

# IMAGE AND VIDEO QUALITY ASSESSMENT BASED ON THE SIMILARITY OF EDGE PROJECTIONS

Dong-O Kim<sup>1</sup>, Rae-Hong Park<sup>1</sup>, and Dong-Gyu Sim<sup>2</sup>

<sup>1</sup>Department of Electronic Engineering, Sogang University, Seoul, Korea  
{hyssop, rhpark}@sogang.ac.kr

<sup>2</sup>Department of Computer Engineering, Kwangwoon University, Seoul, Korea  
dgsim@kw.ac.kr

## ABSTRACT

The goal of image or video quality assessment is to evaluate if a distorted image or video is of a good quality by quantifying the difference between the original and distorted images or videos. In this paper, to assess the visual quality of an arbitrary distorted image or a compressed video, visual features of the image or video are compared with those of the original image or video instead of direct comparison of two images or videos. As visual features, we use directional edge projections that are simply obtained by projecting vertical and horizontal edges detected by vertical and horizontal Sobel masks, respectively. Then, to assess the image or video quality, edge projections are compared using the similarity measures of one-dimensional histograms such as the histogram difference, histogram intersection, Kullback-Leibler divergence,  $\chi$ -square test, and Bhattacharyya distance. Experimental results using LIVE data set and 140 video clips that are compressed with H.263 and H.264/AVC show the effectiveness of the proposed methods through the comparison with conventional algorithms such as the peak signal-to-noise ratio (PSNR), structural similarity, mean singular value decomposition, and edge PSNR (EPSNR) methods.

## KEYWORDS

*Image Quality Assessment, Video Quality Assessment, Edge Projections, Similarity Measures*

## 1. INTRODUCTION

Given a distorted image, objective image quality assessment (IQA) evaluates how much the image is degraded by comparing it with the original image, i.e., how much the image is different from the original image in terms of intensity or color. IQA algorithms have been mainly developed for testing the performance of the image or video coding systems. Currently, commercial multimedia services become common and high quality services become inevitable for commercial success. Thus, IQA algorithms could be a crucial component to improve user satisfaction for the multimedia and network services.

The mean square error (MSE) and peak signal-to-noise ratio (PSNR) can be considered as representative IQA methods. Currently, these two methods are widely used for evaluating the quality of an image or video. However, they simply measure the difference of intensity or color values between an original image and its distorted version without considering the human visual property. Therefore, they may not faithfully reflect the subjective IQA. To alleviate this problem, objective IQA algorithms have been developed and generally, they are classified into two categories according to which information is used for assessing the image quality: structural information-based [1–8] and perceptual information-based [9–13] methods.

First, unlike the PSNR and MSE that consider the only difference of intensity or color values for the assessment of the image quality, the structural information-based IQA methods use the

structural information of the image [1–8]. That is, the assessment of the image quality is to measure how much the structures of the image are distorted. Wang *et al.* [1] presented the structural similarity (SSIM), which uses the structural information such as the mean, variance, and covariance of intensity values of the original and distorted images. The enhanced version of the SSIM, the gradient-based SSIM (GSSIM) [2], was presented, which makes use of the gradient features as the structural information. In a similar way, the edge-based SSIM (ESSIM) was proposed [3]. It uses a histogram with eight bins corresponding to eight edge directions. The IQA algorithm proposed by Wang and Simoncelli [4] is based on the fact that statistical properties of wavelet coefficients in LH, HL, and HH bands are changed when an image is distorted. For measuring the quality of a distorted image, Won adopted the edge histogram descriptor (EHD) of moving picture expert group (MPEG)-7 standard and compared the EHD of the original image with that of the distorted one [5]. Also, Shnayderman *et al.* utilized singular value decomposition (SVD) to assess the quality of an image [6], which is reasonable because SVD well represents texture information as structural features of an image. To assess the image quality, several metrics presented in [7] measure coding artifacts caused by international coding standards such as joint photographic expert group (JPEG), JPEG2000, moving picture expert group (MPEG)-2, MPEG-4, H.263, and H.264/AVC. These IQA methods measure the blurriness, blockiness, jerkiness, and so on.

Second, the perceptual information-based methods measure the distortion that the human vision system can perceive for assessing the image quality [8–12]. An edge PSNR (EPSNR) method was proposed [8], which is motivated by the observation that the human perceives sensitively intensity variation around edges of an image. The EPSNR was adopted as Annex B of ITU-T recommendation J.144. A perceptually-tuned metric based on the wavelet transform and a measure of the intra- and inter-channel visual masking effect was also developed by Charrier and Eude [9]. The perceptual quality significance map was presented in evaluating the visual quality of a distorted image [10], which uses the concept of visual attention in the human visual system (HVS). Kusuma and Zepernick proposed the hybrid image quality metric based on the human visual perception [11], combining various image artifact measures such as the blocking measure, blur measure, edge activity measure, gradient activity measure, and intensity masking detection. Winkler proposed the spatio-temporal contrast gain control model [12], which achieves a close fit to contrast sensitivity and contrast masking data from several psychophysical experiments.

Generally, IQA methods can be simply extended for video quality assessment (VQA) [13]. The VQA can be regarded as the aggregation of evaluation of each frame without considering the relationship between frames. These approaches such as simple extensions of the PSNR, SSIM, and so on, have been often used for VQA. Unlike simple extensions, VQA methods based on temporal information have been presented [14–16], in which the motion clues are explicitly utilized. The temporal information used in [14] is the standard deviation of frame difference values. In [15], the motion level was defined according to the magnitude of motion vectors and was multiplied with SSIM values as weights for assessing the quality of videos. Instead of using motion vectors, the VQA metric using the perceptual motion speed was also presented [16].

In this paper, we propose an IQA method with a simple concept. In the proposed IQA method, to assess the visual quality of an arbitrary distorted image, visual features of the degraded image are compared with those of the original image, instead of direct comparison of two images. As the visual features, we use the directional edge projections, which have been widely used in the pattern recognition or computer vision. Also, by simply extending the proposed edge projection-based IQA method to a VQA method, we quantify the quality of a video compressed with H.263 and H.264/AVC. The proposed VQA method is motivated from the observation that when an image is degraded, the edge projection extracted from the image may be altered. This paper presents the VQA method that assesses the quality of a video by measuring the similarity between

two edge projections of the original and distorted videos, in which it is assumed that two videos are aligned beforehand with the same number of frames.

The rest of the paper is organized as follows. Section 2 introduces the proposed IQA method using the similarity of edge projections and extends it to a VQA method. In Section 3, experimental results of the proposed IQA and VQA methods are shown and discussed. Finally, Section 4 gives conclusions.

## 2. PROPOSED IQA AND VQA METHODS BASED ON THE EDGE PROJECTIONS

It is observed that the human eye well perceives intensity changes in regions around edges if a given image is of good quality, whereas in flat regions if a given image is of poor quality. These observations can be exploited to quantify the image and video quality. To measure these changes in a distorted image is equivalent to assessing the quality of the image. In this paper, we propose an IQA method that utilizes one-dimensional (1-D) directional vertical and horizontal edge projections as visual features that can compute these changes. Figure 1 illustrates the block diagram of the proposed IQA algorithm. As shown in Figure 1, the proposed IQA algorithm mainly consists of three steps: 1) detecting directional edges of an image using Sobel masks, 2) computing edge projections of vertical and horizontal edges of the image and projecting them along the vertical and horizontal directions, respectively, and 3) measuring the similarity between edge projections of a distorted image and transmitted edge projections of the original image. In this section, these three steps are described.

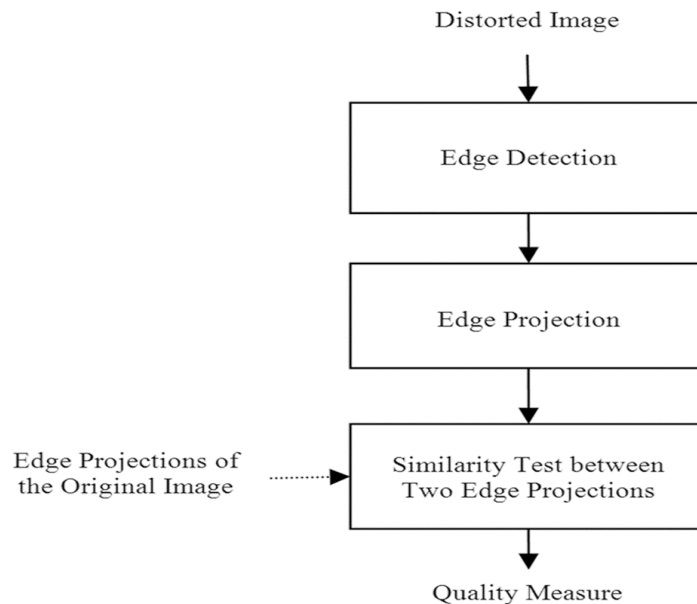


Figure 1. Block diagram of the proposed IQA algorithm.

### 2.1. Edge Detection

An image is frequently represented by edge features that describe the relation between a pixel and its neighboring pixels. Edge features have been often used for assessing the image quality [2][3][5][8]. To obtain edge projections, an edge projection technique considering edge directions is used. In the proposed method, we employ the Sobel masks for detecting vertical and horizontal

edges. For a given image  $I(i,j)$ , the vertical and horizontal gradient magnitude images  $G_v(i,j)$  and  $G_h(i,j)$  are defined as, respectively,

$$G_v(i, j) = |I(i-1, j-1) + 2I(i, j-1) + I(i+1, j-1) - (I(i-1, j+1) + 2I(i, j+1) + I(i+1, j+1))| \quad (1)$$

$$G_h(i, j) = |I(i-1, j-1) + 2I(i-1, j) + I(i-1, j+1) - (I(i+1, j-1) + 2I(i+1, j) + I(i+1, j+1))|, \quad (2)$$

where subscripts  $v$  and  $h$  denote the vertical and horizontal directions, respectively. After obtaining the gradient magnitude images, each directional edge map is obtained by separately thresholding each of the gradient magnitude images with an empirical threshold.

Generally, a single edge map of an image is obtained by considering the gradient magnitude. However, in the proposed IQA method, we utilize two types of directional edge maps: vertical and horizontal ones. By considering the direction of edges in projecting them along the vertical and horizontal directions, a better performance of the proposed IQA method can be expected. For example, blocky artifacts appeared in a compressed image produce intensity changes along both vertical and horizontal directions at pixels around block boundaries. In this case, if only the gradient magnitude is used to obtain a single edge map, it is difficult to faithfully reflect two directional edge changes in an edge map. Whereas, when both the magnitude and directions of the gradient images are used, these changes can be effectively reflected in two directional edge maps.

## 2.2. Edge Projections along the Vertical and Horizontal Directions

Let  $E_v(i,j)$  and  $E_h(i,j)$  denote the binary vertical and horizontal edge maps, respectively, which are obtained by thresholding the gradient images  $G_v(i,j)$  and  $G_h(i,j)$ , respectively. Vertical and horizontal edge projections  $p_v$  and  $p_h$  are defined as [17]

$$p_v(j) = \sum_{i=1}^{N_r} E_v(i, j) \Big/ \sum_{i=1}^{N_r} \sum_{j=1}^{N_c} E_v(i, j) \quad (3)$$

$$p_h(i) = \sum_{j=1}^{N_c} E_h(i, j) \Big/ \sum_{i=1}^{N_r} \sum_{j=1}^{N_c} E_h(i, j), \quad (4)$$

where  $N_r$  and  $N_c$  denote the numbers of rows and columns of the image, respectively. Figure 2 illustrates examples of vertical and horizontal edge projections of the first frame of the Akiyo video sequence. Figures 2(a) and 2(c) show the reference and degraded images, respectively, in which the degraded image is obtained by compressing the original image with H.264/AVC. Figures 2(b) and 2(d) show the vertical and horizontal edge maps and their edge projections of the original and distorted images, respectively. Figure 2 shows that two sets of edge maps of the original and distorted images are different from each other, which give two different edge projections. Thus, it is noted that the vertical and horizontal edge projections can be used for IQA.

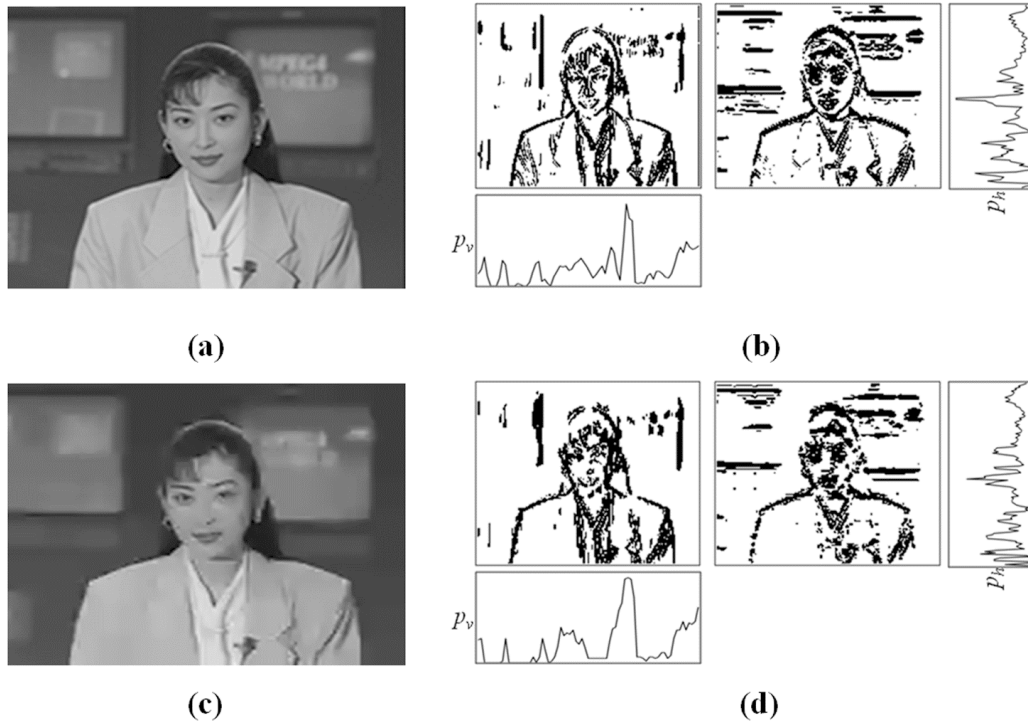


Figure 2. Examples of vertical and horizontal edge projections (Akiyo, 1st frame). (a) reference image, (b) vertical and horizontal edge maps of (a) and their edge projections, (c) degraded image, (d) vertical and horizontal edge maps of (c) and their edge projections.

However, the IQA methods using the edge features have a common drawback, the selection of the number of the edge or equivalently the selection of threshold values. In this paper, the problem is to be resolved. We empirically determine the number of edge pixels, which will be discussed in Section 3. Conclusively speaking, the number of edge pixels that are used for assessing the image quality is increased proportionally to the image size, which will be also discussed in Section 3.

In this paper, for the assessment of the image quality, we compute the vertical and horizontal edge projections of the original and distorted images and measure the similarity between them. For assessing the quality of the distorted video, edge projections of each frame of the distorted video are computed and compared with those obtained for the corresponding frame of the original video. The edge projections can be normalized by the total number of edge pixels, then the normalized edge projections can be regarded as a one-dimensional (1-D) normalized histogram, as defined in (3) and (4). For comparison of edge projections, five similarity measures of 1-D histograms (histogram difference, histogram intersection, Kullback-Leibler divergence,  $\chi$ -square test, and Bhattacharyya distance) [18–20] are used in this paper, which are briefly reviewed in Section 2.3.

### 2.3. Similarity Measures between Two Edge Projections

The similarity between edge projections obtained from the original and distorted images represents how much the distorted image is different from the original image, i.e., how much the quality of the distorted image becomes degraded. In this section, five histogram similarity

measures are introduced [18–20], which are evaluated for assessing the quality of a distorted image with the proposed projection features.

### 2.3.1. Histogram Difference

The simplest similarity measure between two 1-D histograms is the histogram difference [18][19]. This measure can be obtained by computing differences of frequencies of occurrence at every bin between two normalized histograms (i.e., probability density functions (pdfs)) and by accumulating all absolute difference values. The histogram difference  $H_{\text{diff}}$  is defined as

$$H_{\text{diff}} = \sum_{j=1}^{N_c} |p_v^o(j) - p_v^d(j)| + \sum_{i=1}^{N_r} |p_h^o(i) - p_h^d(i)|, \quad (5)$$

where superscripts  $o$  and  $d$  denote the original and distorted images, respectively. Note that we have two types of edge projections: vertical and horizontal.

### 2.3.2. Histogram Intersection

The histogram intersection [18][19] is a general similarity measure between two histograms and is defined as

$$H_{\text{int}} = \sum_{j=1}^{N_c} \min(p_v^o(j), p_v^d(j)) + \sum_{i=1}^{N_r} \min(p_h^o(i), p_h^d(i)). \quad (6)$$

### 2.3.3. Kullback-Leibler Divergence (KLD)

The cross entropy  $H_{\text{CE}}(o, d)$  is used to compute the mutual information, which is defined as

$$H_{\text{CE}}(o, d) = \sum_{j=1}^{N_c} p_v^o(j) \log \frac{p_v^o(j)}{p_v^d(j)} + \sum_{i=1}^{N_r} p_h^o(i) \log \frac{p_h^o(i)}{p_h^d(i)}. \quad (7)$$

$H_{\text{CE}}(d, o)$  can be similarly defined.  $H_{\text{CE}}(o, d)$  and  $H_{\text{CE}}(d, o)$  have different values. Therefore, for a symmetric similarity measure, a KLD  $H_{\text{KLD}}$  is defined as [18][19]

$$H_{\text{KLD}} = A \cdot H_{\text{CE}}(o, d) + B \cdot H_{\text{CE}}(d, o), \quad (8)$$

where  $A$  and  $B$  are constants, with the same value of 0.5 in our experiments. By averaging two cross entropies, the KLD has the symmetry property.

### 2.3.4. $\chi$ -square Test

$\chi$ -square test [18][19] is a good measure to evaluate how much two histograms are different from each other. In this paper, the symmetric  $\chi$ -square test is used, which is computed by dividing the squared difference of two histograms with their sum. That is, the symmetric  $\chi$ -square test  $H_\chi$  is defined as

$$H_\chi(o, d) = \sum_{j=1}^{N_c} \frac{(p_v^d(j) - p_v^o(j))^2}{p_v^d(j) + p_v^o(j)} + \sum_{i=1}^{N_r} \frac{(p_h^d(i) - p_h^o(i))^2}{p_h^d(i) + p_h^o(i)}. \quad (9)$$

### 2.3.5. Bhattacharyya Distance [20]

Before introducing the Bhattacharyya distance, let us introduce the Bhattacharyya coefficient. The Bhattacharyya coefficient is a divergence measure defined as the cosine of the angle between two unit vectors. The Bhattacharyya coefficient  $\rho(p_v^o, p_v^d)$  for the vertical directional edge projections is defined as

$$\rho(p_v^o, p_v^d) = \sum_{j=1}^{N_c} \sqrt{p_v^o(j) p_v^d(j)}. \quad (10)$$

Similarly,  $\rho(p_h^o, p_h^d)$  can be defined. Then, the Bhattacharyya distance  $H_B$  is defined as

$$H_B(o, d) = A \cdot \sqrt{1 - \rho(p_v^o, p_v^d)} + B \cdot \sqrt{1 - \rho(p_h^o, p_h^d)}, \quad (11)$$

where  $A$  and  $B$  are constants as explained in Eq. (8), with the same value of 0.5 used in our experiments.

## 2.4. Proposed IQA Method

For assessing the quality of a distorted image or video, IQA or VQA methods use visual features of an image or video instead of using the full image or video. In general objective IQA methods, metrics that quantitatively compute differences between the original and distorted images are used to evaluate the quality of the distorted image. However, these metrics are not necessarily correlated well with the subjective IQA. Consequently, development of the objective IQA measures that are consistent with those of the subjective IQA is desirable. In this paper, the three types of the regressions (linear, logistic, and non-linear regressions) [21] are performed to explain the relationship between the mean opinion score (MOS), the subjective IQA, and the proposed quality metric (edge projection-based quality metric: EPQM).

First, the proposed EPQM fitted by a linear regression is written as

$$EPQM_X^l = a_l \cdot H_X + b_l, \quad (12)$$

where the superscript  $l$  represents the linear regression and the subscript  $X$  denotes one of five similarity measures ( $EPQM_{diff}$ ,  $EPQM_{int}$ ,  $EPQM_{KLD}$ ,  $EPQM_X$ , and  $EPQM_B$ ) using 1-D histograms, with constants  $a_l$  and  $b_l$  obtained through the linear regression of the EPQM and MOS. Second, the proposed EPQM fitted by the logistic regression is defined as

$$EPQM_X^g = \frac{a_g}{1 + e^{-b_g(H_X - c_g)}}, \quad (13)$$

where the superscript  $g$  represents the logistic regression and  $a_g$ ,  $b_g$ , and  $c_g$  are constants computed by fitting the logistic function. Finally, the proposed EPQM fitted by the non-linear regression is defined as

$$EPQM_X^n = a_n \cdot H_X^3 + b_n \cdot H_X^2 + c_n \cdot H_X + d_n, \quad (14)$$

where the superscript  $n$  represents the non-linear regression and  $a_n$ ,  $b_n$ ,  $c_n$ , and  $d_n$  denote constants obtained through the non-linear regression.

For testing the performance of the proposed EPQM obtained after fitting by each regression method, the fitted versions of the proposed EPQM are compared with the difference MOS (DMOS), the subjective IQA method, through the fidelities such as the Pearson correlation coefficient (CC), the root mean squared error (RMSE), and so on, which were recommended by video quality expert group (VQEG) [21].

## 2.5. Extension to a VQA Method

We simply extend the proposed EPQM IQA algorithm to VQA for evaluating the quality of arbitrarily distorted videos. Similarly, five similarity measures using edge projections proposed in this paper are computed over all frames of a video clip. For example, the proposed EPQM for VQA fitted by the linear regression is simply defined as

$$EPQM_X^l = a \cdot \frac{1}{N_f} \sum_k^{N_f} H_X^k + b, \quad (15)$$

where  $N_f$  denotes the number of frames in a video clip,  $H^k$  is the edge projection of the  $k$ th frame, and the subscript  $X$  represents one of five similarity measures. Also the proposed EPQM for VQA fitted by the logistic and non-linear regressions can be extended in the same manner.

## 3. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this paper, we propose an IQA method and simply extend it to a VQA method. To evaluate the performance of the proposed EPQM as the IQA, we used the Laboratory for Image and Video Engineering (LIVE) data set [22], which includes 29 reference images, 982 degraded images, and their DMOS values. The distorted images (the image size of typically 768x512) are degraded by JPEG2000, JPEG, white noise in the RGB components, Gaussian blur, and transmission errors in the JPEG2000 bit stream using a fast-fading Rayleigh channel model. Also, 140 video clips with common intermediate format (CIF) and quarter CIF (QCIF) video formats compressed with H.263 (TMN 2.0) and H.264/AVC (JM 9.8) are used for evaluating the performance of the proposed VQA method.

Table 1 lists the specification of the video clips used in experiments. QCIF and CIF video clips were compressed at various bit rates from 12 kbps to 270 kbps and at from 64 kbps to 512 kbps, respectively. Thirty evaluators assessed the quality of each video clip to obtain the MOS. By using the double-stimulus continuous quality-scale (DSCQS) algorithm presented in ITU-T recommendation BT.500-11 [23], DMOS values are obtained for subjective IQA. The performance of the proposed EPQM is evaluated by measuring how much the proposed EPQM is correlated or consistent with the DMOS. First, to evaluate the performance of the proposed EPQM as an IQA method, the proposed method was compared with the conventional IQA algorithms such as the PSNR, SSIM [1], and MSVD [6] in terms of the Pearson CC and RMSE, which were recommended by VQEG [23]. Figure 3 shows the Pearson CCs between the DMOS values and EPQM values of the proposed methods as a function of the number of edges for LIVE data set (in case of the linear regression, from 40,000 to 160,000 edges by an increment of 1,000 whereas in cases of the logistic and non-linear regressions, from 40,000 to 90,000 edges by an increment of 1,000). In Figure 3, the maximum Pearson CCs for each type of the regression methods are marked with the vertical lines. Figure 3(a) shows that when the number of edges is equal to about 84,000 (marked by the vertical line), the Pearson CCs of  $EPQM_{diff}^l$  and  $EPQM_{int}^l$  are largest. As shown in Figure 3(b), when the number of edges is about 68,000 (marked by the vertical lines),  $EPQM_{diff}^o$  and  $EPQM_{int}^o$  have the largest Pearson CC values. Also, Figure 3(c) shows that  $EPQM_{diff}^n$  and  $EPQM_{int}^n$  have the largest Pearson CC values with about 67,000 edges (marked by the vertical lines). Table 2 lists the largest Pearson CCs between DMOS values and



EPQM values according to the type of regressions. As shown in Figure 3 and Table 2, in the case of the linear regression, EPQM<sub>B</sub> has the best performance among five EPQMs whereas EPQM<sub>diff</sub> and EPQM<sub>int</sub> have the largest Pearson CC values in the cases of logistic and non-linear regressions. Among five proposed metrics, the EPQM<sub>diff</sub> is selected for the performance comparison with the conventional IQA methods (afterward EPQM<sub>diff</sub> will be called EPQM) since it generally shows the best performance and EPQM<sub>diff</sub> has a lower computational load than EPQM<sub>int</sub>, although both EPQM<sub>diff</sub> and EPQM<sub>int</sub> have the same performance in terms of the Pearson CC.

Table 1. Specification of the video clips used in experiments.

Codec	H.263, H.264/AVC	
Resolution	QCIF (176×144)	CIF (352×288)
Sequences	Carphone Coast guard Container Foreman Mobile News Stefan Table tennis Akiyo Hall monitor Mother and daughter	Carphone Coast guard Container Foreman Mobile News Stefan Table tennis Silent Weather
Bitrate	12 kbps ~ 270 kbps	64 kbps ~ 512 kbps

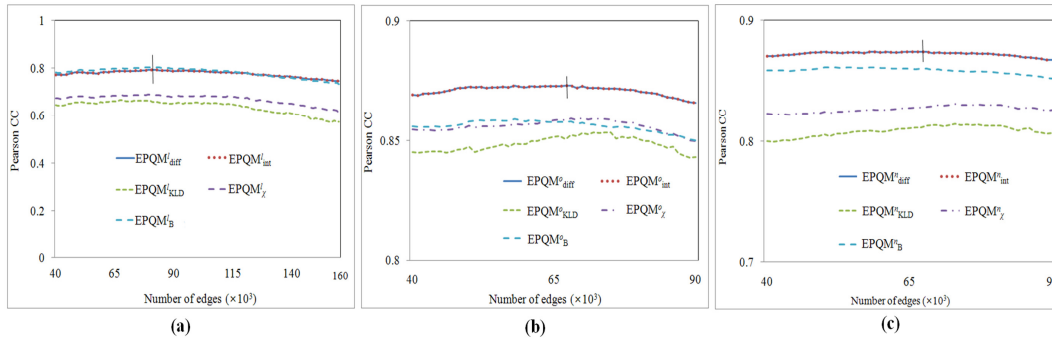


Figure 3. Pearson CC between DMOS and EPQM values as a function of the number of edges (LIVE data set). (a) linear regression, (b) logistic regression, (c) non-linear regression.

Table 2. Performance comparison of EPQMs in terms of the Pearson CC (LIVE data set).

		EPQM <sub>diff</sub>	EPQM <sub>int</sub>	EPQM <sub>KLD</sub>	EPQM <sub>χ</sub>	EPQM <sub>B</sub>
Pearson CC	Linear regression	0.793	0.793	0.666	0.690	0.804
	Logistic regression	0.873	0.873	0.853	0.860	0.860
	Non-linear regression	0.874	0.874	0.839	0.831	0.861

Figure 4 illustrates scatter plots of the DMOS values and fitted versions of four IQA metrics by the non-linear regression for LIVE data set. In the case of Figure 4, if an IQA metric coincides with the DMOS, the scatter plot of them becomes a line. Therefore, as shown in Figure 4, the EPQM shows the best performance among four methods considered. Table 3 shows the performance comparison of the proposed and conventional methods, in terms of the Pearson CC and RMSE, for LIVE data set. As shown in Table 3, in the case of the linear regression, the Pearson CC of the proposed EPQM is larger than those of the SSIM and MSVD whereas the PSNR is more correlated with the DMOS than the proposed EPQM. In terms of the RMSE, the result is similar to that observed in terms of the Pearson CC. However, in cases of both logistic and non-linear regressions, the proposed EPQM has larger Pearson CC values than the PSNR, SSIM, and MSVD. Also, in terms of the RMSE, the proposed EPQM has the best performance among four image quality metrics. All image quality metrics have the best performance in the case of the non-linear regression, except for the MSVD. In the case of the MSVD, the best performance is shown when fitting with the logistic regression.

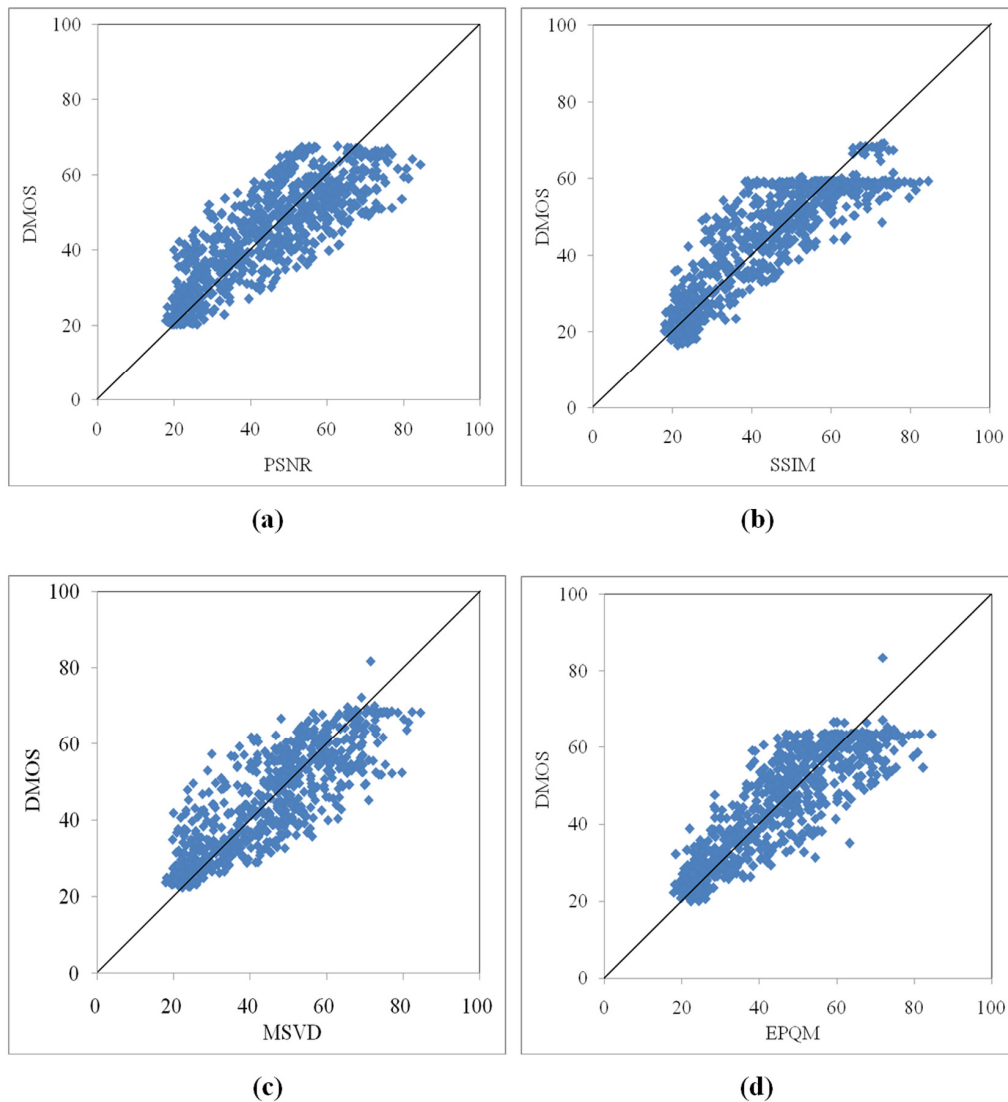


Figure 4. Scatter plots of the DMOS values and fitted versions of four IQA metrics by non-linear regression (LIVE data set). (a) PSNR, (b) SSIM, (c) MSVD, (d) proposed EPQM.

Table 3. Performance comparison of the proposed and conventional methods in the case of the non-linear regression, in terms of the Pearson CC and RMSE (LIVE data set).

		PSNR	SSIM	MSVD	EPQM
Linear regression	Pearson CC	0.800	0.737	0.752	0.793
	RMSE	9.665	10.890	10.621	9.803
Logistic regression	Pearson CC	0.814	0.863	0.846	0.873
	RMSE	9.360	8.136	8.588	7.851
Non-linear regression	Pearson CC	0.819	0.869	0.849	0.874
	RMSE	9.249	7.970	8.520	7.827

Table 4 shows the performance comparison of the proposed EPQM and conventional methods in the case of the non-linear regression, in terms of the Pearson CC, for LIVE data set according to different types of distortion. The proposed EPQM has the best performance for distortion caused by transmission errors in the JPEG2000 bit stream using a fast-fading Rayleigh channel model. In the quality evaluation of the distorted image set, the MSVD gives the best performance for images degraded by JPEG2000 and additive white noise. As shown in Table 4, it is also noted that the SSIM does better reflect JPEG and Gaussian blur artifacts than other methods do. The performance of the proposed EPQM is not significantly better than those of the conventional methods. However, there is one thing to point out, which can be noted from Table 3. To assess the quality of an arbitrary image, an objective IQA algorithm should be similar with DMOS values, regardless of the types of distortion or images. From this point of view, the proposed EPQM is better than the PSNR, SSIM, and MSVD.

Table 4. Performance comparison of the proposed and conventional methods in the case of the non-linear regression, in terms of the Pearson CC according to different types of distortion (LIVE data set).

	PSNR	SSIM	MSVD	EPQM
JPEG2000	0.895	0.928	0.937	0.909
JPEG	0.855	0.926	0.888	0.884
White noise	0.917	0.900	0.968	0.876
Blurring	0.774	0.851	0.803	0.836
Fast fading	0.882	0.923	0.922	0.926

Second, to evaluate the performance of the proposed EPQM as a VQA method, the proposed method is compared with the conventional VQA algorithms such as the PSNR, SSIM, MSVD, and EPSNR [4]. Figure 5 shows the Pearson CC between the DMOS value and each EPQM metric as a function of the number of edges (in the case of QCIF video, from 1,000 to 8,000 by an increment of 100 and in the case of CIF video, from 4,000 to 32,000 by an increment of 400). As shown in Figure 5,  $EPQM_{int}$  and  $EPQM_{diff}$  have the similar performance and the best performance is obtained when the number of edges is equal to about 4,500 (marked by the vertical lines) regardless of the type of regression. These results are similar to those of the proposed IQA, which is previously explained in results of the IQA. As shown in Figures 3 and 5, the larger an image

size is, the larger the number of edges is. Also, in the case of LIVE data set in Figure 3, an image typically has the size of 768×512. That is, an image of LIVE data set is 15.51 times larger than the image size of QCIF video. In the case of the non-linear regression, the number of edges required to obtain the best performance is about 67,000 and 4,500 for the LIVE data set and video clips, respectively. The number of edges for video clips is 14.81 times smaller than that for the LIVE data set. Therefore, Figures 3 and 5 show that the number of edges required for evaluating the image/video quality using the proposed EPQM is approximately proportional to the image size.

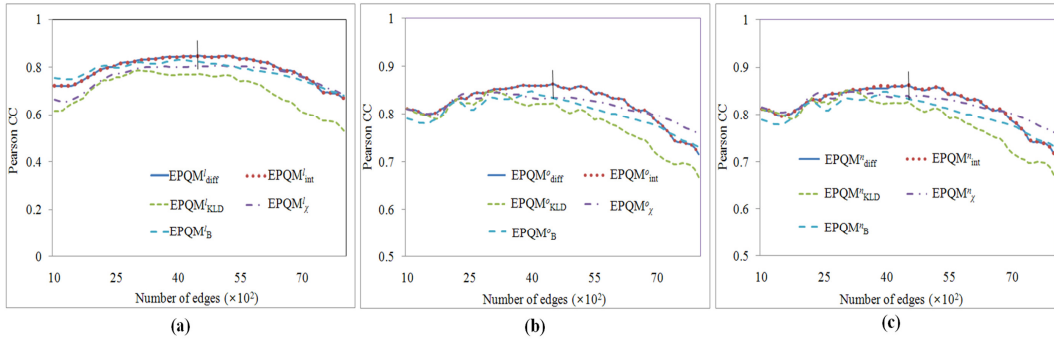


Figure 5. Pearson CC between DMOS and EPQM values as a function of the number of edges (140 video clips). (a) linear regression, (b) logistic regression, (c) non-linear regression.

Table 5 lists the largest Pearson CCs between DMOS values and EPQM values according to the type of regressions. Similarly to experimental results of the IQA, Table 5 shows that EPQM<sub>diff</sub> and EPQM<sub>int</sub> have the best performance among five EPQMs. For evaluating the performance of the proposed EPQM by comparing with the conventional VQAs, EPQM<sub>diff</sub> is selected because of the same reason described in the IQA method.

Table 5. Performance comparison of EPQMs in terms of the Pearson CC (140 video clips).

		EPQM <sub>diff</sub>	EPQM <sub>int</sub>	EPQM <sub>KLD</sub>	EPQM <sub>χ</sub>	EPQM <sub>B</sub>
Pearson CC	Linear regression	0.848	0.848	0.785	0.805	0.831
	Logistic regression	0.863	0.863	0.850	0.846	0.849
	Non-linear regression	0.863	0.863	0.854	0.849	0.849

Figure 6 represents the scatter plots of the DMOS values and fitted versions of five VQA metrics by non-linear regression for 140 video clips. As shown in Figure 6, the EPQM shows the best performance among five methods considered. This is the similar result as given in Figure 4. Table 6 lists the performance comparison of the proposed EPQM and conventional methods in terms of the Pearson CC and RMSE. As shown in Table 6, the proposed EPQM has the best performance among five quality metrics regardless of the type of regression. Table 7 shows the performance comparison of the proposed and conventional methods according to the video codec (H.263 or H.264/AVC) and format (CIF or QCIF) in the case of the non-linear regression. Table 6 shows that the proposed EPQM has the best performance.

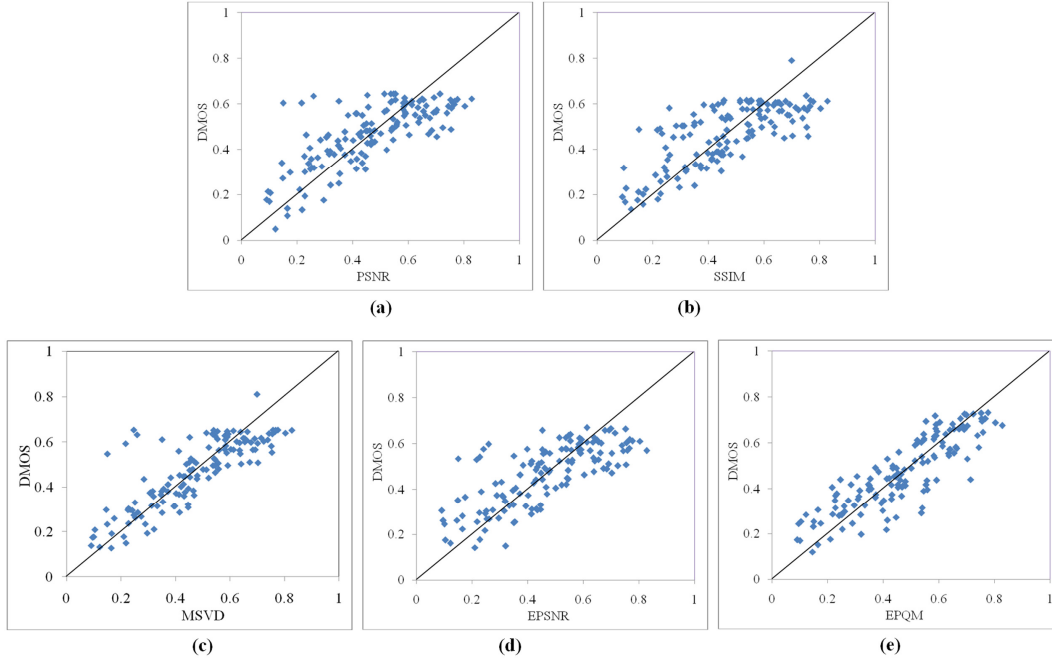


Figure 6. Scatter plots of the DMOS values and fitted versions of five VQA metrics by non-linear regression (140 video clips). (a) PSNR, (b) SSIM, (c) MSVD, (d) EPSNR, (e) proposed EPQM.

Table 6. Performance comparison of the proposed and conventional methods in terms of the Pearson CC and RMSE (140 video clips).

		PSNR	SSIM	MSVD	EPSNR	EPQM
Linear regression	Pearson CC	0.703	0.611	0.637	0.620	0.848
	RMSE	0.128	0.142	0.140	0.141	0.095
Logistic regression	Pearson CC	0.735	0.735	0.818	0.738	0.863
	RMSE	0.122	0.122	0.104	0.121	0.091
Non-linear regression	Pearson CC	0.745	0.739	0.814	0.741	0.863
	RMSE	0.120	0.122	0.105	0.122	0.091

Table 8 lists the performance comparison of the proposed EPQM and conventional VQA methods according to each video clip in terms of the Pearson CC and RMSE. In Table 8, on the whole the Pearson CCs appear large, from which it seems that all five VQA metrics have the good performance. Also as shown in Table 8, the performance of the proposed EPQM is not better than the conventional methods except for two video clips such as Mobile and Stephen video clips. However, as previously pointed out, to assess the quality of an arbitrary image, an objective VQA algorithm should be similar with DMOS values, regardless of the type of video. From this point of view, the proposed EPQM is a better VQA metric than the PSNR, SSIM, MSVD, and EPSNR.

In terms of the Pearson CCs, the proposed EPQM has the largest average value and smallest standard deviation value for all video clips.

Table 7. Performance comparison of the proposed and conventional methods in the case of the non-linear regression, in terms of the Pearson CC and RMSE according to the codec and format (140 video clips).

			PSNR	SSIM	MSVD	EPSNR	EPQM
Codec	H.263	CC	0.726	0.673	0.849	0.734	0.858
		RMSE	0.105	0.114	0.095	0.106	0.092
	H.264 /AVC	CC	0.754	0.766	0.802	0.748	0.870
		RMSE	0.129	0.128	0.118	0.133	0.099
Format	QCIF	CC	0.787	0.793	0.874	0.731	0.886
		RMSE	0.117	0.118	0.094	0.129	0.088
	CIF	CC	0.678	0.699	0.694	0.781	0.803
		RMSE	0.122	0.129	0.125	0.105	0.099

Table 8. Performance comparison of the proposed and conventional methods in the case of the non-linear regression, according to each clip in terms of the Pearson CC (140 video clips).

	PSNR	SSIM	MSVD	EPSNR	EPQM
Akiyo	0.985	0.996	0.985	0.974	0.981
Carphone	0.976	0.985	0.961	0.650	0.969
Coastguard	0.937	0.795	0.951	0.924	0.926
Container	0.903	0.850	0.694	0.943	0.812
Foreman	0.960	0.763	0.929	0.633	0.912
Hall monitor	0.946	0.971	0.963	0.969	0.956
Mobile	0.322	0.817	0.611	0.919	0.929
Mother and daughter	0.976	0.979	0.969	0.959	0.971
News	0.964	0.974	0.984	0.979	0.944
Silent	0.999	0.995	0.998	0.943	0.992
Stephen	0.458	0.885	0.974	0.967	0.977
Table tennis	0.939	0.917	0.921	0.914	0.930
Weather	0.956	0.986	0.895	0.977	0.977
Average	0.821	0.916	0.910	0.903	0.944
Std. deviation ( $\times 10^{-2}$ )	37.110	8.455	11.924	11.856	4.711

In our experiments, we used a PC with Intel Core 2 Duo CPU 2.66GHz and 2 GB RAM and Visual Studio 6.0. For assessing the quality of 768×512 sized image, PSNR, SSIM, MSVD, and EPQM take 0.030, 0.466, 0.767, and 0.152 seconds, respectively. Our IQM is faster than other IQMs except for the PSNR method.

In summary, experimental results with the LIVE data set and various types of videos show that the proposed EPQM shows a good performance because edge projections can effectively reflect the distortion occurred in image compression, transmission, and so on. Especially, it is noted that edge projections can reflect well the artifacts caused by the video compression such as H.263 and H.264/AVC.

#### 4. CONCLUSIONS

To evaluate the quality of a distorted image, we propose an effective image quality metric, EPQM, which is based on edge projections as visual features. We also extend the EPQM into a VQA method for evaluating the quality of a video. Experimental results with the LIVE data set and various video clips show the effectiveness of the proposed EPQM in terms of the Pearson CC and RMSE. Further researches will focus on the development of the VQA method that utilizes temporal information and a motion perception model.

#### REFERENCES

- [1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, *IEEE Trans. Image Processing*, 13 (4) (2004) 600–612.
- [2] G.-H. Chen, C.-L. Yang, and S.-L. Xie, Gradient-based structural similarity for image quality assessment, in: *Proc. Int. Conf. Image Processing*, Atlanta, GA, USA, Oct. 2006, pp. 2929–2932.
- [3] G.-H. Chen, C.-L. Yang, and S.-L. Xie, Edge-based structural similarity for image quality assessment, in: *Proc. Int. Conf. Acoustics, Speech, and Signal Processing 2006*, May 2006, Toulouse, France, vol. 2, pp. 14–19.
- [4] Z. Wang and E. P. Simoncelli, Reduced-reference image quality assessment using a wavelet-domain natural image statistic model, in: *Proc. SPIE Human Vision and Electronic Imaging X*, Jan. 2005, San Jose, CA, USA, vol. 5666, pp. 149–159.
- [5] C. S. Won, Using edge histogram descriptor of MPEG-7 for the measurement of image quality and modifications, in: *Proc. SPIE Multimedia Systems and Applications IX*, Oct. 2006, Boston, MA, USA, vol. 6391, pp. 1–8.
- [6] A. Shnayderman, A. Gusev, and A. M. Eskicioglu, An SVD-based grayscale image quality measure for local and global assessment, *IEEE Trans. Image Processing*, 15, (2), (2006), 422–429.
- [7] J. Caviedes and F. Oberti, No-reference quality metric for degraded and enhanced video, in: *Proc. SPIE Conf. Video Comm. Image Process.*, July 2003, Lugano, Switzerland, vol. 5150, pp. 621–632.
- [8] ITU-R Recommendation J.144, Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference, International Telecommunication Union, 2004.
- [9] C. Charrier and T. Eude, A psychovisual color image quality metric integrating both intra and inter channel masking effect, in: *Proc. European Signal Processing Conference*, Sep. 2006, Vienna, Austria, pp. 2123–2126.
- [10] Z. Lu, W. Lin, E. Ong, X. Yang, and S. Yao, PQSM-based RR and NR video quality metrics, in: *Proc. Visual Comm. Image Process.* 2003, July 2003, Lugano, Switzerland, vol. 5150, pp. 633–640.
- [11] T. M. Kusuma and H.-J. Zepernick, A reduced-reference perceptual quality metric for in-service image quality assessment, in: *Proc. Joint First Workshop on Mobile Future and Symposium on Trends in Communications 2003*, Oct. 2003, Bratislava, Slovakia, pp. 71–74.
- [12] S. Winkler, A perceptual distortion metric for digital color video, in: *Proc. SPIE Human Vision and Electronic Imaging IV*, Jan. 1999, San Jose, CA, USA, pp. 175–184.
- [13] P. Campisi, M. Carli, G. Giunta, and A. Neri, Blind quality assessment system for multimedia communications using tracing watermarking, *IEEE Trans. Signal Process.*, 51 (4) (2003), 996–1002.
- [14] A. A. Webster, C. T. Jones, M. H. Pinson, S. D. Voran, and S. Wolf, “An objective video quality assessment system based on human perception,” in: *Proc. Human Vision, Visual Processing, and Digital Display IV*, 1993, San Jose, CA, USA, pp. 15–26.

- [15] Z. Wang, L. Lu, and A. C. Bovik, Video quality assessment based on structural distortion measurement, *Signal Processing: Image Comm.*, **19** (1) (2004), 1–9.
- [16] Q. Li and Z. Wang, Video quality assessment by incorporating a motion perception model, in: *Proc. Int. Conf. Image Processing*, Sep. 2007, San Antonio, TX, USA, vol. 2, pp. 173–176.
- [17] R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*. McGraw-Hill Inc., 1995, Singapore.
- [18] E. Wahl, U. Hillenbrand, and G. Hirzinger, Surflet-pair-relation histograms: A statistical 3D-shape representation for rapid classification, in: *Proc. Fourth Int. Conf. 3-D Digital Imaging and Modeling 2003*, Oct. 2003, Banff, Alberta, Canada, pp. 474–481.
- [19] D. V. der Weken, M. Nachtgaele, and E. Kerre, Combining neighbourhood-based and histogram similarity measures for the design of image quality measures, *Image and Vision Computing*, **25** (2007), 184–195.
- [20] D. Comaniciu, V. Rameshi, and P. Meer, Kernel-based object tracking, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **25** (5) (2003), 564–577.
- [21] ITU-R Recommendation BT.500-11, Methodology for the subjective assessment of the quality of television pictures, International Telecommunication Union, 2002.
- [22] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, LIVE image quality assessment database release 2, <http://live.ece.utexas.edu/research/quality>.
- [23] VQEG, Final report from the video quality experts group on the validation of objective models of multimedia quality assessment, 2008, <http://www.vqeg.org/>.

## Authors

Dong-O Kim received the B.S. and M.S. degrees in electronic engineering from Sogang University, Seoul, Korea, in 1999 and 2001, respectively. Currently, he is working toward the Ph.D. degree in electronic engineering at Sogang University. His current research interests are image quality assessment and physics-based computer vision for computer graphics.

Rae-Hong Park was born in Seoul, Korea, in 1954. He received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, Korea, in 1976 and 1979, respectively, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1981 and 1984, respectively. In 1984, he joined the faculty of the Department of Electronic Engineering, School of Engineering, Sogang University, Seoul, Korea, where he is currently a Professor. In 1990, he spent his sabbatical year as a Visiting Associate Professor with the Computer Vision Laboratory, Center for Automation Research, University of Maryland at College Park. In 2001 and 2004, he spent sabbatical semesters at Digital Media Research and Development Center, Samsung Electronics Co., Ltd. (DTV image/video enhancement). His current research interests are computer vision, pattern recognition, and video communication. He served as Editor for the Korea Institute of Telematics and Electronics (KITE) *Journal of Electronics Engineering* from 1995 to 1996. Dr. Park was the recipient of a 1990 Post-Doctoral Fellowship presented by the Korea Science and Engineering Foundation (KOSEF), the 1987 Academic Award presented by the KITE, and the 2000 Haedong Paper Award presented by the Institute of Electronics Engineers of Korea (IEEK), the 1997 First Sogang Academic Award, and the 1999 Professor Achievement Excellence Award presented by Sogang University.

Dong-Gyu Sim was born in Chungchung Province, Korea, in 1970. He received the B.S. and M.S. degrees in Electronic Engineering from Sogang University, Seoul, Korea, in 1993 and 1995, respectively. He also received Ph.D. degree at the same University in 1999. He was with the Hyundai Electronics Co., Ltd. from 1999 to 2000, where was involved in MPEG-7 standardization. He was a senior research engineer at Varo Vision Co., Ltd., working on MPEG-4 wireless applications from 2000 to 2002. He worked for the Image Computing Systems Lab. (ICSL) at the University of Washington as a senior research engineer from 2002 to 2005. He researched on the ultrasound image analysis and parametric video coding. He joined the Department of Computer Engineering at Kwangwoon University, Seoul, Korea, in 2005 as an Associate Professor. He was elevated to an IEEE Senior Member in 2004. His current research interests are image processing, computer vision, and video communication.