Convolution Image-based Watermarking for 2D Greyscale Image

LEE Tsz Hong Homer, Department of Mathematics
The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong. *

May 27, 2016

Abstract

Watermarking technique for image is an efficient method for protecting copyright image, and also a huge topic in cryptography. In this paper, two spread spectrum watermarking scheme, the Convolution Image-based Model (CIM) and the Exponential Convolution Image-based Model (ECIM) are going to be formulated and discussed. The watermarking experiment result will be shown and discussed, focusing on the attack scheme, protectability, and the information encryption of the watermark. We will show that the convolution image-based model for invisible watermark is weak of protectability, but it is able to hide the information (the size of watermark must be less than the original image) and store inside the image.

1 Introduction

The information spread like a fire in these few decade. However, the copyright of an image is much more important than before. It is because the more views or querying of the image, the more valuable of the image. The piracy behaviour can be tackled if we can recognise the copyright information.[8] That is the reason we need invisible watermarking for any types of files, since once the user copied the signal, the watermark is also being duplicated.

The main reason to find a watermarking method such that human cannot find it easily is that the downloader/user will not attack the watermark. Moreover, any type of the secure watermark must be selected to make sure that, any attack on the watermark will cause the degradation of the original image.

In this paper, we will discuss our model - Convolution Image-based Model (CIM) and Exponential Convolution Image-based Model (ECIM), which inspired by[5] the image Spread Spectrum watermarking (SSW) technique, which is widely used in 1D signal files, such as audio files and text files. We will discuss the protectability of the models. However, the watermark is always want to be removed by another user, especially for hackers. We will also discuss the information encryption properties of the watermarks under ECIM and CIM.

2 Literature Review

2.1 Discrete Fourier Transform (DFT)

DFT is a discrete extension of Fourier Transformation. [2] Fourier Transform $\mathcal{F}[\,\cdot\,]$ is a transformation for a integrable function, where map the (real) function f to the (complex) frequency domain $\hat{f}$ from the time domain, defined by:

$$\mathcal{F}[f](\xi) = \hat{f}(\xi) = \int f(x)\, e^{-i2\pi x\xi}\, dx$$

And the Inverse Fourier Transform (IFT) is defined by:

$$\mathcal{F}^{-1}[\hat{f}](x) = f(x) = \int \hat{f}(\xi)\, e^{i2\pi x\xi}\, d\xi$$

Therefore the definition of DFT is following[1]:

$$\mathcal{F}[f](\xi) = \hat{f}_m = \sum_{n=0}^{M-1} f_n\, e^{-i2\pi mn/M},$$

and Inverse Discrete Fourier Transform (IDFT) is following:

$$\mathcal{F}^{-1}[\hat{f}](x) = f(x) = \sum_{n=0}^{M-1} \hat{f}_m\, e^{-i2\pi mn/M}$$

Note that DFT provided a periodic extension for the function (i.e., $f_m = f_{m+M}$ for all m and the period is donated by M.).

2.2 Fast Fourier Transform (FFT)

FFT (IFFT) is a fast algorithm method for DFT (IDFT), which was included in the [4] Top 10 Algorithms of 20^{th} Century. [3] The complexity was reduced to $\mathcal{O}(n \log n)$ from $\mathcal{O}(n^2)$. The algorithm was designed based on the properties of:

1. Periodic, where $e^{-i2\pi n/M} = e^{-i2\pi(n+M)/M}$
2. Symmetric, where $e^{-i2\pi n/M} = -e^{-i2\pi(n+\frac{M}{2})/M}$
3. Divisibility, where $e^{-i2\pi mkn/M} = -e^{-i2\pi kn/(\frac{M}{m})}$

Therefore, we separate the function y_n by odd and even part. Repeating the procedure and finally we have a faster algorithm than FT, which is especially important for high dimensional signal.

2.3 Spread Spectrum Watermarking(SSW)

SSW is widely used in the audio watermarking. [5] This is one of the most robust watermarking techniques, which watermark the (audio) signal in the frequency domain.

The watermark is spread over almost the whole spectrum, at least spread over the higher coefficients in the frequency space. Therefore its energy is undetectable in one narrow-band and the high amplitude noise must be added to all frequency band to destroy the watermark.

One of the common method of SSW, apply the watermark function $x = (x_1\, x_2 \cdots x_N)$to the $M \times N$ image $v_{i,j}$ by the function of their DCT coefficients.

3 Model Governing

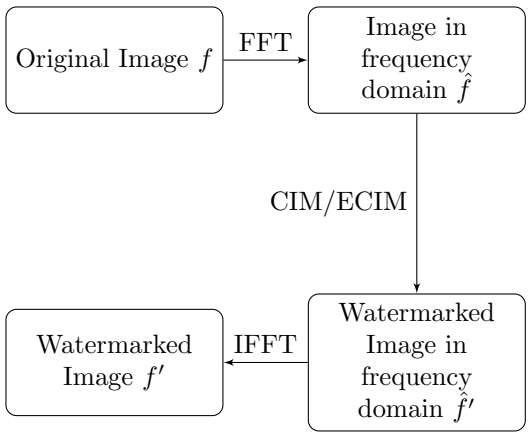
There are some SSW models are used frequently and we will discuss about convolution model and the exponential convolution model. The flow of watermarking is shown in Figure 1.

3.1 The Convolution Image-based Model (CIM)

For the original $M \times N$ image f and its frequency representation $\hat{f}$, donated the watermarking noisy $M_g \times N_g$ image function by g ($\hat{g}$ in the frequency space, where the expected value of $\hat{g}$ is given by $E[\hat{g}] = 0$, with $M_g \leqslant M, N_g \leqslant N$). The watermarked image f' ($\hat{f}'$ in the frequency domain) is formulated by:

$$\hat{f}' = \hat{f}(1 + \alpha\hat{g}) \tag{1}$$

Figure 1: The Flow Chart of SSW Models



pick an appropriate value of α such that the watermark is not obvious by human perception (i.e., by eye-norm). We can see the relationship below clearly.

$$E[\hat{f}'] = E[\hat{f}(1 + \alpha\hat{g})] = E[\hat{f}] + \alpha E[\hat{f}\hat{g}]$$

Note that f is independent of g (i.e., $\hat{f}$ is independent of $\hat{g}$), which means $E[\hat{f}\hat{g}] = E[\hat{f}]E[\hat{g}] = 0$.Therefore we have the following relationship:

$$E[\hat{f}'] = E[\hat{f}]$$

We expect to generate an (statistical) unbiased output image. And recall the convolution theorem state that, the equation [1] is just equivalent that $f' = f * h$, where $h = \delta + \alpha g$, which implies the larger $|\alpha|$, the less likelihood between f and $\hat{f}$. Also, if we pick $g(\vec{t}) = \frac{1}{\sqrt{2\pi}\sigma} \exp(\frac{\vec{t}^2}{2\sigma^2})$, as $\sigma \to 0$, $g(\vec{t}) \to \delta(\vec{t})$. The amplitude of output image f' is just the adjusted (multiply by $(1+\alpha)$) amplitude of the original image. Obviously this watermarking function is not spread over the whole spectrum and it can be attacked easily. This is a generalized SSW model for all image watermarking scheme. The watermarking function g is almost free to be choose. For given alpha, original image and the output image, we can extract the watermark back by the following formula:

$$g = \frac{1}{\alpha}\mathcal{F}^{-1}[\frac{\hat{f}'}{\hat{f}} - 1] \tag{2}$$

3.2 The Exponential Convolution Image-based Model (ECIM)

Similar with CIM, ECIM is defined by the following:

$$\hat{f}' = \hat{f}(e^{\alpha\hat{g}}) = \hat{f}(1 + \alpha\hat{g} + \cdots) \tag{3}$$

By Taylor expansion, we can clearly see that CIM is approximately equivalent with ECIM when $\alpha \approx 0$. However there is slightly different in the formula for getting back the reverted watermark from the image.

$$g = \frac{1}{\alpha}(\mathcal{F}^{-1}[\log(\hat{f}')] - \mathcal{F}^{-1}[\log(\hat{f})]) \tag{4}$$

We will show that the result of ECIM are indifferent with the result of CIM by human

4 Experiment Result

We use FFT for the both model to reduce the runtime. All image was generated within 3 seconds. In the CIM, we take $\alpha = 1 \times 10^{-5}$. It is available to get back the watermark from the output image, when the amplitude of the addition Gaussian noise is 2.5×10^{-3}. As we can observe the watermark in the image is polluted in Figure 2. The result we obtain in ECIM are similar to CIM, where the value of α and the amplitude of the Gaussian noise we pick are the same. We also simulated one common attack - image compression. We perform the compression by the low rank approximation of Singular Value Decomposition (SVD, which are not going to discuss in this paper). In the Figure 3, the original image is just simply k=359. However, start from 353, the noise might be discovered in the reverted watermark (the watermark obtained back from the watermarked image). There is also a new problem, these model required new intensity level to store the information in the image.

Figure 2: Experiment Result of CIM

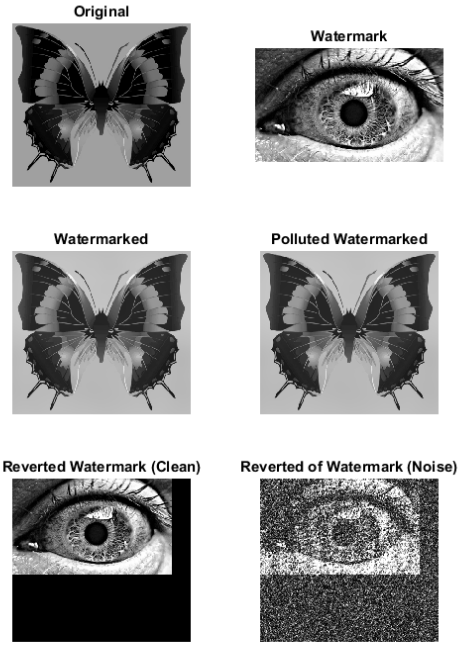
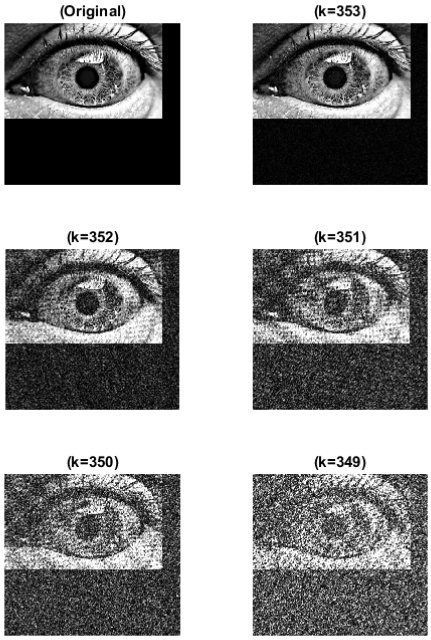


Figure 3: Reverting the watermark under SVD image compression by CIM



5 Discussion

5.1 Detection Scheme

For the CIM model, it is hard to detect the watermarking function without the original image, even the hacker might simply iterate different value of α the appropriate value of α. The CIM / ECIM are extremely strong to avoid encryption information being disclose from hackers. Hackers are able to detect the existence of watermark by finding new intensity level. However it is impossible to know the watermark without original image by equation (2) and (4).

5.2 Attacking Scheme

It is hard to remove the watermark unless the watermarking function is known. The specific explicit attacking scheme do not exist. The following attacking

Figure 4: Result of Watermarked Image and the Reverted Watermark in different value of k

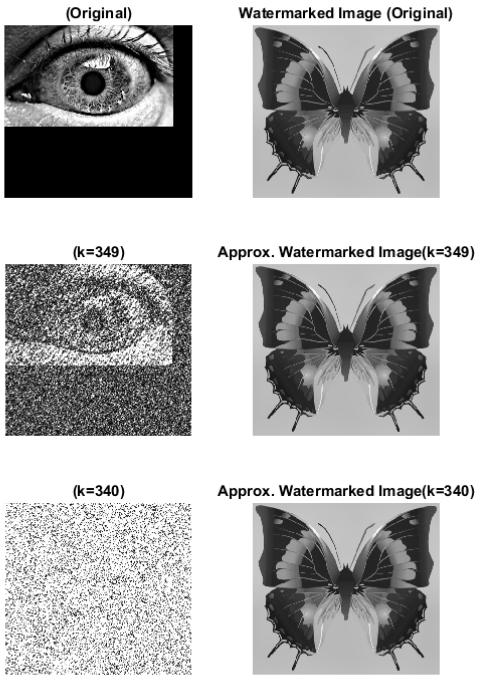
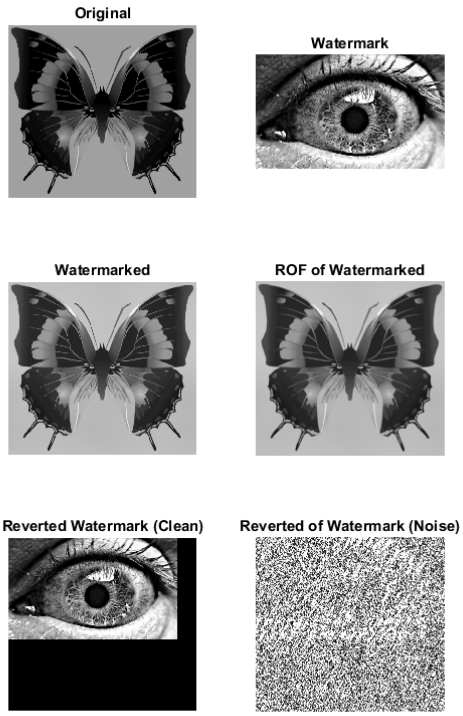


Figure 5: Result of Watermarked Image and the Reverted Watermark under ROF attack, with $\lambda = 1 \times 10^{-20}$



method (no matter watermark removal or extracting) are used quite often.

5.2.1 Image Compression

Image Compression is a common method for attacking watermarks. We applied SVD image compression technique in the experiment. Note that SVD compression is not a frequency, or Fourier transform based technique. Therefore the information lost independent of the frequency in the Figure 3. Moreover, the singular value dropped in low rank approximation of SVD compression is just a small value ($\frac{359-340}{359} = 5.2925\%$) in Figure 4. Clearly the SVD image compression is a successful attacking scheme for watermark removal.

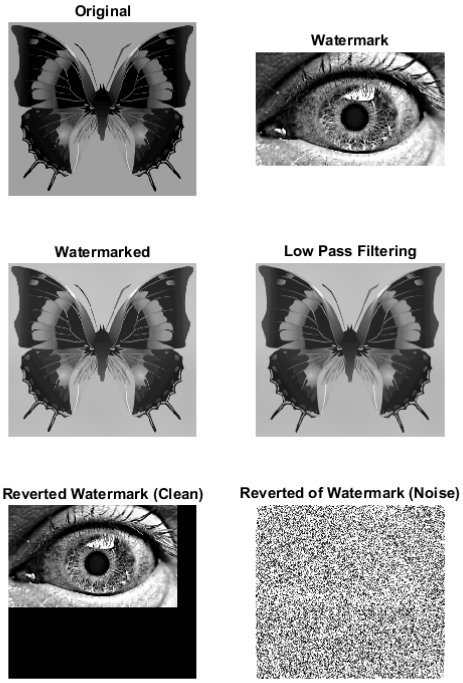
5.2.2 Noise Addition/ multiple watermarking

In the experiment, we simulated Gaussian noise addition (the addition random variable X is following the normal distribution with mean 0 and standard deviation 1, i.e. $X \sim N(0,1)$), which might be used by hackers. In Figure 2, we observed that even the amplitude is 250 times of α, the watermark is still able to be recognised. Therefore, Gaussian noise addition seems is not a well-perform attacking scheme. However, the amplitude of the noise is about 10^{-3} times of the watermarked image intensity, which implies that the noise is not much affect the watermarked image, but affect a lot to the watermark. The protectability is extremely weak under additive Gaussian Noise attack. The additive Gaussian noise can be treated as one of the multiple-watermark attacking scheme. Another possible multiple frequency domain watermarking scheme defined by the following formula might works.

$$\hat{f''} = u(\hat{f'}, \hat{g_0})$$

where $\hat{f''}$, $\hat{g_0}$ are the output and input (another watermark by hacker) watermark, respectively. However we will not discuss more since the function of u might vary and the strength of the attacking scheme is highly depends on the function of u.

Figure 6: Result of Watermarked Image and the Reverted Watermark under Butterworth Low Pass Filtering attack, with $D_0 = 200$, order $= 1$



5.2.3 Low Pass Filtering and TV-denoising

We performed about the TV-denoising by ROF model attack, and the butterworth low pass filtering attack in our experiment, shown as Figure 5 and Figure 6. Both of the denoising attacking results are similar. Although the original image is given, and the affect of the denoising is small (It is indifferent by human eyes.), the destroy of the watermark is sufficiently large. We can only extract the watermark without denoising attack.

5.3 Protectability

Obviously, protectability is highly related with the attacking scheme. As we shown before, the protectability is extremely low under a lot of types of attacking scheme. Clearly we do not use CIM or ECIM if we need a strong watermark for protecting copyright files.

5.4 Information Encryption

The higher protectability, the lower information encryption ability. Users might easily remove the watermark in the image under different attack scheme. However, the information can be contain in the image is very large. The size of secret hidden image can be exactly same as the original image. Also, although the hackers got the value of α, they still cannot detect the watermark without the original image. This is an excellent property of this watermarking model for information encryption. However, no matter CIM or ECIM, they create too much extra intensity of the image. Therefore, the file size might increase exponentially. Hackers may realize the existence of the watermark.

6 Conclusion

The CIM and ECIM are not nice watermark technique to protect the copyright files. However the information encryption properties are useful in some fields (image-based password technique).Thank you for coding advice by Professor Shingyu Leung[6].

Please view appendices for demonstrate the experiment results under CIM model. (We only provided the code of watermarking experiment under simulating additive Gaussian noise attacking scheme. The process of SVD compression, High/Low pass filtering and TV denoising of CIM, and the respective of ECIM code are similar (just replace the function of Bn and replace the watermarking function by equation (3)).

References

[1] *Discrete Fourier transform*, Wikipedia, 2016., https://en.wikipedia.org/wiki/Discrete_Fourier_transform

[2] *Fourier transform*, Wikipedia, 2016., https://en.wikipedia.org/wiki/Fourier_transform

[3] *Fast Fourier transform*, Wikipedia, 2016., https://en.wikipedia.org/wiki/Fast_Fourier_transform

[4] Dongarra, J.; Sullivan, F. , *Guest Editors Introduction to the top 10 algorithms*, August 2002, http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=814652

[5] *Audio watermark*, Wikipedia, 2016., https://en.wikipedia.org/wiki/Audio_watermark

[6] Shingyu LEUNG , *MATLAB demo files for the ideal low/high pass filters and the Butterworth low/high pass filter*, April 2016, http://math4336-2016s.blogspot.hk/2016/04/lecture-19-apr-14.html

[7] Rudin, Stanley Osher and Emad Fatemi, *Nonlinear total variation based noise removal algorithms*, 1992, http://www.csee.wvu.edu/~xinl/courses/ee565/total_variation.pdf

[8] Dr.M.Mohamed Sathik and S.S.Sujatha, *An Improved Invisible Watermarking Technique for Image Authentication*, November 2010, http://www.sersc.org/journals/IJAST/vol24/6.pdf

Appendices

Reference Matlab code (Additive Gaussian Noise):

```
function [B] = CIM(A, g, alpha, noise)
[m,n,k] = size(A);
if k>1
    A = rgb2gray(A/255);
end
A = double(A);
[mg,ng,kg] = size(g);
if kg>1
    g = rgb2gray(g/255);
end
g = double(g);
g1 = g;
if mg>m
    g1(m+1:end,:) = [];
end
if ng>n
    g1(:,n+1:end) = [];
end
if mg<m
    g1(mg+1:m,:) = zeros(m-mg,ng);
end
if ng<m
    g1(:,ng+1:n) = zeros(m,n-ng);
end
%The convolution theorem says we just use pointwise mult
F_A = fftshift(fft2(A));
F_g = fftshift(fft2(g1));
B = abs( ifft2( F_A .* (1 + alpha .* F_g) ));
%Addition of Gaussian Noise
Bn = B + noise.*randn(size(B))/255;
%Getting back the watermark image from known f', f, alph
F_B = fftshift(fft2(B));
F_Bn = fftshift(fft2(Bn));
w = abs( ifft2( (F_B ./ F_A -1) / alpha) );
wn = abs( ifft2( (F_Bn ./ F_A -1) / alpha) );
%Convert to 8-bit and display the images.
B = uint8(B*255);
A = uint8(A*255);
Bn = uint8(Bn*255);
g = uint8(g*255);
w = uint8(w*255);
wn = uint8(wn*255);

subplot(3,2,1); imshow(A); title('Original');
subplot(3,2,2); imshow(g); title('Watermark');
subplot(3,2,3); imshow(B); title('Watermarked');
subplot(3,2,4); imshow(Bn); title('Polluted Watermarked'
subplot(3,2,5); imshow(w); title('Reverted Watermark (CI
subplot(3,2,6); imshow(wn); title('Reverted of Watermark
colormap gray;
```