Journal of Computer and Communications, 2014, 2, 121-126
Published Online July 2014 in SciRes. http://www.scirp.org/journal/jcc
http://dx.doi.org/10.4236/jcc.2014.29016

The Digital Fingerprinting Method for Static Images Based on Weighted Hamming Metric and on Weighted Container Model

Sergey Bezzateev, Natalia Voloshina

Information Security Technologies Department, Saint-Petersburg State University of Aerospace Instrumentation, Saint-Petersburg, Russia
Email: bsv@aanet.ru, natali@vu.spb.ru

Received May 2014

Abstract

The algorithm of fingerprint constructing for still images based on weighted image structure model is proposed. The error correcting codes that are perfect in weighted Hamming metric are used as a base for fingerprint constructing.

Keywords

Fingerprinting, Error Correcting Codes, Perfect Codes in Weighted Hamming Metric, Weighted Container Model, Digital Rights Management

1. Introduction

The author rights protection is one of the most important problems for the multimedia data distribution. Two kinds of methods are used to solve this problem such as digital watermarking (DWM) and digital fingerprinting (DFP) [1] [2]. In the DWM methods the additional information about author (watermark) is embedded into initial data (for example image) to protect it. This information should be resistant against wide class of attack (filtering, compression, etc.) on the watermarked image as long as the level of visual quality is higher than predefined level. The second type of protection means that the additional information (fingerprint) should be formed based on the initial information (image) based on its specific salient features. For example in [3] the edges and corners obtained from an image were used as salient features. The fingerprint is used then to find copies of original image including illegal ones. Fingerprinting method should provide stable fingerprint forming process for different kind of image processing (filtering, compression, etc.).

Main differences between these copy right protection methods are:

Some part of the initial object is changed while watermark embedding process (for example LSB method). Fingerprinting does not imply any changes of initial information.

Digital watermark is constructed by the user (author). It can be represented in different forms (ID information, picture, text, etc.) to prove author rights. Digital fingerprint does not contain any author information. It forms based on initial information.

How to cite this paper: Bezzateev, S. and Voloshina, N. (2014) The Digital Fingerprinting Method for Static Images Based on Weighted Hamming Metric and on Weighted Container Model. *Journal of Computer and Communications*, **2**, 121-126.
http://dx.doi.org/10.4236/jcc.2014.29016

The combination of these two types of author rights protection methods is proposed in this paper. The idea is to add small distortions to the initial image to get stable fingerprint.

In this paper the term salient feature refers to a special vector *b* in weighted Hamming metric obtained from the image and to "nearest" (L, G) code perfect in the weighted Hamming metric [4]. The "nearest" means a code that have codeword *a* on the minimal distance in weighted Hamming metric from the vector *b*.

2. Proposed Method Description

The author rights protection systems are developed in a way they give stable information about the person that has author rights on it. This information should be stable for wide class of distortions that could happen. At the same time author rights protection method should not change the initial information more than it defined for exact marking object. For example it should keep visual quality of the image.

Fingerprinting methods do not change initial objects. They only use salient features of these objects to construct fingerprint. The problem is that not always such features could be found or they could be not resistant to attacks. At the same time it is not necessary to avoid any distortions of initial information (for example for digital images). In this case it is possible to use watermarking approach that changes some part of initial information to embed additional author information. But it is not always possible to embed watermark with acceptable distortion level due to inconsistency of watermark and protected object structure.

The method of digital fingerprint construction based on the property of codeword presence in image (F5 concept [5]) is proposed. If the codeword can't be found then codeword should be added by the watermarking approach.

For example for the still images the essence of the method can be explained as following. The initial image or its part is transformed into bit stream. This bit stream is divided into blocks *a* with length *n*. These blocks *a* are looked at as a code words of code *G* with error vector *e*. If *a* is a codeword of *G* then $e=0$ and syndrome $s=0$. It means that this block has required for fingerprinting property. If $s\neq 0$ then this block does not have required property. It is necessary to decode this codeword i.e. to correct its errors to get this property. As a result the distortions will be added to the initial information. The problem of getting $s=0$ for any block of bit stream could be solved by perfect codes usage [4].

It is possible to use several codes with equal parameters that are defined for weighted Hamming metric. For this approach the code sequence G_i that is formed as a result of the blocks a_i decoding and for which the blocks are the "closest" to their code words can be looked at as a fingerprint.

The structure of the blocks *a* should be weighted because the error corrected codes that are concerted with the weighted Hamming metric are used to construct fingerprint. These weighted blocks could be looked at as a weighted container that is formed for weighted watermarking [6] [7].

It is necessary to dived initial image into several zones to construct weighted blocks *a*. The division is performed based on the influence of the errors that happened during the decoding process on the resulting distortions (visual quality of resulting image). Blocks *a* are formed as a combination of the parts of these different significance zones.

The basic scheme of the proposed fingerprinting method is shown in the **Figure 1**.

3. Error Correcting Codes Consistent with Weighted Hamming Metric

Error correcting codes in weighted Hamming metric are defined as following [4]:

Length *n* of a code word is $n_1, n_2, \cdots, n_l, n=\sum_{i=1}^{l} n_i$.

Any position of the code word $a=\left(a_1^1\, a_2^1 \cdots a_{n_1}^1\; a_1^2 \cdots a_{n_2}^2 \cdots a_1^l \cdots a_{n_l}^l\right)$ have the weight ω_j. Let us consider these weights as growing sequence *i.e.* $\omega_1<\omega_2<\cdots\omega_j<\cdots<\omega_l$. The weight of the code word *a* in weighted Hamming metric is defined as following:

$$wt_{WHM}(a)=\sum_{a_i^j\neq 0}\omega_j .$$

It is obvious that Hamming distance between two vectors in weighted Hamming metric can be defined as $a=\left(a_1^1\, a_2^1 \cdots a_{n_1}^1\; a_1^2 \cdots a_{n_2}^2 \cdots a_1^l \cdots a_{n_l}^l\right)$ and $b=\left(b_1^1\, b_2^1 \cdots b_{n_1}^1\; b_1^2 \cdots b_{n_2}^2 \cdots b_1^l \cdots b_{n_l}^l\right)$:

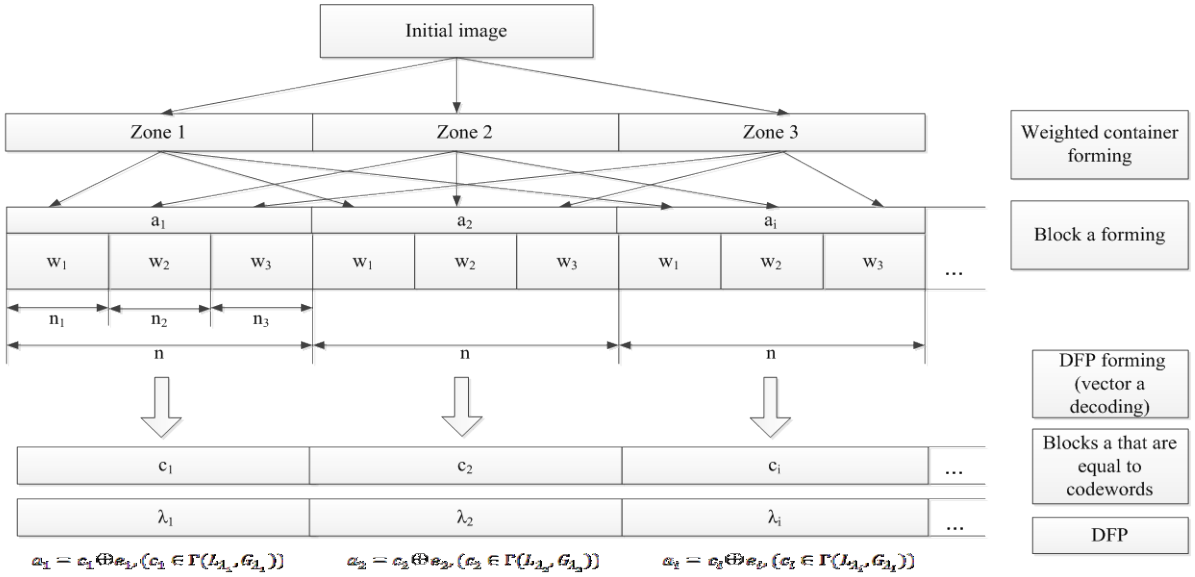

Figure 1. The basic scheme of the fingerprinting method for three zone container.

$$d_{WHM}(a,b) = \sum_{a_i^j \neq b_i^j} \omega_j .$$

For binary codes in weighted Hamming metric it is possible to write the Hamming distance for code with parameters n, M, d_{WHM} the same way as for codes in ordinary Hamming metric

$$M \cdot W_n^{\tau} \leq 2^n,$$

where M -number of a code words in the code (for binary linear code $M = 2^k$)

W_n^{τ} -number of vectors of length n in a sphere of radius τ in weighted Hamming metric.

$$W_n^{\tau} = \sum_{\sum_{j=1}^{l} \tau_j \omega_j \leq \tau} \prod_{i=1}^{l} \binom{n_i}{\tau_i},$$

where $\tau = \dfrac{d_{WHM} - 1}{2}$.

Therefore we can define a perfect code, *i.e.* the code with parameters lies on this bound. For linear binary code we obtain the following equation:

$$2^{n-k} = \sum_{\sum_{j=1}^{l} \tau_j \omega_j \leq \tau} \prod_{i=1}^{l} \binom{n_i}{\tau_i}.$$

Codes that are perfect in a weighted Hamming metric can be constructed by using well-known methods for optimal codes construction (for example Gilbert-Varshamov bound). But such construction method has high calculation complexity and doesn't give any design decoding method for such codes. By using special classes of Goppa codes in weighted Hamming metric it is possible to solve both these problems.

Special Class of Goppa Codes in Weighted Hamming Metric

To construct Goppa codes consistent with a weighted Hamming metric the construction of generalization (L,G) codes with position numerators of different degrees [4] is used. Therefore we use locator set

$L = \left\{ \dfrac{v_i^j(x)}{u_i^j(x)} \right\}_{j=1,l;\, i=1,n_j}$, where $v_i^j(x)$ is formal derivative of denominator $u_i^j(x)$, $\deg u_i^j(x) = \omega_j$ and $u_i^j(x)$

is an irreducible polynomial on $F_{2^m}[x]$. Goppa polynomial for such code is irreducible polynomial

$$G(x), G(x) \in F_{2^m}[x], \deg G(x) = \tau \geq \omega_l,$$
$$\gcd\left(G(x), u_i^l(x)\right) = 1, \forall i : i = 1, n_l$$

Number of different Goppa codes with the same parameters (n, k, d_{WHM}) from this class is determined by the number of different irreducible polynomials with degree τ on $F_{2^m}[x]$.

4. Fingerprint Calculation Algorithm Based on Family of Goppa Codes $\Gamma_{WHM}(n,k,d)$ Perfect in Weighted Hamming Metric

We describe here the simplest version of the fingerprint calculations algorithm for static image that contains different significance zones that is based on the family of Goppa codes perfect in the weighted Hamming metric.

Without loss of generality we consider here a static image that contains three zones.

- Third zone does not assume any distortions, *i.e.* $\omega_3 = \infty$.
- Second zone has relative weight ω_2 of distortions equal to 2.
- First zone allows to make the maximum number of distortions without major consequences for the quality of the resulting image and therefore has minimal relative weight $\omega_1 = 1$.

For such image let us chose a family of Goppa codes perfect in weighted Hamming metric $\Gamma_{WHM}(n,k,d)$ with a following parameters:

$$n = n_1 + n_2 = 2^{2m-1} + 2^{m-1} - 1, n_1 = 2^m,$$
$$n_2 = 2^{2m-1} - 2^{m-1} - 1,$$
$$k = 2^{2m-1} + 2^{m-1} - 2m, d_{WHM} = 5.$$

As mentioned in the previous section the number of different codes in this family is determined by the number of irreducible polynomials of second degree with coefficients from $GF(2^m)$:

$$I_{GF(2^m)}(2) = 2^{2m-1} - 2^{m-1}.$$

Obviously as we consider perfect codes any vector $a = \left(a_1^1\, a_2^1 \cdots a_{n_1}^1\, a_1^2 \cdots a_{n_2}^2\right)$ that is obtained from image that was preliminarily splitted into different significance zones can be represented in the weighted Hamming metric as following:

$$\left(a_1^1\, a_2^1 \cdots a_{n_1}^1\, a_1^2 \cdots a_{n_2}^2\right) = \left(c_1^1\, c_2^1 \cdots c_{n_1}^1\, c_1^2 \cdots c_{n_2}^2\right)_i \oplus \left(e_1^1\, e_2^1 \cdots e_{n_1}^1\, e_1^2 \cdots e_{n_2}^2\right)_i,$$

where

$$\left(c_1^1\, c_2^1 \cdots c_{n_1}^1\, c_1^2 \cdots c_{n_2}^2\right)_i \in \Gamma\left(L_i, G_i\right),$$
$$wt_{WHM}\left(e_1^1\, e_2^1 \cdots e_{n_1}^1\, e_1^2 \cdots e_{n_2}^2\right)_i = t_i \leq 2$$

Thus using for vector *a* various $\Gamma(L_i, G_i)$ codes of the family $\Gamma_{WHM}(n,k,d)$ in decoding procedure we will get error vectors with different weights t_i, $0 \leq t_i \leq 2$.

Fingerprint will consider as a set of numbers $\lambda_1, \lambda_2, \cdots, \lambda_R$ of a codes from the family $\Gamma_{WHM}(n,k,d)$ corresponding to the vectors obtained from the image splitting into zones as follows:

Vector $a = (a_1^1\, a_2^1 \ldots a_{n_1}^1\, a_1^2 \ldots a_{n_2}^2)$ assign the smallest number λ of such $\Gamma(L_i, G_i)$ code for which error vector $e = (e_1^1\, e_2^1 \ldots e_{n_1}^1\, e_1^2 \ldots e_{n_2}^2)_i$ provides the least weight by decoding procedure.

That means

$$\lambda = \min_i i : \; wt_{WHM}\left(e_1^1\, e_2^1 \cdots e_{n_1}^1\, e_1^2 \cdots e_{n_2}^2\right)_i$$
$$= \min_j wt_{WHM}\left(e_1^1\, e_2^1 \cdots e_{n_1}^1\, e_1^2 \cdots e_{n_2}^2\right)_j.$$

The algorithm of fingerprint calculation based on a family of $\Gamma(L_i, G_i)$ codes can be described as:

1) The image is divided into zones of different significance.

2) In accordance with found image zones the blocks are formed and the appropriate parameters of the codes

family are selected.

3) Fingerprint $\lambda_1, \lambda_2, \cdots, \lambda_M$ is formed in following way. Found error vectors $e = (e_1^1\, e_2^1 \ldots e_{n_1}^1\, e_1^2 \ldots e_{n_2}^2)_i$ are corrected and corresponding blocks $a = (a_1^1\, a_2^1 \ldots a_{n_1}^1\, a_1^2 \ldots a_{n_2}^2)$ of source image are converted to code words of the Goppa codes from the family $\Gamma_{WHM}(n,k,d)$. For example the first block is transformed into a codeword $(c_1^1\, c_2^1 \ldots c_{n_1}^1\, c_1^2 \ldots c_{n_2}^2) \in \Gamma(L_{\lambda_1}, G_{\lambda_1})$. The distortion of the source image will be minimal according to the previously described algorithm of obtaining the numbers λ_i.

5. The Fingerprint Checking and Its Resistance to Random and Deliberate Distortions

In accordance with the described algorithm of fingerprint $\lambda_1, \lambda_2, \cdots, \lambda_R$ creation in the processed image *P* there are *R* blocks

$$\left(c_1^1\, c_2^1 \ldots c_{n_1}^1\, c_1^2 \ldots c_{n_2}^2\right)_i \in \Gamma\left(L_{\lambda_i}, G_{\lambda_i}\right),$$
$$i = 1, \cdots, R.$$

Each block should be the codeword of corresponding Goppa code from the family $\Gamma_{WHM}(n,k,d)$. The presence of the fingerprint $\lambda_1, \lambda_2, \ldots, \lambda_R$ is verified by the decoding the blocks *a* by the corresponding code. The fingerprint is decided to be found if all syndromes for all blocks $s = 0$.

The presence of minor distortion in the processed image *P* may cause errors in these blocks. So in distorted image P* there are blocks

$$\begin{aligned}\left(b_1^1\, b_2^1 \cdots b_{n_1}^1\, b_1^2 \cdots b_{n_2}^2\right)_i &= \left(c_1^1\, c_2^1 \cdots c_{n_1}^1\, c_1^2 \cdots c_{n_2}^2\right)_i \oplus \left(e_1^1\, e_2^1 \cdots e_{n_1}^1\, e_1^2 \cdots e_{n_2}^2\right)_i \\ &= \left(\hat{c}_1^1\, \hat{c}_2^1 \cdots \hat{c}_{n_1}^1\, \hat{c}_1^2 \cdots \hat{c}_{n_2}^2\right)_i \oplus \left(\hat{e}_1^1\, \hat{e}_2^1 \cdots \hat{e}_{n_1}^1\, \hat{e}_1^2 \cdots \hat{e}_{n_2}^2\right)_i,\end{aligned}$$
$$wt_{WHM}\left(\hat{e}_1^1\, \hat{e}_2^1 \cdots \hat{e}_{n_1}^1\, \hat{e}_1^2 \cdots \hat{e}_{n_2}^2\right)_i = wt_i \le 2,$$
$$\left(c_1^1\, c_2^1 \cdots c_{n_1}^1\, c_1^2 \cdots c_{n_2}^2\right)_i,$$
$$\left(\hat{c}_1^1\, \hat{c}_2^1 \cdots \hat{c}_{n_1}^1\, \hat{c}_1^2 \cdots \hat{c}_{n_2}^2\right)_i \in \Gamma\left(L_{\lambda_i}, G_{\lambda_i}\right), i = 1, \cdots, R.$$

Using fingerprint $\lambda_1, \lambda_2, \ldots, \lambda_R$ and corrupted image P* it is easy to find an appropriate error vector $(\hat{e}_1^1\, \hat{e}_2^1 \ldots \hat{e}_{n_1}^1\, \hat{e}_1^2 \ldots \hat{e}_{n_2}^2)_i$ and the weight wt_i corresponding to this vector. Obviously that in case of small corruptions $(c_1^1\, c_2^1 \ldots c_{n_1}^1\, c_1^2 \ldots c_{n_2}^2)_i = (\hat{c}_1^1\, \hat{c}_2^1 \ldots \hat{c}_{n_1}^1\, \hat{c}_1^2 \ldots \hat{c}_{n_2}^2)_i$ and therefore $(e_1^1\, e_2^1 \ldots e_{n_1}^1\, e_1^2 \ldots e_{n_2}^2)_i = (\hat{e}_1^1\, \hat{e}_2^1 \ldots \hat{e}_{n_1}^1\, \hat{e}_1^2 \ldots \hat{e}_{n_2}^2)_i$.

Let's define now the penalty function $F = \sum_{i=1}^{R} wt_i \cdot f_{wt_i}$. In the simplest case it is possible to consider that the weight coefficients $f_{wt_i} \in \{f_1, f_2\}$ of the penalty function have the same value and are equal to 1. However the more flexible way where $f_1 \neq f_2$ is possible too. By taking into account the nature of the corruptions and their impact on the perception of the resulting image *i.e.* its visual quality we can chose different values of the weight coefficients.

Decision of the definite fingerprint availability in an existing picture P* is made by the value of the penalty function *F* and threshold values *B*1 and *B*2 that are define the events of the "false alarm" and "skip goal" respectively.

6. Conclusion

A new method of fingerprint construction that uses the features of the original image structure (weighted container) associated with the different degree of its sensitivity to the corruptions in different zones of the container is proposed. In order to use this property of the container it is proposed to use the weighted Hamming metric unlike used in previously known schemes usual Hamming metric. For effective use of this metric it is proposed a family of generalized Goppa codes perfect in weighted Hamming metric. Thanks to the usage of such family of codes it is possible to make the minimum number of distortions when fingerprint is constructed. Additionally the use of error-correcting codes construction allows to correct random corruptions that can occurred during container (image) storing or transmitting. Still open question in optimal choice of the coefficients in the penalty

function provides the minimal probability of “false alarm” and “skip goals” when deciding on the availability of the fingerprint presence in the analyzed image.

References

[1] Duric, Z., Johnson, N.F. and Jajodia, S. (1999) Recovering Watermarks from Images. Technical Report ISE-TR-99-04, Center for Secure Information Systems, George Mason University.

[2] Johnson, N.F., Duric, Z. and Jajodia, S. (1999) A Role for Digital Watermarking in Electronic Commerce. Accepted for Publication ACM Computing Surveys. Publication TBA.

[3] Johnson, N.F., Duric, Z. and Jajodia, S. (1999) On Fingerprinting Images for Recognition. In: *Multimedia Information Systems*, 4-11.

[4] Bezzateev S. and Shekhunova N. (2012) Binary Generalized (L, G) Codes That Are Perfect in a Weighted Hamming Metric. *Problems of Information Transmission*, **48**, 239-242. http://dx.doi.org/10.1134/S0032946012030039

[5] Westfeld, A. (2001) High Capacity Despite Better Steganalysis (F5-A Steganographic Algorithm). In: Moskowitz, I.S., Eds., *Information Hiding*. 4*th International Workshop. Lecture Notes Computer Science*, Vol. 2137, Springer-Verlag, Berlin, Heidelberg, New York.

[6] Bezzateev, S., Voloshina, N. and Zhidanov, K. (2012) Special Class of Error-Correcting Codes for Steganography Systems. *Proceedings TUSUR* (*Novosibirsk*), **1**, 112-118.

[7] Neeta, D. (2006) Implementation of LSB Steganography and Its Evaluation for Various Bits. Digital Information Management, 1*st International Conference.*