ROBUST WATERMARKING IN THE WIGNER DOMAIN

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 this paper, a new watermarking scheme in the joint time-frequency domain is introduced. 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

The past research on image watermarking focused on embedding the watermark in 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]. 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 [2, 3]. 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 [4, 5, 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.

Most work in the time-frequency domain concentrated on using Wigner distribution as the transform domain. In [4], the authors used a two-dimensional chirp signal with a variable spatial frequency as the watermark. The watermark is characterized by a linear frequency change and can be detected by using 2-D space/spatial-frequency distributions. The projections of the 2-D Wigner distribution and the 2-D RadonWigner distribution are used to emphasize the watermark detection process.

In [5], the Wigner distribution is used for watermarking. The watermark is embedded in a subset of the transformed cells selected such that the watermark will survive the JPEG compression. Since the resultant watermarked distribution is not a valid Wigner distribution, a least square error estimate is used in watermark detection. This algorithm detects the presence of the watermark under JPEG attacks.

In [6], a fragile image watermarking method based on time-frequency representation is presented. The watermark is an FM modulated signal which is embedded in the diagonal elements of the image. The particular features of this signal in the time-frequency domain are used to identify the watermark. The Wigner distribution is used to analyze the content of the extracted watermark.

In this paper, a new time-frequency based watermarking method will be introduced using the 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. The robustness for this algorithm under attacks will be tested experimentally.

2. BACKGROUND

For a discrete signal, $s(n)$, the discrete Wigner distribution is given by,

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

where n and $\omega = 2\pi k/N$ are the discrete time and 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 the 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,

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

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

The watermark embedding algorithm can be summarized as follows:

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

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

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:

$$\hat{W}D_x(y, \omega_y) = WD_x(y, \omega_y) + A_x(y, \omega_y) \odot WD_w(y, \omega_y), \quad (5)$$

where $A_x(y, \omega_y)$ is the time-frequency dependent weighting matrix for row x, and $A_x(y, \omega_y) \odot WD_w(y, \omega_y)$ is an element by element multiplication of the two matrices.

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

$$\hat{I}(x, y) = \sqrt{\sum_{\omega_y} \hat{W}D_x(y, \omega_y)}. \quad (6)$$

One important issue is the inversion of $\hat{W}D_x(y, \omega_y)$. It is well-known that not all time-frequency surfaces are valid Wigner distributions. Therefore, some restrictions apply when choosing 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. Assuming that the distribution in (5) is still a valid Wigner distribution, as shown in section 4, the inversion in (6) is still true and the embedding algorithm can be simplified as follows,

$$\hat{I}(x, y) = \sqrt{\sum_{\omega_y} \hat{W}D_x(y, \omega_y)}, \quad (7)$$

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

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

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

where $*$ corresponds to convolution. The simplification in (10) 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 (10) is only true when $\hat{W}D_x(y, \omega_y)$ 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 as discussed in section 4. 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,

$$A_x \propto \begin{cases} \frac{WD_x(y, \omega_y)}{max(WD_x(y, \omega_y))}, & \omega_1 \leq |\omega_y| \leq \omega_2 \\ 0, & elsewhere \end{cases}, \quad (11)$$

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. ERROR INTRODUCED IN WIGNER DISTRIBUTION INVERSION

In the previous section, we assumed that the watermarked Wigner distribution in equation (5) is a valid Wigner distribution. However, this is hard to be satisfied in all cases, and an error is introduced in (5). Our goal is to study the effect of this approximation by looking at how different the Wigner distribution of the signal in equation (6) and the Wigner distribution in equation (5). Let the Wigner distribution of $\hat{I}(x, y)$ be $\overline{WD}_x$. Ideally, $\overline{WD}_x$ and $\hat{W}D_x$ should be identical. However, an error is introduced by equation (6) in the inversion process. We compute the Normalized Mean Square Error (NMSE) between $\overline{WD}_x$ and $\hat{W}D_x$. Table 1. shows the average NMSE for different images over all time-frequency points for all rows. The NMSE is computed from the error introduced in the inversion of the Wigner distribution for different images. The results in Table 1 suggest that the approximation used for the inversion of the Wigner distribution is valid and introduces a small amount of error.

Table 1: The average Normalized Mean Square Error introduced by the approximation of the Wigner distribution.

Image	NMSE	Standard deviation (sd)
Lena	4.6×10^{-11}	2.1×10^{-12}
Barbara	4.5×10^{-11}	2.3×10^{-12}
Camera Man	4.3×10^{-11}	2.2×10^{-12}
Peppers	4.4×10^{-11}	2.3×10^{-12}

We can further study the locations where the error is concentrated by finding the difference between the two Wigner distributions,

$$WD_D = \overline{WD}_x(y, \omega_y) - \hat{W}D_x(y, \omega_y). \quad (12)$$

At each time point, i.e every column in WD_D, we find the histogram of the maximum differences over frequency. Fig. 1 shows that the maximum error is concentrated around the low frequencies. Therefore, in our embedding algorithm we choose the weighting matrix such that the watermark is embedded in the middle frequency range, which is less effected by this approximation error.

5. 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
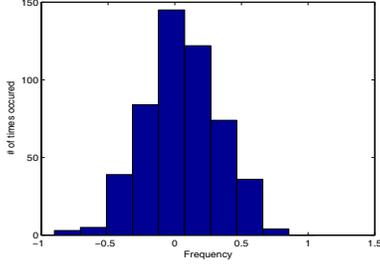
Figure 1: The average histogram for the difference of the two Wigner distributions.

a function that depends on the watermark by squaring (10) and subtracting the square of the original 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 (10) is,

$$\langle A_x(y) * w^2(y), \hat{w}^2(y) \rangle \underset{H_0}{\overset{H_1}{\underset{<}{>}}} \eta, \quad (13)$$

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

Since convolution in time corresponds to multiplication in frequency, we can equivalently rewrite this correlation detector as,

$$\langle C(n)Y_1(n), Y_2(n) \rangle \underset{H_0}{\overset{H_1}{\underset{<}{>}}} \eta, \quad (14)$$

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. For the case that the a priori probabilities of H_0 and H_1 are $\frac{1}{2}$, the probability of error P_e is,

$$P_e = \frac{1}{2}P\left(\sum_n C(n)Y_1(n)Y_2(n) > \eta\right) + \frac{1}{2}P\left(\sum_n C(n)Y_1^2(n) < \eta\right). \quad (15)$$

In order to derive the minimum probability of error detector, we need to find the distributions of $z_1 = \sum_n C(n)Y_1^2(n)$ and $z_2 = \sum_n C(n)Y_1(n)Y_2(n)$. Using the central limit theorem, the pdfs of z_1 and z_2 are assumed to be Gaussian with means and variances defined as,

$$\mu_{z_1} = N\sigma_1^4\left[2\sum_n C(n) + NC(0)\right], \quad (16)$$

$$\sigma_{z_1}^2 = 8N^2\sigma_1^8\left[\sum_n C^2(n) + NC^2(0)\right], \quad (17)$$

$$\mu_{z_2} = N^2\sigma_1^2\sigma_2^2 C(0), \quad (18)$$

$$\sigma_{z_2}^2 = 4N^2\sigma_1^4\sigma_2^4\left[\sum_n C^2(n) + NC^2(0)\right]. \quad (19)$$

The minimum probability of error detector is found by setting $\frac{\partial P_e}{\partial \eta} = 0$. Using the fact that the weighting matrix, $A_x(y, \omega_y)$ has values less than one based on equation (11) to

insure imperceptibility, the threshold η can be approximated by,

$$\eta \approx \frac{N^2 \sum_y A_x(y)\sigma_1^4 \left[\left(\frac{2\sigma_1^2}{\sigma_2^2} - 1\right) \pm \sqrt{2}\left(\frac{\sigma_1^2}{\sigma_2^2} - 1\right)\right]}{\left(\frac{2\sigma_1^4}{\sigma_2^4} - 1\right)}. \quad (20)$$

For the special case when $\sigma_1^2 = \sigma_2^2 = \sigma^2$, equation (20) reduces to,

$$\eta \approx N^2\sigma^4 \sum_y A_x(y). \quad (21)$$

The threshold derived in (20) 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 equation (21) shows that the embedding method depends on N^2 which makes it more dependent on the image size.

6. SIMULATION RESULTS

The watermark embedding algorithm proposed in this paper has been applied to a number of images. The watermark is a normally distributed random sequence. The robustness of the watermark embedding algorithm has been studied experimentally. For the case of the 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.
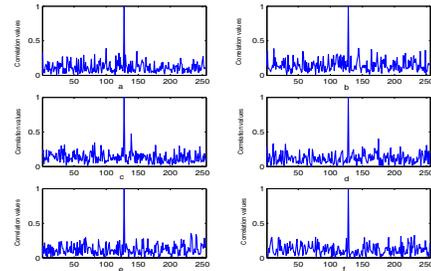

Figure 2: The normalized correlation detector response for the proposed embedding method applied to (256×256) Lena image under different attacks, a. AWGN (variance 20),b. Median filtering (size 4×4), c. Rotation $(10°)$, d. Re-sampling (22%), e. JPEG compression with Quality factor of 80%, f. JPEG compression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

Table 2: 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
Rotation(5°)	0.0536	0.098	0.0724	0.0839
Rotation(10°)	0.0652	0.104	0.0921	0.104
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

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 over the whole image, the resultant watermark is robust to most attacks as it is clear from the simulation results.

7. CONCLUSION

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 proposed 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 spatial domain. The error introduced by this simplification is analyzed and is shown to be concentrated at low frequencies. Therefore, the time- frequency embedding algorithm chooses a weight function that emphasizes the mid-frequencies. This simplification reduces the computational complexity of embedding and detecting the watermark. 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.

REFERENCES

[1] D. P. Mukherjee, S. Maitra, and S. T. Acton, "Spatial domain digital watermarking of multimedia objects for buyer authentication," *IEEE Transactions on Multimedia*, vol. 6, no. 1, pp. 1–15, Feb. 2004.

[2] 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.

[3] Y. Wang, J. Doherty, and R. Van Dyck, "A Wavelet-Based Watermarking Algorithm for Ownership Verification of Digital Images," *IEEE Transactions on Image Processing*, vol. 11, pp. 77–88, Feb. 2002.

[4] 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.

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

[6] B. Barkat and F. Sattar, "A new time-frequency based private fragile watermarking scheme for image authentication," in *IEEE International Symposium on Signal Processing and Applications*, 2003, vol. 2, pp. 363–366.

[7] S. Erkucuk, S. Krishnan, and M. Zeytinoglu, "Robust audio watermarking using a chirp based technique," in *IEEE International Conference on Multimedia and Expo*, July 2003, vol. 2, pp. 513–516.