Information Sciences 179 (2009) 791–808

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Fidelity-guaranteed robustness enhancement of blind-detection watermarking schemes

Chun-Hsiang Huang [a], Ja-Ling Wu [a,b,*]

[a] Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan
[b] Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei 106, Taiwan

ARTICLE INFO

Article history:
Received 19 October 2006
Received in revised form 11 September 2008
Accepted 27 October 2008

Keywords:
Digital watermarking
Fidelity
Robustness
Genetic algorithms (GAs)

ABSTRACT

Watermarking performance enhancement has always been a difficult task since the performance metrics of watermarking systems, i.e., fidelity, robustness, and payload size, inherently conflict with each other. Nowadays, most watermarking schemes hide payloads according to predefined rules or empirical perceptual models. Therefore, the performance of digital watermarking schemes can be determined only passively. In this study, a genetic algorithm-based framework for watermarking performance enhancement is proposed. Watermarked results having better robustness, guaranteed fidelity, and fixed payload size can be obtained. Existing blind-detection watermarking schemes can be improved significantly by incorporating the proposed framework. In addition, the proposed framework has many desirable advantages such as asymmetric embedding/detection overhead, easy integration with existing data-hiding schemes, and direct control over fidelity and robustness.

1. Introduction

In the last decade, various watermarking schemes have been proposed, and significant progress has been made in this field. A comprehensive introduction to existing watermarking technologies and theoretical foundations can be found in [11]. However, designing optimal watermarking schemes is still an open problem since satisfactory and feasible solutions have not been obtained, as stated in [12]. Difficulties in designing optimal watermarking schemes arise from the three conflicting performance metrics of watermarking systems: high fidelity of embedded contents, strong robustness of the hidden information against common processes or malicious attacks, and large payload size. Fig. 1 shows the performance space spanned by these performance metrics. Note that the required payload size often depends on the objective of watermarking schemes. In other words, as long as a predefined amount of embedded data can accommodate necessary information, e.g., author information for copyright protection or the usage rules specified by DRM-enabled consumer-electronic devices, the payload size can be regarded as a fixed parameter. Further, the fidelity-robustness plane is sufficient for discussing watermarking performance. The model illustrated in Fig. 1 can be applied to other important data-hiding schemes also. For example, the designers of steganographic schemes are more concerned about the fidelity and payload size of the schemes than their robustness due to the nature of secret communications. Thus, the plane spanned by fidelity and payload size is sufficient for the performance evaluation of a steganographic system.

However, even for a fixed payload size, deciding an optimal trade-off between fidelity and robustness is not an easy task. Fig. 2 conceptually illustrates the relationship between fidelity and robustness in the proposed fidelity-robustness plane. We indicate the watermarked contents of a certain watermarking scheme by circles. In general, the higher the fidelity, the lower

* Corresponding author. Address: Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei 106, Taiwan. Tel.: +886 2 33664888x214; fax: +886 2 33664898.
E-mail addresses: bh@cmlab.csie.ntu.edu.tw (C.-H. Huang), wjl@cmlab.csie.ntu.edu.tw (J.-L. Wu).

0020-0255/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2008.10.035

Fig. 1. Performance space of various data-hiding schemes.

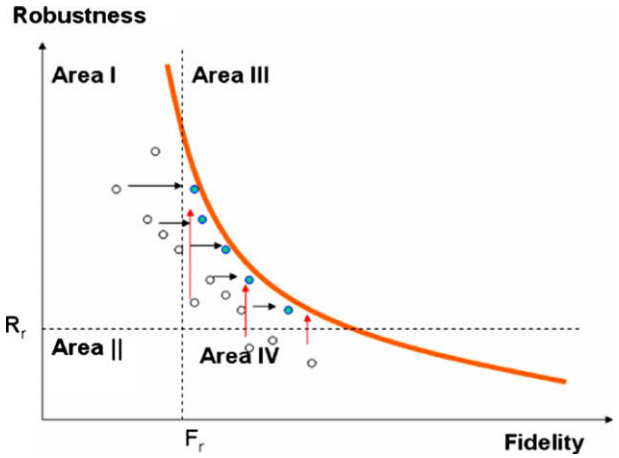

Fig. 2. Fidelity-robustness plane containing watermark embedded results. The curve represents the inherent performance limit of some watermarking algorithms. The watermark-embedded outcomes of the proposed scheme, indicated by empty circles, are determined according to predefined rules or perceptual models. Assuming that the desired embedded results of some watermarking applications must have robustness higher than R_r and fidelity higher than F_r, only works corresponding to the circles locate within Area III and below the curve are feasible. Better embedded results like those indicated by solid circles may be obtained. The vertical and horizontal arrows indicate the possible robustness and fidelity performance enhancements, respectively.

is the robustness of hidden payloads against processing/attacks. In existing watermarking schemes, perceptually acceptable watermarked results are produced according to predefined embedding rules or empirical perceptual models. The robustness of the schemes against various attacks is often evaluated by carrying out extensive experiments. As a result, the region in which embedded outcomes may be located is determined by the adopted watermarking scheme, original/watermarked pairs, and a range of adjustable parameters. Although the chosen watermarking scheme may have potentially better performance, conventional watermarking schemes are unable to produce better embedded outcomes. In these schemes, performance information obtained by an extraction module is often neglected; instead, the information should be used to produce embedded outcomes that approach the performance limit.

Conventionally, optimal watermarking used to be carried out by using theoretical models or devising more elaborate embedding rules. For example, the information-theoretic analysis of general watermarking schemes has been carried out assuming specific stochastic models for hosts, watermarks, and attacks in [24,25]. Quantization watermarking schemes showing good performance and feasible implementations have also been developed in [4–6]. Despite all this progress made in the field of watermarking schemes, the importance of incorporating watermarking performance into embedding modules is still overlooked. Some informed-embedding watermarking schemes in which the original content and potential attacks are regarded as auxiliary information have been introduced in [22,23]. The relationship between our framework and existing informed-embedding schemes will be discussed later.

In order to obtain watermarked results showing better performance, we simulate the embedding process as an optimization problem. Further, to obtain satisfactory solutions in the given computation time or under the resource constraints, genetic algorithm (GA)-based optimization is applied. The detection performances of watermarked results are now actively used as objective function values that drive iterative optimization processes to find better embedding results.

Watermarking schemes can be roughly classified into two categories: informed-detection watermarking schemes (in which watermark detectors require unmarked content) and blind-detection watermarking schemes (in which watermark detectors do not require the unmarked original content). Since the applicability of the latter is considerably wider that of the former and in many important watermarking applications, the unmarked content is unavailable, only the performance enhancement techniques for blind-detection watermarking schemes will be discussed in detail. The performance enhancement of informed-detection watermarking schemes will be briefly covered in the next section.

This paper is organized as follows. Section 2 presents a brief introduction to the adopted evolutionary computational technique—GAs—as well as a literature survey of existing studies on GAs and digital watermarking schemes. Section 3 describes the proposed watermarking performance enhancement framework and compares it with conventional watermarking schemes. A GA-based robustness enhancement scheme having guaranteed fidelity and a fixed payload size is discussed in detail in Section 4. Implementation details, experimental results, and related discussions are also presented therein. In Section 5, some interesting implications and architectural advantages of the proposed framework are discussed. Finally, Section 6 presents our conclusions.

2. Genetic algorithms and digital watermarking

GAs are important optimization techniques belonging to the field of evolutionary computation [13]. They are different from conventional optimization techniques in the following aspects:

- GAs work with a coding of a parameter set, and not with the parameters themselves.
- GAs search a population of points, and not a single point.
- GAs use objective information, rather than the derivatives or other auxiliary information, in common optimization functions.
- GAs use probabilistic transaction rules, and not deterministic transaction rules.

In the beginning of a GA-based optimization process, the natural parameter set of the optimization problem is coded as a finite-length string, named the chromosome, over some finite alphabet set. Since GAs work simultaneously with a population of points in the parameter space, the probability of being trapped in a false optimum is low. In practical GA-based optimization, three operators—reproduction, crossover, and mutation—are usually applied to chromosomes repeatedly. Reproduction is a process in which individual chromosomes are copied according to their objective function (fitness) values. The chromosomes with higher fitness values have higher probability of producing larger number of offsprings in the next generation. It is the objective function that decides the probability of chromosomes' survival or elimination during evolutionary competitions. This reproduction methodology is also called the biased-roulette-wheel approach. Crossover is a procedure in which pairs of chromosomes exchange portions of their genes to form new chromosomes. Consequently, new parameters in addition to the initial parameters can be obtained for further evaluation. Mutation is an occasional random alternation of the value in some positions of chromosomes. Mutation can be treated as a random walk through the parameter space. Sometimes, chromosomes that show good performance but cannot be obtained through reproduction and crossover can be obtained through mutations. While solving problems with constraints, the reproduced chromosomes may violate predefined conditions; hence, checking and discarding operations are included in the framework. These operators are used repeatedly to generate successive generations of chromosomes. Among the chromosomes of the same generation, only a part of them can survive and become parent chromosomes to generate offsprings for the next generation. After a series of computations, the near-optimal solutions of desired parameters for the original problem can be obtained. Fig. 3 illustrates the process for solving a very simple searching problem by using GAs. A detailed explanation of GAs can be found in [10,14]. In the recent years, GAs have been proposed to solve various important problems such as cipher breaking [26], software project management [1], financial computing [2,7], and particle navigation control [3].

In our previous studies, the concept of GA-based enhancement of informed-detection watermarking schemes by searching the best embedding positions has been introduced [17,18]. To the best of our knowledge, these are the earliest publications that have combined GA-based optimization techniques and digital watermarking schemes. Moreover, a lightweight searching algorithm for a small range problem has also been proposed in [21] to reduce the long computation time required in the original scheme. The concept of joint fidelity/robustness optimization and another GA-based spatial domain watermarking algorithm has been proposed in [27]. In this scheme, spatial watermarked image blocks generated on the basis of randomly selected keys are simulated as points to be searched in the solution space and are further analyzed by the GA-based approach. However, the use of this scheme is limited due to the low robustness of its spatial domain embedding nature. The secret key delivery problem could not be solved by using these two schemes. Furthermore, the number of possible solutions is limited to certain embedding results corresponding to random keys, not exactly matching with the characteristics and the limits of the watermarking scheme.

Nevertheless, all the GA-based enhancement schemes have been introduced to be used along with the informed-detection watermarking schemes and inevitably suffer from several serious shortcomings. For example, after GA-based enhancement, either the final embedding positions must be transmitted to the watermark detector as parts of secret information or the watermark detector must perform the time-consuming optimization operation under the condition that the original

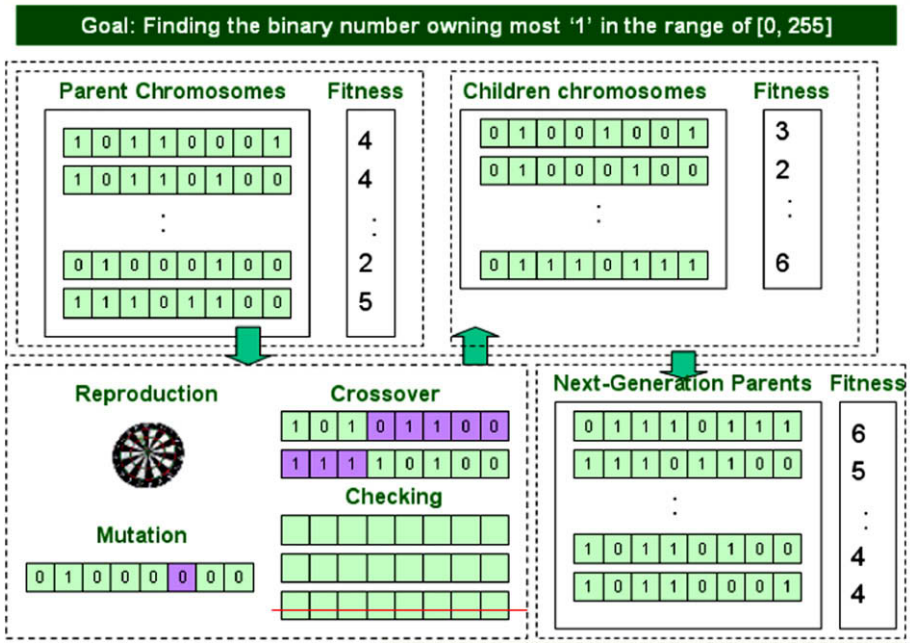

Fig. 3. Process for finding binary number with largest number of "1 s" within range of [0,255] by GA optimization. Candidates are simulated as 8-bit binary strings, and the fitness (objective function) value is defined as the number of "1 s." This figure illustrates only the iteration process of one generation.

content is available. Undoubtedly, both remedies are not applicable in important watermarking applications such as copyright protection of broadcast video or DVD copy prevention. Therefore, another watermarking enhancement scheme having better robustness against various types of attacks, guaranteed fidelity, and application-specific capacity is devised in this study. This new optimization framework is inherently suitable for blind-detection watermarking schemes, and it eliminates the need of any redundant secret information that must be sent to or transmitted by the detector. Furthermore, its asymmetric embedding/detection overhead can effectively reduce the long computation time. Finally, it has several desirable architectural characteristics, e.g., embedding and detection can be carried out in different domains. Flexibility in integrating with existing watermarking schemes is an added advantage of the proposed scheme. Some experimental results can be found in [19,20]; however, the complete conceptual illustrations, implementation details, and experimental results are presented in this paper.

Recently, important GA-based watermarking schemes have been proposed in [28,29]. GAs are used to determine the rule for the conversion of real numbers into integers during the cosine transformation so that the errors in embedded watermarks can be significantly reduced. The applications of GAs in both fragile watermarking and robust watermarking are also discussed.

3. Proposed enhancement framework

Conventional watermarking schemes embed watermarks in a heuristic manner. As shown in Fig. 4a, in the embedding module, predefined rules or complex perceptual models are often adopted as guidelines during watermark embedding. Since the embedded results, relying on the predefined rules, always show the same performance, conventional watermarking schemes cannot approach the inherent performance upper limit, as illustrated in Fig. 2. If the performance of embedded outcomes is not satisfactory in some respects, the only solution is to adjust the embedding parameter empirically so that better watermarked results are obtained. However, empirically setting the parameters is an awkward process that lacks systematic techniques.

Watermarking is often modeled as a communication channel with auxiliary information (information about the host) while discussing the host-interference rejection problem, as stated in [5,6]. However, if we investigate the watermark embedding schemes more thoroughly, it is obvious that the host content is not the only available auxiliary information. As long as the performance evaluation functions can be incorporated, watermark embedders can predict the performance of each embedding operation in advance. In short, the watermarking performance originally evaluated in the detector side can be viewed as another type of "auxiliary information." Thus, this information shall be considered by watermark embedders to produce better embedding results.

The proposed watermarking performance enhancement framework is shown in Fig. 4b. Instead of using predefined rules or complex perceptual models, performance-driven optimization procedures are included in the embedding module. The

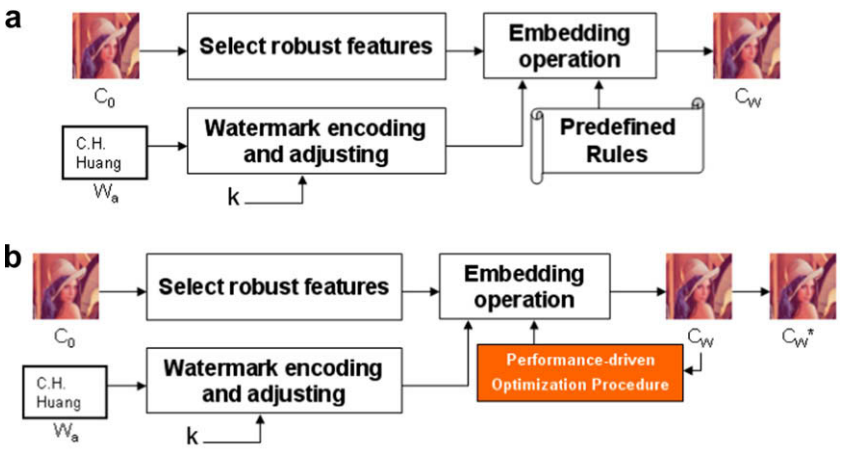

Fig. 4. (a) Traditional watermarking schemes using either predefined rules or empirical perceptual models for embedding operations. (b) Proposed framework employing performance-driven optimization procedures to obtain optimal embedded results.

performance of each embedding operation can be easily predicted and/or obtained by incorporating a performance evaluator in the embedder. Further, the watermark embedder can use the evaluated performance to produce better results. In the extreme cases, the watermark embedder may simply choose the best watermarked results out of all candidates as the final output, rather than simply performing heuristic embedding operations.

Now, the watermarking performance enhancement leads to another problem: can watermark embedders accurately predict watermarking performance? Evaluating fidelity is easy since various subjective or objective quality metrics for digital contents have been proposed. Although whether these metrics are sufficiently close to those of the human vision system (HVS) or the human auditory system (HAS) is still an open problem, this question is not dealt in this study. On the other hand, because various types of attacks should be taken into consideration, a criterion for evaluating robustness against all types of attacks is not feasible. Fortunately, the detection values against certain malicious attacks and common media processes can be measured. For example, as long as a JPEG quantization/dequantization module and a watermark extraction module are included in the embedder, the robustness of an image watermarking scheme against JPEG compression can be precisely measured. This implies that the embedded contents with additional performance information are robust against JPEG compression. The potential methodologies of incorporating various attack models and the difficulties involved therein will be discussed in the following sections.

4. Fidelity-guaranteed robustness enhancement of blind-detection watermarking schemes

4.1. Description of testbed system: a simple blind-detection watermarking scheme

The proposed performance enhancement scheme can be employed to enhance various existing blind-detection watermarking schemes. In order to justify the performance enhancement capability, a blind-detection variant of the block-DCT-based image watermarking scheme, originally introduced in [16], is used to evaluate the performance of the proposed enhancement scheme without loss of generality.

In our testbed watermarking system, watermark sequences are binary strings (with their bit values represented by either 1 or -1) obtained by binarizing real-valued pseudo-random sequences obtained according to the $N(0,1)$ Gaussian distribution. The watermark sequences possess good statistical characteristics such as roughly equal number of 1 s or -1 s and low correlation between two different watermark sequences. Without loss of generality, the watermark bits are embedded at their predefined positions within the middle-band coefficients of each DCT block, as specified in [16]. The inequality relationship between the magnitudes of AC coefficients at selected positions and that of the scaled DC coefficients in the same block is modified to represent the watermark bits. The effect of JPEG quantization is also considered while designing the embedder. The embedding process for the selected AC coefficient within a DCT block is given by

$$V'_{AC} = \begin{cases} sign(V_{AC}) \cdot \left(\frac{V_{DC}}{Q_{DC} \cdot S_{DC}} + \alpha \frac{Q_{AC}}{S_{AC}} \right), & \text{if } w = 1, \\ sign(V_{AC}) \cdot \left(\frac{V_{DC}}{Q_{DC} \cdot S_{DC}} - \alpha \frac{Q_{AC}}{S_{AC}} \right), & \text{if } w = -1, \end{cases} \tag{1}$$

where V_{AC} and V'_{AC} are the original and modified values of the selected AC coefficient, respectively; w denotes the watermark bit that will be embedded in the selected AC coefficient; $sign(\circ)$ is a function that maps an input value to 1 or -1, indicating whether the input value is positive or negative; V_{DC} denotes the DC value of the current block; Q_{AC} is the value in the JPEG luminance quantization table corresponding to the location of V_{AC}; Q_{DC} is the value of the DC coefficient in the JPEG

luminance quantization table; S_{DC} and S_{AC} are predefined scale factors and are set to 8 and 10, respectively, throughout the following experiments; α is the user-specified watermark strength during watermark embedding and is always a positive value. In general, larger the α value, the higher is the robustness against manipulations or attacks and the higher is the number of severe distortions. An adequate number of α values is often determined after performing a series of empirical tests.

The corresponding extraction process is given by

$$\tilde{w} = \begin{cases} 1, & \text{if } |\widetilde{V}_{AC}| > \frac{\widetilde{V}_{DC}}{Q_{DC} \cdot S_{DC}}, \\ -1, & \text{if } |\widetilde{V}_{AC}| \leqslant \frac{\widetilde{V}_{DC}}{Q_{DC} \cdot S_{DC}}, \end{cases} \tag{2}$$

where $\widetilde{V}_{DC}$ and $\widetilde{V}_{AC}$ are the values of the DC and AC coefficients in the DCT block of the watermarked (and potentially distorted) image, respectively, and $\tilde{w}$ is the extracted watermark bit.

The testbed watermarking system is in fact a variant of the quantization watermarking scheme. The magnitude of the chosen AC coefficient will be quantized to a reconstruction point. Note that, in this scheme, the decision threshold is determined by the scaled/quantized DC value and the quantization step is decided by the user-specified watermark strength α and the scaled JPEG quantization value corresponding to the AC coefficient.

Fig. 5 shows the empirical performance of this watermarking system used to embed a watermark sequence in different test images. The horizontal axes represent the peak signal-to-noise ratio (PSNR) value, in decibels, of the watermarked images. The vertical axes represent the normalized correct extraction ratio (CER, i.e., the number of correctly extracted watermark bits divided by the number of total watermark bits). All the images are uncompressed gray-level images of

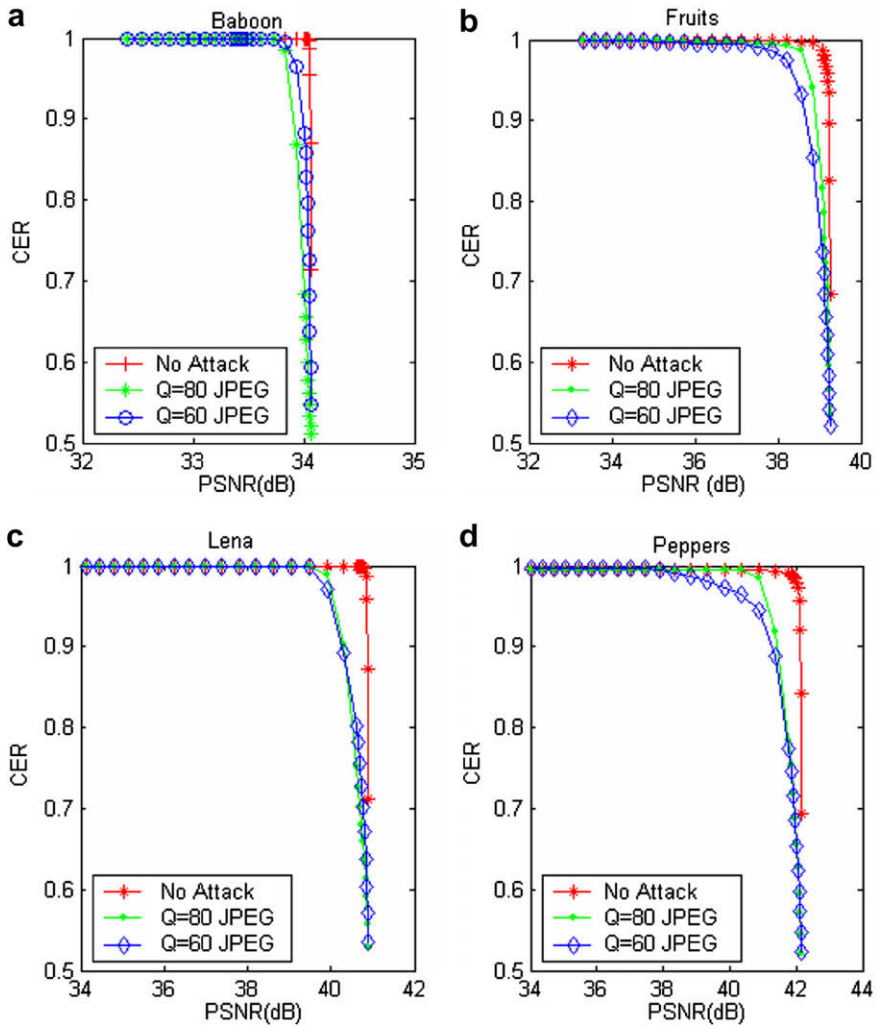

Fig. 5. Empirical performance obtained by using our watermarking scheme to embed 16,384 bits into different 512×512 images. The horizontal axes represent the PSNR value just after watermark embedding, and the vertical axes represent the normalized correct extraction ratio (CER).

512 × 512 pixels. Four watermark bits are embedded in each 8 × 8 DCT block; in other words, a total of 16,384 bits are embedded in each test image. These performance curves are obtained by employing different watermarking strengths, i.e., substituting different values of α into Eq. (1). In these experiments, the watermarked results are obtained by using respective α values ranging from 0 to 5. Note that, in Fig. 5, some experimental results show low CER values even in cases where no attack takes place. This phenomenon is normally observed since the corresponding watermarking energies (α) are extremely small, and the phenomenon implies that the parameter settings may be inadequate to be applied in practical scenarios. Here, these extraction results are incorporated to clearly illustrate the performance of the testbed system for well-known test images. Moreover, since the design of this scheme mainly depends on JPEG quantization, the scheme has good robustness against JPEG compression. The performance of the testbed system against other attacks will be discussed in the later sections.

4.2. Incorporation of proposed GA-based enhancement framework with existing watermarking schemes

Fig. 6 depicts the flowchart of the proposed enhancement framework. The fidelity of the embedded content is specified by users according to different application scenarios and will be guaranteed throughout the entire embedding process. To represent the fidelity constraint concretely, the PSNR value between the embedded image and the original image is used.

The proposed optimization process is carried out in a block-by-block manner. For each 8 × 8 image block, a set of M initial parent chromosomes are generated. Each initial parent chromosome is a randomly generated 8 × 8 block in which the value of each component is uniformly distributed over a symmetric range with respective to the origin. In other words, every parent chromosome represents a potential distortion block, defined as the difference between the original image block and the embedded image block. Next, an energy-shaping module is applied to all initial parent chromosomes so that they satisfy the predefined fidelity requirement. For example, according to the definition of PSNR, if the user requests for a PSNR value higher than 40 dB, the maximum allowable block energy, i.e., the sum of the squared values of the distortion pixels within an 8 × 8 chromosome block, shall be less than 416. If the energy of a randomly generated chromosome block is higher than the maximum allowed energy, the chromosome block will be uniformly scaled down to satisfy the fidelity limit. The difference between the energy of the energy-shaped chromosome and the maximum energy limit will be further decreased by slightly adjusting the randomly selected pixels. In this manner, each pre-processed parent chromosome in fact represents the difference between a potential embedded block satisfying the fidelity constraint and the original block.

Each energy-shaped parent chromosome will be added to the original image block to form an embedded candidate. Then, the fitness value associated with each candidate is calculated. According to the prescribed watermarking scheme, the fitness function value F of a candidate block is defined as

$$F = \sum_{k=1}^{C} F^{(k)}, \tag{3}$$

$$F^{(k)} = \begin{cases} \beta_1 + \beta_2 \cdot \left| |\widetilde{V}_{AC}^{(k)}| - \frac{\widetilde{V}_{DC}}{Q_{DC} \cdot S_{DC}} \right|, & \text{if } sign(\tilde{w}^{(k)}) = sign(w^{(k)}), \\ 0, & \text{otherwise}, \end{cases} \tag{4}$$

where C is the number of watermark bits to be embedded in a single block and $F^{(k)}$ is the fitness value that the kth extracted watermark bit $\tilde{w}^{(k)}$ contributes to F. If the sign of $\tilde{w}^{(k)}$ is same as that of the reference watermark bit $w^{(k)}$, the watermark extractor can successfully extract $w^{(k)}$, and the fitness value shall be increased. The amount of increment is controlled by two weighting factors: β_1 and β_2. β_1 controls the difference in the fitness values of the case that "a watermark bit is correctly extracted" and the case that "the extracted watermark bit is wrong". In other words, β_1 measures the correctness of the embedded watermarks in each coefficient. On the other hand, β_2 controls the dependence of the fitness value on an embedded coefficient above the detection threshold. The other symbol are the same as those given in Eq. (1).

Next, a set of N child chromosomes will be reproduced according to GAs-based rules. Parent chromosomes corresponding to higher fitness values have higher probability to generate more offsprings. Then, the reproduced child chromosomes are randomly grouped into pairs and arbitrary portions of chromosomes are exchanged. The parts of two child chromosomes will be combined to form a new 8 × 8 distortion block. Finally, each pixel component of the child chromosomes has a chance to mutate and generating new candidates.

Although these child chromosomes are generated through their parent chromosomes, the adopted GA operations may result in child chromosomes violating the specified fidelity requirement. Thus, the prescribed energy-shaping process will be applied to these newly generated child chromosomes too.

Finally, a fitness-based selection policy is used to select M next generation parent chromosomes from a set consisting of M parent chromosomes and N child chromosomes. The complete GA-based optimizing processes will be repeated until a certain number of iterations (named as the number of generations in a GA-based approach) have been attained. Finally, the chromosome with the highest fitness value will be added to the original image block; this block is expected to function an embedded block having the best robustness.

Note that the watermark is never explicitly "embedded" in the original block in the proposed scheme. The best candidate under the fidelity constraint is found according to the simulated detection performance. The watermark extraction module is identical to the watermark detector of the original watermarking scheme. This asymmetric embedding/detecting nature of

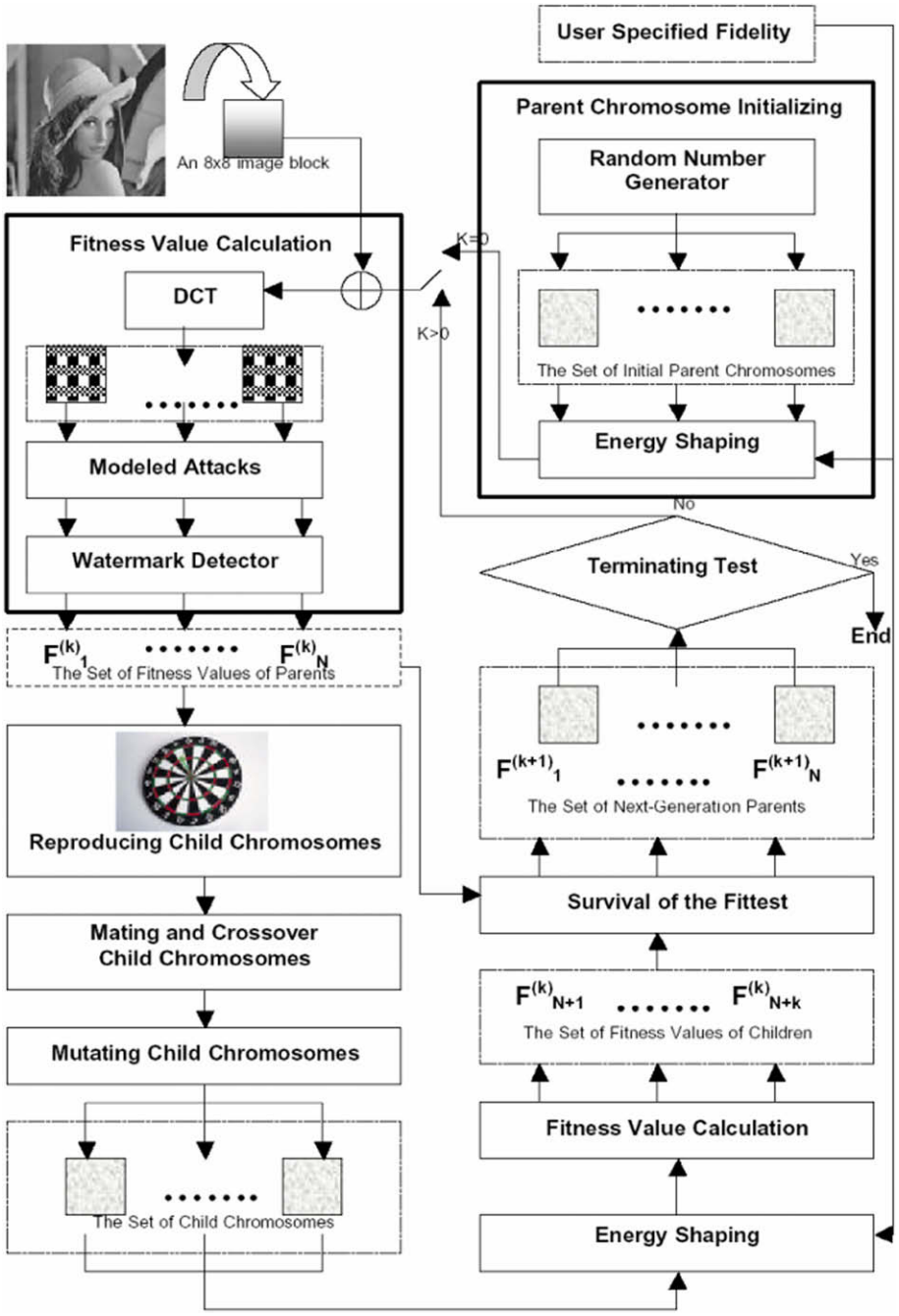

Fig. 6. Integrating GA-based performance enhancement framework with existing watermark scheme. The iterative optimization ends when the iteration number exceeds a predefined limit or an acceptable embedding candidate is obtained.

the modules is quite different from that of conventional watermarking schemes. To be more specific, the proposed embedding module can be regarded as an iterative performance enhancement stage depending on the given watermark detection algorithm and the chosen performance indexes. Similar optimization procedures can be applied to various blind-detection watermarking algorithms as long as watermark detectors are present. Moreover, the performance indexes can be reasonably replaced with suitable indexes, e.g., the objective PSNR can be replaced with subjective perceptual index.

From the viewpoint of the proposed fidelity-robustness model, the roles played by its each operation can be clearly identified. As shown in Fig. 7, the crossover and mutation operators find new embedding candidates, and the energy-shaping module modifies over-distorted images so that the required fidelity constraint is satisfied, and the fitness-based selection process will hopefully generate better parent chromosomes in the long run.

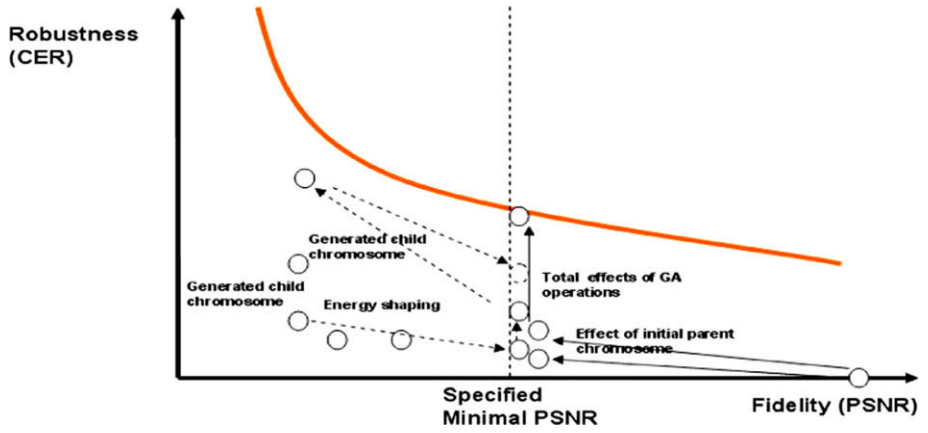

Fig. 7. Effects of components used in proposed enhancing scheme from viewpoints of performance-based watermarking model.

4.3. Implementation details and some experimental results

The implementation of GA modules is based on the MatLab Genetic Algorithm Toolbox developed by Evolutionary Computation Research Group, The University of Sheffield. A detailed introduction to the potential applications and use of this toolbox can be found in [8,9].

In the proposed framework, since the distortion blocks corresponding to each image block in the spatial domain are modeled as candidate chromosomes, real-valued GA optimization is chosen as the basic optimization technique. The number of parent chromosomes (denoted by M) is set to 40 for all experiments after considering the typical GA settings. The generation gap (i.e., the ratio between the number of parent chromosomes and that of child chromosomes) is set to 0.5; thus 20 children (i.e., $N = 20$) will be produced after all GA operations are carried out in one generation. Though the range of values of each gene in the initial chromosomes that corresponds to the distortion of each pixel location is set to $[-128, 128]$, the chromosomes that are subjected to the energy-shaping process will consequently satisfy the fidelity constraint. A discrete recombination is employed to perform the crossover operations on real-valued chromosomes so that the crossover behavior is conceptually similar to that of the uniform crossover operator of binary-valued chromosomes. Since the distortion block corresponds to a large search space, the mutation rate is set to 0.25 and the mutation range for real-coded GAs is set to $[-20, 20]$, for an effective search for all possible candidates.

The weighting factors β_1 and β_2, specified in Eq. (4), are set equal to 100 and 1, respectively. The obvious difference between the two weighting factors will guarantee that candidate blocks corresponding to more correctly extracted payload bits will have great advantages during evolutionary competitions. However, if two candidates possessing similar number of correctly extracted payload bits are compared, the one with higher robustness will have higher chance of survival since the effect of coefficient magnitude over the detection threshold value is also included in the objective function.

Fig. 8 shows the GA-based performance enhancement of several test images against JPEG compression. The generation number is set to 100. The built-in image processing subroutines of MatLab are used to produce JPEG compressed images, and the quality parameter is set to 80. The performance curves of the GA-based enhancement scheme confirm the assumption that robustness increases with decreasing fidelity. Furthermore, it should be noted that for embedded images of excellent visual quality, e.g., when the PSNR value is higher than 40 dB, the percentage of correctly extracted watermark bits is high and sufficient to identify the existence of a watermark. In contrast, the original testbed system with a simple parameter setting fails to produce watermarked contents under such high fidelity requirement.

4.4. Improving performance by increasing computational resources

Though the proposed GA-based scheme can extend the achievable performance, there is still room for improvement. Since the proposed enhancement scheme begins with randomly generated initial parent chromosomes, only improvement in robustness against nonmarked images can be guaranteed. For cases in which a low fidelity requirement is specified, the search space is consequently large; thus, trapping in a local optimum may take place easily when only limited computational resources are employed.

To avoid obtaining solutions corresponding to local optimum, several solutions can be of use. The most intuitive approach is to increase computational resources, e.g., increase the numbers of optimization iterations. The larger the number of iterations, the higher is the probability of obtaining better candidates (according to the definition of objective functions). Fig. 9 shows the obtained robustness performance using the Lena image under the fidelity requirements of 40 and 42 dB against JPEG compression. As the allowable generation number increases, the CER increases accordingly. It should be noted that the embedded results obtained by a large number of iterations can resist attacks better than those obtained by a small number of

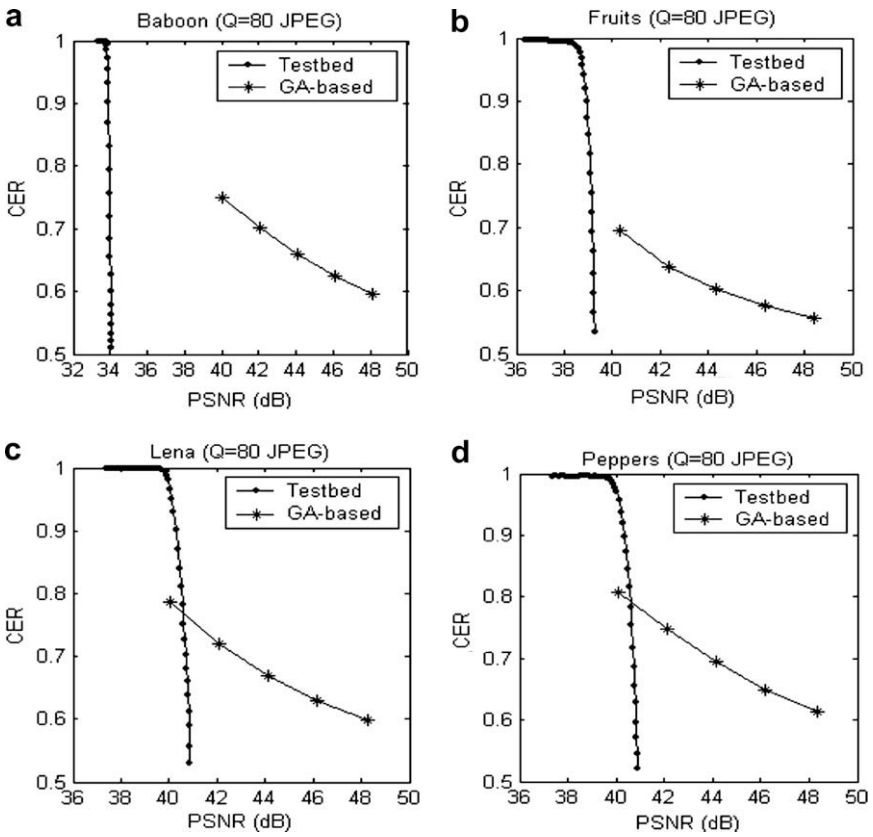

Fig. 8. Comparisons of GA performance enhancement with original performance limit in terms of CER for various test images. The demand for high-quality marked content is successfully addressed.

iterations. This phenomenon also justifies that the settings of the weighting factors defined in Eq. (4) are reasonable, where β_1 guides the search for watermarked candidates possessing more correctly extractable payload bits in subsequent iterations, and β_2 leads to candidates possessing stronger marking energies. The experimental results obtained using other test images are similar to those illustrated in Fig. 9.

Though increasing computational resources may naturally and effectively produce better results, the feasibility of this approach heavily relies on the application scenarios. For scenarios in which increasing computational resources is expensive or not possible, performance enhancement cannot be facilitated only through this approach.

4.5. Better population initialization

Although GAs are generally applicable to a wide range of problems in which very little information is available, there are still many opportunities to incorporate problem-specific heuristics into GA-based systems. As illustrated in [15], the GA stages that are randomly selected, including population initialization, crossover, and mutation, can be replaced by heuristic methods to obtain better candidates. According to Grefenstette [15], incorporating heuristic population initialization can produce good results quickly.

In the case of watermarking performance enhancement, heuristic initial chromosomes can be obtained by incorporating the watermarked results produced by a testbed system. For example, in [28], the original image and the rounded watermarked image are compared to generate initial chromosomes in order to speed up the iterative process. In other words, better watermarked results are assumed to be located in the vicinity of the heuristically obtained results.

To test this assumption, 10 watermarked results for each test image are generated using different watermarking strengths, i.e., different α values, where the α values are uniformly distributed from 1 to 10. In the beginning of the iteration for each block, the distortion blocks are obtained by subtracting the unmarked block from the heuristically generated blocks and are inputted into the proposed system to substitute the portions of the randomly generated initial chromosomes. The energy-shaping process is also applied to these heuristically generated chromosomes to guarantee the minimum PSNR requirement.

Fig. 10 shows the enhancement of CER against JPEG compression after incorporating heuristic initial chromosomes. The enhancement is obvious under the constraint that the same numbers of optimization iterations are performed. Figs. 11–13

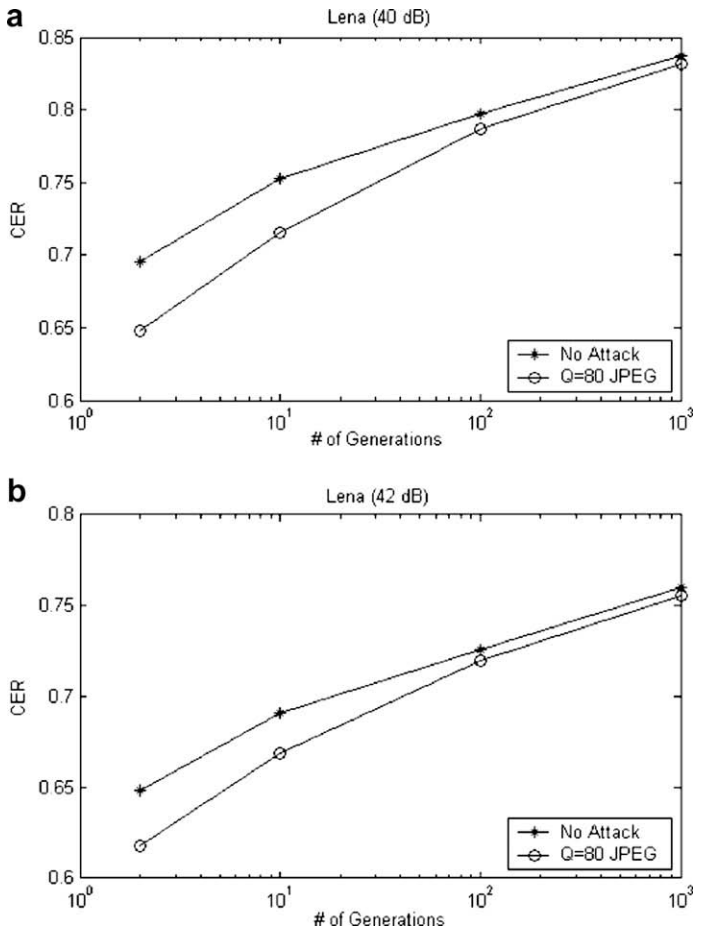

Fig. 9. GA-based robustness enhancement of Lena image under minimum PSNR of (a) 40 dB and (b) 42 dB. In each experiment, the CER increases with respect to the increase in the generation numbers.

show the fidelity-robustness performance of different test images against histogram equalization, resizing, and cropping, respectively. All the attacks are carried out using the MatLab image processing toolbox, where the default settings are used if the settings are not specifically mentioned. The resizing attack is carried out by shrinking the 512×512 test images to 384×384 images and then restoring to the original 512×512 dimensions. The cropping operation is performed by preserving only the 384×384 bottom-right area of each test image. All experimental results show significant enhancement.

4.6. Incorporating attack models as available auxiliary information

The flexible enhancement framework presents a large number of possibilities to enhance robustness against certain types of attacks. As long as attack simulators are incorporated in the proposed system in front of the watermark extractor, i.e., all candidate blocks are subjected to attacks before the evaluation of the objective function, the so-obtained watermarked results have better robustness against the considered attacks than those generated by systems without incorporating attack models.

To justify this assumption, a JPEG compression simulation module, as shown in Fig. 14, is incorporated with the proposed framework. Due to the large amount of memory/computational resources required to perform the complete JPEG simulation on all candidate blocks iteratively, the simulation module is simplified to estimate the loss in information each block due to the JPEG compression. This simulation module uses the standard JPEG quantization table. The images produced by the module are nearly identical to those compressed by common JPEG encoders with a quality factor of 50.

Fig. 15 shows the embedded results obtained by the GA-based scheme and the GA-based scheme incorporated with the JPEG simulation module (both with better population initialization). As expected, the later produces better embedded image. However, we have also found that the performance improvement is not guaranteed for all real-world JPEG compression settings. This inconsistency is mainly caused by the difference in the implementation details of the simulated attack module and the real-world attack model, e.g., adopting different quantization tables or using different DCT subroutines. In fact, this

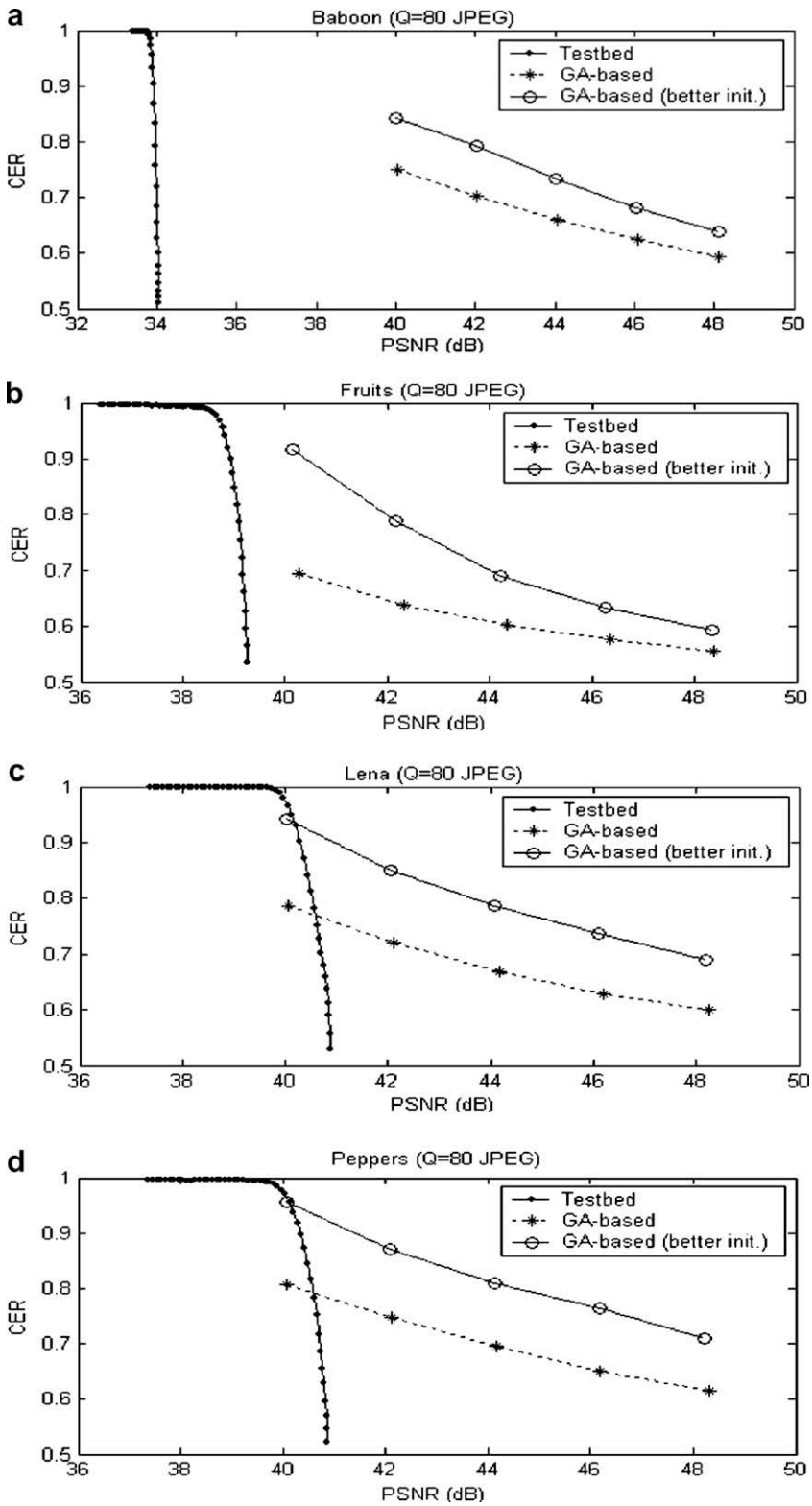

Fig. 10. Robustness performance against JPEG compression of different images.

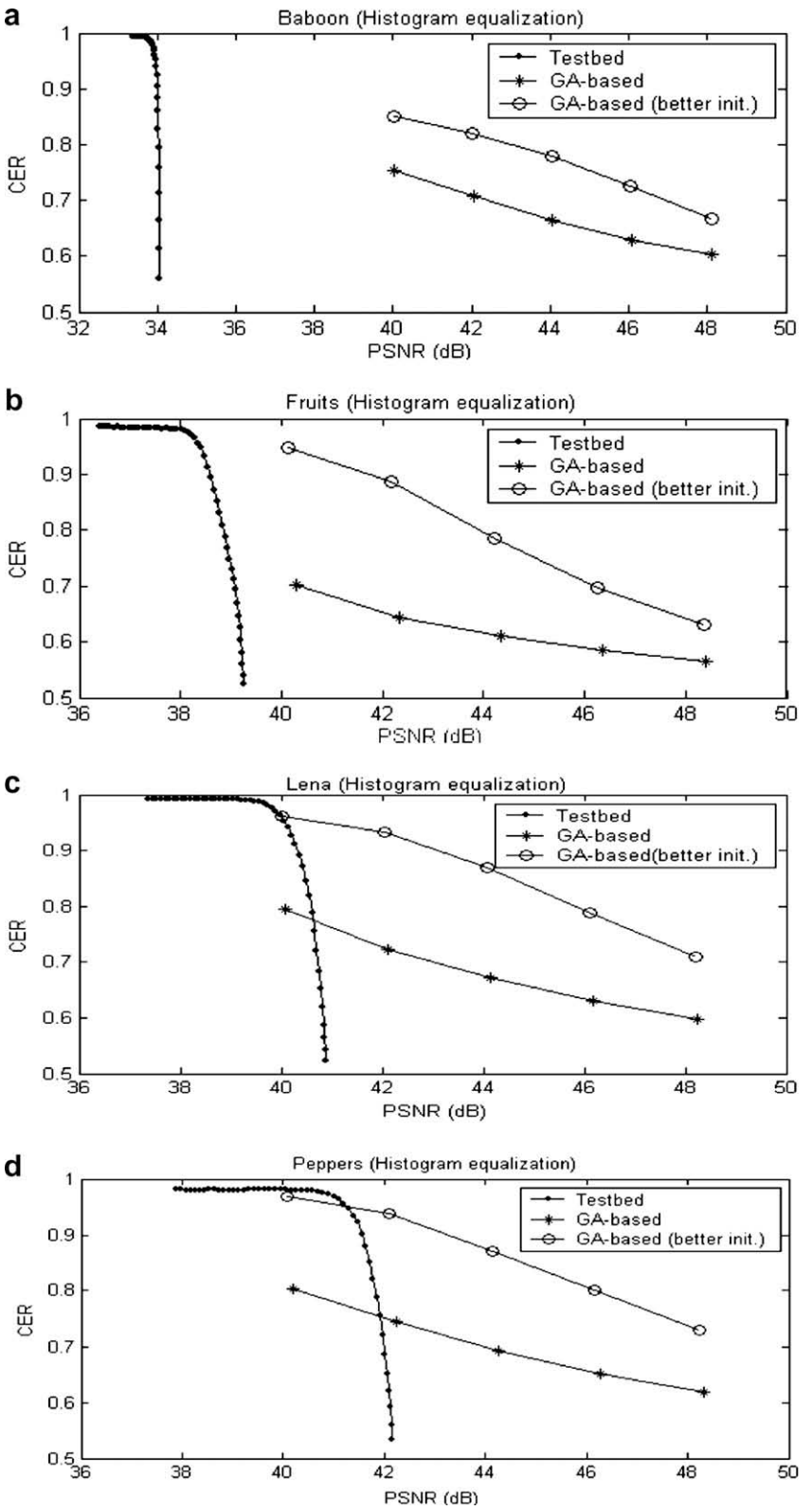

Fig. 11. Robustness performance against histogram equalization of different images.

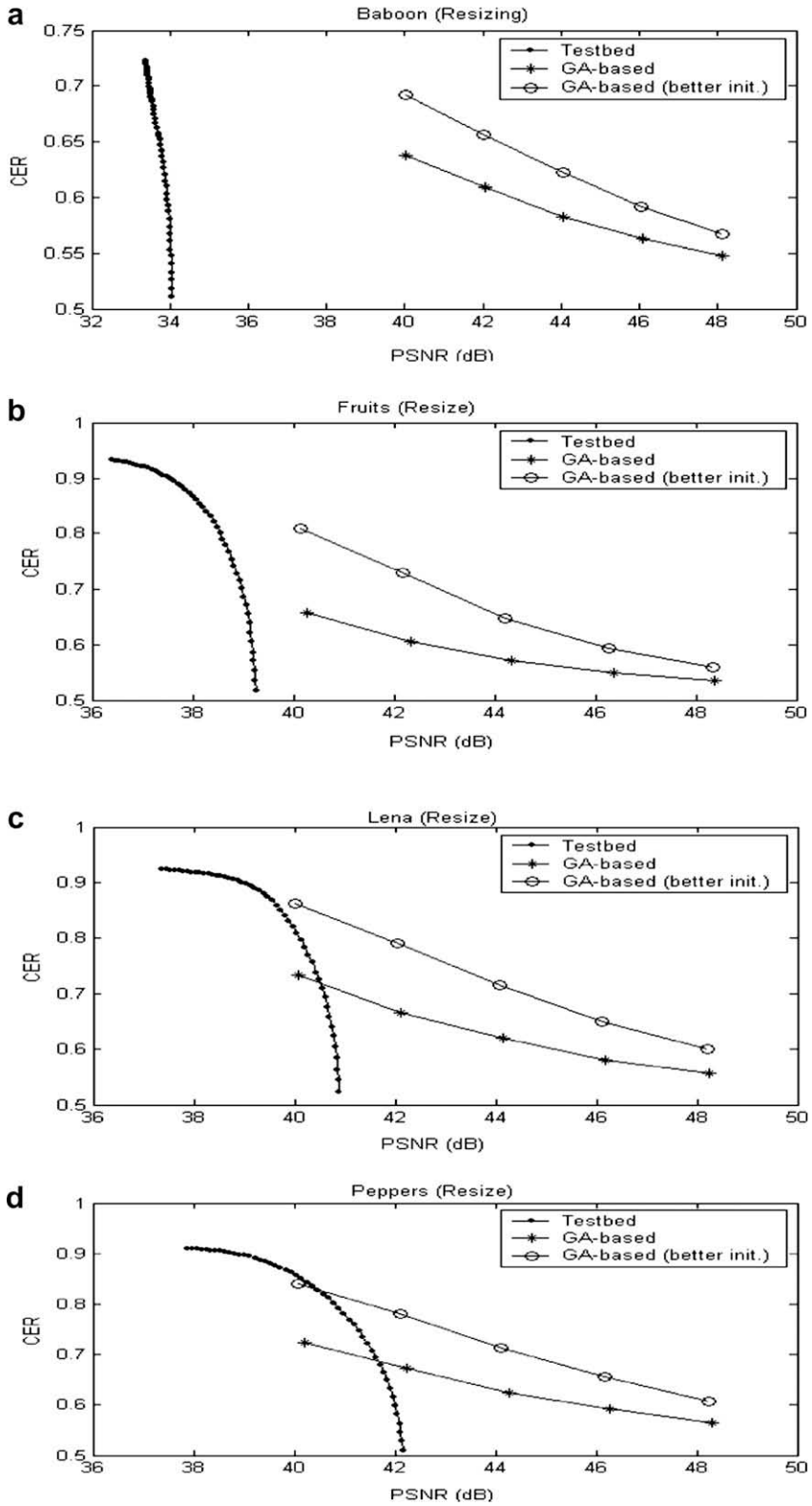

Fig. 12. Robustness performance against resizing of different images. Note that the image Baboon is more vulnerable to resizing attacks because most of this image consists of high-variance details.

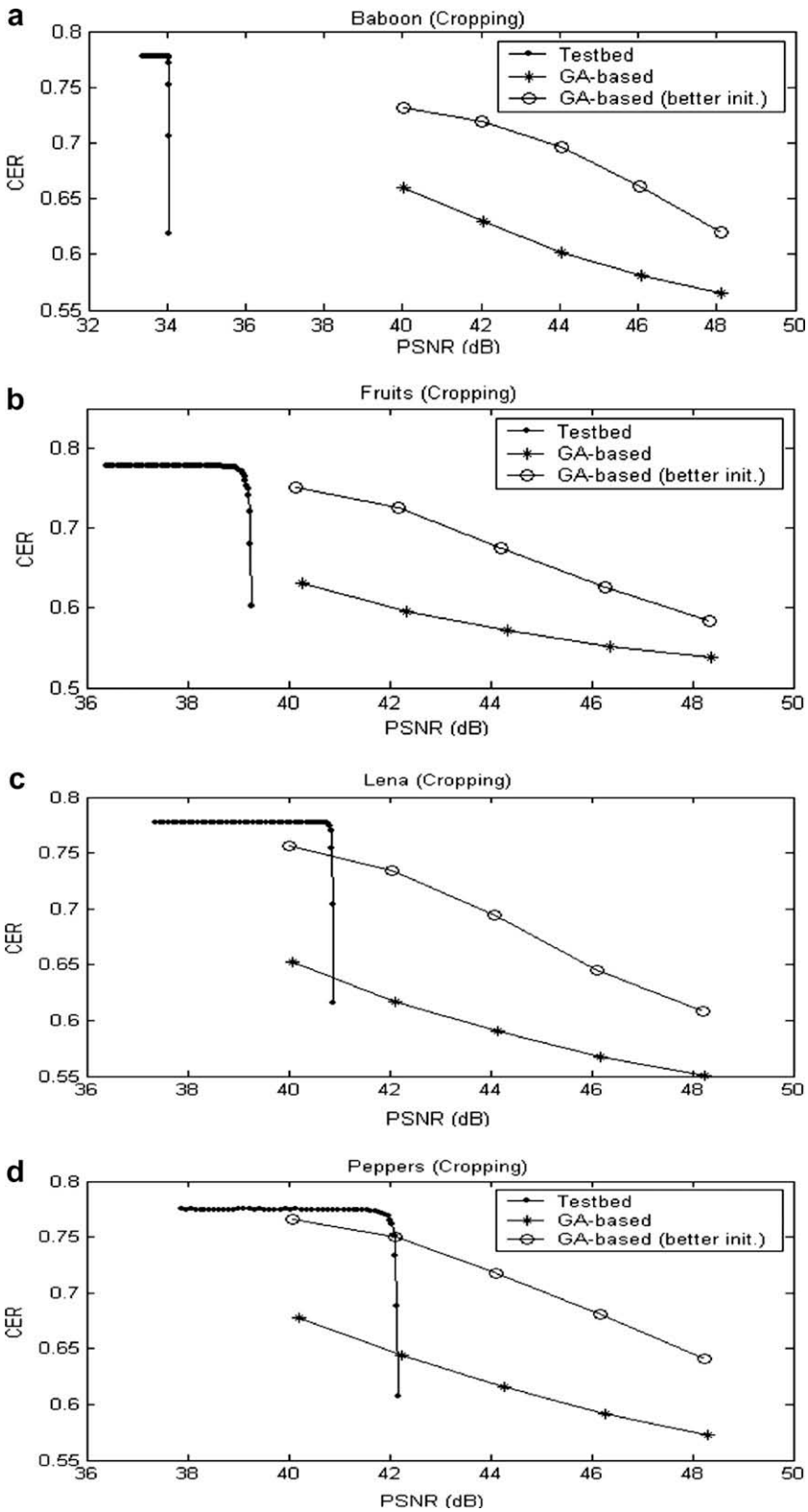

Fig. 13. Robustness performance against cropping for different images. Note that the Baboon image is more vulnerable to resizing attacks because most of the image consists of high-variance details.

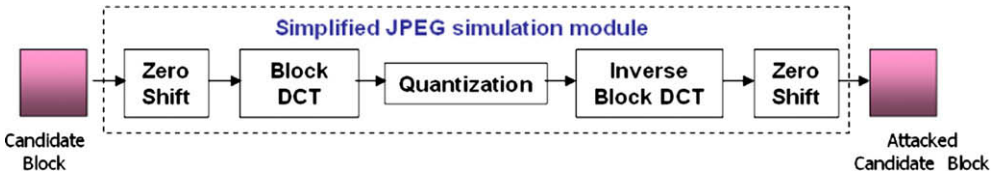

Fig. 14. Simplified JPEG compression module.

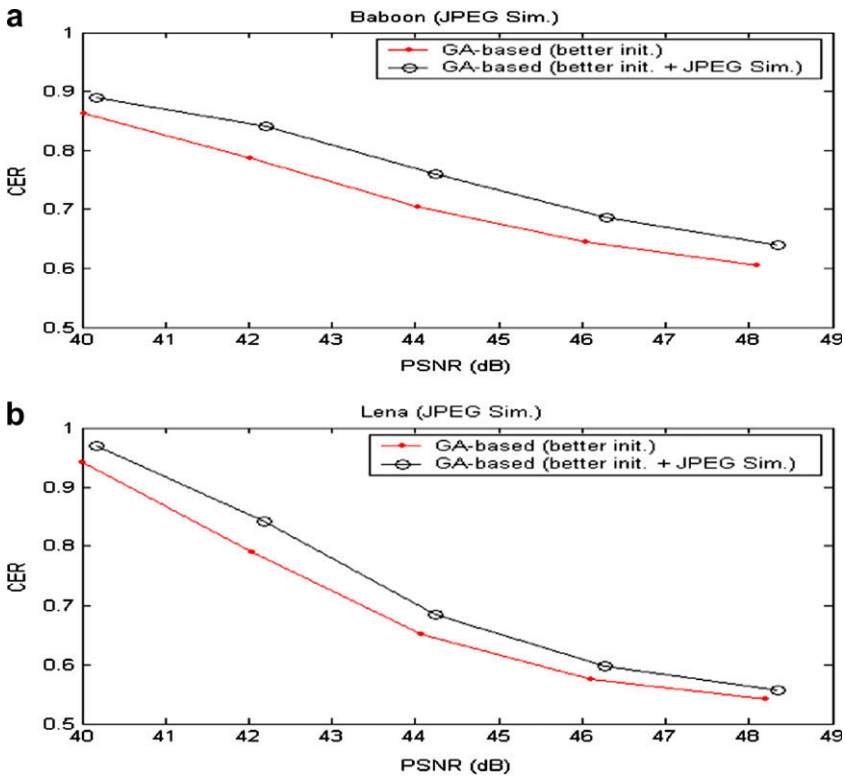

Fig. 15. Performance enhancement by incorporation of JPEG compression simulation modules into proposed scheme for (a) Baboon and (b) Lena images.

is an inevitable constraint for existing informed-embedding watermarking schemes such as the one proposed in [9], and the performance of the informed system against attacks that have not been considered may be worse than that of the original scheme. In other words, though assuming possible attacks as available auxiliary information may be useful for watermarking performance enhancement as compared to the schemes that consider attacks, their performance against other uncontrolled attacks may not improve accordingly. This problem will be further discussed in the next section

5. Discussions

5.1. Architectural advantages of proposed scheme

The architecture of the proposed watermarking performance enhancement has a lot of advantages. First, users can specify the required watermarking fidelity that must be guaranteed according to different application requirements. Next, the asymmetric embedding/detection nature not only fits the system requirement of most blind-detection watermarking schemes but also overcomes the widely criticized shortcoming of evolutionary computation—long computation time. Since the watermark detector is the same as that used in the original watermarking scheme, many common applications of watermarking will not be affected by the required computational resources in the embedding process. Furthermore, the proposed watermarking scheme has a desirable characteristic that embedding and detection can be performed in different domains, facilitating the direct control of fidelity in the spatial domain and strong robustness against attacks in the frequency domain, simultaneously. Moreover, since the obtained results are not generated via explicit embedding, the statistical behavior of the embedded coefficients will be highly random, increasing the difficulty of unintended interception. Finally, most blind-

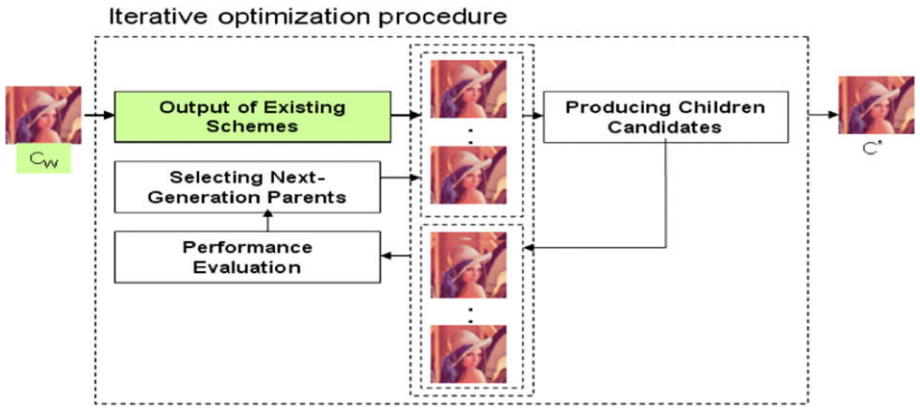

Fig. 16. Incorporation of existing watermarking schemes with GA-based performance enhancement scheme as long as their embedded outputs can be taken as the initial inputs for proposed framework.

detection watermarking schemes can be easily incorporated with the proposed performance enhancement scheme, as shown in Fig. 16.

5.2. Proposed framework as informed-embedding watermarking scheme

The proposed scheme can be regarded as an informed-embedding watermarking scheme since the embedder uses the available auxiliary information. However, because its performance against attacks that have not been considered is not enhanced, the feasibility of the informed-embedding watermarking scheme is limited since a general-purpose watermarking scheme is supposed to be robust against many types of general attacks.

However, there is a digital watermarking application in which all applicable/permissible attacks are supposed to be predicable and even controllable for the embedder; the application is steganography. For steganographic applications, only the host-interference and optional lossy compressions shall be considered. The degree of lossy compression applied to the cover content can even be controlled by the sender of secret messages. Therefore, not only this scheme but also all informed-embedding watermarking schemes can be very useful for steganographic applications.

5.3. Complexity concerns

Currently, generating an embedded image using a non-optimized MatLab implementation with 100 generations requires roughly a computation time of 1 h on a PC with a 3.0 GHz Intel Pentium 4 processor. Considering the performance difference between the MatLab interpreter and real-world compilers, the computational complexity is reasonable but can be further simplified.

5.4. Possible extensions

Simple quality metrics are adopted in this paper to guarantee minimal quality requirements. In fact, finding subjective quality metrics that fit the HVS well has always been an important task in the studies on image and video processing. With advances in visual quality metrics, the proposed scheme may be further enhanced to produce better watermarked results.

6. Conclusions

In this study, a watermark embedder is modeled as an optimization problem. A performance space based model is illustrated to provide more understanding of the watermarking performance enhancement. We have proposed a general watermarking performance enhancement framework on the basis of evolutionary computation techniques. Simulation results show that blind-detection watermarking schemes will be enhanced by incorporating the proposed GA-based framework. The watermarking performance of a scheme can be actively considered as useful information for watermark embedders.

References

[1] E. Alba, J.F. Chicano, Software project management with GAs, Information Sciences 177 (2007) 2380–2401.
[2] G. Armano, M. Marchesi, A. Murru, A hybrid genetic-neural architecture for stock indexes forecasting, Information Sciences 170 (2005) 3–33.
[3] U. Beldek, K. Leblebicioğlu, Strategy creation, decomposition and distribution in particle navigation, Information Sciences 177 (2007) 755–770.
[4] A. Briassouli, M.G. Strintzis, Optimal watermark detection under quantization in the transform domain, IEEE Transactions on Circuits and Systems for Video Technology 14 (2004) 1308–1319.

[5] B. Chen, Design and analysis of digital watermarking, information embedding, and data hiding system, Ph.D. Dissertation, MIT, Cambridge, MA, 2000.
[6] B. Chen, G.W. Wornell, Quantization index modulation: a class of provably good methods for digital watermarking and information embedding, IEEE Transactions on Information Theory 47 (2001) 1423–1443.
[7] S.H. Chen, Y.C. Huang, Relative risk aversion and wealth dynamics, Information Sciences 177 (2007) 1222–1229.
[8] A.J. Chipperfield, P.J. Fleming, C.M. Fonseca, Genetic algorithm tools for control systems engineering, in: Proceedings of the Adaptive Computing in Engineering Design and Control, Plymouth, UK, 1994, pp. 21–24.
[9] A.J. Chipperfield, P.J. Fleming, The MATLAB Genetic Algorithm Toolbox, IEE Colloquium on Applied Control Techniques Using MATLAB, Digest No. 1995/014, 1995.
[10] D.A. Coley, An Introduction to Genetic Algorithms for Scientists and Engineers, World Scientific, Singapore, 1999.
[11] I.J. Cox, J. Bloom, M.L. Miller, Digital Watermarking, Morgan Kaufman Publishers, CA, 2001.
[12] I.J. Cox, M.L. Miller, The first 50 years of electronic watermarking, Journal of Applied Signal Processing 2 (2002) 126–132.
[13] D.B. Fogel, Evolutionary Computation toward a New Philosophy of Machine Intelligence, John Wiley and Sons, Hoboken, NJ, 2006.
[14] D.E. Goldberg, Genetic Algorithm in Search, Optimization and Machine Learning, Addison-Wesley, MA, 1989.
[15] J.J. Grefenstette, Incorporating problem specific knowledge into genetic algorithms, Genetic Algorithms and Simulated Annealing, Morgan Kaufman Publishers Inc., LA, 1987.
[16] C.T. Hsu, J.L. Wu, Hidden digital watermarks in images, IEEE Transactions on Image Processing 8 (1999) 58–68.
[17] C.H. Huang, Implementation and improvement of some digital watermarking algorithms, M.S. Thesis, Department of CSIE, National Taiwan University, Taipei, Taiwan, 1999.
[18] C.H. Huang, J.L. Wu, A watermark optimization technique based on genetic algorithms, in: Proceedings of the SPIE Electronic Imaging 2000, EI'00, San Jose, CA, 2000, pp. 516–523.
[19] C.H. Huang, Performance and security improvements of digital watermarking schemes and a study of DRM frameworks, Ph.D. Dissertation, Department of CSIE, National Taiwan University, Taipei, Taiwan, 2004.
[20] C.H. Huang, J.L. Wu, Fidelity-controlled robustness enhancement of blind watermarking schemes using evolutionary computational techniques, in: Proceedings of the Third International Workshop on Digital Watermarking, IWDW'04, Seoul, Korea, 2004, pp. 271–282.
[21] C.H. Lin, J.L. Wu, C.H. Huang, An efficient genetic algorithm for small search range problems and its applications, Intelligent Multimedia Processing with Soft Computing, Springer, 2005.
[22] M.L. Miller, Watermark embedding for black-box channels, in: Proceedings of the Second International Workshop on Digital Watermarking, IWDW'03, Seoul, Korea, 2003, pp. 18–34.
[23] M.L. Miller, G.J. Doerr, I.J. Cox, Applying informed-coding and embedding to design a robust high-capacity watermark, IEEE Transactions on Image Processing 13 (2004) 792–807.
[24] P. Moulin, J.A. O'Sullivan, Information-theoretic analysis of watermarking, in: Proceedings of the 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'00, Istanbul, Turkey, 2000, pp. 3630–3633.
[25] P. Moulin, J.A. O'Sullivan, Information-theoretic analysis of information hiding, IEEE Transactions on Information Theory 49 (2003) 563–593.
[26] N. Nalini, G.R. Rao, Attacks of simple block ciphers via efficient heuristics, Information Sciences 177 (2007) 2553–2569.
[27] J.S. Pan, H.C. Huang, F.H. Wang, Genetic watermarking techniques, in: Proceedings of the Fifth International Conference on Knowledge-based Intelligent Information Engineering System and Allied Technologies, 2001.
[28] F. Shih, Y.T. Wu, Enhancement of image watermark retrieval based on genetic algorithms, Journal of Visual Communication and Image Representation 16 (2005) 115–133.
[29] F. Shih, Y.T. Wu, Robust watermarking and compression for medical images based on genetic algorithms, Information Sciences 175 (2005) 200–216.