A new adjustable blind Watermarking based on GA and SVD

Hamed Modaghegh, Hossein Khosravi R., Mohammad- R. Akbarzadeh-T
Department of Electrical Engineering, Ferdowsi University Of Mashhad , Iran

Abstract

As information technology and multimedia products become more and more readily available, copyright and other related legal topics become more and more significant. Embedding copyright information as hidden data into the multimedia product –named watermarking- is one of the methods to protect owner rights. Two main concepts in watermarking are imperceptibility and robustness of the watermark. A tradeoff between these two features exists, which can be introduced as an optimization problem. Genetic Algorithm (GA) is applied to solve this optimization problem. In this paper, a new adjustable watermarking method based on singular value decomposition is presented so that SVD parameters are adjusted by using the GA considering image complexity and attack resistance. The proposed watermarking method is also an adjustable solution, so that by changing fitness function (cost function), watermarking method can be converted to each of robust, fragile, or semi-fragile types. Simulation results show that the proposed method has better results from the case where watermarking parameters are adjusted by the user empirically.

Keywords: SVD, blind Watermarking, GA.

1. Introduction

Watermarking was first introduced 700 years ago [1].It was made by impressing a water-coated metal stamp onto the paper during manufacturing. Watermarks have been used by papermakers to identify special products by companies. In general, watermarking is a method for hiding special information (watermark) within cover data in order to save the author ownership [2].

In order to protect the rights of the owner against various attacks, we should insert owner information in the image main parameters; One of these main parameters is singular values [3]. H.zer, et al exploited image singular values with singular value decomposition technique (SVD), and embedded watermark data, then rebuilt the image.

After this presentation several papers used and optimized the technique [4,5], but these techniques suffer from false-positive drawback (the method, recognizes non watermarked image as watermarked).

R. Sun,et al[6] presented another method that solved the problem. The method does not use the original image and is so-called blind. Embedding watermark in image main parameters causes changes in the image, and more information these parameters carry, yields more changes in image. So, apart from the method we used to embed watermark in the image, there exists a tradeoff between perceptibility and robustness against attacks. This tradeoff results in an optimization problem that will be solved in three different conditions. Based on the following three conditions, we calculate optimized parameters for applying watermark using Genetic Algorithm (GA) optimization procedure.

- Maximum robustness and Minimum perceptibility. (Robust Watermarking)
- Minimum robustness and Minimum perceptibility. (Fragile Watermarking)
- Minimum robustness to some attacks while maximum robustness to other attacks and Minimum perceptibility. (Semi-Fragile Watermarking)

In this paper, we exploit image singular values by using SVD technique, calculate optimized parameters for applying watermark by using GA and embed watermark in the image.

Adjustable watermarking method with minimum perceptibility is attained by mixing the SVD and GA techniques.

SVD and GA are introduced in section 2. The proposed solution that uses SVD and GA is presented in section 3; in section 4, simulation results of proposed method within three conditions: Robust, Fragile, Semi-Fragile are compared to simulation results based on empirical adjustment, and we finally offer a conclusion in section 5.

2. Background

2.1. Different Watermarking methods

Considering the amount of robustness to the changes, watermarking techniques can be divided into three categories: [2].

Fragile watermarking: This type of watermarking is destroyed upon small changes in the image, so it is appropriate to review image authentication. Because the smallest change in the watermarked data, will destroy the watermark and the recipient will be aware of the inserted change.

Robust watermarking: This type of watermarking is usually used to apply author ownership on multimedia data and is designed such that it could be robust against small changes, and watermark information could not be easily destroyed, and more, make it feasible to recognize watermark owner even after changes.

Semi-Fragile watermarking: This type of watermarking is fragile to some certain changes and is robust to others. As an example, since white noise exists in every transmission channel, it is better to choose a watermarking method that is robust against white noise and is fragile to other changes, which usually intend abuse.

On the other hand, considering retrieval method, we can categorize watermarking methods to blind and non-blind. In order to retrieve watermark data in non-blind methods, in addition to watermarked image, we also need original image**.** These methods are more exposed to false-positive drawback. In blind retrieval methods, there is no need to original image and watermark can be extracted only from watermarked image. Since the original image is not used in these methods, false-positive problem never occurs. Thus, blind methods have more applications than non-blind methods.

2.2. Watermarking methods by using SVD

Generally for every matrix there exists a decomposition in the form (1), Which is called singular value decomposition where U and V are unitary matrixes so that:

$$A_{MN} = U_{MM} * D_{MN} * V'_{NN} \quad (1)$$

$$I = U * U' \quad (2)$$

$$I = V * V'$$

The matrix D is a diagonal matrix where the diagonal entries are singular values of matrix A. Many methods are proposed, which have embedded watermark by SVD. One of the first was presented by H.zer, et al[3]. By using SVD technique, it first decomposed the matrix in the form of (1).

$$A = U * D * V^T \quad (3)$$

Watermark data (m-by-n matrix) is multiplied to a constant coefficient α and the result is compounded with matrix D.

$$D' = D + aW \quad (4)$$

α shows strength of watermark insertion in the image block (4).). We then decompose this matrix to three matrixes by SVD technique one more time, then reconstruct watermarked image using matrixes Dw, U, and V. (as shown in (5-6)).

$$D' = U_W * D_W * V_W^T \quad (5)$$

$$A_W = U * D_W * V^T \quad (6)$$

Watermark retrieval stages are as follows:

Suppose that matrix $A'w$ is the watermarked image so that some changes are applied by the attacker. First we decompose A'_W matrix by SVD and obtain D'_W:

$$A' = U' * D'_w * V'^T \quad (7)$$

Then we multiply this matrix by U_W and V_W to make D''. Watermark is calculated by subtracting matrix D from D'':

$$D'' = U_W * D'_W * V_W^T \quad (8)$$

$$W' = \frac{D'_W - D}{a} \quad (9)$$

Veysel, Aslantas [4] tried to improve the procedure by merging this method with Genetic Algorithm (GA). The proposed method was robust and non-blind where parameter "a" was calculated empirically by user. The result had very good resistance to attacks, but the main drawback was that the response to non-watermarked images was also positive and declared non-watermarked images as watermarked ones. This drawback is subject to the watermarking method, that is non-blind and to the watermark data that is too low. In this method, major watermark data are calculated by retrieval of matrixes U_W and V_W. These two matrixes are better to be calculated from attacked watermarked image A` rather than original image A.

Other methods were proposed to solve the problem; R.Sun, et al[6] used blocked SVD. Image was fractioned into 8-by-8 simpler blocks. Then SVD is executed for these blocks and we only work on the first element that is the maximum singular value of the matrix. In order to get minimum changes in the image, this number is quantized, as shown below:

$$Z = D(1,1)\ mod\ Q \quad (10)$$

if w = 0 then:

$$z < \frac{3Q}{4} \ \rightarrow\ D(1,1)' = D(1,1) + \frac{Q}{4} - z \quad (11)$$

$$otherwise\ \rightarrow\ D(1,1)' = D(1,1) + \frac{5Q}{4} - z$$

and if w=1:

$$z < \frac{3Q}{4} \ \rightarrow\ D(1,1)' = D(1,1) - \frac{Q}{4} + z \quad (12)$$

$otherwise \rightarrow D(1,1)' = D(1,1) + \frac{3Q}{4} - z$

Q is an arbitrary value to quantize D(1,1)**.** As Q increases, robustness of watermark and so, changes in the image are increased. After applying watermark data to the maximum singular v alue, the watermarked image can be calculated simply by multiplying these 3 matrices:

$$A = U * D' * V^T \quad (13)$$

These steps are taken for other blocks, and as previously mentioned, SVD value is calculated for them and other watermark bits are inserted into maximum singular value. In receiver, we only have to fragment the image into 8-by-8 blocks and run SVD for each of them. Watermark is extracted using:

$$Z = D(1,1)' \ mod \ Q \quad (14)$$

If z<Q/2 then w=0, otherwise w=1.

So all watermark bits can be retrieved from the image; the procedure covered the stated problem and responds correctly to non-watermarked images.

3. Proposed Algorithm (using GA to optimize watermarking algorithm)

The method introduced by R.sun, et al [6] (unlike other proposed methods) does not need the main image in order to extract watermark data (blind). So the problem that occurred in the first method [3] may not happen in this method, and all information can be exploited from watermarked image. In these methods, knowing Q is enough to extract watermark but, using a constant Q, to quantize all blocks, remains as a drawback.

As we know, different image blocks do not have same statistic properties, some have high complexity and some not. The complexity of image block directly influences singular values and increases first singular value. From the other hand, perceptibility of inserted changes, in blocks with higher complexity is lower than blocks with simple details. So it is better to reduce the value of inserted changes in singular values for blocks with lower image pixels complexity. In order to reduce perceptibility of the change, value of Q should be small. In blocks with more complexity, we can select larger values of Q, so that perceptibility of watermark would be still low.

Main disadvantage is absence of a uniform relation between value of Q and perceptibility; because several nonlinear elements in the procedure exist. All these elements have led to lack of ability in computation of optimized Q for each block with analysis. In addition to perceptibility, another factor that is effective in selecting Q, is robustness against attacks. As mentioned in 2-1-2, considering robustness of watermark against attacks, three watermarking types were introduced, so based on the method that we want to implement, we should minimize or maximize robustness. By changing Q, robustness could be changed against certain attack. There is not a simple linear relation between Q and robustness, and the relation depends on different criteria like type of the attack and image block complexity; so that it is feasible that a particular Q causes high robustness for a certain block but the same Q results in low robustness in another block.

Recognizing which block is more affected or less affected by the attack is a difficult task, and it will be different with the specifications of each image. Therefore, the diagnosis of this action is better to be carried out for each image and each attack separately. As we see, obtaining the appropriate Q, to gain maximum or minimum robustness against attacks and minimum perceptibility, can be viewed as an optimization problem. In order to resolve this problem –that has N input variables (N is total number of image blocks)-, we use GA algorithm.

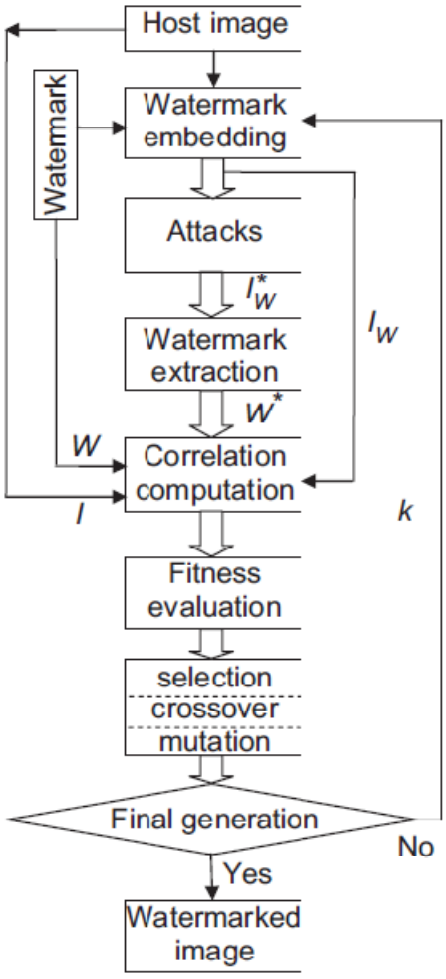

Figure 1. Watermarking algorithm block diagram.

We first produce an initial random population, so that each element of this population (chromosome) consists of N variables, each with B bits; and each of these variables generates value of Q for one block, so we produce a random matrix with dimensions: $N_{pop} * N * B$. We use this matrix to calculate Q for each block and insert $W_1 \times N$ watermark using calculated Q in the block. After obtaining watermarked image, we apply some specific attacks on the image to survey watermark robustness. Then extract watermark from attacked image and compare with original watermark. Our criteria for the amount of perceptibility, is the PSNR[1] criterion, and to obtain signal to noise value it`s enough to compare original image with watermarked image.

Fitness function should be selected in such a way that includes both perceptibility and robustness criterions. In order to gain a robust watermarking method, we should consider minimum perceptibility and maximum robustness. To provide these two conditions, we can define fitness function in various forms. Some of them are as follows:

$$\frac{1}{f_i} = \frac{1}{\frac{1}{t}\sum_{i=1}^{t} corr_w(w, w_i^*)} - corr_I(I_w, I) \quad (15)$$

$$f_i = corr_I(I_w, I) + \frac{1}{t}\sum_{i=1}^{t} corr_w(w, w_i^*) \quad (16)$$

$$f_i = \frac{corr_I(I_w, I)}{\frac{1}{t}\sum_{i=1}^{t}(1 - corr_w(w, w_i^*))} \quad (17)$$

Which f_i is fitness function of i-th population element, t represents number of attacks, W^*_i represents extracted watermark from I_w image and function *corr* represents correlation between the two input matrix. The less $corr_I$ function value, the less perceptibility; and the more $corr_W$, the more robustness. Among fitness functions defined, the third, converges faster and results suitable Qs.

But in order to gain a fragile watermarking method, perceptibility and robustness, both should be minimized so that watermark data may not be extracted by small change in watermarked image. Thus we should define fitness function so that maximizes $corr_I$ and minimizes $corr_W$. We have defined fitness function as follows:

$$f_i = \frac{corr_I(I_w, I)}{\frac{1}{t}\sum_{i=1}^{t} corr_w(w, w_i^*)} \quad (18)$$

[1] Peak Signal to Noise Ratio

As mentioned in section (2), semi-fragile watermarking methods, are robust to some certain attacks and are fragile to others. Therefore the fitness function must be defined so that maximizes $corr_W$ for some attacks and minimizes for others. We use the following fitness function:

$$f_i = \frac{corr_I(I_w, I)}{\frac{1}{t_1}\sum_{i=1}^{t_1} corr_w(w, w_i^*) + \frac{1}{t_2}\sum_{i=1}^{t_2}(1 - corr_w(w, w_i^*))} \quad (19)$$

After calculating fitness function for each population element, considering f_i, we sort them and crossover operator is executed. Mutation operator is executed to generate simple changes and to reconstruct population. The procedure repeats for other populations.

4. Simulation Results

MATLAB software is used for paper simulation and Q is limited between 40 and 100. The attacks used in simulation are Gaussian filter attack (window size 5 and variance 1), mean filter attack (window size 3), noise attack (amplitude 4) and move camera attack (3 pixels with angle 45). In order to review performance of GA algorithm, the software is investigated in 2 stages and optimized GA parameters for applying watermark are obtained.

In the first stage, fitness function value is calculated for population number 50, crossover rate 0.5, number of generation 60 and different values of mutation rate. The best result is obtained where mutation rate is 0.03.

Then in second stage fitness function value is calculated for mutation rate 0.03, same number of population, different crossover rate values. The best result is obtained where crossover rate is 0.5

Table 1. Sample Obtained Q, from best chromosome for first blocks of image in last generation

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	65	93	86	97	85	95	99	93
2	70	48	47	57	80	43	41	46
3	76	98	65	89	95	62	91	75
4	81	98	47	47	57	51	65	69
5	49	72	49	74	63	42	41	46
6	59	61	81	87	97	41	56	46

. Calculated coefficients are shown in Table 1, using 50 for population number, 0.5 for crossover rate, and 0.03 for mutation rate. As shown, Table 1 consists of different coefficient values with average of 70. Obtained values of Q have high dispersion and this is happened as a result of differences between image blocks.

In table 2, results generated using GA are compared to the case with constant value of Q. As we see, the

image obtained from GA method has minimum perceptibility and maximum robustness compared to the results when Q is selected empirically. Since $corr_W$ is unequal to 1, we have some errors in watermark retrieval. In order to remove the error, we use a coding method that corrects errors ($corr_W$ =0.63 equivalent to error=%18.5)

Table 2. calculated result using GA (Robust watermarking), compared with constant Q.

	Image correlation	correlation between W, W'				fitness function
		Camera move	average	guassian	noise	
GA	0.99953	0.66602	0.71875	0.63086	0.76367	3.2752
Q=71.3	0.99955	0.59961	0.65039	0.58789	0.79297	2.9202
Q= 50	0.99979	0.52344	0.53906	0.47656	0.50781	2.0476
Q= 60	0.9997	0.56445	0.5957	0.51758	0.66016	2.4059
Q=80	0.99943	0.62695	0.67969	0.61328	0.84961	3.2489

In table 3, simulation results of fragile watermarking are shown. Fitness function (18) is used to simulate this method. Similar to Table 2, results are compared to the case with constant value of Q. As we see GA calculate Q values so that watermark has shown minimum robustness.

Table 3. calculated result using GA (Fragile watermarking), compared with constant Q.

	Image correlation	correlation between W, W'				fitness function
		Camera move	average	guassian	noise	
GA	0.99956	0.51758	0.59961	0.52539	0.69141	1.7131
Q=68.6	0.99957	0.60352	0.63867	0.57031	0.78125	1.5415
Q=60	0.99969	0.56445	0.58398	0.51172	0.67383	1.7131
Q=80	0.99942	0.61914	0.68359	0.62891	0.85938	1.4323

Table 4. calculated result using GA (semi-fragile watermarking), compared with constant Q.

	Image correlation	correlation between W, W'				fitness function
		Camera move	average	guassian	noise	
GA	0.99951	0.55469	0.68164	0.64063	0.78906	2.77
71.9	0.99955	0.62109	0.65039	0.59961	0.79883	2.543
50	0.99979	0.50781	0.55469	0.48828	0.50781	2.0435
60	0.99967	0.58203	0.60742	0.55078	0.67188	2.2824
80	0.99943	0.62109	0.69141	0.63672	0.86133	2.7924

At last, simulation results of semi-fragile watermarking are shown in Table 4. In this simulation, the goal is robustness against noise attack and camera move. On the other hand, we want the watermarking method to be fragile against Gaussian filter attack and mean filter attack**.** Fitness function: (19) is used to simulate this method.

The only problem we are faced with, is the existence of error in watermark retrieval for noise and camera move attacks. Here, like before we should use coding to correct the error (CorrW=0.78 equivalent to error=%11). It should be noted that, applied coding method should be able to correct up to %16 of error; because if it corrects more than this value, the method will be robust against Gaussian and mean filter attacks so.

5. Conclusion

As we saw in this paper, using GA, helped us to adjust watermark appliance parameters –in presence of attacks- such that we obtained maximum efficiency; more clearly we obtained minimum image change while having maximum or minimum robustness against attacks.

Another advantage of presented method is that watermark applied parameters are optimized with respect to each specific attack and this gives the designer, the ability to generate semi-fragile watermarking algorithms with desired robustness against different attacks.

Further researches could be done in the field by applying better criteria (better then PSNR) with more proximity to human vision.

6. References

[1]Greg Kipper , *Investigator's Guide to Steganography* , Auerbach Publications, 2004.

[2] Stefan Katzenbeisser and Fabien A. P. Petitcolas , *Information Hiding Techniques for Steganography and Digital Watermarking* , Artech House, Inc, 2000.

[3] H. zer, B.Sankur, and N.Memon, "An SVD-Based Audio Watermarking Technique," *proceedings of the 7th workshop on Multimedia and security '05, ACM Press*, August 2005.

[4] Veysel Aslantas, " A singular-value decomposition-based image watermarking using`genetic algorithm", *Int. J. Electron. Commun.* (AE-) 62 (2008).

[5] F. H. Huang, Z.H. Guan, "A Hybrid SVD-DCT Watermarking Method Based on PSNR", *Pattern Recognition Letter*, 2004, pp.1769-1775.

[6] R. Sun, H. Sun, and T. Yao, "A SVD-and quantization based semi-fragile watermarking for image authentication," *in Proc. Int. Conf. Signal Processing (ICSP)*, vol. 2, pp. 26-30, 2002 .