

# 시각적 인지 특성을 고려한 H.264/AVC기반 영상의 무기준법 품질측정방법

\*김요한, \*김현태, \*박진현, \*신지태

\*성균관대학교 정보통신공학부

{dos95, hhtsog, innosoulx, jtshin}@skku.edu,

## A No-Reference Quality Measurement for H.264/AVC Video considering Human Visual System.

\*Yo-Han Kim, \*Hyuntai Kim, \*Jinhyun Park, \*Jitae Shin

\*Sungkyunkwan University

### 요약

영상기술의 발전으로 영상의 품질측정에 관한 연구가 많이 진행되어왔다. 이를 통해 새로운 영상처리, 영상통신등에 사람의 시각적 인지 특성을 고려한 품질측정이 가능해졌다. 무기준법 영상측정방법은 전달된 영상의 정보만을 이용해 품질을 측정하기 때문에, 방송이나 실시간 전송에 주로 이용되어왔다. 이 논문에서 우리는 세가지 비트스트림 정보를 이용한 새로운 무기준법 영상 품질평가 방법을 제안한다. 화면갱신비율, 움직임지수, 양자화계수의 관계를 정립하고, 회기분석을 통해 각각 요소의 가중치를 계산하여냈다. 제안하는 기법을 평가하기 위해, ITU-T P.910에 따라 250개의 영상에 대한 주관적인 영상품질을 측정하였다. 실험결과, 제안하는 화질평가 기법이 기존의 다른 방법들에 비해 주관적인 화질평가와 상관관계가 높게 나타났다.

### 1. 서론

ITU-T는 SG12 워킹그룹을 통해 사용자 체감품질 (Quality of Experience: QoE)에 대하여 사용자가 서비스를 통해 받은 주관적인 인지의 만족도라고 정의하고 있다. 서비스 제공자는 사용자 체감품질을 통해 사용자의 현재 상태를 파악하고, 이를 통해 서비스의 품질을 보장하기 위한 여러 가지 기법들을 도입하여, 사용자에게 최종적인 품질향상을 가져올 수 있다.

최근 DMB, IPTV와 같은 영상기반의 서비스들이 발전하면서, 영상에 대한 객관적인 사용자 체감품질을 측정하기 위한 여러 방법들이 제안되고 있다. 일반적으로 가장 많이 사용되는 지표로 Peak Signal to Noise Ratio (PSNR)이 있지만, PSNR은 인간의 시각적인 인지와 큰 차이를 보인다[2]. 이에 따라 Video Quality Expert Group (VQEG)에서는 사람의 눈이 느끼는 영상의 품질에 대한 객관적인 평가방법에 대해 표준화하고 있다.

객관적인 영상 품질 측정 지표들은 기준이 되는 영상의 유,무에 따라 전기준법 (Full-Reference), 부분기준법 (Reduced-Reference) 그리고 무기준법 (No-Reference)으로 나누어 진다. 전기준법은 위에서 언급한 PSNR과 같이 원본의 모든 정보를 가지고 왜곡된 영상의 품질저하를 측정하는 방법이다. 부분기준법은 원본의 일부 정보를 통해 왜곡된 영상의 품질을 측정하게 되며, 무기준법의 경우 원본의 정보 없이 왜곡된 영상의 정보만으로 품질을 측정하게 된다. 전기준법이나 부분기준법의 경우 무기준법에 비해 높은 정확도를 보여주지만, 원본의 정보가 필요하다는 점에서 사용자의 단말에 적용하기 힘든 단점이 있다.

무기준법의 경우, 두가지로 분류할 수 있는데 무기준화소법 (No-Reference Pixel)과 무기준비트스트림법 (No-Reference Bit-stream)이다. 무기준화소법은 영상이 가지는 각각의 화소를 분석하여 흐림(Blur), 블록화(Blockiness)와 같은 특성을 추정하게 되며 매우 복잡한 계산을 거치게 된다[3][4]. 이에 반해, 무기준비트스트림법

은 부호화(Encoded)된 영상의 비트스트림기반의 정보를 사용하여 비교적 간단하게 측정할 수 있다.

[5]에서 프레임률(Frame Rate)과 PSNR, 움직임정도(Motion Vector)를 통하여 품질측정을 하였고, [6]에서는 [5]를 기반으로 화면 해상도의 변화를 적용한 품질측정방법을 제안하였다. [7]에서는 신경 회로망기반의 부분기준법, 무기준법 품질측정방법을 제안하고 있다.

[5],[6]의 경우 PSNR을 필요로 하기 때문에 부분기준법과 무기준법으로만 적용이 가능하며, 원 영상을 가지는 서비스제공자에서 PSNR정보를 계산하여 사용자에게 전달해야만 품질측정이 가능하다. [7]의 경우 신경회로망 기반으로 복잡한 계산이 필요하며 무기준법의 경우 낮은 상관도로 실제 적용하기 힘들다.

본 논문에서는 사용자의 주관적 화질평가의 결과를 수식적으로 분석하여 새로운 형태의 기존의 방법에 비해 높은 정확도를 가진 무기준법 품질평가방법을 제안하고, 실험을 통하여 성능을 평가해 본다.

### 2. 주관적 화질평가

객관적인 인지적 화질평가를 수행하기 위해 기준이 되는 주관적인 화질평가가 필요하게 된다. ITU-T에서는 이를 위하여 Mean Opinion Score (MOS)라는 지표를 두고 Absolute Category Rate (ACR) 방법을 통해 MOS를 취득하게 된다[1].

ACR 방법은 10초간 영상을 관찰한 후 화질을 매우 나쁨, 나쁨, 보통, 좋음, 매우좋은 5단계로 나누며 이를 각각 1에서 5의 숫자로 표시하는 방법이다.

여러 사람의 MOS를 모아 각 영상의 평균 MOS를 구하게 되며 이 수치와 제안된 방법을 비교하여 그 상관도를 통해 제안된 방법의 정확도를 평가할 수 있다.

그림 1은 ACR 방법을 보여주고 있다.

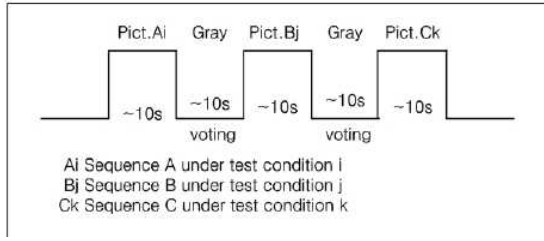


그림1. ACR 방법

MOS를 측정하기 위하여 5개의 영상을 사용하였으며, 10개의 원본영상과 250개의 서로 다른 방법으로 부호화된 영상을 사용하였다.

각각의 원본영상은 28 / 32 / 36 / 40 / 44의 양자화계수, 1.875 / 3.75 / 7.5 / 15 / 30의 프레임률을 사용하였다. 사용된 부호화기는 H.264 base profile이며 Group of Picture (GOP)의 길이는 16이다.

20명의 성균관대학교 학생이 주관적인 화질평가를 수행하였다.



그림1. 사용된 영상

부호화된 영상의 품질을 측정하기 위하여 수행한 주관적 영상평가로부터 Differential Mean Opinion Score (DMOS)를 구하기 위하여 VQEG Hybrid / Bit-stream Group의 실험계획[8]을 사용하였다.

[8]에 의하여 DMOS는

$$DMOS = 5 + MOS(PVS) - MOS(SRC) \quad (1)$$

로 나타낼 수 있다.

PVS는 부호화된 영상이며, SRC는 원본영상이다. 이를 0에서 1사이의 값으로 변환하면 VQM이 되며 아래 수식과 같다.

$$VQM = (DMOS - 1) / 4 \quad (2)$$

최종 왜곡 D는 아래와 같이 정의된다.

$$D = 1 - VQM \quad (3)$$

### 3. 제안된 방법

부호화된 영상은 부호화에 사용되는 계수들에 따라 품질의 저하가 결정된다. 일반적으로 빠른 움직임의 영상은 느린 움직임의 영상에 비해 품질의 저하가 크다. 품질의 저하는 프레임률을 낮출수록 커지고, 양자화계수가 커질수록 커진다.

본 논문에 사용된 비트스트림 정보는 움직임정도, 프레임률, 양자화계수이다. 전체적인 왜곡의 수식화전에 우선 움직임정도(MV)와 양자화계수(QP)를 정규화하였다.

정규화된 움직임정도  $MV_{NORM}$ 은

$$MV_{NORM} = \sqrt{\frac{\sum_{i=0}^M \sum_{j=0}^N \sqrt{x_{i,j}^2 + y_{i,j}^2} \times Fr}{M \times N \times \sqrt{4h^2 + 4w^2}}} \quad (4)$$

로 나타낼 수 있다.

여기서  $M$ 은 프레임률,  $N$ 은 영상의 각 매크로블럭(Macro Block),  $h$ 는 수직해상도,  $w$ 는 수평해상도이다.  $x_{i,j}$ 와  $y_{i,j}$ 는 매크로블럭안의 움직임좌표(Motion Coordinate)이다.

H.264 부호화기에서 양자화계수 QP는 0에서 51사이의 값이 정해지며, 6의 QP차이는 2배의 양자화단차(Quantization Step)을 가진다.

따라서 정규화된 양자화계수  $QP_{NORM}$ 은

$$QP_{NORM} = \log_{\sqrt{6}} QP \quad (5)$$

로 나타낼 수 있다.

그림3에  $QP_{NORM}$ 과 왜곡과의 관계를 나타내었다.

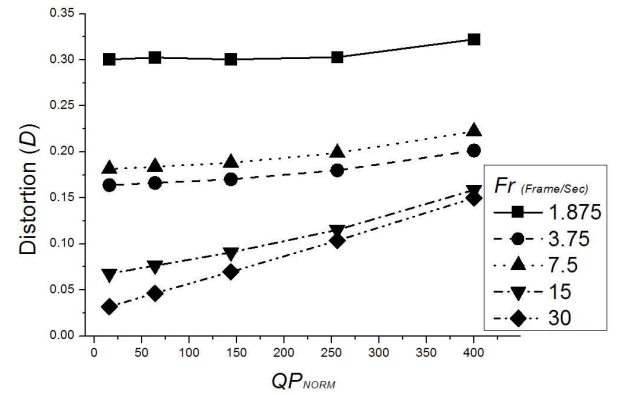


그림3. 왜곡 대 양자화계수

그림3에서 우리는 왜곡을 아래와 같이 수식화 할 수 있다.

$$D = aQP_{NORM} + b \quad (6)$$

a는 변수로하는 왜곡의 기울기, b는 절편이다.

그림 4에서 프레임률(Fr)과 움직임정도(MV)에 따른 기울기의 변화를 볼 수 있다.

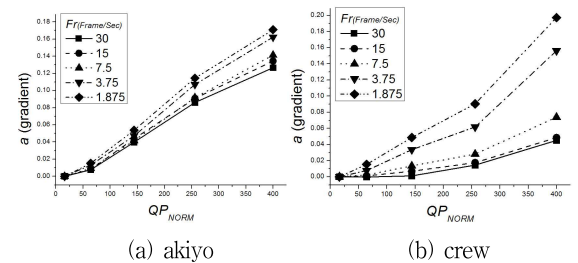


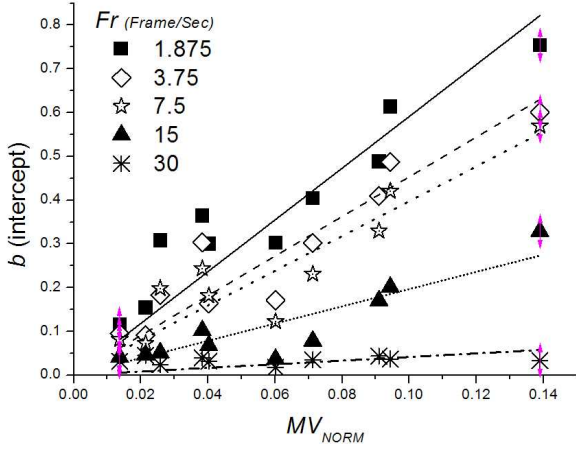
그림4. 기울기 대 양자화계수

그림4에서와 같이, 높은 프레임률에서의 기울기는 움직임정도와 관계없이 유사함을 알 수 있다. 하지만, 낮은 프레임률에서는 움직임이 클수록 기울기가 낮아진다는 것을 알 수 있다. 이를 수식으로 나타내면

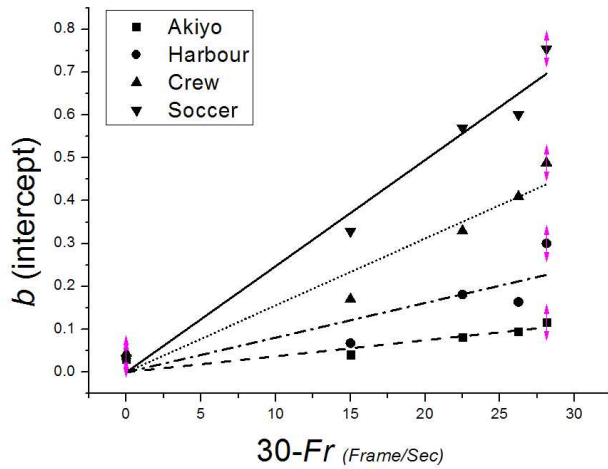
$$a = \alpha_1(\alpha_2 - \alpha_3 MV_{NORM}) \times Fr \quad (7)$$

로 표현할 수 있다.

그림5에서는  $MV_{NORM}$ 과  $30-Fr$ 에 의한 절편의 변화를 나타내었다.



(a)



(b)

그림5. (a) $MV_{NORM}$  (b) $30-Fr$ 에 따른 절편변화

그림 5를 통해서 절편  $b$ 에 대한 관계를 아래와 같이 수식화 할 수 있다.

$$b = \beta_1 MV_{NORM} \times (30 - Fr) + \beta_2 \quad (8)$$

결과적으로 최종 왜곡  $D$ 는

$$D = \alpha_1(\alpha_2 - \alpha_3 MV_{NORM}) Fr \times QP_{NORM} + \beta_1 MV_{NORM} \times (30 - Fr) + \beta_2 \quad (9)$$

로 나타낼 수 있으며 이를 풀면

$$D = \omega_1 Fr \times QP_{NORM} + \omega_2 MV_{NORM} \times Fr \times QP_{NORM} + \omega_3 MV_{NORM} \times (30 - Fr) + \omega_4 \quad (10)$$

가 된다.

측정된 주관적인 화질평가자료를 이용한 최소제곱법 기반의 회기 분석을 통해서 우리는 각각의 계수들을 계산하였다. 그 결과 각각의 계수값은  $w_1=0.0314$ ,  $w_2=-0.3575$ ,  $w_3=0.0737$ ,  $w_4=0$ 으로 산출되었다.

#### 4. 실험방법 및 결과

제안된 화질평가방법의 성능을 평가하기 위해 그림6의 3개의 영상을 선정하였다. 아래 3가지 영상은 움직임 정도의 차이가 큰 대표적인 영상이며, 회기분석을 위해 사용되지 않은 영상들이다.

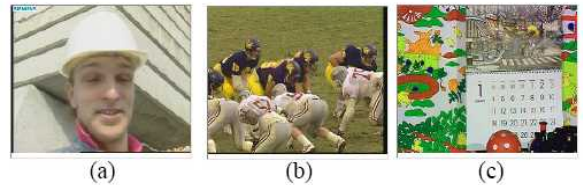


그림6. 실험영상

정확도를 계산하기 위해 피어슨 상관도(Pearson Correlation)을 사용하였다. 표1과 표2에서 제안된 방법이 기존의 방법보다 높은 상관도를 가진다는 것을 보여준다.

수행된 결과를 식(4)을 통해 계산된 각각의  $eMOS$ 값과  $D_{min}$ ,  $D_{max}$ 를 사용하지 않은 기존의 방법으로 계산된 값을 피어슨 상관도(Pearson Correlation)[8]를 통해 연관성을 계산하였다.

계산된 결과를 표1과 그림6,7에 나타내었다.

표1. 제안된 방법의 영상별 상관도

영상	상관도(8)
News	0.805
Foreman	0.950
Football	0.969
Avg	0.908

표2. 기존 방법과의 비교

영상	(7)의 무기준법	(7)의 부분기준법	제안된 방법
Football	0.85	0.94	0.969
Avg	0.86	0.93	0.908

그림 7은 제안된 방법과 주관적 화질평가자료의 분포도를 나타내었다.

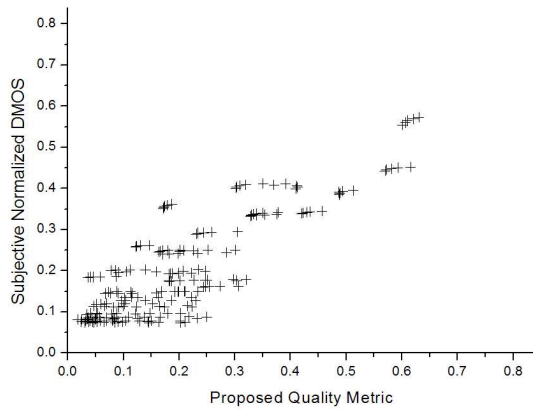


그림7. 제안된 방법과 주관적 화질의 분포도

quality metrics for end and end QoS monitoring", IEICE Trans. Commun., vol.E85-B, Feb, 2006.

[8] VQEG "Hybrid / Bitstream Group Test Plan".

## 5. 결론

본 논문에서는 비트스트림 기반의 무기준법 화질 평가방법을 제안하였다. 실험결과, 제안하는 방법이 비트스트림 정보만을 가지고 기존의 방법에 비해 높은 정확도를 보였다. 하지만 낮은 움직임정도를 가진 영상에서의 정확도가 높은 움직임정도를 가진 영상에서의 정확도에 비해 떨어지는 단점이 존재한다. 차후 연구를 통해 모든 경우의 정확도가 균일해 질 수 있는 방법을 연구하도록 하겠다.

## 감사의 글

“본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2009-(C1090-0902-0011))”

## 참고문헌

- [1] ITU-T P.910 "Subjective Video Quality Assessment Methods For Multimedia Applications"
- [2] Z. Wang, "Blind measurement of blocking artifacts in images" Proc. IEEE Int. Conf. Image Proc, 2000
- [3] P. Marziliano, F. Dufaux, S. Winkler and T. Ebrahimi, "A No-reference perceptual blur metric," Proc. ICIP, vol.3 pp.III57-III60, Jun 2002.
- [4] R. Barland and A. Saadane, "A reference free quality metric for compressed images," Proc. VPQM2006, Jan, 2006.
- [5] R. Feghali, D. Wang, F. Speranza and A. Vincent, "Quality metric for video sequences with temporal scalability", Proc. ICIP, vol.3 pp11-14, Sep, 2005
- [6] C. S. Kim, S.H. Kim, D. J. Seo, and Y. M. Ro, "Measuring video quality on full scalability of H.264/AVC scalable video coding", IEICE Trans. Commun., vol.E91-B, May, 2008.
- [7] P. LE. CALLET, C. V. GAUDIN, S. PECHARD, and E. CAILLAULT, "No reference and reduced reference video