RESEARCH ON THE AUTOCORRELATION DETECTION IN DIGITAL WATERMARKING TECHNIQUES

Cheng-song YANG[a], Chang-qing ZHU[a,b],Qisheng Wang[a,c]

[a]Zhengzhou Institute of Surveying and Mapping, Zhengzhou 450052 - ycsdongshang@163.com
[b]Ministry of Education Key Laboratory of VGE, Nanjing Normal University, Nanjing 210054 - zcq88@263.net
[c]Surveying and Mapping Support Unit of South Xinjiang,Shule,Xinjiang.844200 - qishengw@yahoo.cn

Commission VI, WG IV/1

KEY WORDS: Digital Watermarking, Autocorrelation Detection, Multiple Watermarking, Robustness

ABSTRACT:

Based on the self-characteristic of digital watermarking techniques and the multi-statistic theory, this paper makes a research on the watermark autocorrelation detection. First, the setting of detection threshold, the calculating of false alarm probability and missing detection probability are studied. Second, the relationship between the length of watermark, the watermark embedding strength and the attack strength is established. Third, an improved method of watermark generating is proposed to promote the robustness of watermarking algorithms. Finally, the probability of multiple watermark embedding is studied.

1. INTRODUCTION

The digital watermarking is a new developed technology for information security. By embedding copyright information into and hiding it in the source data, it makes the copyright information as a part of the source data which can not be divided easily. The digital watermarking techniques can be used to protect copyright. The digital watermarking has been applied in many fields such as digital images, video, audio and so on (Sun Shenghe, 2004; Yang Yixian, 2006).

A watermarking system is comprised of watermark generation, watermark embedding and watermark extraction/detection. The main research of watermarking techniques is on the watermark embedding(Li Yuanyuan,2004;Zhong Shangping,2006). However, in the process of watermark extraction, the watermark generation and watermark autocorrelation detection have an important influence on the accuracy and reliability of the watermark extraction/detection result. Based on the reasons above, under the additive random noise, this paper makes a research on the generation of watermark and the autocorrelation detection of watermark based on the statistics theory, studying the setting of detection threshold, the calculating of false alarm probability and missing detection probability, the relationship between the length of watermark、the watermark embedding strength and the attack strength. Finally, the probability of multiple watermark embedding is studied.

2. BASIC AUTOCORRELATION DETECTION THEORY FOR DIGITAL WATERMARKING

Let $W=\{w(k)\}, k=0,1,\cdots,n-1$ be the watermark,

here $P\{w(k)=1\}=\frac{1}{2}$, $P\{w(k)=-1\}=\frac{1}{2}$. Let λ be the watermark strength, the reason why choose such watermark is that this watermark has better statistic characteristic for watermark detection.

Let $G=\{g(k)\}, k=0,1,\cdots,n-1$ be random noise, here

$$E\left[w(k)\right]=0,\quad D\left[w(k)\right]=1$$
$$E\left[g(k)\right]=\mu,\quad D\left[g(k)\right]=\sigma^2$$

$w(k)$ and $g(k)$ is mutually independent, so

$$E\left[w(k)*g(k)\right]=0$$
$$D\left[w(k)*g(k)\right]=\mu^2+\sigma^2$$

From the central limit theorem, we can get that:
When $n\to+\infty$

$$\sum_{k=0}^{k=n-1} g(k)*w(k)\sim N(0,n(\mu^2+\sigma^2))$$

So, when the to-be-detected information is attacked by the random noise, there are two cases:

Case 1: Contain watermark

To-be-detected information:

$$Y=\{y(k)\}=W+G=\{\lambda*w(k)+g(k)\}$$

The result of detection:

$$c = Y * W = \{y(k) * w(k)\} = \sum_{k=0}^{k=n-1} y(k) * w(k)$$

$$c = \sum_{k=0}^{k=n-1} \lambda * w(k) * w(k) + \sum_{k=0}^{k=n-1} g(k) * w(k)$$

When $n \to +\infty$,

$$c = \lambda * n + \sum_{k=0}^{k=n-1} g(k) * w(k)$$

$$c \sim N(\lambda * n, n(\mu^2 + \sigma^2)) \quad (1)$$

Case 2: Non-contain watermark

To-be-detected information:

$$Y = \{y(k)\} = G = \{g(k)\}$$

The result of detection:

$$c = \sum_{k=0}^{k=n-1} y(k) * w(k) = \sum_{k=0}^{k=n-1} g(k) * w(k)$$

$$c = \sum_{k=0}^{k=n-1} g(k) * w(k) \sim N(0, n(\mu^2 + \sigma^2)) \quad (2)$$

Because the formula (1) and (2) have same variance, discrimination rule can be set as follow (Wang Xuemin, 1999):

$$\begin{cases} \text{contain watermark}, & \text{if } c \geq \dfrac{\lambda * n}{2} \\ \text{non-contain watermark}, & \text{if } c < \dfrac{\lambda * n}{2} \end{cases}$$

Then flow detection probability is

$$e_1 = e_2 = \Phi\left(-\frac{\Delta}{2}\right) = \Phi\left(-\frac{\lambda * \sqrt{n}}{2 * \sqrt{\sigma^2 + \mu^2}}\right) \quad (3)$$

When $\sigma = 10, \mu = 5, n = 1000, \lambda = 1.0$, the probability distributing of (1) and (2) is as figure 1 show.

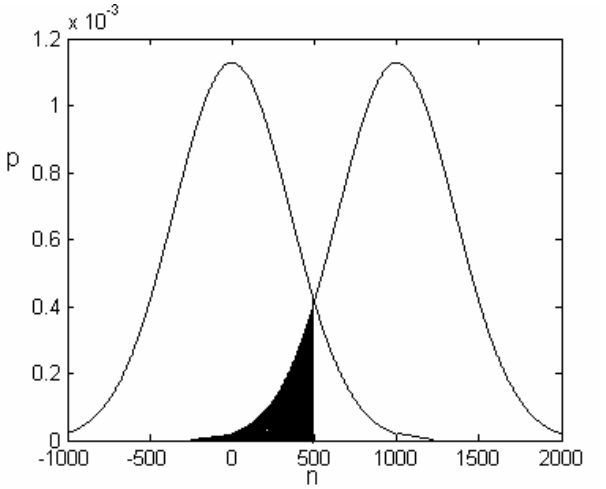

Figure 1. The probability distributing of autocorrelation detection

$\frac{\lambda * n}{2}$ =500, so discrimination rule as follow

$$\begin{cases} \text{contain watermark}, & \text{if } c \geq 500 \\ \text{non-contain watermark}, & \text{if } c < 500 \end{cases}$$

Then missing detection probability

$$e_1 = e_2 = \Phi\left(-\frac{\lambda * \sqrt{n}}{2 * \sqrt{\sigma^2 + \mu^2}}\right) = \Phi(-1.4142) = 0.0787$$

Because λ, n is variable, it is not convenient to compare $\frac{\lambda * n}{2}$ with detect result c to judge whether the to-be-detected information contain watermark or not, we can normalize c to solve this problem.

Let $z = \dfrac{c}{\lambda * n}$, the discrimination rule can be set as follow:

$$\begin{cases} \text{contain watermark}, & \text{if } z \geq 0.5 \\ \text{non-contain watermark}, & \text{if } z < 0.5 \end{cases}$$

3. THE RELATIONSHIP BETWEEN WATERMARK LENGTH、EMBEDDING STRENGTH AND ATTACK STRENGTH

From the formula (3) we know that: on the same condition of attack, watermark length(n)、embedding strength(λ) and attack strength(decided by $\sigma^2 + \mu^2$)decide the rate of false alarms. The rate of false alarms is proportional to the noise

strength ($\sigma^2+\mu^2$) and inversely proportional to the watermark length(n) and embedding strength(λ).

When $\lambda * n$ is fixed, the noise strength ($\sigma^2+\mu^2$) can't be too big, otherwise, the rate of false alarms will be so big that it is impracticable to use the discrimination rule to judge whether there is watermark or not.

When $n=1000, \lambda=1.0$,under the random noise attack which has the statistic characteristic of mean μ =0, variance σ = 2.0、4.0、6.0、8.0, the detection result is as figure 2, here, the 500th to-be-detected information contain watermark. From figure 2 we can get that: the rate of false alarms is small when variance σ is 2.0、4.0 and will increase as the variance σ increase.

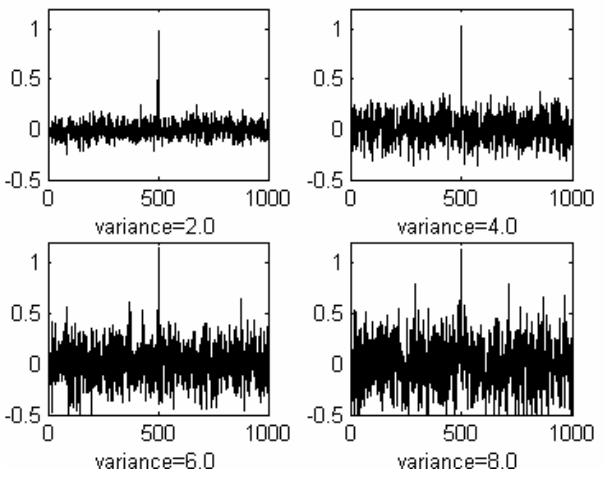

Figure 2. The influence of variance σ to detection

In order to get good attack-resistance capability, the "3σ"rule is erected. It means that there will be no probability of false alarms when $\frac{\mu_2-\mu_1}{2} \geq 3\sqrt{n}\sigma$ .This rule request that:

$$\frac{\lambda * n}{2} \geq 3 * \sqrt{n(\mu^2+\sigma^2)}$$

The $\sqrt{(\mu^2+\sigma^2)}$ can be treated as the random noise strength, when the watermark detection need to resist the random noise below $\sqrt{(\mu_0{}^2+\sigma_0{}^2)}$, the right value of λ, n must be set to satisfy the follow inequality

$$\frac{\lambda * n}{2} \geq 3 * \sqrt{n(\mu^2+\sigma^2)} \qquad (4)$$

At the same time, we can control the false alarms under certain threshold p by control λ 、 n as the formula:

$$e_1 = e_2 = \Phi\left(-\frac{\lambda * \sqrt{n}}{2 * \sqrt{\sigma^2+\mu^2}}\right) \leq p \qquad (5)$$

From the formula (4) and (5), it is known that in order to promote the attack-resistance capability of detection, we can increase the value λ and n ,but considering the central limit theorem, the value of n can't be too small, experiment testify that n should be no less than 800.

4. THE INFLUENCE OF MEAN μ OF RANDOM NOISE TO THE WATERMARK DETECTION

Data's translation has a bad influence on the watermark extraction and detection (YANG Cheng-song, 2007), Data's translation can be treat as the mean (μ) of the random noise. In order to display the influence of the mean (μ) of random noise to the watermark detection, let $\sigma=10, n=1000, \lambda=1.0$, change the value of μ , the variation of false alarms is showed in figure 3.

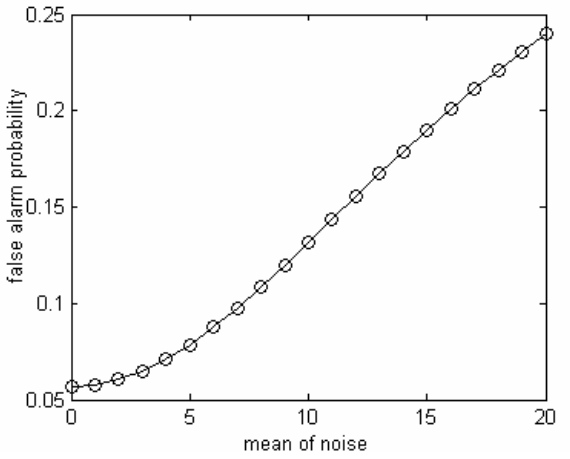

Figure 3. The influence of noise mean to the false alarm probability

From figure 3 we can get that the mean (μ) of random noise has a big influence on the watermark extraction and detection. In order to eliminate the bad influence of mean (μ) to the watermark detection, restriction is set in the process of watermark generation. According to the probability

$P\{w(k)=1\}=\frac{1}{2}, P\{w(k)=-1\}=\frac{1}{2}$, a series of watermark is created, the watermark which meets the inequality $\sum_{k=0}^{k=n-1} w(k) \approx 0$ is chosen.

Random noise can be expressed as $g(k)=\mu+g'(k)$, here

$$g'(k) \sim N(0, \sigma^2)$$

Then $c = \sum_{k=0}^{k=n-1} g(k) * w(k)$

$$c = \mu \sum_{k=0}^{k=n-1} w(k) + \sum_{k=0}^{k=n-1} g'(k) * w(k)$$

$$\sum_{k=0}^{k=n-1} w(k) \approx 0 \quad \Rightarrow \quad c = \sum_{k=0}^{k=n-1} g'(k) * w(k)$$

$$c = \sum_{k=0}^{k=n-1} g'(k) * w(k) \sim N(0, n\sigma^2) \qquad (6)$$

From (6) we can get that the random noise with mean μ and variance σ^2 is equal to the random noise with mean 0 and variance σ^2. So, the bad influence of mean (μ) to the watermark detection is eliminated by adding some restrictions to the process of watermark generation.

5. THE PROBABILITY OF MULTIPLE WATERMARK EMBEDDING

Some time, many watermarks are embedded into a same copy of data (ZHANG Fan,2007; Li Boya,2007). The problem is how to eliminate mutual influence of different watermark. Simple watermark extraction can not solve this problem; the autocorrelation detection can solve this problem, so watermark detection should be executed.

Suppose n watermark $W_k (k = 0,1,\cdots n-1)$ are embedded into a same copy of data. When detecting with W_i, the other watermarks can be treated as the overlap of n-1 random noises with mean 0 and variance $\lambda_k (k = 0,1,\cdots i-1,\cdots,i+1,n-1)$.

There is an example of embedded 10 watermarks into the same data, 100 random noises are used to perform autocorrelation detection, in which the 45^{th} ~ 54^{th} are the embedded watermarks and the others are pseudo random sequence. The figure 4 is autocorrelation detection result.

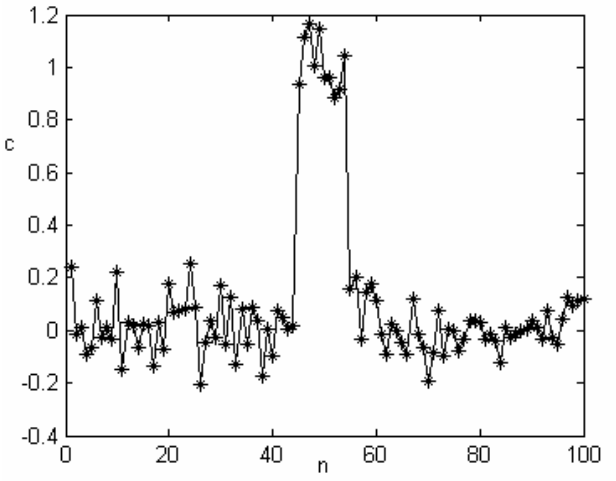

Figure 4. The detection result of multiple watermarks

6. CONCLUSIONS

Based on the Mahalanobis distance discriminant analysis in the statistics theory and he self-characteristic of watermarking techniques, this paper solve the problem of the setting of detection threshold and the calculating of false alarm and flow detection probability, get the relationship between the length of watermark、the watermark embedding strength and the attack strength , prove the probability of multiple watermark embedding, at the same time, a pseudo random sequence watermark generation way with restriction is proposed. Theory and experimentation testify that the research of this paper is very useful to the watermark generation and watermark detection.

REFERENCES

Li Boya,Han Guoqiang,Wo Yan,2007. A Watermarking Based on Chaotic Sequence and HVS. *Control & Automation*, 2007(21), pp.40-42

Li Yuanyuan,Xu Luping,2004.Vector Graphical Objects Watermarking Scheme in Wavelet Domain. *Acta Photonica Sinica*，33(1), pp. 97-100

Sun Shenghe, Lu Zheming，et al，2004. *Digital Watermarking Technology and its Applications*. Sience Press. pp.1-5

Wang Xuemin,1999. *The Applications of Multivariate Analysis.Shanghai University of Finace and Economics Press*,pp.134-148

YANG Cheng-song, ZHU Chang-qing,2007. Watermarking Algorithm for Vector Geo-spatial Data on Wavelet Transformation. *Journal of Zhengzhou Instituite of Surveying and mapping*,24(1),pp.37-39

Yang Yixian,Niu Xinxin，2006. *Theory and Applications of Digital Watermarking.Beijing*, Higher Education Press, pp.13-18

ZHANG Fan, LIU Ya-li, SU Yu-ting, ZHANG Chun-tian,2007. Multiple Watermarking and Capacity Analysis of Digital Image. *Journal of University of Electronic Science and Technology of China*,36(6), pp. 1325-1328

Zhong Shang-ping, GAO Qing-shi，2006. The Feasibility Analysis of Normalized-correlation -based Vector Maps Watermarking Detection Algorithm and the Improved Watermarking Algorithm. *Journal of Image and Graphics*，11(3), pp.401-408

ACKNOWLEDGMENTS

This work was funded by 863 project (2006AA12Z223).