# METHODS FOR DETERMINING DIAGNOSTIC ACCURACY OF LOSSY COMPRESSED MEDICAL IMAGES

Artur Przelaskowski, Marian Kazubek, Tomasz Jamrógiewicz
Institute of Radioelectronics, Warsaw University of Technology
ul. Nowowiejska 15/19, 00-665 Warszawa, Poland
E-mail: arturp@ire.pw.edu.pl

*Summary:* *Some image quality measures are proposed in this paper. The most desired feature of these method is high correlation with diagnostic accuracy of compressed medical images. The chi-square measure (global and local) as a measure of statistical similarity of original and compressed image data, REOBD as a measure of blocking effect and diagnostic quality coefficient give vector measure of diagnostic accuracy of compressed images. Two ways of estimating the diagnostic quality coefficient are presented: objective method of diagnostic accuracy evaluation based on analysis of changes of the diagnostic parameters and proposed subjective (by physician) evaluation of diagnostic quality.*

Lossy compression is connected with necessity of determining acceptable compression ratios by taking into account suitable level of distortion. Ordinarily used rate-distortion theory could not be directly applied in medical image compression because the most important is diagnostic accuracy which could not be determined by computed cybernetics methods. Compression ratios achieved in lossy process are much more higher than the ones in lossless compression even though strict acceptance criterion (without discarding diagnostically valuable information) is used.

Lossy compression techniques are not accepted by many specialists and doctors because of the corruption of the data. The new directions of compression algorithm research (directed to improving the compression efficiency) should be closely connected with elaborating better methods for determining diagnostic accuracy of compressed images. Great effort of research should be performed in the area of characteristics of lossy compression distortion effects. Better objective measures of diagnostic important distortion should be developed. Good criteria and proper tools to construct compression algorithm, based on those measures, are searched. Vendors will not implement unrecoverable image compression until they are sure that they will not be held responsible for an alleged loss of information due to lossy compression beyond that which is caused by the acquisition and digitization scheme.

No single approach to quality and diagnostic accuracy measurement has gained universal acceptance, but three general approaches became dominant: computable objective distortion measures such as mean squared error (MSE) or signal-to-noise ratio (SNR), subjective evaluation based on psychovisual tests or questionnaires with numerical ratings, and clinical simulation and statistical analysis (e.g. ROC) [1].

Subjective ratings and ROC analysis are expensive, time-consuming, burdensome, and usually inaccurate or inconsistent. Objective measures do not have a good correlation with the viewers' response. It could be slightly improved by incorporating into the evaluation process some model of the human visual system (HVS). Also multi-dimensional measures like the Hosaka plots and Eskicioglu [2] charts are applied. They allow to obtain additional information about the loss of fidelity (i.e., type of degradation and distribution error) but they have also disadvantages: difficulty of the compression efficiency comparison and establishing the acceptable compression ratio level. Also they are not directly related to diagnostic accuracy.

We seek a multi-dimentional measure which allow to characterise a kind and a quantity of distortion and to evaluate diagnostic accuracy and acceptable compression level. We tested a vector measure with three items: chi-square measure, block effect measure and diagnostic accuracy coefficient.

Chi-square measure is defined as $c_q^2 = \sum_{i=1}^{NM} \frac{(f_i - f_i^{'})^2}{f_i}$, where $f_i$ - original image pixel value, $f_i^{'}$ - reconstructed image pixel value, $M \times N$ - image block size. We propose to use chi-square value as a measure of global statistical similarity of original and compressed image data - $c_{glob}^2$ and the similarity of local statistics - $c_4^2$ and $c_{16}^2$ (maximum value of chi-square value searched in respectively $4 \times 4$ and $16 \times 16$ blocks). We treated compressed image as a 2D random field generated by compression algorithm. An example of applying chi-square measure is presented on figure 1.

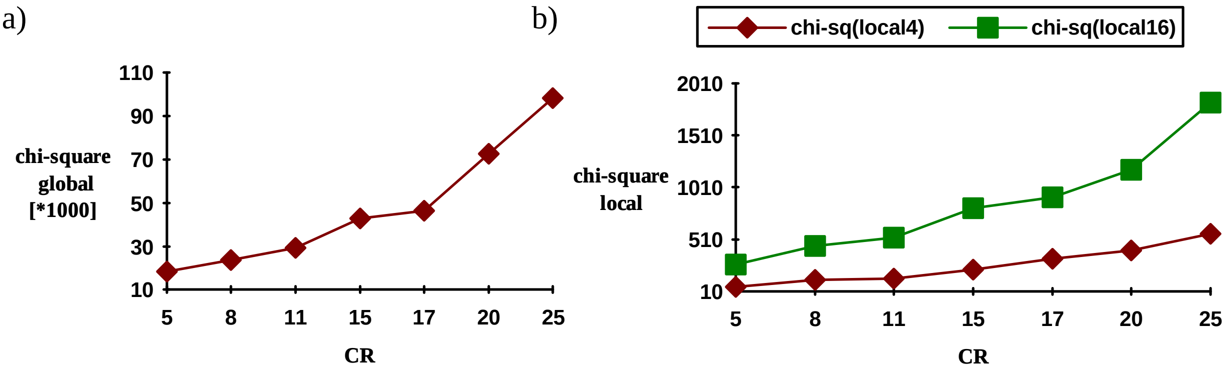


Fig. 1. Values of chi-square measure:chi-square global= $c_{glob}^2$,chi-sq(local4)= $c_4^2$,and chi-sq(local16)= $c_{16}^2$ in dependence of compression ratio CR. 8-bit head MR image was compressed with using Block DCT technique.

We propose to use relative EOBD [3] as a block effect measure - see figure 2, defined as follows:

$$REOBD = \{E[\Delta f(M,n)] + E[\Delta f(m,N)]\}^{1/2}$$

with $$\Delta f(M,n) = [|f(M,n) - f(M+1,n)| - |f^{'}(M,n) - f^{'}(M+1,n)|]^2,$$

$$\Delta f(m,N) = [|f(m,N) - f(m,N+1)| - f^{'}(m,N) - f^{'}(m,N+1)|]^2.$$

The diagnostic quality coefficient (DQC) is defined as a value in range of 0.0 to 1.0 which reflects the diagnostic accuracy of reconstructed image. The value 1.0 means that image is definitely unacceptable for diagnosis and value 0.0 means that the quality of compressed image is comparable or even better the quality of origin (of course diagnostic accuracy is fully preserved). Depending on medical imaging modality and kind of exams, the two ways of evaluating the DQC value is proposed. First, objective method of diagnostic accuracy evaluation is based on analysis of changes of the diagnostic parameters values calculated from original and reconstructed images. DQC is defined as:

$$DQC = \left|\frac{d_R}{d_M}\right|, \quad \text{for } d_R \le d_M; \qquad DQC = 1, \quad \text{for } d_R > d_M, \text{ where}$$

$d_R$ - error of diagnostic parameters evaluation from reconstructed image (in comparison to origin), $d_M$ - error of method of calculation of diagnostic parameters values from images.

From our research the DQC has a value less than 1.0 for compression ratios in range of 1 to 50 - for scintigraphy images (gated heart and thyroid examinations), and for compression ratios less than 30 for CT images used in radiotherapy (chest and brain tomograms).

The second method is subjective (by physician) evaluation of diagnostic quality of the compressed images. Critical criterion used in this method consists with 2 elements: to be acceptable means luck of noticeable degradation of image features which are important in diagnosis, and image quality - evaluated in psychovisual way - is comparable with origin (or better). The specialists work independently in clinical or comparable conditions and give binary answers for each tested image (acceptable or unacceptable). In this case DQC is counted as:

$$DQC = \frac{N_0}{N_B}, \quad \text{for } N_0 \le N_B; \qquad DQC = 1, \quad \text{for } N_0 > N_B, \text{ where}$$

$N_0$ - number of doctor's negations ('image could not be accepted'), $N_B$ - border number of doctor's negations which definitely could not be accepted.

The tests conducted in five medical centers (the results are presented in table 1) show that the compression ratios corresponding to DQC=0 are in range of 1 to 12 for MR images and in range of 1 to 14 for CT images.

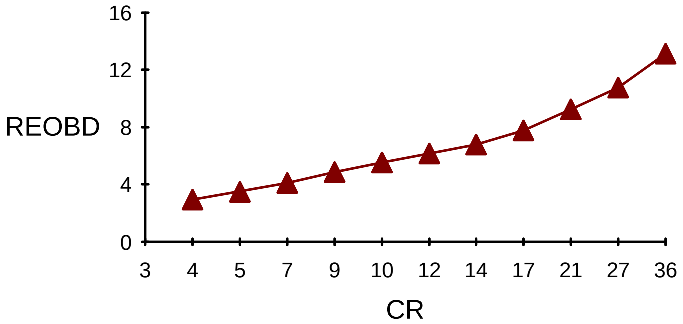


Figure 2. Values of REOBD in dependence of compression ratio CR. 8-bit head MR image was compressed with using Block DCT technique.

Table 1. The results of subjective evaluation tests conducted in 5 warsaw medical centers

| **ACCEPTABLE COMPRESSION RATIOS** | | | | | |
|---|---|---|---|---|---|
| MR images | | | CT images | | |
| brain | liver | breast | chest 1 | chest 2 | brain |
| 14 | 16 | 19 | 12 | 18 | 12 |
| 16 | 17 | - | 18 | - | 12 |
| - | - | - | - | - | 12 |
| **14** | | | **12** | | |

Conclusions: Proposed vector measure of diagnostic accuracy allows to evaluate diagnostically acceptable compression ratios and is more valuable than only subjective methods. The conducted are much more simpler than complex ROC - based techniques. Chi-square measure and REOBD are applied as the characteristic of compressed images distortion quality and quantity). They are easy for comparisons of the different compression techniques effectiveness. These two objective measures are related to main diagnostically important features of compressed images - fidelity of particulars and sharp edges ($c_4^2$ and $c_{16}^2$ ) and local or global structure shapes (REOBD and $c_{glob}^2$).

References:

[1] Cosman P. C., Gray R.M., Olshen R.A.: Evaluating quality of compressed medical images: SNR, subjective rating, and diagnostic accuracy, Proceedings of the IEEE, vol.82, no.6, p.919-932, 1994.

[2] Eskicioglu, A.M.: Application of multi-dimentional quality measures to reconstructed medical images, 8th IEEE Symposium on Computer-Based Medical Systems, Lubbock, 1995.

[3] Kazubek M., Przelaskowski A.,Jamrógiewicz T.: Quality Measurement of Compressed Medical Images: Block Effect Measures, submitted and accepted to NBCBE, Tampere, 1996.