QUALITY MEASURES FOR DECOMPRESSED MEDICAL IMAGES: STATISTICS, OBSERVERS, AND ROCs

Dunling Li and Murray H. Loew

Department of Electrical and Computer Engineering,
George Washington University, Washington DC 20037, USA

ABSTRACT

The diagnostic value of a medical image after it has been decompressed is a very practical measure of the image quality and the compression method. A computational observer can be used to make decisions based on an image and thus can provide an assessment of diagnostic value. This paper provides a closed-form estimate of the channelized Hotelling observer's (CHO) output. In medical applications, CHO has been used successfully to predict human observer performance and to evaluate image quality for detection tasks in various backgrounds. To date, it has been used only empirically – i.e., on a set of images – to evaluate compressed image quality. Using a representation of compression noise this paper derives in closed-form the corresponding CHO test statistics and performance. The performance (as measured by the receiver operating characteristic [ROC]) is verified using decompressed JPEG lumpy-background images. The verification results show that the derived ROCs predict the estimated ROCs very well.

Index Terms — Image coding, quantization, decision-making, image reconstruction, visual system

1. INTRODUCTION

Storage and transmission of large amounts of digitized medical data often requires compression. Since lossless compression generally can yield only a 2:1 to 3:1 compression ratio, lossy image compression must be used in some applications. Image compression algorithms, such as JPEG and JPEG 2000, are employed in medical products already. No standard exists, however, for measurement of the quality of decompressed images. It has become an important research area to estimate the effect of image compression on the accuracy of clinical diagnosis. The most commonly used measurements of image quality, such as mean square error (MSE) or peak signal to noise ratio (PSNR), are not adequate for medical images [1]. Medical image quality can better be measured by human performance in visual tasks that are relevant to clinical diagnosis. The standard method of evaluating diagnostic methods [2] is a receiver operating characteristic (ROC) study, which is time-consuming and costly because it requires a large number of human observations. This is compounded when the set of parameters changes [1].

Computer-model observers are algorithms that attempt to predict human visual performance in noisy images and might represent the desired metric of image quality when the diagnostic decision involves a human observer and a visual task [3]. Among all the model observers, the ideal observer sets an upper bound to the performance of any observer, including the human; it requires knowledge of the PDF of the background noise. The channelized Hotelling observer is one of the most efficient and practical algorithms for prediction of human performance [3,4]; it requires knowledge of the first and second moments of the background noise. Though model observers have been used to estimate decompressed image quality experimentally, there are no closed-form results so far [1,5,6]. Ideally, a closed-form quality measure for decompressed images would allow the user to compare various compression algorithms and parameter sets without extensive image samples. The measure would also provide a way to find optimum compression schemes or parameter sets for clinical diagnostic tasks.

This paper derives the closed-form CHO on decompressed images, shows the mathematical CHO expression for JPEG decompressed lumpy-background images, presents the verification results, and draws conclusions.

2. CHANNELIZED HOTELLING OBSERVER ON DECOMPRESSED IMAGES

Image quality's effect has long been assessed by human-observer performance on two-alternative forced-choice (2AFC) experiments with simultaneous, side-by-side presentation of the two alternative image fields [4,6]. One contains noise only and the other contains signal plus noise. The observer is required to decide which image contains the signal (see Fig. 1). Channelized Hotelling model observers are computational task-performance-based image quality assessment techniques. This paper uses CHO to simulate human behavior on 2AFC data.

A 2AFC can be interpreted as two-class hypothesis test: one hypothesis is an image without signal, and the other is an image with signal. Model observers consider an $N \times N$ image as an $N^2 \times 1$ column vector. If $\vec{\mathbf{S}}$ and $\vec{\mathbf{N}}$ are defined as signal and noise vectors respectively, then the observed image $\vec{\mathbf{X}}$ can be defined as

$$\vec{\mathbf{X}} = \begin{cases} \vec{\mathbf{S}} + \vec{\mathbf{N}} & \text{signal present} \\ \vec{\mathbf{N}} & \text{signal absent} \end{cases} . \quad (1)$$

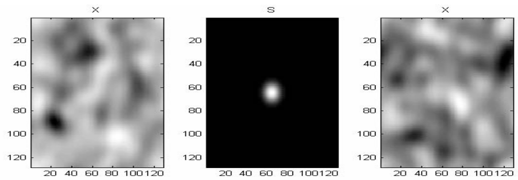

Figure 1. An Example of data for 2AFC Experiments

A model observer performs a detection task by computing a test statistic $\lambda(\vec{\mathbf{g}})$ and comparing it with a threshold λ_t to decide which hypothesis is accepted. The CHO uses a set of spatial frequency channels that are believed to exist in the human visual system in detection tasks [7, 8, 9]. Its detection process operates on the channel outputs, which have the additional benefit of reducing the dimensionality of the problem, making the CHO very computationally effective. The CHO test statistic of the original image is

$$\lambda_{\text{org}} = \vec{\mathbf{S}}^{\mathrm{T}}\mathbf{T}(\mathbf{T}^{\mathrm{T}}\mathbf{S}_2\mathbf{T})^{-1}\mathbf{T}^{\mathrm{T}}\vec{\mathbf{g}} \ , \quad (2)$$

where $\mathbf{T}$ is channel profiles and $\vec{\mathbf{g}}$ is the image vector; S_2 is

$$\mathbf{S}_2 = P_1\mathrm{Cov}(\vec{\mathbf{N}}_1) + P_2\mathrm{Cov}(\vec{\mathbf{N}}_2) = \mathrm{Cov}(\vec{\mathbf{N}}) , \quad (3)$$

where $\mathrm{Cov}(\vec{\mathbf{N}})$ is the covariance matrix of background images. The performance of model observers can be represented by various figures of merit, such as signal to noise ratio (SNR), area under the ROC curve (AUC), detectability, etc. These are all measures that express the tradeoff between true- and false-positives, and it is possible to convert from one to the other.This paper uses AUC as the figure of merit for CHO performance. The AUC can be calculated by

$$\mathrm{AUC} = 1/2 + \mathrm{erf}(d_\lambda/2)/2 , \quad (4)$$

where d_λ is detectability, i.e.,

$$d_\lambda = (\overline{\lambda}_2 - \overline{\lambda}_1)\Big/\sqrt{\sigma_2^2/2 + \sigma_1^2/2} \quad , \quad (5)$$

where $\overline{\lambda}_i$ and σ_i^2 are mean and variance of test statistics for hypothesis i, and i={1,2}. They are defined as

$$\overline{\lambda}_i = \int_\infty \lambda(\vec{\mathbf{g}})P(\vec{\mathbf{g}} \mid T_i)d\vec{\mathbf{g}} \quad (6)$$

and

$$\sigma_i^2 = \int_\infty (\lambda(\vec{\mathbf{g}}) - \overline{\lambda}_i)^2 P(\vec{\mathbf{g}} \mid T_i)d\vec{\mathbf{g}} , \quad (7)$$

where $P(\vec{\mathbf{g}} \mid T_i)$ is the conditional PDF of hypothesis i. The CHO detectability can be derived as

$$d_{\text{org}}^2 = \vec{\mathbf{S}}^{\mathrm{T}}\mathbf{T}(\mathbf{T}^{\mathrm{T}}\mathbf{S}_2\mathbf{T})^{-1}\mathbf{T}^{\mathrm{T}}\vec{\mathbf{S}} . \quad (8)$$

If the decompressed image of image $\vec{\mathbf{X}}$ is defined as $\vec{\mathbf{X}}^{\mathrm{r}}$, then it can be written as

$$\vec{\mathbf{X}}^{\mathrm{r}} = \begin{cases} \vec{\mathbf{S}} + \vec{\mathbf{N}} + \vec{\mathbf{R}}_1 & \text{signal present} \\ \vec{\mathbf{N}} + \vec{\mathbf{R}}_2 & \text{signal absent} \end{cases} , \quad (9)$$

where $\vec{\mathbf{R}}$ is compression noise. We have previously derived a closed-form expression for $\vec{\mathbf{R}}$ and also shown that it has a normal distribution [10]. Since compression noise $\vec{\mathbf{R}}$ and background noise $\vec{\mathbf{N}}$ arise from different sources, this paper assumes that they are independent. If the total noise is defined as

$$\vec{\mathbf{Z}}_i = \vec{\mathbf{N}} + \vec{\mathbf{R}}_i , \quad (10)$$

then its mean vector and covariance matrix are

$$\mathrm{E}[\vec{\mathbf{Z}}_i] = \mathrm{E}[\vec{\mathbf{N}}] + \mathrm{E}[\vec{\mathbf{R}}_i] \quad (11)$$

and

$$\mathrm{Cov}(\vec{\mathbf{Z}}_i) = \mathrm{Cov}(\vec{\mathbf{N}}) + \mathrm{Cov}(\vec{\mathbf{R}}_i) . \quad (12)$$

where i = 1, 2.

If the original image has signal and background known exactly (SKE/BKE), then the total noise statistics of the decompressed image are known for a given compression algorithm. The signal can be the original or the decompressed signal for 2AFC experiments on reconstructed images. If $\vec{\mathbf{S}}$ is the signal in 2AFC, then we can show that the test statistic of CHO on reconstructed images is [11]

$$\lambda_{\text{rec}}(\vec{\mathbf{X}}^{\mathrm{r}}) = \vec{\mathbf{W}}^{\mathrm{T}}\vec{\mathbf{X}}^{\mathrm{r}} , \quad (13)$$

where $\vec{\mathbf{X}}^{\mathrm{r}}$ is reconstructed images and

$$\vec{\mathbf{W}}^{\mathrm{T}} = 2(\vec{\mathbf{S}} + \mathrm{E}[\vec{\mathbf{Z}}_2] - \mathrm{E}[\vec{\mathbf{Z}}_1])^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}(\mathbf{T}^{\mathrm{T}}(\mathrm{Cov}(\vec{\mathbf{Z}}_1) + \mathrm{Cov}(\vec{\mathbf{Z}}_2))\mathbf{T})^{-1}\mathbf{T} . \quad (14)$$

The detectability of CHO on reconstructed images is

$$d_{\text{rec}}^2 = \vec{\mathbf{W}}^{\mathrm{T}}(\vec{\mathbf{S}} + \mathrm{E}[\vec{\mathbf{Z}}_2] - \mathrm{E}[\vec{\mathbf{Z}}_1]) . \quad (15)$$

3. DECOMPRESSED JPEG LUMPY BACKGROUND IMAGES

Lumpy backgrounds are Gaussian blobs (structured noise) at random positions in the image. Due to their mathematical tractability and simplicity, such computer-simulated backgrounds that visually appear similar to real image backgrounds are widely used in medical image quality assessments [12]. Lumpy backgrounds are generated by filtering uncorrelated Gaussian images. The original images used in this paper are generated by the following steps: first, generate uncorrelated Gaussian images with zero mean; then calculate 2D-FFT of the generated images; multiply the Fourier coefficients by low-pass filter coefficients pixel by pixel. Finally, calculate the 2D inverse Fourier transform on filtered Fourier coefficients and take the real parts as lumpy background images. The lumpy background can be expressed in the following mathematical form:

$$\vec{\mathbf{X}} = \sqrt{p}\,\mathrm{Re}(\mathbf{F}^{-1}\mathbf{W}\mathbf{F}\vec{\mathbf{N}}), \tag{16}$$

where p is the power level of uncorrelated Gaussian noise, $\mathbf{F}$ and $\mathbf{F}^{-1}$ are respectively the forward and inverse Fourier transform matrix in a 1D representation, and $\mathbf{W}$ is a diagonal matrix whose diagonal elements are filter coefficients. $\vec{\mathbf{N}}$ is uncorrelated Gaussian noise images with zero means and unit covariance matrices. $\vec{\mathbf{X}}$ is lumpy background images. We can show that $\vec{\mathbf{X}}$ is jointly Gaussian distributed [10]. Its mean is zero, and its covariance matrix is

$$\mathrm{Cov}(\vec{\mathbf{X}}) = p\,\mathrm{Re}(\mathbf{F}^{-1}\mathbf{W}\mathbf{F})\,\mathrm{Re}(\mathbf{F}^{-1}\mathbf{W}\mathbf{F})^{\mathrm{T}}. \tag{17}$$

The left and right images of Figure 1 are the lumpy backgrounds with/without signal respectively. The mean vector and covariance matrices of the decompressed JPEG lumpy background can be derived from the closed-form compression noise statistics [10]. The JPEG image compression algorithm is DCT-based transform coding with uniform partition. Its block size is 8×8. Uniform scalar quantization is used for the quantization of the transform coefficients. A quantization table (QT) is transmitted as part of the encoded bit stream though it provides a default QT. The same QT will be used in all the blocks. The mean vector and covariance matrix of JPEG compression noise are

$$\vec{\mathbf{m}}^{\vec{\mathbf{R}}} = \mathbf{A}^{\mathrm{T}}\vec{\mathbf{m}}^{\vec{\mathbf{Q}}} \tag{18}$$

and

$$\mathrm{Cov}(\vec{\mathbf{R}}) = \mathbf{A}\,\mathrm{Cov}(\vec{\mathbf{Q}})\mathbf{A}^{\mathrm{T}} \tag{19}$$

respectively, where $\vec{\mathbf{m}}^{\vec{\mathbf{Q}}}$ and $\mathrm{Cov}(\vec{\mathbf{Q}})$ are mean vector and covariance matrix of JPEG quantization noise. $\mathbf{A}$ is JPEG transform matrix which is defined by

$$\mathbf{A} = \mathrm{Diag}_{\mathrm{N_b}}\left(\mathbf{A_{64}}, \mathbf{A_{64}}, \ldots, \mathbf{A_{64}}\right), \tag{20}$$

where Nb is the number of 8×8 in the original images; A_{64} is a 1-D transform matrix of 8×8 2-D DCTs; its (m,n) element can be calculated by

$$\mathbf{A}_{64}(m,n) = \mathbf{A}_{64}(8k+o, 8l+p) = \vec{\mathbf{u}}_k(o)\vec{\mathbf{u}}_l(p) \tag{21}$$

and

$$\vec{\mathbf{u}}_j(k) = \begin{cases} \sqrt{1/8} & k = 0 \\ \sqrt{2/8}\cos\left((2k+1)j\pi/(2\mathrm{N})\right) & k \neq 0 \end{cases}, \tag{22}$$

where j,k={0,1,…,7}, where m, n={0, 1,…,63) and k, l, o, p = {0,1,…,7}.

The mean vector and covariance matrix of JPEG quantization noise are functions of the marginal and pairwise PDFs of DCT coefficients of lumpy background images. The calculation can be found in [10].

4. SIMULATION RESULTS

The derived CHO performance is verified using lumpy background images with a circle and a Gaussian disk, each in a set of 2AFC experiments. The variable p in (17) is chosen as 128 for lumpy background generation. The size of the lumpy-background images is chosen as 32×32. The estimated ROC and AUCs were calculated using 2048 images.

The derived ROC and AUC on original lumpy background images are calculated by (2), (4) and (8) while the derived ROC and AUC on JPEG decompressed images are calculated by (13), (14) and (15). Since the signal used in the simulation is symmetrical, and background images are smooth, Laguerre-Gauss functions are chosen to define the channels in CHO simulation. The nth-order Laguerre-Gauss functions are defined as

$$LG_n(r) = 2\sqrt{\pi/a^2}\exp(-\pi r^2/a^2)L_n(2\pi r^2/a^2), \tag{24}$$

where a is width and Ln is nth-order Laguerre polynomial, defined as

$$L_n(x) = \sum_{m=0}^{n}(-1)^m\binom{n}{m}\frac{x^m}{m!}. \tag{25}$$

From left to right of Figure 2 shows first five 32×32 Laguerre-Gauss templates that are used in the simulation. The channel matrix T is constructed by putting the above 5 templates into column vectors and combining, yielding a 1024×5 matrix for T.

Table 1 shows derived and estimated AUCs on original and decompressed JPEG images. In Table 1 C10A10 is a circle signal with radius 10 and amplitude 10. Its derived and estimated AUCs on original images are 0.966. The derived and estimated AUCs on decompressed images against original signal in 2AFC tests are 0.947 and 0.945 respectively; they become 0.958 when the decompressed signal is used in 2AFC tests. Table 1 also lists AUCs for circle signal C10A5 and Gaussian signal G3A10 (standard deviation σ=3 and amplitude A=10) and G10A20 (σ=10 and A=20). All the cases show that the derived AUCs predict the corresponding estimated AUCs very well. Figure 3 shows CHO ROCs on original and decompressed JPEG lumpy backgrounds with a circle signal. In Figure 3, the solid line and star signs are the derived and estimated ROCs on original images. The dashed line and plus signs are the derived and estimated ROCs on decompressed images against original signal, while the dotted line and circle signs are the derived and estimated ROCs on decompressed images against decompressed signals. All the cases show that the derived and estimated ROCs are very close.

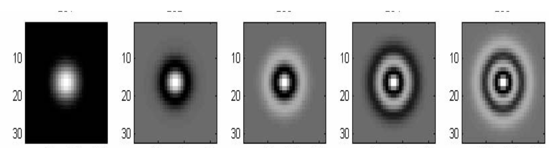

Figure 2. Laguerre-Gauss Channel Profile

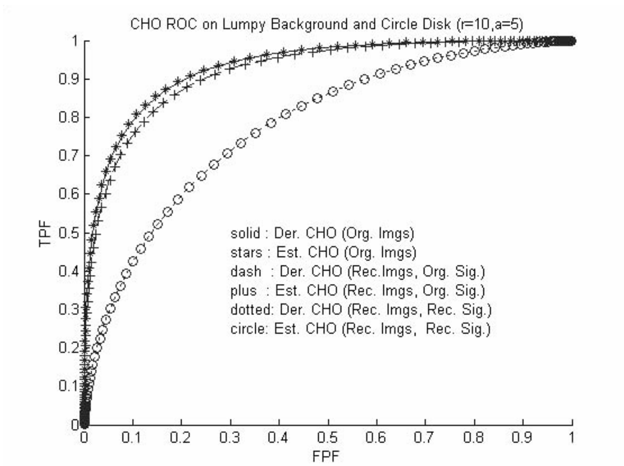

Figure 3. ROC Curves of CHO on Original and JPEG Lumpy Background Images with Circle Signal C10A5

Table 1. AUCs of Channelized Hotelling Observer on Original and Reconstructed Images

Images/ Signal	Original Images		Reconstructed JPEG Images			
	Original signal		Original signal		Reconstructed signal	
Signal	DER.	EST.	DER.	EST.	DER.	EST.
C10A10	0.966	0.966	0.947	0.945	0.958	0.958
C10A5	0.930	0.932	0.918	0.916	0.779	0.778
G3A20	0.966	0.966	0.947	0.945	0.958	0.958
G10A20	0.989	0.989	0.990	0.989	0.987	0.987

5. CONCLUSIONS

Many medical image quality assessment studies show that model-observer based image quality measures can effectively predict human performance on diagnostic tasks. Experimental results have also shown that model-observers can predict human performance on decompressed images as well. In both cases, model observers reduce or eliminated the need for costly and lengthy human-observer studies. The model observers do, however, require extensive examination of variations in data type, parameters, etc.

Based on the derived compression noise statistics, this paper derives the closed-form CHO on decompressed images. The detectability, an image quality measure, is shown, not unexpectedly, to be an inverse function of the covariance matrix of compression noise; i.e., detectability will decrease as compression noise increases for both ideal and channelized Hotelling observers.

The real benefit of these new closed-form image quality measures is that they allow users to calculate the performance of image compression algorithms without going through compression and decompression processes on extensive image samples and various parameter sets. It also is a theoretical approach that enables the user to choose compression parameters that predictably satisfy the tradeoff between the sizes of the lossily compressed image and the preservation of diagnostic information. Moreover, such closed-form quality measures of decompressed images provide a way to optimize compression algorithms subject to a model observer performance criterion. It also provides a theoretical foundation for efforts to create a model observer for decompressed images.

6. REFERENCES

[1] B. Schmanske and M. Loew, “Bit-plane-channelized Hotelling Observer for Assessing Lossy-compressed images”, *SPIE Medical Imaging 2003*, SPIE Vol. 3663, pp. 77-88.

[2] Kyle J. Myers, “Ideal Observer Models of Visual Signal Detection”, in J. Beutel, et al., eds., *Handbook of Medical Imaging*, Vol. 1. pp 559-592, SPIE Press, Bellingham, WA, 2000.

[3] M. Eckstein, C. Abbey, and F. Bochud, “Chapter 10: a practical guide to model observers for visual detection in synthetic and natural noisy images”, in J. Beutel, et al., eds., *Handbook of Medical Imaging*, Vol. 1., pp. 593-628, SPIE Press, Bellingham, Wash., 2000.

[4] A. Burgess, X. Li, and C. Abbey, “Visual signal detectability with two noise components: anomalous masking effects,” *J. Opt. Soc. Am. A*, Vol. 14, No. 9, pp. 2420-2442, Sept. 1997.

[5] C. Abbey, M. Eckstein, and J. Bartroff, “Observer performance for JPEG vs. wavelet image compression of x-ray coronary angiograms”, *Optics Express*, 1999.

[6] M. Eckstein, C. Abbey, et al., “The effect of image compression in model and human performance,” *Proc. SPIE* Vol. 3663, pp. 284-295, Feb. 1999.

[7] H. Barrett, J. Yao, J. Rolland, and K. Myers, “Model observers for assessment of image quality”, *Proc. Natl. Acad. Sci. USA* Vol. 90, pp. 9758-9765, Nov. 1993.

[8] K. J. Myers and H. H. Barrett, “The addition of a channel mechanism to the ideal-observer model,” *J. Opt. Soc. Am. A*, Vol. 4, 2447-2457, 1987.

[9] H. H. Barrett, C. K. Abbey, and E. Clarkson, “Objective image quality III. ROC metrics, ideal observers, and likelihood-generating functions,” *J. Opt. Soc. Am. A*, Vol 15, 1520-1535, 1998.

[10] D. Li and M. Loew, “Closed-form compression noise in images with known statistics,” *Proc. SPIE,* Vol. 5749, pp. 211-222, 2005.

[11] D. Li and M. Loew, “Model-observer based quality measures for decompressed medical images,” *Proc. 2004 IEEE Inl. Sym. Biomed. Imaging* April 2004, pp.832-835.

[12] B. D. Gallas, *Signal Detection in Lumpy Noise*, Ph.D. Dissertation, University of Arizona, 2001.