Quality Analysis for Coded Images with Loss

Lic. Hugo D. Ramón
Full-Time Practical Classes Coordinator, Faculty of Computer Sciences, U.N.L.P.
e-mail: hramon@lidi.info.unlp.edu.ar.

MSc. Claudia C. Russo
Full-Time Co-Chair Professor, Faculty of Computer Sciences, U.N.L.P.
e-mail: crusso@lidi.info.unlp.edu.ar.

Eng. Armando De Giusti
Principal Researcher CONICET, Full-Time Chair Professor, Faculty of Computer Sciences, U.N.L.P.
e-mail: adegiusti@lidi.info.unlp.edu.ar.

Laboratorio de Investigación y Desarrollo en Informática - LIDI
Facultad de Informática - Universidad Nacional de La Plata
Calle 50 y 115, 1er. piso. La Plata (1900). Bs. As. Tel-fax (54) 221 422 77 07
http://lidi.info.unlp.edu.ar

Abstract

A considerable compression rate can only be achieved by means of algorithms with loss, which means that it is not possible to recover the exact original image. This loss of information may have a direct relation to the loss of quality, as well as cause problems related to reliability depending on the area of the application.

The main question to answer is: How do we decide that an image compressed with loss is suitable to be used for a given application?

This question is answered by defining a subjective quality assessment and then relating it to objective values in order to find - by means of a statistic analysis of the data - a relationship between quality and compression ratio.

In particular, the incidence of the type of histogram and its relation to subjective loss and compression ratio are studied.

This paper is related to the Magister Thesis “Análisis del Error en Algoritmos de Transmisión de Imágenes Comprimidas con Pérdida” ("Error Analysis in Transmission Algorithms of Images Compressed with Loss") by Lic. Ramón.

Key Words: Images, Compression, Quality, JPG, Subjective Model, Statistics, Histogram.

Objective Quality Measures

If we are interested in assessing the fidelity of a reconstructed sequence, it is natural to observe the difference between the original values and the reconstructed ones; or, in other words, the distortion introduced by the compression process. Two known measures of the distortion or difference between the original and the reconstructed sequences are the Mean Square Error (1) and the Absolute Difference (2); these are usually called *difference distortion measures*.

$$d(x, y) = (x - y)^2 \qquad (1)$$

$$d(x, y) = |x - y| \qquad (2)$$

In general it is difficult to examine the difference on a term-to-term basis. This is the reason why a set of measures of the average is used to summarize the information in the sequence of differences. The most commonly used average measure is that of square errors, called *mean square error* (3), or, if we are interested in the size of the error in

$$\sigma^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - y_n)^2 \qquad (3)$$

$$SNR(dB) = \frac{\sigma_x^2}{\sigma_d^2} \qquad (4)$$

relation to the signal, the ratio between the mean square value of the source output and the *mse*, which is called *signal-to-noise ratio* (SNR) (4)

There are other objective measures that can be collected by taking into account entropy (5) and conditional entropy.

$$H=-\sum_{image} p(i)\log_2(p(i)) \qquad (5)$$

$$H(X/Y)=-\sum_{image} p(i)\log_2(p(i)) \qquad (6)$$

If X represents the original image and Y the decompressed one, the definition of the conditional entropy of X/Y is the amount of information of the original image that can be known by knowing the reconstructed image only. This can be interpreted by saying that, as we get farther away from the original, there is more uncertainty despite the fact that the reconstructed image is known.

Subjective Quality Measure

The subjective quality of a reconstructed image can be carried out by means of several approaches. A set of random images can be presented to experts or typical users to evaluate them using a scale from 1 to 5. Next, a statistical analysis can be carried out in order to highlight averages, variations and other tendencies in the data.

Formal subjective tests, such as mean Opinion Score and Diagnostic Acceptability Measure are common in the areas of speech processing and audio compression (Quackenbush 1988).

There are considerable variations as regards the numeric range of the answers, the inclusion or not of descriptive set phrases for each value, and in the intention that the numeric range indicates the use of an image for a given application. In general, the classification used asks the user to evaluate the images according to the similitude degree or dissimilitude between two images. This last concept is the one taken for the development of our experience.

Experience Developed

Our main purpose was to draw conclusions about 4 different issues:

1- The relationship existing between the objective measures and the results of the subjective measures by expert and non-expert viewers.

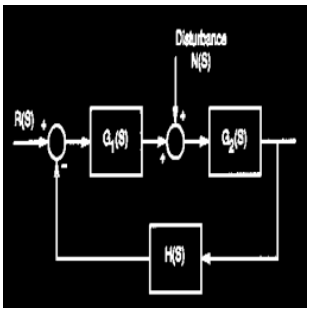
2- The relationship existing between the subjective assessment of the loss with the type of histogram of the image.

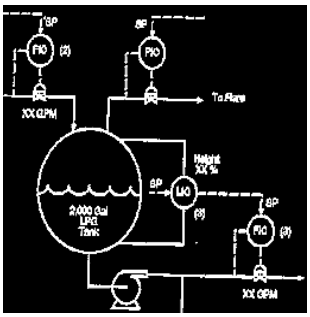
3- To study if there are differences in subjective assessment depending on the image within the same class.

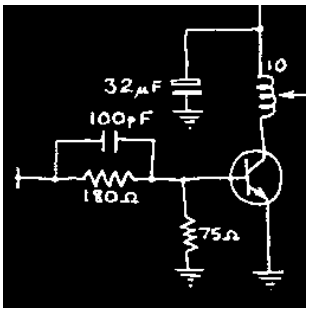
4- To study the incidence of the categories *expert* and *non-expert* in the subjective assessment of quality.

To define subjective assessment, three groups of independent images were selected: the first group included faces, the second group included landscapes - both in gray levels - and the third one included electronic circuits (bi-level).

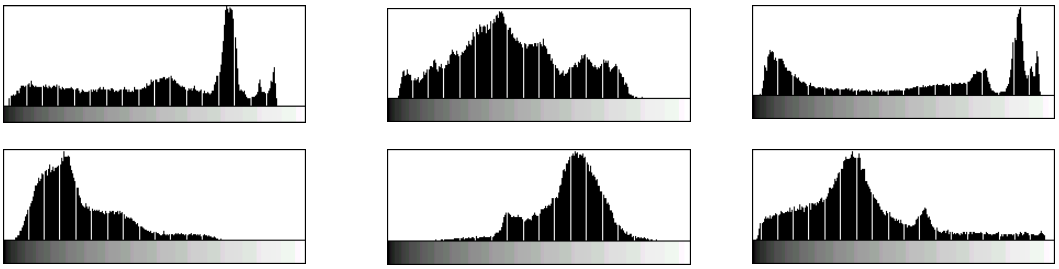
Each group is composed of three images with related histograms or similar detail levels. The originals are below.







The types of images were grouped according to a study of their histograms; the histograms corresponding to the groups of images selected for the study are below.



It should be noted that the histograms of the last set of images (the circuits) are two bars at levels 0 and 255, since these are black and white images. This kind of images are interesting because they are very sensitive to the mirror effect produced by compression schemes.

Each image was compressed with the compression scheme presented by JPEG (Wallace 1991), making three runs with varying thresholds between 3 and 16.

Objective Evaluation

These results present the information corresponding to the objective evaluations; table 1 corresponds to the class "faces", table 2 corresponds to the class "formulas", and table 3 to the class "landscapes".

Table 1 - Class "Faces"

	Ent	E/C	ECM	SNR		Ent	E/C	ECM	SNR		Ent	E/C	ECM	SNR
CARA1	7.58				**CARA2**	7.51				**CARA3**	7.45			
_D1	7.53	3.12	5.50	35.85	**_D1**	7.52	3.50	8.86	31.76	**_D1**	7.37	3.23	7.47	35.26
_D2	7.26	3.59	11.52	32.64	**_D2**	7.51	3.97	17.77	28.74	**_D2**	7.08	3.63	14.80	32.29
_D4	6.91	4.10	26.38	29.04	**_D4**	7.33	4.47	38.44	25.39	**_D4**	6.62	4.10	31.63	28.99

Table 2 - Class "Formulas"

	Ent	E/C	ECM	SNR		Ent	E/C	ECM	SNR		Ent	E/C	ECM	SNR
FORM1	**1.02**				**FORM2**	**0.57**				**FORM3**	**0.34**			
_D1	**1.79**	**0.20**	**31.63**	**28.99**	**_D1**	**1.99**	**0.04**	**55.89**	**20.15**	**_D1**	**1.44**	**0**	**37.18**	**20.54**
_D3	**2.25**	**0.22**	**68.25**	**17.55**	**_D3**	**2.69**	**0.06**	**283.3**	**13.10**	**_D3**	**1.99**	**0.001**	**185.1**	**13.56**
_D4	**1.93**	**0.23**	**89.69**	**16.36**	**_D4**	**2.27**	**0.06**	**359.5**	**12.07**	**_D4**	**1.62**	**0.002**	**226.1**	**12.70**

Table 3 - Class "Landscapes"

	Ent	E/C	ECM	SNR		Ent	E/C	ECM	SNR		Ent	E/C	ECM	SNR
PAIS1	6.96				**PAIS 2**	6.93				**PAIS 3**	7.49			
_D1	6.96	4.61	50.74	20.38	**_D1**	6.91	4.42	39.68		**_D1**	7.50	5.04	98.27	20.47
_D3	6.95	5.08	103.9	17.27	**_D3**	6.77	4.87	78.65	24.79	**_D3**	7.45	5.42	185.8	17.70
_D4	6.90	5.20	125.7	16.44	**_D4**	6.61	4.99	95.11	23.97	**_D4**	7.39	5.50	216.3	17.04

Subjective Assessment

A software system was developed to carry out the evaluation in two different groups of users. The first group was composed by 7 people with experience in handling and processing digital images. The second group was composed by 20 students of the course of studies on Computer Sciences with little or no experience in image treatment. This software allows to classify each of the three images according to a deterioration scale: 1-Very annoying, 2-Annoying, 3-Slightly annoying, 4-Perceptible but not annoying, 5-Imperceptible.

The original image is clearly marked on the screen, and the distribution of the compressed images is at random in order not to introduce any noise in the classification process. The histogram of each image can also be checked in order to make a more detailed classification.

A statistical evaluation of the averages of the interest output variable, called *image quality*, was carried out over the total amount of evaluators (expert and non-expert). The average was used to normalize the answer variable and thus be able to use the variance analysis by means of a linear model within the presupposed concepts (Montgomery 1991).

Conclusions

The purpose of this paper was to define different error measures and then linearly relate them to perceived quality by means of a subjective measure of a set of images. The results are shown according to the objectives of the task.

Objective 1: ***The relationship existing between the objective measure and the results of the subjective assessments by experts and non-experts.***

There is a clear correlation between the proposed subjective scale and the objective measures. In particular, it can be observed that the conditional entropy is a good approximation for the classes "faces" and "landscapes", where there is histogram amplitude.

Objective 2: ***The relationship existing between the subjective assessment of the loss with the kind of histogram of the image.***

From the experimental work carried out, it is clear that for similar compression indices, the subjective assessment of the loss is related to the type of histogram. The classification in three classes proposed in the Thesis is not absolute and a future research work to be done is the analysis of alternative classifications.

Objective 3: ***To study if there are differences in subjective assessment within a class depending on the image.***

No significant differences were found for the different images within the same class of histogram.

Objective 4: ***To study the incidence of the categories "expert" and "non-expert" on the subjective assessment of quality.***

After analyzing the results, it can be concluded that for a pre-set classification, the category "non-expert" is more dispersed than the category "expert".

The results, images, data and graphics, as well as this full paper, can be downloaded from http://lidi.info.unlp.edu.ar

Bibliography

1. Algazi V. R. (1992), “Comparison of Image Coding Techniques with a Picture Quality Scale”. *SPIE* Vol. 1771, *Applications of Digital Image Processing* XV, 1992.
2. Budrikis Z. L. (1972), “Visual Fidelity Criterion and Modeling”. *Proceedings of the IEEE* Vol. 60 Nr. 7, July 1972.
3. Campbell F. W. (1980), “The physics of visual perception”. *Roy. Soc., London, series B* vol. 290, 1980.
4. Cosman P. (1994), “Evaluating Quality of Compressed Medical Images: SNR, Subjective Rating and Diagnostic Accuracy”, *Proceedings of the IEEE* Vol. 82, Nr. 6, June 1994.
5. Cygansky D. (1998), “Information Coding”. www.ece.wpi.edu
6. Davisson L. D., (1981), “Efficient Universal Noiseless Source Codes”. *IEEE Transaction Information Theory*, Vol. IT-27, 1981
7. Fish R., (1991), “A Subjective Visual Quality Comparison of NTSC, VHS and Compressed DS1-compatible Video”. *Proc. SID (Society for Information Display)*, vol. 32 nr.2, 1991.

8. Glassner A.S., (1995, VOL 1), “Principles of Digital Image Synthesis”. *Morgan Kaufman Publishers, Inc.*, Vol. 1 1995.
9. Glassner A.S., (1995, VOL 2), “Principles of Digital Image Synthesis”. *Morgan Kaufman Publishers, Inc.*, Vol. 2 1995.
10. Gray R. M., (1990), “Entropy and Information Theory”. *Springer-Verlang*, 1990.
11. Hall E., (1982), “Evaluating Image Quality: A new formulation computer”. *Bio. Med.* Vol. 2, 1982.
12. Help G., (1991), “Data Compression. Techniques and Applications”. *Hardware a Software Considerations.* John Wiley & Sons, 1991.
13. Huang T. S., (1971), “Image Processing”. *Proc. IEEE* vol. 59 nr. 11, 1971.
14. Jense J., (1996), “Introductory Digital Image Processing”. *A Remote Sensing Perspective.* Prentice Hall, 1996.
15. Luckiesh M., (1965) “Visual Illusions”. 1965.
16. Lukas F., (1982), “Picture Quality Prediction Based on a Visual Model”. *IEEE Transactions on Communications*, Vol. COM-20, Nr. 7, July 1982.
17. Lynch T. J., (1985), “Data Compression: Techniques and Applications”. *Lifetime Learning Wadsworth*, 1985.
18. Meyer P., (1970), “Introductory Probability and Statical Applications”. *Addison-Wesley Publishing*, 1970.
19. Nelson M., (1991), *The Data Compression Book*, Prentice Hall, 1991.
20. Netravalli A. N., (1988), *Digital Pictures: Representation and Compression*, Plenum, 1988.
21. Quackenbush S., (1988), *Objective Measures of Speech Quality*, Prentice Hall Signal Processing Series, 1988.
22. Sayood K., (1996), *Introduction to Data Compression*, Morgan Kaufman Publishers, Inc., 1996.
23. Shannon C. E., (1948), “A Mathematical Theory of Communications”. *Bell System Technologies Journal*, Vol. 27, 1948.
24. Stockham T., (1972), “Image Processing in the Context of a Visual Model”. *Proceedings of the IEEE* Vol. 60, Nr. 7, July 1972.
25. Tucker H., (1962), *An Introduction to Probability and Mathematical Statistics*, Academic Press Inc., 1962.
26. Wallace G., (1991), “The JPEG Still Picture Compression Standard”. *Multimedia Engineering. Digital Equipment Corporation*. Maynard, MASS, 1991.
27. Wang S., (1992), “An Objective Measure for predicting Subjective Quality of Speech Coders”. *IEEE Journal Sel. Areas Commun*. Vol. 10, June 1992.