding the watermark because of the ease of implementation and its proximity to time-frequency embedding. The time-frequency dependence of the watermarking algorithm is ensured through the time-frequency dependent weighting matrix $A_x(y,\omega_y)$. The weighting matrix $A_x(y,\omega_y)$ should be determined based on the Wigner distribution of the corresponding row. In this paper, $A_x(y,\omega_y)$ is chosen such that the watermark is embedded in the mid-frequency range, which ensures a robust and perceptual watermark,

$$A_x \propto \begin{cases} \frac{WD_x(y,\omega_y)}{max(WD_x(y,\omega_y))}, & \omega_1 \le |\omega_y| \le \omega_2 \\ 0, & elsewhere \end{cases}, \tag{12}$$

where ω_1 and ω_2 determine the range of frequencies where a watermark is embedded with typical values of $\omega_1 = \frac{1}{6}$ and $\omega_2 = \frac{1}{3}$.

4. WATERMARK DETECTION

For copyright protection applications, it is important to detect the existence of the watermark even after the watermarked image is attacked. In order to study the performance of the detector, a threshold is derived such that the probability of error is minimized. Let us define two hypotheses: H_1, where the embedded watermark exists and H_0, where the embedded watermark does not exist. In this paper, we assume that we have access to the original image. Therefore, we can extract a function that depends on the watermark by squaring (11) and subtracting the square of the image from it. The extracted function is compared with a series of possible watermarks to determine which watermark has been embedded. The corresponding correlation based detector obtained from (11) is,

$$\left\langle A_x(y) * w^2(y), \hat{w}^2(y) \right\rangle \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{13}$$

where $A_x(y) = \sum_{\omega_y} A_x(y,\omega_y)$.

Since convolution in time corresponds to multiplication in frequency, we can rewrite (13) as,

$$\langle C(n)Y_1(n), Y_2(n) \rangle \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{14}$$

where $C(n)$, $Y_1(n)$ and $Y_2(n)$ correspond to the Fourier transforms of $A_x(y)$, $w^2(y)$ and $\hat{w}^2(y)$, respectively. By defining the probability of false alarm as P_{FA} and the probability of detection as P_D, the probability of error P_e is written as,

$$P_e = p_0 P_{FA} + p_1(1 - P_D). \tag{15}$$

For the case that the priori probabilities of H_0 and H_1 are $\frac{1}{2}$,

$$P_e = \frac{1}{2} P\left(\sum_n C(n)Y_1(n)Y_2(n) > \eta\right) + \frac{1}{2} P\left(\sum_n C(n)Y_1^2(n) < \eta\right). \tag{16}$$

In order to derive the minimum probability of error detector, define,

$$z_1 = \sum_n C(n)Y_1^2(n). \tag{17}$$

The mean and the variance of z_1 are given by,

$$\mu_{z_1} = N\sigma_1^4 \left[2\sum_n C(n) + NC(0)\right], \tag{18}$$

$$\sigma_{z_1}^2 = 8N^2\sigma_1^8 \left[\sum_n C^2(n) + NC^2(0)\right]. \tag{19}$$

Let,

$$z_2 = \sum_n C(n)Y_1(n)Y_2(n). \tag{20}$$

The mean and the variance of this random variable are given by,

$$\mu_{z_2} = N^2\sigma_1^2\sigma_2^2 C(0), \tag{21}$$

$$\sigma_{z_2}^2 = 4N^2\sigma_1^4\sigma_2^4 \left[\sum_n C^2(n) + NC^2(0)\right]. \tag{22}$$

Using the central limit theorem, the pdfs of z_1 and z_2 are assumed to be Gaussians,

$$f_{z_1}(z) \sim N(\mu_{z_1}, \sigma_{z_1}). \tag{23}$$

$$f_{z_2}(z) \sim N(\mu_{z_2}, \sigma_{z_2}). \tag{24}$$

In order to find the minimum probability of error detector, we differentiate P_e with respect to η,

$$\frac{\partial P_e}{\partial \eta} = 0, \tag{25}$$

which yields,

$$f_{z_1}(\eta) - f_{z_2}(\eta) = 0. \tag{26}$$

Since $C(n)$ is the Fourier transform of $A_x(y)$, the following relationships hold,

$$\begin{aligned} C(0) &= \sum_y A_x(y), \\ \sum_n C(n) &= N A_x(0), \\ \sum_n C^2(n) &= N \sum_y A_x^2(y). \end{aligned} \tag{27}$$

Using the fact that $A_x(y, \omega_y)$ has values less than one according to equation (12) to insure imperceptibility,

$$\left(\sum_y A_x(y) \right)^2 >> \sum_y A_x^2(y). \tag{28}$$

Solving (26) for η using the above facts and the assumption in (28), we get,

$$\eta \approx \frac{N^2 \sum_y A_x(y) \sigma_1^4 \left[\left(\frac{2\sigma_1^2}{\sigma_2^2} - 1 \right) \pm \sqrt{2} \left(\frac{\sigma_1^2}{\sigma_2^2} - 1 \right) \right]}{\left(\frac{2\sigma_1^4}{\sigma_2^4} - 1 \right)}. \tag{29}$$

For the special case when $\sigma_1^2 = \sigma_2^2 = \sigma^2$, (29) reduces to,

$$\eta \approx N^2 \sigma^4 \sum_y A_x(y). \tag{30}$$

The threshold derived in (29) shows that the threshold is image dependent. This dependency on the image is reflected through the time-frequency weighting matrix, $A_x(y, \omega_y)$. Therefore, the image's spatial and spectral distribution is taken into account when choosing the appropriate threshold. It is also important to note that this threshold is different for each row of the image and an average threshold can be evaluated for the whole image. The special case threshold given in (30) shows that the embedding method depends on N^2 which makes it more dependent on the image size.

5. SIMULATION RESULTS

The watermark embedding algorithm proposed in this paper has been applied to a number of images (Lena, barbara, camera man). The watermark is a normally distributed random sequence. The robustness of the watermark embedding algorithm has been studied experimentally. For the case of a 256×256 lena image and a watermark of length 256, Fig. 1 shows the difference between the original image and the watermarked one. It is apparent that the embedded watermark is concentrated along the edges of the image, which ensures the imperceptibility of the embedding algorithm.

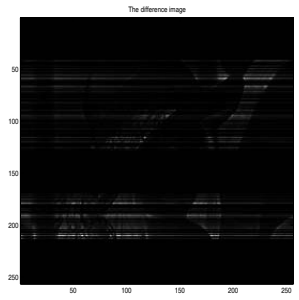

Figure 1: The difference between the original and the watermarked Lena image.

The performance of the proposed algorithm has been tested under different types of attacks including additive white Gaussian noise (AWGN), median filtering, rotation, re-sampling and JPEG compression. The detector for the algorithm described by (13) is tested for a series of watermarks. Fig. 2 shows the normalized correlation function under different attacks. This function reaches its maximum value when the embedded watermark and the tested one are the same and is close to zero for all other watermarks.

Fig. 3 shows the ROC curves for the proposed embedding method under different types of attacks. The proposed algorithm performs very well under most attacks with an average probability of error of 0.03.

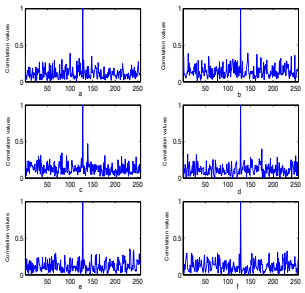

Figure 2: The normalized correlation detector response for the proposed embedding method applied to (256×256) Lena image under different attacks, a. AWGN (variance 20),b. Median filtering (size 4×4), c. Rotation (10°), d. Resampling (22%), e. JPEG compression with Quality factor of 80%, f. JPEG compression with Quality factor of 60%.

The performance of the proposed embedding method for different images has also been studied. We have applied the embedding method to a 256×256 lena1, a 512×512 lena2, a 512×512 barbara and a 256×256 camera man images. Table I summarizes the probability of error in detecting the watermark under different types of attacks for the four test images. The embedding method provides a small probability of error under most of the attacks regardless of the image type. Increasing the image size improves the performance of the embedding algorithm.

The results show that the proposed method performs well under most attacks with a very small error rate. The maximum probability of error occurs under resampling with worst case value of 0.156. Since convolution spreads out the watermark all over the image, the resultant watermark is robust

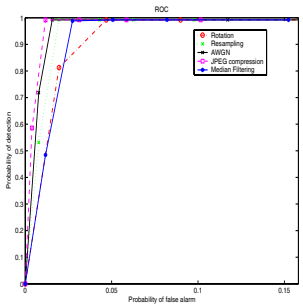

Figure 3: ROC for the watermark detector after different attacks applied to (256×256) Lena image.

Table 1: The probability of error in detecting the watermark under different types of attacks for different images.

Image	Lena1	Lena2	Barbara	Camera
AWGN (Variance=10)	0.0225	0.076	0.0653	0.0429
AWGN (Variance=20)	0.0258	0.086	0.0776	0.0539
Median filtering (4×4)	0.0375	0.0733	0.0687	0.0803
Median filtering (8×8)	0.0282	0.0543	0.0510	0.0691
Median filtering (16×16)	0.0175	0.0472	0.0479	0.0572
Rotation(5°)	0.0536	0.098	0.0724	0.0839
Rotation(10°)	0.0652	0.104	0.0921	0.104
Rotation(15°)	0.0764	0.152	0.123	0.145
Resampling (22%)	0.0306	0.111	0.1095	0.0987
Resampling (32%)	0.0521	0.156	0.153	0.136
JPEG compression (Q=40%)	0.0182	0.118	0.129	0.101
JPEG compression (Q=60%)	0.0163	0.0845	0.102	0.084
JPEG compression (Q=80%)	0.0144	0.0414	0.065	0.0391

to most attacks as it is clear from the simulation results.

6. CONCLUSION

In this paper, a new watermarking algorithm based on the Wigner distribution has been introduced. It has been shown that for positive and real signals, the signal can be retrieved from its Wigner distribution without any error. This realization inspires the use of the time-frequency information for both the image and the watermark. The embedding algorithm in the joint domain is simplified to a nonlinear embedding function in the time domain. This result reduces the computational complexity of embedding and detecting the watermark. For this reason, a time-frequency inspired embedding algorithm in the spatial domain is used. A non-blind correlation based detector is derived using the nonlinear embedding function and the minimum probability of error threshold is found. The proposed algorithm is shown to be transparent and robust under attacks, through experiments.

The weighting matrix $A_x(y, \omega_y)$ plays an important role in this algorithm. One can improve the perceptual quality of the watermarked image by developing human visual models on the time-frequency plane that are similar to the ones developed for DCT [9]. The proposed detection algorithm can also be extended for blind detection assuming that the image and the watermark are independent.

REFERENCES

[1] N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain," *Signal Processing*, vol. 66, pp. 385–403, 1998.

[2] O. Bruyndonckx, J. J. Quisquater, and B. Macq, "Spatial method for copyright labeling of digital images," in *Proc. IEEE Workshop on Nonlinear Signal and Image Processing*, June 1995, pp. 456–459.

[3] M. Barni, F. Bartolini, A. De Rosa, and A. Piva, "Capacity of the watermark channel: How many bits can be hidden within a digital image?," in *Proc. SPIE*, Jan. 1999, vol. 3657, pp. 437–448.

[4] J. J. K. O'Ruanaidh and T. Pun, "Rotation, scale and translation invariant digital image watermarking," *Signal Processing, Special Issue on Copyright Protection and Control*, vol. 66, no. 3, pp. 303–317, May 1998.

[5] S. Pereira, S. Voloshynoskiy, and T. Pun, "Optimal transform domain watermark embedding via linear programming," *Signal Processing*, vol. 81, no. 6, pp. 1251–1260, June 2001.

[6] S. Stankovic, I. Djurovic, and I. Pitas, "Watermarking in the space/spatial-frequency domain using two-dimensional radon-wigner distribution," *IEEE Transactions on Image Processing*, vol. 10, pp. 650–658, Apr. 2001.

[7] B. G. Mobasseri, "Digital watermarking in the joint time-frequency domain," in *IEEE International Conference on Image Processing*, 2002, vol. 3, pp. 481–484.

[8] G. Cristobal, C. Gonzalo, and J. Bescos, "Image filtering and analysis through the Wigner distribution function," in *Advances in Electronics and Electron physics*, W. Hawkes, Ed., pp. 309–397. Academic Press, 1991.

[9] A. B. Watson, "Efficiency of an image code based on human vision," *Journal of Optical Society of America*, pp. 2401–2417, 1987.