A Fuzzy Image Metric With Application to Fractal Coding

Junli Li, Gang Chen, and Zheru Chi, *Member, IEEE*

Abstract—**Image quality assessment is an important issue addressed in various image processing applications such as image/video compression and image reconstruction. The peak signal-to-noise ratio (PSNR) with the L^2-metric is commonly used in objective image quality assessment. However, the measure does not agree very well with the human visual perception in many cases. In this paper, a fuzzy image metric (FIM) is defined based on Sugeno's fuzzy integral. This new objective image metric, which is to some extent a proper evaluation from the viewpoint of the judgment procedure, is closely approximates the subjective mean opinion score (MOS) with a correlation coefficient of about 0.94, as compared to 0.82 obtained using PSNR. Comparing to the L^2-metric, we demonstrate that a better performance can be achieved in fractal coding by using the proposed FIM.**

Index Terms—**Fractal coding, fuzzy integrals, image metrics, image quality assessment, quadtree partition.**

I. Introduction

IN IMAGE quality assessment, a subjective evaluation seems to be more meaningful and practical since there are no satisfactory mathematical models serving this task. In the evaluation of a compression algorithm, where a mathematical model must be used, the L^2-metric is usually chosen as a de facto model for its simplicity. The L^2-metric, except for its mathematical elegance as a metric, has little to do with the human visual system. A few other metrics, for example, the *Hausdorff* metric, have been applied primarily to binary images.

Human vision, which has been studied in many disciplines, is a very complex phnomena. From psychological ophthalmology to optical physics, vision spans the full aspect of understanding the human visual mechanism. Although many efforts have been made to discover the mechanism of "how we see" and establish computational models for the human visual perception [1]–[4], the knowledge of this subject is still quite primitive.

In image quality assessment, many studies have been carried out on subjective measures [5]–[7]. The Mean Opinion Score (MOS) is the most commonly used subjective assessment method, which is equivalent to using the human brain as a tool for quantification. The main drawback of MOS is time consuming and experimentally difficult, and the results obtained may vary depending on the test conditions. Using subjective evaluation and fuzzy aggregation techniques, Tizhoosh *et al.* [8]–[10] developed an observer-dependent five-phase system for image enhancement. At the third phase of processing, the MOS values were mapped into the interval [0, 1], which were regarded as fuzzy density values to construct a fuzzy measure.

As a traditional method for objective image quality assessment, PSNR is merely a good distortion indicator for random errors but not for structured or correlated errors, which are prevalent in image compression and degrade local features and perceived quality much more than do random errors. As a consequence, the correlation between PSNR and visual quality is known to be poor [11].

Research results have been reported on various objective image quality assessment methods in literature. Miyahara *et al.* presented a picture quality scale (PQS) for the coding of achromatic images over the full range of image quality defined by the subjective MOS [12]. The main feature of the measure is that it takes into account the properties of human visual perception for both global and local image characteristics. In modeling a degraded image as an original image that has been subject to linear frequency distortion and additive noise injection, Damera-Venkata *et al.* proposed two parameters, a distortion measure (DM) and a noise quality measure (NQM), to grade the effects of these two distortion sources [13]. Both measures are based on the research results on the human visual system. In an attempt to simulate the subjective evaluation of image quality, Bock *et al.* used fuzzy rules to evaluate image distortion based on a model of human visual system [14]. All these methods mentioned above require complex techniques such as edge extraction or multivariate analysis, so they surely gain some advances while lose the simplicity that PSNR has.

In the n-dimensional vector space $\Re^n$of real numbers, the most frequently used metric is L^p-metric, which is defined as

$$d(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p} \tag{1}$$

for all $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n) \in \Re^n$ and $p \geq 1$. In the extreme case, when p tends to infinity, the metric becomes the *supremum* metric, i.e., the L^∞-metric, which is given by

$$d(x, y) = \sup_{1 \leq i \leq n} |x_i - y_i| . \tag{2}$$

The case $p = 2$ is the most frequently used metric, L^2-metric.

Manuscript received February 6, 2001; revised February 9, 2002. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Robert L. Stevenson.

J. Li is with the Institute of Science, Zhejiang University, Hangzhou, China, and also with the Center for Multimedia Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong.

G. Chen is with the Institute of Science, Zhejiang University, Hangzhou, China (e-mail: gchen@math.zju.edu.cn).

Z. Chi is with the Center for Multimedia Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong (e-mail: enzheru@polyu.edu.hk).

Publisher Item Identifier S 1057-7149(02)04796-6.

Given a metric d on the n-dimensional space $\Re^n$, the *Hausdorff* metric of d, denoted by HD, is defined in a space of subsets of $\Re^n$ by

$$HD(A, B) = \max\left\{\sup_{x\in A}\left\{\inf_{y\in B} d(x, y)\right\}, \sup_{y\in B}\left\{\inf_{x\in A} d(x, y)\right\}\right\} \quad (3)$$

for any subsets of $\Re^n$, A and B.

The conventional image metrics, such as L^∞, L^p, *Hausdorff*, have some drawbacks in measuring the image distortion, which is discussed as follows.

1) The golden locality, e.g., L^∞-metric. The measure of the difference between two images based on these metrics will become much greater when the difference of some pixels is greater. However, human perception will usually be unable to notice the big difference of a small number of image pixels.
2) The computation of a metric such as *Hausdorff* is too complicated to make the metric practical.

Peak Signal-to-Noise Ratio (PSNR) commonly used in evaluating image processing can describe the image quality to a certain extent but fails to produce a measure that is consistent with the human visual perception. Meanwhile, since PSNR is not a metric, its usage is quite limited.

In this paper, we present a new image metric, Fuzzy Image Metric (FIM), which can reflect the human visual perception while remaining the computability and simplicity. Our proposed FIM, which is discussed in Section III, can overcome some drawbacks of the conventional image metrics to a certain extent. Although the proposed FIM is also based on pixels differences, the experimental results reported in Section IV show that as an objective image metric it is more consistent with the human visual perception than PSNR. Moreover, it is simple in computation and it has a good anti-golden-locality capability. Section V demonstrates the application of our fuzzy image metric to fractal coding and compares its performance with that of the L^2-metric. Finally, concluding remarks are drawn in Section VI.

II. Fuzzy Integrals

Fuzzy measures and fuzzy integrals were first introduced by Sugeno [15] to evaluate nonadditive or nonlinear quantity in systems engineering. Since then, they have been applied to various problems such as expert systems, decision making, pattern recognition, image enhancement, risk analysis, and so on [16].

Definition 2.1: Let C be a Borel field of nonempty X. A set function μ: $\Omega \to [0, 1]$ defined on C is called a *fuzzy measure* if μ satisfies

1) boundary conditions: $\mu(\emptyset) = 0$; $\mu(X) = 1$;
2) monotonicity: if $A \subset B \subset X$ and $A, B \in C$, then $\mu(A) \leq \mu(B)$;
3) continuity: if $A_1 \subset A_2 \subset \cdots \subset A_n \subset \ldots, A_n \in C$, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

The triple (X, Ω, μ) is called a fuzzy measure space.

Definition 2.2: Let f be a measurable nonnegative real-valued function on X, and let $A \subset X$. The Sugeno's fuzzy integral $(S)\int_A f d\mu$ of f over A with respect to μ is defined by

$$(S)\int_A f d\mu = \sup_{\alpha>0} \min\left(\alpha, \mu\left(N_\alpha(f) \cap A\right)\right) \quad (4)$$

where

$$N_\alpha(f) = \{x | f(x) \geq \alpha\}. \quad (5)$$

Let X be an object that has n quality factors $x_1, x_2, \ldots x_n$. For a subset A of X, we use $\mu(A)$, which ranges from 0 to 1, to measure the maximum grade that X can achieve for the quality factors of A. Actually, we use $\mu(A)$ to indicate the importance of A. It is reasonable to assume that the value of the importance of X is equal to 1 and the value of the importance of empty set $\emptyset$ is equal to 0. In addition, if a factor set A is included in a factor set B, then the importance of A is lower than that of B, i.e., $\mu(A) \leq \mu(B)$. The set function μ satisfies the following conditions:

1) $\mu(\emptyset) = 0$; $\mu(X) = 1$;
2) If $A \subset B \subset X$, then $\mu(A) \leq \mu(B)$.

Considering a factor vector always consists of a finite number of factors, μ naturally becomes a fuzzy measure. The importance measure plays a key role in the judging procedure. It is a sound way to quantize experts' experience and well received as an accredited judging criterion.

III. Image Metric Based on Fuzzy Integral

We denote two images by

$$X = (x_1, x_2, \ldots, x_K), \quad Y = (y_1, y_2, \ldots, y_K)$$

where x_i, y_i are pixel intensities. Their difference is defined by

$$|X - Y| = (|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_K - y_K|) \quad (6)$$

where $0 \leq x_i, y_i \leq 1$ (after normalization) and K is the number of pixels in the images. We define a fuzzy image metric as follows:

$$\begin{aligned} d_{F(S)}(X, Y) &= (S)\int_D |X - Y| d\mu \\ &= \sup_{\alpha>0} \min\left(\alpha, \mu\left(N_\alpha(|X - Y|)\right)\right) \end{aligned} \quad (7)$$

where D is the image plane and

$$\mu = |\{\cdot\}| / K \quad (8)$$

where $|\{\cdot\}|$ is the number of elements in $\{\cdot\}$ and $\mu(D) = 1$. It is easy to verify that μ is in fact an additive measure, i.e., probability measure. Furthermore, we can easily prove that $d_{F(S)}$ (X, Y) is a metric according to the properties of the fuzzy integral and absolute value inequality. We call such $d_{F(S)}$ *fuzzy image metric* (FIM).

Note 3.1: For gray images, X and Y, we have

$$\begin{aligned} \text{FIM}(X, Y) &= (S)\int_D |X - Y| d\mu \\ &= \max_{0\leq i\leq 255} \min\left(i/255, \mu\left(N_{i/255}\left(|X - Y|\right)\right)\right). \end{aligned} \quad (9)$$

There must exist some $(i_0/255)$ such that FIM achieves its maximum as shown in Fig. 1. Hence FIM can be interpreted as searching for the maximal agreement between the pixel differences $(i/255,\ i = 0, \ldots 255)$ and the proportion of the pixels, on which the errors are no less than $(i/255)$.

Note 3.2: Compared to the conventional image metrics such as L^{∞}, L^{p}, and Hausdorff, and the PSNR measure, FIM not only takes into account the difference between the corresponding pixels, but also the proportion of the pixels whose corresponding differences are no less than a given value.

Note 3.3: FIM is considered as an objective evaluation with a certain nature of subjective evaluation. We regard $|X - Y|$ as an object to be evaluated and each pixel as a quality factor. In addition, we take μ (additive) as the importance measure on the image because

1) The definition of μ is simple and easy to compute. For an image with tens of thousands of pixels, the determination of a nonadditive fuzzy measure on it is in fact infeasible.
2) μ is an additive measure that is the abstract of some important objective concepts such as length, area, value, mass, etc., which is accepted by most people.

We treat the difference of the gray levels of the corresponding two pixels [i.e., $|x_i - y_i|$, $(i = 1, 2, \ldots, K)$] as an intrinsic quality index such that the evaluation of $|X - Y|$ is comprehensive and objective. Then FIM is an objective and ideal evaluation of $|X - Y|$. For the details of the importance measure and subjective judgment, the reader may refer to [15], [17].

Next, we give the three theorems that will be used in the quadtree partition based fractal coding to be discussed in Section V.

Theorem 3.1: For a given $\alpha > 0$, $\text{FIM}(X, Y) < \alpha$ if and only if $\mu(N_\alpha(|X - Y|)) < \alpha$.

Proof: Assume $\text{FIM}(X, Y) < \alpha$ for a given $\alpha > 0$. Then we have

$$\min\left(\alpha,\ \mu\left(N_\alpha\left(|X - Y|\right)\right)\right) < \alpha$$

which implies that $\mu(N_\alpha(|X - Y|)) < \alpha$.

Conversely, assume that $\mu(N_\alpha(|X - Y|)) < \alpha$. For any $\beta \geq \alpha > 0$, since $\mu(N_\beta(|X - Y|)) \leq \mu(N_\alpha(|X - Y|))$ [refer to (5) and (6)], we have

$$\min\left(\beta,\ \mu\left(N_\beta\left(|X - Y|\right)\right)\right) \leq \mu\left(N_\alpha\left(|X - Y|\right)\right) < \alpha.$$

On the other hand, for any $0 < \beta < \alpha$, obviously, we have

$$\min\left(\beta,\ \mu\left(N_\beta\left(|X - Y|\right)\right)\right) \leq \beta < \alpha.$$

Thus, we prove $\text{FIM}(X, Y) < \alpha$. Similarly, we have the following.

Theorem 3.2: For a given $\alpha > 0$, $\text{FIM}(X, Y) \leq \alpha$ if and only if $\mu(N_\beta(|X - Y|)) \leq \alpha$ for any $\beta > \alpha$.

Theorem 3.3: For a given $\alpha > 0$, $\mu(N_\alpha(|X - Y|)) \leq \alpha$ implies that $\text{FIM}(X, Y) \leq \alpha$.

IV. Experimental Results and Discussion

Normally, the smaller the FIM measure, the greater the PSNR measure, and therefore the more similar the two compared images (see Fig. 2). But it is not always the case. Judging from the PSNR measure of Fig. 3(b), the compressed image is unsatisfactory. However, based on the FIM measure, its quality is fair, which is consistent with human visual perception. Based on the PSNR measure, the image shown in Fig. 3(c) is poorer than that shown in Fig. 3(d), which is different from the human perception. The FIM measure does indicate the similar result as the human visual perception. Fig. 3(a) is corrupted by changing the gray level of one pixel (the difference is 200 in 256 grey scales) of the original image and it is visually the same as the original image. However, L^{∞} shows the golden locality in this case, while FIM does not. The behavior of the FIM measure from these experiments is similar to that of the human visual perception.

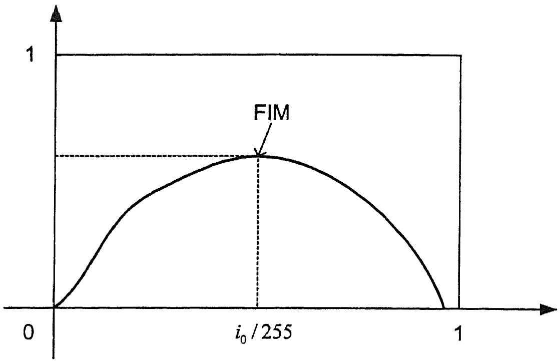


Fig. 1. Relationship between $\min(i/255,\ \mu(N_{i/255}(|X-Y|)))$ and $(i/255)$ with the maximum being the FIM value.

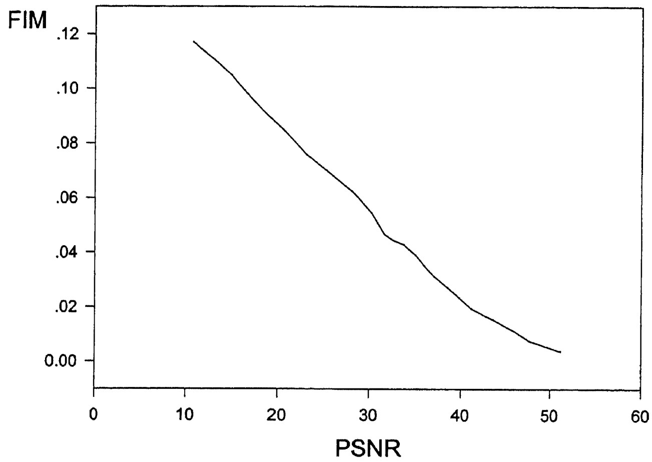


Fig. 2. Fuzzy image metric versus PSNR for the fractal coding of Lena image.

In order to verify the consistence of FIM with Mean Opinion Score (MOS), we selected eight commonly used test images shown in Fig. 4. Normally, for a more reliable image quality assessment, we will use noiseless and good original images. However, for the evaluation of image compression, since we are mainly concerned with the difference (distortion) between the original and the decoded image, the original image has not necessarily to be noiseless and/or good. The test images used for our experiments are noiseless since we want to have a reliable MOS measure. We compressed each image with 15 different compression ratios using the JPEG technique. According to the recommendations of ITU (*International Telecommunication Union*) [6], each compressed image was evaluated by 15

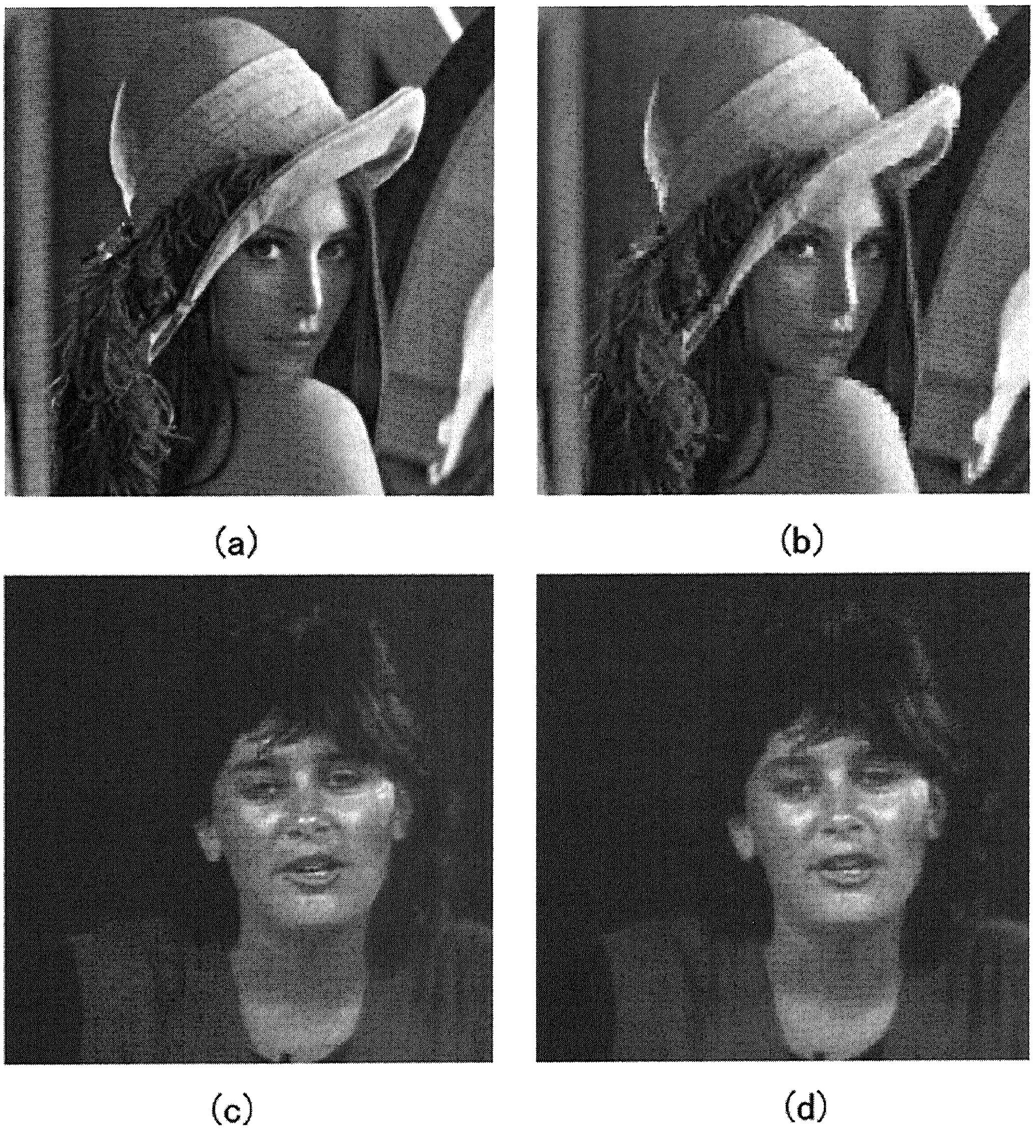


Fig. 3. (a) Lena image with artificial artifact, $L^{\infty} = 0.78$ (maximum is 1) and FIM = 1.5 e–05; (b) Lena image compressed with fractal coding, PSNR = 21.58 and FIM = 0.0586; (c) Missa image compressed with JPEG, PSNR = 38.41 and FIM = 0.0195; and (d) Missa image compressed with fractal coding, PSNR = 38.64 and FIM = 0.0273.

observers who judged the image quality against the original image using a grading scale from five to one where five stands for excellent, four for good, three for fair, two for poor, and one for bad. Finally, the average of 15 scores given by participants was taken as the total quality measurement.

Recommendation ITU-R BT 500-10 [6] proposed two approximation functions, of which one is symmetrical and the other is nonsymmetrical. We found out that the nonsymmetrical function is suitable for fitting the Image Quality Evaluation (IQE) versus FIM as shown in Fig. 5, which is given by

$$\mathrm{IQE}(FIM) = \frac{5}{1 + (\mathrm{FIM}/a)^b} \tag{10}$$

where $a = 0.0647$, $b = 4.438$. The parameters a and b is determined experimentally to fit the MOS-FIM measure of 120 points. Using this fitting curve, image quality can be judged from the FIM value easily.

Finally, we used the curve to verify the agreement between FIM and MOS. A fairly good result is achieved as shown in Fig. 6. Note that a better fitting was achieved at the highest and lowest ends of the quality range than in its middle part. In order to describe the degree of approximation of the FIM to MOS quantitatively, the correlation coefficient R [18] between the FIM and MOS is calculated. Let $x = (x_1, x_2, \ldots, x_n)$ be the FIM measure and $y = (y_1, y_2, \ldots, y_n)$ be the MOS measure, and $\overline{x} = (1/n) \sum_{i=1}^{n} x_i, \overline{y} = (1/n) \sum_{i=1}^{n} y_i$, the correlation coefficient between x and y is given by

$$R = \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) \Bigg/ \left\{ \sum_{i=1}^{n} (x_i - \overline{x})^2 \sum_{i=1}^{n} (y_i - \overline{y})^2 \right\}^{1/2}. \tag{11}$$

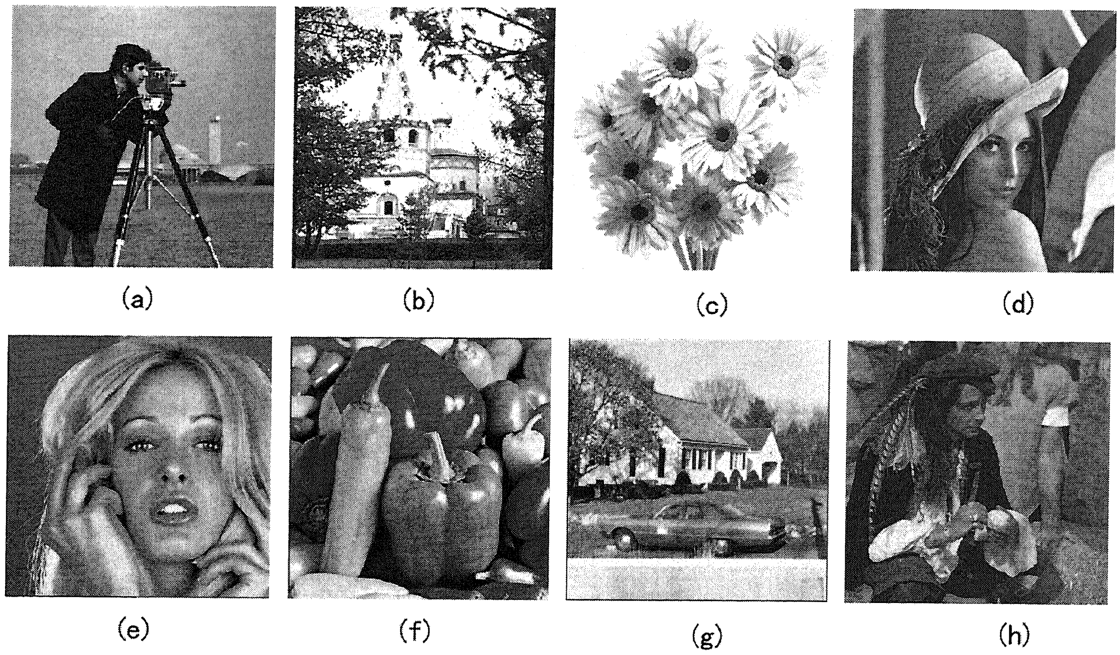


Fig. 4. Test images for FIM and MOS. (a) Cameraman; (b) church; (c) flower; (d) Lena; (e) woman; (f) peppers; (g) house; and (h) man.

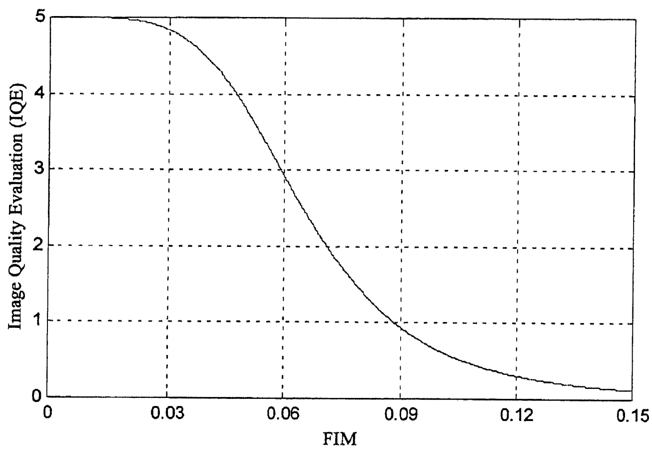


Fig. 5. Image quality evaluation versus FIM.

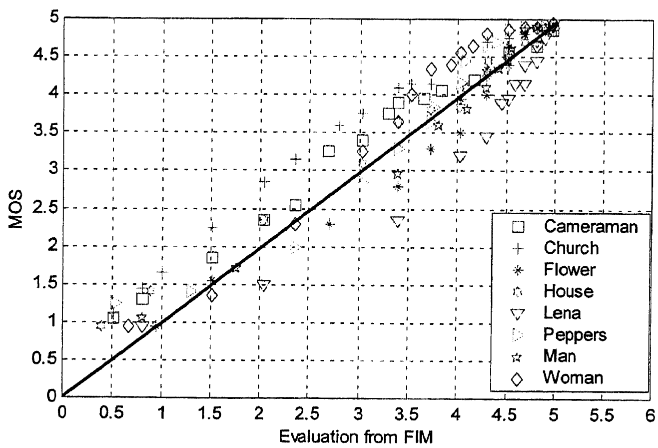


Fig. 6. MOS versus the image quality evaluation from FIM after curve fitting using (10).

TABLE I
PERFORMANCE COMPARISON OF FIM VS THE L^2 METRIC ON THE FRACTAL CODING OF AN 256 × 256 IMAGE

Metric	Range Block Size	PSNR	FIM	Ratio of Processing time to the L^2 metric
FIM	4×4	33.61	0.046	4.1
	8×8	28.18	0.071	3.8
	16×16	23.77	0.101	3.6
	32×32	19.70	0.142	2.9
L^2	4×4	32.66	0.050	1
	8×8	27.43	0.074	1
	16×16	23.31	0.101	1
	32×32	19.38	0.144	1

TABLE II
FIM VS THE L^2 METRIC IN THE QUADTREE APPROACH FOR THE FRACTAL CODING OF THE SAME 256 × 256 IMAGES AS IN TABLE I

Metric	Threshold	FIM	Compression Ratio	Ratio of Processing time to the L^2 metric
L^2	$\varepsilon=5$	0.039063	5.356	1
	$\varepsilon=7$	0.050781	8.938	1
	$\varepsilon=10$	0.0625	15.101	1
FIM	$\varepsilon=10/256=0.0390625$	0.039063	6.332	0.55
	$\varepsilon=13/256=0.050781$	0.050781	9.436	0.58
	$\varepsilon=16/256=0.0625$	0.0625	16.503	0.61

If x and y is highly correlated, R is close to 1. If R is zero, x and y is orthogonal. The correlation coefficient between FIM and MOS is 0.96, which is a slight improvement over that achieved before the curve fitting (0.94) and a good improvement compared to PSNR, whose correlation coefficient is 0.82.

We believe that FIM is a good image metric for applications such as pattern recognition, image retrieval, and fractal coding because the loss of a few trivial details would not affect the entirely judgment as long as the pertinency of the main contents is preserved. However, the loss of these details would cause significant changes to the values of PSNR, and L^∞, L^p and Hausdorff metrics.

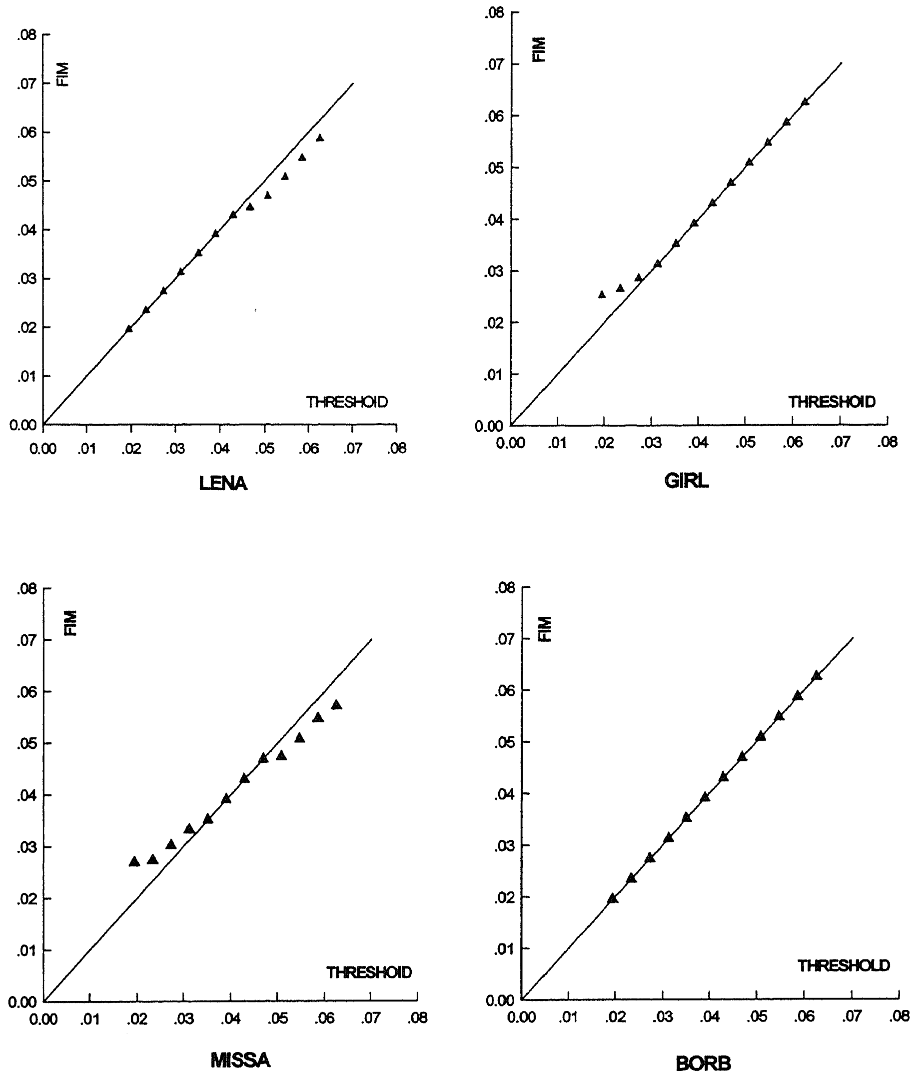


Fig. 7. Threshold versus FIM in the fractal coding of four images. The choice of the threshold ε has a good intuitionistic prediction on the similarity between the original image and the decoded one, and the relationship between the threshold and the FIM are almost linear (identical).

V. Application to Fractal Coding

Barnsley *et al.* [19] proposed to use fractal techniques to code images. The intuitive idea is very appealing: since many images contain complex fractal-like objects, an efficient way of encoding the image is to describe these objects as fractals rather than approximating them in "smooth" ways. However, the methods they described are difficult in implementation, and rely on the use of a computer operator to find correspondences between various parts of an image.

Jacquin [20], [21] presented an automatic coding technique for monochrome images. In his approach, the image is partitioned into squares called range blocks $\{R_k\}$. In each range block R_k, the image is described by the relationship between it and another square on the image, the domain blocks D_k, which has an area of four times of that of R_k. There are eight linear mappings of D_k onto R_k (because of reflection and rotation). For any of these mappings, a pixel of R_k with gray level f corresponds to four pixels of D_k with gray levels f_1, f_2, f_3, f_4. Let $g = (1/4)(f_1 + f_2 + f_3 + f_4)$. For each R_k, the task is to find D_k, a_k (with $|a_k| < 1$) and b_k such that the error between f and $(a_k g + b_k)$ in R_k is the smallest, say ε. We then say that we have achieved a good *collage* on R_k. The local contractive, affine transformation T_k replaces the gray level f by $(a_k g + b_k)$ for each pixel in R_k. Performing this for all range blocks R_k leads to a global transformation $\boldsymbol{T}$. We can consider $\boldsymbol{T}$ as an contractive operator on the image space. By the Banach Fixed Point Theorem, there is a unique image $\boldsymbol{f}^*$ such that $\boldsymbol{T}(\boldsymbol{f}^*) = \boldsymbol{f}^*$.

In Jacquin's fractal coding technique, the L^2 metric is used to evaluate the similarity of range blocks $\{R_k\}$and transformed domain blocks $\{D_k\}$. If the L^2 distance between R_k and transformed D_k is smaller than a given threshold ε, then T_k which transforms D_k into R_k is recorded and a good *collage* on R_k is achieved. However, the L^2 metric has some drawbacks which are discussed as follows.

1) The drawbacks with PSNR, and L^∞, L^p and *Hausdorff* metrics discussed in Section IV are also applied to the L^2 metric.

(a) (b)

Fig. 8. (a) Girl image compressed with a range block size 4 × 4, FIM = 0.0586 and (b) Girl image compressed with a quadtree scheme with threshold ε = 0.039 062 5, FIM = 0.0391.

2) The choice of the threshold ε has a poor prediction on the similarity of the original image and the encoded one. The exact relationship between the threshold ε and the similarity measure is not known.

To compare the performance of using the L^2 metric to that of using the FIM measure in searching for the best matched domain block for a range block, we implement fractal coding on an image (GIRL) by using fixed range blocks with sizes being 4×4, 8×8, 16×16, and 32×32, respectively. The experimental results are shown in Table I.

From Table I, we can see that the results of using FIM is a slightly better than those of using the L^2 metric in terms of PSNR. However, the former took longer time to process than the latter when the block size is 4×4. Based on Theorems 2.1, 2.2, and 2.3, we can use the FIM measure in the fractal coding with the quadtree scheme.

The quadtree is the simplest and most commonly used hierarchical partitioning scheme. In a quadtree approach, one might begin with a regular partitioning of $2^N \times 2^N$ blocks. For each range block, the domain pool is searched for the best matched domain block. If the accuracy of the matching falls within a certain tolerance, it is accepted. If not, the range block is subdivided, and a search is initiated for each sub-block. Various measures were proposed for the control of the image quality in fractal coding with the quadtree partition scheme [22], [23]. Distasi *et al.* presented an entropy based split decision function that can improve image quality and speed up the encoding process [22]. In the approach presented by Saupe and Jacob [23], a test based on block variances was adopted for a splitting criterion to speed up the fractal image compression without harming the rate-distortion performance.

To simplify the discussion, we adopt a full-layer structure, in which the quadtree begins with the whole image and might end with one pixel if a good matched domain block could not be found for a larger range block.

We denote a range block by

$$R_k = \left(r_{k(1)}, r_{k(2)}, \ldots, r_{k(m)}\right)$$

and a domain block by

$$D_k = \left(d_{k(1)}, d_{k(2)}, \ldots, d_{k(m)}\right).$$

For a given threshold, say ε, for each R_k, if $FIM(R_k, D_k) < \varepsilon$ for some D_ks, then we say that we find a good *collage*. Otherwise, R_k is subdivided, and a search is initiated for each sub-block. Note that for Theorems 2.1, 2.2, and 2.3, we only consider whether $\mu(N_\varepsilon(|X - Y|)) < \varepsilon$ is satisfied or not, which can reduce a lot of processing time.

From Table II, we can see that fractal coding using the FIM measure can achieve better compression ratio and use less processing time than that using the L^2 metric. Note that the decoded images have similar visual quality, which is indicated by the same FIM measures. Moreover, the choice of the threshold ε has a good intuitionistic prediction on the similarity between the original image and the decoded one, and the values of the threshold and the FIM measure are almost the same. Fig. 7 shows the relationship between the threshold ε and the FIM measure. As we can see from these results, they are almost linear (identical) functions. Fig. 8 shows two compressed images using the fractal coding without and with the quadtree partition. Judging from the FIM measure, Fig. 8(b) (a smaller FIM value) is better than Fig. 8(a), which is consistent with the human visual perception.

VI. Conclusion

This paper presents a new metric model, FIM, based on the Sugeno's fuzzy integral for objective digital image quality assessment from the viewpoint of human judgment. We use fuzzy integral to bridge the gap between this metric model and pure subjective evaluation. Experimental results show that the proposed FIM measure is highly correlated to the results of the human visual perception and can be successfully used in the fractal coding with a quadtree partition scheme to achieve a better compression performance and to reduce compression time. It is also indicated that a quality measure model like $L^2$and PSNR based on errors' summation cannot reflect the essence of human visual perception. Further improvements can be expected. For example, $|X - Y|$ can be replaced by a better model by incorporating some important properties of the human visual system. By considering background brightness, image contrast, texture variation, and so on, we can adopt the weighted $|X - Y|$ as the difference measure of the two images.

REFERENCES

[1] M. Wertheimer, *Laws of Organization in Perceptual Forms.* London, U.K.: Harcourt, Brace, 1938.

[2] J. J. Gibson, *The Ecological Approach to Visual Perception.* Boston, MA: Houghton Mifflin, 1979.

[3] D. Marr, *Vision.* New York: Freeman, 1982.

[4] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. Image Processing*, vol. 6, pp. 451–462, Mar. 1997.

[5] N. W. Lewis and J. A. Allnatt, "Subjective quality of television pictures with multiple impairments," *Electron. Lett.*, vol. 1, pp. 187–188, July 1965.

[6] "Methodology for the subjective assessment of the quality of television pictures,", Rec. ITU-R BT 500-10, 2000.

[7] *Advanced Methods for the Evaluation of Television Picture Quality.* Eindhoven, The Netherlands: The Mosaic Workshop, Inst. Perception Research, 1995.

[8] H. R. Tizhoosh, G. Krell, and B. Michaelis, "Enhancement of megavoltage images in radiation therapy using fuzzy and neural image processing techniques," in *Fuzzy Systems in Medicine, Studies in Fuzziness and Soft computing.* Berlin, Germany: Physica-Verlag, 1999.

[9] H. R. Tizhoosh and B. Michaelis, "Improvement of image quality based on subjective evaluation and fuzzy aggregation techniques," in *Proc. Eur. Congr. Intelligent Techniques and Soft Computing (EUFIT'98)*, vol. 2, Aachen, Germany, Sept. 7–10, 1998, pp. 1325–1329.

[10] ——, "Image enhancement based on fuzzy aggregation techniques," in *Proc. IEEE Instrumentation and Measurement Technology Conf.*, vol. 3, Venice, Italy, May 24–26, 1999, pp. 1813–1817.

[11] T. Mitsa, "Image quality metrics for halftone images," *Proc. SPIE*, vol. 1778, pp. 196–207, Mar. 1992.

[12] M. Miyahara, K. Kotani, and V. R. Algazi, "Objective picture quality scale (PQS) for image coding," *IEEE Trans. Commun.*, vol. 46, no. 9, pp. 1215–1226, 1998.

[13] N. Damera-Venkata, T. D. Kite, W. S. Geisler, B. L. Evans, and A. C. Bovik, "Image quality assessment based on a degradation model," *IEEE Trans. Image Processing*, vol. 9, pp. 636–650, Apr. 2000.

[14] F. Bock, A. Steudel, H. Clausert, and M. Glesner, "A fuzzy rule-based system for the subjective quality evaluation of distorted images," in *Proc. 6th Int. Conf. Image Processing and its Applications*, vol. 1, July 14–17, 1997, pp. 448–452.

[15] M. Sugeno, "Fuzzy measures and fuzzy integrals—A survey," in *Fuzzy Automata and Decision Processes*, M. M. Gupta, G. N. Saridis, and B. R. Gaines, Eds. Amsterdam, The Netherlands: North-Holland, 1977, pp. 89–102.

[16] M. Grabisch, H. T. Nguyen, and E. A. Walker, *Fundamentals of Uncertainty Calculi With Application to Fuzzy Inference.* Norwell, MA: Kluwer, 1995.

[17] Z. Wang and F. Li, "Applying the fuzzy integral in the judging procedure" (in Chinese), *Fuzzy Math.*, vol. 5, no. 3, pp. 107–144, 1985.

[18] M. G. Kendall, *Multivariate Analysis.* London, U. K.: Griffin, 1975.

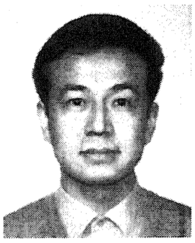
[19] M. F. Bansley and A. D. Sloan, "Methods and apparatus for image compression by iterated function system," U.S. Patent 4 941 193, July 1990.

[20] A. E. Jacquin, "Fractal image coding based on a theory of iterated contractive image transformations," *Proc. SPIE*, vol. 1360, pp. 227–239, 1990.

[21] ——, "Image coding based on a fractal theory of iterated contractive image transformations," *IEEE Trans. Image Processing*, vol. 1, pp. 18–30, Jan. 1992.

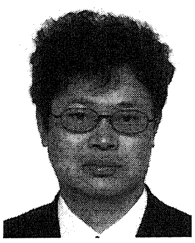
[22] R. Distasi, M. Polvere, and M. Nappi, "Split decision functions in fractal image coding," *Electron. Lett.*, vol. 34, no. 8, pp. 751–753, 1998.

[23] D. Saupe and S. Jacob, "Variance-based quadtrees in fractal image compression," *Electron. Lett.*, vol. 33, no. 1, pp. 46–48, 1997.



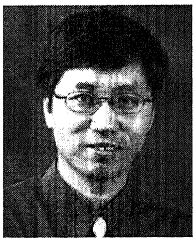
Junli Li received the B.S. degree from Harbin Normal University, China, in 1993 and the M.S. degree from Harbin Institute of Technology in 1996. He is currently pursuing the Ph.D. degree in the Department of Mathematics at Zhejiang University, China.

He was a Research Assistant in the Department of Electronic and Information Engineering, Hong Kong Polytechnic University, from June to October 2001. His research interests include image compression, image quality assessment, image segmentation, image retrieval and fuzzy systems. He has published three technical papers and submitted five others.



Gang Chen received the B.S. degree from Anqing Teachers College, China, in 1983 and the Ph.D. degree in the Department of Applied Mathematics at Zhejiang University, China, in 1994.

From 1994 to 1996, he was a Postdoctoral Researcher in electrical engineering at Zhejiang University. From 1997 to 1999, he was a Visiting Researcher with the Institute of Mathematics at the Chinese University of Hong Kong and the Department of Electronic and Information Engineering at The Hong Kong Polytechnic University. From 1996 to 2000, he was an Associate Professor at Zhejiang University. Since 2001, he has been a Professor. He has been the Deputy Director of the Institute of Imaging and Computer Graphics at Zhejiang University since 2000. His research interests include applied mathematics, image processing, fractal geometry, wavelet analysis and computer graphics. He has co-authored one book and co-edited five technical proceedings. He has published more than 50 technical papers.



Zheru Chi (M'94) received the B.Eng. and M.Eng. degrees from Zhejiang University, China, in 1982 and 1985, respectively, and Ph.D. degree from the University of Sydney, Sydney, Australia, in March 1994, all in electrical engineering.

Between 1985 and 1989, he was on the Faculty of the Department of Scientific Instruments at Zhejiang University. He worked as a Senior Research Assistant/Research Fellow in the Laboratory for Imaging Science and Engineering at the University of Sydney from April 1993 to January 1995. In February 1995, he joined the Hong Kong Polytechnic University, where he is now an Associate Professor in the Department of Electronic and Information Engineering. Since 1997, he has served as a Co-organizer of a special session/session chair/area moderator/program committee member for a number of international conferences. He was one of contributors to *Comprehensive Dictionary of Electrical Engineering* (Boca Raton, FL: CRC, 1999). His research interests include image processing, pattern recognition, neural networks, and fuzzy systems. He has co-authored one book and five book chapters, and published more than 70 technical papers.