# ROTATION, SCALING, AND TRANSLATION-INVARIANT MULTI-BIT WATERMARKING BASED ON LOG-POLAR MAPPING AND DISCRETE FOURIER TRANSFORM

*Wilson Wai Lun FUNG and Akiomi KUNISA*

Digital Systems Research Center
SANYO Electric Co., Ltd.
3-10-15 Hongo, Bunkyo-ku, Tokyo, 113-8434 Japan.
wilson@tk.hm.rd.sanyo.co.jp and kunisa@tk.hm.rd.sanyo.co.jp

## ABSTRACT

This paper proposes an RST-invariant multi-bit watermarking system based on LPM and DFT. This system embeds the watermark in a rotation and translation invariant domain obtained by first acquiring log-polar mapped DFT (LPM-DFT) magnitudes of an image, and then taking the magnitude of the 1-dimensional DFT along the angular axis of this LPM-DFT magnitudes. Neither the original cover image nor a presumed watermark is required during watermark extraction. Scaling invariant is accomplished by a simple search along the log-radius axis. Simulation results indicate that this new system has high robustness against RST distortion and moderate robustness against image compression such as JPEG.

## 1. INTRODUCTION

Many watermarking schemes have been proposed to improve the robustness of image watermarking against geometric distortions. Some of them are based on combining log-polar mapping (LPM) and discrete Fourier transform (DFT) to attain rotation, scaling, and translation (RST) invariant watermark[2, 4]. Most of these schemes, however, are zero-bit payload watermarking methods[1], which require the advance knowledge of the watermark during detection.

In this paper, we propose a new watermarking system using log-polar mapping (LPM) and discrete Fourier transform (DFT) to achieve robustness against RST distortions. This system embeds the watermark in a rotation and translation (RT) invariant domain, and handles scaling via a simple search along the log-radius axis. The original cover image is not required during watermark extraction. This proposed system is a multi-bit watermarking system, which has considerably broader applications than other RST-resilient zero-bit watermarking techniques[2, 4].

---

[1] Although many watermarking system generates a multi-bit watermark

## 2. RT-INVARIANCE VIA LPM AND DFT

The use of Fourier-Mellin transform (FMT) to achieve RST-invariant watermarking was first proposed in [3], which also suggested to implement FMT via performing DFT on a log-polar mapped image. According to [2], RST-invariance via FMT can be illustrated as follows:

$$
\begin{aligned}
i'(x,y) &= i(\sigma(x\cos\alpha + y\sin\alpha) + x_o, \\
&\quad \sigma(-x\sin\alpha + y\cos\alpha) + y_o)
\end{aligned} \tag{1}
$$

where $i'(x,y)$ is the version of an image $i(x,y)$ after being rotated by $\alpha$, scaled by $\sigma$, and translated by $(x_o, y_o)$.

The DFT magnitude of $i'(x,y)$ is given by

$$
\begin{aligned}
|I'(f_x, f_y)| &= |\sigma|^{-2}\big|I(\sigma^{-1}(f_x\cos\alpha + f_y\sin\alpha), \\
&\quad \sigma^{-1}(-f_x\sin\alpha + f_y\cos\alpha))\big|.
\end{aligned} \tag{2}
$$

As DFT magnitude of any image is invariant to translation, it is referred as translation-invariant (T-invariant) domain.

By applying LPM $(f_x, f_y) = (e^\rho \cos\theta, e^\rho \sin\theta)$ to (2), then rewriting the result in log-polar coordinates, we have

$$
|I'(\rho, \theta)| = |\sigma|^{-2}|I(\rho - \log\sigma, \theta - \alpha)|. \tag{3}
$$

The DFT magnitude of $|I'(\rho, \theta)|$ is RST-invariant, because it is independent of the shift $(\log\sigma, \alpha)$ from scaling and rotation. While [3] reported a successful attempt in embedding and extracting a watermark in this RST-invariant domain, other researchers had difficulty repeating this process and started pursuing alternatives using LPM only[2, 4].

In the proposed system, a watermark is embedded into the rotation and translation (RT) invariant domain, which is the magnitude spectrum of the 1-dimensional DFT (1D-DFT) along the angular ($\theta$) axis of log-polar mapped DFT

---

during embedding, the detector only determine the presence of this watermark. As this watermark is given to the detector in advance, it carries zero-bit worth of information, and is considered a zero-bit watermark[1].
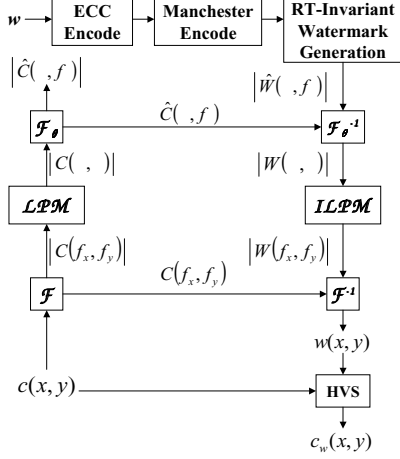
**Fig. 1**. The proposed embedding algorithm.

(LPM-DFT) magnitude. Given an image $f(x,y)$, its LPM-DFT magnitude $|F(\rho,\theta)|$ can be obtained from (2) and (3), and its representation in RT-invariant domain is defined as

$$|\hat{F}(\rho,f_\theta)| \triangleq |\mathcal{F}_\theta\{|F(\rho,\theta)|\}| \quad (4)$$

where $\mathcal{F}_\theta\{\cdot\}$ denotes 1D-DFT along $\theta$-axis.

Using (4), (3) is transformed into RT-invariant data by

$$
\begin{aligned}
|\hat{I}'(\rho,f_\theta)| &= |\mathcal{F}_\theta\{|I'(\rho,\theta)|\}| \\
&= |\sigma|^{-2}|\hat{I}(\rho - \log\sigma, f_\theta)|. \quad (5)
\end{aligned}
$$

The translation $\log\sigma$ due to scaling is handled by a simple search along $\rho$-axis.

## 3. EMBEDDING ALGORITHM

Figure 1 shows the block diagram of the proposed watermarking algorithm, where $|C(\cdot,\cdot)|$ and $\angle C(\cdot,\cdot)$ indicate the magnitude and phase of complex data $C(\cdot,\cdot)$ respectively. It consists of the following steps:

1. Obtain phase $\angle\hat{C}(\rho,f_\theta)$ and $\angle C(f_x,f_y)$ from cover image $c(x,y)$ as follows:

   (a) Perform DFT on $c(x,y)$ to obtain $|C(f_x,f_y)|$ and $\angle C(f_x,f_y)$.

   (b) Log-polar map $|C(f_x,f_y)|$ to obtain $|C(\rho,\theta)|$.

   (c) Perform 1D-DFT on $|C(\rho,\theta)|$ along $\theta$-axis to obtain phase $\angle\hat{C}(\rho,f_\theta)$.

2. Encode a watermark sequence $\mathbf{w} = (w_0,\ldots,w_{l-1})$ of $l$ bits with error correction code (ECC) to obtain $\mathbf{t} = (t_0,\ldots,t_{n-1})$ of $n$ bits. Then modulate $t_i, i = 0,\ldots,n-1$, with Manchester code to obtain

$$(m_{2i}, m_{2i+1}) = \begin{cases} (1,0) & \text{for } t_i = 1 \\ (0,1) & \text{for } t_i = 0. \end{cases} \quad (6)$$

3. From $\mathbf{m}$, generate 2-dimentional RT-invariant watermark $|\hat{W}(\rho,f_\theta)|$ by

$$\hat{W}(\rho, f_b + j) = \begin{cases} m_j & \text{for } \rho = \rho_1,\ldots,\rho_2 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

   where $j = 0,\ldots,2n-1$, and $\rho_1$, $\rho_2$ and $f_b$ are positive integers.

4. Inverse transform $|\hat{W}(\rho,f_\theta)|$ to spatial domain watermark $w(x,y)$ as follows:

   (a) Obtain the LPM-DFT magnitude domain watermark $|W(\rho,\theta)|$ by performing 1D-IDFT on $|\hat{W}(\rho,f_\theta)|$ with phase $\angle\hat{C}(\rho,f_\theta)$.

   (b) Obtain the DFT magnitude domain watermark $|W(f_x,f_y)|$ by applying inverse log-polar mapping (ILPM) to $|W(\rho,\theta)|$.

   (c) IDFT $|W(f_x,f_y)|$, using phase $\angle C(f_x,f_y)$, to its spatial form $w(x,y)$.

5. Combine the spatial watermark $w(x,y)$ and the cover image $c(x,y)$ via a human visual system (HVS).

In Step 3, the same watermark data is embedded $s$ times from $\rho_1$ to $\rho_2 = \rho_1 + s - 1$. According to [4], watermark should be embedded at the middle frequencies where it is less distorted by LPM and compression attacks. Also, watermark embedded at the low frequencies in RT-invariant domain tends to be overwhelmed by the energy of cover images. As a result, the watermark is not embedded to frequencies below $f_b$ in RT-invariant domain.

As mentioned in [3], using phase from the cover image during IDFT yields watermark that resembles the cover image. We extend this idea to the 1D-IDFT applied to RT-invariant watermark in Step 4a. Comparing to other watermarks that are generated using random phase, this watermark can be embedded at higher power to achieve higher robustness while the amount of distortion remains the same.

## 4. EXTRACTION ALGORITHM

Figure 2 illustrates the proposed watermark extraction algorithm, which is outlined as follows:

1. Transform a test image $c_t(x,y)$ into RT-invariant data $|\hat{C}_t(\rho,f_\theta)|$.

2. For each $\rho$, obtain a soft value sequence $\mathbf{v}^{(\rho)} = (v_0^{(\rho)},\ldots,v_{n-1}^{(\rho)})$ by

$$v_i^{(\rho)} = |\hat{C}_t(\rho, f_b + 2i)| - |\hat{C}_t(\rho, f_b + 2i + 1)|. \quad (8)$$

3. For all $\rho$, ECC-decode $\mathbf{v}^{(\rho)}$ into antipodal signal $\mathbf{d}^{(\rho)} = (d_0^{(\rho)},\ldots,d_{l-1}^{(\rho)})$.
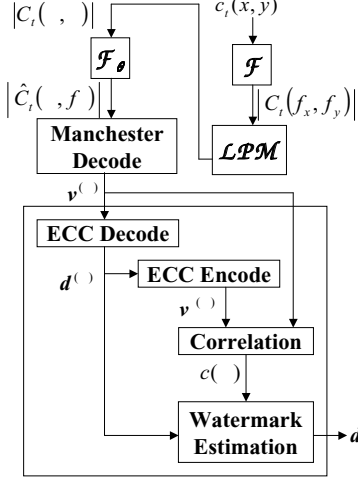
**Fig. 2**. The proposed extraction algorithm.

4. For all $\rho$, encode $\mathbf{d}^{(\rho)}$ into $\mathbf{v'}^{(\rho)}$ of $n$ bits with the ECC and calculate the normalized correlation $c(\rho)$ between $\mathbf{v}^{(\rho)}$ and $\mathbf{v'}^{(\rho)}$.

5. Estimate the watermark $\mathbf{d} = (d_0, \ldots, d_{l-1})$ with

$$
d_i = \begin{cases} 1 & \text{if } \sum_{\rho|c(\rho)>\tau} c(\rho)d_i^{(\rho)} > \mu \\ 0 & \text{if } \sum_{\rho|c(\rho)>\tau} c(\rho)d_i^{(\rho)} < -\mu \\ \text{invalid} & \text{otherwise} \end{cases}
$$

(9)

If any bit in $\mathbf{d}$ is invalid, $c_t(x, y)$ is considered a non-watermarked image.

The exhaustive search described in Step 3 to 5 is based upon the assumption that the soft value sequence $\mathbf{v}^{(\rho)}$ of any non-watermarked image consists of random data, which has a large distance from its nearest ECC codeword $\mathbf{v'}^{(\rho)}$ with high probability. Under this assumption, any watermark candidate $\mathbf{d}^{(\rho)}$ with $c(\rho) > \tau$ is considered valid and are summed up in (9) to estimate the watermark $\mathbf{d}$. For non-watermarked image, most, if not all, watermark candidates should be invalid, yielding a weighted sum with its absolute value smaller than $\mu$ in (9), and thereby allowing the extractor to identify non-watermarked images.

## 5. IMPLEMENTATION STRATEGIES

### 5.1. DFT on DFT Magnitude Spectrum

During watermark embedding, transformation of the RT-invariant watermark back to spatial domain involves performing 1D-IDFT first, followed by ILPM and then another IDFT. The first 1D-IDFT transforms the RT-invariant watermark into a real, DC-free[2] signal $W(f_x, f_y)$ in DFT spectrum domain, which violates the definition of magnitude of

---

As $|\hat{W}(\rho, f_\theta = 0)| = 0$ from (7), $W(f_x, f_y)$ contains the same amount of positive and negative data, or $E\{W(f_x, f_y)\} = 0$.

being strictly positive. If both magnitude $|W(f_x, f_y)|$ and phase $\angle W(f_x, f_y)$ of $W(f_x, f_y)$ is embedded, after translation distortion, only $|W(f_x, f_y)|$ remains valid. In this case, the watermark is severely distorted in RT-invariant domain from the fact: $|W(f_x, f_y)| \neq W(f_x, f_y)$.

One solution to this problem is to add a DC offset to the watermark such that the signal from 1D-IDFT is above zero in DFT spectrum domain. However, to produce such a DC offset requries an excessive amount of watermark power, which introduces significant distortion to the cover image.

Another solution is to set all of the negative samples of $W(f_x, f_y)$ to zero, resulting a signal $W_0(f_x, f_y)$ with $W_0(f_x, f_y) = |W_0(f_x, f_y)|$. We can express $W_0(f_x, f_y)$ as the average of $W(f_x, f_y)$ and $|W(f_x, f_y)|$, i.e.

$$
W_0(f_x, f_y) = \frac{W(f_x, f_y) + |W(f_x, f_y)|}{2}.
$$

(10)

In RT-invariant domain, (10) becomes

$$
|\hat{W}_0(\rho, f_\theta)| = \frac{|\hat{W}(\rho, f_\theta)| + |\mathcal{F}_\theta\{|W(\rho, \theta)|\}|}{2}.
$$

(11)

During watermark extraction, $|\mathcal{F}_\theta\{|W(\rho, \theta)|\}|$ can be regarded as distortion to the watermark signal $|\hat{W}(\rho, f_\theta)|$ that can be corrected by ECC. This is not the best solution, but the simplest one to be applied to the proposed system.

### 5.2. LPM and ILPM Using Forward and Backward Mappings

Any mapping from a source image $\mathbf{s}$ to a destination image $\mathbf{d}$ can be represented by a continuous mapping function $M$ and its inverse $M^{-1}$ such that $p' = M(p)$ and $p = M^{-1}(p')$, where $p$ is a continuous point[3] on $\mathbf{s}$ and $p'$ is its corresponding location on $\mathbf{d}$. During implementations, continuous points $M(p)$ and $M^{-1}(p')$ are quantized to their nearest discrete points[4] by quantization function $Q$. Forward and backward mappings are two implementation techniques approximating the continuous mapping.

Forward mapping operates on the coordinate system of $\mathbf{s}$, copying each sample on point $p$ of $\mathbf{s}$ to its corresponding discrete point $p'_d = Q(M(p))$ on $\mathbf{d}$. If $N \geq 1$ samples on points $\mathbf{p} = (p_0, \ldots, p_{N-1})$ of $\mathbf{s}$ are mapped to one sample $d(p'_d)$ at point $p'_d$, or $\forall p_i \in \mathbf{p} : p'_d = Q(M(p_i))$, $d(p'_d)$ is approximated by an $N$ weighted average of $s(p_i)$. This allows forward mapping to use all available redundant information to approximate $d(p'_d)$ accurately. In contrast, when no sample is mapped to $d(p'_d)$, $d(p'_d)$ is undefined, or a "hole", which distorts the watermark severely.

In contrast, backward mapping operates on the coordinate system of $\mathbf{d}$, computing the value for sample $d(p'_d)$ on $\mathbf{d}$ by sampling at its corresponding location $M^{-1}(p'_d)$ on $\mathbf{s}$.

---

[3]A continuous point is a point $(x, y)$ with $x$ and $y$ being real numbers.
[4]A discrete point is a point $(x, y)$ with $x$ and $y$ being integers.

**Table 1**. Parameters used in current implementation of the proposed watermarking system.

| Message length | $l = 32$ |
|---|---|
| ECC (Turbo code) | rate 1/3 with memory size 6 |
| Code length | $n = 114$ |
| Frequency offset | $f_b = 32$ |
| Embedding redundancy | $s = 13$ |
| Embedding frequency range in T-invariant domain[5] | $\omega_1 = 0.33 \ (\rho_1 = 455)$ $\omega_2 = 0.43 \ (\rho_2 = 468)$ |
| Detection threshold | $\mu = 0, \tau = 0.53$ |

If $s(M^{-1}(p'_d))$ does not exist, it is interpolated using samples at discrete points neighbouring $M^{-1}(p'_d)$. In this way, every sample on $\mathbf{d}$ has its value computed from $\mathbf{s}$ via interpolation. The interpolated value of $s(M^{-1}(p'_d))$, however, may not be the optimal approximation of $d(p'_d)$. This is especially significant when many samples on $\mathbf{s}$ are mapped to one sample on $\mathbf{d}$, causing downsampling to occur. In this case, interpolation using only discrete samples neighbouring $M^{-1}(p'_d)$ is insufficient to eliminate aliasing.

In the proposed system, backward and forward mapping is used for ILPM and LPM respectively. During ILPM, the watermark data is slightly oversampled to lessen the interference of neighboring watermark data samples. Backward mapping produces more stable results by eliminating occurrences of "hole". During LPM, the watermarked frequency in Cartesian DFT spectrum is unknown to the extractor, thereby requiring LPM of the whole spectrum. The many-to-one mapping at high frequencies favors forward mapping. Although a lack of source samples at low frequencies leads to occurrences of "hole", sample density in DFT spectrum can be increased using zero-padding method suggested by [2], which eliminates this problem.

## 6. SIMULATION RESULTS

To verify the robustness of the proposed system, two sets of simulations are conducted using parameters as shown in Table 1. Watermark robustness is measured by bit error rate (BER) and watermark error rate (WER).

In the first set of simulations, 1000 different watermarks are embedded into the image "lena" of size $512 \times 512$ at PSNR=43.2dB. These watermarked images are then tested for their robustness against various rotation, scaling, translation and JPEG compression attacks. Table 2 summarizes the results for this set of simulations.

In the other set of simulations, robustness of the watermark across multiple images is evaluated. The same 1000 watermarks are embedded into 4 different cover images. All watermarked images are rotated by $15°$, scaled by 0.75, translated by $(5, 10)$, and then JPEG compressed at QF=90.

---

[5]Normalized frequency $\omega$ of Cartesian frequencies $(f_x, f_y)$ is defined as $\omega = \sqrt{f_x^2 + f_y^2} / \sqrt{f_W^2 + f_H^2}$ where $f_W$ and $f_H$ are the highest horizontal and vertical frequency respectively.

**Table 2**. Simulation results for rotation, scaling, translation and JPEG compression attack.

| Attack type | Parameter | BER | WER |
|---|---|---|---|
| Rotation | $0.5°$ | $3.1 \times 10^{-5}$ | $1.0 \times 10^{-3}$ |
| | $1.0°$ | 0 | 0 |
| | $2.0°$ | $4.7 \times 10^{-4}$ | $3.0 \times 10^{-3}$ |
| | $3.0°$ | $3.1 \times 10^{-5}$ | $1.0 \times 10^{-3}$ |
| | $10.5°$ | 0 | 0 |
| | $15.0°$ | 0 | 0 |
| | $20.5°$ | 0 | 0 |
| | $26.0°$ | 0 | 0 |
| | $34.5°$ | 0 | 0 |
| | $45.0°$ | 0 | 0 |
| Scaling | 0.5 | $9.4 \times 10^{-1}$ | $1.0 \times 10^{0}$ |
| | 0.7 | 0 | 0 |
| | 0.9 | 0 | 0 |
| | 1.1 | 0 | 0 |
| | 1.3 | 0 | 0 |
| | 1.5 | $5.0 \times 10^{-4}$ | $7.0 \times 10^{-3}$ |
| Translation | (100,100) | 0 | 0 |
| JPEG | 40 | $1.0 \times 10^{-1}$ | $2.5 \times 10^{-1}$ |
| | 50 | $1.9 \times 10^{-3}$ | $1.5 \times 10^{-2}$ |
| | 60 | $3.1 \times 10^{-5}$ | $1.0 \times 10^{-3}$ |

**Table 3**. Simulation results with various cover images.

| Image | PSNR | BER | WER |
|---|---|---|---|
| Lena | 43.2dB | $1.8 \times 10^{-3}$ | $1.5 \times 10^{-2}$ |
| F16 | 42.2dB | $2.5 \times 10^{-4}$ | $4.0 \times 10^{-3}$ |
| Mandrill | 38.4dB | $6.9 \times 10^{-4}$ | $1.0 \times 10^{-2}$ |
| Peppers | 42.7dB | $3.8 \times 10^{-4}$ | $3.0 \times 10^{-3}$ |

The simulation results are summarized in Table 3.

## 7. CONCLUSION

In this paper, we proposed a new RST-invariant multi-bit watermarking system based on LPM and DFT. This system pertains high robustness against RST-distortion as well as JPEG compression, while quality of the watermarked image is well preserved.

## 8. REFERENCES

[1] I. J. Cox, M. J. Miller, and J. A. Bloom, *Digital Watermarking*, Morgan Kaufmann Publishers, 2002.

[2] C. Lin, M. Wu, J. A. Bloom, I. J. Cox, M. L. Miller, and Y. M. Lui, "Rotation, scale, and translation resilient public watermarking for images", *IEEE Trans. on Image Proc.*, vol. 10, no. 5, pp. 767–782, May 2001.

[3] J. J. K. Ó Ruanaidh and T. Pun, "Rotation, scale and translation invariant digital image watermarking", *Proc. IEEE Inter. Conf. on Image Proc.*, Santa Barbara, CA, vol. 1, pp. 536–539, Oct. 1997.

[4] D. Zheng, J. Zhao, and A. El Saddik, "RST-invariant digital image watermarking based on log-polar mapping and phase correlation", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 753–765, Aug. 2003.