Multimedia and Security Workshop at ACM Multimedia 99, Orlando, FL, USA, Oct 1999

Public Watermarking Surviving General Scaling and Cropping: An Application for Print-and-Scan Process

Ching-Yung Lin

Department of Electrical Engineering
Columbia University
500 W120th St. #1312
New York, NY 10027, USA
(212) 854-4393
cylin@ctr.columbia.edu

ABSTRACT

Scaling and cropping are very common in today's image processing software. When an image is printed-and-scanned, the final image is generally a cropped version of the rotated, scaled original image, with additional noises. The cropped image usually does not have the same aspect ratio as the original, so the DFT coefficients of the cropped image and the original would be quite different. In this paper, we propose an algorithm embedding spread spectrum watermarks in the DFT magnitudes of the log-log map magnitudes in the Fourier domain. These watermarks would be resistant to any aspect ratio of scaling and cropping and pixel value distortions as in the print-and-scan process.

KEYWORDS: Public Watermarking, Rotation, Scaling, Cropping, DFT, Scanning

1 Introduction

Watermarking methods embed information in a multimedia object in which the modification should be imperceptible. The embedded information can be used as a proof of ownership or as a kind of secret data transmission. Several types of watermarking systems have been proposed since 1990 [1]. Among them, public watermarking is considered to have a broader application value, because it can detect the watermarks without the original object. There are two types of public watermarking systems: one-bit watermark and multiple-bit watermark [2]. The one-bit public watermarking system (also referred to as *semi-private watermarking*) detects the existence of a specific identification watermark in the multimedia content. It usually serves as evidence of ownership. The multiple-bit public watermarking system (or *blind watermarking*) extracts the embedded information of the watermark. It is usually used for data hiding or ownership declaration.

Today, the print-and-scan process is commonly used for image reproduction and distribution. It is popular to transform images between the electronic digital format and the printed pictures. Therefore, for copyright protection, an effective watermarking system should be able to detect or extract watermarks, regardless the media format of the image. However, while most of previous research on watermarking focused on the electronic digital image, it was not clear how the print-and-scan process affects the image, nor how a watermarking system can survive it.

The rescanned image is generally distorted in both the pixel values and the geometric boundary. The distortion of pixel values is caused by (1) the luminance, contrast, gamma correction and chromnance variations, and (2) the blurring of adjacent pixels. These distortions are the typical effects of the printer and scanner.

The distortion of the geometric boundary in the print-and-scan process is caused by rotation, scaling, and cropping (RSC). We would like to point out that the geometric distortion in the scanning process could not be adequately modeled by the well-known rotation, scaling, and translation (RST) effects, because of the design of today's Graphic User Interface (GUI) for the scanning process as in Fig 1. The RST is usually used to model the geometric distortion on the image of an observed object. It has been widely used in pattern recognition. In those cases, the capturing window of camera is usually predetermined, *i.e.* the size of the captured image is usually determined by the system. However, in the scanning process, the scanned image may only cover a part of the original picture and may have an arbitrary cropped image size. Cropping introduces large changes to the image. Detailed discussion of the modeling of the Print-and-Scan process can be found in [3].

Some watermarking methods have been proposed to solve the related problems. O'Ruanaidh and Pun [4] first suggested a watermarking method based on the log-polar map of the Fourier coefficients (also known as the Fourier-Mellin Transform). They proposed that the Discrete Fourier Transform (DFT) magnitudes of the Fourier-Mellin coefficients can be used to embed the watermark, because their well-known shifting property to RST distortions. However, the Fourier-Mellin transform can only deal with uniform scaling (*i.e.*, the same scaling factor in both horizontal and vertical direction), as well as cropping which keeps the original aspect ratio. Therefore, if the image is cropped with arbitrary aspect ratio as in the print-and-scan process, the Fourier-Mellin-based methods would become invalid. A detailed discussion of the Fourier-Mellin-based watermarking method can be found in [5], where we proposed an RST resilient watermarking method and solved many implementation difficulties. Another method proposed by Pereira *et.al.* [6] embeds a registration pattern as well as a watermark into an image. This is an effective solution but can have the problems of reducing the fidelity and tampering the watermark.

Most previous work presented theoretical solutions based on the continuous Fourier Transform without addressing the practical implementation constraints in the DFT domain. Although DFT coefficients are the sampling values of the continuous Fourier coefficients, their properties are not simply the same, because the sampling rate, the aliasing effect, and the discontinuity on the boundary of periods affect DFT coefficients. The DFT coefficients are, in fact, sampled from a repeated discrete image. Their sampling positions in the continuous Fourier domain are determined by the repetition period, which is, in many cases, the size of image or the smallest radix-2 size larger than the size of image. It may be noticed that, without special considerations, DFT-based robust watermarking methods can automatically survive scaling with any aspect ratio and cropping with a fixed aspect ratio of the original image. This phenomenon comes from the fact that the DFT coefficients are at the same sampling positions in the continuous Fourier domain after these manipulations [3], if the sizes of DFT points are always the sizes of image (*e.g.*, using 256x256 DFT for 256x256 images, and 128x128 DFT for their down-sampled 128x128 versions). However, if the image is cropped with an arbitrary aspect ratio, its DFT coefficients will be quite different because they are sampled at different positions in the Fourier domain. The changes of the DFT coefficients on the cropped image are similar to the changes of the continuous Fourier coefficients of the scaled and translated original image. We will discuss them further in Section 2.

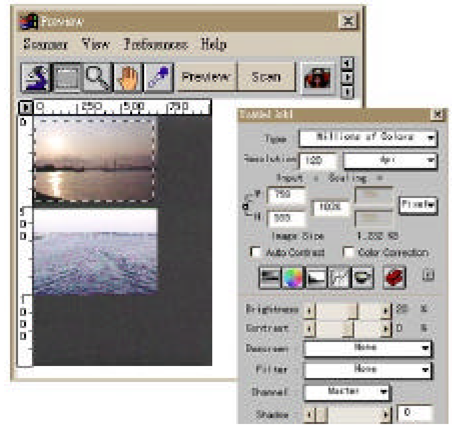

Figure 1: The control windows of scanning process. The scanned image only includes the cropped area.

Acknowledging the RSC effects during the print-and-scan processes, in this paper, we propose an algorithm embedding spread spectrum watermarks in the DFT magnitude of the log-log map coefficients on the DFT domain. We will show in Section 2 that scaling and cropping an image with arbitrary aspect ratio results in a simple two-dimensional translation in the log-log map of DFT coefficient magnitudes. Therefore, the DFT magnitudes of this map will be invariant after scaling and cropping. Because, so far, we could not find a reliable transformation which simultaneously provides applicable properties to RSC, a drawback of the proposed system is that it is not rotation invariant and one would have to test the scanned image several times within the possible range of rotation. In practice, this may not be a serious problem because the rotation angle of scanned image would not be too much, while the range of cropping and scaling are usually unbounded. The proposed algorithm can also be applied to the images edited by general image software, in which scaling and cropping are more common than rotation. We will explain the watermarking method in detail in Section 3, and show some preliminary experimental results in Section 4.

2 Modeling of the Print-and-Scan Process

When a user scans a picture, at the first step, he/she has to place the picture on the flatbed of the scanner. This may introduce a small orientation, if the picture is not well placed. Then, the scanner scans the whole flatbed to get a previewed low-resolution image. After this process, the user subjectively selects a cropping window to decide an appropriate range of the picture. Then, the scanner scans the picture again with a higher resolution to get a scanned image. The scanned image is usually a different size because the resolution in the scanner and the printer are generally different. Usually, it includes only a part of the original image. In our tests, the image is not generally rotated because users usually place the picture or document along the corner of the flatbed. Even if the picture is not well placed, the rotation angle is generally within a small degree, *e.g.*, ±3 degrees.

Assume we have a continuous finite support image,

$$x(t_1,t_2) = \begin{cases} x_0(t_1,t_2), & t_1 \in [-\frac{T_1}{2},\frac{T_1}{2}], t_2 \in [-\frac{T_2}{2},\frac{T_2}{2}] \\ 0, & elsewhere \end{cases} \quad (1)$$

If this image is scaled by l_1 in the t_1-axis and l_2 in the t_2-axis, then

$$x_S(t_1,t_2) = x(\tfrac{t_1}{l_1},\tfrac{t_2}{l_2}) \xleftrightarrow{F} X(l_1 f_1, l_2 f_2) = X_S(f_1,f_2) \quad (2)$$

From Eq. (2), if we assume a transformation, *LL*, which maps the original Cartesian coordinate points to their log-log coordinate points [6], *s.t.*,

$$(LL \circ X)(f_1', f_2') = X(e^{f_1'}, e^{f_2'}) \quad (3)$$

then we can get,

$$(LL \circ X_S)(f_1', f_2') = (LL \circ X)(f_1' + \log l_1, f_2' + \log l_2) \quad (4)$$

If the image is translated, then we should change the X_S and X to their magnitudes $|X_S|$ and $|X|$ in Eq. (4).

For cropping, we can consider the cropped image, x_C, as a subtraction of the discarded area, $x_{\bar{C}}$, from the original image, x. Then, this equation,

$$X_C(f_1, f_2) = X(f_1, f_2) - X_{\bar{C}}(f_1, f_2) \quad (5)$$

represents the cropping effect in the continuous Fourier domain. If the discarded area is much smaller than the original image, then the Fourier coefficients of the discarded area, $X_{\bar{C}}$, can be considered as noises in Eq. (5).

Because we only have the discrete images before and after scanning, in practical cases, DFT is usually used as a sampling method on the frequency domain and takes advantage of FFT. The relationships of (continuous) Fourier transform (FT), Fourier Series (FS), and DFT are:

-- The FS coefficients are the samples of the FT coefficients of a finite support continuous signal. They are calculated by repeating the signal in the time domain. Assume the repetition period is *T*. Once the signal becomes periodic in the time domain, its FT coefficients will have non-zero values only in the *n/T* positions. These values are the multiplication of the FS coefficients and a delta function. We should notice that the FS coefficients are always proportional to the FT values of the original non-periodic signal in the n/T positions.

--The DFT coefficients represent the FS coefficients of the discretized original signal. After the original signal is sampled, its FS coefficients would become periodic. The DFT coefficients are the FS coefficients in a period. Smaller sampling frequency in the time domain would introduce an aliasing effect in the frequency domain. That can be considered as additive noises to the DFT coefficients.

From the above descriptions, we know that the repetition period of the original signal decides the sampling positions of DFT coefficients in the frequency domain. In the scaling cases, if the repetition is always the same as the image size, then the FS of the original continuous image, $\tilde{X}$, and the scaled image, $\tilde{X}_S$, should be the same. That is,

$$\begin{aligned} \tilde{X}_S(n_1,n_2) &= X_S(\tfrac{n_1}{T_{S1}},\tfrac{n_2}{T_{S2}}) = X(\tfrac{n_1 l_1}{T_{S1}},\tfrac{n_2 l_2}{T_{S2}}) \\ &= X(\tfrac{n_1}{T_1},\tfrac{n_2}{T_2}) = \tilde{X}(n_1,n_2) \end{aligned} \quad (6)$$

where T_{S1} and T_{S2} are the sizes of the scaled image. Adding the concern of discretization in the spatial domain, we can get the DFT coefficients in the scaled case, $\hat{X}_S$ as

$$\hat{X}_S(n_1,n_2) = \hat{X}(n_1,n_2) + N_{sampling} \quad (7)$$

where $\hat{X}$ is the DFT of original image. In Eq. (7), the sampling noises happen when images are down-sampled.

If an image is cropped, then the changes of DFT coefficients are introduced from three factors: (1) *the change of image size*, (2) *the information loss of the discarded area*, and (3) *the translation of the origin point of the image*. Assume the size of cropped image is $\alpha_1 T_1$ x $\alpha_2 T_2$. If the size of DFT is the same as the size of the cropped image, then we can obtain the DFT coefficients after scaling and cropping,

$$|\hat{X}_{SC}(n_1,n_2)| = |\hat{X}(\tfrac{n_1}{a_1},\tfrac{n_2}{a_2}) + \hat{N}_{SC}(n_1,n_2)| \quad (8)$$

where

$$\hat{N}_{SC}(n_1,n_2) = -\hat{X}_{\bar{C}}(\tfrac{n_1}{a_1},\tfrac{n_2}{a_2}) + N_{sampling} \quad (9)$$

In Eq. (9), if the cropped area include the entire original image, i.e., α_1, α_2 >= 1, then the effect of the discarded area can be ignored. If the cropping ratios are too small, then the power loss in the discarded area may not be just ignored as noises. In our experiments, the reliable minimum thresholds are at about 0.8, which may be small enough for most scanned images [3]. In Eq. (9), strictly speaking, there is no definition in $\hat{X}$ at the non-integer positions. But, since $\hat{X}$ are samples of X, we can set $\hat{X}(\frac{n_1}{a_1}, \frac{n_2}{a_2}) = X(\frac{n_1}{a_1 T_1}, \frac{n_2}{a_2 T_2})$ directly from the original Fourier coefficients. In practical applications, these values are generally obtained from interpolation.

In addition to using the same size DFT of the scaled and cropped image, some cases use the smallest radix-2 FFT that are larger than the image size. In that case, Eq. (8) and (9) are still applicable, but α_1 and α_2 should be replaced by other values. Detailed modeling description of cropping and scaling can be found in [3].

Comparing Eq. (2) and Eq. (8), we can find the changes of the DFT coefficients after scaling & cropping and those of the continuous Fourier coefficients after scaling are similar. Therefore, after scaling & cropping, as in Eq. (3) and Eq. (4), the log-log map of the DFT coefficient magnitudes will suffer simple shift. Then, the DFT magnitudes of the log-log map of DFT magnitude should be similar. We can use this property for watermarking.

3 Algorithm

The embedding algorithm is shown in Figure 2. At the first step, we scale the image to a fixed size of 256x256 pixels. As shown in Eq. (7), scaling with arbitrary aspect ratio does not affect the DFT coefficients, if all images are resized to a small standard size, it can reduce both the computational cost and implementation cost. We only use this standard size image for calculating the amount of additive watermarks on this scale. Then, these additive watermarks are resized to the original size and added to the original image. In this way, the watermarked image will not suffer the fidelity loss of down-sampling. Since the image size is generally larger than this standard size, the additive watermarks are usually scaled up, which will not introduce the sampling noise in Eq. (7).

The second step is to get the log-log map from the DFT magnitudes. We use the bilinear interpolation because it is easier to implement and can get reasonable results. We noticed that the system has to interpolate the DFT magnitudes, instead of interpolating the complex DFT coefficients and then get their magnitudes. Intuitively, the latter method seems to be more correct. However, because the image may be translated by cropping, the phase change of DFT coefficients will result in incorrect interpolated log-log map magnitudes [3].

We then use the spread spectrum method to embed watermarks [4][7]. Two kinds of watermarks are tested

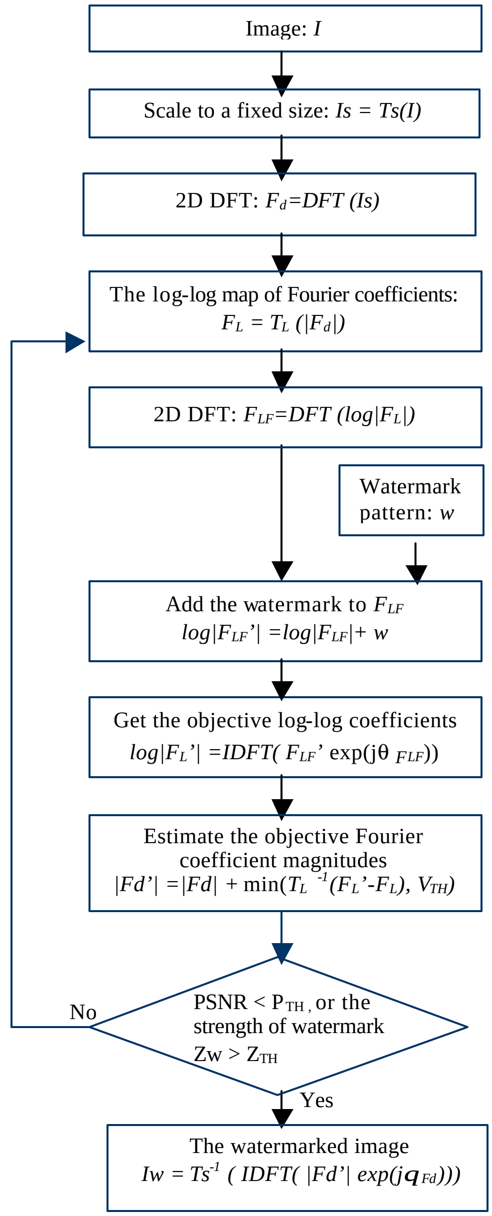

Figure 2: The embedding process

in our system. For the 1-bit watermark, we generate a user identification code and take its convolution with pseudo-noise patterns as an additive watermark on the log values of DFT magnitudes of the log-log map magnitudes. For the multiple-bit watermark, we use the same method as in [4], which uses the shift position of the fixed pseudo-noise pattern to represent the embedded information. The watermarks are only embedded in the mid-band areas, e.g., 33 <= n_1, n_2 <= 128 in the 256x256 DFT, to avoid significant effect on the fidelity and to be robust to other manipulations such as compression. We embed the same pattern in

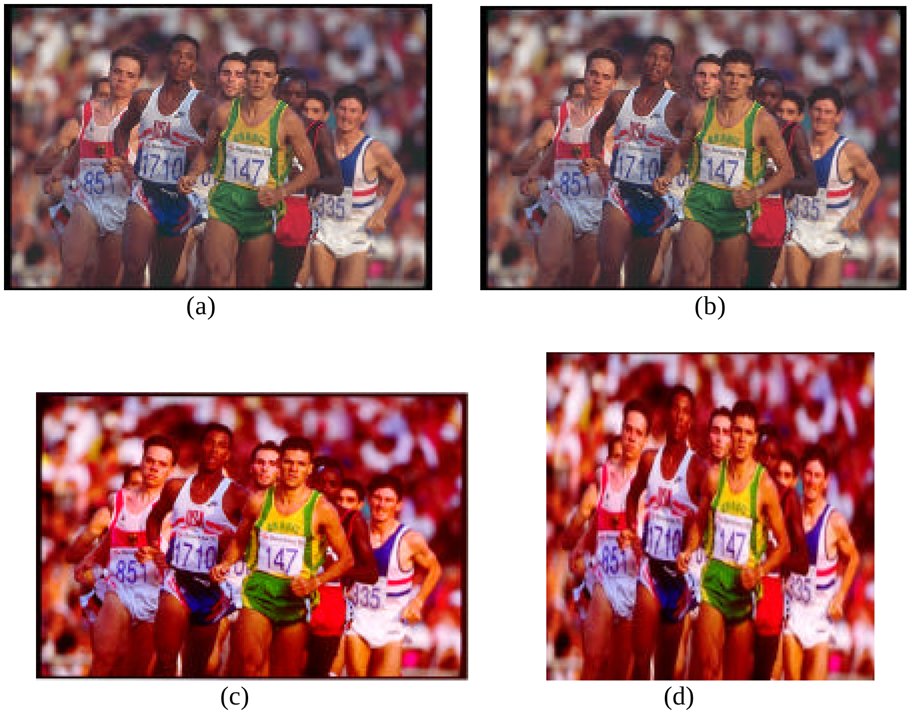

Figure 3: Experimental results: (a) original image [384x256]; (b) watermarked image, PSNR= 45.14dB, Z=16.26; (c) print-and-scanned image [401x268], Z=6.37; (d) after cropping, resizing and JPEG compression [300x300, CR=10:1], Z=3.16

the four quadrants of the map to satisfy the real pixel value constraint, and to make consistent watermark detection if the scanned image is rotated for 90, 180, 270 degrees.

Because the log-log map coefficients and the DFT coefficients are not 1-to-1 mapping, the embedding process could not be done directly. We can use the iterative method to estimate the change in the log-log map, but manipulate the original DFT coefficients to make the mapping results as close as possible. The iteration process will be continued until either the strength of watermark or the fidelity loss is larger than a threshold. We use the Z-statistic as a measure of the strength of watermark [8].

The first few steps of watermark detection process are the same as in the embedding process. For the 1-bit watermark, the Z-statistic between the DFT magnitudes of log-log map and the known watermark pattern is calculated as an indication of the existence of watermark. In general, if Z>3, then it is considered as the existence of watermark with a false positive rate of 10^{-3}. For the multiple-bit watermarks, the embedded symbols are extracted by calculating the largest Z-statistic value of the position of shifted PN patterns.

4 Experiments

In our experiments, we tested the watermarked images at multiple stages of manipulation, because that is usually the real case.

We tested the robustness of watermark on a color image randomly chosen from the Corel Stock Photo Library. The original image size is 384x256. After the embedding process, the PSNR of watermarked image is 45.14 dB and the original strength of watermark, Z, is 16.26. There is no visible degradation on the watermarked image. These images are shown in Fig. 3(a) and 3(b).

We first test the watermarked image with general cropping and resizing functions using the Paint Shop Pro. The results are shown here.

Manipulation (step by step)	**Z value**
(1) Cut off borders (cropping) to 365x248	7.60
(2) Resize it to the original size: 384x256.	7.67
(3) Crop the image again to 339x253 (84%x96%) of the original area ratio.	5.00
(4) Resize the image: 256x256	4.45
(5) JPEG compression (Paint Shop Prop default QF, CR=10.6:1), Z=3.88	3.88

We print the watermarked image on an inkjet paper using the EPSON Stylus Photo EX. The physical size of printed image is 13.5x9cm^2. Then, we use the HP Scanjet 4C to scan this picture with defaulted resolution. The scanner automatically adjusts the brightness, contrast, gamma correction, and all other settings. We then test several manipulations on the scanned image. The images after step (1) and step (4) are shown in Fig 3(c) and 3(d).

Manipulation (step by step)	Z value
(1) Print & Scan: Image Size: 401x268.	6.37
(2). Crop the image to 385x259	4.70
(3) Resize the image: 300x300	4.43
(4) JPEG compression (Paint Shop Prop default QF, CR=9.97:1).	3.16

We use another image, which is a parrot with trees as background, to test the multiple-bit watermarking in the proposed algorithm. An information of "SIGNAFY" represented by 14 3-bit symbols is embedded to the image. The original image size is 512x768. After watermarking, the PSNR is 46.43 dB. Then, this image is cropped to 486x740, resized to 512x512, and compressed by JPEG. We find that all the bits can be extracted correctly after these attacks.

5 Summary and Future Direction

In this paper, we proposed a public watermarking algorithm that is robust to the print-and-scan and general scaling and cropping processes. Preliminary experiments have shown the effectiveness of this algorithm. In the future, we will go on to a test of large image database with the proposed method, and look for a method which can deal with the rotation, scaling, and cropping simultaneously.

Acknowledgements

This work was performed at and funded by the NEC Research Institute and Signafy, Inc. The author would like to thank Yuiman Lui, Matt Miller, and Jeff Bloom for sharing ideas, and Dr. I. J. Cox, and Prof. S.-F. Chang for helpful discussions and reviewing this paper.

References

[1] M. Kutter and F. A. P. Petitcolas, "A Fair Benchmark for Image Watermarking Systems," *SPIE Security and Watermarking of Multimedia Content*, San Jose, CA, Jan 1999.

[2] J. Fridrich and M. Goljan, "Comparing Robustness of Watermarking Techniques," *SPIE Security and Watermarking of Multimedia Content*, San Jose, CA, Jan 1999.

[3] C.-Y. Lin, S.-F. Chang, and I. J. Cox, "Modeling Image Print and Scan Processes," submitted to *Intl. Symp. on Multimedia Information Processing (ISMIP 99)*, Taipei, Taiwan, Dec. 1999.

[4] J. O'Ruanaidh and T. Pun, "Rotation, Scale and Translation Invariant Spread Spectrum Digital Image Watermarking," *Signal Processing 66 (1998)*.

[5] C.-Y. Lin, M. Wu, M. L. Miller, I. J. Cox, J. Bloom and Y. M. Lui, "Geometric Distortion Resilient Public Watermarking for Images," submitted to *SPIE Security and Watermarking of Multimedia Content II,* San Jose, CA, Jan 2000.

[6] S. Pereira, J. O'Ruanaidh, F. Deguillaume, G. Csurka and T. Pun, "Template Based Recovery of Fourier-Based Watermarks Using Log-polar and Log-log Maps," *IEEE ICMCS 99*, Florence, Italy, June 1999.

[7] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Trans. on Image Processing*, Dec. 1997.

[8] H. S. Stone, "Analysis of Attacks on Image Watermarks with Randomized Coefficients," *NECI Technical Report*, 1996.