Image Quality Scale (IQS) for Compressed Images Quality Measurement

Nuntapong Yamsang and Somkait Udomhunsakul

Abstract— **In this paper, an objective measurement to evaluate the quality of gray scale compressed images, Image Quality Scale (IQS), has been proposed. It is defined by 5 measurements, which are Mean Square Error, Edge Measurement, Correlation Measurement, Human Visual System Measurement and Spectrum Measurement. The evaluation result is rated into 5-level grading scale, 1 to 5, which is comparable to Mean Opinion Score (MOS). The objective of this paper is to provide defining method, definition, and reliability of IQS. From the experiments, we demonstrate that the reliability (correlation coefficient) of IQS is higher than PSNR, MSE, Edge Measurement, Correlation Measurement, HVS Measurement and Spectrum Measurement. Therefore, the IQS is reliable and can be used to evaluate the quality of compressed images, JPEG and JPEG2000.**

Keywords— Image quality assessment, Image Compression, Least Square Method

I. Introduction

Nowadays, image compression is an encoding process to reduce the storage and transmission requirements in many applications. As long as image compressions are meaningful, so compression evaluations are very essential. To achieve a good compression evaluation, it could bring out a good compression technique. In general, compressed image quality can be evaluated by objectively and subjectively. Objective methods are defined by mathematical definition, such as Peak Signal to Noise Ratio (PSNR), Human Visual System (HVS), etc. On the other hand, subjective ones are by human perception specified by Mean Opinion Score (MOS) [1, 2]. Though objective measurements are not tedious as subjective measurements but they do not correlate well with the subjective measurement, which provides truly definitive measure [4].

During a few decades, many researchers had attempted to develop some objective measurements. In 1998, Miyahara et al [3] had proposed a Picture Quality Scale (PQS) that is closely approximates the MOS (Mean Opinion Score). However, for very high quality images, it is possible to obtain values of PQS larger than five. At the low end of the image quality scale, PQS can obtain negative values (meaning less result) [4].

Nuntapong Yamsang and Somkait Udomhunsakul, Assistant Professor, are with Faculty of Engineering, Department of Information Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand.

In addition, Universal Quality Index (UQI) and Structural Similarity Index Measurement (SSIM) were proposed by Bovik et al [5,6]. Their results demonstrated that they were grater accuracy and consistency than MSE and PSNR. However, SSIM is failed to measure badly blurred images. Practically, the UQI, and SSIM measurement results were rate the image quality from 0 to 1 (unacceptable to excellent quality).

In this work, we propose an objective measurement to evaluate the quality of gray scale compressed images denoted as Image Quality Scale (IQS). It is defined by 5 measurements, which are Mean Square Error (MSE), Edge Measurement (Edge), Correlation Measurement (C), Human Visual System (HVS) Measurement and Spectrum Measurement (Spt.). It gives us the evaluation result as 5-level grading scale, 1 to 5, which is comparable to Mean Opinion Score (MOS).

In the next sections, section 2 we describe the image quality measurement. The reliability, Least square method and defining of IQS are contained in section 3, 4 and 5, respectively. Finally, the experimental results and conclusion will be provided.

II. Image Quality Measurement

Image quality measurements are classified into subjective or objective assessment.

2.1 Subjective Assessment

In fact, in image compression systems, the truly definitive measure of image quality is perceptual quality. The distortion is specified by MOS, which is result of perception based subjective evaluation [1, 2]. The meaning of the 5-level grading scale of MOS is 5-imperceptible, 4-perceptible, but not annoying, 3-slightly annoying, 2-annoying and 1-very annoying.

2.2 Objective Assessment

The objective assessments are save more time [7,8]. In this paper, to justify measurement from bit per pixel depth, the gray-scale value of the original image sample ($f(m,n)$) and compressed image sample ($\hat{f}(m,n)$) would be normalized by peak signal, as equation 1 and 2,

$$x(m,n) = \frac{f(m,n)}{2^b - 1} \qquad (1)$$

$$\hat{x}(m,n) = \frac{\hat{f}(m,n)}{2^b - 1} \quad (2)$$

where, b is bit per pixel.

2.2.1 Mean Square Error (MSE)

The simplest of distortion measurement is Mean Square Error (MSE), defined as,

$$MSE = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} e^2(m,n) \quad (3)$$

where $e(m,n)$ is difference of the $x(m,n)$, original image, and $\hat{x}(m,n)$, reconstructed image. The higher of MSE value refers to the lower image quality.

2.2.2 Edge Measurement (Edge)

This type of quality measure can be obtained from,

$$Edge = \frac{1}{MN}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(Q(i,j) - \hat{Q}(i,j)\right)^2 \quad (4)$$

where $Q(i,j)$ and $\hat{Q}(i,j)$ are the edge gradients of the original and compressed image using a Sobel operator. The higher of Edge measurement means the lower of image quality.

2.2.3 Correlation Measurement (C)

The similarity between two digital images could be quantified by correlation function. Each image is normalized by its root power. So the Correlation measurement is defined as,

$$C = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} f(m,n)\hat{f}(m,n)}{\sqrt{\sum_{m=1}^{M}\sum_{n=1}^{N} f^2(m,n)\sum_{m=1}^{M}\sum_{n=1}^{N}\hat{f}^2(m,n)}} = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} x(m,n)\hat{x}(m,n)}{\sqrt{\sum_{m=1}^{M}\sum_{n=1}^{N} x^2(m,n)\sum_{m=1}^{M}\sum_{n=1}^{N}\hat{x}^2(m,n)}} \quad (5)$$

The higher values of correlation measurement imply more similarity between the original image and compressed image.

2.2.4 Human Visual System (HVS) Measurement

The HVS is too complex to be fully understood, but it is able to be explained into a simplified HVS model. One of the modes for the human visual system is given as a band pass filter with a transfer function, $H(\rho)$ where $\rho = (u^2+v^2)^{1/2}$.

$$H(\rho) = \begin{cases} 0.05e^{\rho^{0.554}} & ,\rho < 7, \\ e^{-9\left[\left|\log_{10}\rho - \log_{10}9\right|\right]^{2.3}} & ,\rho \geq 7 \end{cases} \quad (6)$$

Then, HVS is defined as,

$$HVS = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(U\{x(m,n)\} - U\{\hat{x}(m,n)\}\right)^2 \quad (7)$$

Where $U\{\cdot\}$ operator define as : $U\{x(m,n)\} = DCT^{-1}\{H(\rho)\cdot DCT\{x(m,n)\}\}$.

HVS would be lower if the image quality is better.

2.2.5 Spectrum Measurement (Spt.)

Discrete Fourier Transform, F(i, j), of an normalized image, x(m, n), could provide the magnitude as $M(i,j) = \left|\frac{F(i,j)}{MN}\right|$ and phase as $\phi(i,j) = \arctan(F(i,j))$, hence Spectral Measurement is given as,

$$Spt. = (1-\lambda)\cdot D_M(i,j) + \lambda \cdot D_\phi(i,j) \quad (8)$$

Where $\lambda = 3.845 \times 10^{-10}$, Magnitude distortion, $D_M(i,j) = \sum_{i=1}^{I}\sum_{j=1}^{J}\left(M(i,j) - \hat{M}(i,j)\right)^2$, Phase distortion, $D_\phi(i,j) = \sum_{i=1}^{I}\sum_{j=1}^{J}\left(\phi(i,j) - \hat{\phi}(i,j)\right)^2$.

Same as other distortions, the higher spectrum measurement value due to the image quality is lower.

III. Reliability of Objective Image Quality Measurement

The reliability of an objective measurement could be evaluated by correlation between objective measurement and subjective measurement. The Correlation coefficient (r) can be expressed as, [9]

$$r = \frac{\sum_i (s_i - \bar{s})(o_i - \bar{o})}{\sqrt{\sum_i (s_i - \bar{s})^2 \sum_i (o_i - \bar{o})^2}} \quad (9)$$

where, s_i and o_i are the series of subjective and objective measurements, respectively. The possible values of correlation coefficient are between -1 and 1, the better correlation make the correlation coefficient closer to -1 or 1.

If an objective measurement provides output which is comparable to MOS (5-level grading) then the standard error (S) could be evaluated the reliability too. The standard error is defined as, [9]

$$S = \sqrt{\frac{1}{n}\sum_i (o_i - s_i)^2} \quad (10)$$

where n is the number of sample. The standard error value should be smaller to reach the better measurement.

IV. LEAST SQUARE METHOD

Conception of least square method is to find an equation function, f(x), that provides the smallest square error [10]. The function should be defined as a linear, exponential, polynomial or etc. So, its parameters (p_n) are adjusted to make the function be a least square function and the parameters that make it be could accord with,

$$\frac{\partial}{\partial p_n}\left(\sum_i \left(f(x_i)-y_i\right)^2\right)=0 \quad (11)$$

where f(x) is the least square function, y is a sample function and p_n is n^{th} parameter.

In this paper, we define objective measurement as variable x and MOS as variable y. Then we can find the function f(x) which refers to subjective measurement.

V. DEFINING METHOD

The IQS model is separated into three steps shown in figure 1. First, the gray-scale values of original and compressed images, $(f(m,n), \hat{f}(m,n))$, are justified by peak signal are normalized that is divided by peak signal. Second, each measurement calculates the distortion and maps it into scale (1 to 5) by least square function calculated by holding to subjective measurement's principles. Finally, each scale is weighted and summed for providing IQS. In addition, to weight each scale of objective measurements, its reliability should be considered.

VI. EXPERIMENTAL RESULTS

This section is to provide the defining method, definition, and reliability of IQS.

6.1 Defining and Definition

In figure 2, eight original images (Airplane, Baboon, Goldhill and Lena for 512×512 pixel-size, 8-bpp resolution, Earth, Omaha, Sena and Sinsin for 256×256 pixel-size, 8-bpp resolution), are coded and decoded with JPEG and JPEG2000 algorithm. Each one uses 10 rates. JPEG use quantized parameter (Q) at 0.14, 0.2, 0.27, 0.37, 0.52, 0.72, 1, 1.4, 2 and 2.7. The JPEG 2000 uses percentage threshold at 2.5%, 3.2%, 4%, 5%, 6.3%, 8%, 10%, 12.5%, 16% and 20% remaining rates. Then there are one hundred and sixty distorted images used for defining process.

MSE, Edge Measurement, Correlation Measurement, HVS Measurement, Spectrum Measurement and MOS are used for measure the image quality. The relationship between values of objective measurements and subjective measurements (the truly definitive measure) were considered. Table 1 shows the reliabilities of these objective measurements by correlation coefficient values. Figure 3 shows the exponential distribution graphs of the relationship between objective and subjective measurements.

TABLE I
CORRELATION COEFFICIENT OF OBJECTIVE MEASUERMENTS
(FROM IMAGES MEASURING IN SUBTOPIC 6.1)

Measurements	Correlation Coefficient
MSE	-0.5146
Edge Measurement	-0.5325
Correlation Measurement	0.6390
HVS Measurement	-0.4566
Spectrum Measurement	-0.5469

The model function of least square equation for MSE, Edge Measurement, HVS Measurement and Spectrum Measurement are expressed in equation 12, where M_i is MSE, Edge Measurement, HVS Measurement or Spectrum Measurement and p_i is parameter of each measurement. Additionally equation 13 is a model for Correlation Measurement, where C and p_c are Correlation Measurement value and parameter of correlation measurement scale, respectively.

$$f_{M_i} = 4\cdot e^{p_i M_i} + 1 \quad (12)$$

$$f_C = 4\cdot e^{p_C(1-C)} + 1 \quad (13)$$

Exponential least square functions were calculated and the results of parameters are shown in equation 14 to 18.

$$f_{MSE} = 4\cdot e^{-1242.8MSE} + 1 \quad (14)$$

$$f_{Edge} = 4\cdot e^{-75.766Edge} + 1 \quad (15)$$

$$f_C = 4\cdot e^{-501.22}\cdot e^{501.22C} + 1 \quad (16)$$

$$f_{HVS} = 4\cdot e^{-794390HVS} + 1 \quad (17)$$

$$f_{Spt.} = 4\cdot e^{-967.17Spt.} + 1 \quad (18)$$

TABLE II
STANDARD ERROR AND CORRELATION COEFFICIENT OF MAPPING FUNCTIONS

Equation	Standard Error	Correlation Coefficient
f_{MSE}	0.9887	0.7133
f_{Edge}	0.8328	0.8026
f_C	0.9508	0.7255
f_{HVS}	1.0562	0.7373
f_{Spt}	0.9977	0.6673

The five mapping equations provide standard error, max error and correlation coefficient as shown is table 2.

To weight and sum the mapping function, we determine the weight value (w_i) by considering the reliability which is standard error of each function. The weight value is shown as,

$$w_i = \frac{1/S_i}{\sum_j 1/S_j} \quad (19)$$

where S_i is standard error of f_{MSE}, f_{Edge}, f_C, f_{HVS} or f_{Spt}.

Equation 14 to 18 were weighted and summed to reach IQS, defined as,

$$\begin{aligned} IQS = \; & 0.77607 \cdot e^{-1242.8MSE} + 0.92133 \cdot e^{-75.766Edge} \\ & + 0.80703 \cdot e^{-501.22} \cdot e^{501.22C} + 0.72651 \cdot e^{-794390HVS} \\ & + 0.76906 \cdot e^{-967.17Edge} + 1 \end{aligned} \quad (20)$$

From IQS definition, the standard error is evaluated at 0.8016. So the standard error value of IQS is lower than f_{MSE}, f_{Edge}, f_C, f_{HVS}, f_{Spt}.

6.2 Image Quality Scale Reliability

Six other original images (Jug, Women and People for 512×512 pixel-size, 8-bpp resolution, Cameraman, Face and Sphinx for 256×256 pixel-size, 8-bpp resolution) are used for evaluate the reliability of measurements. They are coded and decoded same as subtopic 6.1. Then IQS, PSNR, MSE, Edge Measurement, Correlation Measurement, HVS Measurement, and Spectrum Measurement are used for measure one hundred and twenty distorted images. The correlation coefficients are represented in table 3. As can be seen, IQS provides the highest correlation coefficient. Therefore, IQS is the most reliable.

TABLE III
CORRELATION COEFFICIENT OF OBJECTIVE MEASUERMENTS
(FORM IMAGES MEASURING IN SUBTOPIC 6.2)

Measurements	**Correlation Coefficient**
IQS	0.8998
PSNR	0.7848
MSE	-0.7645
Edge Measurement	-0.7734
Correlation Measurement	0.7529
HVS Measurement	-0.6691
Spectrum Measurement	-0.7189

VII. CONCLUSION

In this work, we propose an optimal objective measurement led to an ability to subjectively judge the compressed image quality. Because of non-linear relation between objective measurement and subjective measurement, so the correlation coefficient of objective measurements could be applied by nonlinear mapping functions. From the experiments, we prove that IQS is reliable and practical to measure the quality of JPEG and JPEG 2000 compressed images.

REFERENCES

[1] Method for the Subjective Assessment of the Quality of Television Pictures, CCIR Rec. 500-2, 1982.
[2] ITU, Methods for the Subjective Assessment of the Quality of Television Pictures, ITU-R Rec. BT. 500-7, August 1998.
[3] M. Miyahara, K. Kotani, and V.R. Algazi, “Objective Picture Quality Scale (PQS) for Image Coding,” IEEE Transactions on Communication, vol. 46, no. 9, 1998, pp. 1215-1226.
[4] S. Grgic, M. Grgic, and M. Mrak, “Reliability of Objective Picture Quality Measurement,” Journal of Electrical Engineering, vol. 55, no. 1-2, 2004, pp. 3-10.
[5] Z. Wang and A.C. Bovik, “A universal image quality index,” IEEE Signal Processing Letters, vol. 9, no. 3, 2002, pp. 81-84.
[6] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, “Image quality assessment: from error measurement to structural similarity,” IEEE Transactions on Image Processing, vol. 13, no. 4, 2004.
[7] A. M. Eskicioglu and P. S. Fisher, “Image Quality Measures and Their Performance”, IEEE Transactions on Communications, vol.43, no.12, pp. 2959-2965, 1995
[8] I. Avcibas, B. Sankur and K. Sayood, "Statistical Evaluation of Image Quality Measure", Journal of Electronic Image, vol. 11(2), pp. 206-223, Apr. 2002.
[9] X. Rong Li, Probability, Random Signals, and Statistics, New York: CRC Press, 1999.
[10] James L. Buchanan and Peter R. Turner, Numerical Methods and Analysis, New York: McGraw-Hill, Inc., 1992.

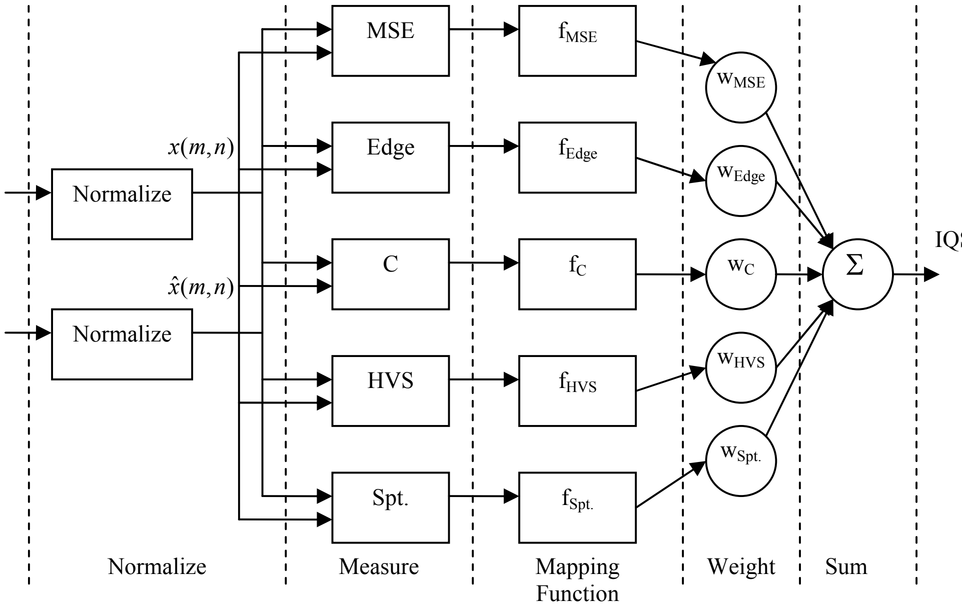

Fig. 1 The IQS System.

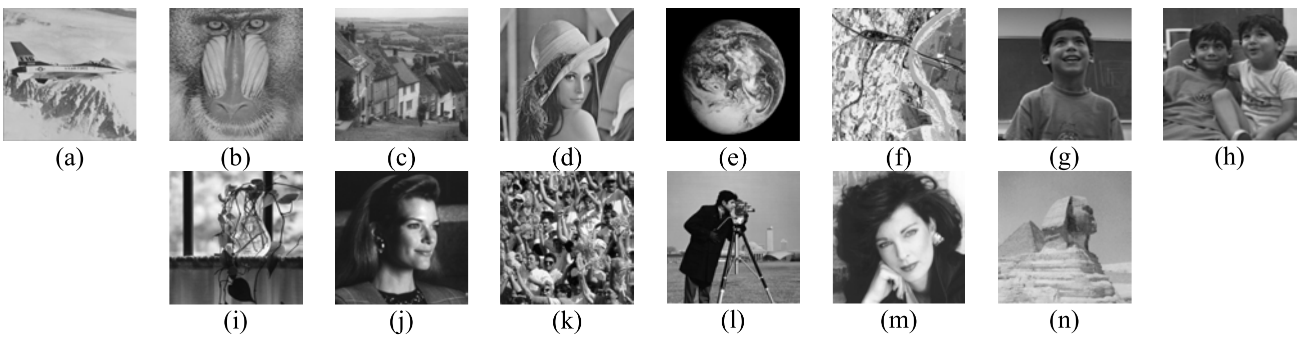

Fig. 2 Original Images, Used in Experiment. (a) Airplane. (b) Baboon. (c) Goldhill. (d) Lena. (e) Earth. (f) Omaha. (g) Sena. (h) Sensin. (i) Jug. (j) Woman. (k) People. (l) Cameraman. (m) Face. (n) Sphnix. Airpane, Baboon, Goldhill, Lena, Jug, Women and People are 512×512 pixel-size and Earth, Omaha, Sena, Sensin Cameraman, Face and Sphnix are 256×256 pixel-size.

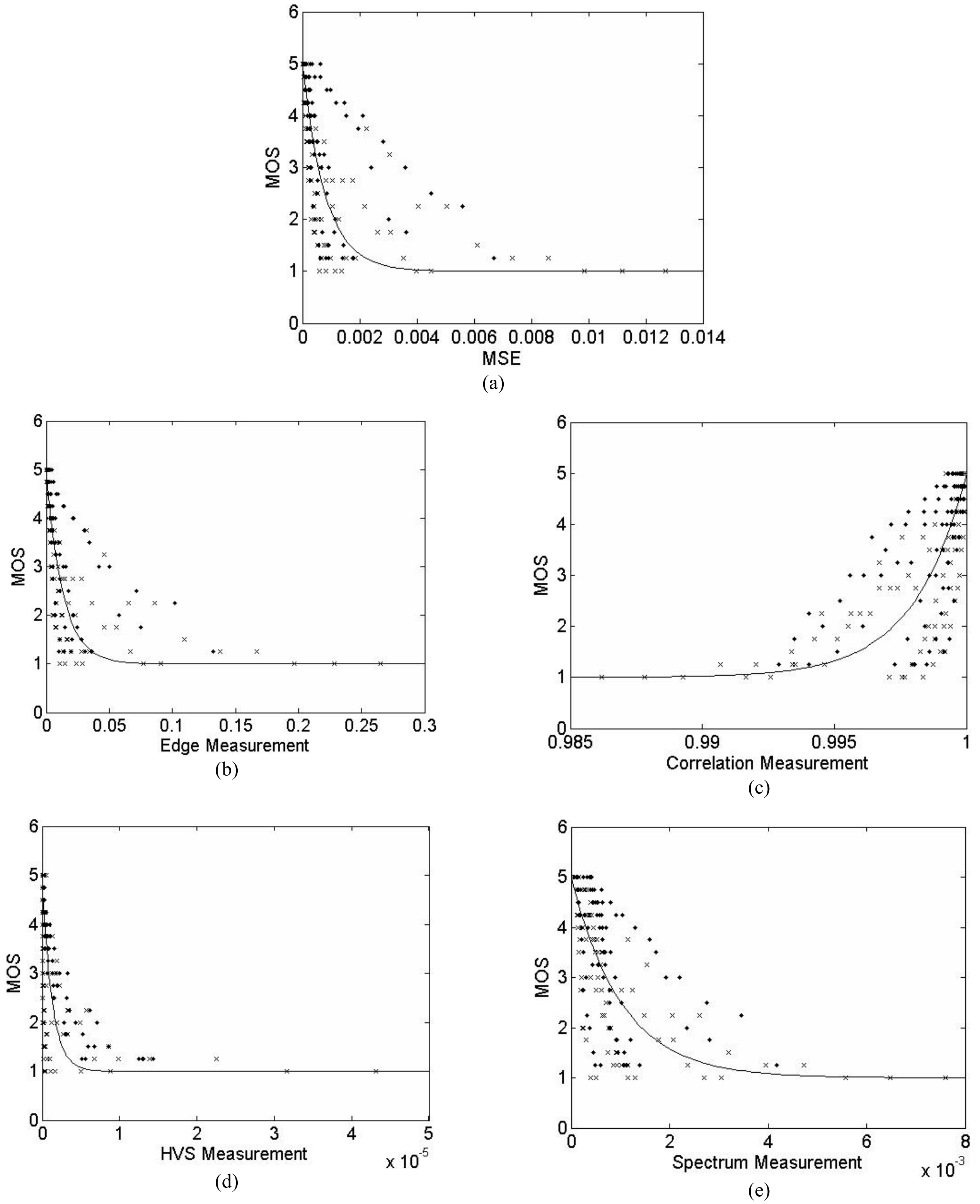

Fig. 3 Relationships between objective measurements and subjective measurement. Each sample point represents one test image.
(a) Relationship between MSE and MOS (b) Relationship between Edge Measurement and MOS (c) Relationship between Correlation Measurement and MOS (d) Relationship between HVS Measurement and MOS (e) Relationship between Spectrum Measurement and MOS