

Automated Optimization of JPEG 2000 Encoder Options Based on Model Observer Performance for Detecting Variable Signals in X-Ray Coronary Angiograms

Yani Zhang*, *Member, IEEE*, Binh T. Pham, and Miguel P. Eckstein

Abstract—Image compression is indispensable in medical applications where inherently large volumes of digitized images are presented. JPEG 2000 has recently been proposed as a new image compression standard. The present recommendations on the choice of JPEG 2000 encoder options were based on nontask-based metrics of image quality applied to nonmedical images. We used the performance of a model observer [non-prewhitening matched filter with an eye filter (NPWE)] in a visual detection task of varying signals [signal known exactly but variable (SKEV)] in X-ray coronary angiograms to optimize JPEG 2000 encoder options through a genetic algorithm procedure. We also obtained the performance of other model observers (Hotelling, Laguerre–Gauss Hotelling, channelized-Hotelling) and human observers to evaluate the validity of the NPWE optimized JPEG 2000 encoder settings. Compared to the default JPEG 2000 encoder settings, the NPWE-optimized encoder settings improved the detection performance of humans and the other three model observers for an SKEV task. In addition, the performance also was improved for a more clinically realistic task where the signal varied from image to image but was not known *a priori* to observers [signal known statistically (SKS)]. The highest performance improvement for humans was at a high compression ratio (e.g., 30:1) which resulted in approximately a 75% improvement for both the SKEV and SKS tasks.

Index Terms—Automated optimization, detection performance, image compression, JPEG 2000, model observer, psychophysical study, SKEV/SKS task.

I. INTRODUCTION

THE INITIAL motivation to digitize medical images is to support image transfer, archiving, and image manipulations that might improve image quality including filtering techniques, feature motion stabilization, image enhancement, and volume rendering. Hospitals and the clinical environment are moving toward digitization [1] because digitally stored images are more readily displayed and more easily reviewed than their

analog counterparts. To be comparable with current analog film-based medical images, digitized images must be of high quality and high resolution and, therefore, have a very large volume in general. To represent large medical images with the smallest possible size, image compression provides immediate and substantial reduction in the cost of medical image storage and transmission. There has been considerable interest in applying lossy or irreversible compression for medical images [2]–[4].

A. The New Still-Image Compression Standard: JPEG 2000

As a new still-image compression algorithm, JPEG 2000 is being designed to address the requirement of a diversity of applications, e.g., internet, printing, medical imagery, and digital library [5]. Part one of the standard supports the following set of features: superior low-bit rate performance, lossy to lossless compression, multiple resolution representation, embedded bit-stream [progressive decoding and signal-to-noise ratio (SNR) scalability], tiling, region-of-interest coding, and flexible file format [5]. Various JPEG 2000 encoder options are provided to implement these new features. Fig. 1 shows the general block diagram of the JPEG 2000 encoder. The JPEG 2000 algorithm starts with a preprocessing stage which includes component decomposition, tiling, DC shifting, and inter-component transform to decorrelate the color data. Following this, the discrete wavelet transform (DWT) is applied. The DWT can be *irreversible* (Daubechies 9-tap/7-tap filter [6]) or *reversible* (5-tap/3-tap filter). After that, all coefficients are quantized. In the final coding stage there are two steps: namely Tier-1 coding also called arithmetic coding and Tier-2 coding also called bit-stream coding. We use *Jasper* which is a C implementation of the JPEG 2000 standard initially developed by Image Power and the University of British Columbia [7] (<http://www.ece.uvic.ca/mdadams/jasper/>).

Previous work [8] has made specific recommendations about encoder settings (tile sizes, code block sizes, and type of DWTs) to achieve higher quality in the images undergoing compression. However, these recommendations were based on the peak signal-to-noise ratio (PSNR); a monotonic transformation of the root-mean-square error (RMSE) between the original image and the image that underwent compression). The PSNR metric is not perceptual nor task-based but rather mathematically tractable and easy to measure.

An early study by Mannos and Sakrison [9] demonstrated that simple computational distortion measures, such as the mean-square error (MSE), cannot reliably predict the difference of

Manuscript received May 7, 2003; revised November 9, 2003. This work was supported by the National Institute of Health (NIH) under Grant RO1-HBL 53455. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was S. Aylward. *Asterisk indicates corresponding author.*

*Y. Zhang is with the Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106 USA (e-mail: zhang@psych.ucsb.edu).

B. T. Pham and M. P. Eckstein are with the Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106 USA (e-mail: pham@psych.ucsb.edu; eckstein@psych.ucsb.edu).

Digital Object Identifier 10.1109/TMI.2004.824153

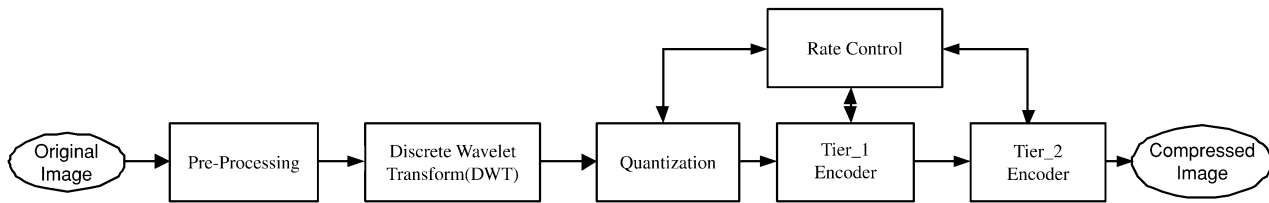


Fig. 1. The block diagram of JPEG 2000 encoder.

the perceived quality of one image with another. More recently Eckstein *et al.* [10] have shown that the MSE can lead to erroneous evaluation of task-based medical image quality. Furthermore, when the images are for medical use, image quality can be defined objectively in terms of diagnostic task performance [11]. If the diagnostic procedure involves physicians visually scrutinizing images, then image quality can be defined in terms of human visual performance in the clinically relevant tasks. In this context, one might evaluate image quality by conducting psychophysical studies assessing human visual performance for the display or image modality conditions of interest. However, such studies become impractical and/or nonfeasible when the study parameter set is large since the number of conditions to be evaluated increases geometrically as a function of the number of parameters [11]. One potential solution is to employ a computer model observer that reliably reflects human observer performance. Image quality can then be measured in terms of model observer performance.

B. Model Observer Performance on Signal Detection Task in Medical Images

Task-based model observers, also called numerical observers or computer observers, involve mathematical algorithms processing the input image/images and making a decision about the presence/absence, location and/or class of a signal. The ideal observer [12] uses prior knowledge about the signal and statistical properties of the noise to make decisions optimally. With clinical images, since the statistical properties of the images are not known *a priori*, the calculation of ideal observer is too difficult [13]. Another set of suboptimal model observers have been proposed to specifically predict human visual detectability of signals embedded in noisy backgrounds and have been proposed as a standard for evaluation of medical image quality [14]–[23]. In this study, we focus on this latter class of model observers. These have been used to predict human performance with computer generated signals embedded and noise [14]–[23]. In addition, they have been used with hybrid images that combine simulated signals and real patient backgrounds [24] to do automated evaluation and optimization on JPEG lossy image compression algorithms [10], [25], [26]. In this paper, we use performance of model observers to optimize the encoder options of the newer JPEG 2000 compression scheme on a signal detection task in X-ray coronary angiograms. In addition, previous studies used a task in which the signal does not vary from trial to trial and is known to the observer [signal known exactly (SKE) task] [10], [11], [14], [25]. The SKE task is simple but not representative of clinical scenarios in which signals vary in size and shape across images and the

physician does not know *a priori* the signal to appear. In the present study we attempt to compensate for the lack of signal variability in an SKE task by using a signal known exactly but variable (SKEV) task in which the signal varies in shape and size across trials but is known to the observer. An even more clinically realistic task is the signal known statistically task (SKS), in which the signal varies across trials and the observer is not informed about the particular signal presented on each trial. However, the SKS task is computationally demanding precluding extensive investigation of the parameter space in a reasonable time frame. For this reason, we use model observer performance on an SKEV task to optimize JPEG 2000 encoder options. A subsequent psychophysical study is used to test whether the JPEG 2000 encoder settings optimized with respect to model observer performance lead to improved human performance. Furthermore, we investigate whether optimizing performance in the computationally tractable SKEV task leads to improved model and human performance in the computationally complex yet more clinically realistic SKS task.

Our recent study [27] obtained similar SKEV task performance as a function of JPEG 2000 encoder settings for several kinds of model observers including the nonprewhitening matched filter observer with an eye filter (NPWE), Hotelling observer (HO), channelized-Hotelling observer (CHO), and Laguerre-Gauss Hotelling observer (LGHO). Performance of all model observers resulted in good prediction of human performance as a function of JPEG 2000 encoder options for a limited number of conditions. In this paper, we choose the NPWE model observer for the automated optimization to take advantage of the computation simplicity and its computing time economy. We will test whether the NPWE performance optimized JPEG 2000 encoder settings lead to improved performance for the other model and human observers. In addition, we will assess whether the optimal settings are the same for different compression ratios.

C. Automated Optimization and Genetic Algorithms (GAs)

The performance of a model observer can reliably predict human performance degradation in the compressed images [10], [28], but it is very impractical and time consuming to compute the performance of a model observer for each possible combination of JPEG 2000 encoder options. For example, in our test set we have a total of 541 angiograms of size 512×512 pixels. One compression/decompression cycle takes around 5–7 min on a Pentium 4, 1.8-GHz computer. This results in several months of computer time to get a good encoder setting if we use the exhaustive search method going through more than 1 000 000

encoding setting possibilities¹. A GA is, thus, implemented to optimize the encoder settings.

The GAs first investigated by Holland [29] use properties of the evolution of natural systems and produce the optimal or near-optimal solutions in the complex search space for the optimization problems in diverse fields. GAs are attractive since they can provide an efficient, parallel exploration of the large, complex search space without getting stuck in a local optimum, while only minimum knowledge is available about the objective function [30]–[34]. It has been theoretically proven that GAs provide robust search even if the search space is not continuous [33]. GAs have been used in a number of different medical image processing areas such as medical diagnosis, medical image registration, the left ventricle of the heart detection and edge detection [35]–[39]. The GA being used in this study is an iterative procedure (Fig. 2) that searches for better solutions by using a collection of JPEG 2000 encoder settings as a population to maximize the NPWE model observer performance on an SKEV task.

II. METHODS

A. Generation of Test Images

The method for generating the test images is guided by two main considerations. The first is to include factors affecting real clinical images. The second consideration is to have a generating process for which ensemble (population) statistical properties can be estimated with reasonable effort.

The clinical digital coronary angiograms were acquired at 30 frames/s with a 7-in image intensifier field size (Advantx/DXC, General Electric Medical Systems). The parameters used in the image acquisition were a standard exposure control at $0.3 \mu\text{Gy}$ per frame and extended dynamic range enabled video circuitry. The images were digitized with a linear analog amplification and lookup table to achieve a 512×512 pixel matrix with a resolution of 0.3 mm/pixel and 256 gray levels. Totally 541 images which were extracted from 50 different image sequences of 17 different patients were used as the backgrounds. These patients were a random sample from a large pool getting an X-ray coronary angiogram at the Catherization Laboratory with no regard to the patients' state of disease.

We have previously used a mathematical method to generate realistic looking computer simulated arteries. The simulated artery model is discussed in detail and validated in a previous publication [40]. Here, we outline the general principles of the method. For our test images the projected simulated arteries are 3-D right circular cylinders with a diameter of 12 pixels (3.6 mm). The arteries also include a sinusoidally modulated narrowing in diameter toward the center (minimum diameter of 8 pixels and a length of 50 pixels). The attenuation coefficient μ was set to $0.16/\text{mm}$ to produce simulated arteries with the same projected intensity as real angiogram coronary arteries of the same diameter. The signals were projected ellipsoids with the vertical axis ranging from 3 to 25 pixels and the horizontal

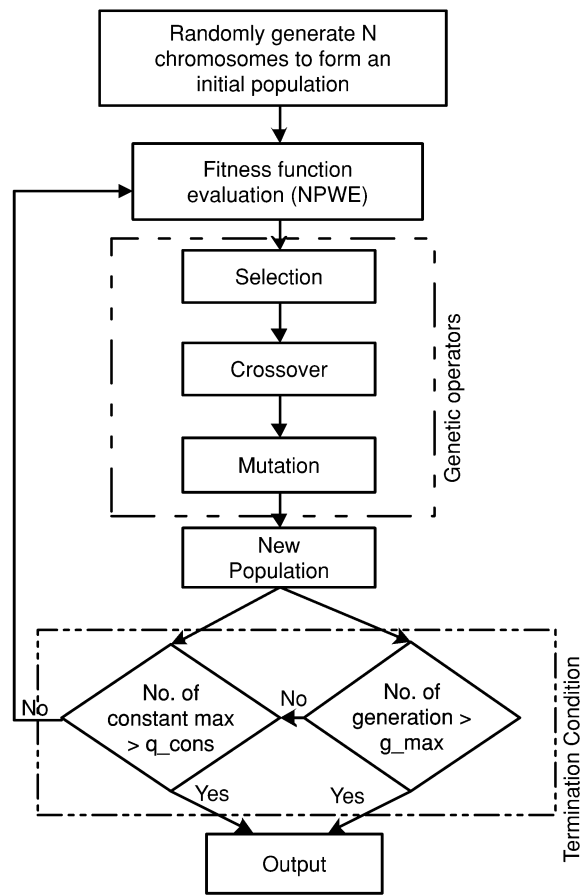


Fig. 2. The flow chart of GA.

axis ranging from 3 to 10 pixels. This resulted in a total of 184 possible signals.

We generated four simulated arteries and projected them at a distance of 32 pixels apart into the backgrounds. To embed a simulated object into a real angiogram, we assumed that the original image can be decomposed into a primary image due to the primary beam and a secondary image due to scattered rays. We first subtracted an estimate of the secondary from the original image, then attenuated the remaining estimated primary corresponding to the simulated artery, and then added back the secondary image [40]. Since more than one group of simulated arteries may be inserted into one background, we obtained a total of 900 test images of size 256×256 pixels. These test images are centered around a group of four simulated arteries with a signal inserted in one of the four arteries. The average signal contrast calculated from all of the 900 test images was 0.17, where the signal contrast is defined as: $(\text{peak signal luminance} - \text{background luminance}) / \text{background luminance}$.

B. Task

For each test image, the model and human observers' task is to identify the artery containing the signal out of the four arteries [four alternative forced choice (4 AFC)].

C. NPWE Model Observer Methodology

The general form for the linear model observer is the inner (or dot) product between a template and the data at each of the

¹The $512(\text{tile size}) \times 2(\text{mode}) \times 8(\text{number of resolution levels}) \times 10(\text{code block size}) \times 7(\text{precinct size}) \times 8[\text{miscellaneous (lazy, vcausal, terminal, pterm, resetprob, sop, eph, and segsym)}] = 4587520$.

possible signal locations. Model observers in an SKEV task can be expressed in the matrix formulation [11] as

$$\lambda_{i,j} = \sum_{n=1}^{N^2} \omega_{n,j} g_{n,i} = \omega_j^t g_i \quad (1)$$

where $\lambda_{i,j}$ is the scalar response of the model at the i th location to the j th template ω_j , an $N \times N$ 2-dimensional template represented by an $N^2 \times 1$ column vector. The subscript j refers to the j th template corresponding to the j th signal presented on that trial or image. The superscript t refers to transpose. The vector g_i is the data at the i th location represented in a column vector format.

The NPWE model uses information about the signal but takes into account the human visual sensitivity to different spatial frequencies. The template corresponding to the j th signal is obtained by filtering templates matching the signals with the square of the contrast sensitivity function (CSF). This can be achieved by multiplying in the Fourier domain the j th signal and the eye filter [41]–[43]

$$\tilde{\omega}_j = \tilde{s}_j |\tilde{E}|^2 \quad (2)$$

where \tilde{E} is the eye filter in the Fourier domain and \tilde{s}_j is the signal in the Fourier domain. The spatial domain template ω_j can be obtained by using an inverse Fourier transform of $\tilde{\omega}_j$.

The eye filter used for the NPWE model is given by

$$E_\rho = \rho^\eta \exp(-c\rho^\gamma) \quad (3)$$

where $\rho = \sqrt{u^2 + v^2}$ is the radial spatial frequency in cycles per degree. In this paper we used an eye filter with parameters close to those used by Burgess [43]. We selected eye filter parameters to be $c = 0.013$, $\eta = 1.4$, and $\gamma = 2.6$ with the peak at 4 cycles/degree.

D. Hotelling, Channelized-Hotelling and Laguerre-Gauss Hotelling Models for SKEV Tasks

Although the NPWE model observer is used to optimize the JPEG 2000 encoder settings, other model observers commonly used in the medical imaging literature are also evaluated on their performance for the default and NPWE-optimized encoder settings. The HO involves a prewhitening operation followed by a matched filter. Unlike the NPWE model observer, the HO derives a template that takes into account knowledge about not only the signal profile but also the background statistics [17], [44]. The LGHO proposed by Barrett *et al.* [44] uses a set of linear preprocessing filters (Laguerre-Gauss functions) that capture the main features of the Hotelling template. Thus, the Laguerre-Gauss Hotelling model observer does not act on the pixels but rather on the set of scalar values resulting from the dot product of the Laguerre-Gauss function and the data. The CHO was introduced by Myers and Barrett [45] as a way to incorporate elements from human vision system into models of signal detection for medical images. Here, we use a channelized-Hotelling with orientation and spatial frequency tuned Gabor channels which reflect the properties of the cells in the primary visual cortex [46]. Details about the theory and the application of these model observers to the current task are discussed in Appendix I.

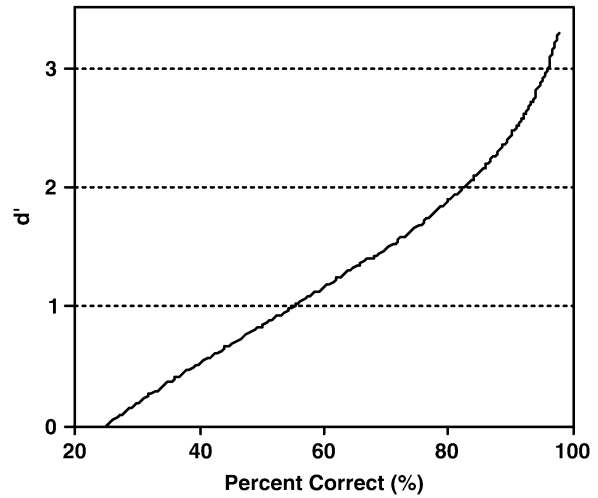


Fig. 3. d' versus percent correct (%) for a 4AFC task.

E. Model Observer Performance for an SKEV Task

A general method to measure performance in a multiple alternative forced choice (MAFC) task is to compute the probability of the correct outcome in a trial [47], [48]. A correct outcome occurs when the response of the j th template to the signal location exceeds the maximum response to the noise only location. This can be mathematically expressed as

$$P_c = \frac{1}{H} \sum_{h=1}^H \text{step}(\lambda_{s,j,h} - \max(\lambda_{n,j,h})) \quad (4)$$

where H is the total number of trials, $\lambda_{s,j,h}$ is the template response in trial h to the signal present location, and $\lambda_{n,j,h}$ is the template response in trial h to the noise locations. The step function is 1 where the argument is larger than 0 and is 0 when the argument is less than 0. The percent correct (P_c) can be then converted to an empirically obtained index of detectability (d') by generating a look up table for P_c versus the index of detectability from the following relationship [49]:

$$P_c(d', M) = \int_{-\infty}^{+\infty} \phi(x - d') [\Phi(x)]^{M-1} dx \quad (5)$$

where $\phi(x) = (1/\sqrt{2\pi})e^{-x^2/2}$, $\Phi(x)$ is the cumulative Gaussian distribution function, $\Phi(x) = \int_{-\infty}^x \phi(y) dy$, M is the number of possible signal location in the experiment, and d' is the index of detectability.

The relationship between the detectability (d') and the percent correct (P_c) [49] for a 4AFC [$M = 4$ in (5)] task is shown in Fig. 3. The 25% P_c on the 4AFC task equals a detectability of 0 because then the observer is essentially guessing. As P_c approaches 100%, d' asymptotically approaches $+\infty$.

F. GA Optimization Procedure

One of the important technical terms in GAs is the chromosome which represents possible solutions to the given problem. Each chromosome is assigned a fitness value based on the quality of the solution to the problem. The GA starts with an initial set of N randomly-generated chromosomes which is called a population as shown in Fig. 2, where N is the population size. After evaluating the fitness values of the

TABLE I
 GENES AND THEIR VALUE SPACES

genes	encoder options	range
gene 1	tile size	32, 64, [88-92], [108-114], 128, [148-152], [216-240], 256, [436-484], 512 (default)
gene 2	resolution level	2, 3, 4, 5, 6 (default), 7, 8
gene 3	mode	int (default), real
gene 4	code block size	32 × 64, 32 × 32, 64 × 32, 64 × 64 (default), 256 × 16, 16 × 256, 16 × 128, 128 × 16, 128 × 32, 32 × 128, 128 × 8, 8 × 128, 256 × 8, 8 × 256,
gene 5	precinct size	256 × 256, 128 × 128, 512 × 256, 64 × 64, 32 × 32, 256 × 512, default.

chromosomes, genetic operators are carried out to generate new populations and the evolution continues until a predetermined stopping criterion is satisfied. The GA converges to the best chromosome, which represents the optimal or suboptimal solution to the studied problem. A GA is composed of the following four basic components [29]: a fitness function, a representation scheme, genetic operators (selection, crossover, and mutation), and control parameters. Among them, the fitness function and the representation scheme form the bridge between the original problem context and the problem-solving framework [50].

1) *Fitness Function*: The fitness function describes how well a chromosome represents the solution to the problem. Here, the NPWE performance, P_c for an SKEV task, is used as the fitness function. The performance is obtained for the 900 test angiograms with simulated arteries and signals inserted.

2) *Representation of Solution*: We use a binary base-2 coding to represent each chromosome as a string of elements of the set $\{0; 1\}$ (bits). The effect of binary coding is to increase the number of search space dimensions and decrease the probability of the convergence to a local optimum. Based on our previous evaluation on the JPEG 2000 encoder options [27], we choose the following encoder options as genes of the chromosome to reduce the search space.

- The encoding options $\text{tilewidth} = w$ and $\text{tileheight} = h$ set the nominal tile width and height to w and h , respectively. Here, we set $\text{tilewidth} = \text{tileheight}$ to reduce the possible search range.
- Encoding option mode determines which multicomponent and wavelet transforms (if any) are employed. In *int* mode, all transforms employed are integer to integer (e.g., RCT, 5/3 DWT). In *real* mode, real to real transforms are employed (e.g., ICT, 9/7 DWT).
- Encoder option numrlvs corresponds to the resolution level which must be an integer greater than or equal to one.
- $\text{cblkwidth} = w$ and $\text{cblkheight} = h$ set the codeblock size.
- $\text{prcwidth} = w$ and $\text{prcheight} = h$ set the precinct size.

The parameter space is shown in Table I.

An initial population with a set of chromosomes is generated randomly [34]: each gene is given a random value chosen from the parameter space. Fig. 4 is an example of a binary represented chromosome and the corresponding JPEG 2000 encoder setting.

Chromosomes subsequently evolve through the genetic operators until the stopping criterion is satisfied. These operators are illustrated below.

3) *Genetic Operators: Selection, Crossover, and Mutation*: The chromosome with the highest fitness value is

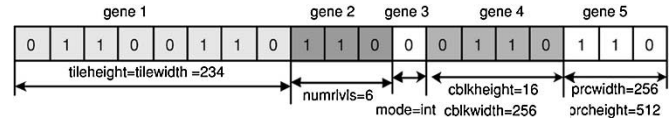


Fig. 4. An example of a chromosome and the corresponding JPEG 2000 encoder setting.

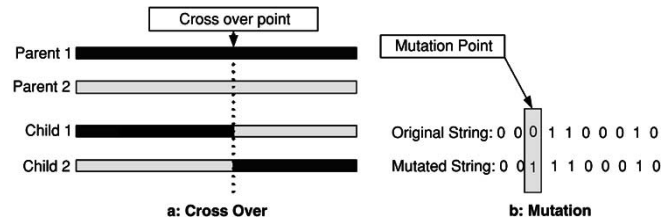


Fig. 5. Crossover (left) and Mutation (right).

selected as an elitist [33]. The elitist strategy is a mechanism which can be employed in a GA to ensure that the most highly fit chromosomes of the population are passed on to the next generation without being altered by genetic operators. This operation is used to guarantee the nature’s survival of the fittest mechanism. Then the chromosomes are selected by a “roulette wheel” [51] selection scheme where each current chromosome in the population has a roulette wheel slot sized in proportion to its fitness value. The probability that a chromosome C_i is selected as a member of the next generation is

$$P_i = \frac{F(C_i)}{\sum_{j=1}^N F(C_j)} \quad (6)$$

where $F(\cdot)$ denotes the fitness function, which is the NPWE performance for an SKEV task in our case. To reproduce, we simply spin the weighted roulette wheel $N - 1$ times [33] to select the chromosomes that will compose the mating pool for the subsequent crossover and mutation operations.

The widely used one-point crossover operator [29], [31], [33] is chosen. In this operation $(N - 1)/2$ parent pairs of strings are formed randomly from the mating pool. A single crossover point is positioned randomly between two consecutive bits, and all the bits coming after that point are exchanged between the two strings involved in the crossover process, yielding two new strings (children), as shown in Fig. 5(a). The parent chromosomes are then replaced by the two offsprings. Crossover is not applied to all pairs of chromosomes selected for mating, but rather controlled by the crossover rate which is defined as the

ratio of the number of offspring produced by crossover in each generation to the population size.

A mutation operator involves modification of the values of the genes in a chromosome and increases the variability of a population. At every gene position within the chromosome string, a random number in $\{0, 1\}$ is generated and checked in the mutation process. If the random number is smaller than the mutation rate then the bit value of the gene is altered (e.g., 0 is replaced by 1 and vice versa) as shown in Fig. 5(b), otherwise, it is kept unchanged.

After evolving through these operators, N chromosomes form the new population of the next generation.

4) *Control Parameters*: Based on the results of several preliminary runs we set the control parameters as follows: population size: $N = 41$; crossover rate: 0.8; mutation rate: 0.04. These values are appropriate both in preventing premature convergence and in reaching convergent results within reasonable time.

5) *Termination*: The first termination condition is the maximum generation time G_{\max} . We set G_{\max} to 60, thus, the maximum GA computation time for one compression ratio is about 8 days. The variable q_{cons} counts the generations where the same best chromosomes resulted and is set as the second condition to terminate the generation. If q_{cons} is larger than a predefined threshold, for example 15, the algorithm stops. After the algorithm stops, the chromosome with the maximum fitness represents the optimized solution. The final result is a string representing a JPEG 2000 encoder setting.

The software prototype implementing the above GA procedure is described in Appendix II.

6) *Comparison of Computation Time*: The total number of encoder settings evaluated by the algorithm proposed in this paper is very straightforward: $N \times G_{\max} = 2460$. For the same parameter space as shown in Table I, an exhaustive algorithm which computes all the possible combinations of encoder options will evaluate 95 (tile sizes) $\times 2$ (mode selection) $\times 15$ (codeblock sizes) $\times 6$ (precinct size) $\times 7$ (resolution levels) = 119 700 JPEG 2000 encoder settings. One cycle of compression/decompression of our test images for one compression ratio takes 5–7 min. Thus, it would take around 300 days to exhaust all of the possible encoder options. The GA based algorithm's computation efficiency is much better than the exhaustive method: 8 days for the GA versus 300 days for the exhaustive method.

G. Optimization Procedure

We ran the GA optimization programs (using the NPWE performance in the SKEV task as the figure of merit) for the five compression ratios: 7:1, 10:1, 15:1, 20:1, and 30:1 to obtain the optimal encoder settings for each compression ratio.

H. Psychophysical Experiments

After obtaining the NPWE-optimized JPEG 2000 encoder settings for the SKEV task, we conducted a human psychophysical study to determine whether the NPWE-optimized encoder settings led to improved human performance in both the SKEV and SKS task.

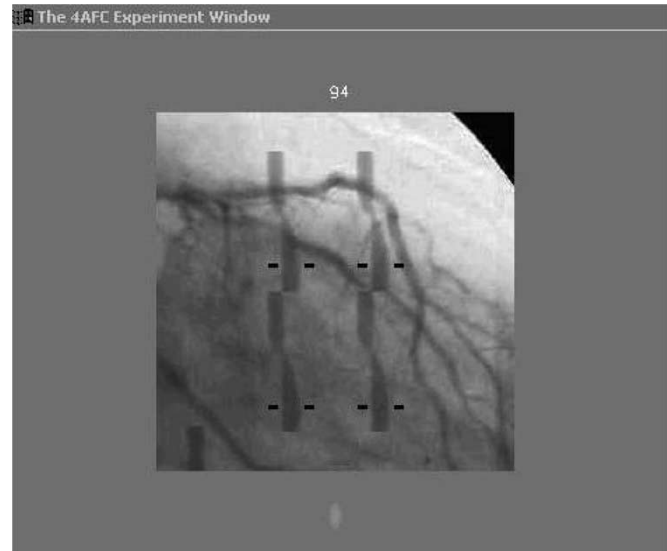


Fig. 6. GUI for an SKEV task (4AFC): Black bars in the image indicate the locations where a signal might present. Below the test image shows the signal.

1) *Image Display*: The images were displayed on an image system M17LMAX monochrome monitor with a maximum resolution of 1664×1280 pixels (Image Systems, Minnetonka, MN). The luminance versus gray level relationship was set according to the DICOM standard [52]. Before each experiment, the monitor was calibrated. The luminance range was from 0.00 cd/m^2 to 120.00 cd/m^2 .

2) *Observer Tasks*: A four-alternative forced choice (4-AFC) paradigm was used where a filling defect (signal) was placed in one of four locations, and the human observer chose the position where she or he thought the signal appeared. Experiments were conducted in a darkened room with 0.3 mm/pixel and a viewing distance of 40 cm. Signals with the vertical axis ranging from 3 to 25 pixels and the horizontal axis ranging from 3 to 10 pixels correspond to 0.128 to 1.067 degrees of visual angle vertically and 0.128 to 0.422 degrees of visual angle horizontally. On each trial an image was randomly sampled from the 900 test image database and displayed. No time limitation was imposed on the observer to make a decision. Four observers (initials: KF, LF, BP, and YZ) participated in the experiment. Observers were instructed to give accurate responses and were given feedback about their decisions (correct/incorrect). An experimental session consisted of 100 trials and lasted approximately 10 min. Each observer had at least 9 sessions per compression condition resulting in a total of 900 trials per experimental condition. The total number of conditions was 10 including both the default and the optimized conditions at the compression ratios of 7:1, 10:1, 15:1, 20:1, and 30:1. All observers participated in a training session with a highly detectable signal (1 session of 100 trials per condition resulting in a total of 1000 training trials).

Human observers participated in two different tasks: the SKEV and SKS tasks. In the SKEV task, on each trial a copy of the signal was shown at the bottom of the test image to inform the observer about the size and shape of the signal. In the SKS task, there was no information about the signal presented except that observers knew that the presented signal would be one from the pool of 184 signals. Fig. 6 shows the

TABLE II
OPTIMIZED ENCODER SETTINGS FOR FIVE DIFFERENT COMPRESSION RATIOS

compression ratios	tile size	resolution level	mode	code block size	precinct size
7:1	447	6	real	8×128	512×512
10:1	469	2	int	64×32	128×128
15:1	448	8	real	64×32	512×512
20:1	448	8	real	64×32	512×512
30:1	447	6	real	64×32	256×256

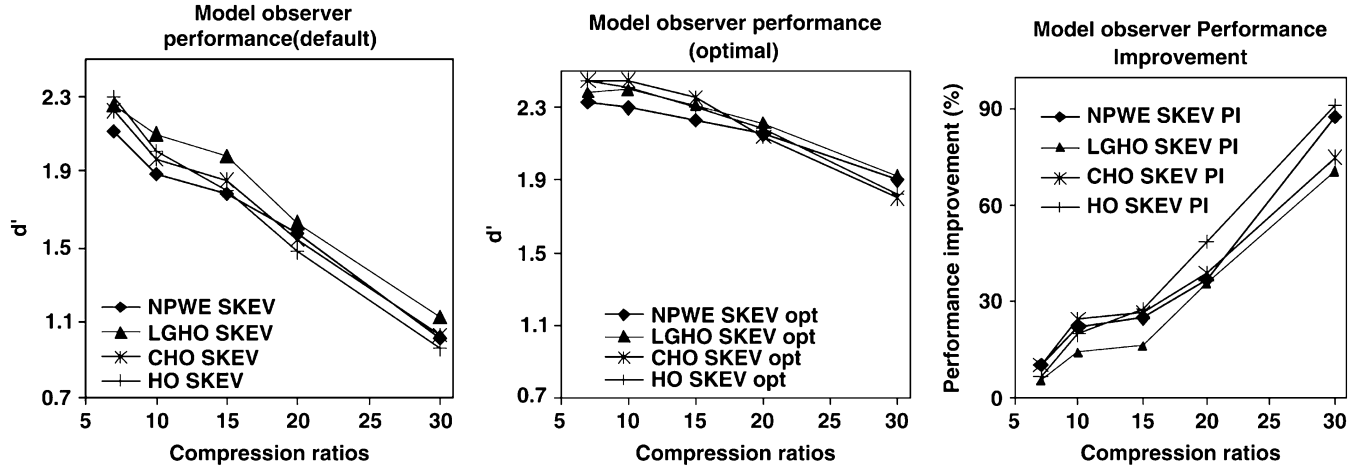


Fig. 7. Model observer performance for an SKEV task at five compression ratios 7:1, 10:1, 15:1, 20:1, and 30:1. Left: Default JPEG 2000 encoder settings; Middle: Optimal JPEG 2000 encoder settings; Right: Performance improvement.

graphic user interface (GUI) configuration developed for this test in an SKEV task.

3) *Human Performance Measurement*: Human performance (P_c) was measured by calculating the proportion of trials in which the observer correctly localized the signal. An index of detectability (d') was then calculated from P_c using (5).

I. Figure of Merit for Optimization of Compression

In order to show numerically the performance improvement from the default encoder setting to the optimal one, we defined the performance improvement (PI) percentage as follows:

$$PI(\%) = \frac{d'_{\text{optimal}} - d'_{\text{default}}}{d'_{\text{default}}} \times 100 \quad (7)$$

where d'_{optimal} denotes the detectability obtained using the NPWE-optimized encoder settings and d'_{default} denotes detectability using the default ones.

III. RESULTS

The optimized JPEG 2000 encoder settings for each of the five specific compression ratios are shown in Table II.

These encoder settings led to the highest performance of the NPWE for an SKEV task through a GA procedure. We will evaluate the NPWE-optimized encoder settings using several other model observers such as the HO, CHO, and LGHO and human observers on both the SKEV and SKS tasks.

A. Model Observer and Human Observer Performance

1) *Default Versus NPWE-Optimized Encoder Settings in an SKEV Task*: Fig. 7 compares the performance of the NPWE with that of HO, CHO, and LGHO for the SKEV task. The left graph represents the performance of model observers on images compressed with the default JPEG 2000 encoder settings. The middle graph shows the performance of model observers on images compressed with the NPWE-optimized encoder settings. The right graph shows the performance improvement as defined in (7) from the default to the NPWE-optimized encoder settings. As expected the higher the compression ratio, the higher the performance improvement. The model observer performance improvement reaches a maximum of approximately 90% (in d' units) at a compression ratio of 30:1 and a minimum of around 10% at lower compression ratios (7:1 and 10:1). Furthermore, our results indicate that all four investigated model observers show performance improvement for the five evaluated compression ratios. Interestingly, all four model observers result in similar performance improvement trends.

Fig. 8 shows the average human observer (AVE) performance (the left graph) and individual performance of the four observers for the default encoder settings and the NPWE-optimized encoder settings for an SKEV task. Error bars represent the standard error of the mean. All observers present higher performance for the NPWE-optimized encoder settings than that for the default ones.

Fig. 9 (left graph) compares NPWE model observer performance and the average human observer performance for the SKEV tasks. Although humans perform consistently lower than the NPWE model observer, the performance follows similar

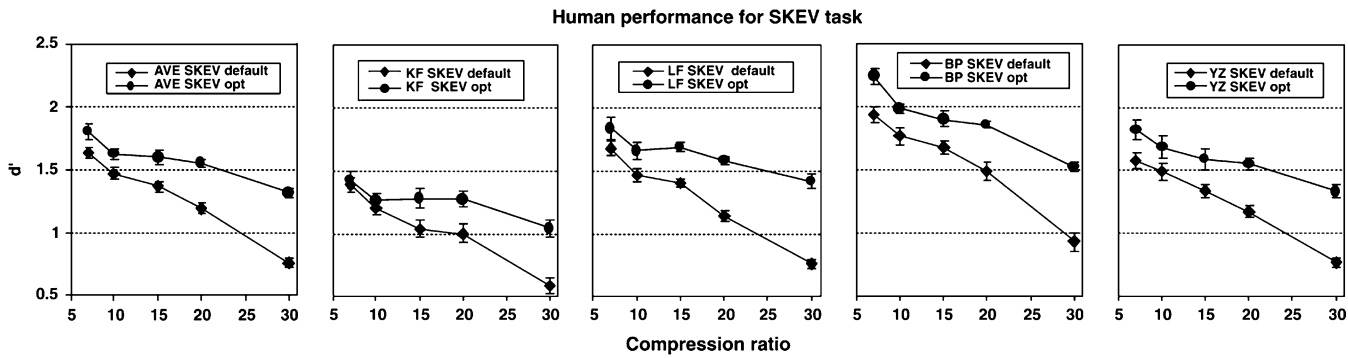


Fig. 8. Human observer performance for NPWE-optimized and default JPEG 2000 encoder settings for SKEV tasks (from left to right): AVE and individual observers (KF, LF, BP, and YZ).

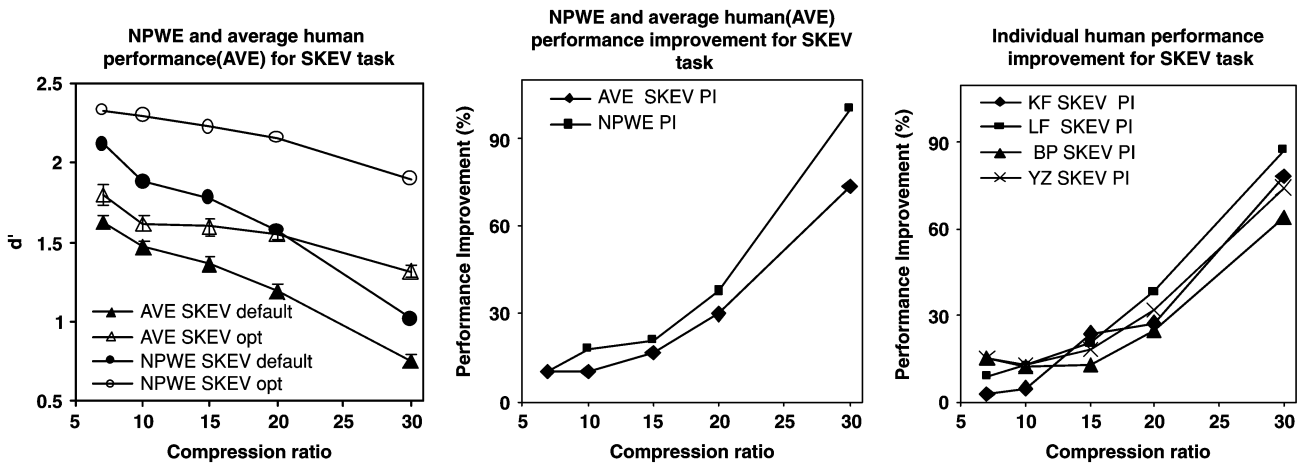


Fig. 9. NPWE and human observer performance for SKEV tasks using default and NPWE-optimized encoder settings. Left: NPWE versus AVE performance; Middle: NPWE versus AVE performance improvement; Right: Individual human observer (KF, LF, BP and YZ) performance improvement.

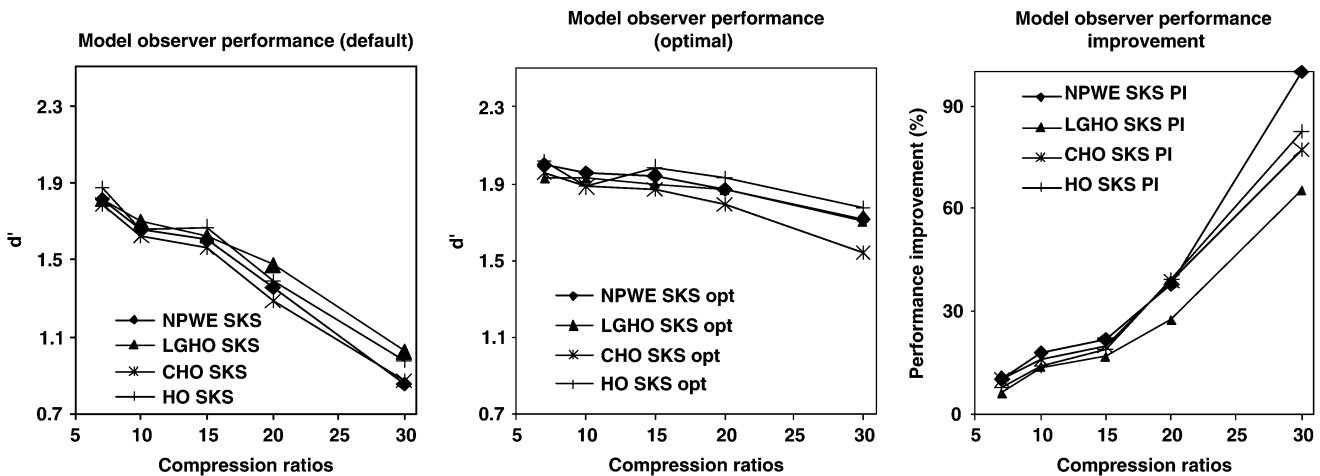


Fig. 10. Model observer performance for an SKS task at five compression ratios 7:1, 10:1, 15:1, 20:1, and 30:1 (from left to right): Default JPEG 2000 encoder setting, optimal JPEG 2000 encoder setting, performance improvement.

trends. The middle and the right graphs of Fig. 9 compare the performance improvement from the default encoder settings to the NPWE-optimized settings. Both average and individual human observers show detection performance improvement comparable to the NPWE model observer. Note that at a compression ratio of 30:1, the average human observer performance improvement is around 73% and the highest individual performance improvement is around 90% (LF). Even at the

lower compression ratios (7:1 and 10:1) the average human observer performance improvement can reach 10%.

2) *Default Versus NPWE-Optimized Encoder Settings in an SKS Task:* The results in the previous section show how an optimization based on NPWE performance in an SKEV task leads to improved performance for both human and other model observers in an SKEV task. One may ask whether similar performance improvement would be obtained for the

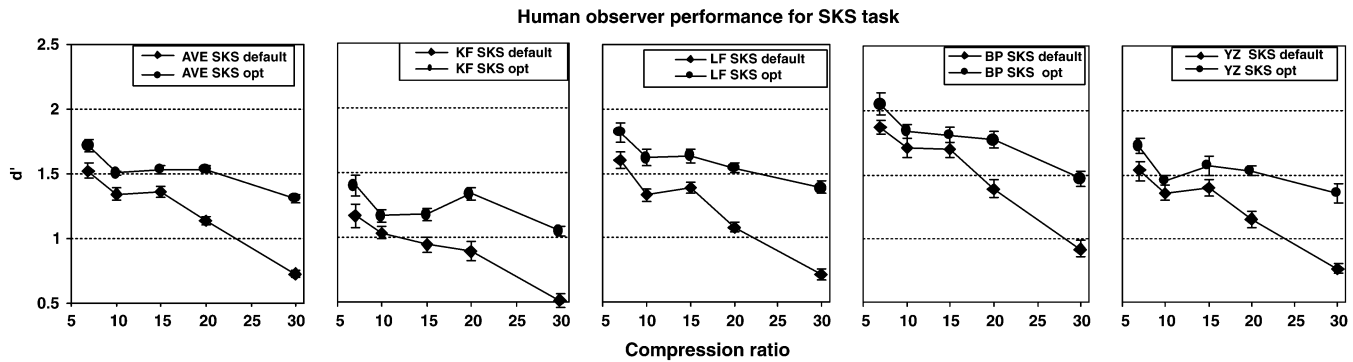


Fig. 11. Human observer performance for NPWE-optimized and default JPEG 2000 encoder settings for SKS tasks (from left to right): Average human observer (AVE) and individual observers: KF, LF, BP, and YZ.

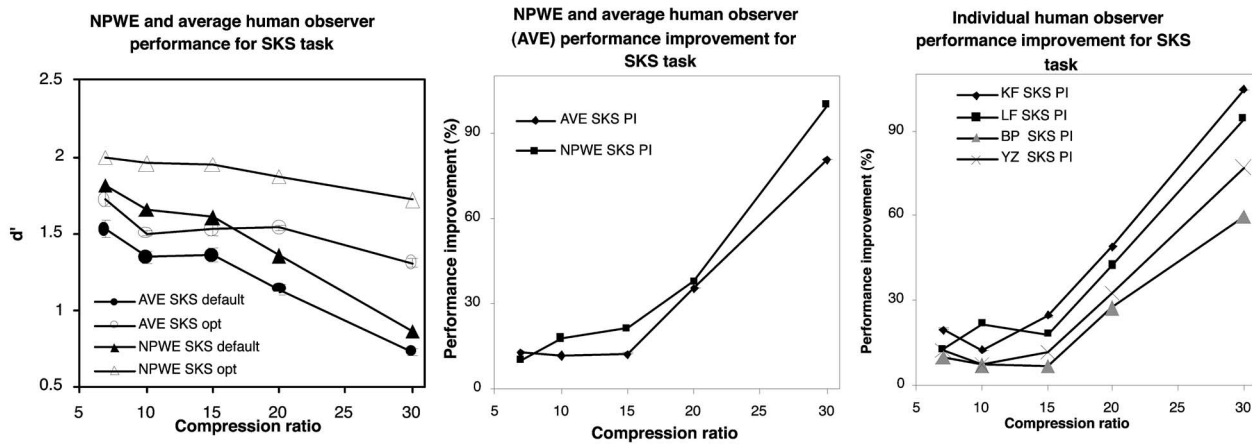


Fig. 12. NPWE and human observer performance for SKS tasks using default and NPWE-optimized encoder settings. Left: NPWE versus AVE performance; Middle: NPWE versus AVE performance improvement; Right: Individual human observer (KF, LF, BP and YZ) performance improvement.

more clinically relevant SKS task even though the optimization was based on an SKEV task. Fig. 10 shows model observer performance for the SKS tasks (see Appendix II for the calculation of model observer performance for SKS tasks). The left and the middle graphs of Fig. 10 compare the model observer performance for the SKS tasks with the default and NPWE-optimized JPEG 2000 encoder settings, respectively. Performance of different model observers follows similar trends. Fig. 10 (right graph) shows model observer performance improvements. The results show that the NPWE-optimized encoder settings for an SKEV task also result in better model observer performance for an SKS task. All four model observers indicate a performance improvement for the evaluated five compression ratios and result in similar performance improvement trends. The performance improvement ranges from 10% (at compression ratio 7:1) to 70% (at compression ratio 30:1).

Fig. 11 shows the average and individual human observer performance using the default and the optimal encoder settings for an SKS task. Average and all four individual human observers present higher performance for the NPWE-optimized encoder settings. For the optimized encoder settings, observers YZ and KF performed worse with compression ratio 10:1 than with compression ratio 15:1. Note that this reverse rank order of performance for 10:1 and 15:1 compression ratios is not unique to the optimized conditions. Some observers (LF and YZ) also show that trend for the default encoder setting suggesting that the reverse rank order is not due to a faulty optimization.

The left graph of Fig. 12 compares the NPWE model observer performance with the average human observer performance for the SKS tasks. Performance follows similar trends although the humans perform consistently lower than the model observer. The middle and the right graphs of Fig. 12 compare the performance improvement from the default encoder settings to the NPWE-optimized settings. Both average human and individual human performance improvements are similar to NPWE model observer performance improvements. At the compression ratio of 30:1, the average human observer performance improvement is around 80% and the highest individual performance improvement is around 100%. Even at the lower compression ratios (7:1 and 10:1) the average human observer performance improvement can reach 12%. These results show that NPWE-optimized encoder settings based on the more computationally tractable SKEV task leads to improved human and model observer performance in the more clinically relevant SKS task.

3) *SKEV Versus SKS*: Fig. 13 compares model observer performance for SKEV and SKS tasks using the NPWE-optimized JPEG 2000 encoder settings. In general, model observer performance in the SKS task is lower than that in the SKEV task (the first row of Fig. 13). This is to be expected from the theory of signal detectability given that uncertainty about relevant parameters of the signal to be detected will degrade performance [49]. Interestingly, the performance improvements (from the default to the NPWE-optimized encoder option settings) for the SKEV and SKS tasks are similar (the second row of Fig. 13).

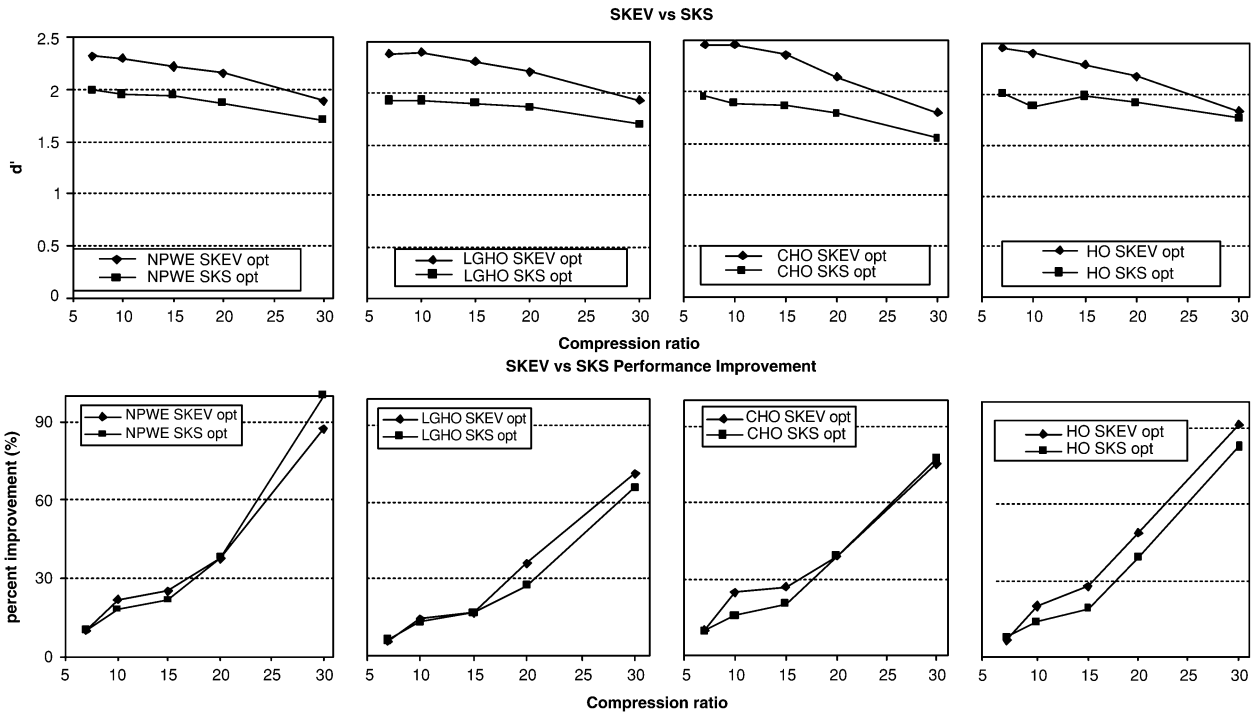


Fig. 13. Model observers (from left to right: NPWE, LGHO, CHO, and HO) performance comparison for SKEV and SKS tasks. Top: Model observer performance in d' ; Bottom: Performance improvement (%) from the default to NPWE-optimized JPEG 2000 encoder settings.

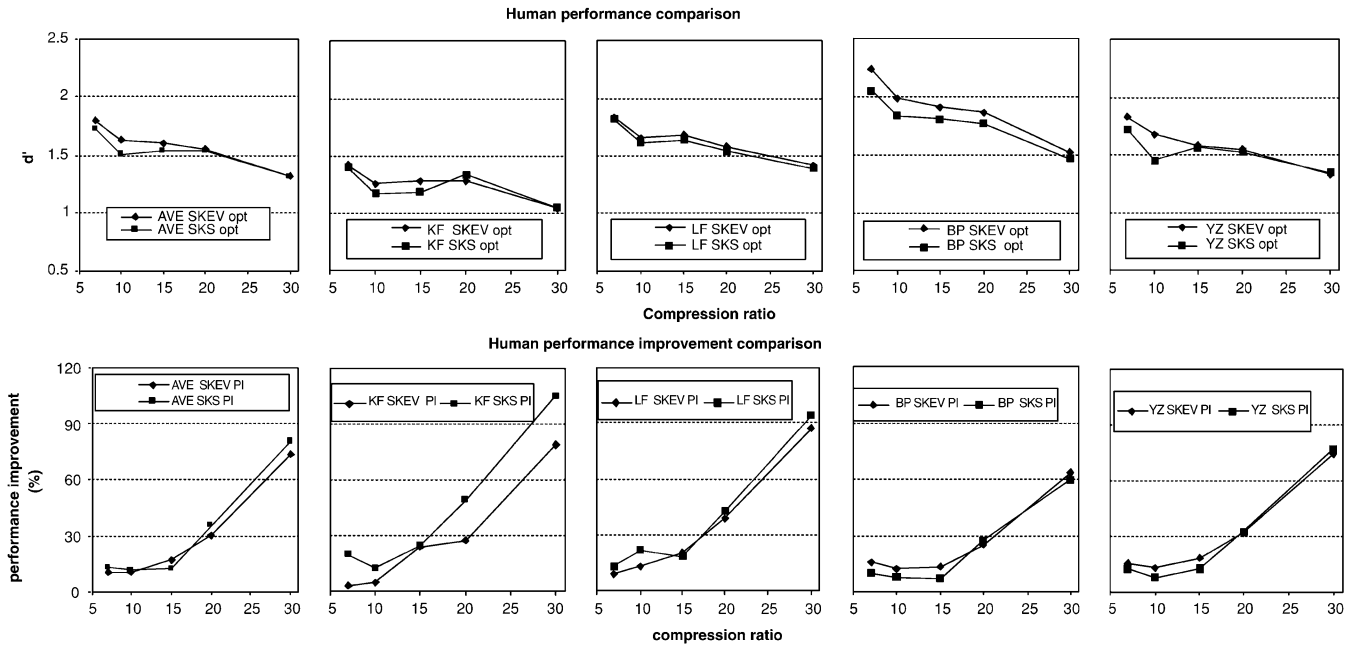


Fig. 14. Human observer performance comparison for SKEV and SKS tasks (from left to right): Average human observer (AVE), KF, LF, BP, and YZ. Top: Human observer performance in d' ; Bottom: Performance improvement (%) from the default to NPWE-optimized JPEG 2000 encoder settings.

The first row of Fig. 14 shows the average and the individual human observer performance for the SKEV and SKS tasks using the NPWE-optimized JPEG 2000 encoder settings. Note that performance degradation of human observer from the SKEV task to SKS task is smaller than that of model observers (Fig. 13). The performance differences between the SKEV and SKS tasks decrease with increasing compression ratios for both the human and model observers. In addition, the performance improvements (from the default to the NPWE-optimized

encoder settings) of all four human observers are very similar for the SKEV and SKS tasks (the second row of Fig. 14).

B. Performance Equivalent Compression Ratio

A potentially useful measure when comparing different compression schemes is to identify compression ratios that lead to similar observer performance across two different compression schemes. Fig. 15 (top-left table) shows arrows linking compression ratios that lead to equivalent NPWE performance (d') for

Compression Ratio	NPWE SKEV default (d')		NPWE SKEV opt (d')		Compression Ratio	NPWE SKS default (d')		NPWE SKS opt (d')
7	2.11	↔	2.33	↔	7	1.81	↔	2.00
10	1.89		2.30		10	1.66		1.96
15	1.79		2.23		15	1.60		1.94
20	1.57		2.16		20	1.36		1.87
30	1.01		1.90		30	0.86		1.71

Compression Ratio	Ave Human SKEV default (d')		Ave Human SKEV opt (d')		Compression Ratio	Ave Human SKS default (d')		Ave Human SKS opt (d')
7	1.63	↔	1.80	↔	7	1.53	↔	1.72
10	1.47		1.62		10	1.34		1.50
15	1.37		1.60		15	1.36		1.53
20	1.19		1.55		20	1.13		1.53
30	0.76		1.32		30	0.73		1.31

Fig. 15. Task performance equivalent compression ratio.

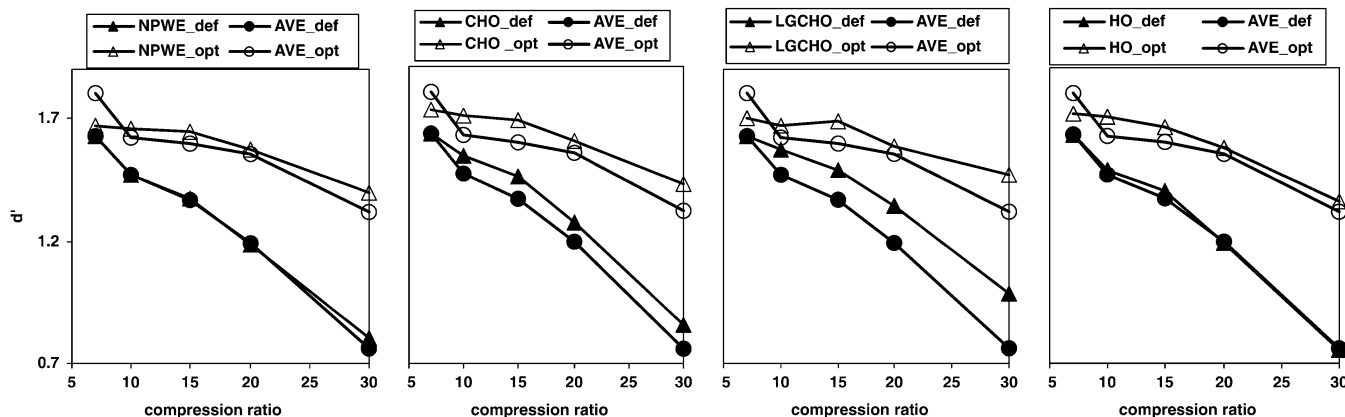


Fig. 16. Model observer (with internal noise) and human observer performance for the default and the optimized JPEG 2000 encoder settings for SKEV task. From left to right: NPWE, CHO, LGCHO, HO. The internal noise level was adjusted to match the model observer performance to the average human observer performance at compression ratio 7:1 using default JPEG 2000 encoder setting.

the SKEV tasks. For example, a compression ratio of 20:1 with the NPWE-optimized encoder results in equivalent NPWE performance to a 7:1 compression ratio with the default encoder setting. Fig. 15 (top-right table) shows arrows linking compression ratios that lead to equivalent NPWE performance (d') for the SKS tasks.

The bottom-left table and bottom-right table in Fig. 15 show arrows linking human performance equivalent compression ratios for the SKEV and SKS tasks, respectively.

For the NPWE model observer, a compression ratio of 30:1 with the optimized setting leads to equivalent performance to a compression ratio of 10:1 with the default setting (for both the SKEV task and SKS tasks). From the average human observer data, performance at a compression ratio of 20:1 with the optimized settings leads to an equivalent performance to a compression ratio of 10:1 with the default setting (for both the SKEV task and SKS tasks). These findings strongly support the idea that results obtained from the SKEV task can be extended to the SKS tasks.

C. Internal Noise

Some previous studies [10], [23], [25], [54], [53] have used the addition of internal noise to degrade model observer performance to a level that is comparable to human performance. For the present work, internal noise for model observers in the SKEV tasks was implemented by adding a random variable (ϵ) to their scalar decision variables [λ in (1)]. The random variable was sampled independently for each location and trial from a normal distribution with zero mean and a variance proportional to the decision variable's variance due to the external noise ($\sigma_{\lambda_e}^2$). This proportional internal noise has been used by other investigators [53], [23]. However, unlike previous treatments with SKE tasks, in the SKEV task there are different model templates corresponding to the different signals. Making the internal noise proportionality parameter specific to each signal/template would give rise to too many free parameters. We, therefore, used the same proportionality constant (κ) for all templates. The models' decision variable after inclusion of

internal noise is given by: $\lambda_{in,con} = \lambda_{e,con} + \epsilon_{con}$, where the subscript con represents compression condition, $\lambda_{in,con}$ is the decision variable after injection of internal noise, $\lambda_{e,con}$ is the decision variable due to the external noise [prior to internal noise injection, (1)], and ϵ_{con} is sampled from $N(0, \kappa\sigma_{\lambda_{e,con}}^2)$. The constant κ was iteratively changed so that the model observer performance matched the average human performance for the 7:1 compression ratio (default JPEG 2000 encoder setting) for the SKEV task. The same proportionality constant of internal noise was then used for the other 9 compression conditions (10:1, 15:1, 20:1, and 30:1 using default encoder setting and 7:1, 10:1, 15:1, 20:1, and 30:1 using optimized encoder setting) in the SKEV task. Fig. 16 shows that model observer performance with the inclusion of internal noise does not change the rank order of the encoder settings. The optimized encoder settings result in better model observer performance for every evaluated compression ratios. The results also show a very good correspondence between model and human observer performance.

D. Implementation of the Compression Ratio Specific Encoder Settings

Fig. 17 compares the NPWE performance for an SKEV task where each of the optimized encoder settings obtained for the specific compression ratios (7:1, 10:1, 15:1, 20:1, and 30:1) is used for all other studied compression ratios. The results show that there is little variability in NPWE performance across the compression ratio specific optimized encoder settings with one exception. The optimized encoder setting specific for the compression ratio 10:1 results in much lower performance for other compression ratios. Table II shows that the optimized JPEG 2000 encoder setting for compression ratio 10:1 is different from the optimized encoder settings for other compression ratios. This finding might suggest that the implementation of the 10:1 optimization was faulty or a statistical anomaly and did not result in the optimized parameters. To test this possibility we explicitly compared model observer performance for the 10:1 compression ratio. We evaluated the 10:1 and 15:1 NPWE optimized encoder settings (Table II). Table III shows that all model observers perform better using optimized encoder setting for the compression ratio 10:1 rather than that for the compression ratio 15:1.

IV. DISCUSSION

A. Medical Image Compression Optimization: SKE, SKEV, and SKS Tasks

We optimized JPEG 2000 encoder settings based on NPWE model observer performance for an SKEV task. Previous studies have shown that the NPWE, CHO, and HO model observers can accurately predict human detection performance for different image compression algorithms and parameters for the simpler SKE task [25], [26]. The SKE task is simple but not representative of the array of possible filling defects (signals) in a clinical scenario in which signals vary in size and shape across images. Moreover, the physician does not know *a priori* the shape and size of the signal to appear. For this reason we investigate SKEV tasks that attempt to compensate for the lack of signal variability of the SKE task. In addition, the computational economy of the

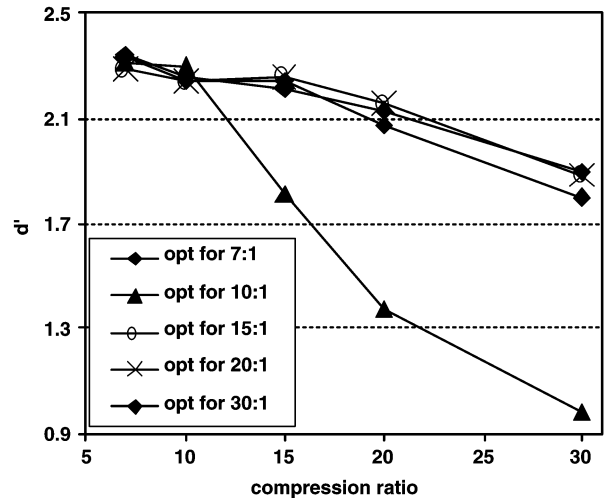


Fig. 17. NPWE model observer performance for an SKEV task using compression ratio specific encoder settings: “opt for x:1” represents the encoder settings optimized for the specific compression ratio “x:1.”

TABLE III
MODEL OBSERVER PERFORMANCE (d') COMPARISON SPECIFIC FOR THE 10:1 COMPRESSION RATIO IN AN SKEV TASK

	NPWE	CHO	HO	LGHO
Optimized encoder setting for 10:1	2.30	2.46	2.42	2.40
Optimized encoder setting for 15:1	2.23	2.37	2.38	2.38

SKEV task becomes important since we are interested in conducting model observer based optimizations through iterative methods. However, the SKEV approach is useful if and only if the SKEV task performance is highly correlated with the SKS task performance. For example, the optimized JPEG 2000 encoder settings should lead to better model/human observer performance than the default ones whether an SKEV or an SKS task is chosen. Our results in Section III show a close relationship between the two tasks. In general, performance is expected to be worse in the SKS task because the model/human cannot use prior knowledge to ignore templates corresponding to signals that are not present in that trial. This is consistent with the concept of stimulus uncertainty on human and model performance [49]. Note that in Figs. 13 and 14 the difference across tasks for the human observers is smaller than that for the model observers. This is explained by the fact that humans cannot use precise knowledge of the signal information even when the signal is shown to them (intrinsic uncertainty [55], [56]). In summary, our results indicate that for the present set of signals the computationally simpler SKEV task can be reliably used to optimize task based image quality for the more clinically realistic SKS task.

B. NPWE Performance Optimization Versus Other Model Observers and Human Observer Performance

The NPWE performance results were compared with human observer and other model observer (HO, CHO, and LGHO) performance for both the SKEV and the more clinically realistic SKS tasks. Our results demonstrate that all four model observers evaluated show similar signal detection performance trends on default versus NPWE-optimized JPEG 2000 encoder settings

for the SKEV and SKS task. The results imply that the conclusions about the optimized JPEG 2000 encoder settings for the SKEV task would not change if we used any of the four model observer performance as the fitness function of the GA. The fact that the simpler NPWE model is a good predictor of human performance saves a great amount of computation time in the optimization procedure. Obtaining NPWE performance in the SKEV task for the 900 test images takes around 5 min, while our approach to obtain CHO performance for the same task takes about 4 h [27]. Please refer to Appendix I for the CHO template computation.

C. Performance Equivalent Compression Ratio for Default and Optimized Encoder Settings

The results of our study show that we can increase the compression ratio by a factor of 2 by using the optimized JPEG 2000 encoder settings rather than the default ones without degrading signal detection performance. This indicates that through optimization we can significantly mitigate the storage and transmission problems that arise due to the large volume of digital angiograms. However, our data indicates that the NPWE model observer predicts larger benefits for equivalent compression ratios than that shown by human observers. For example for both the SKEV and SKS task the NPWE model observer performance at a compression ratio of 20:1 using the NPWE-optimized encoder setting was equivalent to its performance at a compression ratio of 7:1 using the default encoder setting. In contrast, the average human observer (for both the SKEV and SKS task) performance at a compression ratio of 15:1 using the NPWE-optimized encoder setting was equivalent to performance at a compression ratio of 7:1 using the default encoder setting.

D. Compression Ratio Specific Encoder Settings Versus Compression Ratio Independent Optimized Encoder Setting

One question of interest when optimizing JPEG 2000 encoder options is whether the settings that lead to best model observer performance vary across compression ratios. Our results (Table II) show that the optimized encoder settings for each specific compression ratio are different from each other. However, in general it is possible to choose one general optimal encoder setting rather than a compression ratio specific encoder settings for a range of compression ratios (Fig. 17). There are some exceptions such as the case observed for the 10:1 optimized JPEG 2000 encoder setting that leads to inferior performance when applied to the other compression ratios (Table II).

E. Internal Noise

The addition of internal noise [23], [53], [54] to the model observers typically allows for better prediction of human absolute performance. In this paper, we were mostly concerned about the rank order of different JPEG 2000 encoder settings. The optimization based on NPWE performance was implemented without inclusion of internal noise. In this context, an important question is whether the addition of internal noise to the models in the evaluation stage changes the rank order of the encoder settings. Our results showed that the rank order of the encoder settings did not change with the inclusion of internal

noise (Fig. 16). Thus, for all model observers and conditions the JPEG 2000 optimized encoder settings led to higher performance than the default encoder settings irrespective of whether internal noise was included. A separate question is whether inclusion of internal noise in the NPWE model for the optimization would lead to better encoder settings. This question was not evaluated and needs to be addressed in the future.

F. Relation to Clinical Studies

Coronary angiograms consist of dynamic image sequences. However, in the present optimization model and human detection performance is evaluated for static images extracted from X-ray coronary angiographic sequences. This choice is based on computational simplicity. Previous results have shown that although absolute human performance is higher for the dynamic X-ray angiograms, the rank order of two compression schemes is the same for both the static and dynamic image displays [57].

Previous research has reported clinical studies evaluating the effect of JPEG lossy compression on diagnostic decisions for X-ray coronary angiograms [2], [4], [58], [59]. A growing number of commercial vendors have been granted Food and Drug Administration approval to market digital angiographic storage and review systems that use lossy JPEG [60]. With the emergence of wavelet based compression algorithms, many researchers have investigated these wavelet based compression algorithms [10], [61].

The recently developed JPEG 2000 standard serves as a powerful tool for providing a set of features that are of importance to many high-end and emerging applications by taking advantage of new technologies [62]. In this paper, we studied the effect of the new JPEG 2000 standard on diagnostic decisions for X-ray coronary angiograms. The obtained experimental results might serve to guide the choice of JPEG 2000 encoder settings in future clinical studies evaluating the effect of JPEG 2000 on diagnostic performance with X-ray coronary angiograms.

V. CONCLUSION

A GA was implemented to optimize the JPEG 2000 encoder settings with respect to the nonprewhitening matched filter with an eye filter (NPWE) model observer detection performance in a task with varying signals (SKEV). The NPWE-optimized encoder settings led to large improvements in performance for human observers as well as other model observers (Hotelling, channelized-Hotelling and Laguerre-Gauss Hotelling) for both the SKEV task and the more clinically realistic signal known statistically task (SKS). The human performance improvements reached a maximum of 75% (d') for the higher compression ratio like 30:1. The results show how optimization schemes can be used in conjunction with model observers to optimize task-based medical image quality using clinical backgrounds and a complex task that includes signal variability.

APPENDIX I

THEORY OF MODEL OBSERVERS

A. Hotelling and Laguerre-Gauss Hotelling Observer

The HO uses information about the signal and the spatial correlation of the noise to derive its corresponding templates. In the

absence of internal noise and assuming an invertible image covariance matrix, the Hotelling-observer template is given by

$$\omega_{Hot} = \mathbf{K}_g^{-1}[\langle \mathbf{g}_s \rangle - \langle \mathbf{g}_n \rangle] \quad (\text{A.I})$$

where \mathbf{K}_g is the image covariance matrix, $\langle \mathbf{g}_s \rangle$ is the mean signal plus background, and $\langle \mathbf{g}_n \rangle$ is the mean background. One potential problem in the evaluation of the Hotelling model is the inversion of the covariance matrix in (A.1). If the image is large then the covariance will require prohibitively large number of samples to be inverted. There have been two solutions in the literature [24], [44]. The first solution is to constrain the covariance matrix to a square region around the possible signal locations. In the current study, 30×30 pixel regions were selected. The second approach used a set of functions to extract some main features from the images, such as Laguerre-Gauss polynomials [24], [44], [48], $L_n(x) = \sum_{m=0}^n (-1)^m \binom{n}{m} (x^m/m!)$. The channels are expressed in polar coordinates with origin at the signal location

$$\mathbf{C}_{LG} = \exp(-\gamma/2)L_n(\gamma), \quad (\text{A.II})$$

where $\gamma = 2\pi(((x - x_c)^2)/(a^2) + ((y - y_c)^2)/(b^2))$, x_c , and y_c denote the center of the signal, and a and b are used to determine the orientation of the channels. For our implementation we used up to the sixth order of the Laguerre-Gauss polynomials and three orientations: vertically oriented ($a = 5, b = 14$), horizontally oriented ($a = 14, b = 5$), or rotationally invariant ($a = b = 8$) resulting in a total of 18 channels. The use of oriented channels is due to the fact that our signals are oriented. Note that the Laguerre-Gauss channels are not intended to reflect a feature extraction stage in the human visual system but are solely used to mitigate the computational problems of calculating the Hotelling template for a limited number of samples. We then used the channel response covariance matrix rather than the pixel covariance matrix as in HO to get the template for LGHO.

The best linear template that can be obtained from the channel weights and the channel profiles is calculated as in [10]

$$\omega_{LGHO} = \mathbf{W}_c^t \mathbf{C}_{LG} \quad (\text{A.III})$$

where \mathbf{W}_c is a vector containing the optimal linear weights for each of the channels and defined by [18], [45], [46]

$$\mathbf{W}_c = \mathbf{K}_v^{-1}[\langle \mathbf{g}_{C/s} \rangle - \langle \mathbf{g}_{C/b} \rangle] \quad (\text{A.IV})$$

where \mathbf{K}_v is a $N \times N$ matrix describing the covariance of the output of the channels to the images. For our particular implementation, the covariance matrix consists of an 18×18 matrix. The vector $\langle \mathbf{g}_{C/s} \rangle$ contains the mean signal plus backgrounds as seen by each channel. For each of the 184 signals, 400 additional samples are created as a training set to estimate $\langle \mathbf{g}_{C/s} \rangle$. Estimation of $\langle \mathbf{g}_{C/b} \rangle$ is based on 2700 samples. Thus, a total of $184 \times 400 + 2700$ test locations needs to be computed for the LGHO model to estimate the best linear weights. There might be alternative approaches to reduce the computation time but these were not investigated here.

B. Channelized-Hotelling Observer

Unlike the LGHO, the CHO does use a set of channels that are intended to reflect a preprocessing stage in the human visual system. The set of channels reduce the amount of information available to the model [45]. The channelized Gabor Hotelling template is the best template that can result from a linear combination of the spatial frequency and orientation channels that supposedly reflect the response of neurons in the primary visual cortex [18], [24], [25]. The Gabor channels are given by

$$C_{CHO}(x, y) = \exp\left(\frac{-4 \ln 2(x^2 + y^2)}{W_s^2}\right) \times \cos[2\pi f_c(x \cos \theta + y \sin \theta) + \beta], \quad (\text{A.V})$$

where f_c is the spatial frequency, θ is the orientation, W_s is the width, and β is the phase.

In this study we used an 80-channel model with five spatial frequencies (central frequencies, 16, 8, 4, 2, and 1 cycle per degree), eight orientations (equally spaced), and two phases (odd, 0 and even, $\pi/2$). The spatial frequency bandwidth of the channels was approximately 1 octave. The template calculation for CHO is similar to that for LGHO as described in Appendix I-A.

APPENDIX II

CALCULATION OF THE MODEL OBSERVER PERFORMANCE FOR SKS TASKS

SKS tasks are different from the SKEV tasks in that the models do not know *a priori* the signal type presented in the trial and, therefore, require computing a dot product between a template corresponding to each possibly present signal and the data at each location [63], [64]

$$\lambda_{i,j} = \sum_{x=1}^N \sum_{y=1}^N \omega_j(x, y) g_i(x, y) = \omega_j^t \mathbf{g}_i, \quad (\text{A.VI})$$

where $\lambda_{i,j}$ is the scalar response of a template of the j th signal type to the data at the i th location. The result of this first linear stage of SKS models consists of J scalar responses (one per template) per location or alternative. The second stage in the SKS model requires combining the J scalar responses per location into a single scalar response that can then be used as a decision variable to choose the signal location. Here, we use the sum of likelihood rule to combine information across the different scalar outputs [49]. We assume that the internal responses of each template/filter is Gaussian distributed. On each trial, the likelihood of the filter response (l_i) is calculated given that any one of the signal is present at the i th location and is not present at the remaining $M - 1$ locations. The model selects the location with the highest sum of likelihoods (l_i). The expression for the sum of likelihood is given by

$$l_i = \sum_{j=1}^J \left[\frac{1}{\sqrt{2\pi\sigma_j^2}} \right]^M \exp\left(-\frac{(\lambda_{i,j} - \mu_{s,j})^2}{2\sigma_j^2}\right) \times \prod_{m=1}^M \left[\exp\left(-\frac{(\lambda_{m,j} - \mu_{n,j})^2}{2\sigma_j^2}\right) \right]^{(1-\delta_{m,i})} \quad (\text{A.VII})$$

where the summation is over all J possible signal types, $\mu_{s,j}$ is the expected response of the j th template to the j th signal type, $\mu_{n,j}$ is the expected response of the j th template to the j th signal absent locations, σ_j is the standard deviation of the response of the j th template, M is the total number of locations, and $\delta_{m,i}$ is the Kronecker delta which takes a value of 1 for $i = m$ and 0 otherwise. The exponent $(1 - \delta_{m,i})$ allows computing the probability of the signal being absent at all but the i th location (when $i = m$ then $1 - \delta_{m,i} = 0$) where the probability of the signal being present is being computed.

Performance for the model observers in the SKS task can be mathematically expressed as

$$P_c = \frac{1}{H} \sum_{h=1}^H \text{step}(l_{s,h} - \max(l_{n,h})) \quad (\text{A.VIII})$$

where $l_{s,h}$ is the sum of likelihood as in (A.VIII) at the signal location in trial h and $\max(l_{n,h})$ is the maximum of the likelihoods at the nonsignal locations.

APPENDIX III

GENETIC ALGORITHM SOFTWARE PROTOTYPE

The programs are written using MATLAB 6.1 and simulations are conducted on a Pentium III 1.8-GHz platform.

The proposed algorithm is summarized as follows.

- 1) Randomly generate $N = 41$ chromosomes (JPEG 2000 encoder settings) for the first population P , where each gene takes values from the parameter space shown in Table I. Let $f^* = 0, g = g^* = 0$, where g represents the current generation number, g^* denotes the generation number when the overall best chromosome, among all generations, is first found, and f^* denotes the best chromosome associated fitness value.
- 2) Run the compression/uncompression for the test images using each chromosome in the current population P and calculating the NPWE performance. If $F(C) > f^*, C^* = C, f^* = F(C), g^* = g$, where C^* represents the best chromosome in the current generation, i.e., the one with highest fitness value in P and C represents the chromosome in the current population.
- 3) Select the best chromosome to the next generation directly and then using roulette wheel selection scheme to select other $N - 1$ chromosomes from the current population P . Apply crossover operator at the probability 0.8 and mutation operator at the probability 0.04 to the $N - 1$ selected chromosomes and creating new chromosomes for population P' of the next generation. Updating the generation number $g = g + 1$.
- 4) If $g - g^*$ is greater than the threshold $q_{\text{cons}} = 15$, Goto Step 5);
 else if $g > 60$, Goto Step 5);
 else $P = P'$, Goto Step 2).
- 5) Report the best chromosome C^* that indicates the best solution for our problem.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments on improving this paper. They would also like to thank L. Furchtgott and K. Friedman for participation as observers in the study. Finally, they would also like to thank B. Monte for his assistance in proofreading this paper.

REFERENCES

- [1] J. Oh and S. I. Woolley, "Diagnostic quality testing for wavelet-compressed digital angiogram images," in *Proc. Inst. Elect. Eng. Image Processing and Applications Conf. (IPA'99)*, Manchester, U.K., 1999, pp. 512–516.
- [2] G. Koning, P. Bretta, P. Zwart, E. Hekking, and J. H. Reiber, "Effect of lossy data compression on quantitative coronary measurements," *Int. J. Cardiac Imag.*, vol. 13, pp. 261–270, Aug. 1997.
- [3] B. J. Erickson, "Irreversible compression of medical images," *J. Digital Imag.*, vol. 15, pp. 5–14, Apr. 2002.
- [4] K. R. Persons, N. J. Hangiandreou, N. T. Charboneau, J. W. Charboneau, E. M. James, B. R. Douglas, A. P. Salmon, J. M. Knudsen, and B. J. Erickson, "Evaluation of irreversible JPEG compression for a clinical ultrasound practice," *J. Digital Imag.*, vol. 15, pp. 15–21, Apr. 2002.
- [5] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG 2000 still image coding system: An overview," *IEEE Trans. Consumer Electron.*, vol. 46, pp. 1103–1127, Nov. 2000.
- [6] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, pp. 205–220, Apr. 1992.
- [7] M. D. Adams and F. Kossentini. Jasper: A software-based JPEG-2000 codec implementation. presented at Int. Conf. Image Processing. [Online]. Available: <http://www.ece.uvic.ca/mdadams/jasper/>
- [8] M. Rabbani and R. Joshi, "An overview of the JPEG 2000 still image compression standard," *Signal Processing: Image Commun.*, vol. 17, pp. 3–48, 2002.
- [9] J. L. Mannos and J. D. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 525–536, Apr. 1974.
- [10] M. P. Eckstein, J. L. Bartroff, C. K. Abbey, J. S. Whiting, and F. O. Bochud, "Automated computer evaluation and optimization of image compression of X-ray coronary angiograms for signal known exactly detection tasks," *Opt. Express*, vol. 11, pp. 460–475, Mar. 2003.
- [11] H. H. Barrett, J. Yao, J. P. Rolland, and K. J. Myers, "Model observers for assessment of image quality," *Proc. Nat. Acad. Sci.*, vol. 90, pp. 9758–9765, 1993.
- [12] K. J. Myers, "Ideal observer models of visual signal detection," in *Handbook of Mescal Imaging: Physics and Psychophysics*. Bellingham: SPIE, 2000, pp. 559–592.
- [13] M. A. Kupinski, J. W. Hoppin, E. Clarkson, and H. H. Barrett, "Ideal-observer computation in medical imaging with use of markov-chain monte carlo techniques," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 20, pp. 430–439, 2003.
- [14] A. E. Burgess, R. F. Wagner, R. J. Jennings, and H. B. Barlow, "Efficiency of human visual signal discrimination," *Science*, vol. 214, pp. 93–94, 1981.
- [15] A. E. Burgess and H. Ghandeharian, "Visual signal detection. II. Signal location identification," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 1, pp. 906–910, 1984.
- [16] K. J. Myers, H. H. Barrett, M. C. Borgstrom, D. D. Patton, and G. W. Seeley, "Effect of noise correlation on detectability of disk signals in medical imaging," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 2, pp. 1752–1759, 1985.
- [17] R. D. Fiete, H. H. Barrett, W. E. Smith, and K. J. Myers, "Hotelling trace criterion and its correlation with human-observer performance," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 4, pp. 945–963, May 1987.
- [18] J. Yao and H. H. Barrett, "Predicting human performance by a channelized Hotelling observer model," in *Math. Meth. Med. Imag., SPIE*, vol. 1768, 1992, pp. 161–168.
- [19] J. P. Rolland and H. H. Barrett, "Effect of random background inhomogeneity on observer detection performance," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 9, pp. 649–658, 1992.
- [20] A. E. Burgess, "Visual signal detection with two-component noise: Low-pass spectrum effects," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 16, pp. 694–704, 1996.

- [21] "Medical Imaging—The Assessment of Image Quality," Tech. Rep.54, 1996.
- [22] A. E. Burgess, X. Li, and C. K. Abbey, "Visual signal detectability with two noise components: Anomalous masking effects," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 14, pp. 2420–2442, 1997.
- [23] C. K. Abbey and H. H. Barrett, "Human- and model-observer performance in ramp-spectrum noise: Effects of regularization and object variability," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 18, pp. 473–488, 2001.
- [24] M. P. Eckstein and J. S. Whiting, "Lesion detection in structured noise," *Acad. Radiol.*, vol. 2, pp. 249–253, 1995.
- [25] M. P. Eckstein, C. K. Abbey, F. O. Bochud, J. L. Bartroff, and J. S. Whiting, "Effect of image compression in model and human performance," in *Proc. SPIE: Medical Imaging 1999: Image Perception and Performance*, vol. 3663, E. A. Krupinski, Ed., 1999, pp. 243–252.
- [26] M. P. Eckstein, C. K. Abbey, and J. L. Bartroff, "Model observer based optimization of image compression algorithms," in *Proc. SPIE: Medical Imaging: Image Perception*, vol. 3981, K. M. Hanson, Ed., 2000, pp. 106–115.
- [27] Y. Zhang, B. Pham, and E. P. Eckstein, "Evaluation of JPEG 2000 Encoder Options: Human and Model Observer Detection of Variable Signals in X-Ray Coronary Angiograms," *IEEE Trans. Med. Imag.*, 2004, to be published.
- [28] Y. Zhang, B. Pham, and M. P. Eckstein, "Investigation of jpeg 2000 encoder options on model observer performance in signal known exactly but variable tasks (SKEV)," in *Proc. SPIE: Medical Imaging: Image Perception, Observer Performance and Technology Assessment*, vol. 5034, D. P. Chakraborty and E. A. Krupinski, Eds., 2003, pp. 371–382.
- [29] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: Univ. Michigan Press, 1975.
- [30] P. Andrey and P. Tarroux, "Unsupervised segmentation of Markov random field modeled textured images using selectionist relaxation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20, no. 3, pp. 252–262, 1998.
- [31] D. Beasley, D. R. Bull, and R. R. Martin, *An Overview of Genetic Algorithms*, 1993.
- [32] S. M. Bhandarkar and H. Zhang, "Image segmentation using evolutionary computation," *IEEE Trans. Evol. Comput.*, vol. 3, pp. 1–21, 1999.
- [33] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.
- [34] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Design*. New York: Wiley, 1997.
- [35] M. Anastasio, H. Yoshida, R. Nagel, R. Nishikawa, and K. Doi, "A genetic algorithm-based method for optimizing the performance of a computer-aided diagnosis scheme for detection of clustered microcalcifications in mammograms," *Medical Physics*, vol. 25, pp. 1613–1620, 1998.
- [36] J. Yang and V. Honavar, "Feature subset selection using a genetic algorithm," *IEEE J. Intell. Syst.*, vol. 13, pp. 44–49, Mar./Apr. 1998.
- [37] M. Gudmundsson, E. A. El-Kwae, and M. R. Kabuka, "Edge detection in medical images using a genetic algorithm," *IEEE Trans. Med. Imag.*, vol. 17, pp. 469–474, June 1998.
- [38] C. K. Chow, H. T. Tsui, T. Lee, and T. K. Lau, "Medical image registration and model construction by genetic algorithms," in *Proc. Int. Workshop Medical Imaging and Augmented Reality*, 2001, pp. 174–179.
- [39] J. M. Rouet, J. J. Jacq, and C. Roux, "Genetic algorithms for a robust 3-D MR-CT registration," *IEEE Trans. Inform. Technol. Biomed.*, vol. 4, pp. 126–136, June 2000.
- [40] C. A. Morioka, C. K. Abbey, M. P. Eckstein, R. A. Close, J. S. Whiting, and M. LeFree, "Simulating coronary arteries in X-ray angiograms," *Med. Phys.*, vol. 27, pp. 2438–2443, 2000.
- [41] M. Ishida, K. Doi, L. N. Loo, C. E. Metz, and J. L. Lehr, "Digital image processing: Effect of detectability of simulated low-contrast radiographic patterns," *Radiology*, vol. 150, pp. 569–575, 1984.
- [42] R. F. Wagner and K. E. Weaver, "An assortment of image quality indices for radiographic film-screen combinations-can they be resolved?," in *Proc. SPIE: In Application of Optical Instrumentation in Medicine*, vol. 35, 1972, pp. 83–94.
- [43] A. E. Burgess, "Statistically defined backgrounds: Performance of a modified nonprewhitening matched filter model," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 11, pp. 1237–1242, 1994.
- [44] H. H. Barrett, C. K. Abbey, B. G. Gallas, and M. P. Eckstein, "Stabilized estimates of Hotelling-observer detection performance in patient-structured noise," in *Proc. SPIE, Medical Imaging 1998: Image Perception*, vol. 3340, H. L. Kundel, Ed., 1998, pp. 27–43.
- [45] K. J. Myers and H. H. Barrett, "Addition of a channel mechanism to the ideal-observer model," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 4, pp. 2447–2457, 1987.
- [46] M. P. Eckstein, C. K. Abbey, and F. O. Bochud, "A practical guide to model observers for visual detection in synthetic and natural noisy images," in *Handbook of Medical Imaging: Physics and Psychophysics*. Bellingham, WA: SPIE Press, 2000, pp. 593–628.
- [47] F. O. Bochud, C. K. Abbey, and M. P. Eckstein, "Visual signal detection in structured backgrounds. III. Calculation of figures of merit for model observers in statistically nonstationary backgrounds," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 17, pp. 193–205, 2000.
- [48] M. P. Eckstein, C. K. Abbey, and F. O. Bochud, "Visual signal detection in structured backgrounds IV. Figures of merit for model performance in multiple-alternative forced-choice detection tasks with correlated responses," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 17, pp. 206–217, 2000.
- [49] W. W. Peterson, T. G. Birdsall, and W. C. Fox, "The theory of signal detectability," *Trans. IRE Inform. Theory PGIT*, vol. 4, pp. 171–212, 1954.
- [50] A. E. Eiben, R. Hinterding, and Z. Michalewicz, "Parameter control in evolutionary algorithms," *IEEE Trans. Evol. Comput.*, vol. 3, pp. 124–141, July 1999.
- [51] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, 2nd ed. Berlin, Germany: Springer-Verlag, 1994.
- [52] DICOM Part 14: Grayscale Standard Display Function (2003). [Online]. Available: <http://medical.nema.org/dicom/2003.html>
- [53] A. E. Burgess and B. Colborne, "Visual signal detection. IV. Observer inconsistency," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 5, pp. 617–627, 1988.
- [54] M. P. Eckstein, A. J. Ahumada, and A. B. Watson, "Visual signal detection in structured backgrounds. II. Effects of contrast gain control, background variations, and white noise," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 14, pp. 2406–2419, 1997.
- [55] D. G. Pelli, "Uncertainty explains many aspects of visual contrast detection and discrimination," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 2, pp. 1508–1532, 1985.
- [56] W. P. J. Tanner, "Physiological implications of psychophysical data," *Ann. New York Acad. Sciences*, vol. 89, pp. 752–765, 1961.
- [57] C. A. Morioka, M. P. Eckstein, J. L. Bartroff, J. Hausleiter, G. Aharanov, and J. S. Whiting, "Observer performance for JPEG vs. wavelet image compression of X-ray coronary angiogram," *Opt. Express*, vol. 5, pp. 8–19, July 1999.
- [58] S. Silber, R. Drr, G. Zindler, H. Mhling, and T. Diebel, "Impact of various compression rates on interpretation of digital coronary angiograms," *Int. J. Cardiol.*, vol. 60, pp. 195–200, June 1997.
- [59] W. A. Baker, S. E. Hearne, L. A. Spero, K. G. Morris, R. A. Harrington, M. H. Sketch, V. S. Behar, Y. Kong, R. H. Peter, T. M. Bashore, J. K. Harrison, and J. T. Cusma, "Lossy (15:1) JPEG compression of digital coronary angiograms does not limit detection of subtle morphological features," *Circulation*, vol. 96, pp. 1157–1164, 1997.
- [60] J. L. Elion and J. S. Whiting, "Clinical use of lossy image compression in digital angiography," *Amer. J. Cardiol.*, vol. 78, pp. 219–220, 1996.
- [61] B. Zhao, L. H. Schwarz, and P. K. Kijewski, "Effects of lossy compression on lesion detection: Predictions of the nonprewhitening matched filter," *Med. Phys.*, vol. 25, pp. 1621–1624, 1998.
- [62] A. N. Skodras, C. A. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Processing Mag.*, pp. 36–58, Sept. 2001.
- [63] M. P. Eckstein, B. Pham, and C. K. Abbey, "Effect of image compression for model and human observers in signal-known-statistically tasks," in *Proc. SPIE: Medical Imaging: Image Perception, Observer Performance and Technology Assessment*, vol. 4686, D. P. Chakraborty and E. A. Krupinski, Eds., 2002, pp. 13–24.
- [64] M. P. Eckstein, Y. Zhang, and B. Pham, "Optimization of model observer performance for signal known exactly but variable tasks leads to optimized performance in signal known statistically tasks," in *Proc. SPIE: Medical Imaging: Image Perception, Observer Performance and Technology Assessment*, vol. 5034, D. P. Chakraborty and E. A. Krupinski, Eds., 2003, pp. 123–134.