GENETIC PERCEPTUAL SHAPING: UTILIZING COVER IMAGE AND CONCEIVABLE ATTACK INFORMATION DURING WATERMARK EMBEDDING

ASIFULLAH KHAN, ANWAR M. MIRZA
Faculty of Computer Science & Engineering,
Ghulam Ishaq Khan (GIK) Institute of Engineering Science & Technology,
Topi-23460, Swabi, PAKISTAN
{akhan, mirza}@giki.edu.pk

Abstract

We describe a new watermarking scheme based on intelligent shaping of a digital watermark using Genetic Programming (GP). The proposed method, in addition to achieving a superior tradeoff between watermark robustness and imperceptibility, is also able to structure the watermark in accordance with an anticipated attack. This has been achieved by simultaneously hiding the watermark as well as spreading and fusing it in such a way to resist the conceivable attack. Robustness versus imperceptibility tradeoff and increase in bit correct ratio after attack, have been employed as the optimization criteria in the GP search. The concept of bonus fitness has been used to implement multi-objective fitness based GP evolution. Experiments on standard images indicate that such watermark shaping functions could be developed that are cover image independent and enhance imperceptibility. They offer high resistance against removal and interference attacks of Checkmark benchmark.

***Keywords**:* Watermarking, Genetic Programming, Perceptual Model, Discrete Cosine Transform (DCT), Bit Correct Ratio (BCR), JPEG and Human Visual System (HVS).

1. Introduction

Due to the rapid growth in the use of digital media, there is an increasing concern about unauthorized handling, copying and reuse of information. Watermarking which is being considered as the practice of imperceptibly altering data to embed a message about that data, is an effective way to counter these types of problems [1]. Digital watermarking is performed upon a variety of different digital materials, like audio, images, text, movies and 3D models. It has also a broad range of applications, like ownership assertion, authentication, broadcast monitoring, and integrity control [2]. In a watermarking system, there is an intrinsic relation between two of its most important, but contradicting properties: robustness and imperceptibility. Imperceptibility means that the watermarked data should be perceptually equivalent to the original, unwatermarked data. On the other hand, robustness means that the watermark should not be rendered undetectable, unless damaging the usefulness of the cover data itself [3]. If we try to improve the watermark imperceptibility, robustness decreases and vice versa. Consequently, one needs to make a tradeoff according to the application domain. For this purpose, different methods, both in spatial as well as transformed domain, have been used to tailor a watermark according to the cover image [4-11].

Watermarks are rendered undetectable with an attack, where the attack is defined as any processing of the watermarked data that might damage the watermark [1, 3]. Thus watermarking can be viewed as a reliable mode of communication to transfer important information (i.e. a watermark) embedded in a signal (e.g. a cover image) safely through a hostile environment [12]. Attacks can be intentional such as watermark estimation using Wiener filtering or unintentional such as JPEG compression. An extensive list of attacks appears in [1, 13-17].

Due to the nature of diverse types of attacks, there is no generic watermarking scheme that could resist all sorts of attacks. However, it can be assumed that many applications are not concerned with all conceivable attacks, but with specific attacks that might occur before decoding [1]. Investigators have addressed this problem in various ways. One way is to develop watermarking approaches suitable for the anticipated attack [18]. For example, in case of rotational attack, alteration in the phase, rather than the amplitude of the Fourier component, is performed to embed a watermark [19]. Another possibility is to achieve robustness against the probable processing of the watermarked image, by restructuring the watermark. In this scenario, robustness is often achieved at the expense of imperceptibility, computational cost, data payload, or even robustness to some other processing.

To defend attacks, efforts have been made to increase robustness at low cost of imperceptibility. For instance Jonathan et al. [3] have taken a theoretical approach to answer the complex question of "how should a watermark be structured to maximize its robustness". They have proposed that the watermark power spectrum should be proportional to that of the original signal. Liang et al. [21] propose robust watermarking using robust coefficients for embedding. Huang et al. [8, 20], on the other hand, have used Genetic Algorithms for the selection of coefficients to be altered for watermark embedding. However, these efforts concentrate on tailoring just the choice of specific coefficients, not the whole watermark, to a cover image and intended attack. In fact, they are not using perceptual models; rather a fixed strength of the alteration is used for each selected DCT coefficient.

Perceptual models [23-26], as those of Watson's, which have been frequently used in image compression are used to compute the strength of the alteration for each selected coefficient. These perceptual models make a tradeoff between robustness and imperceptibility according to the cover image. However, they do not take into consideration the watermark application and thus the intended attacks. For instance, when the watermarked image is expected to be JPEG compressed, it is judicious to structure the watermark in view of the JPEG compression. Pertinent examples exist in literature [27], where appropriate watermarking approaches as well embedding domains have been studied to achieve robustness against JPEG compression.

One way to restructure a watermark in view of the anticipated attack is to keep high watermark strength for those selected coefficients that are less affected by the attack. However, firstly this requirement needs to consider limitations imposed by imperceptibility. Secondly, this requirement varies for different types of attacks. Consequently, our aim in this work is to propose and study an automatic system that can restructure the watermark in accordance to the cover image and intended attack. Specifically, to propose a system for developing suitable watermark shaping functions, which are image independent and intended attack-resistant.

We address these requirements through the following contributions:

1. We consider the perceptual shaping of a watermark to be vital, not only for imperceptibility enhancements, but we realize it to be a method of structuring the watermark in accordance to the anticipated attack.
2. We introduce the concept of developing complex and appropriate watermark shaping functions from the existing ones. Specifically, we consider Watson's perceptual model, characteristics of the HVS and information about the distortion caused by the anticipated attack, as independent variables and genetically search for application-specific watermark shaping functions.

The idea used is analogous to combining classifiers for developing complex, but appropriate classifier for a certain application of pattern recognition [28]. We call this technique as Genetic Watermark Shaping Scheme (GWSS) and the genetically developed watermark shaping functions as Genetic Watermark Shaping Functions (GWSF).

In section 2, we discuss perceptual shaping of a digital watermark including discussion about perceptual models. We discuss attacks and their countermeasures in section 3, while imperceptibility and robustness measures in section 4. Section 5 explains our proposed technique GWSS. This includes description of various modules of GWSS and explains our bonus fitness idea used in the multi-objective based GP evolution. This section also describes the testing and comparison phase of the evolved GWSF. Section 6 presents implementation details and section 7 gives results and discussions. Conclusion and future work are discussed at the end.

2. Perceptual Shaping of a Digital Watermark

A watermark is generally embedded in a cover image with a high strength in areas where it is well hidden and with a low strength in places where it is clearly discernible. This type of strategy is known as perceptual shaping of a watermark [1]. For this purpose, usually perceptual models that are used in compression techniques are employed. These perceptual models are able to learn the content of a cover image by exploiting the sensitivities/insensitivities of an HVS. They take advantage of frequency sensitivity models that are based on viewing conditions as well as the cover image dependent, luminance sensitivity and contrast masking effects. Frequency sensitivity describes the HVS sensitivity to sine wave gratings at different spatial frequencies and depends only on the surrounding conditions. Luminance sensitivity on the other hand, is a measure of the effect of detectability threshold of a signal on a constant background. It depends on the average luminance value of the background as well as on the signal's luminance level. In block-based DCT case, the DC coefficient of each block dictates the luminance sensitivity for that block. The third important property of HVS that is exploited for hiding a watermark is the contrast masking. It represents the detectability of one signal in presence of another signal. This masking (hiding) effect increases when the masking signal and the signal to be masked have same spatial frequency, orientation and location. In block-based DCT, the AC coefficients dictate this behavior.

Watson's Perceptual Model (WPM) [23, 24] has been used in JPEG compression and watermark shaping [4-6]. This model is based on DCT domain and has been originally proposed by Ahumada et al. [29]. On the other hand, Lambrecht et al. [25] have proposed a perceptual model that is based on Gabor filters. Watson et al. [26], have also developed a wavelet domain perceptual model. Recently Kutter et al. [30] have presented a perceptual model that takes into account the sensitivity and masking behavior of HVS, by means of a local isotropic contrast measure and a masking model. In our present investigations, we are comparing the developed GWSF with that of WPM.

2.1 Watson's Perceptual Model (WPM)

Consider an image matrix $\mathbf{x}$ in spatial domain. The image is transformed to matrix $\mathbf{X}$ by applying 8x8 block DCT. According to the WPM, we define the visibility threshold $T(i,j)$ for every (i,j) DCT coefficient of 8x8 block as follows:

$$\log T(i,j) = \log\left(\frac{T_{\min}\,(f_{i,o}^2 + f_{o,j}^2)^2}{(f_{i,o}^2 + f_{o,j}^2)^2 - 4\,(1-r)\,f_{i,o}^2 f_{o,j}^2} \right) + u\left(\log\sqrt{(f_{i,o}^2 + f_{o,j}^2)} - \log f_{\min} \right)^2 \tag{1}$$

where $f_{i,o}$ and $f_{o,j}$ denotes the vertical and horizontal frequencies (cycles/degree) of the DCT basis functions respectively. $T_{\min}$ is the minimum value of $T(i,j)$ corresponding to $f_{\min}$. The rest of the parameters are also set empirically [16-17]. The effect of luminance sensitivity is considered by correcting this threshold corresponding to average luminance of each block:

$$T'(i,j) = T(i,j)\left(\frac{X_{o,o}}{\overline{X}_{o,o}} \right)^{a_T} \tag{2}$$

where $X_{o,o}$ is the DC coefficient of each block and $\overline{X}_{o,o}$ represents the average screen luminance =1024 (for an 8-bit image). The following relation incorporates the effect of contrast masking:

$$T^*(i,j) = \max\left[\, T'(i,j)\ ,\ \left|T'(i,j)\right|^{1-\omega} X(i,j)^{\omega} \right] \tag{3}$$

where $X(i,j)$ is AC DCT coefficient of each block and ω has been empirically set to a value of 0.7.

The corrected threshold $T^*(i,j)$ is finally used to compute the allowed alteration of the DCT coefficient instead of luminance. These allowed alterations represent the perceptual mask denoted by $\boldsymbol{\alpha}$:

$$\alpha(k_1,k_2) = 4\cdot(1+(\sqrt{2}-1)\,\delta(l_1))\cdot(1+(\sqrt{2}-1)\,\delta(l_2))\cdot\gamma\cdot T^*(l_1,l_2) \tag{4}$$

where $\gamma < 1$ is a scaling factor used to incorporate certain degree of conservativeness in the watermark corresponding to the effects like spatial masking in the frequency domain that have been overlooked. While $l_1 = k_1 \bmod 8,\ l_2 = k_2 \bmod 8$ and $\delta(.)$ is the Kronecker delta function.

3. Attacks and their Countermeasures

Digital watermarks can be attacked in a variety of different ways and each application requires its own type of robustness. Cox et al. [1] have discussed in detail the types and levels of robustness that might be required for a particular watermarking application. They have discussed some of the attacks as well as their countermeasures. Voloshynovsky et al. [13] have classified attacks into four basic categories: removal and interference attacks, geometrical attacks, cryptographic attacks and protocol attacks. Intentional tempering, as opposed to the common signal processing attacks are difficult to survive. However, watermark attacks as well as their countermeasures are complex and still a topic of research. Therefore in evaluating the potential of a watermarking technique to meet the robustness requirements, many assumptions are made especially about the attacker. For example, does the attacker know the watermarking algorithm, has he got a detector that he can modify, what tools are available to him etc. Once the watermarking system is specified publicly, an attacker usually has more freedom as compared to a watermarker because the attacker is free to develop extra and more intricate attacks, while the watermarker can no longer amend it [1].

4. Watermark Robustness and Imperceptibility Measures

The imperceptibility of a watermark is generally measured in terms of weighted Peak Signal to Noise Ratio (wPSNR) [15], Watermark to Document Ratio (WDR) [31] and Structural Similarity Index Measure (SSIM) [32]. SSIM measure uses the hypotheses that HVS is highly adopted for extracting structural information. It is argued that natural image signals are highly structured, as the nearby pixel exhibit strong dependencies [32]. These dependencies provide information about the structure of the object in an image, which are overlooked by the error-based measures. To estimate robustness during GP simulation, we use watermark power. We represent watermark power by mean squared strength (MSS) given as:

$$MSS = \frac{1}{N_b N_d} \sum_{u_1=1}^{N_b} \sum_{u_2=1}^{N_d} \alpha(u_1, u_2)^2 \tag{5}$$

where, N_b is the total number of 8×8 blocks in the cover image and N_d is the number of bandpass (low and mid frequency) DCT coefficients.

We have used watermark power as an estimate of robustness, because in the testing and comparison phase of best evolved GWSF (section 5.3), the underlying watermarking technique is the same for both WPM and GWSF based perceptual shaping schemes. Hence, as in our previous work [22], we assume that the MSS will provide a suitable measure of the estimated robustness at the embedding stage of the GP simulation.

5. Proposed Technique for Developing a GWSF

The basic architecture of our proposed scheme for developing GWSF is shown in figure 1. Five modules work in a cyclic fashion. We first explain the overall working of the basic architecture. Details of the individual modules are given in section 5.1.

The GP module produces a population of GWSF. Each GWSF is presented to the perceptual shaping module, where it is applied to the cover image in DCT-domain, generating a perceptual mask. In the watermarking stage, the watermark is shaped using the perceptual mask. The conceivable attack is performed on the watermarked image in the attack module. In the decoding module, the embedded message is retrieved from the corrupted image. The watermark imperceptibility at the embedding stage and BCR at the decoding stage, are then used in the scoring criterion of the GP module. In this way, the GP module evaluates the performance of its several generated GWSFs.

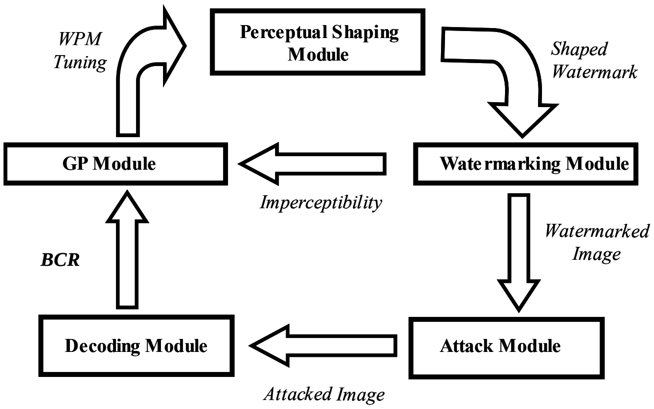

Figure 1 Basic architecture of GWSS

5.1 Evolution of watermark shaping functions

5.1.1 The GP Module

Genetic programming is a type of Evolutionary Algorithms that are based on the mechanism of natural selection and natural genetics. In GP, a candidate solution is represented using a data structure such as a tree. Initially, a random population of such candidate solutions is created. Every candidate solution is evaluated and scored using application dependent fitness function. The survival of fittest is implemented by retaining the best individuals. The rest are deleted and replaced by the offspring of the best individuals. The retained ones and the offspring make a new generation. Some offspring may have high score than their parents in the previous generation.

The whole process is repeated for the subsequent generations. With the scoring and selection procedure in place, each new generation has, on average, a slightly higher score than the previous one. The process is stopped when a single individual in a generation gets a score that exceeds a desired value. In this way the solution space is refined generation by generation and thus converges to the optimal/near optimal solution. For a detailed study one may refer to [33]. In this present work, we search for superior watermark shaping functions— watermark shaping functions that are able to make superior tradeoff between robustness and imperceptibility with respect to existing tradeoff techniques.

To represent a possible solution with a GP tree, one needs to define suitable functions, terminals, and fitness criteria according to the optimization problem. These settings for evolving GWSF are as under:

GP Function Set: Function set in GP is a collection of functions available to the GP system. In our GP simulations, we have used simple functions, including four binary floating arithmetic operators (+, -, *, and protected division), *LOG, EXP, SIN* and *COS*.

GP Terminals: To develop initial population of GWSF, we consider GWSF as watermark shaping function and the characteristics of HVS as independent variables. By doing this, in essence, we are letting GP exploit the search space representing different possible forms of dependencies of the watermark shaping function on the characteristics of HVS. Therefore, the current value of WPM-based perceptual mask, DC and AC DCT coefficients of 8x8 block are provided as variable terminals (equation 12 and figure 2). Random constants in the range [-1,1] are used as constant terminals.

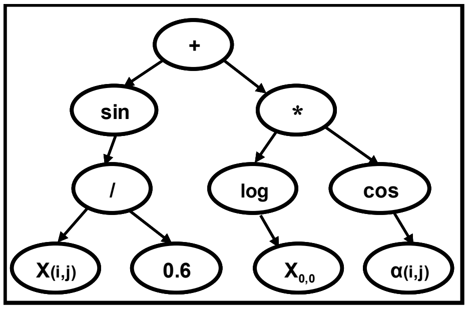

Figure 2 An example GP tree representing a GWSF

Fitness Function: A fitness function in GP is supposed to grade each individual of the population. It is designed to provide feedback about how well an individual of the GP population is performing at the given task. More details are given in section 5.1.3. Every watermark shaping function of a GP population is evaluated in terms of structuring the watermark. The evaluation is based on how well is the SSIM measure at a certain level of watermark power as well as how high the BCR value is:

$$Fitness = W_1 * SSIM_{E.S} + W_2 * BCR_{attack} \tag{6}$$

where $SSIM_{E.S}$ denotes the structure similarity index measure of the marked image at a certain level of estimated robustness. W_1 and W_2 represent the corresponding weightage of the two terms in the fitness. If W_1 and W_2 are set to 1.0, the fitness attains a maximum value of 2.0. BCR is given as:

$$BCR(\mathbf{M}, \mathbf{M}') = \sum_{i=1}^{L_m} \overline{(m_i \oplus m_i')} \Big/ L_m \tag{7}$$

where $\mathbf{M}$ represents the original, while $\mathbf{M}'$ represents the decoded message, L_m is the length of the message and $\oplus$ represents exclusive-OR operation. It should be noted that $(1 - BCR)$ represents bit incorrect ratio.

Thus, each individual GWSF of a GP population is scored using equation 6 as a fitness function. The greater the fitness is, the better the individual has performed.

Termination Criterion: The GP simulation is ceased when one of the following conditions is encountered:

1. The fitness score exceeds 1.99 with $\mathrm{MSS} \geq 20.0$.
2. The number of generations reaches the predefined maximum number of generations.

5.1.2 Perceptual Shaping Module

A watermark shaping function tailors a watermark according to the cover image by exploiting the characteristics of HVS. This enables us to embed a large energy watermark at low cost of resultant distortion to the cover image. The perceptual shaping module receives the individual GWSF provided by the GP module as an input. Each GWSF is operated on the cover image in DCT-domain. Corresponding to the selected DCT coefficient of a block, the GWSF returns a value. The magnitude of this value represents the perceptual strength of the alteration made to that coefficient. The functional dependency of the perceptual model on the characteristics of HVS can be represented as follows:

$$\alpha(k_1, k_2) = f\left(T(i,j),\ X_{0,0},\ X(i,j) \right) \tag{8}$$

where the first variable, T is the visibility threshold representing frequency sensitivity of HVS. $X_{0,0}$ is the DC DCT coefficient, while $X(i, j)$ is the AC DCT coefficient of the current block. They represent the luminance sensitivity and contrast masking characteristics of HVS respectively.

Operating the GWSF on all of the DCT coefficients, we obtain the perceptual mask for the current cover image. The product of the spread-spectrum sequence and expanded message bits is multiplied with this perceptual mask to obtain the watermark. The 2-D watermark signal $\mathbf{W}$ (see figure 3) is given as:

$$\mathbf{W} = \boldsymbol{\alpha} \cdot \mathbf{S} \cdot \mathbf{b} \tag{9}$$

where $\mathbf{S}$ is a pseudo random sequence and $\mathbf{b}$ is the repetition-based expanded code vector, corresponding to the message to be embedded. Adding this watermark to the original image in transformed domain performs the embedding:

$$\mathbf{Y} = \mathbf{X} + \mathbf{W} \tag{10}$$

Here the watermark $\mathbf{W}$ is our desired signal, while the cover image $\mathbf{X}$ acts as an additive noise. As we are genetically tuning WPM whose corresponding perceptual mask is represented by $\boldsymbol{\alpha}$, therefore, equation 9 will be modified as follows:

$$\mathbf{W} = \boldsymbol{\alpha}_G \cdot \mathbf{S} \cdot \mathbf{b} \tag{11}$$

where $\boldsymbol{\alpha}_G$, representing perceptual mask corresponding to GWSF, incorporates the dependencies from WPM, AC and DC coefficients and the intended attack.

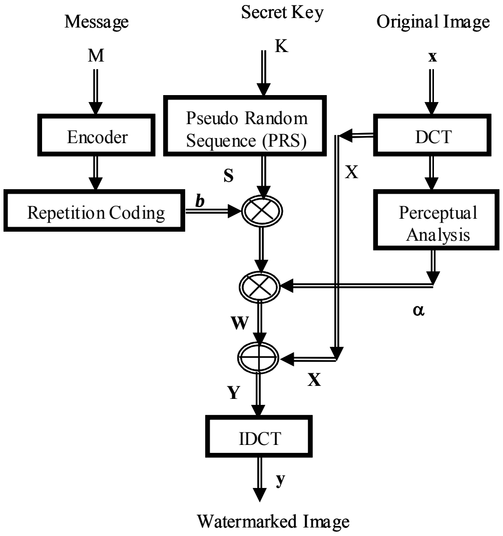

Figure 3 Hernandez's [5] watermark embedding scheme

If A denotes the information about the intended attack, then equation 8 is modified to include the resultant changes in the distribution of the DCT coefficients caused by the attack as follows:

$$\alpha_G(k_1,k_2) = f\left(\alpha(k_1,k_2),\ X_{0,0},\ X(i,\ j),\ A\right) \tag{12}$$

5.1.3 Watermarking Module

In order to evaluate the performance of each individual GWSF of the GP population, the watermarking module implements the spread spectrum based watermarking technique proposed by Hernandez et al. [5]. This watermarking technique is oblivious and embeds message into the low and mid frequency coefficients of 8×8 DCT blocks of a cover image. The employed watermarking scheme performs the statistical modeling of DCT coefficients using generalized Gaussian distribution. This fact helps in constructing better detector/decoder structures than the simple Gaussian correlation receiver that is mostly used. One of the reasons for using this watermarking scheme is that the DCT is applied in blocks of 8x8 pixels, in a manner similar to that used in JPEG algorithm. Hence, it is easy to use and compare WPM with that of the GWSF. Secondly, this watermarking scheme has strong theoretical foundations [5]. The embedding in DCT-domain is performed using equation 10.

The watermarking module of our proposed technique provides the imperceptibility of the resultant watermark as a feedback to the GP module. The structure of how different sub-modules work within the GWSS is shown in figure 4.

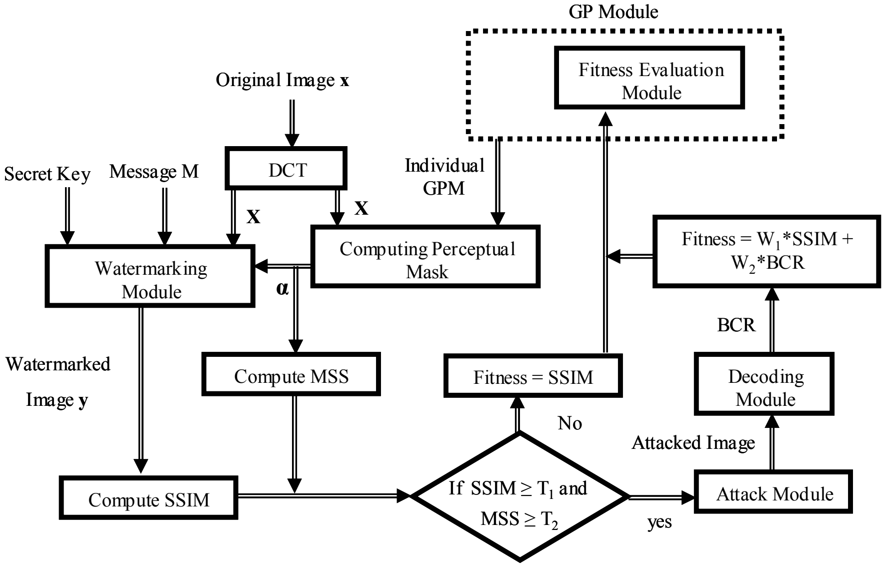

Figure 4 Block diagram of the proposed Genetic Watermark Shaping Scheme (GWSS).

5.1.4 Attack Module

In this module, the anticipated attack is performed on the watermarked image. We assume that the decoding module is fixed and does not modify in accordance to the attack. Specifically, to develop Wiener attack-resistant GWSF, before decoding the embedded message, we perform Wiener attack. Similarly to develop JPEG, Median filtering, and Gaussian attack-resistant GWSFs, GP simulations are carried out separately with each attack being performed before decoding the message.

5.1.5 Decoding Module

The decoding module receives the corrupted image after an attack as an input. It performs decoding of the embedded message as discussed in [5]. The same GWSF, as used in the embedding stage, is used to obtain the perceptual mask for the received image. The perceptual mask is then used to obtain sufficient statistics for the Maximum Likelihood based decoder.

5.2 Bonus Fitness-based Evolution

In the decoding stage, both imperceptibility and robustness requirements of a watermark are implemented through the use of multi-objective fitness function [33-34]. One way to perform this is to use equation 6. However, the drawback of this type of fitness function is that due weightage for learning the distribution of the DCT coefficients of each block of a cover image is not incorporated. In other words, instead of searching for a superior and image independent GWSF, main effort of the GP search is spent on searching a GWSF that results in high BCR value. Consequently, optimization of robustness versus imperceptibility tradeoff is belittled. This type of GWSF is not image adaptive and might have very poor performance for attacks other than the intended attacks. This problem is solved by using the idea of bonus fitness that we have used in our earlier work [35]. As can be examined from figure 4, those GWSF that make a better tradeoff between robustness and imperceptibility, are given bonus fitness. The bonus fitness is the amount of resistance against the intended attack in terms of BCR_{attack}. Thus equation 6 is modified as follows:

$$Fitness = \begin{cases} W_1 * Fitness_1 + W_2 * Fitness_2 & if\ SSIM \geq T_1\ and\ MSS \geq T_2 \\ W_1 * Fitness_1 & otherwise \end{cases} \tag{13}$$

where $Fitness_1 = SSIM_{E.S}$ while $Fitness_2 = BCR_{attack}$ and T_1 , T_2 are lower bounds of $SSIM_{E.S}$ and MSS respectively.

In this way, the second driving force is separated from the first and basic driving force through the concept of bonus fitness. Otherwise, the GP simulation will usually tend to focus on the second requirement and will altogether neglect the basic requirement. Figure 5 elaborates this idea of bonus fitness incorporated in the GP search. We can observe that in each generation, those GWSF that make a good tradeoff are tagged (they are represented with star symbol and thus conceptually separated from the main GP search beam). A competition in terms of the 2nd fitness among these tagged GWSF then starts immediately. The overall fitness is improved with improvement in both types of fitness. The selection of when to tag an individual GWSF, by judging the tradeoff, is of crucial importance. It is implemented by requiring the MSS and SSIM values to lie above certain lower bounds. The smaller these lower bounds for fulfilling the first fitness criteria are, the larger is the diversity among the tagged GWSFs.

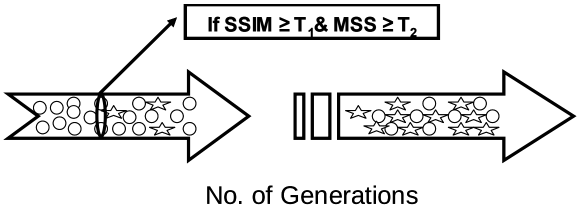

Figure 5 A representation of GP search beam illustrating the bonus fitness concept

5.3 Testing Performance of the Best-evolved GWSF

In order to assess the performance of the best-evolved GWSF, its expression is saved at the end of the GP simulation. The best-evolved GWSF is then compared with that of WPM in terms of watermark shaping. Where by, the watermark shaping ability is assessed by computing watermark imperceptibility as well as robustness measures. Figure 6 shows the details of the testing phase for the evolved GWSF.

6. Implementation Details

We have used MATLAB environment for our experimental studies. To employ GP, we use GPLAB toolbox [36-37]. The GP parameter settings are shown in table 1, while the remaining parameters are used as default in the software.

Lena image of size 256x256 is used as a cover image with $N_d = 22$ (7 to 29 in zigzag order) during the GP simulation. Message size is kept equal to 64 bits. Following [23-24], the parameters of WPM are set as $r = 0.7$, $T_{\min} = 1.1548$, $u = 1.728$, $f_{\min} = 3.68$ cycles/degree and $a_T = 0.649$. To estimate the value of parameter c for generalized Gaussian Distribution-based modeling of each (i,j) DCT sequence [5], we have considered its range [0.02, 2.0] with grid step of 0.02. The watermark power, represented by MSS, is constrained to lie above a certain lower bound for all the individuals. To assign bonus fitness, we have taken T_1, T_2, W_1 and W_2 as 0.96 , 20.0, 1.0 and 1.0 respectively. The values of T_1, T_2 are set empirically.

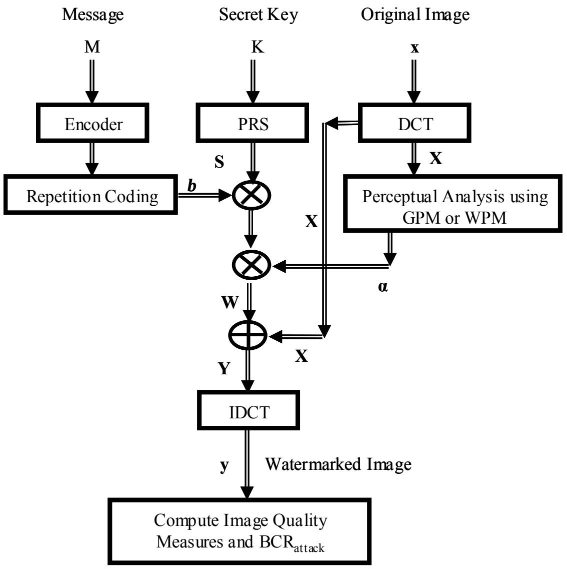

Figure 6 Block diagram of the testing and comparison phase

In the testing phase, all images except Baboon and Boat are of size 256x256. The attacks for which specific GWSF are developed, include adaptive Wiener filtering of window size 3x3, JPEG compression (QF = 80), Median filtering of window size 3x3 and Gaussian noise of σ = 50. In order to develop GWSF, keeping population size equal to 300 and no. of generations 30, the GP simulation consumes about one hour on a Pentium IV machine (2.0 GHz speed and 256 Mb RAM). In the testing phase, the watermarking scheme using the best-evolved GWSF spends about 30 sec to watermark Lena image.

TABLE 1 GP Parameters Setting

Objective:	**To evolve conceivable attack-resistant GWSF**
Function Set:	+, -, *, protected division, *SIN, COS,* and *LOG*
Terminal Set:	Constants: *random constants in range* of [-1, 1] Variables : $X_{0,0}/1024, \lvert X(i,j) \rvert$ *and* $\alpha(i,j)$
Fitness :	*SSIM*
Selection:	Generational
Population Size:	260
Initial max.Tree Depth	6
Initial population:	Ramped half and half
Operator prob. type	Variable
Sampling	Tournament
Expected no. of offspring	rank89
Survival mechanism	Keep best
Real max level	31
Termination:	Generation 32

7. Results and Discussion

7.1 Perceptual Shaping Using GWSF

In figure 7, watermarking strength corresponding to each bandpass DCT coefficient of block-based DCT is shown. These strengths are produced by the Wiener attack-resistant GWSF for Lena image. It is observed that depending upon the current AC and DC coefficient, it provides suitable imperceptible alterations according to the spatial content of that block. This fact indicates that GWSF is able to exploit HVS for shaping the watermark according to any cover image. In other words, GWSF makes the watermarking technique adaptive with respect to the cover image. The resultant watermark is shown in figure 8.

7.2 Imperceptibility of the resultant watermark

In figure 11, we have shown the difference image, obtained by subtracting the original image (figure 9) from the watermarked image (figure 10) in spatial domain. The pixel intensity of the difference image is amplified ten times for illustration purpose. Although, DCT domain is used for embedding, still GWSF is able to learn the spatial distribution of the Lena image, as most of the strong embedding is performed in highly textured areas.

7.3 GWSF developed for Wiener Attack

In table 2, both WPM and Wiener attack-resistant GWSF are compared in terms of the marked image quality and $(1-BCR_{attack})$ for 8 different standard images. The perceptual masks corresponding to both WPM and GWSF are multiplied with some scaling factor to achieve equal distortion of the resultant watermarked image (in terms of approximately equal SSIM value for each image). It is observed that although evolved using Lena image, Wiener attack-resistant GWSF is image independent. This is because its imperceptibility measures are comparable to that of WPM for the entire test images. However, in terms of $(1-BCR_{attack})$ performance, the Wiener attack-resistant GWSF has superior performance as compared to that of WPM, for almost all of the test images. The Wiener attack-resistant GWSF is given below:

$$\boldsymbol{\alpha}_{\mathbf{G}}(k_1,k_2) = \cos(\sin(\boldsymbol{\alpha}(k_1,k_2)) + \boldsymbol{\alpha}(k_1,k_2)) + (\log(\cos(0.22897)) + 0.22897) * X(i,j) \qquad (14)$$

7.4 GWSF developed for Gaussian Noise Attack

Table 3, shows the same comparison in case of Gaussian noise attack (σ = 50). Again, Gaussian attack-resistant GWSF has comparable performance to that of WPM in terms of imperceptibility, while superior performance in case of robustness $(1-BCR_{attack})$. Figure 12, shows the watermarked image after being attacked by the Gaussian noise. Whereas, figure 13 demonstrates the $(1-BCR_{attack})$ versus standard deviation performance of both WPM and GWSF. It can be observed that Gaussian noise attack-resistant GWSF has low $(1-BCR_{attack})$ values corresponding to different standard deviations.

7.5 GWSF developed for JPEG Compression Attack

Figure 14, 15 and table 4 show the same comparison in case of JPEG attack. It is observed that imperceptibility performance of the JPEG attack-resistant GWSF is low as compared to that of WPM (low SSIM values corresponding to less energy watermark embedding). But on the other hand, the improvement in $(1-BCR_{attack})$ performance in this case is far better from the previous two cases.

7.6 GWSF developed for Median Filtering Attack

Table 5 compares the evolved Median attack-resistant GWSF to that of WPM. In this case the imperceptibility performance at a certain level of watermark power is comparable, but $(1-BCR_{attack})$ performance is again superior. The reason behind this is that the attack-resistant GWSF spreads the watermark energy in such areas, where the attack as well as the distortion affect is less.

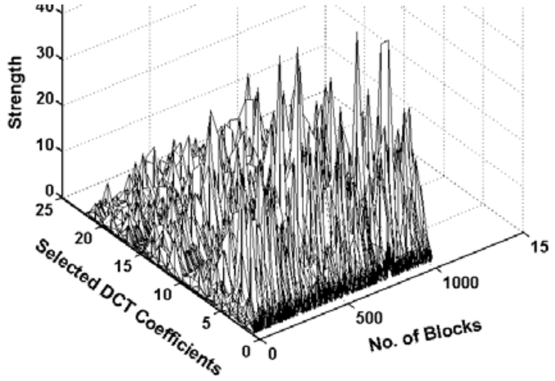

Figure 7 Watermarking strength for Lena image using evolved GWSF

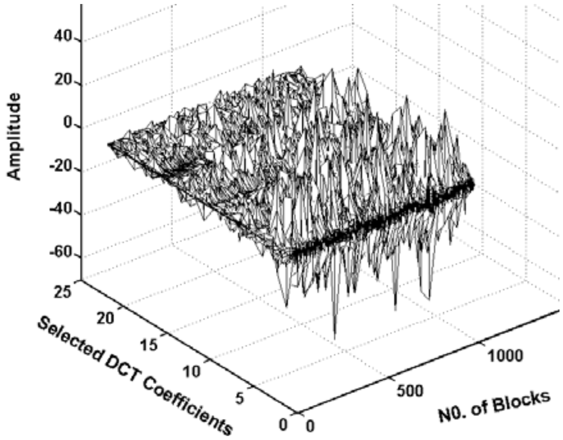

Figure 8 Watermark generated for Lena image using evolved GWSF

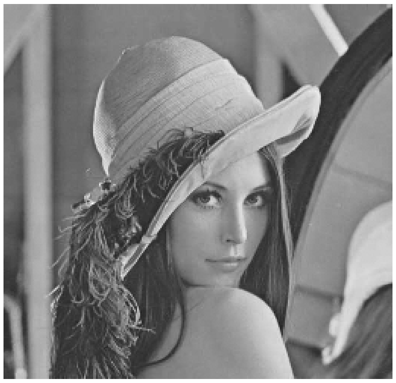

Figure 9 Original Lena image

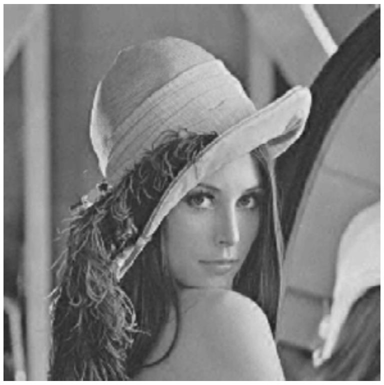

Figure 10 Watermarked Lena image with Wiener attack-resistant GWSF

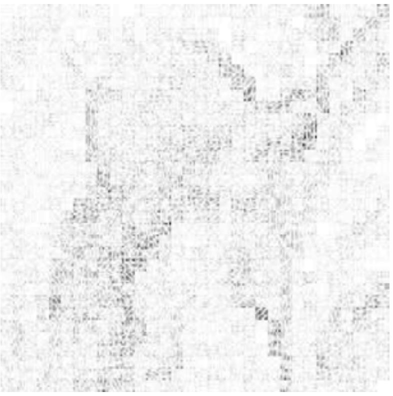

Figure 11 Difference image of Lena

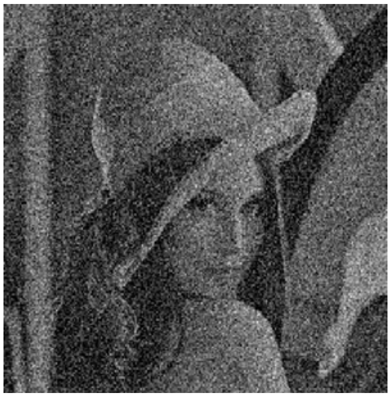

Figure 12 Gaussian noise attacked image (σ = 50)

TABLE 2 Wiener Attack-Resistance Performance Comparisons. Note that the attack-resistance performance is compared by keeping the imperceptibility of the watermark almost the same in both cases.

Images	Perceptual Model	Scaling factor	MSS	SSIM	WDR	*w*PSNR	(1-BCR)
Trees	WPM	0.30123	63.4553	0.9737	-29.604	40.617	0.125
	GWSF	0.48	83.991	0.9738	-28.386	39.951	0.0469
Lena	WPM	0.30123	23.696	0.9816	-33.3294	43.245	0.2656
	GWSF	1.282	17.1462	0.9816	-34.7357	44.0532	0.0
Baboon (232x248)	WPM	0.30123	72.2899	0.9779	-27.3715	43.477	0.0469
	GWSF	0.519	89.7937	0.9779	-26.4274	43.277	0.0313
Couple	WPM	0.30123	62.3051	0.9711	-28.913	40.2752	0.0938
	GWSF	0.552	80.3543	0.9711	-27.8227	39.68	0.0781
Boat (232x248)	WPM	0.30123	55.9301	0.9730	-29.908	40.6763	0.125
	GWSF	0.52	67.9235	0.9731	-29.065	40.144	0.0313
Fruits	WPM	0.30123	36.278	0.9737	-33.6307	41.3071	0.2344
	GWSF	0.513	71.0392	0.9771	-30.818	39.7361	0.1719
House	WPM	0.30123	26.643	0.9745	-33.6307	41.3071	0.0781
	GWSF	0.5524	39.1316	0.9745	-31.9648	39.7503	0.0781
Chemical Plant	WPM	0.30123	42.3627	0.9778	-29.4101	41.556	0.0781
	GWSF	0.494	56.589	0.9778	-28.1518	40.8834	0.0561

TABLE 3 JPEG Attack-Resistance Performance Comparisons. Note that the attack-resistance performance is compared by keeping the imperceptibility of the watermark almost the same in both cases.

Images	Perceptual Model	Scaling factor	MSS	SSIM	WDR	*w*PSNR	(1-BCR)
Trees	WPM	0.30123	63.4553	0.9737	-29.604	40.617	0.0625
	GWSF	1.505	67.8033	0.9737	-29.3161	40.4706	0.0469
Lena	WPM	0.30123	23.696	0.9816	-33.3294	43.245	0.2031
	GWSF	1.554	25.126	0.9816	-33.0744	43.1152	0.1875
Baboon (232x248)	WPM	0.30123	72.2899	0.9779	-27.3715	43.477	0.0781
	GWSF	1.59	74.627	0.9778	-27.232	43.45	0.0313
Couple	WPM	0.30123	62.3051	0.9711	-28.913	40.2752	0.0625
	GWSF	1.694	67.173	0.9712	-28.588	40.1253	0.0469
Boat (232x248)	WPM	0.30123	55.9301	0.9730	-29.908	40.6763	0.0938
	GWSF	1.594	57.696	0.973	-29.772	40.606	0.0938
Fruits	WPM	0.30123	36.278	0.9737	-33.6307	41.3071	0.1563
	GWSF	1.53	42.734	0.9771	-33.025	41.0104	0.1406
House	WPM	0.30123	26.643	0.9745	-33.6307	41.3071	0.1719
	GWSF	1.609	27.1527	0.9745	-33.547	41.288	0.1563
Chemical Plant	WPM	0.30123	42.3627	0.9778	-29.4101	41.556	0.0781
	GWSF	1.51	46.917	0.9778	-28.966	41.289	0.0781

TABLE 4 Gaussian Attack-Resistance Performance Comparisons. Note that the attack-resistance performance is compared by keeping the imperceptibility of the watermark almost the same in both cases.

Images	Perceptual Model	Scaling factor	MSS	SSIM	WDR	*w*PSNR	(1-BCR)
Trees	WPM	0.30123	63.4553	0.9737	-29.604	40.617	0.2344
	GWSF	0.787	40.533	0.9737	-31.55	40.7437	0.0469
Lena	WPM	0.30123	23.696	0.9816	-33.3294	43.245	0.1719
	GWSF	1.252	18.83	0.9816	-34.329	43.313	0.0
Baboon (232x248)	WPM	0.30123	72.2899	0.9779	-27.3715	43.477	0.3438
	GWSF	1.163	78.475	0.9778	-27.014	41.488	0.2188
Couple	WPM	0.30123	62.3051	0.9711	-28.913	40.2752	0.2188
	GWSF	1.072	104.39	0.971	-26.659	36.9406	0.0781
Boat (232x248)	WPM	0.30123	55.9301	0.9730	-29.908	40.6763	0.1563
	GWSF	1.178	63.283	0.973	-29.362	38.104	0.0781
Fruits	WPM	0.30123	36.278	0.9737	-33.6307	41.3071	0.25
	GWSF	1.563	49.919	0.9771	-32.353	40.3825	0.0625
House	WPM	0.30123	26.643	0.9745	-33.6307	41.3071	0.1719
	GWSF	1.818	24.124	0.9745	-34.055	41.562	0.0
Chemical Plant	WPM	0.30123	42.3627	0.9778	-29.4101	41.556	0.1875
	GWSF	0.978	20.622	0.9778	-32.535	42.4616	0.0781

TABLE 5 Median Filtering Attack-Resistance Performance Comparisons. Note that the attack-resistance performance is compared by keeping the imperceptibility of the watermark almost the same in both cases.

mages	Perceptual Model	Scaling factor	MSS	SSIM	WDR	*w*PSNR	(1-BCR)
Trees	WPM	0.30123	63.4553	0.9737	-29.604	40.617	0.1563
	GWSF	1.413	26.869	0.9737	-33.34	42.5537	0.0156
Lena	WPM	0.30123	23.696	0.9816	-33.3294	43.245	0.1563
	GWSF	1.128	10.2681	0.9816	-36.9644	45.06	0.0
Baboon (232x248)	WPM	0.30123	72.2899	0.9779	-27.3715	43.477	0.3438
	GWSF	1.544	35.026	0.9779	-30.527	44.149	0.25
Couple	WPM	0.30123	62.3051	0.9711	-28.913	40.2752	0.2188
	GWSF	1.65	34.089	0.971	-31.53	41.4407	0.0
Boat (232x248)	WPM	0.30123	55.9301	0.9730	-29.908	40.6763	0.1563
	GWSF	1.382	23.9525	0.973	-33.6025	42.4273	0.0
Fruits	WPM	0.30123	36.278	0.9737	-33.6307	41.3071	0.125
	GWSF	1.066	12.3244	0.9771	-38.4293	41.865	0.0313
House	WPM	0.30123	26.643	0.9745	-33.6307	41.3071	0.2031
	GWSF	1.003	10.0407	0.9745	-37.8633	44.3606	0.0469
Chemical Plant	WPM	0.30123	42.3627	0.9778	-29.4101	41.556	0.0156
	GWSF	1.405	24.8827	0.9778	-31.7236	42.4763	0.0781

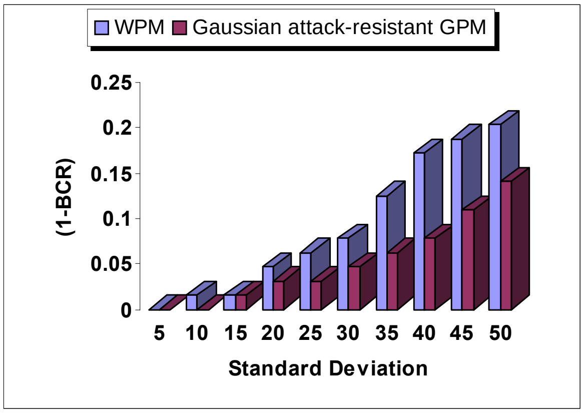

Figure 13 (1-BCR) vs. std. deviation of Gaussian noise attack.

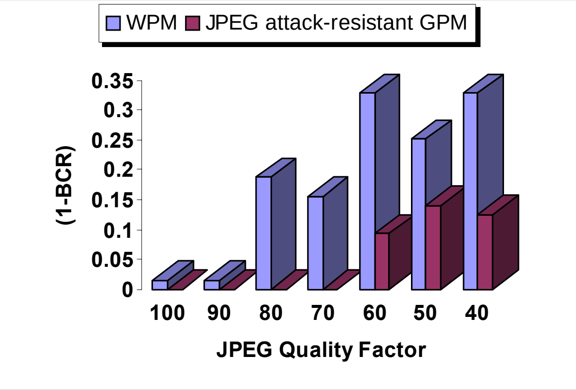

Figure 14 (1-BCR) vs. quality factor of JPEG compression attack.

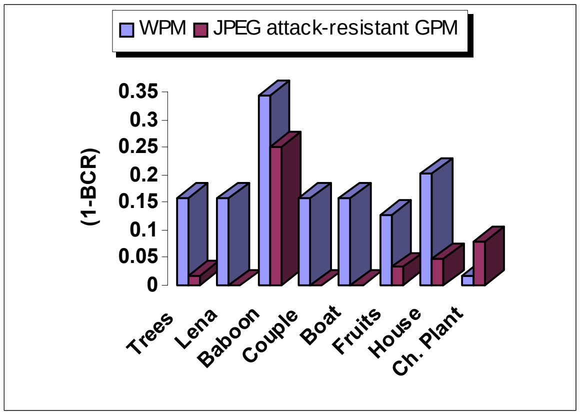

Figure 15 Bar chart of (1-BCR) vs. JPEG attack (QF=70) for different images.

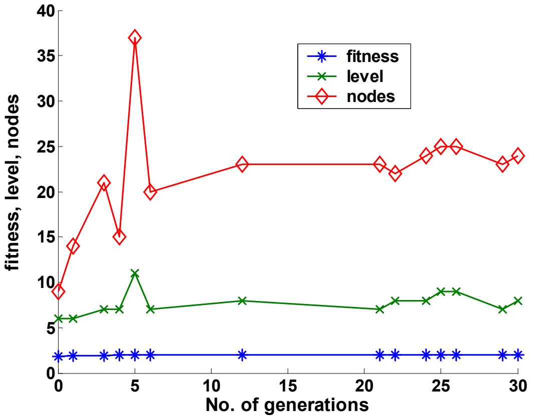

Figure 16 Accuracy vs. complexity plot of GP simulation for evolving median filtering attack-resistant GWSF

Figure 16 shows the accuracy versus complexity plot of GP simulation. It is observed that as generations pass by, improvement in fitness of the best Median attack-resistant GWSF is achieved at the cost of its complexity. That is, with increase in fitness of the best GWSF of a generation, its genome's total number of nodes as well as its average tree depth increases. The above analysis of the various evolved GWSFs indicate that GWSS develops GWSF that results in cover image as well as attack dependent restructuring of the watermark.

Conclusions

In this paper we have considered the GP-based perceptual shaping of a digital watermark in accordance to the cover image and anticipated attack. The GP tuned GWSFs are image adaptive and the GWSS as a whole is attack adaptive. A significant improvement in resistance against the intended attack is achieved by letting the GP search exploit the attack information. This is in essence, like attack-informed embedding. Both these attributes of a GWSF; superior tradeoff and high resistance against an anticipated attack, are obtained by incorporating the concept of bonus fitness in multi-objective fitness function. Developing GWSF needs considerable execution time (about one hour). However, once the best GWSF is developed, then employing GWSF for watermark shaping is quite straight forward and easy to implement. Even in the development phase, with the use of fast and parallel processing based implementations of GP [38-39], it is possible to use GP-based watermarking to real business applications. The proposed GWSS is applicable for tuning other perceptual models as well. In addition to the selection of suitable strength, the selection of DCT coefficients for embedding as proposed in [8] may also be performed. This will require the whole 63 AC coefficients of a DCT block to be considered for embedding, instead of the middle frequency coefficients. This may further improve the resistance against the intended attack, as different attacks usually affect different frequency bands in DCT block. Work is in progress to develop GWSF for restructuring of a watermark against a battery of attacks.

Acknowledgements: The authors greatly acknowledge the financial support provided by the Higher Education Commission, Government of Pakistan, under the indigenous PhD scholarship program (No. 17-6 (183)/Sch / 2001).

References

[1] I. J. Cox, M. L. Miller, and J. A. Bloom, Digital Watermarking and fundamentals, Morgan Kaufmann, San Francisco, 2002.

[2] I. J. Cox, M. L. Miller and J. A. Bloom, Watermarking Applications and Their Properties, Proceedings of the International Conference on Information Technology: Coding and Computing - ITCC2000, 2000, pp. 6-10.

[3] K. Su. Jonathan and B. Girod, Power-spectrum conditions for energy–efficient watermarking, IEEE Trans. on Multimedia, 4 (4), Dec. 2002.

[4] C. Podilchuk and W. Zeng, Image-adaptive watermarking using visual models, IEEE Journal on Selected Areas in Communications, 10(4), (1998), 525-540.

[5] J. R. Hernandez, M. Amado, and F. Perez-Gonzalez, DCT-Domain watermarking techniques for still images: Detector performance analysis and a new structure, IEEE Trans. on Image Processing, 9(1), (2000), 55-68.

[6] A. Briassouli , P. Tsakalides and A. Stouraitis Alpha stable DCT hidden messages in heavy-tails: DCT domain watermark detection using alpha-stable models, IEEE Trans. on Multimedia, 2005 (to appear).

[7] S. Voloshynovskiy, A. Herrigel, N. Baumgaetner, and T. Pun, A stochastic approach to content adaptive digital image watermarking, In third international workshop on Information Hiding, (Dresden, Germany), Sep. 29, 1999.

[8] Shieh, Huang, Wang and Pan, Genetic watermarking based on transform domain technique, Pattern Recognition, vol. 37, 2004, pp. 555-565.

[9] Ki-Ryong Kwon, Ji-Hwan Park, Eung-Joo Lee, Ahmed H. Tewfik, Highly Reliable Stochastic Perceptual Watermarking Model Based on Multiwavelet Transform, Digital Watermarking, Second International Workshop, IWDW 2003, Seoul, Korea, October 20-22, 2003, pp. 423-434.

[10] P. Meerwald, and A. Uhl, A survey of wavelet-domain watermarking algorithms, Proceedings of SPIE, Security and Watermarking of Multimedia Contents III, v. 4314, San Jose, CA, USA, January 22 - 25,200.

[11] D. Kundur, D. Hatzinakos, Towards Robust Logo Watermarkin Using Multiresolution Image Fusion Principles, IEEE Trans. on Multimedia, 6(1), 2004, pp. 185-198.

[12]] A. Sequeira, D. Kundur, Communication and Information Theory in Watermarking: A Survey, Multimedia Systems and Applications IV, A. G. Tescher, B. Vasudev, and V. M. Bove, eds., Proc. SPIE (vol. 4518), pp. 216-227, Denver, Colorado, August 2001.

[13] S. Voloshynovskiy, S. Pereira, V. Iquise, and T. Pun, Attack modeling: towards a second generation watermarking benchmark, Signal Processing, 81, 6, pp. 1177-1214, June 2001. Special Issue: Information Theoretic Issues in Digital Watermarking, 2001. V. Cappellini, M. Barni, F. Bartolini, Eds.

[14] M. Kutter and F. A. P. Petitcolas, A fair benchmark for image watermarking systems, Electronic Imaging '99. Security and Watermarking of Multimedia Contents, vol. 3657, Sans Jose, CA, USA, the International Society for Optical Engineering, Jan. 1999, pp. 25-27.

[15] S. Voloshynovskiy, S. Pereira, T. Pun, J.J. Eggers and J. K. Su, Attacks on Digital Watermarks: Classification, Estimation Based Attacks, and Benchmarks, IEEE Commun. Mag. 39 (8) (2001) 118-126.

[16] F. Hartung, J. K. Su and B. Girod. Spread Spectrum Watermarking: Malicious Attacks and Counterattacks, in Proc. SPIE, Security & Watermarking Multimedia Contents, vol. 3657, San Jose, CA, Jan. 1999, pp. 147-158.

[17] Bum-Soo Kim, Jae-Gark Choi, Chul-Hyun Park et al., Robust digital image watermarking method against geometrical attacks, Real-Time Imaging, 9 (2003), 139-149.

[18] E. Praun, H. Hoppe and A. Finkelstein, Robust Mesh Watermarking, Proceedings of SIGGRAPH 1999, Computer Graphics Proceedings, Annual Conference Series, ACM, pp.49-56.

[19] J.J.K. O'Ruanaidh, W.J. Dowling, F.M. Boland, Phase Watermarking of Digital Image, Proc. IEEE Int. Conf. on Image Processing, Vol. 3, Lausanne, Switzerland, 1996, pp. 239-242.

[20] Hsiang-Cheh Huang, Lakhmi C. Jain, Jeng-Shyang Pan "Intelligent Watermarking Techniques", World Scientific Pub Co Inc, 2004.

[21] T. Liang and J.J. Rodriguez. Robust Watermarking Using Robust Coefficients, Security and Watermarking Multimedia Contents II, SPIE,3971, 2000, pp. 326-335.

[22] A. Khan, A. M. Mirza, A, Majid, Optimizing perceptual shaping of a digital watermark using genetic programming, Iranian Journal of Electrical and Computer Engineering (IJECE), vol. 3, no. 2, 2004, pp. 144-150.

[23] A. B. Watson, Visual optimization of DCT quantization matrices for individual images, in Proc. AIAA Computing in Aerospace 9, San Diego, CA, (1993), pp. 286–291.

[24] J. A. Solomon, A. B. Watson, and A. J. Ahumada, Visibility of DCT basis functions: Effects of contrast masking, in Proc. Data Compression Conf., Snowbird, UT, 1994, pp. 361–370.

[25] Christian J. van den Branden Lambrecht and Joyce E. Farrell, Perceptual Quality Metric for digitally coded color images,Proc. EUSIPCO, 1996, pp. 1175-1178.

[26] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, Visibility of Wavelet Quantization noise, IEEE Trans. on Image Processing 6, 1997, pp. 1164-1175.

[27] C. Fei, D. Kundur and R. H. Kwong, Analysis and Design of Watermarking Algorithms for Improved Resistance to Compression, IEEE Trans. on Image Processing, vol. 13, No. 2, 2004, pp.126-144.

[28] G. Brown, J. Wyatt, R. Harris and X. Yao, Diversity Creation Methods: A Survey and Categorization, Journal Information Fusion 6, 2005, 5-20.

[29] A. J. Ahumada and H. A. Peterson, Luminance-model-based DCT quantization for color image compression, Proc. SPIE on Human Vision, Visual Processing, and Digital Display III, vol. 1666, 1992, pp. 365–374.

[30] M. Kutter and S. Winkler. A Vision-based Masking Model for Spread-Spectrum Image watermarking. IEEE Trans. on Image Processing, 11(1), 2002, pp. 16-25.

[31] F. I. Koprulu, Application to low-density parity-check codes to watermark channels, Ms. Thesis, 2001, Electrical and Electronics Bogaziei University, Turkey.

[32] Z. Wang, A. C. Bovik H. R. Sheikh, Image quality assessment: From error measurement to structure similarity, IEEE Trans. on Image Processing, Vol. 13, No. 1, 2004.

[33] W. Banzhaf, P. Nordin, R. E. Keller and F. D. Francone, Genetic Programming An Introduction: On the Automatic Evolution of Computer Programs and Its Applications, Morgan Kaufmanns Publishers, Inc. San Francisco, California, 1998.

[34] Mengjie Zhang, William D. Smart, Multiclass Object Classification Using Genetic Programming. EvoWorkshops 2004, pp. 369-378.

[35] A. Khan, A. Majid and Anwar. M. Mirza, Combination and Optimization of Classifiers in Gender Classification Using Genetic Programming, KES journal, Netherlands, Vol. 8, 2004, pp. 1-11.

[36] http://gplab.sourceforge.net

[37] Sara Silva, Jonas Almeida, Dynamic Maximum Tree Depth - A Simple Technique for Avoiding Bloat in Tree-Based GP, in Proc. of the Genetic and Evolutionary Computation Conference (GECCO-2003), Chicago, Illinois, USA, July-2003, pp. 1776-1787.

[38] P. Nordin, M. Brameier, F. Hoffmann, F. Francone, and W. Banzhaf, AIM-GP and Parallelism, Proceedings of Congress on Evolutionary Computation, Washington, 1999, IEEE Press, Piscataway, NJ, pp. 1059 – 1066.

[39] W. Kantschik and W. Banzhaf, Linear-Graph GP—A new GP Structure, EuroGP 2002, Kinsale, Ireland, Springer LNCS 2278, Berlin, 2002, pp. 83-92.