Rationale and Objective. The set partitioning in hierarchical trees (SPIHT) wavelet image compression algorithm with the human visual system (HVS) quantization matrix was investigated using x-ray coronary angiograms. We tested whether the HVS quantization matrix for the SPIHT wavelet compression improved computer model/human observer performance in a detection task with variable signals compared to performance with the default quantization matrix. We also tested the hypothesis of whether evaluating the rank order of the two quantization matrices (HVS versus default) based on performance of computer model observers in a signal known exactly but variable task (SKEV) generalized to model/human performance in the more clinically realistic signal known statistically task (SKS).

Materials and Methods. Nine hundred test images were created using real x-ray coronary angiograms as backgrounds and simulated arteries with filling defects (signals). The task for the model and human observer was to detect which one of the four computer simulated arterial segments contained the signal, four alternative-forced-choice (4 AFC). We obtained performance for four model observers (nonprewhitening matched filter with an eye filter, Hotelling, Channelized Hotelling, and Laguerre Gauss Hotelling model observers) for both the SKEV and SKS tasks with images compressed with and without the HVS quantization matrix. A psychophysical study measured performance from three human observers for the same conditions and tasks as the model observers.

Results. Performance for all four model observers improved with the use of the HVS quantization scheme. Improvements ranged from 5% (at compression ratio 7:1) to 50% (at compression ratio 30:1) for both the SKEV and SKS tasks. Human observer performance improvement averaged across observers ranged from 6% (at compression ratio 7:1) to 35% (at compression ratio 30:1) for the SKEV task and from 2% (at compression ratio 7:1) to 38% (at compression ratio 30:1) for the SKS task. Addition of internal noise to the model observers allowed for good prediction of human performance.

Conclusions. Use of the HVS quantization scheme in the SPIHT wavelet compression led to improved model and human observer performance in clinically relevant detection tasks in x-ray coronary angiograms. Model observer performance can be reliably used to predict the human observer performance for the studied tasks as a function of SPIHT wavelet image compression. Our results further confirmed that model observer performance in the computationally more tractable SKEV task can be potentially used as a figure of merit for the more clinically realistic SKS task with real anatomic backgrounds.

Key Words. Signal detection task; model observer; medical image compression; human visual system.

© AUR, 2005
procedures per 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, whereas adult cine angiograms must be stored for 7 years after the procedure (2). Image compression techniques offer an efficient and cost-effective means to reduce the cost of storage and increase the speed of transmission. There are two types of compression methods: lossless and lossy. Lossless compression can typically provide a compression ratio of 2:1 to 3:1, and furnishes a reconstructed image that is identical to the original image; however, this kind of compression still cannot satisfy the requirement for transmission (4). Lossy compression allows higher compression ratios at the cost of a reconstructed image that differs from the original one. These image differences depend on the compression ratio used and on the adequacy of the compression algorithm. The application of optimized lossy compression methods would be acceptable if they could enable perceptually high quality image reconstruction for specific tasks like detection or recognition of disease (5). Researchers have investigated the effect of lossy compression algorithms including the first still image compression standard Joint Photographic Experts Group (JPEG), the new standard JPEG 2000, and other wavelet based algorithms for x-ray coronary angiograms (2,5–8).

In this article, we will evaluate the set partitioning in hierarchical trees (SPIHT) wavelet compression algorithm particularly using a human visual system (HVS)-based quantization matrix. The wavelet transform has emerged as a powerful mathematical tool in many years of science and engineering specially for image compression (9,10) because of its flexibility in representing images and its ability to take into account human visual system characteristics. The discrete wavelet transform (DWT) is a member of this family that operates on discrete sequences and has proven to be an effective tool for image compression (11). Among the wavelet compression, the SPIHT technique is an award-winning method and has received widespread attention since its introduction by Said and Pearlman (12). The method provides the following features: progressive image transmission, fully embedded coded file, simple quantization algorithm, fast coding/decoding, completely adaptive, exact bit rate coding, and error protection. SPIHT employs the following three key concepts: exploitation of self-similarity of the image wavelet transform by using a tree-based organization of the coefficients; partial ordering by magnitude of the transformed coefficients with a set partitioning sorting algorithm; and ordered bit plane transmission of refinement bits for the coefficient values. This results in a compressed bitstream in which the most important coefficients are transmitted first, the values of all coefficients are progressively refined, and the relationship between the coefficients representing the same location at different scales is fully exploited for compression efficiency. In this study, we used the SPIHT wavelet implementation provided by researchers from the Mayo Clinic and Foundation (13). In wavelet-based image compression procedures, the image is subjected to a two-dimensional DWT whose coefficients are then quantized and entropy coded. Typically, a uniform quantizer is implemented by dividing by a factor and rounding to the nearest integer. The wavelet encoder provided by the Mayo Clinic and Foundation (13) provides an optional HVS for the user to choose a quantization matrix that is in theory specific to the human visual system. With the HVS rule, the quantization matrix is adjusted according to wavelet subbands and orientations to improve the perceptual quality of compressed images. The human visual system adaptation multiplies the DWT coefficients in certain wavelet subbands by the values in the HVS quantization matrix to increase the relative importance of coefficients at different frequency scales and orientations. Here, we investigated whether the HVS quantization matrix leads to improved human and model observer detection performance in a task detecting variable simulated signals in x-ray coronary angiograms.

Model observers are task-based algorithms that attempt to predict human observer performance in clinically relevant visual tasks such as detection or classification (14–17). Previous research has used model observers to predict human visual detection performance in a variety of computer generated noise (eg, white, band pass, low-pass, lumpy backgrounds) (18–23) and real anatomic backgrounds (eg, x-ray coronary angiograms, mammograms) (7,24,25). The majority of previous studies using model observers involve tasks in which the signal never varies in shape or size and the observer knows a priori the size and shape of the signal (ie, signal-known-exactly tasks, SKE) (18–20,23). In the clinical practice, the signals to be visually detected vary in shape and size across images. Thus a task in which signals vary in size and shape and observers know which signal will be present has been proposed to overcome the signal size/shape specificity in the SKE task. This task is typically referred to as the signal known exactly but variable task (SKEV). The SKEV task is computationally tractable as the SKE task but is arguably closer to the clinical scenario than the SKE task.
However, in the clinical setting, physicians do not know a priori which particular signal will appear. Therefore, a more clinically realistic task is the signal known statistically task (SKS) in which signals vary in size and shape and the observers do not know a priori which of the signals will be present. One shortcoming of the SKS task is that it requires computationally more complex model observers given that one of many signals could appear on each image. The additional computational complexity makes optimization based on model performance in SKS tasks time consuming. It would be useful if optimizing performance based on the simpler SKEV task led to optimized performance in the SKS task.

In this article, we assessed whether evaluation of model/human observer performance in the simpler SKEV task led to the same conclusions about the rank order of compression encoder parameters in model and human observer performance as in an SKS task. In this study, we first computed model observer performance for an SKEV task. We then obtained model observer performance for an SKS task for comparison. Finally, a human psychophysical study was conducted to verify whether the model observer predicted human performance and whether the HVS quantization scheme resulted in improved human detection performance for both the SKEV and SKS tasks.

MATERIALS AND METHODS

Image Preparation

Test-images used in this study consisted of simulated filling defects (signals) and arteries embedded in real x-ray coronary angiographic backgrounds.

The backgrounds were clinical digital coronary angiograms acquired at 30 frames/second with a 7-inch image intensifier field size (Advantx/DXC, General Electric Medical Systems Milwaukee, WI). These 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. A total of 541 images extracted from 50 different image sequences of 17 different patients were used as the backgrounds.

The projected simulated arteries were obtained by tracing the contours of a real artery, and generating three-dimensional right circular cylinders with varying diameters. The filling defects (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 with the same contrast and different energies. For further details about the algorithm to create the computer-simulated arteries and signals please refer to previous treatments (26).

Four simulated arteries were generated and projected 32 pixels apart as a group into 512 × 512 pixel images. A signal was randomly inserted into one of the arteries. Then we compressed/uncompressed these images and extracted 256 × 256 images centered around each group of inserted arteries from the 512×512 images. The final test set consisted of a total of 900 256 × 256 pixel images with different backgrounds extracted from the patient’s images.

Model Observers

Linear model observers have a common procedure for the SKEV task. On each trial, a template corresponding to the signal present in the image is applied to each of the possible signal locations. For each trial, the model chooses the location with the largest scalar response (14). The response is expressed as

$$\lambda_{ij} = \sum_{n=1}^{N^2} w_{nj}g_{nj} = w^t g_j$$  \hspace{1cm} (1)$$

where $\lambda_{ij}$ is the scalar response of the model at the $i^{th}$ location, $w^t$ is the template, $g_j$ is the data vector, the subscript $j$ on $w$ refers to the $j^{th}$ template, the superscript $t$ refers to transpose, the subscript $i$ for $g$ refers to the $i^{th}$ location, and the subscript $n$ represents the element of the template($w$) and the data vector ($g$).

Four different model observers were studied: 1) non-prewhitening matched filter with an eye filter (NPWE). For the NPWE model used in this article, the template corresponding to a signal was obtained by filtering templates matching the signal with the square of the contrast sensitivity function. This can be achieved by multiplying in the Fourier domain the signal and the eye filter (27,28); 2) the Hotelling observer (HO) derived a template that took into account knowledge about not only the signal profile but also the background statistics. The template was calculated from the inverse of the image covariance matrix. If the image is large then the covariance requires prohibitively large number of samples to be inverted. In the current study, 30 × 30 pixel regions were selected to constrain the covariance matrix to a square region around the possible signal locations (24); 3) the channelized Hotelling observer (CHO) was introduced as a way to incorporate elements from human visual system
into models of signal detection for medical images (20). Here, a CHO with orientation and spatial frequency tuned Gabor functions (23,24) was chosen. The Gabor channel mechanism used in this article has 80 channels with five spatial frequencies (central frequencies: 16, 8, 4, 2, and 1 cycle per degree), eight orientations (equally spaced in orientation), and two phases (odd 0; even $\pi/2$). The spatial frequency bandwidth of the channels was approximately one octave. The channelized Hotelling template was the linear combination of Gabor channels that maximized signal detectability ($d'$); 4) The Laguerre Gauss Hotelling observer (LGHO) used functions that were not intended to reflect the human visual system, but instead to reliably estimate the best linear template in a computationally efficient manner. For more detail on the LGHO model, please refer to (29). In this implementation we

**Figure 1.** Signals and the Corresponding Model Observer Templates. NPWE = nonprewhitening matched filter with an eye filter; HO = Hotelling observer; CHO = channelized Hotelling observer; LGHO = Laguerre Gauss Hotelling observer. Row one: four different signals. Row two: signal templates. Row three: NPWE templates. Row four: CHO templates. Row five: HO templates. Row six: LGHO templates.
used up to the sixth order of the Laguerre-Gauss polynomials and three orientations (vertically oriented, horizontally oriented, and rotationally invariant) resulting in a total of 18 channels. The use of oriented channels was due to the fact that our signals were oriented.

Figure 1 shows several samples of signals and the corresponding model observer templates (NPWE, CHO, HO, LGHO).

For a task with four possible signal locations (four alternative forced choice [AFC]), model observer performance was computed by calculating the probability of a correct outcome. Figure 2 shows the model observer decision calculation for each trial for the SKEV task. For a given image, model observer picks up the $j^{th}$ template corresponding to the known signal and calculates the responses for the four possible locations. A correct outcome occurs when the response of the $j^{th}$ template to the signal location exceeds the maximum response to the noise-only locations. An estimate of the proportion correct ($P_c$) is obtained from samples by applying the template to the different locations in all test images and tallying the proportion of trials where the model correctly identifies the signal location:

$$
P_c = \frac{1}{H} \sum_{h=1}^{H} O_h
$$

where $H$ is the total number of trials/images, $\lambda_{s,j,h}$ is the template response in trial $h$ to the signal present location, and the max function takes the maximum among the $M-1$ ($M = 4$ is the number of possible signal locations in the present task) responses to the noise locations, $\lambda_{n,j,h}$. Equation 2 computes the proportion of trials in which the model observer correctly identifies the signal location, $P_c$.

$P_c$ can then be converted to an empirically obtained index of detectability ($d'$) by generating a look up table for $P_c$ versus the index of detectability ($d'$) from the following relationship (30):

$$
P_c(d', M) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Phi(x)^{M-1} dx,
$$

where $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$, $\Phi(x)$ is the cumulative Gaussian distribution function, $\Phi(x) = \int_{-\infty}^{x} \phi(y) dy$, 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 (19,31). For example, as shown in Figure 3, the $d' = 1.2$ equals a $P_c$ of 0.8 for a 2 AFC task. With the same $d'$, the $P_c$ is 0.62 for a 4 AFC task.

Model observer performance in the SKS task can be computed similarly by calculating the proportion of trials in which the sum of likelihood for the signal location takes the largest value. We briefly describe the mathematical expression for model observers in the SKS task in the Appendix.

Psychophysical Experiment Setup

Human observers participated in the same task to which the model observers were applied: detection of a
signal (filling defect) in one of the four computer-simulated arterial segments. In the SKEV condition, on each trial, a high-contrast copy of the signal was presented below the test image to inform observer the signal size/shape on that trial. In the SKS condition, no information was provided about the presented signal except that the observer knew it was sampled from a group of 184 possible signals and the ranges of possible sizes/shapes. The images were displayed on an image system M17LMAX monochrome monitor with a maximum resolution of 1664 × 1280 pixels (Image Systems, Minnetonka, MN). Before the experiment, the monitor was calibrated according to the Digital Imaging and Communications in Medicine standard (http://medical.nema.org/). For each trial, the human observer chose the location where he or she thought the signal appeared from four possible locations. Experiments were conducted in a darkened room with a viewing distance of 40 cm. On each trial an image was randomly selected from the 900 test image database and displayed. No time limitation was imposed on the observer to make a decision. Three observers participated in the experiment. Each observer participated in nine sessions (100 trials per session) per compression condition resulting in a total of 900 trials per experimental condition. All observers were trained for 900 trials with the same types of signals but with higher contrast. The experimental sequence of the SKEV and SKS tasks was randomized across observers and experimental days to avoid bias and learning effects. Human performance (Pc) 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 Pc using Equation 3. The total number of conditions was 11 including both the compression with and without HVS quantization scheme at the compression ratios of 7:1, 10:1, 15:1, 20:1, 30:1, and the uncompressed condition.

RESULTS

Model Observer Performance

Figure 4 shows performance ($d'$) of the four studied model observers NPWE, HO, CHO, and LGHO for compression ratios 7:1, 10:1, 15:1, 20:1, and 30:1 for both the SKEV and SKS tasks. For comparison, the graph also shows model observer performance for the uncompressed condition, compression ratio 1:1. As with other compression algorithms—JPEG (32), JPEG 2000 (33)—model observer performance is a continuous decreasing function of compression ratio (5,34). Results show that all model observers result in higher performance with the HVS quantization scheme (squares) than the default quantization scheme (triangles) for both the SKEV and SKS tasks for all five studied compression ratios.

Table 1 gives model observer performance improvement for both the SKEV and SKS tasks from the default...
to the HVS quantization scheme. Performance improvement increases with the increasing of compression ratios. For all model observers, performance improvement with the HVS quantization matrix was comparable for SKEV and SKS tasks.

Human Observer Performance

Figure 5 shows human observer performance for the SKEV and SKS task for the uncompressed condition and the five compressed conditions with or without the HVS quantization scheme. Performances for individual observers (initialed by K.J., M.G., and Y.Z.) and the averaged human performance are presented. Error bars represent the standard error of the mean across sessions (nine sessions of 100 trials for each compression condition). Table 2 gives the human observer performance improvement from the default quantization scheme to the HVS quantization scheme. The performance improvement in the SKEV task at the 30:1 compression ratio represents an improvement of 34% over performance without the HVS quantization (Table 2). At 20:1 compression ratio, the performance improvement was 15% for the SKEV task. For the SKS task (Table 2), the performance improvement was 38% at the 30:1 compression ratio and 20% at the 20:1 compression ratio. Results show that the performance improvement increases with increasing compression ratios. The difference across tasks (SKEV, SKS) for the human observers is smaller than that for the model observers (Fig. 4 vs. Fig. 5). This occurs because humans have intrinsic (35,36) uncertainty and cannot use precise knowledge of the signal information even when the signal is shown to them.

Table 3 shows the t-test results comparing performance for the images compressed with and without the HVS quantization scheme for SKEV and SKS tasks for each individual observer and compression ratio. Although we observed higher performance for the HVS quantization scheme at low compression levels (7:1 and 10:1), these differences did not reach significance ($P > .05$; Table 3). The differences in human performance across quantization schemes (averaged across observers) were significant for 15:1, 20:1, and 30:1 compres-

![Figure 5](image-url)
Table 2

<table>
<thead>
<tr>
<th>CR</th>
<th>Average</th>
<th>KJ</th>
<th>MG</th>
<th>Y. Zhang</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SKEV</td>
<td>SKS</td>
<td>SKEV</td>
<td>SKS</td>
</tr>
<tr>
<td>7:1</td>
<td>6.10</td>
<td>1.81</td>
<td>8.88</td>
<td>1.57</td>
</tr>
<tr>
<td>10:1</td>
<td>7.41</td>
<td>6.85</td>
<td>7.96</td>
<td>15.35</td>
</tr>
<tr>
<td>15:1</td>
<td>6.24</td>
<td>12.84</td>
<td>9.87</td>
<td>9.42</td>
</tr>
<tr>
<td>20:1</td>
<td>13.34</td>
<td>21.48</td>
<td>10.92</td>
<td>29.40</td>
</tr>
<tr>
<td>30:1</td>
<td>34.84</td>
<td>37.87</td>
<td>40.83</td>
<td>53.32</td>
</tr>
</tbody>
</table>

CR = compression ratio; SKEV = signal known exactly but variable task; SKS = signal known statistically task.

Table 3

<table>
<thead>
<tr>
<th>CR</th>
<th>Y Zhang</th>
<th>Jorgensen K</th>
<th>Gonzalez M</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SKEV</td>
<td>SKS</td>
<td>SKEV</td>
<td>SKS</td>
</tr>
<tr>
<td>7:1</td>
<td>1.664</td>
<td>1.144</td>
<td>0.670</td>
<td>1.180</td>
</tr>
<tr>
<td>10:1</td>
<td>0.524</td>
<td>1.217</td>
<td>1.506</td>
<td>1.995</td>
</tr>
<tr>
<td>15:1</td>
<td>0.297</td>
<td>1.440</td>
<td>*4.530</td>
<td>1.766</td>
</tr>
<tr>
<td>30:1</td>
<td>*5.022</td>
<td>*2.931</td>
<td>*2.774</td>
<td>*5.725</td>
</tr>
</tbody>
</table>

* Significant values

SKEV = signal known exactly but variable task; SKS = signal known statistically task; CR = compression ratio.

Each compression has two compression conditions: with HVS and without HVS. Each compression condition has nine experimental sessions to get the statistics for the t-test. The fifth and the ninth columns show t values for performance averaged across observers. Significant t values (P < .05) are highlighted.
formance with model observer performance after including internal noise. We performed the calculation 100 times and obtained a mean Pc for each condition, thus the resulted error bars were much smaller than the plotting symbols. Results show that inclusion of internal noise does not change the rank order of the model observer across quantization schemes (default vs. HVS) performance. The ability of models after inclusion of the internal noise to predict human performance was assessed by a chi-square goodness of fit. Table 4 shows that there is a good agreement between the average human performance and the model observers with internal noise. In addition, using the proportional method, the LGHO and the NPWE models are the best predictors (lowest $\chi^2$), whereas the HO results in the highest $\chi^2$. Using the quadratic method, the HO and the NPWE models are the best predictors (lowest $\chi^2$) among the four studied model observers, whereas the LGHO results in the highest $\chi^2$. Depending on the particular method adopted to introduce internal noise, different models present differences on the ability to predict human performance. Although some differences exist in the ability of the models to predict

![Figure 6](image-url)

**Figure 6.** Model observer (with internal noise) and human observer performance for SKEV task. From left to right: NPWE, HO, CHO, LGHO. Row one: proportional internal noise method. Row two: quadratic internal noise method.

<table>
<thead>
<tr>
<th>Model</th>
<th>NPWE</th>
<th>HO</th>
<th>CHO</th>
<th>LGHO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square (proportional method)</td>
<td>0.17</td>
<td>7.02</td>
<td>2.33</td>
<td>0.96</td>
</tr>
<tr>
<td>Chi-square (quadratic method)</td>
<td>0.24</td>
<td>0.21</td>
<td>0.70</td>
<td>2.24</td>
</tr>
</tbody>
</table>

NPWE = nonprewhitening matched filter with an eye filter; HO = Hotelling observer; CHO = channelized Hotelling observer; LGHO = Laguerre Gauss Hotelling observer.

Reduced chi-square is defined as $1/(n - p) \sum_{i=1}^{n} (d_{hi} - d_{mi})^2/\sigma_i^2$, where $d_{hi}$ and $d_{mi}$ denote the detectability for human and model observer at the $i^{th}$ condition respectively. Also, $\sigma_i^2$ is the observer variability of the human $d_{hi}$, $n$ is the number of studied conditions ($n = 11$, including the uncompressed one and the 7:1, 10:1, 15:1, 20:1, 30:1 compression for the case with or without using HVS quantization scheme), and $p$ is the number of fitting parameters (for our case $p = 1$, the internal noise). For computational simplicity, the internal noise was adjusted to match Pc of the uncompressed condition rather than minimizing the overall $\chi^2$ across all data points.
human performance, it should be emphasized that all four models correctly predict the superiority of the HVS quantization scheme over the default (comparing the first row with the second row of Fig 6), except the CHO model with the quadratic internal noise injection method.

**DISCUSSION**

**Compression Ratio and Scheme Selection**

Users should select the appropriate compression ratio/scheme according to their objective and applications. It would be unrealistic to claim that one compression ratio can be used for all types of image display, transfer, and storage (40). Higher compression ratios can be used for applications that do not have an immediate impact on patient care and studies that are relatively old and have a low probability of being used for clinical purposes. Original images (or lossless compressed images) should be used for short-term interpretation and for long-distance consultation with therapeutic implications. However, our present and previous results (34) show that careful consideration of the quantization schemes of compression algorithms can greatly mitigate the impact of image compression on task performance in clinically relevant tasks such as detection. In this article, we investigated the effect of using the HVS quantization scheme with the SPIHT wavelet compression on human and model performance detecting signals varying in size within x-ray coronary angiograms. Our results showed that performance degradation as a function of compression can be greatly mitigated by using the HVS quantization scheme (up to a 38% improvement for 30:1 compression ratio). The present result joins a number of studies in which compression encoding schemes different from the default ones have led to improved human detection performance (5,34). Together these studies emphasizes the importance of considering the observer and the task when choosing parameters in image processing algorithms.

**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 or encoding parameters for the simpler SKE task (7,41), which is simple but not representative of the signals in a clinical scenario in which signals vary in size and shape across images. For this reason, we investigated SKEV and SKS tasks. However, the SKEV approach is useful if and only if the SKEV task performance is highly correlated with the SKS task performance. Our results confirm the close relationship between the two tasks for the range of size/shape uncertainty studied. Performance is worse in the SKS task than the SKEV because the model/human cannot use prior knowledge in the SKS task to ignore templates corresponding to signals that are not present in that trial. This can be explained by the concept of stimulus uncertainty on human and model performance (42). The close relationship between the NPWE, HO, CHO, and LGHO models for detection tasks with x-ray coronary angiograms is consistent with our previous results comparing model performance for different compression algorithms (5,6). Model observers can also accurately predict the improvement of human performance for both the SKEV and SKS task using the HVS quantization scheme in the wavelet compression algorithm. The present results therefore confirm previous success in predicting human performance with model observers with other compression schemes (JPEG and JPEG 2000) for detecting varying signals in x-ray coronary angiograms.

**Internal Noise**

An important question is whether the addition of internal noise to the models in the evaluation changes the rank order of the performance. Our present results showed that for all model observers and conditions the wavelet compression algorithm with HVS quantization scheme led to higher performance than that using the default quantization scheme regardless of whether internal noise was included or not. Thus the inclusion of internal noise did not change the conclusions about the compression conditions. We found that the agreement between model observer (with added internal noise) and human performance was best for NPWE and LGHO models followed by the CHO model when using the proportional internal noise method to inject the internal noise and the agreement was best for the Hotelling and NPWE models followed by the CHO model when using a quadratic internal noise method. Results showed that the agreement between model and human performance was specific to the method used to introduce internal noise. Furthermore previous results have shown that the CHO model with channel internal noise (rather than the decision variable internal noise) is a good prediction of human performance. Future research will systematically evaluate different methods of including internal noise in model observers.
Limitations and Relation to Clinical Studies

Our present work evaluated the SPIHT wavelet based compression algorithm with a human visual system quantization scheme on model and human observer performance with computer simulated filling defects. Previous studies have shown a general agreement between results obtained with trained observers and simulated signals in x-ray angiograms and physician’s performance with real morphological features in clinical images (2,43,44). Given the close correspondence between the simulated morphological features versus real ones, we suggest that the present results might be used to guide choice of compression parameters in future clinical studies. However, detection of filling defects is only one piece of clinical information obtained from coronary angiograms. Other tasks include the assessment of flow characteristics, lesion calcification, detection of collaterals, dissection, flaps, and assessment of stent deployment. It may be necessary to evaluate image compression schemes/ratios for every type of image and lesion seen clinically (45). However, some previous results have shown that compression schemes that lead to improved detection of filling defects also lead to improved performance with other tasks such as ulcer and bridging lumen detection as well as stenosis grading (2). In addition, the 4 AFC task is not really what physicians do in clinical practice in which they decide on whether a lesion (signal) is present or absent and its location is unknown (yes/no, location uncertainty).

CONCLUSIONS

Model observers were used to quantitatively evaluate the effect of wavelet compression algorithm with human visual system quantization scheme on the visual signal detection task consisting of simulated signals embedded in real x-ray coronary angiographic backgrounds. We investigated the SKEV and SKS tasks. We also compared model observer performance with human performance in a subsequent psychophysical study. We found good agreement between human observer and model observer performance and among different model observers (non-preshrinning 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 for some tasks as a reliable predictor on model and human performance for the computationally more complex but clinically more realistic SKS task. Most importantly for practical purposes the SPIHT wavelet compression algorithm with the HVS quantization scheme may greatly improve human signal detection performance. The average human performance improvements reached approximately 38% (d’) for the higher compression ratio such as 30:1.

ACKNOWLEDGMENT

The authors thank A. Manduca from the Mayo Clinic Foundation for providing the SPIHT wavelet compression code. The authors also thank K. Jorgensen and M. Gonzalez for their participation as observers in the psychophysical studies.

REFERENCES


APPENDIX

Calculation of Model Observer Performance for SKS Task

The following section briefly discusses the mathematical framework of the SKS model observers used in this article. A more detailed treatment of the topic has been previously presented (5.6).

The first stage in SKS models is to obtain a set of templates, one for each of the possible signals that might be present. The models compute a dot product between each template and the data at each location:

$$\lambda_{ij} = \omega_s g_i$$ (A.1)

where $\lambda_{ij}$ is the scalar response of a template corresponding to the $j^{th}$ signal type ($J$ different signals in total) to $g_i$ the data at the $i^{th}$ location. The next step is to combine the $J$ scalar responses per location into a single response. The sum of likelihood rule (assuming equal variance Gaussian internal responses) is used to combine the J scalar responses per location (6,46). On each $j^{th}$ trial, sum of the likelihoods of the filter responses ($\lambda$) is given by (6,46):

$$I_j = \sum_{j=1}^{M} \left[ \frac{1}{\sqrt{2\pi\sigma_j^2}} \right] \exp \left( -\frac{(\lambda_{ij} - \mu_{ij})^2}{2\sigma_j^2} \right) \times \max_{m=1}^{M} \left[ \exp \left( -\frac{(\lambda_{mj} - \mu_{ij})^2}{2\sigma_j^2} \right) \right]^{(1-\delta_{im})}$$ (A.2)
where the summation is over all $J$ possible signal types, $\mu_{j,i}$ is the expected response of the $j^{th}$ template to the $j^{th}$ signal type, $\mu_{j,n}$ is the expected response of the $j^{th}$ template to the signal absent locations, $\sigma_j$ 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.

Model observer performance in the SKS task can be computed by calculating the proportion of trials in which the sum of likelihood for the signal location takes the largest value. This can be written as:

$$O_h = \text{step}(l_{s,h} - \max(l_{n,h}))$$

$$= \begin{cases} 
1 & \text{if } l_{s,h} \geq \max(l_{n,h}) \\
0 & \text{if } l_{s,h} < \max(l_{n,h}) 
\end{cases}$$

$$Pc = \frac{1}{H} \sum_{h=1}^{H} O_h \quad (A.3)$$

where $l_{s,h}$ is the likelihood (Eq A.2) at the signal-present location in trial $h$, $\max(l_{n,h})$ is the maximum of the likelihoods at the signal-absent locations, and $H$ is the total number of trials. The obtained proportion correct can be transformed to an index of detectability measure using Equation 3.