

# Evaluation of JPEG 2000 Encoder Options: Human and Model Observer Detection of Variable Signals in X-Ray Coronary Angiograms

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

**Abstract**—Previous studies have evaluated the effect of the new still image compression standard JPEG 2000 using nontask based image quality metrics, i.e., peak-signal-to-noise-ratio (PSNR) for nonmedical images. In this paper, the effect of JPEG 2000 encoder options was investigated using the performance of human and model observers (nonprewhitening matched filter with an eye filter, square-window Hotelling, Laguerre-Gauss Hotelling and channelized Hotelling model observer) for clinically relevant visual tasks. Two tasks were investigated: the signal known exactly but variable task (SKEV) and the signal known statistically task (SKS). Test images consisted of real X-ray coronary angiograms with simulated filling defects (signals) inserted in one of the four simulated arteries. The signals varied in size and shape. Experimental results indicated that the dependence of task performance on the JPEG 2000 encoder options was similar for all model and human observers. Model observer performance in the more tractable and computationally economic SKEV task can be used to reliably estimate performance in the complex but clinically more realistic SKS task. JPEG 2000 encoder settings different from the default ones resulted in greatly improved model and human observer performance in the studied clinically relevant visual tasks using real angiography backgrounds.

**Index Terms**—Filling defect (signal), JPEG 2000, model observer, SKEV/SKS task.

## I. INTRODUCTION

ANGIOGRAMS constitute one of the most important data sources for cardiovascular radiology. Conventional angiograms are stored on film (hard copy) and take about 30 min to develop after each procedure [1]. A digital video record of the angiogram generated during the procedure enables clinicians to quickly review the data and perform a diagnosis without waiting for the film to be developed. Digital image data has the further advantage of being suitable for multiuser access, image retrieval, network transmission and image-manipulations as well as computer analysis [1]–[4].

However, one of the practical limitations in the storage and transmission of digital angiograms is the excessive size of patient

images. An individual cardiac catheterization procedure can occupy up to 1 Gigabyte [5]. Long term storage requirements are challenging because of the number of patients performed in one year. For example, a typical four procedure room catheterization laboratory can perform approximately 3000–4000 procedures a year. The shelf life of image sequences is long for children's catheterization procedures which must be stored until they reach 21 years of age and adult cine angiograms must be stored for 7 years after the procedure [6]. Moreover, very high data transfer rates for display are required for dynamic display of image sequences. A raw video stream tends to be quite demanding when it comes to storage requirements, and demand for network capacity when being transferred between computers. For example, a simple patient digital angiogram video typically requires 7.5 Mbytes/sec for  $512 \times 512 \times 8$ -bit resolution at 30 frames/sec, resulting in 0.25 Mbytes/frame [3]. These large data volumes can quickly fill available storage media and are cumbersome to transfer between sites over communications links on which data rates are limited.

The storage and transmission problems that arise due to the large volume of digital angiograms can be significantly mitigated by the use of image compression techniques. Before being stored or transferred, the raw stream can be transformed to a representation using compression. Researchers have investigated the effect of compression algorithms including the first still image compression standard JPEG and some wavelet based algorithms for X-ray coronary angiograms [6]–[10]. A new still image compression standard, the JPEG 2000, was developed recently [11]. By taking advantage of new technologies, JPEG 2000 provides a wide range of functionalities for many high-end and emerging applications, such as medical imagery, remote sensing, mobile application, and digital library. The JPEG 2000 standard offers a number of encoder options that directly affect the coding results. Rabbani and Joshi [12] compared some of the JPEG 2000 encoder options using the test images “bike,” “cafe,” and “woman” which were chosen from the JPEG 2000 test set. Santa-Cruz *et al.* [13] evaluated and assessed JPEG 2000 performance from various points of view, such as compression efficiency, complexity and the set of supported functionalities. They also used the images from the official JPEG 2000 test set. The metric of image quality used in [12], [13] was the peak signal to noise ratio (PSNR) which is a monotonic transformation of the root mean square error (RMSE) between the original image and the image that underwent compression. When the images are for medical use, the image quality can be defined objectively in terms of

Manuscript received December 3, 2003; revised February 5, 2004. This work was supported in part 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 A. Manduca. *Asterisk indicates corresponding author.*

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

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

Digital Object Identifier 10.1109/TMI.2004.826359

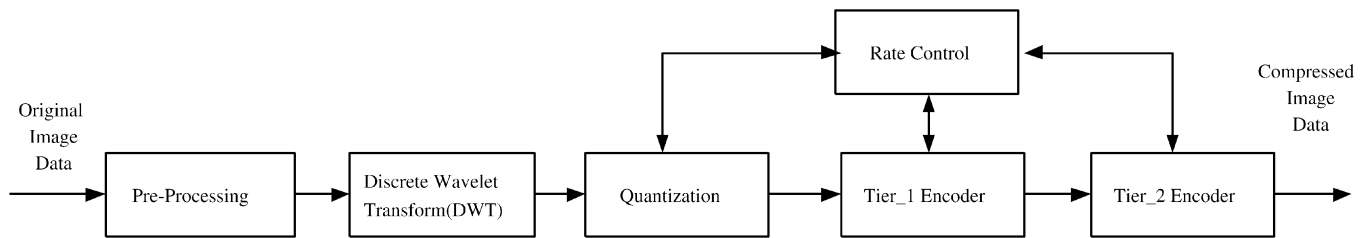


Fig. 1. General block diagram of the JPEG 2000 encoder.

diagnostic task performance [14]. 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 option to evaluate the effect of the encoder options is to conduct human observer psychophysical studies. However, such studies become overwhelming when the study parameter set is large since the number of images required increases geometrically as a function of the number of parameters. This problem can be potentially overcome by employing a computer model observer that can reflect the human visual detection degradation in the compressed medical images for any arbitrary compression algorithm and compression ratio. Model observers attempt to predict human performance in visual detection tasks [9], [14]–[17]. Specially model observers have been used with computer generated signals embedded in real patient backgrounds [18] to automatically evaluate and optimize JPEG lossy image compression algorithms [9], [10], [19]. It might be argued that the simulated signals are an over simplification of the real complex signals in real clinical images. However, the signal simulation approach assumes that the human observers' visual and cognitive processes mediating the detection of the simpler simulated signals are the same as those mediating the detection of the more complex real signals. This hypothesis has been supported by previous findings. Eigler *et al.* [20] found that image processing improving nonphysician observers' performance on detecting simulated signals embedded in X-ray coronary angiograms led to improved physicians' performance with real lesions.

In this paper, we use the performance of model observer to evaluate the JPEG 2000 encoder options' influence on a signal detection task for X-ray coronary angiograms. In particular, we are interested in determining whether encoder options different from the default ones which maximize peak signal to noise ratio (PSNR) lead to better human and model observer task based performance. Furthermore, we are interested in determining whether rank order of encoder settings based on human/model observer performance is invariant with compression ratios.

Previous studies have investigated model observer performance for tasks in which the signal is known to the observer and does not vary from trial to trial (signal known exactly, SKE [9], [14], [19], [21]). The SKE task has been extensively studied because it allows for easy design of human psychophysical studies and for computationally simple model observer predictions [9], [14], [17], [21]. The SKE task is simple but not representative of the array of possible signals that could be present in medical images in the clinical scenario in which signals vary in size and shape across images. In addition, the physician does not know *a*

*priori* the shape and size of the signal to appear in an image. For this reason we investigate two tasks that attempt to overcome the shortcomings of the SKE task. We first use the signal known exactly but variable task (SKEV) in which the signal varies in shape and size across trials and is known to the observer. The SKEV task overcomes the lack of signal variability in the SKE task. The SKEV task however, is not realistic in that the observer is informed about which particular signal is present on each trial. A more general and clinically realistic task is one in which the signal varies across trials and the observer is not informed about the particular signal presented on each trial. Instead the observer knows that the presented signal will be sampled from a known pool of signals, i.e., the signal is known statistically. The signal known statistically (SKS) task is closer to the real clinical task. The disadvantage is that SKS task is computationally expensive precluding extensive investigation of the parameter space (see Section III-E for the SKS task computation and parameter estimation). Therefore, one important question is whether human and model observer performance in the computationally more tractable SKEV task can be used as a figure of merit to estimate performance in the more clinically realistic SKS task. Prior to the description of the model observers and human psychophysical tasks, we briefly describe the JPEG 2000 standard.

## II. THE JPEG 2000 CODING ENGINE

JPEG 2000 provides a set of encoder options to support the coding functions from preprocessing to coding sections. Each encoder option can be set with different parameters. In the following, an encoder setting denotes a combination of encoder options used in a particular compression session. Currently, two implementations of JPEG 2000 are available. One is a Java implementation by the JJ2000 group [22] consisting of Canon Research France, Ericsson and EPFL (<http://jj2000.epfl.ch>). The other is a C implementation by Image Power and University of British Columbia [23], which is called Jasper (<http://www.ece.uvic.ca/mdadams/jasper>). In this paper we used the latter implementation. Fig. 1 shows the general block diagram of the JPEG 2000 encoder. Next, we describe each encoder step and the related encoder options provided by Jasper.

### A. Preprocessing

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 [24]. The most important feature in this stage is tiling. Tiling partitions the input image into rectangular and nonoverlapping

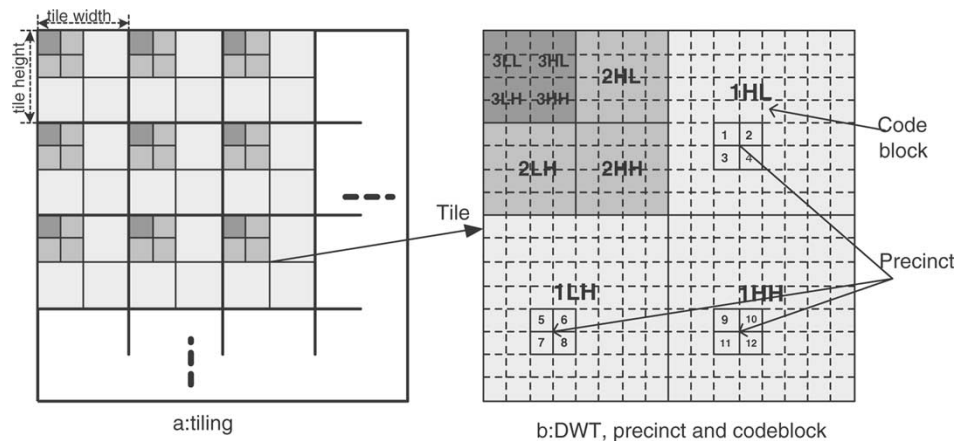


Fig. 2. Left: Tiling, Right: DWT on each tile.

tiles of equal size (except possibly for those tiles at the image borders). Each tile is compressed independently using its own set of specified compression parameters, as though it was entirely distinct image (Fig. 2). Note that all of the following operations, wavelet transform, quantization and coding are performed independently on the tiles.

#### 1) Preprocessing Encoder Options:

a) *Tile size*: The encoding option,  $tilewidth = w$  and  $tileheight = h$ , sets the nominal tile width and height to  $w$  and  $h$  respectively.

b) *Nomct*: Disallows the use of any multicomponent transform. In our case, the image is 8-bit grey level so it is unnecessary to evaluate this option later on.

#### B. The Discrete Wavelet Transform (DWT)

Following the preprocessing stage, DWT is applied.

1) *DWT*: The DWT can be *irreversible* or *reversible*. The *reversible transform* is implemented by means of the Daubechies 5-tap/3-tap filter. The *irreversible transform* is implemented by means of the Daubechies 9-tap/7-tap filter [25]. For the coefficients of the filters, please refer to [11]. The wavelet transform is applied on each tile. Each wavelet transform yields four subbands: horizontally and vertically lowpass (LL), horizontally lowpass and vertically highpass (LH), horizontally highpass and vertically lowpass (HL), horizontally and vertically highpass (HH). Fig. 2(b) shows a 3-level DWT decomposition. DWT decomposition provides a natural solution for the multiresolution requirement.

#### 2) DWT Encoder Options:

a) *mode = m*:  $m$  can be set to “*int*” (default) or “*real*.” The choice of mode determines which multicomponent and wavelet transforms (if any) are employed. In integer (“*int*”) mode, all transforms employed are integer to integer [e.g., reversible color transform (RCT), 5/3 DWT]. In real mode, real to real transforms are employed [e.g., irreversible color transform (ICT), 9/7 DWT].

b) *numrlvls = n*: The argument  $n$  corresponds to the resolution levels which must be an integer greater than or equal to one. The default value is 6.

c) *numgbits = n*: The number of guard bits is set to  $n$ . Guard bits is used to avoid the overflow of the subband value [12].

#### C. Quantization

After DWT transformation, all coefficients are quantized. Scalar quantization is used in Part I of the standard. Quantization is the process by which the coefficients are reduced in precision. This operation is lossy, unless the quantization step is 1 and the coefficients are integers, as produced by the reversible integer 5/3 wavelet. Following is the kernel of the compression algorithm, the coding stage.

#### D. Coding

The coding stage includes two steps: Tier-1 coding also called arithmetic coding and Tier-2 coding also called bit-stream coding. All the complexity of coding is concentrated in Tier-1 which generates embedded block bit-streams. Tier-2 plays a vital role in efficiently representing the individually coded blocks in a full featured bit-stream.

1) *Tier-1 Encoder*: The quantizer indexes for each subband are partitioned into codeblocks. The nominal dimensions of a codeblock are free parameters specified by the encoder but are subject to the following constraints: a) they must be an integer power of two and the total number of coefficients in a codeblock can not exceed 4096; b) the height of the codeblock cannot be less than 4. An embedded code stream is produced for each block independently using bit-plane coding. There are three coding passes: significant pass, refinement pass and cleanup pass per bit plane. Each coding pass is associated with a particular component, resolution level, subband, and code block. The output of tier-1 encoding pass is a collection of coding passes for the various code blocks.

2) *Tier-2 Encoder*: The input of the tier-2 encoding process is the set of bit-plane coding passes generated during tier-1 encoding. Code blocks are grouped into *precincts* as shown in Fig. 2(b). JPEG 2000 organizes the compressed data from the code blocks into units known as *packets* and *layers* during the tier-2 coding step. One packet is generated for each component, resolution level, layer and precinct. More than one ordering of packets in the code stream is supported. Such orderings are called *progressions*. The default one is layer-resolution-component-position. The resulting packets are then output to the final code stream.

TABLE I  
PARAMETER RANGE AND THE DEFAULT VALUES OF SEVERAL ENCODER OPTIONS

	tilesize	mode	numrlvls	codeblock	prcsize
Parameter range	1→image size	int/real	1→ 8*	$h(> 4), w = 2^{n\dagger}$	$h, w = 2^n$
Default	image size	int	6	64	32768

\* upper limit depends on the particular application

†  $n$  is an unsigned integer

### 3) Related Encoder Options:

a) *Codeblock Size*:  $cbkwidth = w$  and  $cbkheight = h$  set the codeblock size. The argument  $w$  and  $h$  must be an integer power of two. The default value is  $64 \times 64$ .

b) *Prcsize*:  $prcwidth = w$  and  $prcheight = h$  set the precinct width and height respectively and  $w$  and  $h$  must be an integer power of two. The default value is 32768.

c) *Miscellaneous Options*: **lazy** mode allows the computational complexity of bit-plane coding to be significantly reduced, by decreasing the number of symbols that must be arithmetically coded. This comes at the cost of coding efficiency [26]; **vcausal** uses vertically stripe causal contexts. This option is aimed at enabling the parallel decoding of the coding passes as well as reducing the external memory utilization [12]; **termall** terminates all coding passes, thus, it facilitates improved error resilience at the expense of decreased coding efficiency; **pterm** uses predictable termination; **resetprob** resets the probability models after each coding pass; **sop** generates SOP marker segments; **eph** generates EPH marker segments, and **segsym** uses segmentation symbols, etc.

Table I summarizes the main encoder parameters and their default settings.

## III. MODEL OBSERVERS

Model observers are proposed on the premise that they can predict the performance of human observers for a given visual detection task. There are numbers of models of human visual detection in noise that generate explicit predictions about the detectability of a signal embedded in a noisy background [17], [21], [27]–[30]. The most commonly used models are the non-prewhitening matched filter observer with an eye filter (NPWE) [31], [32], the Hotelling observer (HO) [14], [17], [30], the channelized Hotelling observer (CHO) [14], [17], [30], [33], [34], and the Laguerre-Gauss Hotelling observer (LGHO) [35], etc.

The general form for all these linear models is the inner (or dot) product between a template and the data at each of the possible signal locations. The model then selects the location associated with the maximum scalar response. For the case where there are many possible signals (SKEV, SKS) the models use different templates for each signal. The models' responses for an SKEV task can be expressed in the matrix formulation [14] as

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

where  $\lambda_{i,j}$  is the scalar response of the model at the  $i$  location using the  $j$ th template.  $\omega_j$  is an  $N \times N$  two-dimensional (2-D)

TABLE II  
EYE FILTER PARAMETERS USED BY PREVIOUS AND CURRENT STUDIES

	c	$\eta$	$\gamma$
Bochud [37]	2	1.9	0.5
Burgess [32]	0.04	1.3	2
Current paper	0.013	1.4	2.6

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 being present on that trial or image. The superscript  $t$  refers to transpose. The vector  $\mathbf{g}_i$  is the data at the  $i$ th location represented in a column vector format.

### A. NPWE Model Observer

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 [31], [32], [36]

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

where  $\tilde{\mathbf{E}}$  is the eye filter in the Fourier domain and  $\tilde{\mathbf{s}}_j$  is the signal in the Fourier domain. The spatial domain template  $\omega_j$  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 that used by Burgess [32]. Also note that the eye filter in this paper departs from a previously used eye filter based on contrast matching experiments [37].<sup>1</sup> This latter eye filter had been successfully used in previous work studying SKE task performance [9]. Table II shows the parameters used previously [32], [37] and in this paper. The one-dimensional (1-D) eye filters with different parameters are shown in Fig. 3.

### B. Hotelling Observer

The Hotelling observer (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

<sup>1</sup>With parameters used in [37], the NPWE model observer results in much lower detection performance than the human observer.

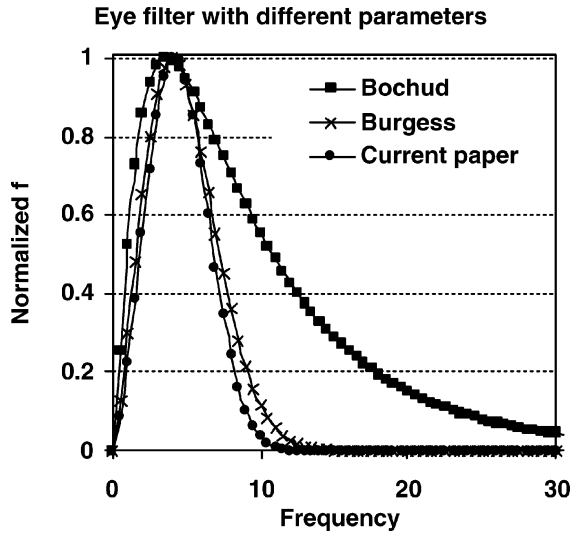


Fig. 3. One-dimensional eye filters with different parameters.

assuming an invertible image covariance matrix, the Hotelling-observer template is given by

$$\omega_{\text{Hot}} = \mathbf{K}_g^{-1}[\langle \mathbf{g}_s \rangle - \langle \mathbf{g}_n \rangle] \quad (4)$$

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 (4). If the image is large then the covariance will require a large number of samples to be inverted. There have been two solutions in the literature [18], [35]. The first solution is to constrain the covariance matrix to a square region around the possible signal locations [18]. In the current study,  $30 \times 30$  pixel regions were selected considering the largest signal size is  $10 \times 25$ . The second approach uses a set of functions to extract some main features from the images, such as Laguerre-Gauss polynomials [18], [35], [38],  $L_n(x) = \sum_{m=0}^n (-1)^m \binom{n}{m} (x^m/m!)$ . The channels are expressed in polar coordinates with origin at the signal location

$$\mathbf{C}_{\text{LG}} = \exp\left(\frac{-\gamma}{2}\right) L_n(\gamma) \quad (5)$$

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 6th 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 (Hotelling observer) 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 [19], [35]

$$\omega_{\text{LGHO}} = \mathbf{W}_c^t \mathbf{C}_{\text{LG}} \quad (6)$$

where  $\mathbf{C}_{\text{LG}}$  is the channel profile defined in (5),  $\mathbf{W}_c$  is a vector containing the optimal linear weights for each of the channels and defined by [15], [33], [39]

$$\mathbf{W}_c = \mathbf{K}_v^{-1}[\langle \mathbf{g}_C/s \rangle - \langle \mathbf{g}_C/b \rangle] \quad (7)$$

where  $\mathbf{K}_v$  is an  $N \times N$  matrix describing the covariance of the output of the channels to the images. For our particular implementation the covariance matrix consists of an  $18 \times 18$  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 were created to estimate  $\langle \mathbf{g}_C/s \rangle$ . Estimation of  $\langle \mathbf{g}_C/b \rangle$  was based on 2700 samples. Thus, a total of  $184 \times 400 + 2700$  test locations were used 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.

### C. 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 [33]. 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 [9], [15], [18]. The Gabor channels are given by

$$C_{\text{CHO}}(x, y) = \exp\left(\frac{-4 \ln 2(x^2 + y^2)}{W_s^2}\right) \cdot \cos[2\pi f_c(x \cos \theta + y \sin \theta) + \beta] \quad (8)$$

where  $f_c$  is the spatial frequency,  $\theta$  is the orientation,  $W_s$  is the width and  $\beta$  is the phase.

In this study we used an 80-channel model with 5 spatial frequencies (central frequencies, 16, 8, 4, 2, and 1 cycle per degree), 8 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 Section III-B.

### D. Calculation of the Performance for Models in SKEV Tasks

The most general way to measure performance in a multiple alternative forced choice (MAFC) task is to compute the probability of correct outcome in each trial [37], [38]. 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

$$O_h = \text{step}(\lambda_{s,j} - \max(\lambda_{n,j})) = \begin{cases} 1, & \text{if } \lambda_{s,j} \geq \max(\lambda_{n,j}) \\ 0, & \text{if } \lambda_{s,j} < \max(\lambda_{n,j}) \end{cases}$$

$$P_c = \frac{1}{H} \sum_{h=1}^H O_h \quad (9)$$

where  $H$  is the total number of trials,  $O_h$  is the outcome in trial  $h$ ,  $\lambda_{s,j}$  is the template response in trial  $h$  to the signal present location, and  $\lambda_{n,j}$  is the template response in trial  $h$  to the noise locations.  $P_c$  is the percent of tasks in which the model correctly identifies the signal location.

### E. Calculation of the Performance for Models in 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. Therefore, SKS tasks require computing a dot product between a template corresponding to each possibly present signal and the data at each location [40], [41] as expressed by (1). 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 [42]. We assume that the internal response 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 locations. The model selects the location with the highest sum of likelihoods ( $l_i$ ). The expression for the sum of likelihood is given by [42]

$$l_i = \sum_{j=1}^J \left[ \frac{1}{\sqrt{2\pi\sigma_j^2}} \right]^M \exp \left( -\frac{(\lambda_{i,j} - \mu_{s,j})^2}{2\sigma_j^2} \right) \cdot \prod_{m=1}^M \left[ \exp \left( -\frac{(\lambda_{m,j} - \mu_{n,j})^2}{2\sigma_j^2} \right) \right]^{(1-\delta_{m,i})} \quad (10)$$

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  $j$  template to the  $j$ th signal absent locations,  $\sigma_j$  is the standard deviation of the response of the  $j$  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  $i$  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

$$O_h = \text{step} (l_s - \max(l_n)) \begin{cases} 1, & \text{if } l_s \geq \max(l_n) \\ 0, & \text{if } l_s < \max(l_n) \end{cases} \\ P_c = \frac{1}{H} \sum_{h=1}^H O_h \quad (11)$$

where  $l_s$  is the sum of likelihood as in (10) at the signal location in trial  $h$  and  $\max(l_n)$  is the maximum of the likelihoods at the signal absent locations.

## IV. GENERATION OF TEST IMAGES AND PERFORMANCE MEASUREMENT

### A. Backgrounds

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  $\mu$ Gy per frame and extended dynamic range enabled video circuitry. The images were digitized with a linear analog ampli-

fication and lookup table to achieve a  $512 \times 512$  pixel matrix with a resolution of 0.3 mm/pixel and 256 gray levels. A total of 541 images 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 of patients getting an X-ray coronary angiogram at the Catherization Laboratory with no regard to the patients' state of disease.

### B. Generation of Simulated Arteries and Signals

The simulation of arteries attempts to mimic the image generation process of X-ray coronary angiograms [43]. This method has been validated using phantom studies [43]. First, we model the attenuation of the X-rays as a function of path length through the contrast filled vessel. Second, we model the imaging system blur (focal spot and image intensifier distortion) as a spatial convolution of the simulated artery with a 2-D isotropic Gaussian point spread function. Finally, we estimate the scattering and veiling glare and correct the image based on the convolution technique of Love and Kruger [44]. For our test images the projected simulated arteries are three-dimensional 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  $\mu$  was set to 0.16/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 axis ranging from 3 to 10 pixels. This resulted in a total of 184 possible signals. The average signal contrast calculated from all of the 900 test images was 0.17, where signal contrast is defined as: (peak signal luminance—background luminance)/ background luminance.

### C. Inserting the Arteries to the Real Background

We generated four simulated arteries projected 32 pixels apart as a group into  $512 \times 512$  pixel backgrounds which were extracted from real patient digital X-ray coronary angiograms. 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. Note that more than one group of arteries may be inserted into one  $512 \times 512$  backgrounds. Then those images were compressed/uncompressed. Following this, we extracted  $256 \times 256$  test images centered on each group of inserted arteries from the  $512 \times 512$  images. The final test set consisted of a total of 900 test images with  $256 \times 256$  pixel size for each compression condition. Fig. 4 shows samples of test images with additional white arrows indicating the signal present location.

### D. Task

For each test image, model and human observers identify the artery containing the signal out of the four arteries. This task is known as four alternative forced choices, 4 AFC.

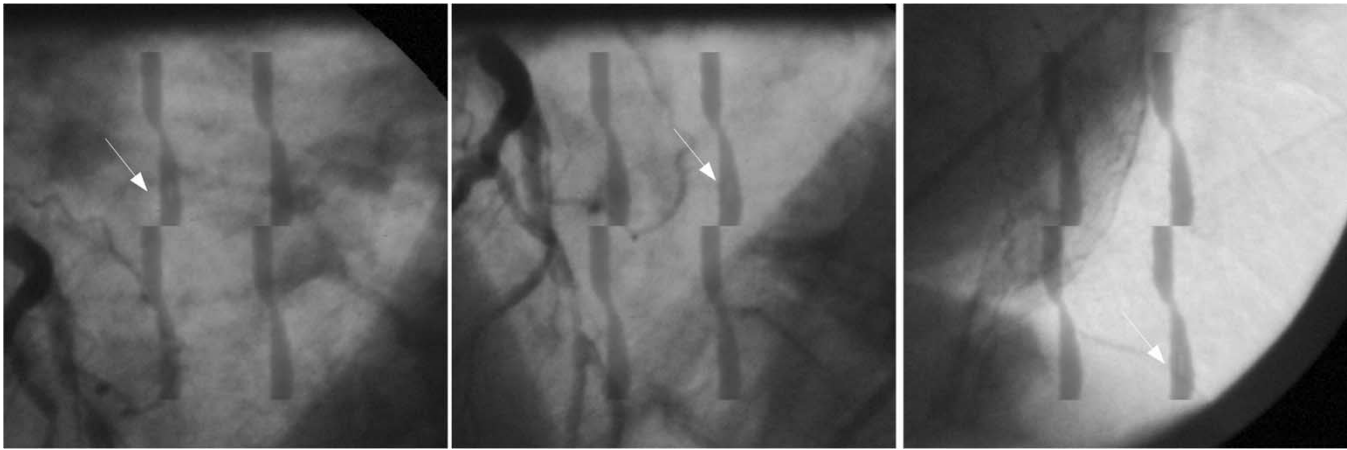


Fig. 4. Samples of test images. White arrows indicate the signal location.

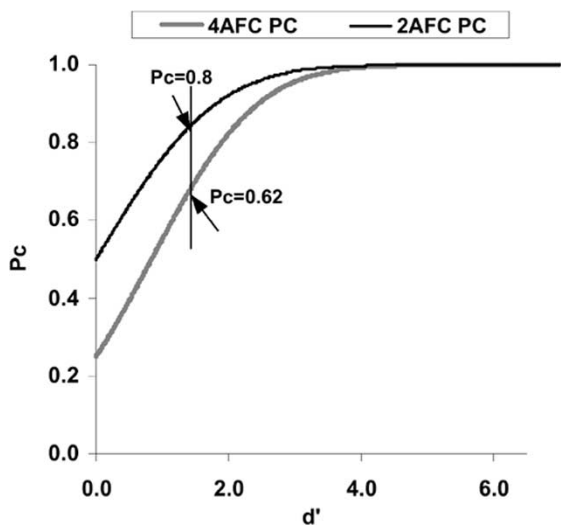


Fig. 5. The relationship between Pc and  $d'$  for 4AFC and 2AFC tasks.

#### E. Pc to $d'$

Percent correct (Pc) can be then converted to an empirically obtained index of detectability ( $d'$ ) by generating a look up table for Pc versus the index of detectability from the following relationship [42]:

$$Pc(d', M) = \int_{-\infty}^{+\infty} \phi(x - d') [\Phi(x)]^{M-1} dx \quad (12)$$

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 advantage of using  $d'$  is that it does not depend on the number of possible locations in the tasks [45], [46]. For example as shown in Fig. 5,  $d' = 1.2$  corresponds to an 80% correct for a 2AFC task. For the same  $d'$ , the percent correct is 62% for a 4AFC task.

### V. MODEL OBSERVER PERFORMANCE EVALUATION

In this section, we evaluated the JPEG 2000 encoder options using model observers. We first compare performance of four model observers (NPWE, HO, CHO, LGHO) in the SKEV task as a function of image compression ratio with three different JPEG 2000 encoder settings. Following, given that we obtained

similar dependence for all model observer performance for an SKEV task, we use the NPWE model observer to extensively investigate the JPEG 2000 encoder option parameter space.

#### A. Performance Comparison of NPWE, HO, CHO and LGHO

Fig. 6 shows performance ( $d'$ ) for the four model observers as a function of 6 compression ratios. Each graph corresponds to a particular JPEG 2000 encoder setting: default (left graph), mode = real (middle graph), and tilewidth = tileheight = 470 (right graph). The results show that although small differences exist, for the present task the NPWE results in similar trends as a function of compression ratio to the other model observers. Based on these and past results [19], the NPWE model observer is selected to extensively investigate the JPEG 2000 encoder options.

#### B. NPWE Performance for JPEG 2000 Encoder Options

We evaluated NPWE performance on the SKEV task for the following JPEG 2000 encoder options: *tilesize*, *mode*, *numrlvls*, *numgbits*, *codeblock size*, *prcsize*, and a set of miscellaneous options.

1) *Tilesize*: In our experiment, we selected the same value for *tilewidth* and *tileheight* and we used *tilesize* to represent that value. Fig. 7(a) shows the NPWE model observer performance for different *tilesize*s at a compression ratio of 15:1. *Tilesize*s begin from 32 and increase by 2 till 512. We found that NPWE performance reaches its local peak values (roughly  $d'$  larger than 1.8) when the tile size is within the ranges of [88, 92], [108, 114], [148, 152], [216, 240], and [436, 484]. Fig. 7(b) is the NPWE performance for the following five compression ratios: 7:1, 10:1, 15:1, 20:1 and 30:1. The *tilesize*s evaluated are 36 sampled (not equally spaced) values between 32 and 512. NPWE performance curves result in similar shape for different compression ratios.

We then define the performance improvement (PI) percentage as

$$PI(\%) = \frac{d'_{\text{new}} - d'_{\text{default}}}{d'_{\text{default}}} \times 100 \quad (13)$$

where  $d'_{\text{new}}$  denotes the detectability obtained using JPEG 2000 encoder parameters other than the default ones and  $d'_{\text{default}}$  denotes detectability using the default encoder setting. Table III

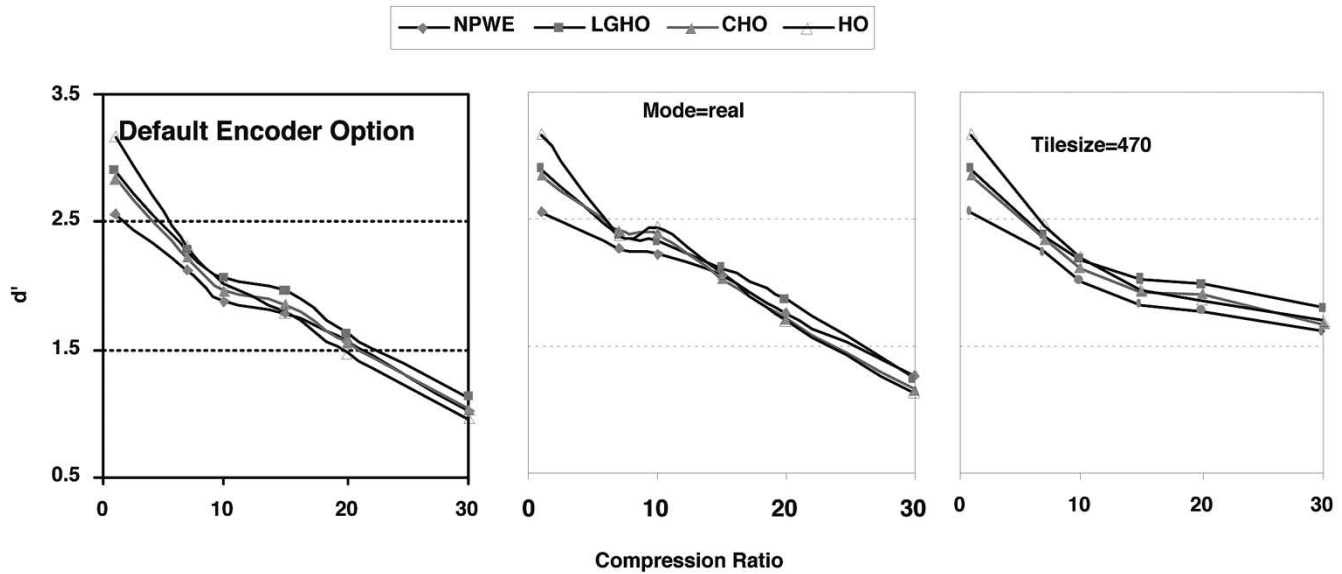


Fig. 6. Performance comparison of NPWE, HO, CHO, LGHO on an SKEV task. From left to right: default; mode = real; tilewidth = 470, and tileheight = 470.

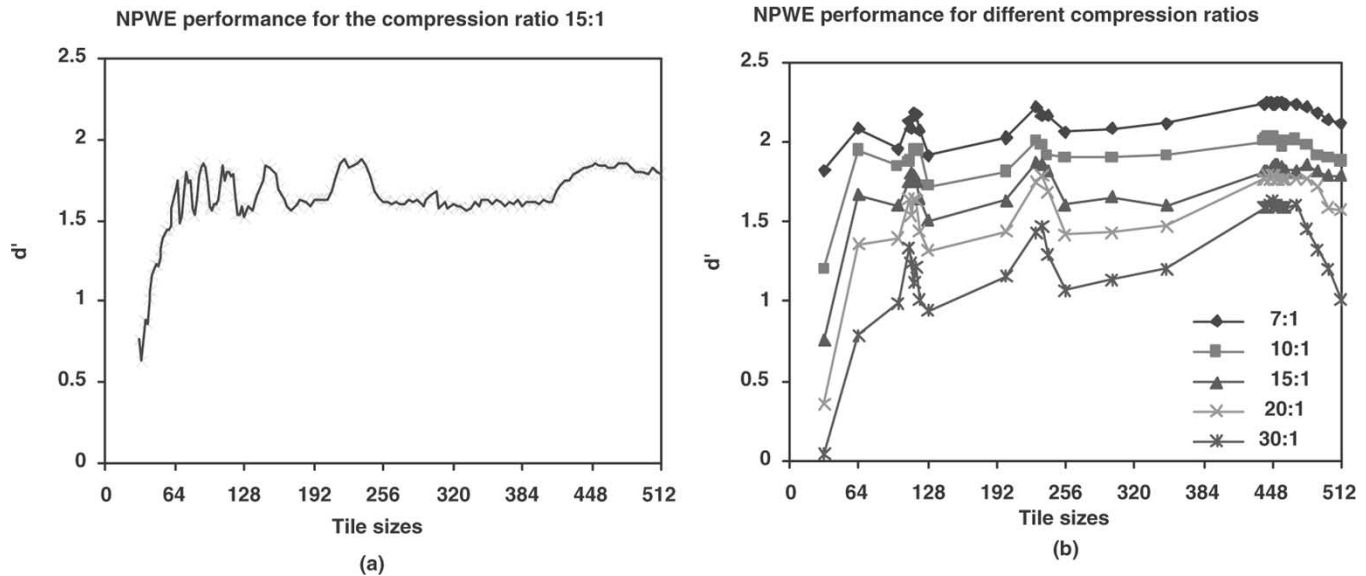


Fig. 7. NPWE model observer performance ( $d'$ ) in an SKEV task as a function of tile sizes: (a) densely sampled tilesizes at compression ratio 15:1 (b): roughly sampled tilesizes for 5 compression ratios: 7:1, 10:1, 15:1, 20:1, and 30:1.

summarizes the performance improvement based on the 36 tile-sizes evaluated. For each compression ratio, the highest performance improvement occurred at different tilesizes which were highlighted in Table III. The maximum NPWE performance for different compression ratios occurs at different tilesizes. The higher the compression ratio is, the worse the performance with small tilesizes (i. e. 32 and 64).

2) *Mode*: Fig. 8 shows the NPWE performance comparison between *int* (default) and *real* mode for the SKEV task. Results demonstrate that NPWE performance for images compressed with the *real* mode consistently exceeds that for images compressed with the *int* mode.<sup>2</sup>

3) *Number of Resolution Levels*: Figs. 9 and 10 show the NPWE performance on the SKEV task for images compressed with different resolution levels for four tilesizes: 128, 256, 470,

<sup>2</sup>However, it should be noted that the *int* mode can also perform lossless compression. Thus, for a particular task, when the compression ratio is around 2:1 to 3:1, integer mode would result in better NPWE performance.

TABLE III  
PERFORMANCE IMPROVEMENT FOR DIFFERENT TILESIZES

Tilesizes	ratio 7:1	ratio 10:1	ratio 15:1	ratio 20:1	ratio 30:1
32	-13.51	-36.36	-57.60	-77.27	-95.77
64	-1.35	3.03	-6.40	-13.64	-22.54
115	3.38	3.79	0.00	4.55	9.86
128	-9.46	-8.33	-15.20	-16.36	-7.04
230	4.73	6.06	<b>4.80</b>	10.91	40.84
235	2.03	5.30	4.00	<b>13.64</b>	45.07
256	-2.70	0.76	-9.60	-10.00	5.63
444	<b>6.76</b>	<b>7.58</b>	2.40	<b>13.64</b>	56.34
448	6.08	<b>7.58</b>	3.20	<b>13.64</b>	<b>60.56</b>
470	6.08	6.82	2.40	12.73	59.15
500	1.35	0.76	0.00	0.91	18.31

512 (from left to right) using the 5/3 DWT (*int* mode) and 9/7 DWT (*real* mode). The results show that the NPWE perfor-

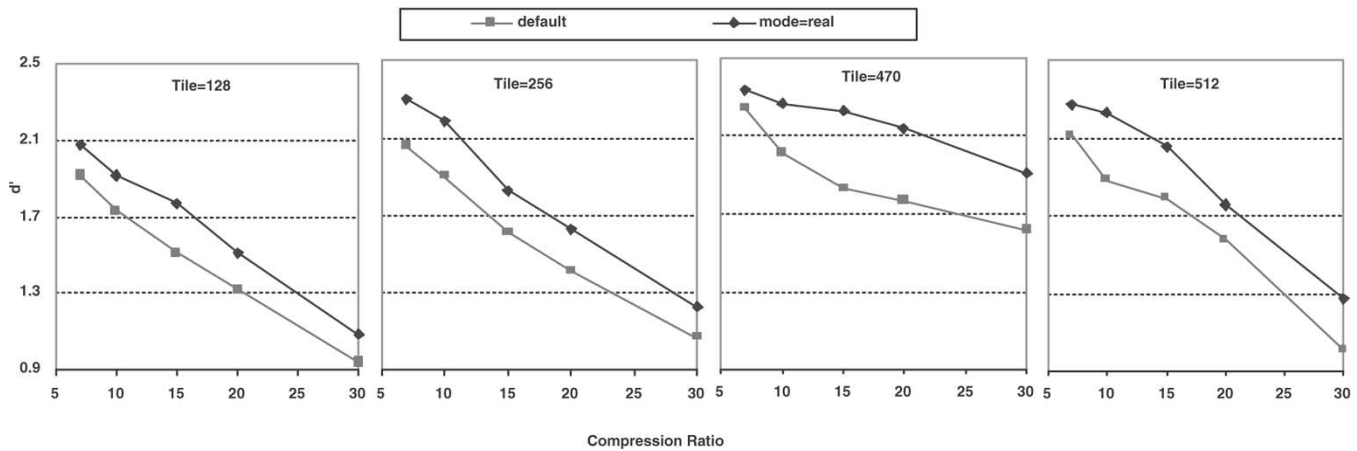


Fig. 8. NPWE model performance ( $d'$ ) for real and integer (default) mode for 4 tile sizes (left to right: 128, 256, 470, and 512).

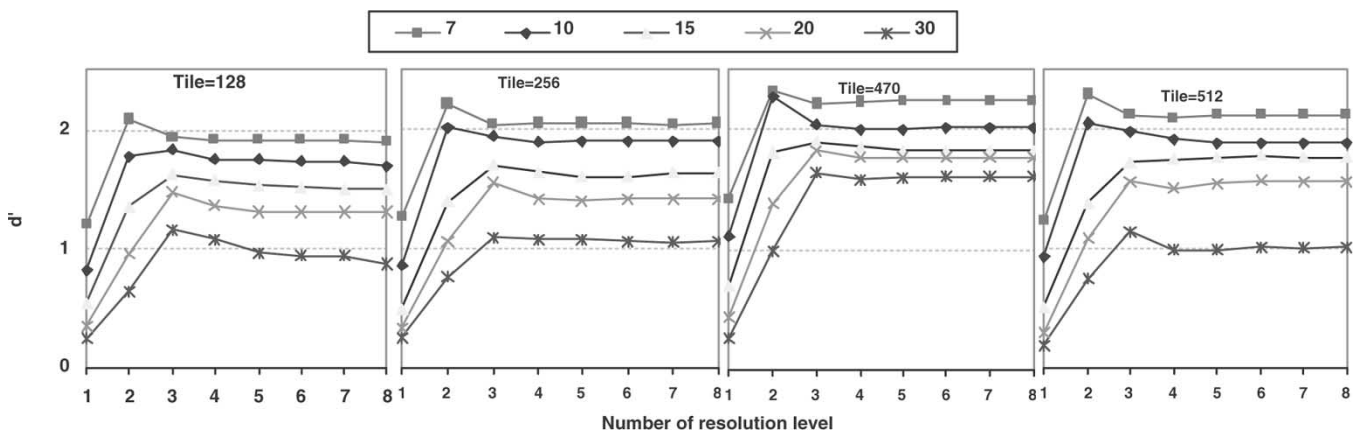


Fig. 9. NPWE model observer performance ( $d'$ ) for encode option *number of resolution levels* with *int* mode for four tile sizes (from left to right: 128, 256, 470, and 512).

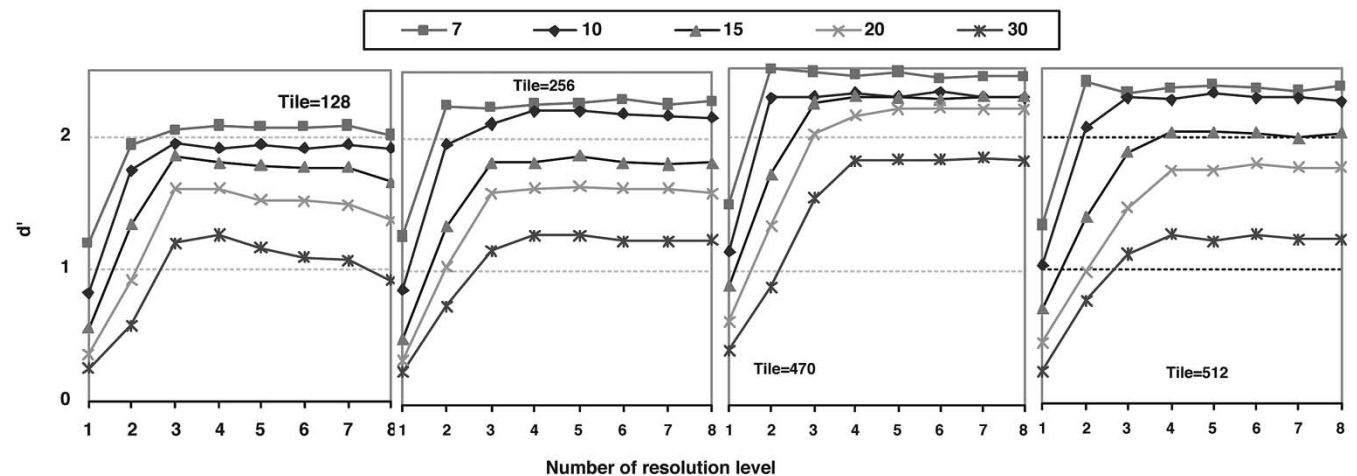


Fig. 10. NPWE model observer performance ( $d'$ ) for encode option *number of resolution levels* with *real* mode for four tile sizes (from left to right: 128, 256, 470, and 512).

mance increases from 1 to 3 resolution levels but then remains approximately constant from 4 to 8 resolution levels for all four tilesizes. An exception was obtained at tilesize 128. In addition, the value of 2 for the number of resolution levels leads to improved NPWE performance at compression ratios of 7:1 and 10:1, and a value of 3 leads to improved NPWE performance at compression ratio 30:1.

4) *Number of Guard Bits*: Fig. 11 shows that the number of guard bits ranging from 1 to 4 does not have much effect on the NPWE model performance in our SKEV task with all four evaluated tile-sizes.

5) *Code-Block Size*: We evaluated the code block sizes of  $128 \times 8$ ,  $8 \times 128$ ,  $128 \times 32$ ,  $32 \times 128$ ,  $64 \times 64$ ,  $64 \times 32$ ,  $32 \times 64$ ,  $32 \times 32$ ,  $16 \times 16$ , and  $8 \times 8$ . Fig. 12 shows that there

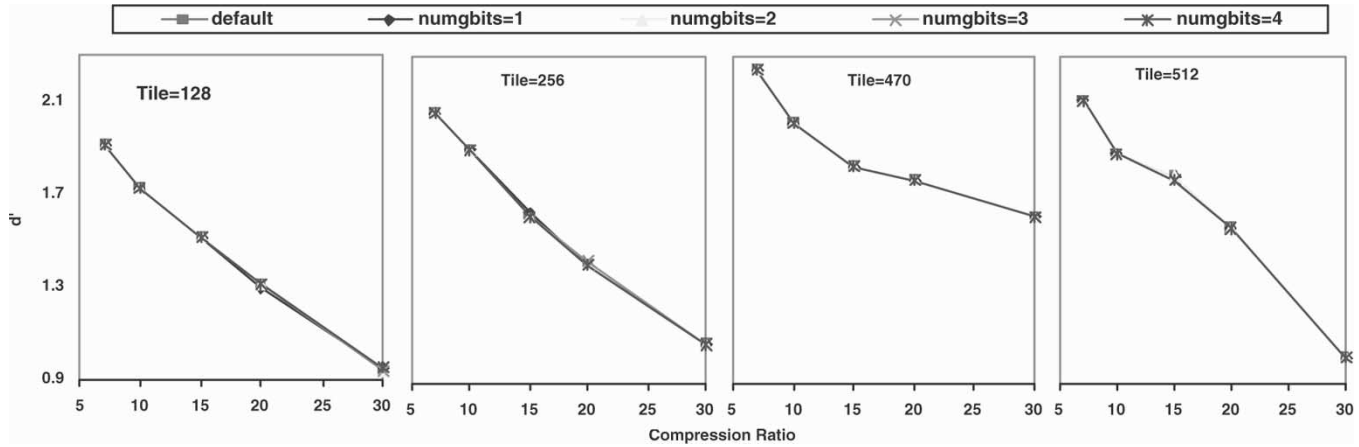


Fig. 11. NPWE model observer performance ( $d'$ ) for encode option *number of guard bits* for four tile sizes (from left to right: 128, 256, 470, and 512). All conditions are superimposed.

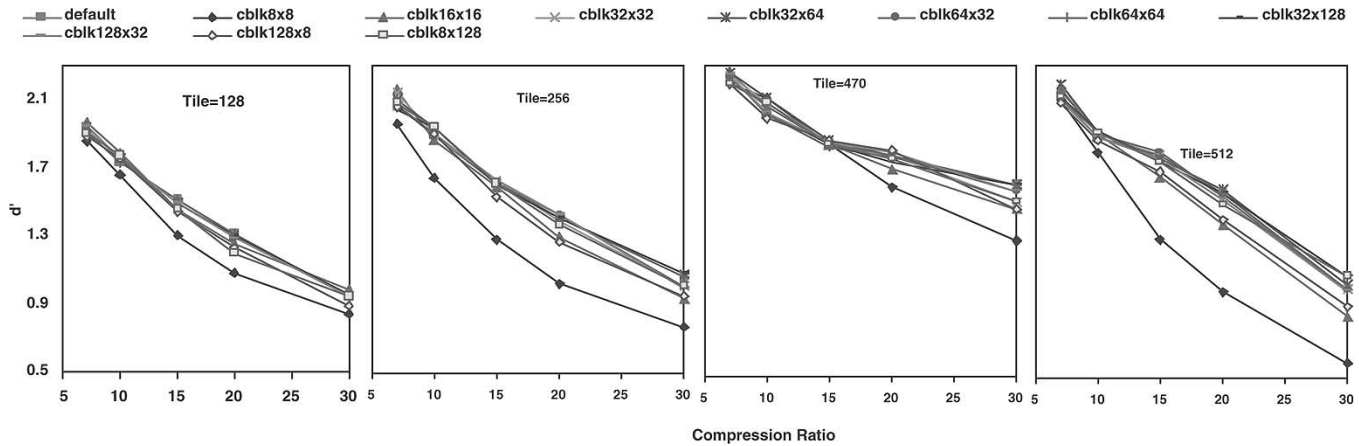


Fig. 12. NPWE model observer performance ( $d'$ ) for the encoder option *Code-block size* for four tile sizes (from left to right: 128, 256, 470, and 512).

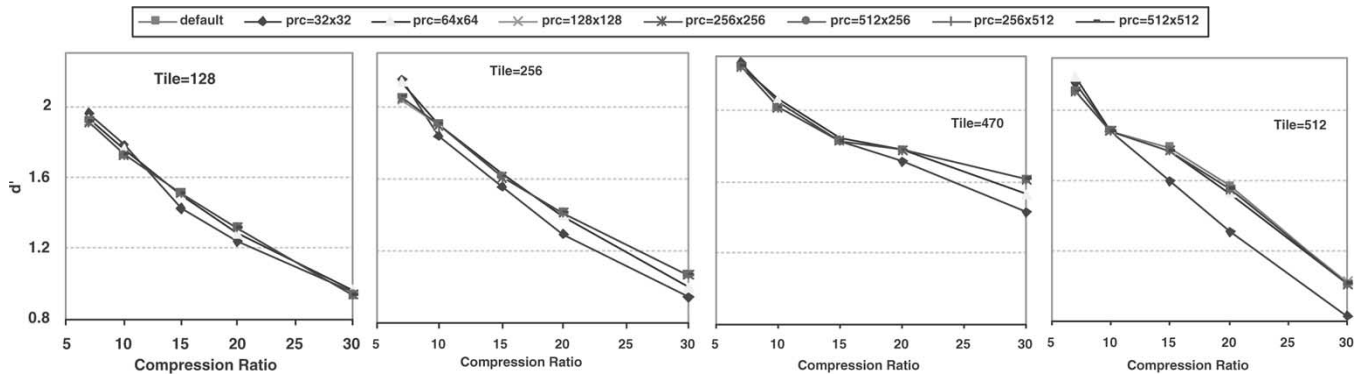


Fig. 13. NPWE model observer performance ( $d'$ ) for the encoder option *precinct size* for four tile sizes (from left to right: 128, 256, 470, and 512).

is little difference in NPWE performance with code block sizes except for the small code block size  $8 \times 8$  which results in significant degradation (bottom curve in Fig. 12).

6) *Prcsize*: Fig. 13 shows that except for the prcsize  $32 \times 32$  there is not much performance difference among the selected prcsizes: default,  $512 \times 256$ ,  $256 \times 512$ ,  $256 \times 256$ ,  $128 \times 128$ , and  $64 \times 64$ .

7) *Miscellaneous Options*: Fig. 14 summarizes the NPWE performance for the following coding options: *lazy*, *vcausal*, *termall*, *pterm*, *resetprob*, *sop*, *eph*, and *segsym* (see [12] for the

details on these parameters). The NPWE model observer performance in the SKEV task for these options is approximately the same as the default one. Thus, users can choose these options freely according to their application requirements.

### C. NPWE Performance: SKEV Versus SKS

In the previous section, we evaluated NPWE model observer performance for different JPEG 2000 encoder options in an SKEV task. One may ask whether similar performance

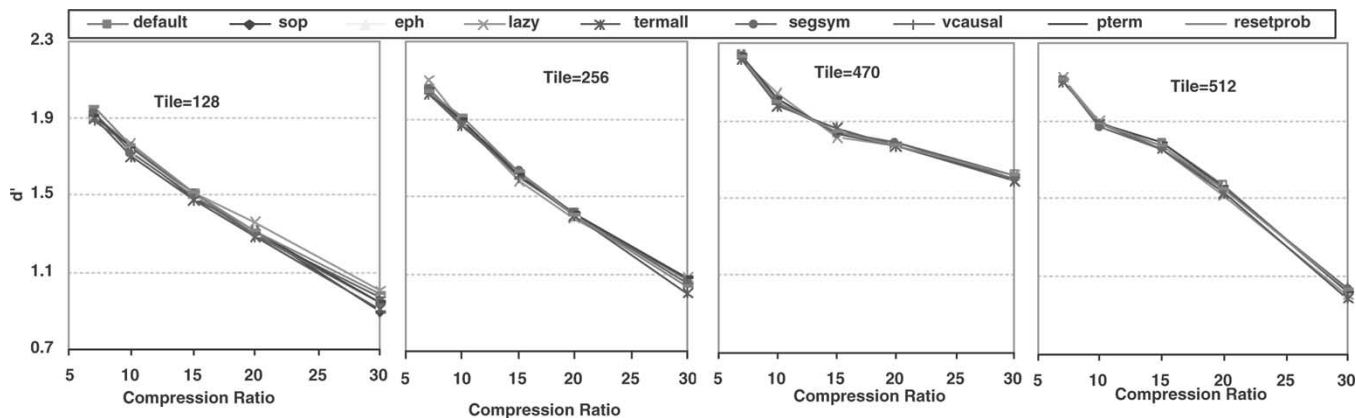


Fig. 14. NPWE model observer performance ( $d'$ ) for miscellaneous encoder options for four tile sizes (from left to right: 128, 256, 470, and 512).

dependence on encoder options would be obtained for the more general SKS task. In this section, we select several encoding options to compare the NPWE performance for both SKEV and SKS tasks.

Fig. 15(a)–(e) shows NPWE performance for the SKEV and SKS tasks with different encoder option manipulations: the number of resolution levels (a), the effect of tile size for the real mode (b), and the int mode (c), the effect of code block sizes for  $\text{tile size} = 128$  (d), and  $\text{tile size} = 470$  (e). NPWE model observer performance follows the same trends for both the SKEV and SKS tasks. The performance in the SKS task is lower than that in the SKEV task. This is expected from the theory of signal detectability given that uncertainty about relevant parameters of the signal to be detected will degrade performance [47]. More importantly, a comparison between the SKEV and SKS model performance shows that even though there are quantitative differences, the performance rank order of encoder option conditions is *always* preserved across tasks. For example,  $\text{tile size} = 470$  results in better NPWE performance than  $\text{tile size} = 128$  for both the SKEV and SKS task. Real mode leads to improved NPWE performance for both SKEV and SKS tasks.

## VI. HUMAN PSYCHOPHYSICAL STUDIES

In the previous sections we have evaluated the effect of different parameters of JPEG 2000 encoder options on a specific model observer performance (NPWE). In this section, human observer performance is obtained for a number of encoder settings.

### A. Psychophysical Experiments

The human observer task was the same as that given to the model observer: to detect the signal at one of four simulated arteries (4 AFC). Images were displayed on an image systems M17LMAX monochrome monitor with a maximum resolution of  $1664 \times 1280$  pixels (Image Systems, Minnetonka, MN). The luminance versus gray level relationship was set according to the DICOM standards and controlled by a DOME board. The maximum luminance was set to be around  $120 \text{ cd/m}^2$  and the minimum luminance was set to be around  $0.00 \text{ cd/m}^2$ . Experiments were conducted in a darkened room. Observers viewed

the images binocularly from a distance of 40 cm and had unlimited time to reach a decision.

Signals varied in size and shape across trials (SKEV, SKS). On each trial in the SKEV task a high contrast 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 no information about the signal presented except observers knew that the presented signal would be one from the pool of 184 signals. When a decision was reached the observer placed the cursor near the location of the signal and clicked the mouse key to indicate their choice for that trial.

Four nonphysician trained observers participated in the experiment. Observers were first trained in 1 session of 100 trials for each experimental condition and then participated in 9 sessions of 100 trials per compression condition. Different conditions consisted of different encoder settings. On each trial an image was randomly sampled from the same 900 test image-set used to compute model observer performance.

We tested two sets of conditions: One set included the uncompressed images (compression ratio 1:1) and the default encoder setting with five compression ratios: 7:1, 10:1, 15:1, 20:1, and 30:1. The second set of conditions included manipulations of the encoder options: code block size, number of resolution levels, mode and tile size. Most of them are under compression ratio 20:1. The order of the conditions was randomized.

Human performance ( $P_c$ ) was determined 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 (12).

### B. Results

1) *Effect of Compression Ratio With Default Encoder Setting for SKEV and SKS Tasks:* Fig. 16 shows the average human and NPWE performance ( $d'$ ) as a function of image compression ratio using the default JPEG 2000 encoder setting in both SKEV and SKS tasks (the performances of four individual human observer, YZ, BP, LF, and KF, are shown in Appendix I). NPWE and human observer performance degraded with increasing compression ratios. Our results show a good agreement of the dependence of performance on compression ratio between humans and the NPWE model

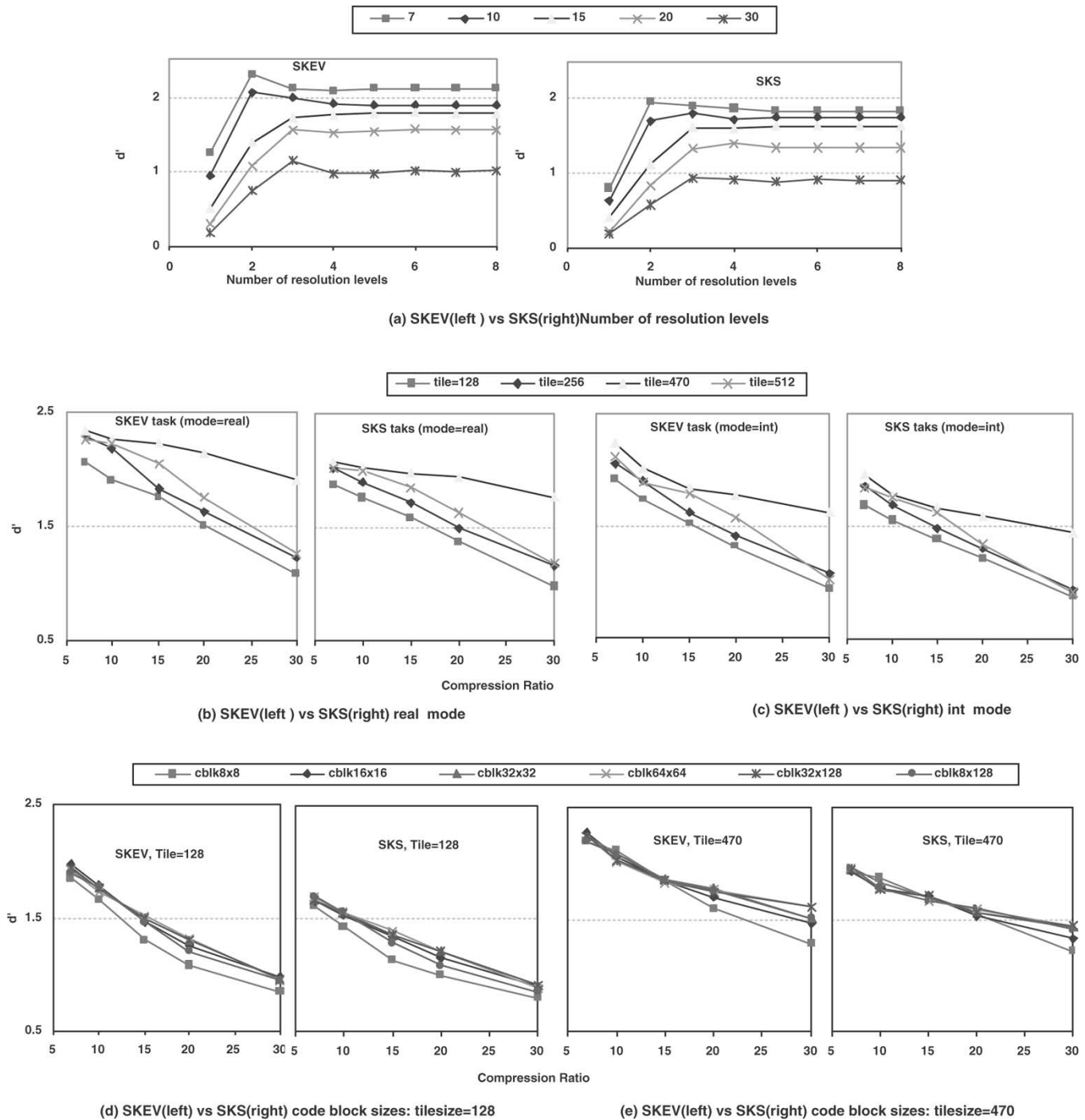


Fig. 15. SKEV versus SKS for the encoder parameter *number of resolution level* (a), *real mode* (b), *int mode* (c), and *code block size* (d, e).

for both the SKEV and SKS tasks. Error bars represent the standard error of the mean.

2) *Multiple Encoder Options:* Fig. 17 shows that performance for all four model observers (NPWE, CHO, LGHO, HO) is similar. Thus, we only show NPWE performance for comparison with the human observer performance in the following figures. Fig. 18 shows the average human (see Appendix II for the results of each of the four individual observers) and NPWE performance as a function of different encoder options: code block size, number of resolution levels, mode and tile size in the SKEV (the first row) and the SKS (the second row) tasks.

a) *Code Block Size:* The first column of Fig. 18 shows the results of different code block sizes. Performance for both the NPWE and human observers is not sensitive to the four different code block sizes evaluated for both the SKEV and SKS tasks.

b) *Number of Resolution Levels:* The second column of Fig. 18 shows the results of varying the number of resolution levels. Average human performance is better for 6 and 8 levels of resolution than that for 2 levels of resolution. In addition, there is no significant difference between human performance for 6 and 8 levels of resolution. The results agree quite well with the NPWE model observer performance.

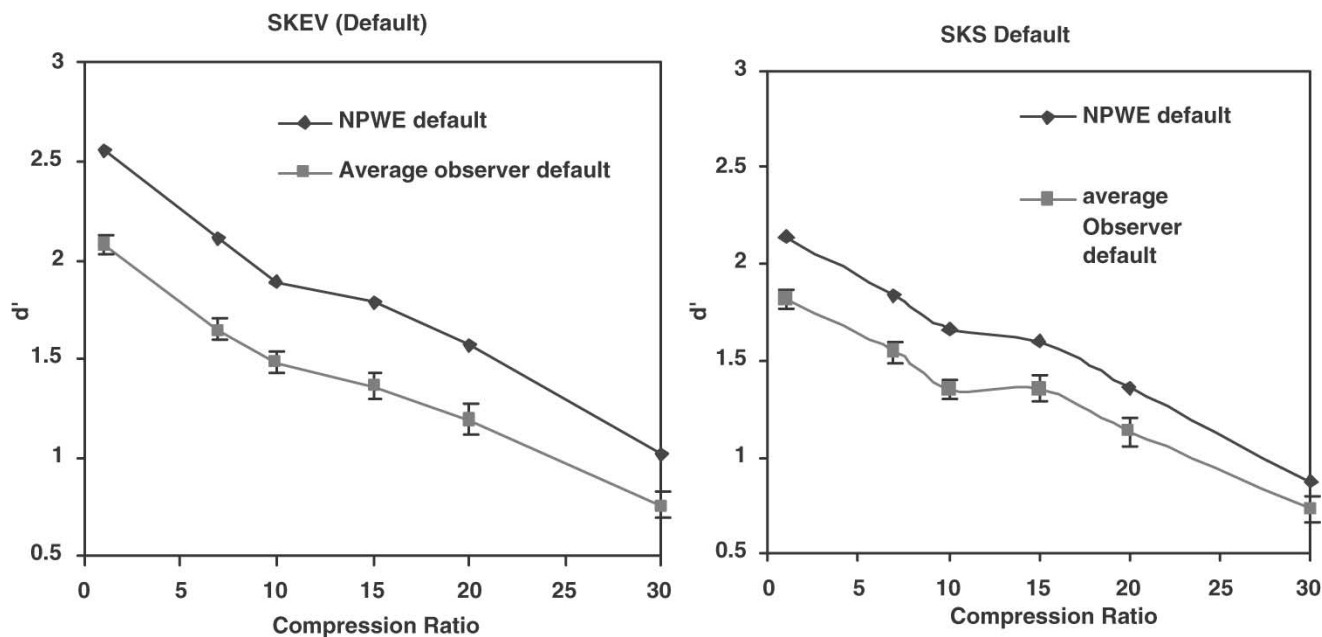


Fig. 16. Human observer versus NPWE model observer performance on default encoder setting for the SKEV(left) and SKS task(right).

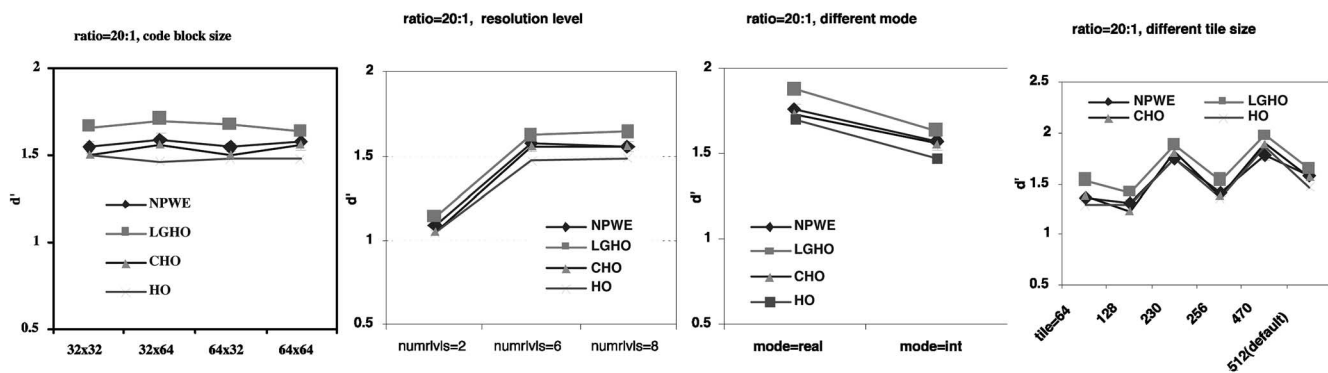


Fig. 17. Model observer performance for SKEV tasks: Column 1: different code block size; Column 2: different resolution level; Column 3: different mode; Column 4: different tile-size.

c) *Mode*: The third column of Fig. 18 shows the results of the different modes. Both average human observer and NPWE performance is higher for images compressed with the *real* mode than the *int* mode.

d) *Tile Size*: The fourth column of Fig. 18 shows human and NPWE performance for different tilesizes. Here, we only evaluated six tilesizes: 64, 128, 230, 256, 470, and 512 (default). Human observer performance presents a similar trend to that of the NPWE model observer with the maximum performance achieved at tile size 470.

Overall, the NPWE model observer gives a good prediction of human performance on both of the SKEV and SKS tasks.

### C. Model Observer Versus Human Observer for SKEV and SKS Tasks

Figs. 19 and 20 show that the difference across tasks for the human observers is smaller than that for the NPWE model. The effect of image compression ratio is similar for both tasks for human and model observers (Fig. 19). In addition, the effect of the different encoder settings for a given compression ratio is very similar for the SKEV and SKS tasks for the human

and model observers (Fig. 20). Our results indicate that for the present set of signals, the computationally simpler SKEV task can be reliably used to evaluate and optimize task based image quality for the more clinically realistic SKS task.

## VII. DISCUSSION

### A. Model Observer Versus Human Observer

Previous studies have shown that model observers (NPWE, CHO, HO, and LGHO) can accurately predict human detection performance as a function of image compression algorithm and/or encoding parameters for the simpler SKE task [9], [10]. In the present study, we compared human and model observer (NPWE, HO, CHO, and LGHO) performance for a number of JPEG 2000 encoder settings for the more clinically realistic SKEV and SKS tasks. Our results show that model observers can accurately predict the effect of different JPEG 2000 encoder settings on human performance for the SKEV task. For the SKS task, we evaluated the NPWE model observer performance and found it to be a good predictor of human performance (Fig. 18). Our results also demonstrated

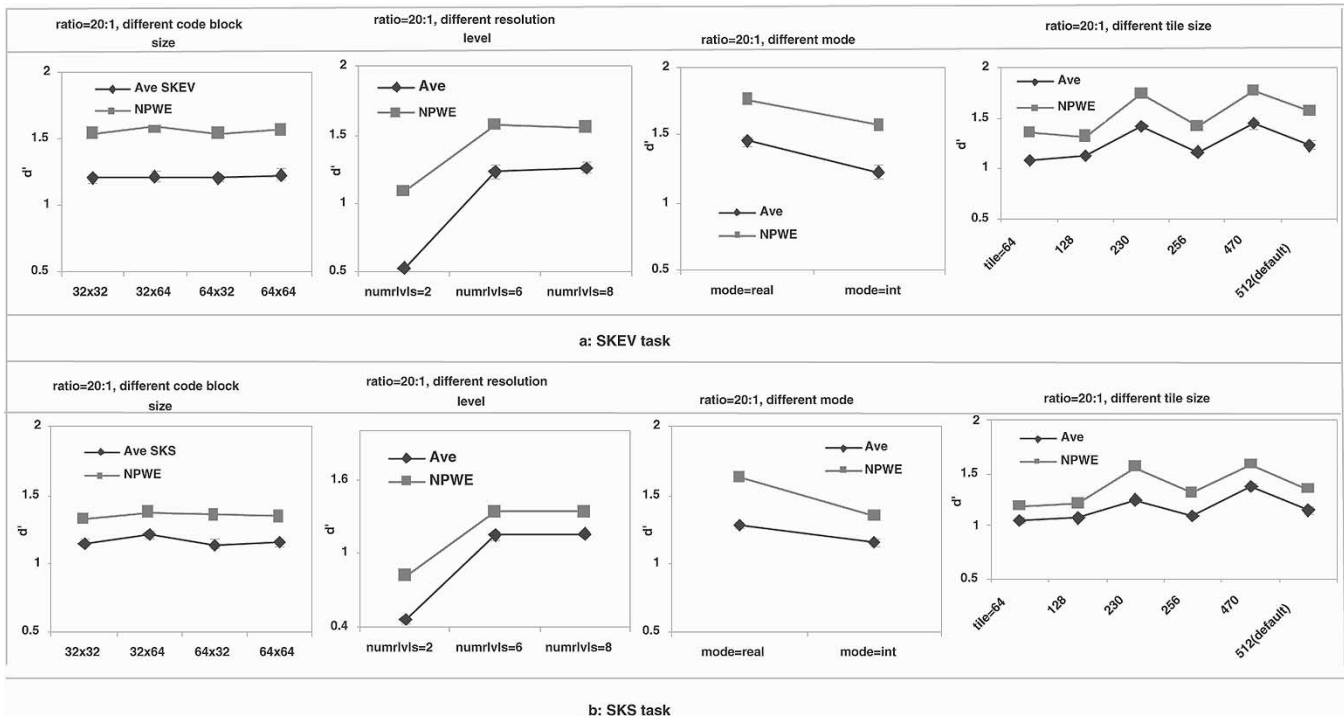


Fig. 18. Average human observer versus NPWE model observer performance for SKEV (the first row) and SKS (the second row) task at compression ratio 20:1. Column 1: different code block size; Column 2: different resolution level; Column 3: different mode; Column 4: different tile-size.

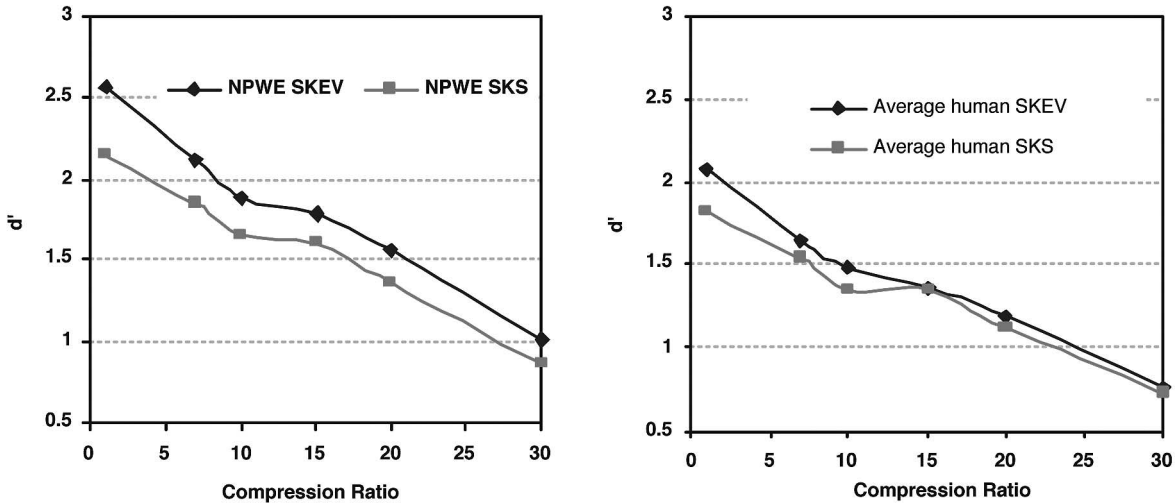


Fig. 19. SKEV versus SKS on NPWE and human observer performance using default encoder setting for five compression ratios.

that all four model observers evaluated showed similar dependencies of performance on JPEG 2000 encoder settings for the SKEV task (Fig. 17). Table IV summarizes the best encoder option values from those evaluated in Section VI-B2 for human and the four model observers in the SKEV task. The results illustrate that the conclusions about JPEG 2000 encoder values for the SKEV task would not change if we used any of the four model observers to evaluate task performance because all four models correctly predicted the rank order of the different encoder options for human performance. The fact that the simpler NPWE model is a good predictor of human performance is of important value for future iterative optimization of JPEG 2000 encoder settings where economy in computational time becomes crucial.

### B. Interaction Between Encoder Option Setting and Compression Ratio

One question of interest when evaluating JPEG 2000 encoder options is whether the values leading to best model observer performance vary across compression ratios. Table V summarizes the JPEG 2000 encoder parameters resulting in the best NPWE performance in the SKEV task at five different compression ratios based on the evaluation results described in Section V.<sup>3</sup> Our results show that the JPEG 2000 encoder parameters resulting in the best NPWE performance vary across compression ratios in the SKEV task. Thus, there might be no single optimized encoder setting that will maximize model observer performance

<sup>3</sup>Note that in Section V during the procedure of evaluating one particular encoder parameter, we set other encoder options to be the default parameters.

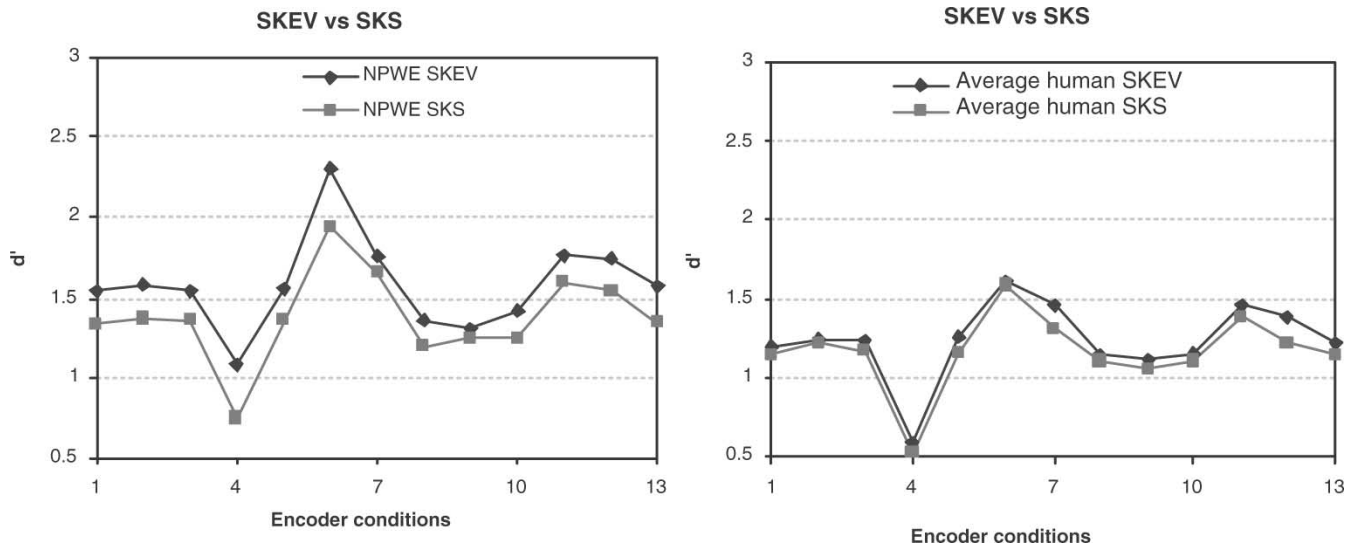


Fig. 20. SKEV versus SKS on NPWE and human observer performance for multiple encoder conditions at a compression ratio of 20:1.

TABLE IV  
MODEL OBSERVER PERFORMANCE COMPARISON ON THE ENCODER OPTIONS EVALUATED BY HUMAN OBSERVER FOR SKEV TASK

	NPWE	Square Window Hotelling	Channelized Hotelling	LG Channelized Hotelling	Human Observer
Codeblock Size ( $32 \times 32$ , $32 \times 64$ , $64 \times 32$ , $64 \times 64$ )	not sensitive	not sensitive	not sensitive	not sensitive	not sensitive
Number of resolution level(2, 6, 8)	$d'_2 < (d'_6 = d'_8)$	$d'_2 < (d'_6 = d'_8)$	$d'_2 < (d'_6 = d'_8)$	$d'_2 < (d'_6 = d'_8)$	$d'_2 < (d'_6 = d'_8)$
mode(real vs. int)	$d'_{real} > d'_{int}$	$d'_{real} > d'_{int}$	$d'_{real} > d'_{int}$	$d'_{real} > d'_{int}$	$d'_{real} > d'_{int}$
Tile size	$d'_{230,470} >$ $d'_{64,128,512}$	$d'_{230,470} >$ $d'_{64,128,512}$	$d'_{230,470} >$ $d'_{64,128,512}$	$d'_{230,470} >$ $d'_{64,128,512}$	$d'_{230,470} >$ $d'_{64,128,512}$

TABLE V  
BEST NPWE PERFORMANCE ENCODER SETTINGS AT FIVE COMPRESSION RATIOS FOR SKEV TASK

Parameter	Compression ratio 7:1	10:1	15:1	20:1	30:1
Codeblock Size	$32 \times 64$	$32 \times 128$	$64 \times 32$	$32 \times 64$	$8 \times 128$
Number of resolution level	2	2	default(6)	default(6)	3
Mode (real vs. int)	real	real	real	real	real
Tile size	444	444/448	230	235/446/448	448
Precinct size	prc= $64 \times 64$	prc= $64 \times 64$	default(32768)	default(32768)	prc= $128 \times 128$

across all compression ratios. These findings suggest that results may not be generalized across compression ratios and that compression ratio specific optimizations are required to achieve the best model observer performance.

C. Task Based Image Quality Metrics

Another important issue is to compare the evaluation results of JPEG 2000 encoder options through task-based metrics (model observer/human observer) with those obtained by using nontask based and more general approaches to evaluate image quality. One common metric to evaluate the image quality of an image that has undergone compression is PSNR. Rabbani

and Joshi [12] investigated the effect of JPEG 2000 encoder options using the PSNR as the metric of image quality for a set of nonmedical images. They reported that image quality (PSNR) decreased with decreasing tile size. This is contrary to our results with the clinically relevant task where we find that model and human observer performance varies across tilesizes. In addition, Rabbani and Joshi found that the  $64 \times 64$  code block size is the best choice while our study did not find major differences on human and model observer performance across code block sizes. With respect to number of resolution levels, their study found that values above 5 to be adequate while our study found NPWE performance to be best at a value of 2 for lower compression

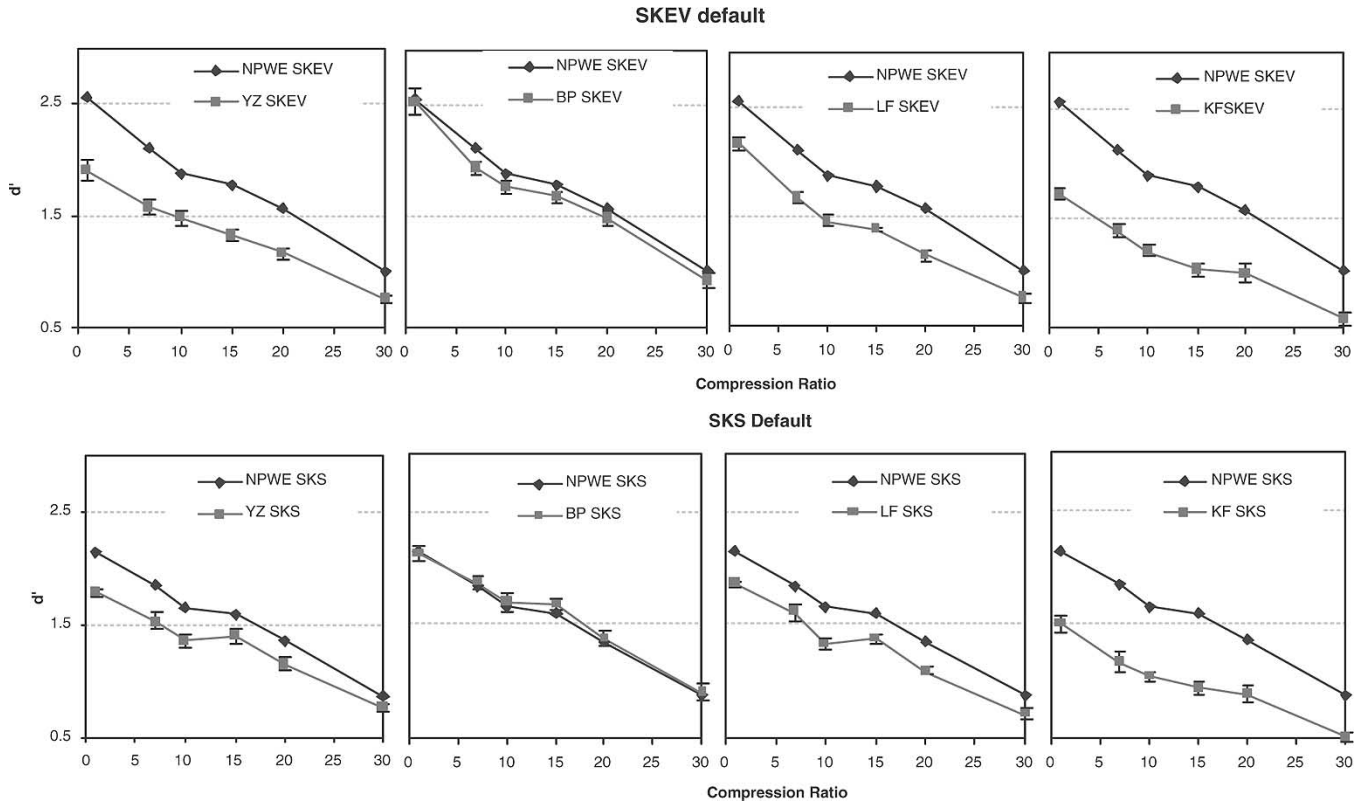


Fig. 21. Human observer (YZ, BP, LF, and KF) versus NPWE model observer performance on default encoder setting for the SKEV and SKS task: First row (from left to right): Performance comparison on SKVE tasks. Second row (from left to right): Performance comparison on SKS tasks.

ratios ( $\leq 10:1$ ) and to be the best at a value of 3 for compression ratio 30:1. On the other hand, both studies found that the real mode (9-tap/7-tap filter) leads to a higher quality than the integer mode (5-tap/3-tap filter). Overall, our findings emphasize that recommendations/evaluations about JPEG 2000 encoder option settings based on nontask based metrics of image quality with nonmedical images (e.g., PSNR) should not be used in guiding the encoder setting selection for medical images where the figure of merit is performance in clinically relevant tasks.

#### D. SKEV Versus SKS Tasks

In this paper, we evaluated model and human performance for both the SKEV and SKS task. The SKS task is clinically more realistic because it does not assume that the human/model observer knows the shape/size of the signal presented in the current trial. The SKEV is less realistic because it assumes that the human/model observer knows the shape and size of the signal presented in the current trial. The advantage of the SKEV task is that it requires model observers that are computationally less time-consuming than those used for the SKS task. Thus, in practice it would be desirable to use the simpler SKEV tasks to perform model observer evaluations and optimizations of medical image quality. However, the SKEV approach is useful if and only if the SKEV task performance is highly correlated with SKS task performance. For example, the rank order of a number of JPEG 2000 encoder settings based on task performance should not change depending on whether an SKEV or an SKS task is chosen. In this respect, our results consistently 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 (Figs. 19 and 20). This is consistent with concept of the effect of stimulus uncertainty on human and model performance [42]. Because humans cannot use precise knowledge of the signal information even when the signal is shown to them (intrinsic uncertainty [47], [48]), the difference across tasks for the human observers is smaller than for the NPWE model.

#### E. Studies on the Compression of X-Ray Coronary Angiograms and Limitations of Present Study

Coronary angiography remains the gold standard for the diagnosis of coronary artery disease [49]. Previous studies have evaluated the effects of lossy data compression algorithm JPEG on the clinical image quality of coronary angiography using both computer simulated [19], [43] and real morphological features and stenosis of arteries in clinical images [6]–[9], [50]. In general, there has been an agreement between the results of both of these approaches. Our present results have evaluated the effect of different JPEG 2000 encoder options on human performance with computer simulated filling defects. One observation is that the degradation of performance with low compression ratios (i.e., 7:1) is greater in the present study evaluating JPEG 2000 than in previous studies evaluating the JPEG algorithm [6], [19], [40].

There has been no clinical study to our knowledge that has evaluated JPEG 2000 compression algorithm or the effect of JPEG 2000 encoder options in diagnostic performance with X-ray coronary angiograms. Given the previous close correspondence between the two approaches (simulated morphological features versus real ones), we suggest that the present results might be used to guide the selection of JPEG 2000 encoder option settings in future clinical studies.

Compression Ratio 20:1

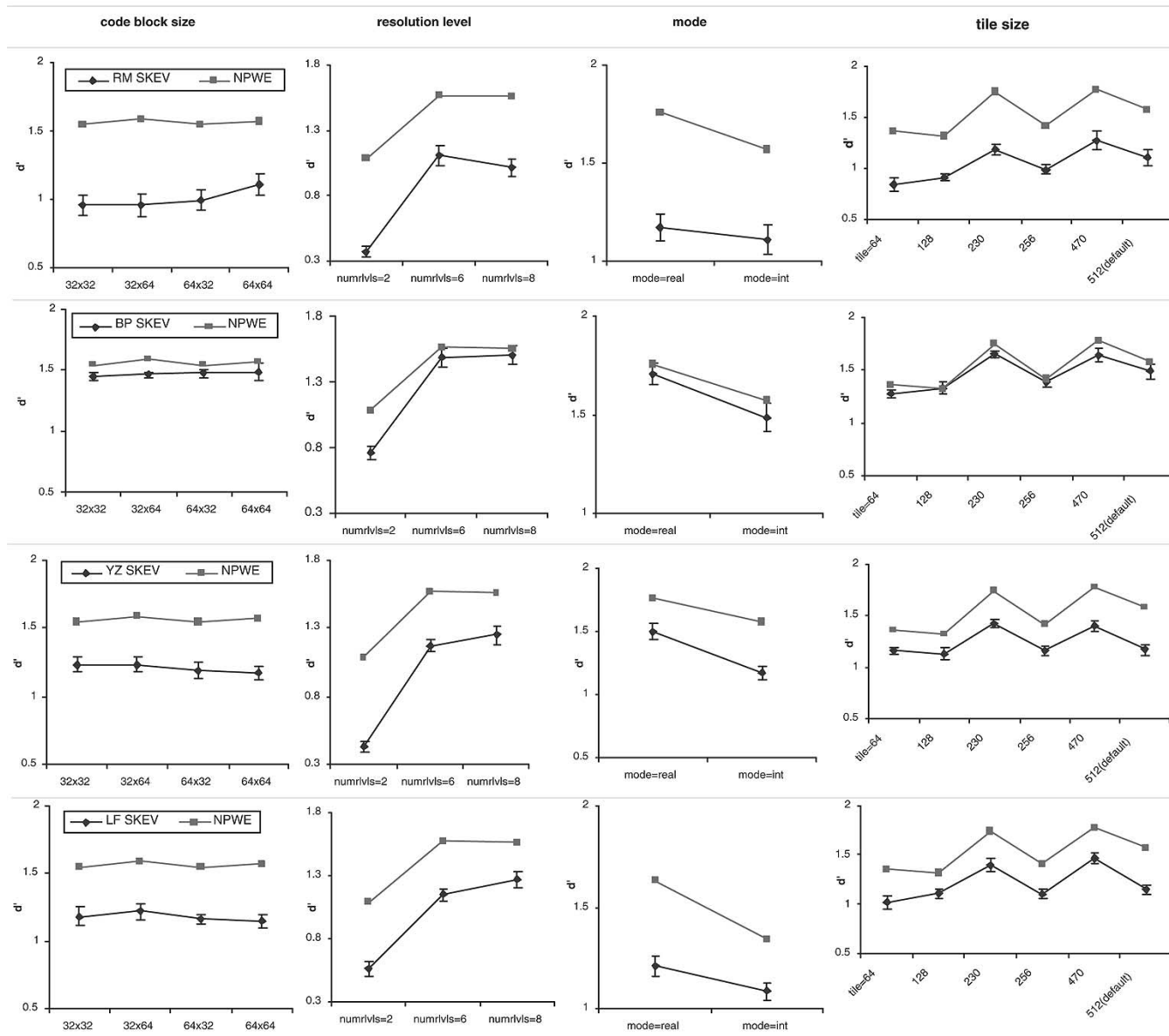


Fig. 22. Human observer versus NPWE model observer performance for an SKEV task: Each row shows the model and individual observer (from top to bottom, RM, BP, YZ, and LF) performance with error bars. Column 1: different code block size; Column 2: different resolution level; Column 3: different mode; Column 4: different tile-size.

There are some limitations of the present study. Although image acquisition with image intensifiers is still the standard of care, many catheterization laboratories across the country are currently replacing their systems by flat panel detectors. The image noise characteristics are different for flat panel detectors because different signal processing is applied to the flat panel images. The present study used images acquired with image intensifiers and it is unknown whether these results will generalize to the newer detectors. Future studies will have to assess this question. In addition, the current study evaluated the detection of a filling defect which is one of many clinically relevant tasks. Other important tasks include detection of dissection and assessment of the stent deployment. In previous studies we have found that although there are differences across tasks, rank order of conditions is preserved across visual tasks [6]. For example, in a previous study [6] performance with JPEG compression led

to better performance than a Crewcode wavelet compression for all four clinically relevant tasks (detection of filling defect, ulceration, lumen and stenosis grading). Based on these results we find likely that our results will generalize to other tasks but future work needs to directly investigate this issue. The present study follows previous work evaluating still image compression of X-ray coronary angiograms. However, given the dynamic nature of angiograms, video compression algorithms [51] need to be evaluated and compared to the still image standards.

VIII. CONCLUSION

We have employed model observers to quantitatively evaluate the effect of JPEG 2000 encoder options on the visual signal detection task consisting of simulated signals embedded in real X-ray coronary angiographic backgrounds. We have

Compression Ratio 20:1

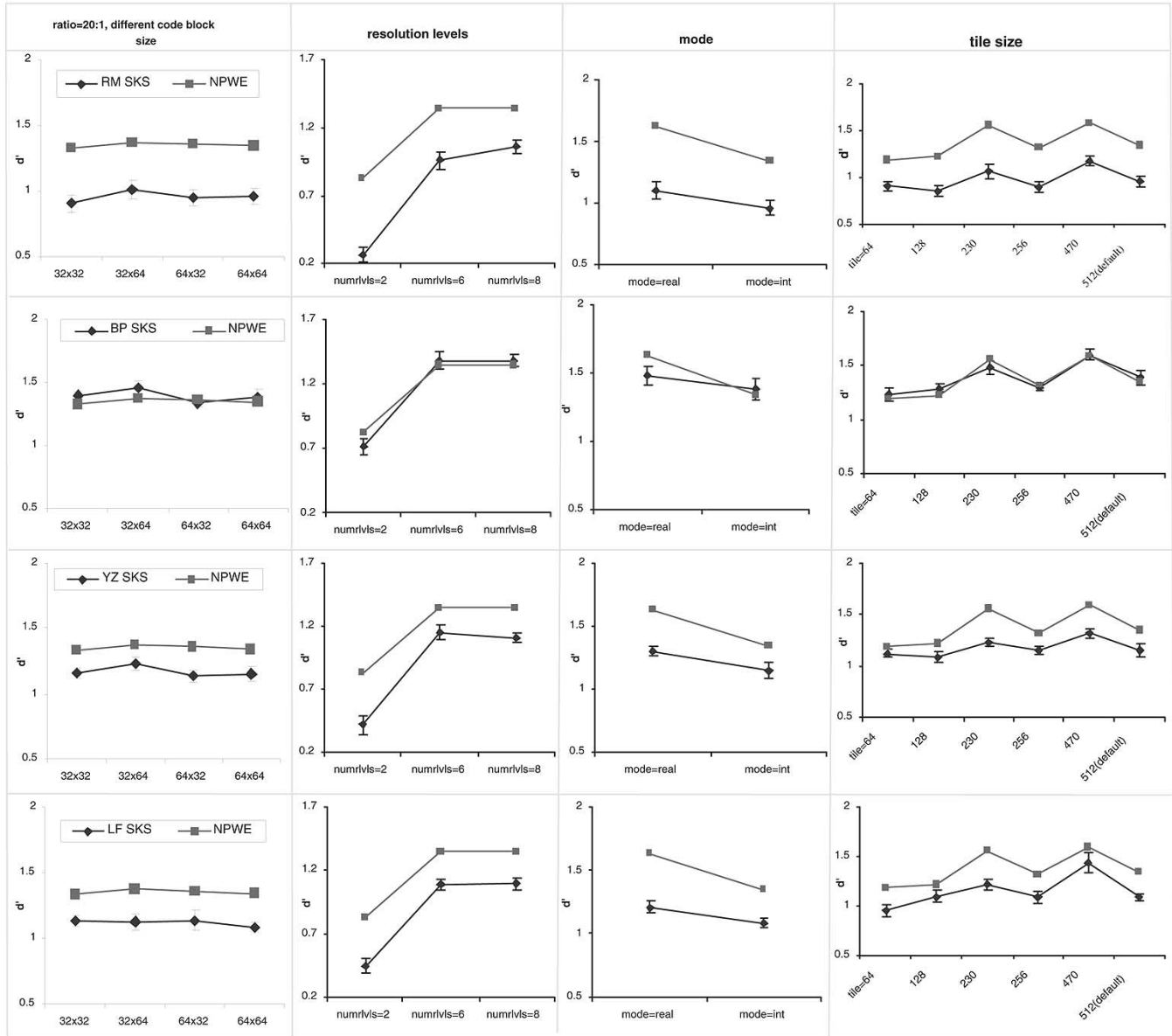


Fig. 23. Human observer versus NPWE model observer performance for an SKS task: Each row shows the model and individual observer (from top to bottom, RM, BP, YZ, and LF) performance with error bars. Column 1: different code block size; Column 2: different resolution level; Column 3: different mode; Column 4: different tile-size.

investigated the clinically more realistic signal known exactly but variable (SKEV) and signal known statistically (SKS) tasks rather than the simpler signal known exactly (SKE) task. We also compared model observer performance with human performance in a subsequent psychophysical study. We found good agreement between human observers and model observer performance and among different model observers (nonprewhitening matched filter with an eye filter, channelized Hotelling, Hotelling, Laguerre-Gauss Hotelling). Our results suggest that the computationally more tractable SKEV task can be used as a reliable predictor on model and human performance for the computationally more complex but clinically more realistic SKS task. In addition, the results suggest that the NPWE model in an SKEV task could be used for future automated optimization of JPEG 2000 encoder options by taking advantage of the NPWE's computation simplicity and

its good prediction capability. Most importantly for practical purposes JPEG 2000 encoder options different from those based on nontask based metrics of image quality (PSNR; JPEG 2000 default setting) may greatly improve the performance and make it possible to obtain similar image quality with higher compression ratio by selecting appropriate encoder options. Together, our work demonstrates how task-based model observers can be used to accurately predict human performance in a signal known statistically task using real anatomic backgrounds.

#### APPENDIX I

##### INDIVIDUAL HUMAN OBSERVER PERFORMANCE UNDER DEFAULT ENCODER SETTING FOR SKEV AND SKS TASKS

Fig. 21 shows the individual human observer: YZ, BP, LF, and KF's performance for images compressed using the default JPEG 2000 encoder setting.

## APPENDIX II

## INDIVIDUAL HUMAN OBSERVER PERFORMANCE UNDER MULTIPLE ENCODER CONDITIONS FOR SKEV AND SKS TASKS

Figs. 22 and 23 show the individual human observer: RM, BP, YZ, and LF's performance for images compressed using multiple JPEG 2000 encoder settings for SKEV and SKS tasks respectively.

## ACKNOWLEDGMENT

The authors would like to thank J. S. Whiting for helpful suggestions. The authors would also like to thank L. Furchgott, K. Friedman, and R. Mancina for participation as observers in the study.

## REFERENCES

- [1] T. Min-Jen, J. D. Villasenor, and B. Ho, "Coronary angiogram video compression," in *Conf. Rec. Nuclear Science Symp. Medical Imaging*, vol. 4, 1995, pp. 1847–1851.
- [2] J. Oh and S. I. Woolley, "Diagnostic quality testing for wavelet-compressed digital angiogram images," in *Proc. Inst. Elect. Eng. IPA'99, Image Processing and Applications Conf*, vol. 465, Manchester, U.K., 1999, pp. 512–516.
- [3] J. Oh, S. I. Woolley, T. N. Arvanitis, and J. N. Townend, "A multi-stage perceptual quality assessment for compressed digital angiogram images," *IEEE Trans. Med. Imag.*, vol. 20, pp. 1352–1361, Dec 2001.
- [4] H. K. Huang, *PACS: Picture Archiving and Communication Systems in Biomedical Imaging*. New York: VCH1, 1996.
- [5] S. Nissen *et al.*, "Cardiac angiography without cine film: Erecting a "Tower of Babel" in the cardiac catheterization laboratory," *J. Amer. Col. Cardiol.*, vol. 24, pp. 834–837, 1994.
- [6] 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. Exp.*, vol. 5, pp. 8–19, July 1999.
- [7] G. Koning, P. Béretta, P. Zwart, E. Hekking, and J. H. Reiber, "Effect of lossy data compression on quantitative coronary measurements," *The Int. J. Cardiac Imag.*, vol. 13, pp. 261–270, Aug. 1997.
- [8] R. Brennecke, U. Bürgel, G. Rippin, F. Post, H.-J. Rupperecht, and J. Meyer, "Comparison of image compression viability for lossy and lossless JPEG and wavelet data reduction in coronary angiography," *The Int. J. Cardiac Imag.*, vol. 17, pp. 1–12, Feb. 2001.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] D. Santa-cruz, R. Grosbois, and T. Ebrahimi, "JPEG 2000 performance evaluation and assessment," *Signal Processing: Image Commun.*, vol. 17, pp. 113–130, 2002.
- [14] H. H. Barrett, J. Yao, J. P. Rolland, and K. J. Myers, "Model observers for assessment of image quality," in *Proc. Nat. Acad. Sci.*, vol. 90, 1993, pp. 9758–9765.
- [15] J. Yao and H. H. Barrett, "Predicting human performance by a channelized Hotelling observer model," in *Proc. SPIE Mathematical Method in Medical Imaging*, vol. 1768, 1992, pp. 161–168.
- [16] 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*, vol. 4, pp. 945–963, May 1987.
- [17] A. E. Burgess, X. Li, and C. K. Abbey, "Visual signal detectability with two noise components: Anomalous masking effects," *J. Opt. Soc. Amer. A*, vol. 14, pp. 2420–2442, 1997.
- [18] M. P. Eckstein and J. S. Whiting, "Lesion detection in structured noise," *Academic Radiol.*, vol. 2, pp. 249–253, 1995.
- [19] 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. Exp.*, vol. 11, pp. 460–475, Mar. 2003.
- [20] N. L. Eigler, M. P. Eckstein, K. Mahrer, and J. S. Whiting, "Improving detection of coronary morphological features from digital angiograms. Effect of stenosis-stabilized display," *Circulation*, vol. 89, pp. 2700–2709, 1994.
- [21] 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.
- [22] JJ2000: An Implementation of JPEG2000 in JAVA™. [Online]. Available: <http://jj2000.epfl.ch>
- [23] M. D. Adams and F. Kossentini, "Jasper: A software-based JPEG-2000 codec implementation," in *Int. Conf. Image Processing*, Sept. 2000.
- [24] 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.
- [25] 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.
- [26] M. D. Adms, H. Man, F. Kossentini, and T. Ebrahimi. JPEG 2000: The Next Generation Still Image Compression Standard. [Online]. Available: [http://www.imagepower.com/technology/research/JPEG2000\\_Tutorial.pdf](http://www.imagepower.com/technology/research/JPEG2000_Tutorial.pdf)
- [27] A. E. Burgess and H. Ghandeharian, "Visual signal detection. II. Signal location identification," *J. Opt. Soc. Amer. A*, vol. 1, pp. 906–910, 1984.
- [28] 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*, vol. 2, pp. 1752–1759, 1985.
- [29] A. E. Burgess, "Visual signal detection with two-component noise: Low-pass spectrum effects," *J. Opt. Soc. Amer. A*, vol. 16, pp. 694–704, 1996.
- [30] J. P. Rolland and H. H. Barrett, "Effect of random background inhomogeneity on observer detection performance," *J. Opt. Soc. Amer. A*, vol. 9, pp. 649–658, 1992.
- [31] 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.
- [32] A. E. Burgess, "Statistically defined backgrounds: Performance of a modified nonprewhitening matched filter model," *J. Opt. Soc. Amer. A*, vol. 11, pp. 1237–1242, 1994.
- [33] K. J. Myers and H. H. Barrett, "Addition of a channel mechanism to the ideal-observer model," *J. Opt. Soc. Amer. A*, vol. 4, pp. 2447–2457, 1987.
- [34] 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*, vol. 18, pp. 473–488, 2001.
- [35] 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.
- [36] 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.
- [37] 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*, vol. 17, pp. 193–205, 2000.
- [38] 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*, vol. 17, pp. 206–217, 2000.
- [39] ———, "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.
- [40] 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.
- [41] 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.

- [42] 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.
- [43] 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.
- [44] L. A. Love and R. A. Kruger, "Scattering estimation for digital radiographic system using convolution filtering," *Med. Phys.*, vol. 14, pp. 178–185, 1987.
- [45] A. E. Burgess and H. Ghandeharian, "Visual signal detection. I. Ability to use phase information," *J. Opt. Soc. Amer. A*, vol. 1, pp. 900–905, 1984.
- [46] M. P. Eckstein and J. S. Whiting, "Visual signal detection in structured backgrounds I: Effect of number of possible spatial locations and signal contrast," *J. Opt. Soc. Amer. A*, vol. 13, pp. 1777–1787, 1996.
- [47] D. G. Pelli, "Uncertainty explains many aspects of visual contrast detection and discrimination," *J. Opt. Soc. Amer. A*, vol. 2, pp. 1508–1532, 1985.
- [48] W. P. J. Tanner, "Physiological implications of psychophysical data," *Ann. New York Acad. Sci.*, vol. 89, pp. 752–765, 1961.
- [49] S. Silber, R. Dörr, G. Zindler, H. Mühling, and T. Diebel, "Impact of various compression rates on interpretation of digital coronary angiograms," *Int. J. Cardiol.*, vol. 60, pp. 195–200, June 1997.
- [50] 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.
- [51] D. Gibson, M. Spann, and S. I. Woolley, "Diagnostically lossless 3d wavelet compression for digital angiogram video," in *Proc. IEEE Data Compression Conf.*, 2002.