

# **An Efficient Observer Model for Assessing Signal Detection Performance of Lossy-Compressed Images**

Brian M. Schmanske and Murray H. Loew

The George Washington University, School of Engineering and Applied Science  
Institute for Medical Imaging and Image Analysis  
Department of Electrical and Computer Engineering  
801 22nd Street NW, Washington, DC 20052  
schmansk@gwu.edu

**Abstract.** A technique for assessing the impact of lossy wavelet-based image compression on signal detection tasks is presented. A medical image's value is based on its ability to support clinical decisions, including detecting and diagnosing abnormalities. However, image quality of compressed images is often stated in terms of mathematical metrics such as mean square error. The presented technique provides a more suitable measure of image degradation by building on the channelized Hotelling observer model, which has been shown to predict human performance of signal detection tasks in noise-limited images. The technique first decomposes an image into its constituent wavelet subbands. Channel responses for the individual subbands are computed, combined, and processed with a Hotelling observer model to provide a measure of signal detectability versus compression ratio. This allows a user to determine how much compression can be tolerated before image detectability drops below a certain threshold.

## **1 Introduction**

Lossy image compression used to reduce the size of large medical image files degrades image quality. Often, quality metrics such as mean square error or peak signal-to-noise ratio are used to assess image quality. However, these metrics do not adequately measure the true value of a medical image, namely, its ability to support clinical decisions, including detecting and diagnosing abnormalities.

Human observers can be used to determine image quality for performing signal detection tasks, but the trials require many human observers assessing many images to make the results statistically relevant. These trials are thus very time consuming and expensive. Model observers can be used as a human surrogate if the models can be shown to accurately predict human performance.

## **2 Model Observers**

Model observers are used to predict human observer performance in detecting signals in noise limited images that are representative of nuclear-medicine images [1],[2]. Model observers reduce an image to a decision variable that is compared to a decision

threshold to decide whether a signal is present or not. For a signal in Gaussian noise, the ideal observer is given by:

$$\lambda(\mathbf{x}) = \left( \mathbf{K}_x^{-1} \Delta\mu_x \right)^t \mathbf{x} \quad (1)$$

where  $\mathbf{x}$  is the image represented as a vector,  $\mathbf{K}_x$  is the image covariance matrix, and  $\Delta\mu_x$  is a vector containing the difference between the mean “signal present” image and the mean “signal not present” image.  $\lambda(\mathbf{x})$  is the decision variable produced by this decorrelating matched filter.

If the image’s probability density function is unknown, the first and second order statistics can still be used to generate the covariance matrix and mean difference vector of (1). Though no longer an ideal observer, the decision variable can still predict human observer performance. This model observer is known as the Hotelling observer.

A difficulty with using the Hotelling observer is the generation of an invertible covariance matrix. For an image with  $k$  pixels, the covariance matrix will have  $k^2$  elements, and require  $k$  sample images to ensure the matrix is nonsingular [2]. Even for modest sized images, the required number of sample images becomes quite large and inverting the covariance matrix becomes computationally intensive.

A channelized Hotelling observer solves the problem associated with the covariance matrix. The channelized Hotelling observer operates on channel responses of the image rather than on the image itself [2]-[4]. The channel responses are generated by

$$\eta = \mathbf{T}^t \mathbf{x} \quad (2)$$

where the columns of  $\mathbf{T}$  contain the channel templates associated with different spatial frequency bands. The channel templates can be chosen to mimic the human visual system or to optimize results relative to the ideal observer. The channel responses,  $\eta$ , are then processed with the Hotelling observer, using  $\mathbf{K}_\eta^{-1}$  and  $\Delta\mu_\eta$  to generate the decision variable.

In a two-alternative-forced-choice test, samples of “signal present” and “signal not present” images are used to compute observer performance. Decision variables for the samples are obtained, and the decision threshold is varied to obtain receiver operating characteristic (ROC) curves. The area under the ROC curves (AUC) is a measure of observer performance and can be used to assess image quality.

The channelized Hotelling observer model has been shown to predict human observer performance in a number of tasks, including predicting degraded performance when lossy image compression is employed [5]. The compression studies looked at a limited number of compression ratios by compressing the entire image and evaluating the degraded image with the model. Since all image samples must be compressed/decompressed for each value of compression ratio, most practical problems can examine only a limited number of ratios.

### 3 Wavelet Compression

JPEG 2000 uses wavelets to compress images [6]. A wavelet compression algorithm first analyzes a signal, decomposing it into high-pass and low-pass subbands. The

high-pass subband coefficients have lower entropy than the original signal samples, and entropy coding these coefficients yields compression gain [7]. Following the wavelet analysis, the bitmaps of the subband coefficients are entropy coded separately. The encoded bitmaps are then arranged and transmitted to deliver the spatial (subband) resolution and SNR (bitmaps) required. As subbands and bitmaps are discarded from the compressed file, the compression ratio increases and image quality decreases.

In one dimension, wavelet analysis is implemented with linear subband filters:  $\mathbf{h}_0[n]$  represents the low-pass subband filter and  $\mathbf{h}_1[n]$  represents the high-pass subband filter.

$$\mathbf{X}[2k] = \mathbf{h}_0[n] * \mathbf{x}[n] \Big|_{n=2k} = \sum_{l \in \mathbb{Z}} \mathbf{h}_0[2k-l] \mathbf{x}[l] \quad (3)$$

and

$$\mathbf{X}[2k+1] = \mathbf{h}_1[n] * \mathbf{x}[n] \Big|_{n=2k+1} = \sum_{l \in \mathbb{Z}} \mathbf{h}_1[2k+1-l] \mathbf{x}[l] \quad (4)$$

where  $\mathbf{x}[n]$  are the original signal samples,  $\mathbf{X}[n]$  are the wavelet coefficients, and  $\mathbb{Z}$  is the set of all integers.

Evaluating a convolution at even indices is equivalent to a filter followed by down-sampling by 2 [7]. Therefore,  $\mathbf{X}[2k]$  and  $\mathbf{X}[2k+1]$  can be obtained from a two-channel analysis filter bank with filters  $\mathbf{H}_0$  and  $\mathbf{H}_1$  followed by downsampling by 2. The channel filter responses are denoted by  $\mathbf{v}[k]$  and  $\mathbf{u}[k]$ , where

$$\mathbf{v}[k] = \mathbf{X}[2k+1] \quad (5)$$

$$\mathbf{u}[k] = \mathbf{X}[2k]. \quad (6)$$

$\mathbf{v}[k]$  and  $\mathbf{u}[k]$  are the high-pass and low-pass coefficients respectively.

The original signal is synthesized, or reconstructed, from the subband coefficients using filters  $\mathbf{g}_0[n]$  and  $\mathbf{g}_1[n]$ :

$$\mathbf{x}[n] = \sum_{k \in \mathbb{Z}} \mathbf{u}[k] \mathbf{g}_0[n-2k] + \sum_{k \in \mathbb{Z}} \mathbf{v}[k] \mathbf{g}_1[n-2k]. \quad (7)$$

where  $\mathbf{g}_i[n] = \mathbf{h}_i[-n]$ ,  $i=0,1$ .

For two levels of analysis where the low-pass coefficients are analyzed a second time, the coefficients are found as follows:

$$\mathbf{v}_1[k] = \sum_{l \in \mathbb{Z}} \mathbf{h}_1[2k-l] \mathbf{x}[l] \quad (8)$$

$$\mathbf{u}_1[k] = \sum_{l \in \mathbb{Z}} \mathbf{h}_0[2k-l] \mathbf{x}[l] \quad (9)$$

$$\mathbf{v}_2[k] = \sum_{l \in \mathbb{Z}} \mathbf{h}_1[2k-l] \mathbf{u}_1[l] \quad (10)$$

$$\mathbf{u}_2[k] = \sum_{l \in \mathbb{Z}} \mathbf{h}_0[2k-l] \mathbf{u}_1[l]. \quad (11)$$

The subscripts on the subband coefficient vectors indicate the number of analysis steps the signal has experienced.  $\mathbf{v}_1$ ,  $\mathbf{v}_2$ , and  $\mathbf{u}_2$  would then be coded and transmitted in a compressed file.

Following the multiple analysis steps, the original signal is traditionally reconstructed by synthesizing the coefficients in the reverse order that they were analyzed. For the above example, the low-pass coefficients for the first subband are synthesized from the high-pass and low-pass coefficients of the second subband

$$\mathbf{u}_1[n] = \sum_{k \in Z} \mathbf{u}_2[k] \mathbf{g}_0[n-2k] + \sum_{k \in Z} \mathbf{v}_2[k] \mathbf{g}_1[n-2k]. \quad (12)$$

The original signal can then be recovered using the high-pass and low-pass coefficients of the first subband:

$$\mathbf{x}[n] = \sum_{k \in Z} \mathbf{u}_1[k] \mathbf{g}_0[n-2k] + \sum_{k \in Z} \mathbf{v}_1[k] \mathbf{g}_1[n-2k]. \quad (13)$$

The convolution is a linear operation and follows the distributive property. That is:

$$\mathbf{f}[n] * (\mathbf{a}[n] + \mathbf{b}[n]) \Big|_{n=2k} = \sum_{l \in Z} \mathbf{f}[2k-l] (\mathbf{a}[l] + \mathbf{b}[l]) \quad (14)$$

$$= \sum_{l \in Z} \mathbf{f}[2k-l] \mathbf{a}[l] + \sum_{l \in Z} \mathbf{f}[2k-l] \mathbf{b}[l] \quad (15)$$

The low-pass coefficients of the first subband can then be computed as follows:

$$\mathbf{u}_1[n] = \sum_{k \in Z} \mathbf{u}_2[k] \mathbf{g}_0[n-2k] + \sum_{k \in Z} \mathbf{v}_2[k] \mathbf{g}_1[n-2k] \quad (16)$$

$$= \mathbf{u}_{21}[n] + \mathbf{v}_{21}[n]. \quad (17)$$

$\mathbf{u}_{jk}[n]$  and  $\mathbf{v}_{jk}[n]$  represent the high-pass and low-pass coefficients for the  $j^{\text{th}}$  subband, synthesized to the  $k^{\text{th}}$  subband level. When  $k$  is zero, the coefficients represent the subband contributions to the original signal.

Substituting (17) into (13), and taking advantage of the distributive property of the convolution, the original signal can be reconstructed as follows:

$$\mathbf{x}[n] = \sum_{k \in Z} \mathbf{u}_1[k] \mathbf{g}_0[n-2k] + \sum_{k \in Z} \mathbf{v}_1[k] \mathbf{g}_1[n-2k] \quad (18)$$

$$= \sum_{k \in Z} (\mathbf{u}_{21}[k] + \mathbf{v}_{21}[k]) \mathbf{g}_0[n-2k] + \sum_{k \in Z} \mathbf{v}_1[k] \mathbf{g}_1[n-2k] \quad (19)$$

$$= \sum_{k \in Z} \mathbf{u}_{21}[k] \mathbf{g}_0[n-2k] + \sum_{k \in Z} \mathbf{v}_{21}[k] \mathbf{g}_0[n-2k] + \sum_{l \in Z} \mathbf{v}_1[k] \mathbf{g}_1[n-2k] \quad (20)$$

$$= \mathbf{u}_{20}[n] + \mathbf{v}_{20}[n] + \mathbf{v}_{10}[n]. \quad (21)$$

Hence, the original signal can be reconstructed by summing the individual subband contributions to the original signal.

Images are two-dimensional data sets. They can be analyzed with 1-D wavelets with a two-step method. First, the rows of the image are analyzed with filters **H0** and **H1**. The resulting coefficients then are analyzed across the columns. This analysis produces four subbands for each level of analysis. One subband contains the low-

pass coefficients (LL). Others contain high-pass coefficients in the horizontal direction (HL), the vertical direction (LH), and the diagonal direction (HH). Subsequent analysis can be performed on the low-pass coefficients [7]. The analysis of an image using two levels of decomposition is shown in Fig. 1.

Taking advantage of the distributive property of the convolution, the subband coefficients that result from image analysis can be reconstructed so the high-pass and low-pass coefficients from each subband are synthesized separately. Fig. 2 illustrates how the high-pass coefficients from each subband can be synthesized so their contribution to the original image can be obtained. The individual subband contributions can then be summed to recover the original image:  $\mathbf{x} = \mathbf{u}_{20} + \mathbf{v}_{20} + \mathbf{v}_{10}$ . Since the subband coefficients are themselves linear combinations of bitmaps, a synthesis could also be constructed to provide bitmap contributions rather than subband contributions.

## 4 Subband-Channelized Hotelling Observer

The standard channelized Hotelling observer, as mentioned in Section 2, is implemented by first passing a signal through a set of channels to obtain channel responses. The channel responses are then analyzed with a Hotelling observer to determine whether a signal is present.

Consider an image  $\mathbf{x}$ . The channel responses are given by

$$\boldsymbol{\eta} = \mathbf{T}^t \mathbf{x}. \quad (22)$$

where the column vectors of  $\mathbf{T}$  each represent channel templates with different spectral profiles. The channel responses,  $\boldsymbol{\eta}$ , are then processed by a Hotelling observer to obtain a decision variable,  $\lambda(\boldsymbol{\eta})$ ,

$$\lambda(\boldsymbol{\eta}) = (\mathbf{K}_\eta^{-1} \Delta \boldsymbol{\mu}_\eta)^t \boldsymbol{\eta}. \quad (23)$$

$\mathbf{K}_\eta$  is the covariance matrix of the channel responses  $\boldsymbol{\eta}$ .  $\Delta \boldsymbol{\mu}_\eta$  is the difference between the channel responses of the two classes of images: “signal present” and “signal not present”.

Section 3 showed an image could be analyzed and synthesized to obtain the individual subband or bitmap contributions. For the case that considers only subband contributions, the image is represented as a sum of subband contributions:

$$\mathbf{x} = \mathbf{u}_{40} + \mathbf{v}_{40} + \mathbf{v}_{30} + \mathbf{v}_{20} + \mathbf{v}_{10} \quad (24)$$

for four levels of wavelet analysis.

Substituting (24) into (22),

$$\boldsymbol{\eta} = \mathbf{T}^t (\mathbf{u}_{40} + \mathbf{v}_{40} + \mathbf{v}_{30} + \mathbf{v}_{20} + \mathbf{v}_{10}) \quad (25)$$

$$= \mathbf{T}^t \mathbf{u}_{40} + \mathbf{T}^t \mathbf{v}_{40} + \mathbf{T}^t \mathbf{v}_{30} + \mathbf{T}^t \mathbf{v}_{20} + \mathbf{T}^t \mathbf{v}_{10} \quad (26)$$

$$= \boldsymbol{\eta}_{\mathbf{u}40} + \boldsymbol{\eta}_{\mathbf{v}40} + \boldsymbol{\eta}_{\mathbf{v}30} + \boldsymbol{\eta}_{\mathbf{v}20} + \boldsymbol{\eta}_{\mathbf{v}10} \quad (27)$$

Equation (27) shows the channel response of an image can be represented as a sum of channel responses of the image’s subband components. Effects of lossy compression can be measured by removing channel responses of component subbands. For

example, if the first subband high-pass coefficients are not included in the compressed image file, use the following to compute  $\mathbf{K}_\eta^{-1}$ ,  $\Delta\mu_\eta$ , and  $\lambda(\eta)$ :

$$\eta = \eta_{u40} + \eta_{v40} + \eta_{v30} + \eta_{v20}. \quad (28)$$

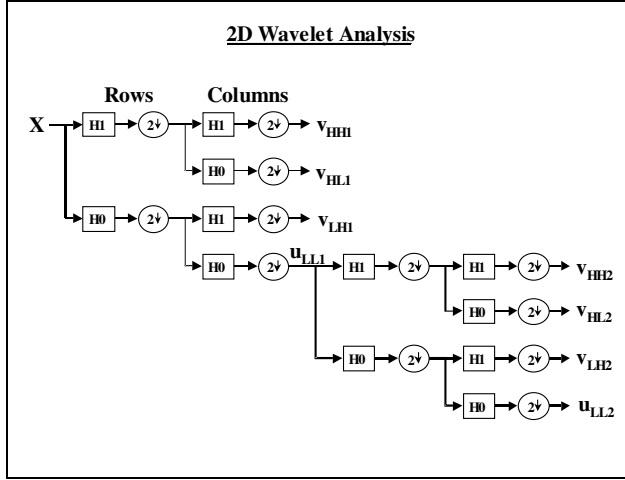


Fig. 1. Analysis of 2-D images

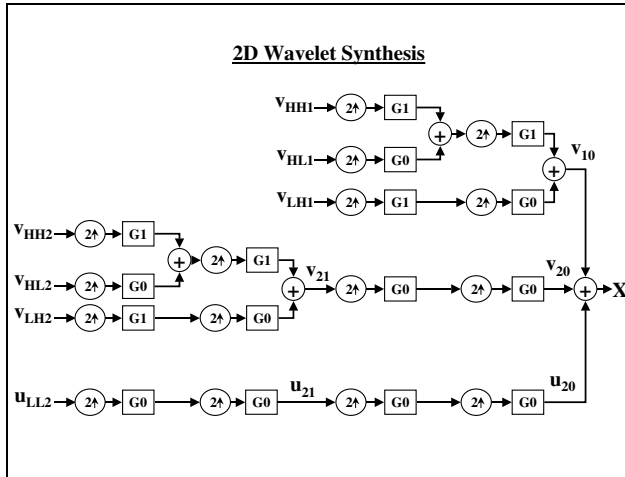


Fig. 2. Subband synthesis of 2-D image

## 5 Simulation and Results

To demonstrate the utility of the subband-channelized Hotelling observer model, the AUCs of sample images are computed for various compression ratios. Four hundred simulated images were used with a correlated Gaussian noise background. Half of the images contained a signal, represented as a disk with a Gaussian intensity profile.

Five channel profiles were used, defined by Laguerre-Gauss functions as described in [8]. A simple Haar wavelet was used to generate the subbands; four levels of analysis were performed. It was assumed that spatial resolution progression was used to arrange the bitmaps, so compression is achieved by sequentially removing the four high-pass subbands. Compression ratios were computed using the bitmap entropies of each subband. Fig. 3 shows the AUC vs. compression ratios for the images. The figure clearly shows the degradation in image quality relative to the signal detection task as spatial resolution decreases and compression ratios increase, allowing the user to choose the tradeoff between compression ratio and signal detection performance.

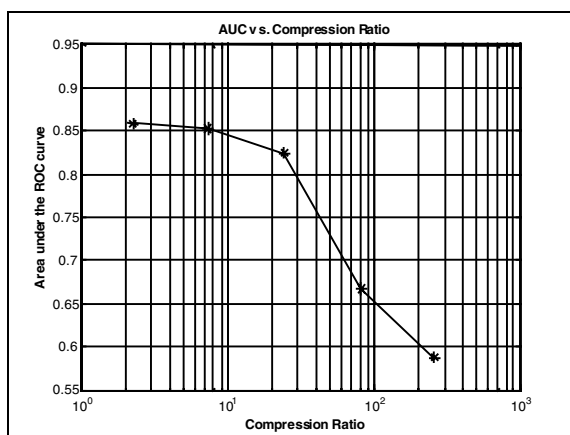


Fig. 3. Model observer performance

## 6 Conclusions

A technique for efficiently calculating observer performance of nuclear medicine images degraded by lossy compression has been presented. The technique is based on the channelized Hotelling observer model, which has been shown to predict human observer performance for detection tasks with noise-limited images. The new technique uses channel responses of the subbands to compute observer performance rather than the channel responses of the image itself.

Plotting the AUC versus compression ratio for compression that decreases spatial resolution shows the utility of this technique. The plot allows one to determine an acceptable compression ratio for an image based on the required signal detectability. While results were shown for a subband-channelized Hotelling observer, the technique is extendable to a bitmap-channelized model since the subband coefficients themselves are linear combinations of bitmaps. The bitmap-channelized model will provide greater flexibility and control over the image quality assessment. It will permit the assessment of image compression based on either SNR or mixed progression offered by JPEG 2000, in addition to the compression based on spatial resolution progression shown in Section 5.

Should the standard channelized Hotelling observer be used for this assessment, each compression ratio value would require a separate compression/decompression of the image samples. But by having to synthesize the wavelet coefficients only once,

this subband technique, and its bitmap extension, offers a computationally efficient means of assessing image quality for the optimization of lossy compression in signal detection tasks.

## References

1. Jacob Beutel, Harold L. Kundel, and Richard L. Van Metter ed., *Handbook of Medical Imaging*, Vol. 1. Physics and Psychophysics, SPIE, Bellingham Washington, 2000.
2. Harrison H. Barrett, Jie Yao, Jannick P. Roland, and Kyle J. Myers, "Model Observers for assessment of image quality," *Proc. Natl. Acad. Sci. USA*, Vol. 90, November 1993.
3. Kyle J. Myers and Harrison H. Barrett, "Addition of a channel mechanism to the ideal-observer model," *J. Opt. Soc. Am. A*, Vol. 4, No. 12, December 1987.
4. Craig K. Abbey and Harrison H. Barrett, "Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability," *J. Opt. Soc. Am. A*, Vol 18, No. 3, March 2001.
5. Miguel Eckstein, Craig K. Abbey, Francois O. Bochud, Jay L. Bartoff, James S. Whiting, "The effect of image compression in model and human performance," *SPIE Conference on Image Perception and Performance*, SPIE Vol. 3663, Feb 1999.
6. Athanassios Skodras, Charilaos Christopoulos, and Touradj Ebrahimi, "The JPEG 2000 Still Image Compression Standard," *IEEE Signal Processing Magazine*, Vol 18, No. 5, Sept 2001.
7. Martin Vetterli and Jelena Kovacevic, *Wavelets and Subband Coding*, Prentice Hall PTR, Upper Saddle River, New Jersey, 1995.
8. Brandon D. Gallas, "Signal Detection in Lumpy Backgrounds", PhD. Dissertation, Program in Applied Mathematics, University of Arizona, 2001.