A MEASURE FOR IMAGE QUALITY

Le Phu Dung, Srinivasan Bala, Mohammed Salahadin,
Kulkarni Santosh and Wilson Campbell
Faculty of Computing and Information Technology
Monash University, Australia.
{pdle,srini,salah,santosh,cwilson}@broncho.ct.monash.edu.au

Key Words: Multimedia, image quality measure, image compression.

ABSTRACT

This paper describes a new measure for calculating the error that has been introduced when an image is modified which can be used to compare the quality of images. It is based on the following principle:

The quality of an image should be measured locally and globally. This implies that:

(1) Since the change of each pixel value affects the quality of the image, and the overall quality is dependent on the total number of pixels that were changed, the total number of pixels that were changed should hence be used in evaluating the quality of the image.
(2) The change of each pixel value first affects the quality of a small part of the image that directly contains the pixel, and the change of that part of the image affects the quality of the whole image. This suggests that image quality should be evaluated part by part rather than pixel by pixel.

Experimental results show that this proposed measure performs better than the signal-to-noise ratio and colour-based methods. It not only works well when the other two methods work but also works where the other two methods fail.

1 INTRODUCTION

Multimedia systems need to handle diverse data types that include text, numeric data, audio, image and video, etc. Data such as image and video require a great amount of memory space for storing and more time for transmitting hence they need be compressed or changed [9]. The high demand for reducing storage and transmission time for still and moving images has made image and video compression research a very active area. Many successful image compression algorithms have been devised and used, such as image compression based on Discrete Cosine Transform JPEG [10, 13], sub-band coding [14], fractal image compression [3, 7, 8, 11], and wavelet image compression [2, 5, 12]. When the original image is modified (whether by lossy compression or by other image processing applications), the modified image often needs to be compared with the original image to determine how much distortion has occurred or how much quality is lost. For instance, in image compression, the reconstructed image is compared to the original in terms of quality and compression ratio to evaluate the efficiency of the compression algorithm. The most commonly used method for measuring the quality of a modified image against the original is the signal-to-noise ratio measure. Unfortunately it is well-known that this method does not always work and as we will show why it does not always work in the later sections of this paper.

There are other methods for measuring image quality such as the colour-based technique [1], however none of the reported methods gives the correct answer for all cases. There are at least two factors which contribute to the difficulty in finding a complete algorithm:
(1) There are many different kinds of noises and

each can affect the quality of the image differently. (2) It is not simple to mathematically prove the quality of an image without human judgement.

This paper describes a new method for measuring the quality of images. The method is based on the idea that it will be more accurate if:

(a) image quality is evaluated by unit, such as a square, that contains more than one pixel,

(b) the number of pixels that are changed is taken into account since the quality of the whole image is dependent on it, and

(c) the total pixel-value difference is taken into account since such amount affects the quality of the image.

In the following sections, we first describe the class of measures that can possibly be used for evaluating image quality, then describe the proposed measure, its implementation and the experimental results of the method in comparison to the other two methods.

2 IMAGE QUALITY MEASURES

Image quality or error measures use as their basis the *residue* image, that is, the modified image subtracted from or compared to the original by some mathematical means. All the measures try to map the residue image to some quantity that is expected to have the following properties:

(i) not negative,
(ii) zero only when the original and modified images are identical, and
(iii) monotonically increasing as the modified image *looks worse*.

In looking for such a measure, we may be confronted with the question 'With what method should the error be measured?', since there are many possible choices.

The most common is the class of measures that use the $L^p(A)$ norm which is defined by:

$$\| f - g \|_{L^p(A)} = \left(\int_A |f(x) - g(x)|^p dx\right)^{\frac{1}{p}}$$

where $0 < p < \infty$.

This class of measures, including the mean-absolute (when $p = 1$) and mean-square (when $p = 2$) errors as special cases, gives much flexibility for selecting a quality or error measure. However, currently there has no way to prove which value of p is best for use in computing image quality or the error. Devore et al. [6] showed by examples that with $p = 1$ the attempt to minimize the error can lead to more pleasing pictures than $p = 2$. However, such result does not imply that the measure with $p = 1$ is a correct method for evaluating image quality or computing errors.

3 THE PROPOSED METHOD

In this section we describe the new method, called *Image Quality Measure Error* (IQME), for measuring the quality of images.

Let μ be the Lebesgue measure on R^2 and $\int$ denote the Lebesgue integral.
Let f and g, without loss of generality we assume that f and g are continuous on A = [a,b]*[c,d] $\subset R^2$, in turn represent the original and modified image:

$$f,g : A \longrightarrow [0,\infty)$$

For any $x \in A$ and $0 < e_x \in R$ and $E_{e_x} = [x - e_x, x + e_x] * [x - e_x, x + e_x] \bigcap A$, we define the local absolute mean error (*lame*) function as follows:

$$lame(\mathrm{x}) = \frac{\int_{E_{e_x}} |f(x)-g(x)|}{\mu(E_{e_x})}.$$

Let $\vartheta = \{x \in A\colon \mathrm{f(x)} \neq \mathrm{g(x)}\}$.

Since the quality of the modified image is affected by ϑ and the total difference between f and g, we call such quality effects together *norm_effect*, defined by:

$$norm\_effect = \begin{cases} \frac{\mu(A)}{\mu(\vartheta)} * \int_A lame(x)dx, & \text{if } \mu(\vartheta) > 0\ ; \\ 0, & \text{otherwise.} \end{cases}$$

Each element of ϑ first affects the quality of a small part of the image that directly contains that element, and logically any change of that part of the image affects the quality of the whole image. We call such a factor *part_effect*.

For any $x \in A$ and $0 < e_x \in R$ and $E_{e_x} = [x - e_x, x + e_x] * [x - e_x, x + e_x] \bigcap A$, there exist constants $m_f^{e_x}$ and $m_g^{e_x}$, by the mean-value theorem, such that:

$$m_f^{e_x} = \frac{\int_{E_{e_x}} f(x)dx}{\mu(E_{e_x})}.$$

$$m_g^{e_x} = \frac{\int_{E_{e_x}} g(x)dx}{\mu(E_{e_x})}.$$

To compute *part_effect*, we first compute the $pmld_f^{e_x}$ and $pmld_g^{e_x}$, called *partial mean local deviations*, for each E_{e_x}. They are defined as follows:

$$pmld_f^{e_x} = \frac{\int_{E_{e_x}} (f(x) - m_f^{e_x})dx}{\mu(E_{e_x})}.$$

$$pmld_g^{e_x} = \frac{\int_{E_{e_x}} (g(x) - m_g^{e_x})dx}{\mu(E_{e_x})}.$$

The *part_affect* is then defined as:

$$part\_affect = \frac{|\int_A pmld_f^{e_x} - \int_A pmld_g^{e_x}|}{(\int_A pmld_f^{e_x} + \int_A pmld_g^{e_x})^2}.$$

Now the image quality measure error is defined as:

$$IQME = norm\_affect * part\_affect$$

4 THE IMPLEMENTATION

This section describes one of the possible implementations of the proposed measure.

Let $\{x_n\}$ and $\{y_n\}$ represent the original and the modified images in discrete form respectfully.

Given an odd number $m \in N$ and $m > 1$, for every pixel p select the greatest square of odd size l such that $l \leq m$ and p is the central pixel of that square. Since the size of the images is n, there are n such squares, including the particular case that we used in this paper: the smallest square is of size 1 consisting of 1 pixel for the special case where the pixel is on the boundary.

The *norm_effect* is calculated as follows:

$$norm\_effect = \frac{n}{\vartheta} * (\sum_{i=1}^{n}(\frac{1}{k}\sum_{j=1}^{k}|x_j - y_j|))$$

where ϑ is the total number of pixels of the original image that are different from the corresponding pixels of the modified image, and $\frac{1}{k}\sum_{j=1}^{k}|x_j - y_j|$ is the local absolute mean error (*lame*).

Let $m_x^{s^i}$ and $m_y^{s^i}$ in turn be the mean values of the square s^i of $\{x_n\}$ and $\{y_n\}$, the original and the modified images. They are defined as:

$$m_x^{s^i} = \frac{1}{k}\sum_{j=1}^{k} x_j$$

$$m_y^{s^i} = \frac{1}{k}\sum_{j=1}^{k} y_j$$

and the partial mean local deviations are defined as:

$$pmld_x^{s^i} = \frac{1}{k}\sum_{j=1}^{k}(x_j - m_x^{s^i})$$

$$pmld_y^{s^i} = \frac{1}{k}\sum_{j=1}^{k}(y_j - m_y^{s^i})$$

where k is the size of the square.

The *part_effect* is calculated as follows:

$$part\_effect = \frac{|\sum_{i=1}^{n} pmld_x^{s^i} - \sum_{i=1}^{n} pmld_y^{s^i}|}{(\sum_{i=1}^{n} pmld_x^{s^i} + \sum_{i=1}^{n} pmld_y^{s^i})^2}$$

The image quality measure error is:

$$IQME = norm\_effect * part\_effect$$

5 EXPERIMENTAL RESULTS

We carried out two intensive tests on a number of images. In the first test, the original images are compressed by a good algorithm from [5] at different compression ratios. The results showed that in the normal case when signal-to-noise ratio and colour-based methods work, the proposed method also works. Some of the results for selected images and compression ratios are listed in table 1.

In the second test, modified images are generated either by applying an unreliable compression algorithm (we used several unreliable algorithms in the second test. These algorithms are originally based on fractal, sub-band and transform coding. They are modified to either work with one format but not with others, or perform well for smooth images with low compression ratio but do poorly with complicated images or with high compression ratio. These modified algorithms can produce unexpected results.) to the original images, or by inserting different kinds of noises into the original images. The results of this test showed that the proposed method works well while the other two methods either do not work at all or inconsistent. Some selected results are listed in table 2.

Image	*Compression Ratio*	*PSNR*	*RMSE*	*CDIF*	*IQME*
Lenna	4:1	43.70	1.664277	0.000013	0.003211
"	8:1	39.41	2.727389	0.000020	0.015584
"	16:1	36.17	3.960342	0.000030	0.035185
"	32:1	33.17	5.596424	0.000036	0.065199
"	64:1	30.22	7.855745	0.000044	0.121023
Goldhill	4:1	41.12	2.240863	0.000015	0.002339
"	8:1	35.93	4.070622	0.000018	0.016405
"	16:1	32.62	5.962691	0.000022	0.047030
"	32:1	30.07	7.991423	0.000030	0.114563
"	64:1	28.21	9.901258	0.000035	0.205744
Barbara	4:1	40.59	2.379967	0.000017	0.001740
"	8:1	34.55	4.772578	0.000019	0.007755
"	16:1	29.53	8.505284	0.000020	0.031316
"	32:1	26.63	11.874612	0.000031	0.078412
"	64:1	24.30	15.533400	0.000041	0.182962
Mandrill	4:1	32.46	6.072964	0.000554	0.004020
"	8:1	27.70	10.506272	0.000559	0.026107
"	16:1	24.80	14.671946	0.000562	0.077234
"	32:1	22.86	18.336484	0.000570	0.171743
"	64:1	21.58	21.245285	0.000572	0.309748

Table 1: Selected experimental results for the three methods for the normal case.

Image	*Changed by*	*PSNR*	*RMSE*	*CDIF*	*IQME*
L2	*GoodComp* 8:1	39.41	2.727389	0.000020	0.015584
L3	*GoodComp* 64:1	30.22	7.855745	0.000044	0.121023
L4	Insertion of noise	20.16	25.013991	0.000228	0.000448
L5	*Unr1Comp* 4:1	11.83	65.279321	0.000020	0.236275
M2	*GoodComp* 4:1	32.46	6.072964	0.000554	0.004020
M3	Insertion of noise	33.14	5.617281	0.000002	0.297950
M4	*Unr2Comp* 23:1	14.22	49.578578	0.000544	1.141327

Table 2: Selected experimental results showing poor performances of signal-to-noise ratio and colour-based methods, while IQME has good performance

We include peak-signal-to-noise ratio (PSNR), root-mean-square error (RMSE), the colour-based (CDIFF) which is based on [1], and IQME results in the tables for comparison. For PSNR, the larger the value of PSNR the better the quality of the image, and for the others, RMSE, CDIFF and IQME, the smaller the value the better the quality (smaller error size implies better quality).

The selected modified images are listed in table 2 and re-produced (figure 1 to 7) in the following pages. These images correspond to the popular Lenna and Mandrill images, and are generated by the above-mentioned good algorithm [5], called *GoodComp*, and two selected unreliable algorithms of those above-mentioned, called *Unr1Comp* and *Unr2Comp*. The selected modified images are:

L2 : Lenna's reconstructed images, compressed by *GoodComp* algorithm with compression ratio 8:1.
L3 : Lenna's reconstructed images, compressed by *GoodComp* algorithm with compression ratio 64:1.
L4 : Lenna's modified image with the insertion of noise. This image looks very much the same as the original.
L5 : Lenna's reconstructed image, compressed by *Unr1Comp* algorithm.
M2 : Mandrill's reconstructed image, compressed by *GoodComp* algorithm with compression ratio 4:1.
M3 : Mandrill's modified image with the insertion of noise.
M4 : Mandrill's reconstructed image, compressed by *Unr2Comp* algorithm.

Figure 1: Lenna's reconstructed image compressed by *GoodComp* 8:1, L2

Figure 2: Lenna's modified image with the insertion of noise, L4

Figure 3: Lenna's reconstructed image compressed by *GoodComp* 64:1, L3

Figure 4: Lenna's reconstructed image compressed by the unreliable algorithm *Unr1Comp* 4:1, L5

Figure 5: Mandrill's reconstructed image compressed by *GoodComp* 4:1, M2

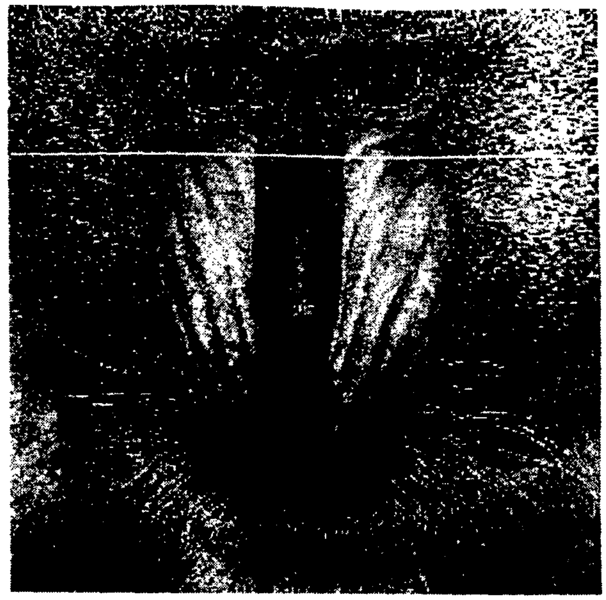

Figure 7: Mandrill's modified image with the insertion of noise, M3

Figure 6: Mandrill's reconstructed image compressed by the unreliable algorithm *Unr2Comp* 23:1, M4

Any image quality measure, objective or subjective, has to finally use human perception in evaluation. From the visual inspection of the printed images (figures 1 to 7) one can observe that both the signal-to-noise ratio and colour-based methods do not work when images L3 and L4 are compared. They also do not work when images M2 and M3 are compared. In addition, the colour-based method does not work when images L2 and L5 are compared or when images M2 and M4 are compared.

When two images are compared in terms of quality, one desires a measure that has the following properties:
(i) The measure should parallel the human visual system with the expectation that the differences in quality judged by human eye to be large are also mathematically large.
(ii) If the differences are insignificant to the human eye, the error size should be small.

Although signal-to-noise ratio method is not a perfect technique, it is still the choice in image compression field due to its simplicity. However, it does not always work as shown by [4] and by our experimental results.

The colour-based technique is claimed to be more accurate than the signal-to-noise method. It works well when signal-to-noise method works, however, it also suffers similar problems to the signal-to-noise ratio method, as shown by our experimental results.

Our experiments showed that the signal-to-noise and colour-based methods do not always satisfy the above two properties while the proposed method exhibits those properties.

6 CONCLUSIONS

We have proposed an algorithm for measuring image quality when a modified image is compared against the original. The main idea is to take into consideration the effect of the change of each pixel value to the local area which contains that pixel and the effect of that area to the whole image. Also the number of pixels that were changed should also be taken into account since the overall quality of the image depends on it. Since the units used in this method can contain more than one pixel and they can be overlapped, computationally this method is more complicated than the other two. However, our experiments on different sets of test images showed that the proposed algorithm performs better than the two most commonly used methods. It not only works well when the other two methods work, but also works when the other two methods perform poorly.

REFERENCES

[1] Annamalai M., Sundaram A. and Bhargava B., 'A Colour-Based Technique for Measuring Visible Loss for Use in Image Data Communication', *Proceedings of the 7th International Conference (DEXA96)*, pp. 39-48, 1996.

[2] Antonini M., Barlaud M., Mathieu P. and Daubechies I., 'Image Coding Using Wavelet Transform', *IEEE Transactions on Image Processing*, Vol. 1, No. 2, pp. 205-220, 1992.

[3] Barnsley M. F. and Hurd L. P., *'Fractal Image Compression'*, Wellesley, MA: AK Peters, Ltd., 1993.

[4] Bernd G., 'What's Wrong with Mean-squared Error?', *In Digital Images and Human Vision*, (Ed.) Andrew B. Walson, MIT Press, pp. 207-220, 1993.

[5] Davis G. M., 'Self-quantization of Wavelet Subtrees', *Proceedings of the SPIE Wavelet Applications II*, Vol. 2491, pp. 141-152, 1995.

[6] DeVore R. A., Jawerth A. B. and Lucier B. J., 'Image Compression Through Wavelet Transform', *IEEE Transactions on Information Theory*, Vol. 38, No. 2, pp. 719-746, 1992.

[7] Fisher Y., *'Fractal Image Compression: Theory and Applications to Digital Images'*, (Ed.) Fisher Y., Springer Verlag, 1994.

[8] Jacquin A. E., 'Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations', *IEEE Transactions on Image Processing*, Vol. 1, No. 1, pp. 18-30, 1992.

[9] John B. F. K., *'Multimedia Systems'*, Ed. John B. F. K., ACM Press, Addison-Wesley, Reading, MA., 1994.

[10] Le Gall D., 'MPEG: A Video Compression Standard for Multimedia Applications', *Communications of the ACM*, 34(4), pp.46-58, 1991.

[11] Monro D. M. and Dudbridge F., 'Fractal Approximation of Image Blocks', *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 485-488, 1992.

[12] Sharpio J., 'Embedded Image Coding Using Zerotrees of Wavelet Coefficients', *IEEE Transactions on Signal Processing*, 1993, Vol. 41, No. 12, pp. 3445-3462.

[13] Wallace G. K., 'The JPEG Still Compression Standard', *Communications of the ACM*, 34(4), pp.31-44, 1991.

[14] Westerink P., Biemond J., Boekee D., and Woods J. W., 'Sub-band Coding of Images Using Vector Quantisation', *IEEE Transactions on Communications*, 36(6), pp. 713-719, 1988.