

인간 시각 특성과 영상통신 파라미터를 이용한 동영상 품질 메트릭 개발

Development of an Perceptual Video Quality Assessment Metric Using HVS and Video Communication Parameters

이원균*, 장성환, 박희철, 이주용, 서창렬, 김정준

(Won-Kyun Lee*, Seong-Hwan Jang, Heui-Cheol Park, Ju-Yong Lee, Chang-Ryul Suh, Jung-Joon Kim)

Abstract : In this paper, we solved the underestimation problem of PSNR, which is caused by repeated frames, by easily synchronizing original and decoded frames using the proposed marks. Also we propose full-reference system which can be applied for measuring the quality of various kinds of video communication systems, e.g. wireless handsets, mobile phones and applications for PC. In addition, we propose a new video quality assessment metric using video communication parameters, i.e. frame rate and delay. According to the experiments, the proposed metric is not only appropriate for real-time video communication systems but also shows better correlation with the subjective video quality assessment than PSNR. The proposed measuring system and metric can be effectively used for measuring and standardizing the video quality of future communications.

Keywords: video quality assessment, quality metric, video communication parameters

I. 서론

최근 영상통화 서비스가 대중화 되는 등 멀티미디어 서비스가 다양하게 제공되고 사용자의 영상 품질에 대한 기대치가 높아짐에 따라 사용자의 시각을 반영하는 동영상 품질 측정 기술에 대한 요구가 증가하고 있다. 이에 따라 영상 서비스의 품질 측정과 기준 정립을 위한 영상 품질 평가 방법에 대한 연구가 활발하게 이루어지고 있다.

동영상 품질 평가는 크게 주관적 평가와 객관적 평가의 두 가지 방법으로 측정된다. 주관적 평가 방법은 다수의 평가자들에게 동영상을 보여주고 평가자들이 직접 영상 품질을 평가하는 방식으로, 사람이 인지하는 영상 품질을 총체적으로 평가할 수 있는 장점이 있지만 시간과 비용, 결과의 항시성 측면에서 효율적이지 못하다. 특히 실시간 품질 측정이 요구되는 영상 통신 시스템에서는 더욱 수행하기 어렵다.

객관적 평가 방법은 원본 영상과 복원 영상의 차이를 비교하여 수학적으로 계산하는 방법이며, MSE(Mean Square Error)와 PSNR(Peak Signal to Noise Ratio)이 가장 많이 사용된다.[2] 이 방법은 픽셀단위로 프레임간의 차이를 구하는 것이기 때문에 쉽고 빠르게 계산될 수 있지만, 인간시각특성(HVS: Human Visual System)을 정확히 반영하여 측정하지 못하는 단점이 있다. 예를 들면 노이즈가 생기는 위치에 따라 주관적 품질이 달라지는 경우이다. 이는 HVS 특성상 이미지가 복잡하게 이루어진 영역의 왜곡보다는 공간적 변화가 매끄러운 구조를 갖고 있는 배경과 같은 부분의 왜곡에서 더 민감하게 인지되기 때문이다. 따라서 같은 양의 노이즈를 추가하여 PSNR 값이 동일하다 하더라도 공간적 변화가 매끄러운 배경과 같은 부분에 노이즈가 생기면 복잡한 텍스처 부분에 생기는 것보다 인간의 눈에는 상대적 품질 저하가 심하게 느껴진다. 이와 같은 이유로 PSNR은 HVS를 정확히 반영하지 못한다.[2]

현재 국제전기통신연합(ITU: International Telecommunication Union)과 VQEG(Video Quality Expert Group)를 중심으로 HVS를 반영할 수 있는 객관적 품질 측정 방법에 대한 표준화가 진행 중이다. 객관적 화질평가 방법이 HVS와 부합됨을 검증하기 위해서 VQEG에서는 ACR(Absolute Category Rating)이나 DSCQS(Double Stimulus Continuous Quality Scale)방법 등을 사용하고 있다.[4]

본 논문에서는 프레임율과 지연시간 등의 영상통신 파라미터를 이용하여 HVS를 반영하고 실시간 영상 통신 시스템의 품질 측정에 보다 적합한 객관적 품질 측정 메트릭을 제안한다. 또한 유지 프레임으로 인한 PSNR 측정의 저하를 본 논문에서 제안하는 마크를 이용하여 해결하는 방법과, 유/무선 단말기간의 화상통화뿐만 아니라 메신저 등의 PC용 화상대화 서비스까지 다양한 영상통화 품질 측정에 활용 가능한 full-reference형 영상 품질 측정 기술을 제안한다.

II. 새로운 영상 품질 측정법

1. 테스트 시퀀스 생성

원본 영상과 복원 영상을 비교하여 영상 품질을 측정하는 full reference형 측정 기술을 위해서는 원본 영상을 테스트에 최적화 시켜야 한다. 지연시간과 프레임율, PSNR 등을 측정하기 위해서는 원본영상과 복원영상의 동기화가 이루어져야 하므로 본 논문에서 제안하는 측정 방법에서는 원본 영상의 모든 프레임마다 프레임 번호를 의미하는 마크를 표시한다. 그러나 기존 보편화 된 아라비아 숫자를 사용하게 되면, 복원 영상이 캡처될 때 매 프레임을 정확하게 샘플링할 수 없으므로 프레임과 프레임의 사이에 숫자가 변하는 중간 값이 캡처되어 잔상이 생길 수 밖에 없다. 이로 인하여 그림 1과 같이 프레임 번호를 식별하기 어려운 문제가 발생하므로 본 논문에서는 숫자를 대신하여 그림 2에서와 같이 제안된 새로운 부호로 표시한다. 제안된 새로운 부호를 사용하면 숫자와 숫자가 변하는 도중이 캡처되어 잔상이 생겨도 전후 숫자를 예측할 수 있기 때문에 식별 문제를 효율적으로 해결할 수 있다.

* 책임저자(Corresponding Author)

논문접수 : 2007. 7. 21., 채택확정 : 2007. 7. 30.

이원균 : 성균관대학교 정보통신공학부, KT 인프라연구소
(wonnylee82@gmail.com)

장성환, 박희철, 이주용, 서창렬, 김정준 : KT 인프라연구소

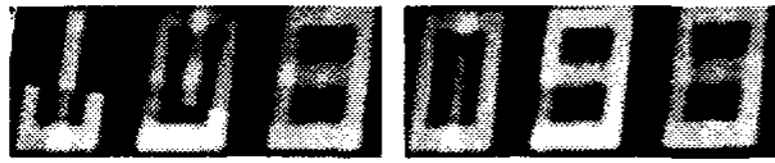


그림 1. 아라비아 숫자를 사용하면 발생하는 잔상 문제
Fig. 1. Afterimage problems in case of Arabic numerals

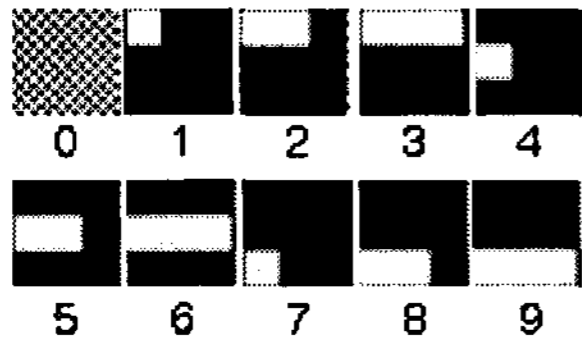


그림 2. 본 논문에서 제안하는 잔상문제를 해결한 마크
Fig. 2. Proposed marks without afterimage problems.

2. Full reference형 영상 품질 측정 모델

그림 3은 본 논문에서 제안하는 full reference형 영상 품질 측정 모델이다. 설계한 마크를 씌운 원본 영상을 skype, nateon 등의 메신저와 휴대폰, 비디오폰 등의 유/무선 영상 단말기의 다양한 인코더에 입력하며(①) 네트워크를 통해 디코더로 송신한다.(②) 디코더는 이 영상을 복원하게 되고, 이 영상을 비디오 카메라로 캡처하여 프레임별로 데이터 획득 장치에 저장하여 영상 품질 측정장치에 보낸다.(③) 이와 동시에 인코더가 현재 처리하고 있는 원본 영상 프레임을 영상 품질 측정장치에 보낸다.(④) 영상 품질 측정장치는 원 영상과 복원된 영상의 동기를 맞추어 화질을 비교할 수 있고, 미리 원본 영상의 프레임마다 표시한 마크를 이용하여 지연시간과 프레임율의 측정이 가능하다.(⑤)

3. 지연시간 측정

지연시간은 코덱에 의하여 발생하는 인코딩/디코딩 지연시간과 네트워크에 의하여 생기는 지연으로 나눌 수 있다.[7] 지연시간 측정은 인코더와 디코더에서 같은 시간에 처리하고 있는 프레임들간의 프레임 번호를 비교하여 계산한다.(그림 3의 ⑤) 예를 들어 그림 4에서와 같이 디코더가 n번째 프레임을 처리하고 있을 때 인코더는 n+5번째 프레임을 처리하고 있다면, 지연시간은 5프레임만큼 나게 되고, 30fps 영상이었다면 이 값을 30으로 나누어 지연시간을 구한다.(식 (1)) 원본 프레임과 복원 프레임의 프레임 번호는 그림 1과 같이 미리 설계한 마크를 확인하여 구할 수 있다.

$$Delay = (N_O - N_R) / FR \tag{1}$$

N_O : 원본 영상 프레임 번호
 N_R : 복원 영상 프레임 번호
 FR : 원본 영상의 프레임율

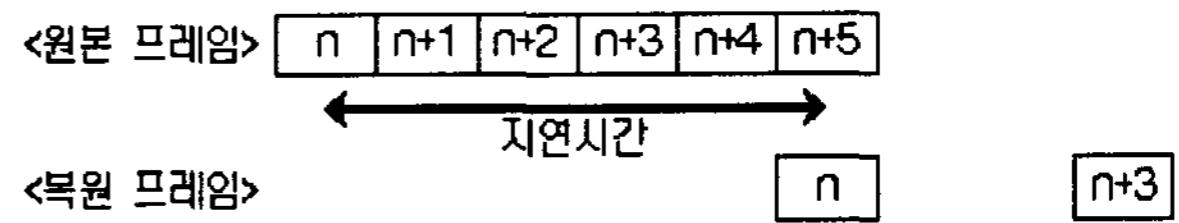


그림 4. 영상의 지연시간 측정
Fig. 4. Measurement of video delay

그림 5는 마크를 확인하면 알 수 있듯이 126번째 인코딩하고 있는 원본 프레임과 119번째 디코딩 된 복원 프레임을 보여준다. 두 개의 프레임을 비교하면 프레임 간의 차이는 7프레임이 되고 영상이 30 fps이므로 지연시간은 233ms 이다.

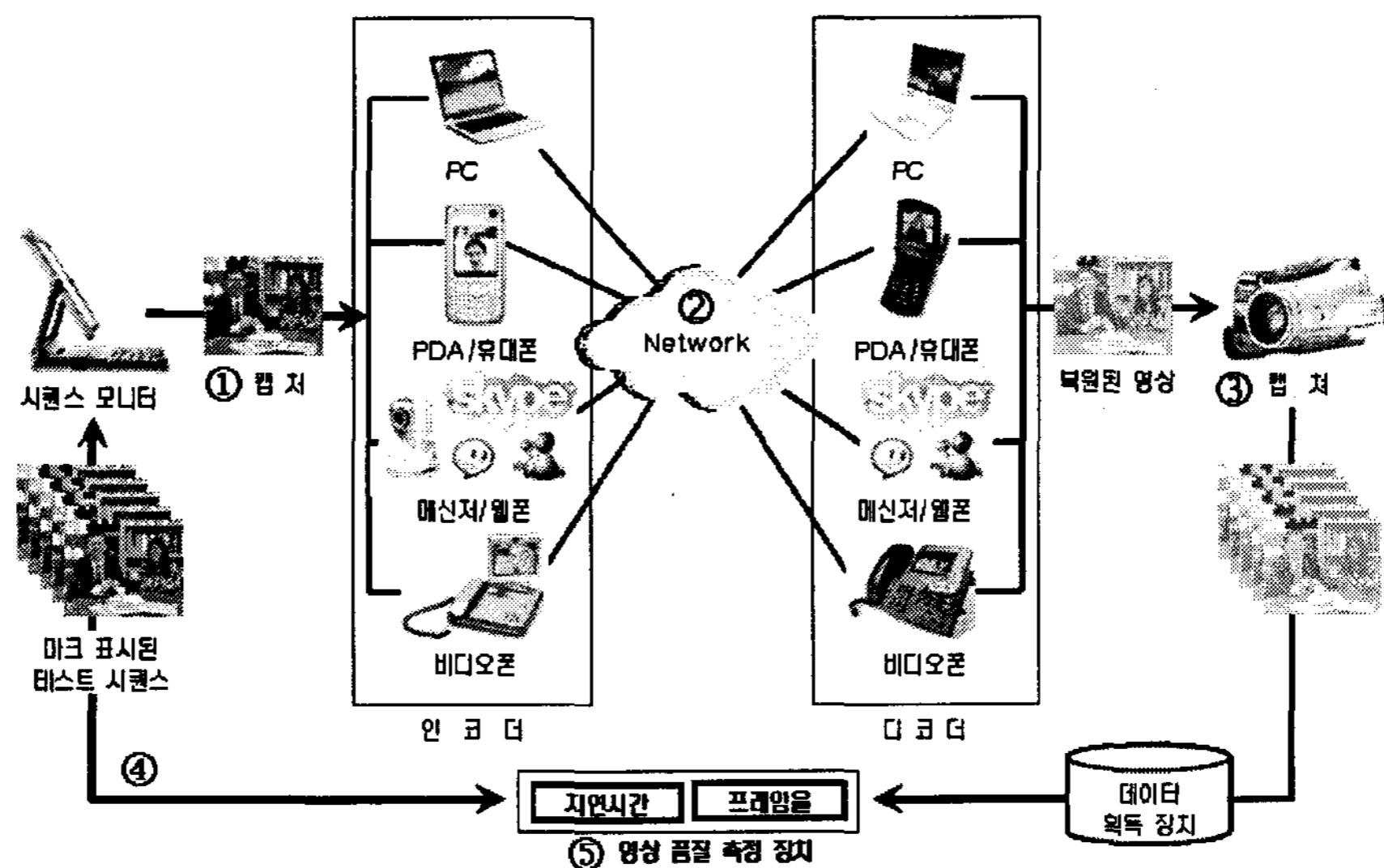


그림 3. 본 논문에서 제안하는 full reference형 영상 품질 측정 모델
Fig. 3. Proposed full-reference video quality assessment system



그림 5. 126번째 원본 프레임(좌)과 119번째 복원 프레임(우)의 이미지 캡처

Fig. 5. The captured images of 126th original frame(left) and 119th decoded frame(right)

4. 프레임을 측정

프레임율은 1초에 디스플레이되는 프레임의 개수를 나타내며 같은 품질의 프레임이라면 프레임율이 높을수록 움직임이 자연스러워지므로 사용자가 인지하는 영상 품질은 높아진다.[3]

프레임을 측정은 데이터 획득 장치에 저장된 30장의 프레임 중 인접한 프레임간의 차이가 발생하는 개수를 세어서 구할 수 있다.

III. 영상통신 파라미터를 이용한 새로운 품질 측정 메트릭

1. 새로운 메트릭의 필요성

객관적 품질 평가 방법 중 가장 널리 쓰이고 있는 PSNR 값과 주관적 평가 방법의 결과와 비교했을 때 고정된 프레임율에서 서로의 결과값의 관계가 상당히 선형적으로 나타난다.(상관도: 0.75) 하지만 시간적 요소를 추가적으로 고려하여 움직임의 부드러움 항목에 중점을 둔 주관적 방법과 PSNR을 비교하였을 때의 결과는 서로간의 상관도가 상당히 낮게 나왔다.(상관도: 0.14)[3] 이는 영상 품질 평가에서는 공간적 요소와 시간적 요소를 모두 고려하여 측정하여야 하나, PSNR은 각각의 프레임간 사이의 화질만을 평가하여 공간적 요소만을 고려하고 지연시간이나 프레임율 등의 시간적 요소는 고려하지 않기 때문에, 동작의 자연스러움과 지연도를 고려한 주관적 평가와의 상관성이 떨어지게 된다.

따라서 영상 자체의 화질만을 평가하는 PSNR 자체로는 HVS에 기반한 영상 품질 평가가 불가능 하기 때문에 공간과 시간적 요소를 모두 고려한 새로운 영상 품질 메트릭이 필요하다. 본 논문에서 제안하는 새로운 영상 품질 측정 메트릭은 공간적 요소인 PSNR과 시간적 요소인 지연시간 및 프레임율을 모두 고려한다. 복원된 영상의 특성을 분석하여 메트릭에 적용하므로 실시간 영상 품질 측정에 적합할 것이고, 나아가 주관적 품질 평가로 인하여 발생할 수 있는 추가적인 인력과 비용 소요도 줄일 수 있다.

2. 원본 프레임과 복원 프레임의 동기화 방법

본 논문에서는 원본영상에 표시하는 마크를 이용하여 프레임간의 동기를 간단하고 효과적으로 찾는 방법을 제안한다. 그림 4에서 복원 프레임 n이 원본 프레임 n과 동기임을 찾아 내려면 복원 프레임 n과 지연시간내의 구간인 원본 프레임 n~n+5번째 프레임의 마크 부분에 해당하는 MSE값을 각각 비교해야 한다. 가장 작은 MSE 값을 갖는 프레임들이 서로 동

기이다. 제안한 마크를 이용하면 CIF(352x288)의 테스트 시퀀스 중 96x32(32x32의 한자리) 크기의 MSE값을 비교하기 때문에 계산 영역을 크게 줄일 수 있다.

$$MSE(a) = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (OF_{k-a}(i,j) - DF_k(i,j))^2 \quad (2)$$

a: 동기(0 ≤ a ≤ 지연된 프레임수)
 m,n: 제안한마크사이즈(96 × 32)
 OF: 원본프레임(복원프레임과 비교하여 a만큼 차이남)
 DF: 복원프레임

원본 프레임과 복원 프레임간의 MSE 측정법을 식 (2)에서 나타내었다. 복원 프레임과 원본 프레임을 a만큼 이전에 인코딩된 프레임과 비교하여 측정된 MSE 값 중 최소가 나올 때의 a값이 동기가 된다. a는 식(1)에서 측정된 지연시간 보다 크게 나올 수 없으므로, 지연 프레임이 없는 0에서 지연 프레임까지의 값으로 제한된다.

이 방법을 사용하면 테스트 시퀀스의 길이가 제한적이지 않고 연속적이라 하더라도 언제든지 빠르고 간단하게 프레임간의 동기화가 가능하기 때문에 품질 측정의 효율성을 높여 주고 실시간 영상 품질 측정에 적합하다.

3. 유지 프레임으로 인한 PSNR 문제 해결

영상 통신에서는 네트워크 상황이나 코덱 성능 등으로 인해 인코더로 입력되는 프레임 모두를 인코더가 처리해서 전송할 수 없기 때문에 프레임 스킵이 일어날 수 밖에 없다. 대부분의 디코더는 프레임 스킵이 일어날 때 다음 프레임이 복원되어 수신될 때까지 바로 전(前)의 프레임을 유지하여 반복적으로 보여준다.(그림 6) 하지만, PSNR 계산시 이렇게 유지되고 있는 프레임들과 원본 프레임을 비교하게 되면 실제 프레임간의 동기화가 이루어지지 않았기 때문에 PSNR 수치는 현저히 떨어질 수 밖에 없다.[1] 예를 들어 n ~ n+5까지의 원본 영상이 인코딩 됐을 때 복원되는 프레임은 n과 n+4의 2개의 프레임이라면, 3개의 프레임 구간 동안 바로 전(前)의 프레임인 n이 유지될 것이고, 이 구간 동안에는 동기화가 이루어지지 않았기 때문에 원본 프레임과 유지 프레임간의 차이가 크게 나므로 PSNR 수치는 감소하게 된다. 만약 유지 프레임 구간 동안 장면의 전환이 일어나게 되면 원본 프레임과 복원 프레임간의 차이가 더욱 커지게 되어 측정 결과의 신빙성이 더욱 떨어질 것이다.

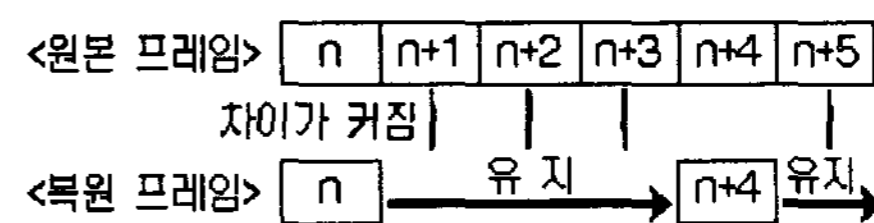


그림 6. 유지된 프레임으로 인한 PSNR 저하

Fig. 6 An underestimation of PSNR due to repeated frames

이러한 문제를 해결하기 위해서 본 논문에서는 PSNR을 복원 프레임을 기준으로 원본 프레임과 비교하여 동기를 찾아 계산하는 방법을 제안한다. 그림 7에서 볼 수 있듯이 n~n+3번째까지의 프레임과 유지 프레임은 원본 프레임 n과 동기이므로 서로 동기화를 시켜준다. 이렇게 동기화한 후 PSNR을 측정하면 유지프레임으로 인한 PSNR저하를 방지할 수 있다.

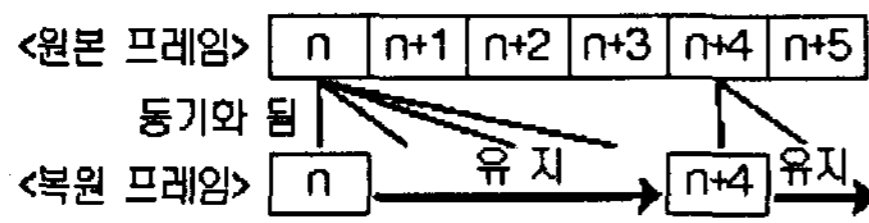


그림 7. 동기화로 유지 프레임으로 인한 문제 해결
 Fig. 7. Solving the problem due to repeated frames by synchronizing original and decoded frames

4. 새로운 품질 메트릭

HVS에 근접한 객관적 품질 측정 메트릭을 위해서는 공간적 요소인 PSNR 뿐 아니라 시간적 요소인 프레임율과 지연 시간을 총체적으로 고려해야 한다.

유지되는 프레임으로 인한 문제점을 해결하기 위해 본 논문에서 제안한 PSNR 측정법으로 PSNR을 측정하게 되면 복원 프레임을 기준으로 동기화가 되어 비교하기 때문에 기존 PSNR 측정법에 비해 수치가 올라간다. 이렇게 측정된 PSNR을 영상통신 파라미터인 프레임율과 지연시간으로 보정시키면 HVS에 적합한 인지적 품질에 가까워 지게 된다. 보정폭은 프레임율과 지연시간 값에 의해 달라지는데 여러가지 계수를 적용하여 메트릭을 연구한 결과, 다음과 같은 새로운 메트릭이 가장 좋은 결과를 보여주었다.

$$PSNR.DF = PSNR - 0.21(30 - FR) - 0.84(DE - 0.15)^{1.38} \quad (3)$$

FR: 프레임율 (fps)
 DE: 지연시간 (s), 최소값 = 0.15

새로운 메트릭인 PSNR.DF에서 FR이 30 fps(full frame)이라면 유지프레임이 없어지기 때문에 논문에서 제안한 PSNR과 기존 PSNR값이 같아진다. 또한, 지연시간의 최소값을 0.15초로 제한하였으므로 150ms 이하이면 유지프레임 문제를 해결한 PSNR과 FR값에 따라 품질 측정값이 결정된다. 프레임율이 떨어지게 되면 유지프레임이 생기는데 앞서 제안한 방법으로 기존보다 높아진 PSNR값을 FR과 DE로 상쇄시키면 HVS에 더 근접한 메트릭이 완성된다.

IV. 실험 및 결과 분석

주관적 화질 평가 방법은 HVS를 좀 더 정확히 반영할 수 있기 때문에 객관적 화질 평가 방법을 검증하는데 사용되며, 객관적 방법이 주관적 방법과 연관성이 커서 선형적인 관계를 보일수록 객관적 화질 평가모델의 우수성이 높다고 볼 수 있다.[1]

테스트 시퀀스인 원본 영상은 본 논문에서는 30 fps(full frame)을 기준으로 'paris', 'football', 'susie' 영상을 사용한다. 주관적 평가 방법은 [1]에서와 같이 DSCQS를 사용하는 총체적인 평가 이외에 동작의 자연스러움과 지연도를 고려한 주관적 평가를 추가적으로 실시하였다. 이렇게 실시한 주관적 평가 결과와 기존 PSNR 측정치, 유지된 프레임 문제를 해결한 PSNR를 이용한 새로운 메트릭(PSNR.DF)의 결과와 각각 상관도를 비교하였다. 영상 통신 환경은 유선 단말기와 PC용 어플리케이션을 이용하였다.

실험결과는 표 1에서 볼 수 있듯이 PSNR.DF로 측정된 각

각의 영상의 품질 결과와 주관적 영상품질 측정결과와의 상관관계수가 모두 0.73 이상이었다. 기존 PSNR 측정치의 상관관계수와 비교해도 모두 0.10 이상 향상됐음을 확인할 수 있다. 따라서 PSNR.DF는 기존 PSNR보다 사용자가 인지하는 품질 결과에 더 근접하였다고 할 수 있다.

표 1. 주관적 화질 평가와의 상관계수

Table 1. Correlation coefficients with the assessed subjective quality

	PSNR	PSNR.DF
Paris	0.64	0.76
Football	0.55	0.73
Susie	0.69	0.80

V. 결론

본 논문에서는 마크를 이용하여 복원 프레임을 기준으로 PSNR을 측정하는 방법과 공간과 시간적 요소를 모두 고려한 새로운 품질 측정 메트릭인 PSNR.DF를 제안하였다. 기존 PSNR은 영상의 화질 자체만을 고려하기 때문에 HVS에 기반한 주관적 화질 측정 결과와 동떨어진 면이 있었고, 프레임율이 떨어지게 되면 발생하는 유지프레임으로 인하여 PSNR수치가 낮게 평가되는 문제점이 있었다. 제안한 메트릭인 PSNR.DF는 프레임율과 지연시간의 영상통신 파라미터를 PSNR과 접목하였기 때문에 실시간 영상 통신 시스템의 품질 측정에 적합하고, 실험 결과 기존 PSNR보다 인지적 품질과의 상관도가 높아졌다.

또한 제안한 마크를 사용하여 유선 단말기 뿐 아니라 무선 및 메신저 등의 PC용 화상대화 서비스까지 다양한 영상통화 품질 측정에 활용 가능한 full-reference형 영상 품질 측정 기술을 제안하였다. 이 기술을 사용하면 향후 다양한 영상통화 서비스의 품질 측정 및 품질 기준 정립에 효율적으로 사용 가능하다.

참고문헌

- [1] R. Feghali, D. Wang, F. Speranza and A. Vincent, "Quality Metric for Video Sequences with Temporal Scalability," *IEEE Intl. Conf. on Image Processing*, vol. 3, pp. III- 137-40, September, 2005.
- [2] S. Winkler, "Digital Video Quality," John Wiley & Sons, Ltd, Switzerland, 2005.
- [3] G Hauske, T. Stockhammer and R. Hofmaier, "Subjective Image Quality of Low-Rate and Low-Resolution Video Sequences," *Momuc*, March, 2003.
- [4] Video Quality Experts Group, "Multimedia Group Test Plan," Ver. 1.16, 2007.
- [5] 서동준, 김천석, 배태면, 노용만, "비디오 스케일러빌리티를 고려한 영상 품질 메트릭 모델링," 한국방송공학회 학술발표대회 논문집, pp. 81~84, November, 2006.
- [6] 부소영, 이병욱, "저속 영상부호화를 위한 최적 프레임율과 공간 양자화 결정," 한국통신학회논문지, vol.29, pp.842~847, June, 2004.
- [7] S.Mohamed and G Rubino, "Real-time Video Quality Assessment in Packet Networks: A Neural Network Model," Inria Rennes, France, May, 2001.