

논문 2008-45SP-3-5

에지 투영의 유사도를 이용한 압축된 영상에 대한 Reduced-Reference 화질 평가

(Reduced-Reference Quality Assessment for Compressed Videos Based
on the Similarity Measure of Edge Projections)

김 동 오*, 박 래 홍**, 심 동 규***

(Dong-O Kim, Rae-Hong Park, and Dong-Gyu Sim)

요 약

화질 평가는 원영상과 열화된 영상 간의 차이를 측정함으로써, 열화된 영상의 화질이 좋고 나쁨을 판단하는 것을 목표로 한다. 본 논문에서는 열화된 영상의 화질 평가를 위해, 원영상과 열화된 영상 전체를 비교하는 것 대신, 원영상과 열화된 영상, 각각의 특징으로 에지 투영을 이용하는 방법을 제안하였다. 여기서 에지 투영은 에지 맵에서 수직, 수평 방향으로 투영시킴으로써 얻을 수 있다. 에지 투영 시 수직, 수평 방향에 대한 그래디언트 크기를 고려함으로써, 보다 나은 화질 평가 방법을 제안하였다. 제안한 방법의 타월함을 기존의 화질 평가 방법인 structural similarity (SSIM), edge peak signal-to-noise ratio (EPSNR), 그리고 edge histogram descriptor (EHD) 방법과 비교 실험을 통해 보였다.

Abstract

Quality assessment aims to evaluate if a distorted image or video has a good quality by measuring the difference between the original and distorted images or videos. In this paper, to assess the visual quality of a distorted image or video, visual features of the distorted image are compared with those of the original image instead of the direct comparison of the distorted image with the original image. We use edge projections from two images as features, where the edge projection can be easily obtained by projecting edge pixels in an edge map along vertical/horizontal direction. In this paper, edge projections are obtained by using vertical/horizontal directions of gradients as well as the magnitude of each gradient. Experimental results show the effectiveness of the proposed quality assessment through the comparison with conventional quality assessment algorithms such as structural similarity (SSIM), edge peak signal-to-noise ratio (EPSNR), and edge histogram descriptor (EHD) methods.

Keywords : Quality assessment, MOS, VQEG, DSCQS, Edge projection

I. 서 론

객관적 화질 평가는 하나의 영상이 주어졌을 때 그 영상이 원영상에 비해 얼마나 많이 열화 되었는지, 즉

원영상과 얼마나 많이 차이가 나는 지를 이용한다. 객관적 화질 평가 방법은 일반적으로 원영상과 열화된 영상 간의 밝기값 차이를 이용한 것이 주를 이룬다. 객관적 화질 평가는 원영상의 사용 여부나 그 정도에 따라 full-reference (FR), reduced-reference (RR), 그리고 no-reference (NR) 화질 평가로 나눌 수 있다^[1~8].

FR 화질 평가 방법의 대표적인 예로 mean square error (MSE)와 peak signal-to-noise ratio (PSNR)을 들 수 있다. 현재까지 화질을 측정하는 방법으로 이 두 가지 방법이 가장 많이 사용된다. 하지만 MSE나 PSNR은 단순히 두 영상 사이의 에러만을 측정하는 것

* 학생회원, ** 평생회원, 서강대학교 전자공학과
(Department of Electronic Engineering, Sogang University)

*** 평생회원, 광운대학교 컴퓨터공학과
(Department of Computer Engineering, Kwangwoon University)

※ 이 연구에 참여한 연구자(의 일부)는 『2단계 BK21 사업』의 지원비를 받았음

접수일자: 2007년8월15일, 수정완료일: 2008년4월7일

으로, 사람의 시각 특성이 전혀 고려되지 않아 사람에게 의한 주관적 화질 평가 결과와의 연관성이 떨어진다. 이러한 문제를 해결하기 위해 영상의 밝기 차이 뿐 아니라 영상의 구조적 차이나 인지적 (perceptual) 특징 등을 이용한 방법들이 제안되었다. Wang 등은 특정 창틀 (window) 내의 밝기값의 평균과 분산 그리고 두 윈도우 간의 공분산 (covariance), 즉 창틀 내의 구조적 특성을 이용한 방법인 structural similarity (SSIM)을 제안하였다^[1]. 또한 그래디언트 (gradient) 정보를 이용하여 SSIM 방법을 개선한 gradient-based SSIM (GSSIM)이 제안되었다^[2]. 이와는 달리 사람의 시각이 영상의 경계 (edge)에 민감한 특성을 이용하여 경계 주위에서 열화가 일어난 정도를 측정하는 edge PSNR (EPSNR) 방법이 제안되었다^[3]. EPSNR 방법은 International Telecommunication Union-Telecommunication Standardization Sector (ITU-T) recommendation J.144의 Annex B로 채택되었다. 엄민영 등은 인간 시각 시스템 (human visual system)이 영상을 한 번에 전체적으로 인지하지 못하고 부분적으로 인지한다는 점을 이용해 블록 기반으로 Gabor 웨이블릿 (wavelet)을 사용하여 화질을 평가하는 방법을 제안하였다^[4]. 이러한 방법들은 모두 원영상이 존재하는 경우에만 사용 가능하기 때문에, 실질적인 멀티미디어 서비스의 수신 품질에 대한 평가가 불가능하다.

FR 화질 평가 방법과는 달리, 원영상을 분석하여 얻은, 표현량이 적은 정보만을 이용하여, 화질을 평가하는 RR 화질 평가 방법들이 제안되었는데^[5~6], 이러한 방법을 이용하면 FR 방법의 문제를 다소 해결할 수 있다. 그림 1은 일반적인 RR 화질 평가 방법에 대한 블록도이다. 영상 전체를 전송하는 대신, 영상을 대표하는 특징을 추출하고 그 특징들을 ancillary channel을 통해 전송함으로써, 원영상에 대한 전체 정보 없이 열화된 영상의 화질을 평가할 수 있다. Wang과 Simoncelli는 영상이 열화될 때, 웨이블릿 도메인 (wavelet domain)에서 LH, HL 그리고 HH 밴드에서 나타나는 계수들의 통계적 특성이 달라지는 점을 이용하는 화질 평가 방법을 제안하였다^[5]. Won은 MPEG-7 표준안의 edge histogram descriptor (EHD)를 이용하여 화질 저하로 인한 원영상과 열화된 영상의 EHD 간 차이를 비교함으로써, 화질을 평가하는 방법을 제안하였다^[6]. 에지 특징들을 이용하는 방법 외에도 Kusuma와 Zepernick은 사람의 시각적 인지를 기반으로 하는 hybrid image quality metric (HIQM)을 제안하였다^[7].

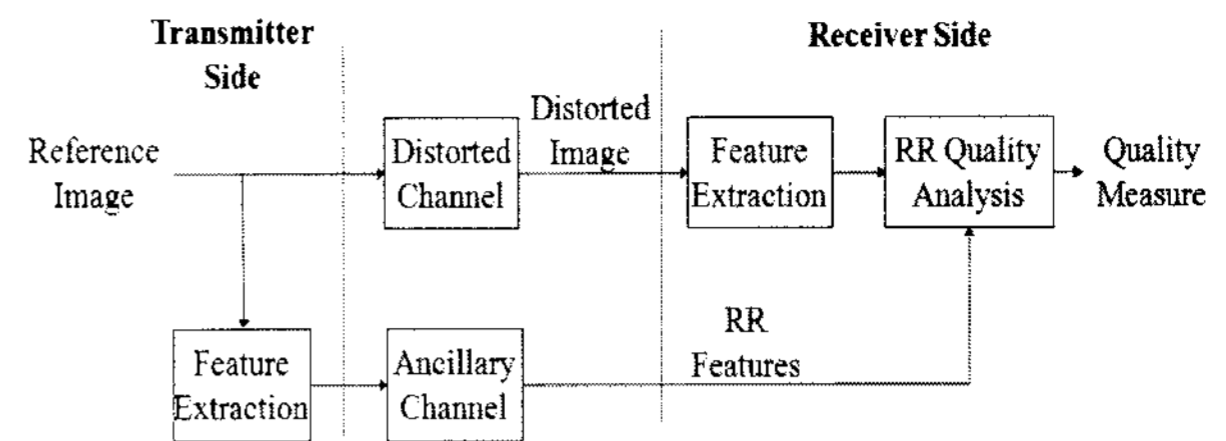


그림 1. 기존의 RR 화질 평가 방법의 블록도

Fig. 1. Block diagram of conventional RR image quality assessment.

NR 화질 평가 방법의 경우, 열화된 영상 자체의 화질을 평가하는 방법으로 원영상에 대한 정보를 전혀 사용하지 않는 화질 평가 방법을 일컫는다. 이 방법의 경우, 영상 자체에 있는 coding artifacts나 패킷 전송 채널에서 발생하는 에러 등을 측정함으로써 화질을 평가한다^[8]. 이러한 방법들에서는 영상의 희미한 (blurry) 정도나 블록화 현상 (blockiness), 급격히 움직이는 정도 (jerkiness) 등이 화질의 평가 기준이 된다.

본 논문에서는 RR 화질 평가 방법에 초점을 맞추고 있으며, 따라서 영상을 표현할 수 있는 유효한 특징을 찾아내는 것이 가장 중요하다. 본 논문에서는 에지 투영을 이용하는 화질 평가 방법을 제안한다. 이것은 압축된 영상이 가지는 블러링 현상이나 블록 현상 때문에 에지 투영이 바뀐다는 점에서 착안한 것이다. 이를 통해 수직, 수평 에지 성분을 고려하여 에지 투영을 얻고, 이 히스토그램들을 비교하여 화질을 평가하는 방법을 제안한다.

본문의 구성은 다음과 같다. II장에서는 제안한 에지 투영의 유사도 측정을 이용한 화질 평가 방법에 대해 기술한다. III장에서는 실험 결과에 대해 토의하며, IV장에서는 결론을 맺는다.

II. 제안한 에지의 1차원 히스토그램을 이용한 화질 평가 방법

본 논문에서는 사람의 시각이 에지 영역에서의 변화와 평탄한 영역에서 일어나는 화질 열화에 민감하게 반응하고, 이러한 열화들은 에지 투영에 잘 나타나는 점을 이용하여 주관적 화질 평가를 객관화 하는 방법을 제안한다. 본 논문에서 제안한 화질 평가 방법은 두 단계로 구성된다. 먼저 원영상과 열화된 영상에서 에지의 1차원 히스토그램을 얻고, 이렇게 얻은 두 히스토그램의 유사도를 측정한다. 이를 각각 II.1절과 II.2절에서 설명하였다.

1. 방향성을 고려한 에지 투영 (edge projection)

영상 $I(i,j)$ 의 수직, 수평 방향에 대한 그래디언트 영상은 Sobel 연산자를 이용해 얻을 수 있다. 수직, 수평 그래디언트 크기 영상 $G_v(i,j)$ 와 $G_h(i,j)$ 는 각각

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

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

와 같이 정의된다. 일반적으로 영상의 그래디언트 크기를 이용하여 하나의 에지 맵을 구성한 후, 그 에지 맵을 수직, 수평 방향에 대해 투영함으로써 에지 투영이 얻어진다. 그러나 이와 달리, 본 논문에서는 수직, 수평 방향에 대한 각각의 그래디언트 크기 영상을 독립적으로 이진화하는 과정을 수행하여 두 개의 에지 맵을 구한다. 이렇게 함으로써, 에지의 방향성까지 고려할 수 있고, 하나의 에지 맵을 사용할 때보다 나은 성능을 기대할 수 있다. 블록 현상의 경우, 수직, 수평에 대한 변화를 모두 갖고 있기 때문에, 방향성에 대한 고려 없이 하나의 에지맵을 생성하면 이러한 열화에 대한 요소를 화질 평가 시 적용하기 어렵다.

수직, 수평 방향에 대한 각각의 그래디언트 영상을 임의의 임계치를 적용하여 얻은 수직, 수평 에지 영상을 각각 $E_v(i,j)$ 와 $E_h(i,j)$ 라고 하면, 각각의 수직, 수평 에지 투영 p_v 와 p_h 는

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

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

로 정의되며, 여기서, N_r 과 N_c 는 각각 영상의 열과 행의 수를 나타낸다. 그림 2은 Akiyo 비디오 클립의 첫 번째 프레임 영상과 해당 프레임의 열화된 영상, 그리고 그것들로부터 얻은 수직, 수평 에지 맵과 투영을 나타낸다. 그림 2(a)는 원 Akiyo 비디오 클립의 첫 번째 프레임 영상과 H.264/AVC 코덱으로 압축된 Akiyo 비디오 클립의 해당 프레임 영상이며, 그림 2(b)는 원영상의 첫 번째 프레임에 대한 수직, 수평 에지 맵과 투영을 나타내고, 그림 2(c)는 열화된 영상에 대한 에지 맵과 투영을 나타낸다. 영상 화질이 나빠질 때 에지 맵이 달라지는 것을 확인할 수 있으며, 이로 인해 에지 투영

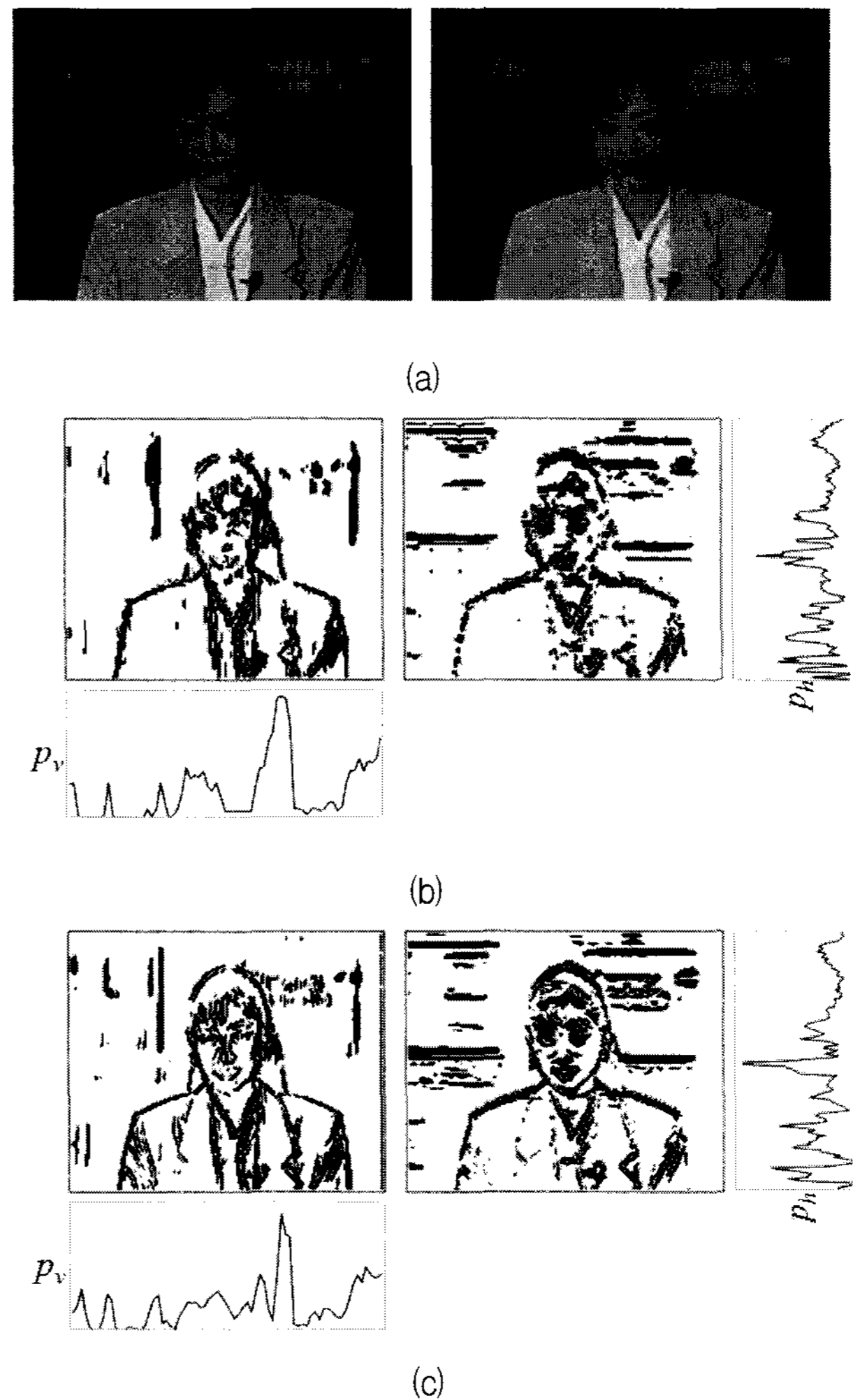


그림 2. 수직, 수평 에지 투영의 예 (Akiyo 첫 번째 프레임). (a) 원영상과 열화된 영상, (b) 원영상의 수직, 수평 에지 맵과 각각에 대한 에지 투영, (c) 열화된 영상의 수직, 수평 에지 맵과 각각에 대한 에지 투영

Fig. 2. Examples of vertical and horizontal edge projections (Akiyo, 1st frame). (a) the reference (left) and distorted (right) images (b) vertical and horizontal edge maps of the original image and their edge projections, (c) vertical and horizontal edge maps of the distorted image and their edge projections.

역시 다른 것을 볼 수 있다.

원영상에서 구한 수직, 수평 에지 투영은 ancillary channel을 통해 전송되며, 열화된 영상에 대한 두 개의 에지 투영 또한 위의 과정을 통해 구할 수 있다. 원영상과 열화된 영상, 각각에서 얻은 수직, 수평 에지 투영들의 유사도를 측정함으로써, 영상의 열화 정도를 평가할 수 있다. II.2 절에서는 유사도 측정을 위한 방법에 대해 설명한다.

2. 에지 투영의 유사도 측정

원영상과 열화된 영상으로부터 구한 에지 투영들 사이의 유사도를 측정하는 것은 열화된 영상이 원영상에 비해 얼마나 달라졌는지, 즉 열화된 영상의 화질이 얼마나 나빠졌는지를 의미한다. 이 절에서는 에지 투영들 간의 유사도를 측정하기 위해 사용한 네 가지 척도에 대해 설명한다^[9]. 정규화 된 에지 투영은 1차원 히스토그램으로 여겨질 수 있다. 따라서 1차원 히스토그램의 유사도를 측정하는 것이 에지 투영의 유사도를 측정하는 것과 동일한 문제라 할 수 있다.

가. 히스토그램 차이 (histogram difference)

두 개의 1차원 히스토그램을 비교하는 가장 간단한 방법으로 히스토그램 차이를 꼽을 수 있다. 히스토그램 차이는 말 그대로 두 히스토그램에서 같은 계급 (bin)의 빈도수 (frequency)의 차이를 모든 계급에 대해 계산하고, 그 결과를 더함으로써 얻을 수 있다. 히스토그램 차이 H_{diff} 는

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

와 같다. 여기서 위첨자 o , d 는 각각 원영상과 열화된 영상을 나타낸다.

나. 히스토그램 인터섹션 (histogram intersection)

히스토그램 인터섹션 방법은 히스토그램 간의 유사도를 나타내는 척도로서 보편적으로 사용되는 방법이며, 히스토그램 인터섹션 H_{int} 는

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

와 같이 정의된다.

다. Kullback-Leibler divergence

히스토그램 차이나 인터섹션 방법과는 달리 상호 정보 (mutual information)를 측정하는 방법으로 많이 사용되는 크로스 엔트로피 (cross entropy)를 들 수 있다. 크로스 엔트로피 H_{CE} 는

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

와 같다. $H_{CE}(d,o)$ 또한 위와 같은 방식으로 구할 수 있다. $H_{CE}(o,d)$ 와 $H_{CE}(d,o)$ 은 다른 값을 가지므로 본 논문

에서는 크로스 엔트로피에 대칭성 (symmetry)을 부여한 Kullback-Leibler divergence (KLD) 방법을 사용하여 히스토그램 간의 유사도를 측정한다. H_{KLD} 는

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

로 정의된다. 여기서, A 와 B 는 상수로서, 실험에서는 각각 0.5의 값으로 사용하였다. 이렇게 두 개의 크로스 엔트로피를 평균을 냄으로써, KLD는 대칭성을 갖는다.

라. χ -square test

χ -square test는 두 개의 히스토그램 분포가 얼마나 다른 지를 평가하는 좋은 척도가 된다. 본 논문에서는 Pearson χ -square test를 사용한다. Pearson χ -square test는 두 개의 히스토그램의 차이를 제공하고 그것을 하나의 히스토그램의 빈도수로 나눔으로써

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

와 같이 나타낼 수 있다. $H_{\chi}(d,o)$ 또한 위와 같은 방법으로 정의되며, 크로스 엔트로피의 경우와 동일하게 대칭적 χ -square test $H_{\chi sym}$ 은

$$H_{\chi sym}(o,d) = A \times H_{\chi}(o,d) + B \times H_{\chi}(d,o) \quad (10)$$

와 같다. 여기서, A 와 B 는 상수로서, 실험에서는 KLD에서 사용한 것과 동일하게 0.5의 값을 사용한다.

3. 비주얼 화질 평가 척도 (visual quality metric)

열화된 영상의 화질을 평가할 때, 영상의 모든 정보를 이용하는 방법 대신, 열화된 영상과 원영상에 대한 에지의 1차원 히스토그램 유사도를 이용한다. 대개 객관적 화질 평가 방법의 경우, 원영상과 열화된 영상 간의 차이를 정량적으로 측정함으로써, 화질 평가를 위한 척도로 사용되지만, 이러한 평가 방법들이 항상 주관적 화질 평가와의 연관성이 있다는 것을 의미하지는 않는다. 따라서 주관적 화질과 일치하는 객관적 화질 평가 방법이 필요하다. 이를 위해 본 논문에서는 주관적 화질 평가 (subjective quality assessment) 방법인 mean opinion score (MOS) 값과의 선형 회귀 (linear regression)를 통해 MOS 값과 제안한 에지 투영 기반 화질 평가 척도 (edge projection-based quality metric: EPQM)가 선형 관계에 있도록 한다. 제안한 EPQM은

$$EPQM_X = a \times H_X + b \quad (11)$$

로 정의되며, 여기서 a 와 b 는 EPQM과 MOS 값의 선형 회귀를 통해 얻을 수 있는 상수이다. 또한 아래 첨자 X 는 화질을 평가하는 척도로 사용한 각각의 히스토그램 유사도 측정 방법을 나타낸다.

III. 실험 결과 및 토의

본 논문에서 제안한 화질 평가 방법의 성능을 평가하기 위해, 표준 영상 코덱인 H.263 codec과 H.264/AVC codec를 이용하여 압축된 140개의 CIF (common intermediate format)와 QCIF (quarter CIF) 비디오 클립 (video clip)들을 사용하였다. 표 1은 실험에 사용한 140개의 비디오 클립에 대한 명세를 나타낸 것이다. 실험에 사용한 영상은 Akiyo, Carphone, Foreman 등의 표준영상들이다. 그리고 이 영상들을 압축하기 위해 사용한 H.263과 H.264/AVC는 각각 TMN 2.0과 JM 9.8 소프트웨어를 사용하여 열화된 영상을 제작하였다. 비트율의 경우, CIF 영상의 경우, 64 kbps부터 512 kbps까지, QCIF 영상의 경우는 12 kbps부터 270 kbps까지, 다양한 비트율을 사용하여 열화된 영상을 제작하였다.

표 1. 140개 비디오 클립에 대한 명세
Table 1. Specification of 140 video clips.

코덱	H.263 (63), H.264/AVC (78)	
해상도	QCIF (176×144) (94)	CIF (352×288) (46)
시퀀스	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
비트율	12 kbps ~ 270 kbps	64 kbps ~ 512 kbps
프레임율	15fps 비디오 (48), 30fps 비디오 (92)	

H.263 코덱으로 압축한 영상은 전체 140개 중에 63개의 비디오 클립이며, H.264/AVC 코덱으로 압축한 비디오 클립은 78개이다. 또한 QCIF 영상은 94개, CIF 영상은 46개로 전체 비디오 클립을 구성하였다. 또한 프레임율에 따라 전체 비디오 클립은 15fps 영상이 48개, 30fps 영상이 92개로 이루어졌다. 이러한 비디오 클립들에 대해 30여명의 평가자들로부터 MOS 값을 얻었고, 주관적 화질 측정 방법으로 ITU-T BT.500-11 권고안^[10]의 double-stimulus continuous quality-scale (DSCQS) 방법을 이용해 difference MOS (DMOS) 값을 얻었다^[11].

제안한 화질 평가 방법의 성능을 평가하기 위해 기존의 FR 화질 평가 방법인 SSIM^[1]과 EPSNR^[3], 그리고 EHD^[6] 방법들과 비교하였다. 비교 방법은 SSIM, EPSNR, EHD, 그리고 제안한 EPQM을 이용하여 각각의 비디오 클립에 대해 객관적 화질 평가 수치를 얻은 다음, 각각의 결과를 DMOS 값과의 Pearson 상관 계수와 sum of absolute error (SAE) 측면에서 비교하였다. 또한 제안한 EPQM이 RR 화질 평가 방법이므로 FR 화질 평가 방법에 비해 요구되는 데이터량의 감소 측면에서도 제안한 EPQM의 성능을 비교하였다. SSIM과 EHD 방법들은 둘 모두 정지영상에 대한 화질 평가 방법으로 제안된 것들이다. 따라서 동영상 화질 평가를 위해 단순히 매 프레임마다 화질 평가 수치를 얻고 비디오 클립 전체의 프레임 수로 정규화하여 동영상 화질 평가 수치를 얻었다. 이것은 제안한 방법의 경우와 동일하게 적용된 것으로, 제안한 EPQM의 경우에도 매 프레임마다 화질 평가 수치를 계산하고 그것을 전체 프레임 수로 정규화하는 방식으로 동영상 화질을 평가한다.

본 논문에서 제안한 네 개의 EPQM의 성능 평가를 위해 예지 수에 따른 EPQM 값과 DMOS 값과의 Pearson 상관관계를 그림 3에 그래프로 도시하였다. 이 그래프는 원영상과 열화된 영상에서 수직, 수평 에지의 개수를 1000개에서 7000개까지 100개씩 늘리면서 얻은 결과이다. 그림 3에서 볼 수 있듯이, EPQM_{diff}나 EPQM_{int}는 비슷하게 좋은 성능을 가지며 EPQM_{KLD}나 EPQM_{sym}에 비해 좋은 성능을 내는 것을 볼 수 있고, 또한 예지 수가 4500개 정도 (실선으로 표기된 부분)에서 가장 좋은 성능을 나타내는 것을 볼 수 있다. EPQM_{int} (이 후부터 EPQM은 EPQM_{int}를 칭함)와 기존의 화질 평가 방법을 비교하여 제안한 화질 평가 방법의 우수성을 확인하였다. 그림 4는 DMOS 값과 제안한 EPQM 값의 상관성을 나타낸다. 그림 4를 통해 제안한

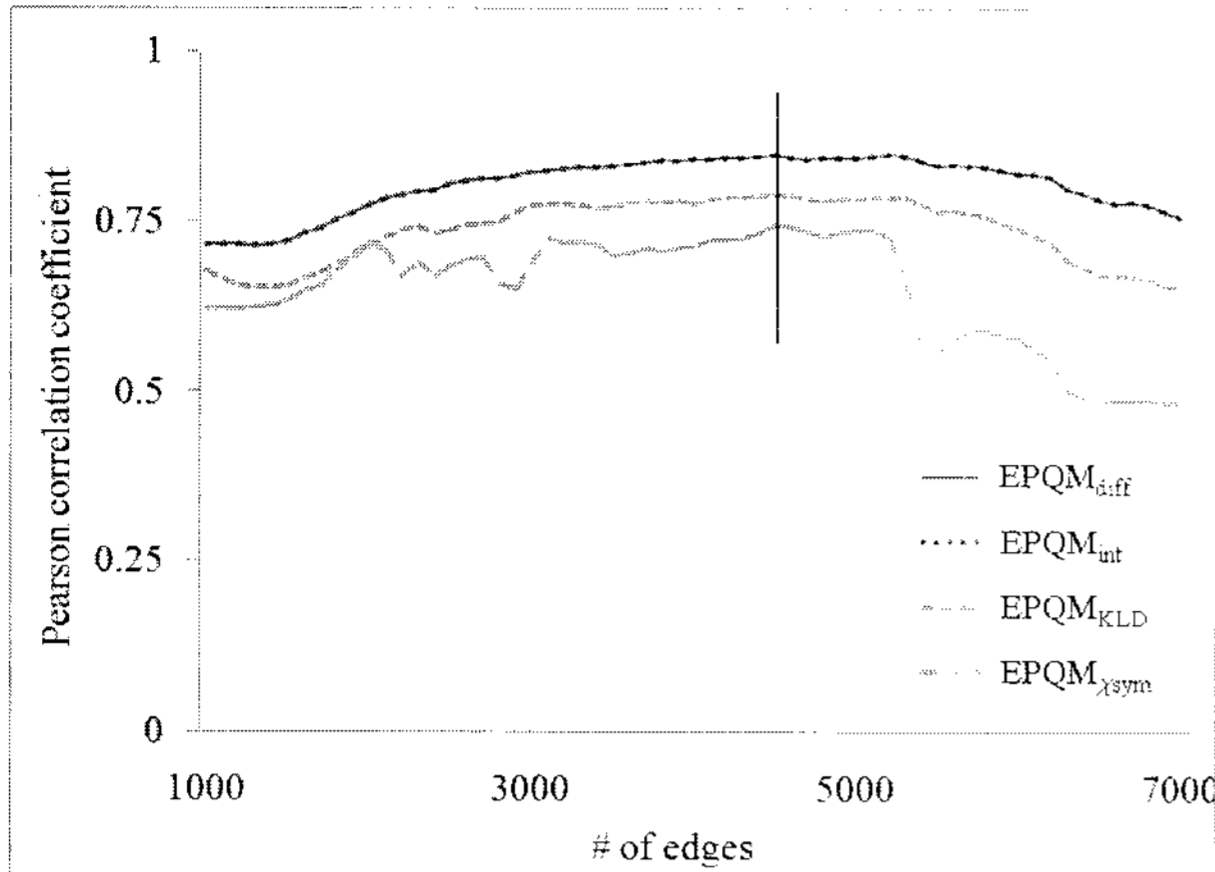


그림 3. 에지 수에 따른 DMOS 값과 EPQM 값 사이의 Pearson 상관 계수

Fig. 3. Pearson correlation coefficients between DMOS and EPQM values as a function of the number of edges.

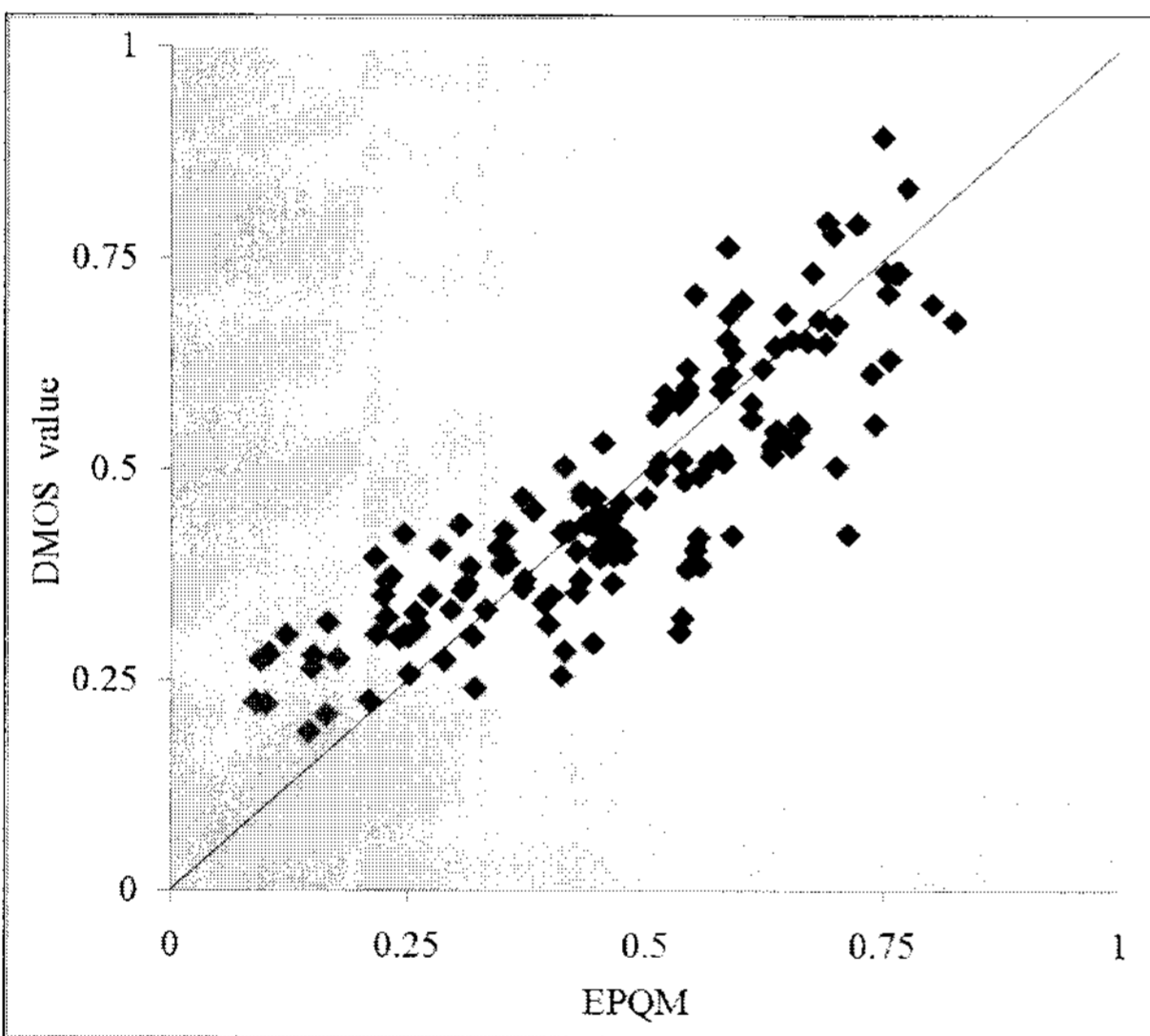


그림 4. DMOS 대비 제안한 EPQM의 분포도

Fig. 4. Scatter plot of the proposed EPQM and DMOS.

EPQM 방법이 주관적 화질 평가인 DMOS와 실선 근처에서 선형적인 관계를 보인다는 것을 알 수 있다. 이것은 제안한 화질 평가 방법의 결과가 인간의 주관적 화질 평가 결과와 유사하다는 것을 의미하며, 이는 제안한 알고리즘이 유효한 평가 결과를 낸다는 것을 나타낸다.

표 2는 제안한 EPQM의 성능을 평가하기 위해, 기존의 방법인 SSIM, EPSNR, 그리고 EHD 방법을 이용하였다. 제안한 방법과 기존 방법들의 Pearson 상관 계수와 SAE, 그리고 화질 평가 시 필요한 데이터량의 세 가지 측면에서 성능을 비교하였다. 우선, 제안한 EPQM과 SSIM을 비교하면, Pearson 상관 계수는 0.236 (38.56%)이 증가하였으며, SAE는 5.707 (34.19%)이 감

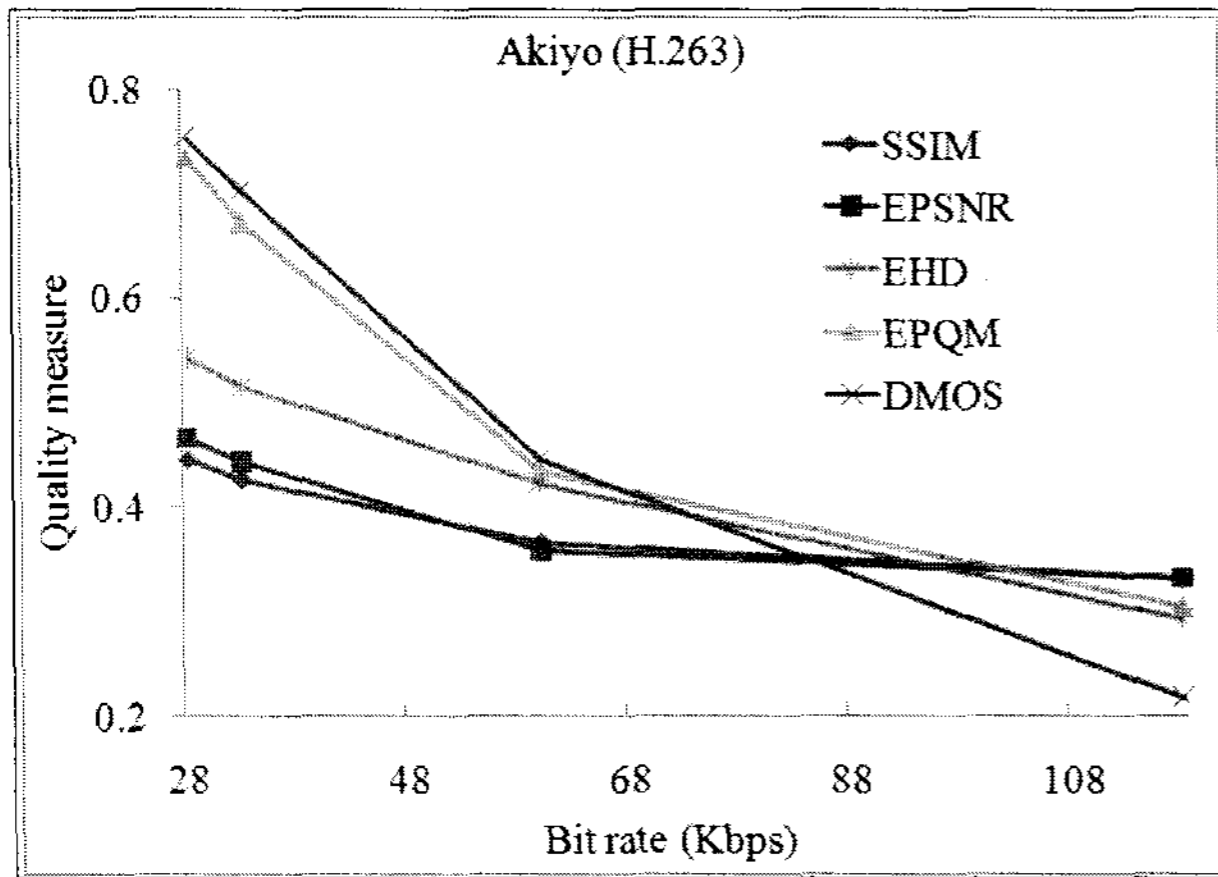
표 2. 제안한 방법과 기존 방법들 사이의 성능 비교
Table 2. Performance comparison of the proposed EPQM with conventional methods.

	SSIM	EPSNR	EHD	EPQM
Pearson 상관 계수	0.612	0.620	0.772	0.848
SAE	16.690	15.446	13.232	10.983
Data size (%)	100	100	0.632 (QCIF) 0.158 (CIF)	2.533 (QCIF) 1.265 (CIF)

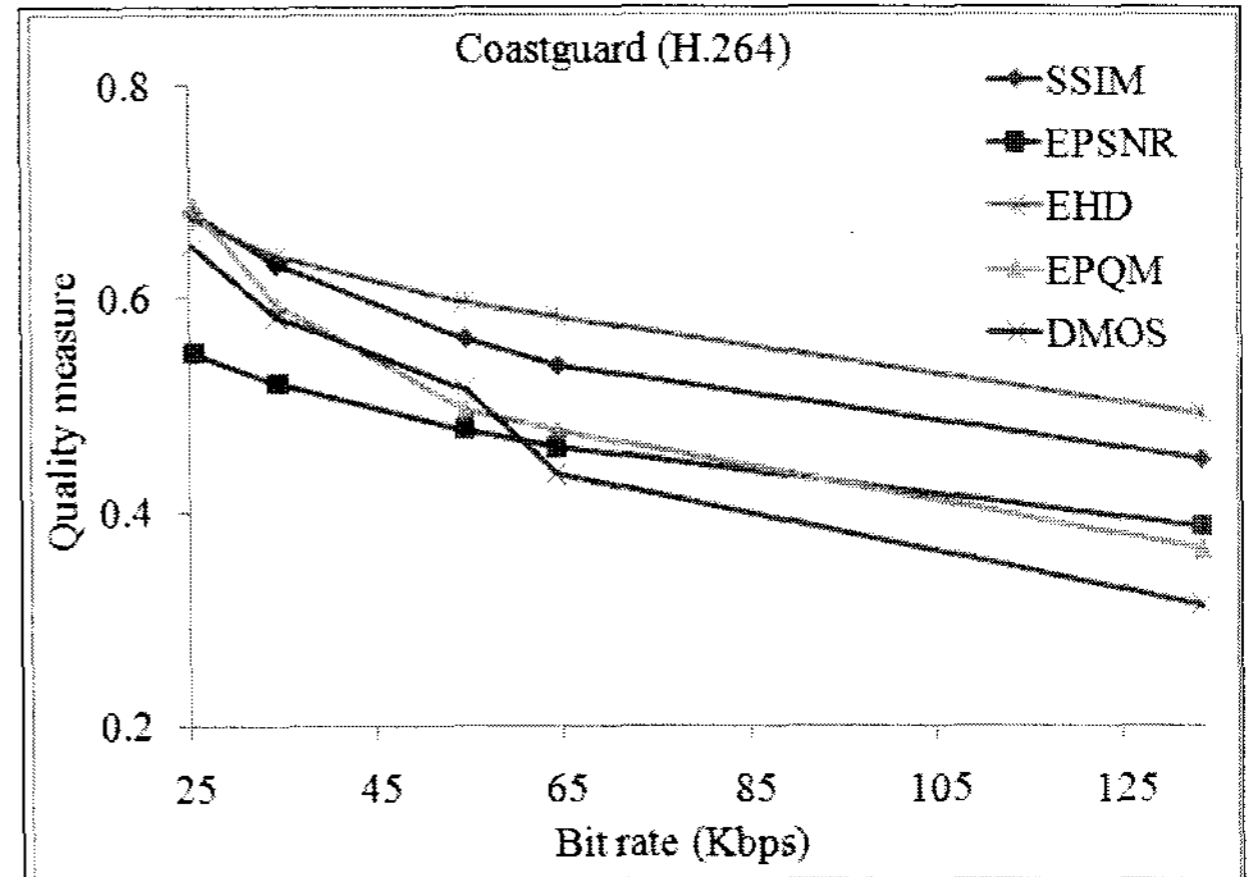
소하였다. 그리고 제안한 EPQM과 EPSNR을 비교하면, Pearson 상관 계수는 0.228 (36.77%)가 증가하였고, SAE는 4.463 (28.89%)가 감소하였다. 또한 EHD와 비교했을 때, Pearson 상관 계수는 0.076 (9.84%)가 증가하였고, SAE는 2.249 (17.00%)가 감소하였다. 데이터량 측면에서 제안한 EPQM이 기존의 FR 방법인 SSIM과 EPSNR에 비해 데이터량이 약 40배나 90배 정도 줄어들었지만, EHD 방법에 비해서는 4배나 8배의 데이터량이 필요한 것을 알 수 있다. 여기서, EHD의 경우, 메타 데이터로 영상에 부착돼 있어 추가로 전송해야 할 데이터량이 없다. 하지만 영상을 원영상 없이 영상 특징을 이용해 화질을 평가할 수 있다는 측면에서 볼 때, 영상을 표현하기 위해 필요한 데이터량을 기준으로 비교한 것이다.

그림 5~6은 영상별로 비트율이 다를 때, SSIM, EPSNR, EHD, 그리고 제안한 EPQM을 DMOS 값과 함께 도시한 그래프이다. 비트율이 낮아지면 화질이 나빠지므로, DMOS나 다른 방법들의 값들은 높아지며, 반대로 비트율이 높아지면 화질이 좋아지므로 DMOS나 다른 방법들의 값들은 낮아진다. 또한 DMOS 그래프와 유사한 형태나 근접해 있는 정도가 SSIM, EPSNR, EHD, 그리고 EPQM의 성능을 나타낸다.

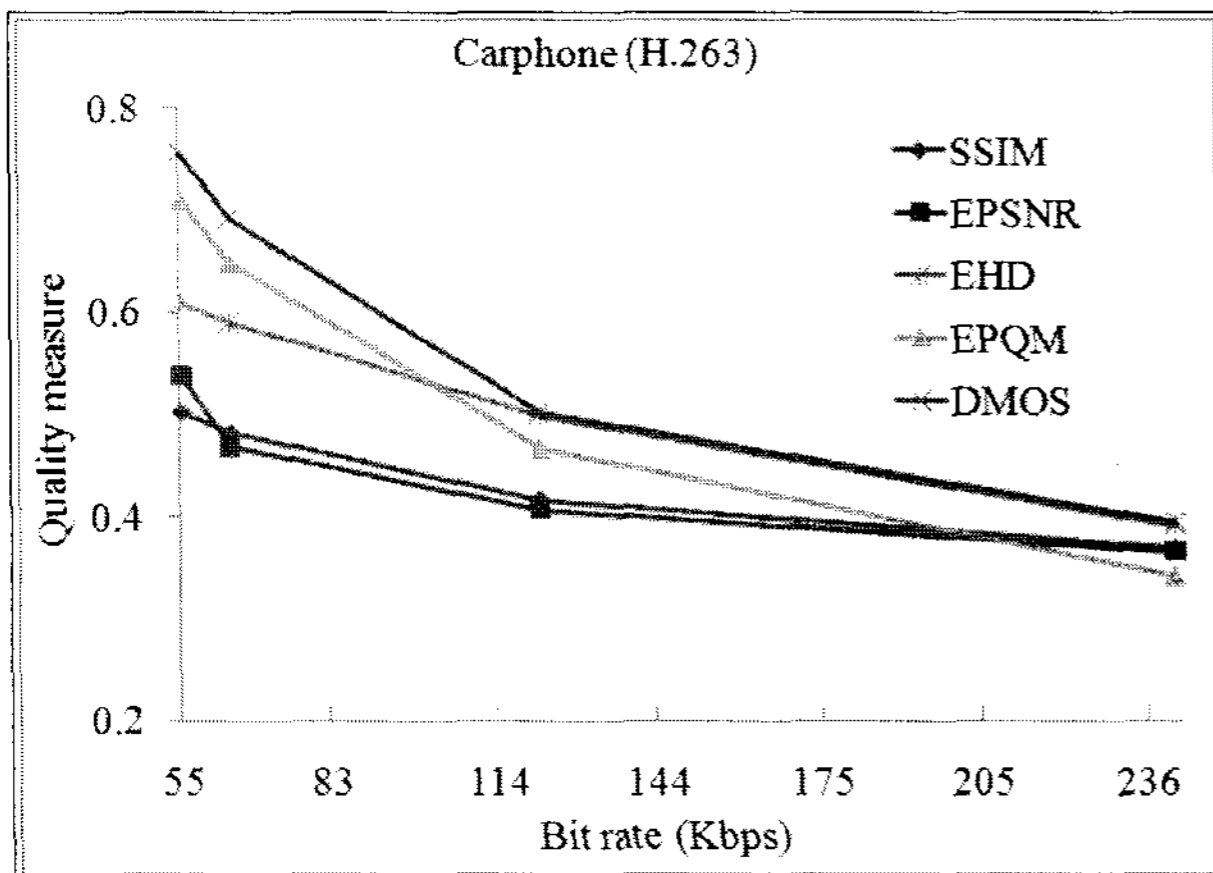
그림 5는 H.263 코덱으로 압축된 QCIF 비디오 클립에 대한 화질 평가 결과이다. 그림 5(a)는 H.263 코덱으로 압축된 QCIF Akiyo 비디오 클립에 대한 화질 평가 결과이며, 이를 통해 SSIM이나 EPSNR보다 제안한 EPQM이 DMOS에 더 근접해 있고 유사한 형태를 갖는 것을 알 수 있다. 그림 5(b)는 H.263 코덱으로 압축된 QCIF Carphone 비디오 클립에 대한 화질 평가 결과이며, 그림 5(a)처럼 다른 방법들보다 제안한 EPQM이 DMOS에 더 근접해 있는 것을 볼 수 있다. 이와 유사한 결과는 그림 6에서도 볼 수 있다. 그림 6은 H.264 코덱으로 압축된 비디오 클립에 대한 결과를 도시한 것이



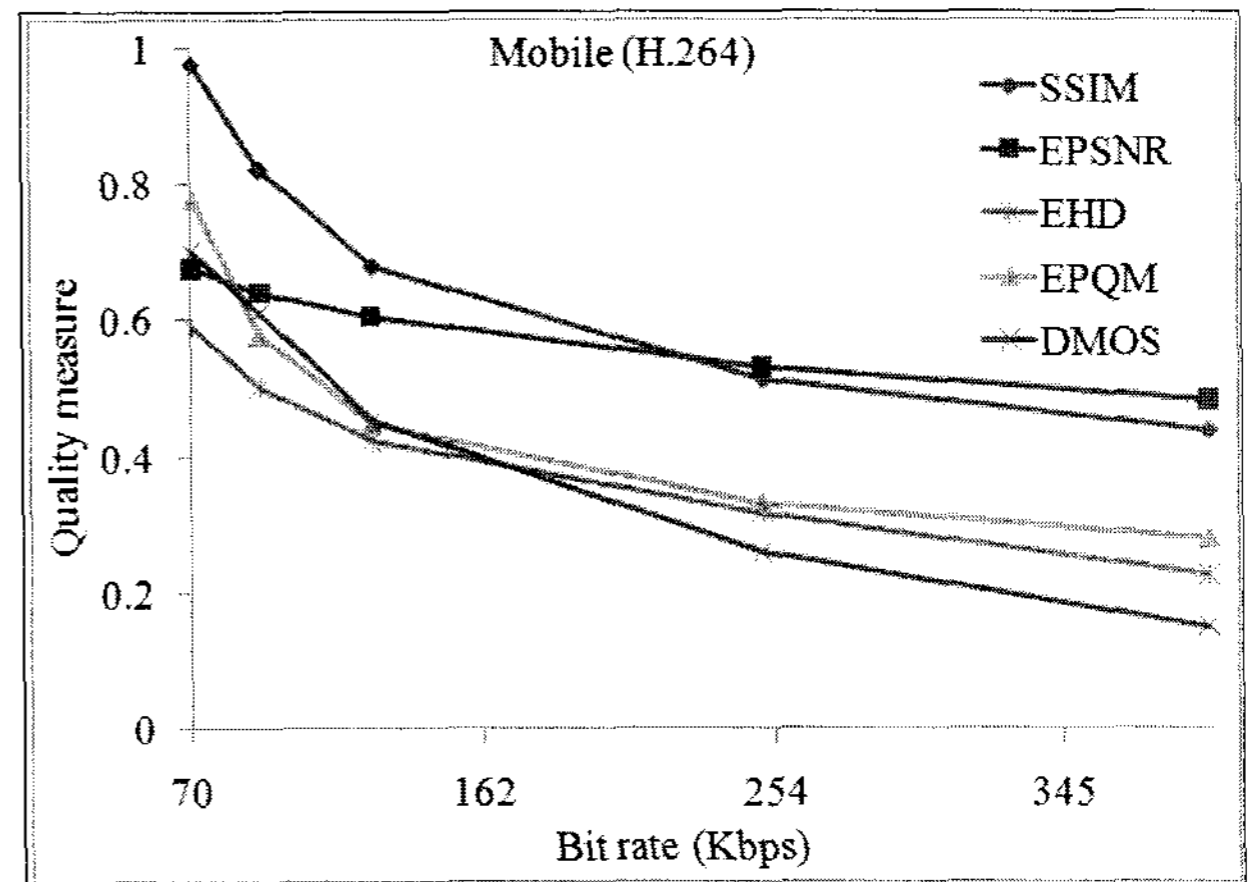
(a)



(a)



(b)



(b)

그림 5. H.263 코덱으로 압축된 QCIF 비디오 클립에 대한 화질 평가 결과
(a) Akiyo, (b) Carphone

Fig. 5. Quality evaluation results for the QCIF video clips compressed with H.263 codec.
(a) Akiyo, (b) Carphone.

그림 6. H.264 코덱으로 압축된 비디오 클립에 대한 화질 평가 결과
(a) QCIF Coastguard, (b) CIF Mobile

Fig. 6. Quality evaluation results for video clips compressed with H.264 codec.
(a) QCIF Coastguard, (b) CIF Mobile,

다. H.264 코덱으로 압축된 QCIF Coastguard 비디오 클립에 대한 결과를 도시한 그림 6(a)에서도 그림 5(a)와 5(b)에서처럼 다른 방법들보다 EPQM이 나은 결과를 보인다. 또한 그림 6(b)는 H.264 코덱으로 압축된 CIF Mobile 비디오 클립에 대한 화질 평가 결과를 도시한 것이다. 그림 6(b)에서도 제안한 EPQM이 DMOS에 기존의 방법들보다 더 근접해 있음을 알 수 있다. 그림 5와 6의 결과를 통해, 예지의 수를 모든 영상에 대해 확일적으로 정하여 사용하였지만, 영상의 종류와 내용에 관계없이 좋은 성능을 보이고 있음을 알 수 있다.

표 3은 코덱과 영상 포맷을 기준으로 영상들을 분류하여 화질 평가 방법들의 성능을 비교한 것이다. 표 3에서 코덱을 기준으로 살펴보면, 우선 H.263 코덱으로 압축된 비디오 클립에 대해 제안한 EPQM이 가장 높은

Pearson 상관 계수를 가지며, 이것은 기존의 SSIM과 EPSNR, 그리고 EHD에 비해 각각 0.281 (50.18%)와 0.240 (39.93%), 그리고 0.113 (15.52%) 증가한 수치이다. SAE 또한 제안한 방법이 SSIM과 EPSNR, 그리고 EHD보다 각각 2.145 (32.95%)와 1.672 (27.70%), 그리고 1.213 (21.75%) 감소한 수치를 갖는다. H.264 코덱으로 압축된 비디오 클립에 대한 결과도 제안한 EPQM이 훨씬 더 좋은 성능을 보이는데, Pearson 상관계수의 경우, 제안한 EPQM이 SSIM과 EPSNR, 그리고 EHD보다 각각 0.213 (33.13%)와 0.215 (33.54%), 그리고 0.054 (6.73%) 증가하였으며, SAE 관점에서 볼 때, EPQM이 SSIM과 EPSNR, 그리고 EHD보다 각각 3.562 (34.99%), 2.791 (29.66%), 그리고 1.035 (13.52%) 만큼 감소하였다.

표 3. 제안한 EPQM과 기존 방법들 간의 코덱별, 포맷별 성능 비교

Table 3. Performance comparison of the proposed EPQM with conventional methods according to the classification.

			SSIM	EPSNR	EHD	EPQM
코덱	H.263	상관 계수	0.560	0.601	0.728	0.841
		SAE	6.510	6.037	5.578	4.365
	H.264	상관 계수	0.643	0.641	0.802	0.856
		SAE	10.180	9.409	7.653	6.618
포맷	QCIF	상관 계수	0.644	0.582	0.832	0.870
		SAE	11.478	11.185	8.255	7.242
	CIF	상관 계수	0.609	0.757	0.682	0.793
		SAE	5.212	4.262	5.035	3.741

영상 포맷을 기준으로 성능을 비교해 보면, 코덱을 기준으로 평가한 결과와 유사한 결과를 재확인할 수 있다. QCIF 영상에 대해 제안한 EPQM의 Pearson 상관 계수가 SSIM과 EPSNR, 그리고 EHD보다 각각 0.226 (35.09%), 0.288 (49.48%), 그리고 0.038 (4.57%) 증가한 것을 볼 수 있다. 또한 SAE 측면에서 볼 때, EPQM은 SSIM, EPSNR, 그리고 EHD보다 각각 4.236 (36.91%), 3.943 (35.25%) 1.013 (12.27%) 만큼 감소한 것을 알 수 있다. 그리고 CIF 비디오 클립에 대한 결과에서도 제안한 방법이 Pearson 상관 계수 측면에서, SSIM과 EPSNR, 그리고 EHD보다 각각 0.184 (30.21%), 0.036 (4.76%), 그리고 0.111 (16.28%)만큼 증가하였다. 또한 SAE 측면에서는, 제안한 EPQM이 SSIM과 EPSNR, 그리고 EHD보다 각각 1.471 (28.22%), 0.521 (12.22%), 그리고 1.294 (25.70%)만큼 감소한 결과를 보였다.

이러한 결과들을 종합하여 볼 때, 제안한 방법인 EPQM이 기존의 방법들인 SSIM과 EPSNR보다 좋은 성능을 보이며, 이것은 EPQM이 기존의 방법들보다 사람의 주관적 화질을 잘 반영해 준다는 것을 의미한다.

IV. 결 론

영상 압축으로 인해 열화된 영상의 화질을 평가하기 위해 예지의 1차원 히스토그램을 특징으로 이용하는 방법을 제안하였다. 예지 투영을 이용함으로써, 원영상의 전체 정보를 이용하는 것 보다 훨씬 적은 (약 1/40배나 1/90배) 정보만으로 화질 평가를 가능하게 하는 RR 화질 평가 방법인 EPQM을 제안하였다. 실험 결과를 통

해 MOS값과의 Pearson 상관 계수, SAE, 그리고 평가를 위해 필요한 데이터량 측면에서, EPQM의 성능이 기존의 방법들 보다 우수함을 보였다. 추후로 EPQM의 차원을 줄임으로써 추가로 전송되는 데이터량 측면에서 보다 효율적인 RR 화질 평가 방법과 임계치에 영향을 받지 않는 방법에 대해 연구할 것이다.

참 고 문 헌

- [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*, vol. 13, no. 4, pp. 600-612, Apr. 2004.
- [2] G.-H. Chen, C.-L. Yang, and S.-L. Xie, "Gradient-based structural similarity for image quality assessment," in *Proc. International Conf. Image Processing*, pp. 2929-2932, Atlanta, GA, USA, Oct. 2006.
- [3] ITU-T Recommendation J.144, "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference," *International Telecommunication Union*, Mar. 2004.
- [4] 엄민영, 최윤식, 장석각, 조봉관, "Gabor 웨이블릿 기반 객관적 화질 평가," 대한전자공학회 논문지, 제41권 SP편, 제6호, 81-88쪽, 2004년 11월
- [5] 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*, vol. 5666, pp. 149-159, San Jose, CA, USA, Jan. 2005.
- [6] 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*, vol. 6391, pp. 1-8, Boston, MA, USA, Oct. 2006.
- [7] 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*, pp. 71-74, Bratislava, Slovakia, Oct. 2003.
- [8] J. Caviades and F. Oberti, "No-reference quality metric for degraded and enhanced video," in *Proc. SPIE Conf. Video Comm. Image Process.*, pp. 621-632, Lugano, Switzerland, July 2003.
- [9] E. Wahl, U. Hillenbrand, and G. Hirzinger, "Surflet-pair-relation histograms: A statistical

3D-shape representation for rapid classification,” in *Proc. Forth International Conf. 3-D Digital Imaging and Modeling 2003*, pp. 474-481, Banff, Alberat, Canada, Oct. 2003.

[10] ITU-T Recommendation BT.500-11, “Methodology for the subjective assessment of the quality of television pictures,” *International Telecommunication Union*, Jan. 2002.

[11] 이선오, 김현오, 심동규, “디지털 비디오에 대한 주관적 화질 측정과 분석,” 제19회 신호처리합동학술대회논문집, 제19권, 제1호, 215쪽, 안산, 2006년 9월

저 자 소 개



김 동 오(학생회원)
1999년 서강대학교 전자공학과
학사 졸업.
2001년 서강대학교 전자공학과
석사 졸업.
2001년~현재 서강대학교
전자공학과 박사 재학.

<주관심분야: 영상신호처리, 컴퓨터비전>



박 래 홍(평생회원)-교신저자
1976년 서울대학교 전자공학과
학사 졸업.
1979년 서울대학교 전자공학과
석사 졸업.
1981년 Stanford University
전기공학과 석사 졸업.

1984년 Stanford University 전기공학과 박사
졸업.

1984년~현재 서강대학교 전자공학과 교수.

1990년 University of Maryland (College Park,
MD, USA) 방문교수.

<주관심분야: 영상처리, 컴퓨터 비전, 패턴인식>



심 동 규(정회원)
1999년 서강대학교 전자공학과
공학박사.
1999년~2000년 (주) 현대 전자
2000년~2002년 (주) 바로 비전
2002년~2005년 University of
Washington

2005년~현재 광운대학교 컴퓨터공학과 (부교수)

<주관심분야: 영상신호처리, 컴퓨터비전>