



HAL
open science

Evaluation Of Medical Image Compression

Cécile Delgorge, Christophe Rosenberger

► **To cite this version:**

Cécile Delgorge, Christophe Rosenberger. Evaluation Of Medical Image Compression. Handbook of Research on Advanced Techniques in Diagnostic Imaging and Biomedical Applications, IGI Global, pp.447-459, 2009, 10.4018/978-1-60566-314-2.ch028 . hal-00989898

HAL Id: hal-00989898

<https://hal.science/hal-00989898v1>

Submitted on 11 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Evaluation of Medical Image Compression

C. Delgorge-Rosenberger
Université d'Orléans, France

C. Rosenberger
GREYC – ENSICAEN – Université Caen Basse Normandie – CNRS, France

ABSTRACT

The authors present in this chapter an overview on evaluation of medical image compression. The different methodologies used in the literature are presented. Subjective evaluation uses some a priori knowledge such as the judgment of experts or the ability to realize a correct diagnosis. Objective evaluation generally takes into account the value of metrics: the PSNR is an example of such a criterion. The goal of hybrid evaluation is to realize a reliable judgment while having a simple computation. The authors discuss on the benefits and drawbacks of these approaches. The European Project called OTELO in which they were involved, gives feedback on ultrasound image compression.

INTRODUCTION

Medical imaging is an important and a powerful technique whose goal is to facilitate the expert diagnosis. Many image processing algorithms can be used within this context such as: image filtering, compression, segmentation, interpretation or retrieval... One important issue concerns the evaluation of different image processing results

for the medical expert: as for example, image filtering can improve the image quality but can disturb the ability of a medical expert to make a diagnosis.

The proposed chapter deals with the particular field of the evaluation of medical image compression. Image compression for medical applications is an important topic as many image acquisitions are transmitted and stored

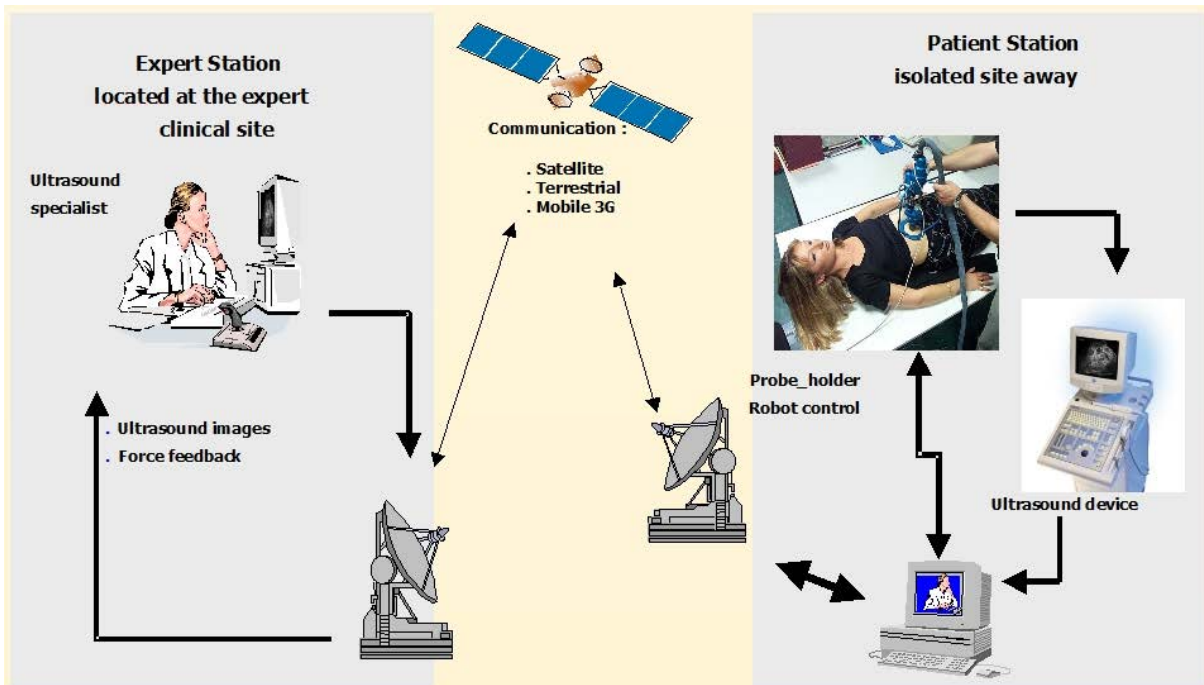
for a further analysis. In this context, we want to minimize the size of the compressed image while keeping a sufficient quality for the diagnosis. Several evaluation methods have been proposed in the state of the art. We propose in this chapter to make an overview of these approaches. We present different evaluation techniques for either an expert in image processing or for a medical expert. We discuss the advantages and drawbacks of each method. The European OTELO project (Delgorge *et al.*, 2005) in which we have been involved, provides a good experience feedback in the evaluation of medical image compression.

The main objective of the European OTELO project (mObile Tele-Echography with an ultra Light rObot) was to develop a robotic tele-echographic system. A light weight robot holds and moves a real probe on a distant patient according to the expert gesture and permits an image

acquisition using a standard ultrasound device (see Figure 1). Ultrasound images constitute the only feedback information available to the medical expert to remotely control the distant robotized system. The expert controls the remote probe holder robot by using a dedicated input device and based on the quality of the received information. The diagnosis made by the specialist strongly depends on the quality of these images. An important task also concerns the evaluation of the quality of the compressed images. Many experimental results are presented in this chapter in order to illustrate the behaviors of the different evaluation methods.

We can distinguish three types of evaluation methods in the state of the art. The first one concerns the subjective evaluation. The quality of a compression result, for any medical types of images, is traditionally evaluated by considering a visual test where many experts examine a large

Figure 1. the OTELO tele-echographic system



set of images and score each one based on their quality or the ability to make a correct diagnosis. Second, we present the objective evaluation methods. Many statistical criteria have been proposed in the literature to automatically evaluate different compression results of a single image. We propose to compare many of them and discuss their efficiency. Subjective and objective evaluation methods have many advantages and drawbacks. These two approaches are complementary. That is why many works propose an hybrid approach. These methods are presented in the main trust of this article. The goal of these methods is to make a reliable judgment (similar to a medical expert) while using some statistical criteria to make the evaluation of a large set of compression results possible. The future trends in the domain are then proposed and a conclusion is given.

BACKGROUND

Image compression is an important issue in medical imaging as the distant visualization of medical images is now possible through high bandwidth networks for different applications (discussion between experts on a difficult case, storage of medical images of a patient...) and as

telemedecine becomes an emergent technology nowadays (Delgorge et al., 2005).

Image compression is an image processing algorithm whose objective is to decrease the size of storage of the image while preserving as much as possible its visual quality (see Figure 2). Even if this definition is quite simple, the main problem is to evaluate the quality of a compression result. As for example, the result given in Figure 2 is clearly not very good but it is difficult to say if this quality would be satisfactory for a medical expert to make a diagnosis.

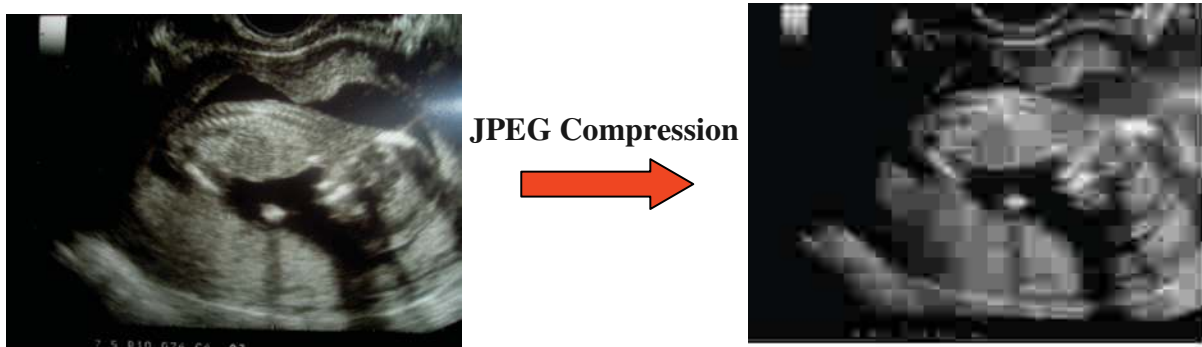
We present in the following sections two classical approaches for the evaluation of medical image compression results.

Subjective Evaluation

The evaluation of image processing results can be realized in a subjective way. It can be done through different approaches:

- The quality of a result can be visually appreciated by an user on a benchmark image. The well known « Lena » image is an example. All researchers in image processing can rather easily compare the results they

Figure 2. Example of a JPEG compression result (ratio 5%) of an ultrasound image



obtain using their own algorithm with some from the state of the art;

- It is also possible to use a synthetic ground truth in order to compare the results of different algorithms. As for example, synthetic images can be used in order to evaluate the quality of a medical image segmentation result;
- The use of a phantom (synthetic material) also permits to evaluate some image processing results while keeping all acquisitions artifacts;
- Another solution consists in asking as many experts as possible to quantify the quality of the obtained results. For medical applications, many approaches were defined in the literature such as the simple visualization of results, the visual comparison of results or the performance quantification in term of ability to realize a correct medical diagnosis.

Nowadays, the two first methods are rarely used because the achieved evaluation reliability is poor. The two last methods are more and more used with many alternatives. We present in this section, the possible solutions and some associated feedbacks.

In the multimedia world, the quantification of the subjective quality perceived by an observer remains the only reference as regards to the quality of images (Klein, 1993). The International Consultative Committee of Radio communication (CCIR) defined multiple standardized evaluation methods to define the image quality. The principle of a subjective test is contingent upon a standardized structure, which consists of an observers group, a quantification method for the appreciation of the visual quality (i.e. a subjective scale of quality notation), a set of test images, a display material, a test room and a protocol defining the notation process of the images as well as the experimental conditions.

The three most current tests are the following (International Telecommunication Union, 2002):

- **Method with one stimulus:** It makes it possible to evaluate many imagery systems. A set of images degraded by various treatments is presented in a random order. The appreciation of the image quality is given by the expert without any reference image. The scale of notation can be defined or built by the observer itself as the test advances. Figure 3 shows an example of such a subjective evaluation with 5 possible scores (from unacceptable to very good). The main problem of this approach is that it requires a certain number of images for an observer in order to stabilize the scale of scores. Note also that the original image is not given, that is not very comfortable for the observer.
- **Method with double stimulus:** Images are evaluated per pair: an original image (the reference) and a compression result. The observer must score this result according to the preset scale. In general, two categories of scales are proposed: the first translates the feeling of the visual quality of the image (that can be considered as a criterion of “nice image”), the second translates the degree of degradation (that is a criterion of “good image” for the application).
- **Comparative method:** Different images representing the multiple treatments we want to compare are presented at the observer. The comparison is realized considering the different compression results. The observer must sort all of them considering the alteration associated to the original image. Figure 4 gives an example where an expert is asked to sort 5 compression results given the original image.

Figure 3. Example of an evaluation using only one stimulus

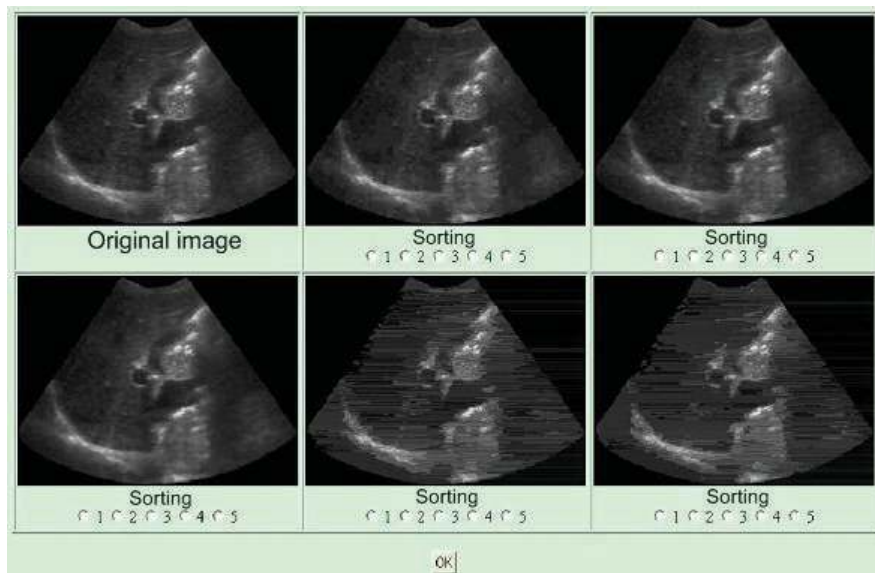


A large quantity of recommendations is published (International Telecommunication Union, 2002). They relate mainly to:

- the environment of the test: displaying conditions, the visualization material, distance of the expert to the screen,
- the resolution and contrast of the screen,
- the image database used for the test,
- observers (or experts): they must be at least 15, their vision must be measured and corrected if necessary, their experiment must be raised (specialty, age),
- the protocol of the test: it must be precisely explained to the observers,
- the duration of test: it must last at least 30 minutes,
- the analysis and the presentation of the results.

This kind of tests implies that all the experts follow a very rigorous protocol. This approach requires important human and financial fundings. This is why, it is not rare to find in the literature some studies taking as a starting point these recommendations but not following them all completely (Naegele-Jackson *et al.*, 2002; Azpiroz-Leehan *et al.*, 2004). Moreover, such qualitative and subjective evaluations depend of

Figure 4. Example of an evaluation using the comparative approach



the medical specialty of the expert, its motivation, its availability, its tiredness or its interest in such tests.

Receiver Operating Characteristic (ROC) curves provide many tools to analyze the behavior of models, algorithms or human judgments. It can be used to compare different compression methods by considering the medical application (Weatherburn *et al.*, 2003). In order to choose a compression method or a parameter, we can evaluate the number of cases corresponding to a correct diagnosis. Figure 5 shows an example of ROC curves. We look for a compression method that maximizes the correct positive rate while minimizing the false positive rate. In this case, the Test A has to be preferred than the other one.

Objective Evaluation

A subjective evaluation permits to take into account the medical expertise. The main drawback of this approach is that these tests are very time and manpower consuming. In the literature, many statistical criteria are available and offer a simple tool to evaluate the quality of a compression result according to the original one.

Statistical Criteria

There are several methods to evaluate an image quality. The quality of a compression result can be represented thanks to the *pixel distance* between the compression result and the original image. The most known distance is the Minkowski one. From this measure, we obtain many distances: the Euclidean distance, the Manhattan distance, the Chebychev distance and the mean square error (MSE). These measures can be calculated considering the neighborhood of a pixel and not only with one pixel (Tamtaoui *et al.*, 1999). In the image processing literature, the most frequently used measures are the mean square error (MSE) and the peak signal to noise ratio (PSNR) (Yang, 2005; Zhong, 2005).

The similarity between two images can be translated thanks to *correlation measures*. Several criteria have been proposed in (Linfoot, 1958) and are based on the power spectral density: in particular, the fidelity, the structural content and the normalized cross correlation. The Czekanowski coefficient (Andreautos *et al.*, 1998) measures the correlation between two images considering parameters extracted from both of them.

Figure 5. Example of ROC curves

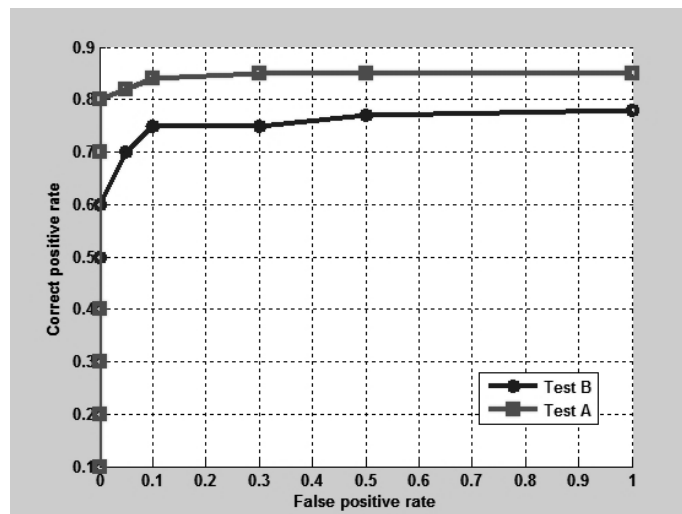
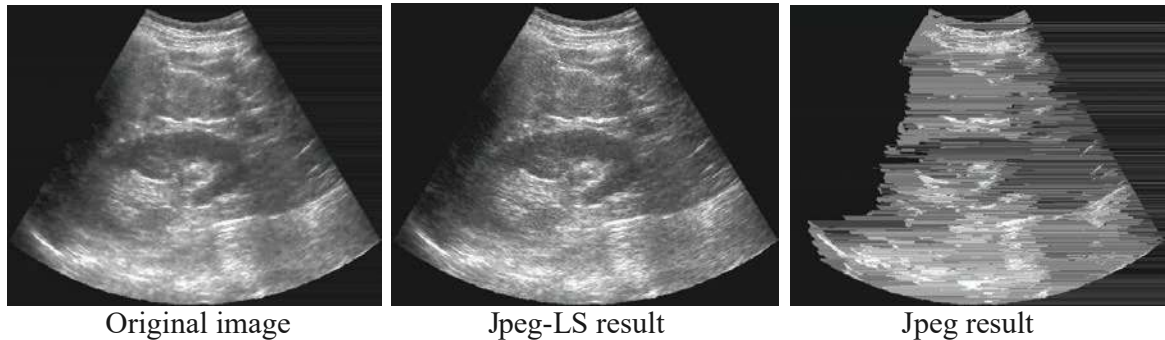


Figure 6. Comparison of two compression results with same PSNR



The above mentioned distances can be calculated in the *spectral* field. The goal is also to isolate the image distortion on particular frequency bands or to separate the distortions due to the spectral phase or magnitude (Avcibas *et al.*, 2002).

T-J Chen proposed in (Chen 2003) a new quality measure based on the Moran *statistics*. The Moran I coefficient translates the studied area clearness, it measures a structural distortion and not a gray level pixel variation. The compressed image statistical properties are used in (Turaga *et al.*, 2004) to evaluate the distortion due to the compression.

Some *graphical criteria*, based on the histograms of images, have been proposed. The graphical distance also measures the dissimilarity between two histograms or uses the difference image histogram to represent a fidelity measure between the original image and the compression result (Sundersingh, 2000).

Discussion

The main advantage of objective criteria is the simplicity of computation. It is then possible to automatize the evaluation of a large set of compression results. The important drawback of all these statistical criteria is the fact that they do not always correspond to the human visual sys-

tem (HVS) – representing the observer’s visual perception.

The definition of the distance measures depends on the kind of degradations introduced in the image by the compression algorithm. The problem of a graphical measure based on the histogram is the fact that the histogram is not a single function: the same histogram can represent two completely different images.

Moreover, the MSE and PSNR distance measures are very popular due to their simplicity of mathematical definition, even if many studies showed that these criteria are not enough relevant. As for example, Figure 6 shows the bad reliability of the PSNR measure. Two different compression methods (obtained by using the Jpeg-LS and Jpeg algorithms) give two compression results with different compression rate but a similar PSNR value (PSNR=32.5dB). We can clearly see that the Jpeg result quality is visually worst than the Jpeg-LS one, whereas the PSNR evaluates as equal their quality. In this example, we understand that the PSNR is not an efficient criterion for a medical application.

In the OTELO project framework, we studied some statistical criteria and compared them with respect to the results of an expert evaluation. We selected several criteria of different types according to some previous works in the literature (Avcibas *et al.*, 2002): distance mea-

Table 1. Statistical criteria chosen for the OTELO study

<i>D1</i>	Minkowski distance - Mean absolute error
<i>D2</i>	Minkowski distance - Mean square error
<i>D3</i>	Minkowski distance - Modified infinity no
<i>D4</i>	Neighborhood error - 8 neighbors
<i>D5</i>	Neighborhood error - 24 neighbors
<i>D6</i>	Multi-resolution error
<i>C1</i>	Normalized cross correlation
<i>C2</i>	Image fidelit
<i>C3</i>	Czekonowski correlation
<i>S1</i>	Spectral phase error
<i>S2</i>	Spectral phase-magnitude error
<i>S3</i>	Block spectral magnitude error
<i>S4</i>	Block spectral phase error
<i>S5</i>	Block spectral phase-magnitude error
<i>S6</i>	Block spectral error
<i>PI</i>	Peak signal to noise ratio

asures, denoted Dx ; correlation measures Cx ; spectral measures Sx ; PSNR measure PI (see Table 1). The three best criteria are $D5$, $S2$ and $S1$, reaching for $D5$ a maximal value of 65.3%. This means that this criterion is able to reproduce the ability of a medical expert to compare two

compression results in **65.3%** of the cases. One can notice that the PSNR criterion, often used for the comparison of compression results, ranks only at the ninth place.

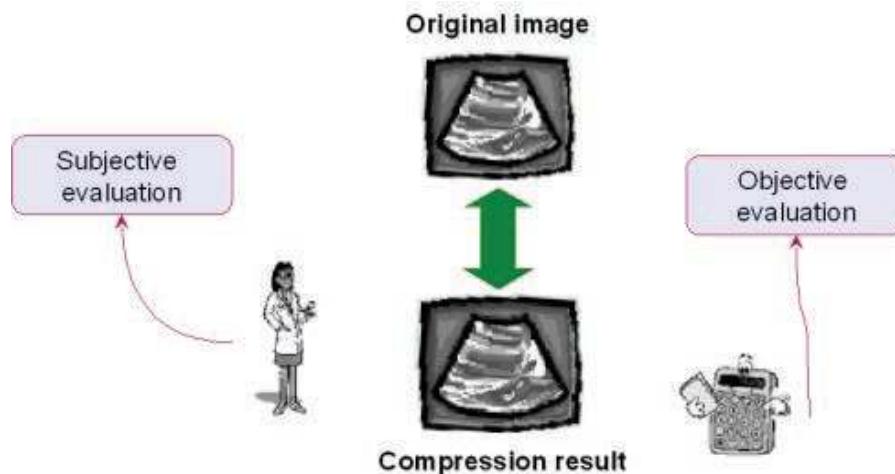
MAIN FOCUS OF THE CHAPTER

We focus in this section on the hybrid evaluation approach. This methodology has for objective to realize a reliable judgment while keeping a simple computation (see Figure 7).

Human Visual System (HVS) Based Criteria

The objective of these criteria is to give some evaluation results close to those we would obtain with a subjective evaluation while having the same calculation simplicity as the objective criteria. Since 1950, many models have been proposed with the goal to simulate the human visual system (HVS), representing the visual perception. The first HVS models were based on an approach called single channel. In this case, the HVS is considered as a simple space filter

Figure 7. Hybrid evaluation



whose characteristics are defined by a function of sensitivity to contrast.

Le Callet and Barba (Le Callet & Barba, 2001) proposed on this principle, a criterion for the evaluation of a compression result. It carries out the combination of perceptual errors between the original image and the compression result to obtain a quality score. The performances of the criterion (evaluated using a subjective evaluation) allow a benefit of about 30% of effectiveness (measured by the coefficient correlation factor) compared to the PSNR. Mathematical models of observers were also developed for the nuclear medicine (Pommert & Höhne 2002). The compression result, the error image or a simple statistical criterion are then balanced by one of these models. The definition of artificial models of observers concerns a rather long and expensive procedure and is dependent on the type of tested images.

Some studies model the HVS and then define some criteria from this model (Carnec, 2004). One can also find in the literature many criteria proposing to combine statistical metrics and some characteristics of the HVS. In the following study (Miyaji *et al.*, 2000), the SNR is balanced by local factors reflecting visual perception (the effect of mask, frequential characteristics). When some specificities of the image are taken into account, one can note that the correlation of the criterion with subjective quality increases (compared to the only use of statistical metrics). The error of evaluation noted by this HVS based criterion decreases about 30% compared to the PSNR.

For the OTELO project, we studied also some HVS based criteria: human visual system based measures Hx ; contrast measure TI (see Table 2). The $T1$ and $H1$ criteria gave good results, obtaining respectively the second and fifth place considering the results of all the statistical and HVS based criteria. The $T1$ criterion is able to reproduce the ability of a medical expert to compare two compression results in **65.3%** of the cases.

Fusion of Evaluation Criteria

The models proposed in the literature to represent the HVS are either too simple or too complex: they do not represent in an efficient way the HVS and they are too complex to obtain a criterion easily usable. In order to mitigate these various disadvantages, some works propose to combine several statistical measures (Cane, 1997) and bring them closer to the HVS (see Figure 8). One can find some works on combinations of criteria by linear regression, analysis of the variance and construction of a Kohonen chart (Avcibas *et al.*, 2002) or with a genetic algorithm (Olivès, 1998; Delgorge *et al.*, 2006).

A possible method for the fusion of evaluation criteria consists in combining linearly the best ones. A combined criterion can be written as follows:

$$C_{FUSION} = \sum_{i=1}^N a_i \cdot C_i \quad (1)$$

Where a_i are coefficients that permit to define the importance of the criterion C_i in the computation of C_{FUSION} and to take into account its variation interval. These coefficients have to be determined by optimization by considering some evaluation examples given by experts. The value N corresponding to the number of criteria to fuse and can be set by the user. The choice of criteria to fuse can be done according to their efficiency (used alone) or by selecting them in the optimization process (Delgorge *et al.* ; 2006).

Table 2. HVS based criteria chosen for the OTELO study

$H1$	Absolute norm Human Visual System
$H2$	L2 norm Human Visual System
$H3$	Similarity
$H4$	DCTune error
$T1$	Contrast measure

Figure 8. Fusion of evaluation criteria

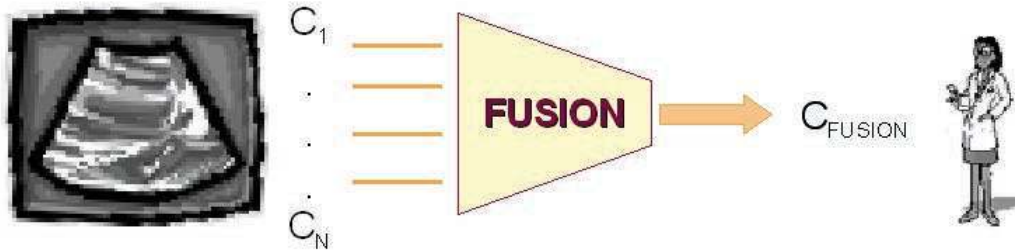
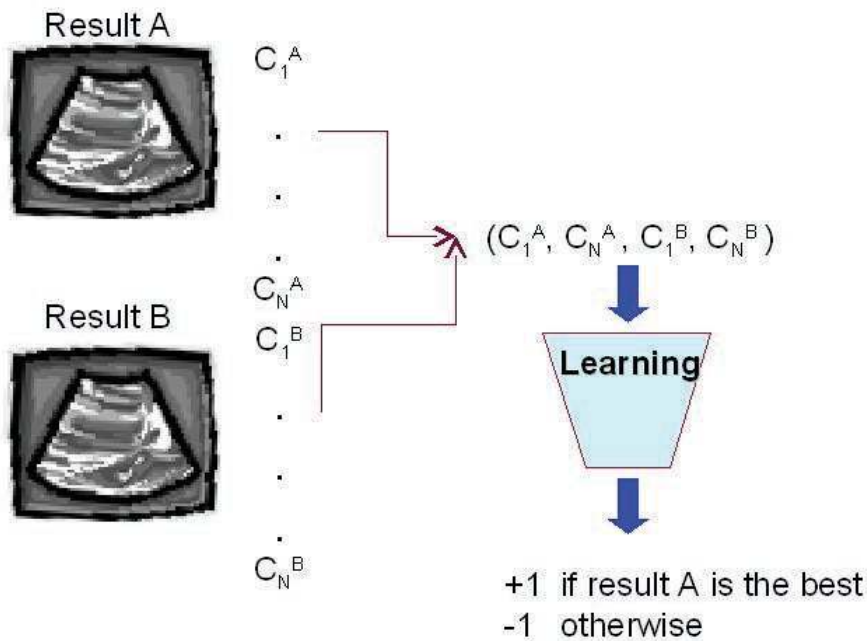


Figure 9. One solution to learn the medical expertise



The optimization method can be a classical linear method (such as simplex) without selection of criteria to fuse. In the other case, a genetic algorithm can be used.

This last approach has been applied in the OTELO project, in order to improve the evaluation of compression results quality and to perform better than the statistical criteria. We combined the minimal number of statistical criteria while obtaining an evaluation as close as possible to

the medical judgment. The highest similarity rate of correct comparison obtained is **75.3%** for the fusion of 9 selected criteria. The selection process of criteria to fuse permits also to determine complementary ones.

Medical Expertise Learning

The fusion of statistical criteria can be done according to another approach: the learning of the

medical assessment data. In this case, the objective is to define an algorithm that is able to compare two compression results similarly as a medical expert. In the previous section, the decision function for the comparison of two compression results was set (as for example, induced by the choice of a linear combination). With this new approach, the decision function is estimated thanks to a learning algorithm (see Figure 9). The support vector machine (SVM) has been proposed as the learning algorithm for the OTELO application (Delgorge *et al.*, 2006).

Suppose we have two compression results A and B we want to compare. We compute on each compression result N different evaluation criteria (that can be selected previously). As for example, for the compression result A, we have a vector C_1^A, \dots, C_N^A where C_i^A is the value of the criterion number i. We have so two vectors for each compression result we can merge in a single one denoted $COMP = C_1^A, \dots, C_N^A, C_1^B, \dots, C_N^B$. This pattern represents the comparison of the compression results A and B. We can use some judgments given by experts on the comparison of different compression results as reference. Given this reference, if the result A is better judged than the result B by experts, the COMP vector must lead to an exit of the comparison algorithm equals to 1.

In the OTELO project experiments, when 95% of the whole data set is used in the learning database, the system obtains a successful recognition rate of **92.8%**, for the fusion of 5 criteria computed for each of the two compression results to compare.

Discussion

As for us, hybrid evaluation is nowadays the best approach to quantify the performance and the quality of an image processing result. The definition of a single evaluation tool for any medical applications is impossible. It is so necessary to take into account the particular context of the application and also different evaluation approaches.

This is also the case for different domains such as paper reviewing. The evaluation of a paper is based on different criteria (clarity, adequacy to the literature review, contribution...) that are taken into account by the evaluator based on its experience and expertise.

We presented three possible methods in the medical image compression. Table 3 sums up the results we obtained in the context of the OTELO project. Note that the best objective evaluation criterion gave a performance equals to 65,3% based on medical experts judgment. The medical expertise learning shows clearly its benefit.

FUTURE TRENDS

Image processing is now a mature technology. The evaluation of a processing result is an important task especially in medical imaging. Different approaches have been proposed in the literature. Nevertheless, few techniques are used for the validation of a new algorithm. Generally, the PSNR is used to compare multiple compression results even if different studies showed that this evaluation criterion does not give a reliable judgment.

We think that two main future trends have to be considered.

The first one concerns the analysis of evaluation methodologies. Even if some new criteria are proposed in the literature, none validation are generally provided. The systematic comparison of evaluation criteria is necessary to increase the reliability of the judgment that is realized. It will have also an impact on the progress in defining new image compression methods.

Table 3. Hybrid evaluation results

Approaches	Performance
Human Visual System criterion	65,3%
Fusion of criteria	75,3%
Medical expertise learning	92,8%

The second issue concerns the proposition of new hybrid evaluation criteria. They combine the advantages of objective and subjective evaluation methods. An hybrid evaluation criterion, through integrating some visual perception considerations, is an elegant and interesting solution. The expertise learning is another solution which is similar to a medical expert training. New criteria can be developed integrating many aspects such as the medical application, the type of image or the type of displaying tool.

CONCLUSION

Evaluation of medical image compression is a great challenge as this processing is more and more used for storage and telemedicine applications. The difficulty for the medical application is that a medical expert does not see the same thing than a non expert. The main interest of a medical expert is to quantify in which measure the processing will facilitate its diagnosis. A non expert will appreciate the quality of an image considering the alterations of the compression result. Many studies have to be done in order to define a metric that embeds this kind of information.

ACKNOWLEDGMENT

This work was funded by the European Commission under OTELO project (IST 2001-32516).

REFERENCES

- Andreutos, D., Plataniotis, K. N., & Venetsanopoulos, A. N. (1998). Distances Measures for Color Image Retrieval. *IEEE ICIP International Conference on Image Processing*.
- Avcibas, I., Sankur, B., & Sayood, K. (2002). Statistical evaluation of image quality measures. *Journal of Electronic imaging*, 11(2), 206-223.
- Azpiroz-Leehan, J., Leder, R., & Lerallut, J.-F. (2004). Quantitative and qualitative evaluation of filter characteristics for wavelet packet compression of MR images. *In 26th Annual International Conference of the Engineering in Medicine and Biology Society*, (pp. 1537-1540).
- Cane, B. Mc (1997). On the Evaluation of Image Segmentation Algorithms. *DICTA'97 and IVCN'Z'97*, (pp. 455-459).
- Carnec, M. (2004). *Critères de qualité d'images couleur avec référence réduite perceptuelle générique*. Phd Thesis de l'Ecole Polytechnique de l'Université de Nantes.
- Chen, T.-J., Chuang, K.-S., Wu, J., Chan, S. C., Hwang, I.-M., & Jan, M.-L. (2003). A novel image quality index using Moran I statistics. *Physics in Medicine and Biology*, 48, 131-137.
- Delgorge, C., Courreges, F., Al Bassit, L., Novales, C., Rosenberger, C., Smith-Guerin, N., Bru, C., Gilabert, R., Vannoni, M., Poisson, G., & Vieyres, P. (2005). A tele-operated mobile ultrasound scanner using a light weight robot. *IEEE Transactions on Information Technology in BioMedicine, special issue mHealth*, 9(1).
- Delgorge, C., Rosenberger, C., Poisson, G., & Vieyres, P. (2006). Towards a new tool for the evaluation of the quality of ultrasound compressed images. *IEEE Transactions on Medical Imaging*, 25, 1502 – 1509.
- International Telecommunication Union (2002). ITU-R-BT.500-11 Methodology for the subjective assessment of the quality of television pictures. *Recommendation on Broadcasting Service*.
- Klein, S. A. (1993) Image quality and image compression: a psychophysicist's viewpoint. *Digital and human vision*, (pp. 73–88).
- Le Callet, P., & Barba, D. (2001) Critère de qualité visuelle d'images couleur - application à l'évaluation réaliste de codeurs d'images. *Colloque GRETSI*.

Linfoot, E. H. (1958). Quality Evaluation of Optical Systems. *Optica Acta*, 5(1-2), 1-13.

Miyaji, S., Hamada, T., & Matsumoto, S. (2000). Development of a Digital Compressed Picture Quality Assessment System Considering Human Visual Perception. *Systems and Computers in Japan*, 5, 97–109.

Naegele-Jackson, S., Holleczeck, P., Rabenstein, T., Maiss, J., & Sackmann, M. (2002). Influence of compression and network impairments on the picture quality of video transmissions in telemedicine. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, (pp. 2060–2068).

Olivès, J.-L. (1998). *Optimisation globale d'un système imageur à l'aide de critères de qualité visuelle*. PhD thesis de l'Ecole Nationale Supérieure de l'Aéronautique et de l'Espace ENSAE.

Pommert, A., & Höhne, K.H. (2002). Evaluation of Image Quality in Medical Volume Visualization: The State of Art. *Lecture Notes in Computer Science*, 2489, 598–605.

Sundersingh, B.J. (2000). *Qualitative Evaluation of Data Compression in Real-Time Ultrasound Imaging*. PhD Thesis University of Tennessee, Memphis.

Tamtaoui, A., Ribagnac, M., & Aboutajdine, D. (1999). Elaboration de critères de mesure de qualité objective pour la qualité subjective. *Colloque GRETSI*, (pp. 845-848)

Turaga, D. S., Chen, Y. & Caviedes, J. (2004). No reference PSNR estimation for compressed pictures. *Signal Processing Image Communication*, 19, 173-184.

Weatherburn, G. C., Ridout, D., Strickland, N. H., Robins, P., Glastonbury, C. M., Curati, W., Harvey, C., & Shadbolt, C. (2003). A comparison

of conventional film, CR hard copy and PACS soft copy images of the chest: Analyses of ROC curves and inter-observer agreement. *European Journal of Radiology*, 47, 206-214.

Yang S.-B. (2005). Smooth side-match weighted vector quantizer with variable block size for image coding. *IEEE Proceedings Vision and Image and Signal Processing*, 152, 763-770.

Zhong, J., & Ning, R. (2005). Image Denoising based on wavelets and multifractals for singularity detection. *IEEE Transactions on Image Processing*, 14, 1435-1447.

KEY TERMS

Fusion: Combination of different data in order to improve decision making.

Hybrid Evaluation: The goal of this approach is to obtain a judgment as reliable as the subjective one while having an easy computation.

Human Visual System (HVS): It refers to the visual perception of humans that is simulated by researchers in evaluation.

Objective Evaluation: It is a quantitative evaluation generally based on statistical criteria. None *a priori* knowledge is used for the evaluation.

Peak Signal-to-Noise Ratio (PSNR): The PSNR is most commonly used as a measure of quality of reconstruction in image compression.

Receiver Operating Characteristic (ROC): ROC curves provides tools to analysis the behavior of models, algorithms or human judgments.

Subjective Evaluation: It is a quantitative or a qualitative evaluation involving experts or some *a priori* knowledge.