

# 패킷비디오 네트워크에서의 패킷손실을 고려한 무기준 화질열화 예측기법

한정현, 김요한, 김현태, 신지태

성균관대학교 정보통신공학부

{kazikai, dos95, htthsog, jtshin}@skku.edu

## No-Reference Video-Quality Assessment of Packet Loss Effect over Packet Video Network

Junghyun Han, Yo-han Kim, Hyuntai Kim, Jitae Shin

Sungkyunkwan University, School of Information and Communication Engineering

### 요약

인터넷의 발전과 비디오 압축 코덱의 발전으로 인해 실시간 영상 스트리밍 (Streaming) 서비스가 가능해졌다. 하지만, 패킷 손실로 인한 비디오 화질 열화는 여전히 문제가 되고 있다. 이 때문에 부호화기에서는 에러에 강한 내성을 가지는 부호화방법으로 압축하며, 복호화기에서는 에러 은닉기법을 통하여 패킷 손실에 의한 화질열화를 최소화한다. 이런 기법들을 활용하기 위해서는 패킷손실을 보정한 정확한 화질 열화를 측정 할 수 있는 방법이 필요하다. 영상의 화질을 평가하는 방법은 어떤 정보를 기준으로 평가하느냐에 따라서 전체 기준법 (Full Reference), 부분기준법 (Reduced Reference) 그리고 무기준법 (No Reference)으로 나눌 수 있다. 이 중에서 무기준법은 수신 영상만을 가지고 측정하기 때문에, 다른 방법보다 활용범위가 넓다. 본 논문에서는 패킷 비디오 네트워크 환경에서의 발생할 수 있는 다양한 패턴의 패킷 손실에 따른 비디오의 화질 열화를 무기준법에 의해 측정하는 방법을 제안한다. 제안한 방법은 수신된 영상의 움직임 정보를 이용하여 패킷 손실에 따른 각 프레임당 초기 에러 값을 예측 한다. 실험 결과를 통하여 제안한 방법이 상당히 높은 피어슨 상관도 (Pearson Correlation)를 보여준다.

### 1. 서론

현대 사회에서 인터넷과 압축 코덱의 발전에 따른 실시간 영상 스트리밍(Streaming) 서비스의 사용비중이 높아지고 있다. 하지만 패킷 비디오 특성상 패킷 손실이 발생 하게 되고 이것은 화질 열화를 발생 하게 하여 영상의 화질은 떨어지게 된다. MSE (mean square error) 기준인 PSNR (peak signal-to-noise ratio) 평가는 원본 영상을 요구 하기 때문에 실시간 영상 스트리밍 서비스에서의 화질 평가 방법으로는 적절치 않다. 특히 ITU (International Telecommunication Union) 와 VQEG (Video Quality Expert Group)를 중심으로 멀티미디어, 특히 동영상의 객관적 품질측정을 위한 국제표준화 과정이 진행 중에 있다 [1][2].

영상의 화질을 평가하는 방법에 있어서 어떤 정보를 사용 하느냐에 따라 전체기준법 (Full Reference), 감소기준법 (Reduced Reference), 무기준법 (No Reference)으로 나눌 수 있다 [3]. 전체 기준법은 일반적으로 사용 하는 PSNR 이 있고, 원본영상의 전체 정보와 비교영상의 전체정보를 모두 가지고 화질을 평가 하는 방법이다. 감소 기준법의 경우는 원본영상의 일부분이나 가공된 정보와 수신 영상과의 비교를 통해서 영상 품질을 측정하는데 이는 전체 기준법에 비해 영상전송량의 부하가 적다는 장점이 있다. 하지만 전체 기준법과 감소 기준법은 실시간 영상품질 평가를 위한 적합한 방법이 아니다. 그 이유는 수신 영상 품질 평가를 위해 원본 데이터나 추가적인 정보가 필요 하고, 이런 과정들은 전송측면에서 부하로 작용하게 된다. 따라서 수신된 동영상만 이용 하여 품질을 측정 하는 무기준법이 필요하다.

따라서 본 논문에서는 패킷 비디오 네트워크에서 실시간 영상 스트리밍 서비스에 적용할 수 있는 무기준법 동영상 화질 평가 방법을 제안하고자 한다.

### 2. 관련연구

기존의 무기준법에 관한 연구는 다양한 분야로 진행 되고 있다. 기존의 품질 평가 방법은 패킷 손실을, 패킷 지연 등 QoS (Quality of Service) 측면에서 이루어져 왔다. 그 중 객관적으로 화질 열화를 예측 하는 기법으로는 대표적으로 MSE가 알려져 있다. MSE는 다음과 같은 식으로 나타내어 진다.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

하지만 이 MSE는 수신되어서 복호화 된  $\hat{y}_i$  신호뿐 만 아니라 원본 정보인  $y_i$  를 요구하기 때문에 무기준법으로는 적절치 않다. 원본 영상 없이 수신된 영상만을 통하여 화질을 측정하는 무기준 화질 평가 방법에는 픽셀 단 에서 화질 평가를 수행하는 NR-P (No Reference Pixel) 방법과 압축된 비트 스트림 단 에서 화질 평가를 수행하는 NR-B (No Reference Bitstream) 방법이 있다. 대표적으로 사용하는 무기준법 메트릭 으로는 블로킹(blockiness), 블러링(blurring) 등이 있고 이 요인들은 인간 시각 체계에 민감하게 반응을 하기 때문에 주로 무기준법 접근 방식에 많이 쓰이고 있다 [4][5]. 또한 다른 방법으로는 양자화 파라미터, 에러 전파 모델, 에러은닉기법 등을 고려하여 비트 스트림 단 에서 화질 평가를 수행 하는 방법 들이 있다

[6][7]. 하지만 이러한 방법들은 수식이 복잡하고 요구하는 파라미터 또한 많아 실시간 영상 스트리밍 상에서의 무기준 화질 열화 측정 기법으로는 적합하지 않다. 즉 간단한 파라미터를 사용 하는 새로운 객관적 무기준 화질 열화 예측 기법이 필요하다.

### 3. 제안된 방법

본 논문에서 패킷 손실에 따른 화질 열화에 대한 영상 품질을 측정 하는 객관적 화질 측정 방법을 제안한다. 대부분의 동영상 압축 알고리즘은 움직임 보상 예측을 기반으로 이루어져 있다. 동영상 부호화기의 경우 각 프레임 간의 상관도를 이용한 인터 프레임 예측 방법을 사용 하는데 패킷 손실이 발생 하였을 경우 다음 그림 1과 같이 에러 전파가 나타난다.

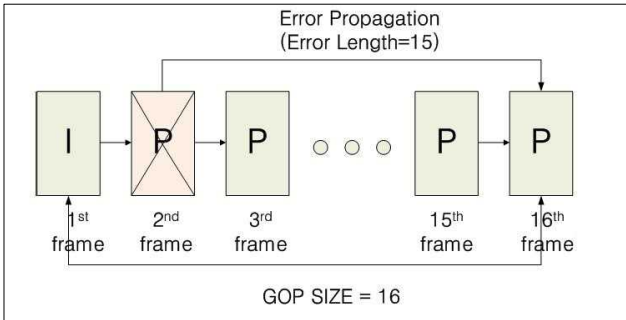


그림 1. 패킷 손실이 발생하였을 경우 나타나는 에러 전파

에러 전파에 따른 한 프레임 내에서 k번째 프레임의 MSE는 다음과 같은 식으로 나타낸다.

$$MSE_{kth} = \sum_{i=0}^{EL} (1 - \frac{i}{n}) \times MSE_{initial} \quad (2)$$

식(2)에서 EL은 패킷손실이 발생한 프레임부터 GOP의 마지막 프레임 까지의 프레임 수를 말하면 Error Length 라고 한다. 또한 n은 GOP 사이즈를 말한다.

하지만, 만약 채널 환경이 좋지 않을 때는 한 GOP 내에서도 패킷 손실이 여러 번 발생 할 수 있고 그 상황 또한 고려 해줘야 한다. 아래 식의 NL을 한 GOP 프레임 안에서 일어나는 LOSS의 개수라고 하면 식 (2)는 다음과 같이 변한다.

$$MSE_{kth} = \sum_{j=0}^{NL} \sum_{i=0}^{EL} (1 - \frac{i}{n}) \times MSE_{initial} \quad (3)$$

식 (3)은 한 GOP 안에서만 고려 할 수 있기 때문에 동영상 시퀀스를 고려한 식으로 확장이 필요하다. NG가 시퀀스 안에 있는 총 GOP의 개수라고 하면 최종적으로 구할수 있는 k번째 MSE는 다음과 같다.

$$MSE_{kth} = \sum_{k=0}^{NG} \sum_{j=0}^{NL} \sum_{i=0}^{EL} (1 - \frac{i}{n}) \times MSE_{initial} \quad (4)$$

이렇게 되면 최종적으로 패킷 손실로 인한 MSE의 값을 구할 수 있다. 하지만 여기서 인수로 사용 되는 MSE의 초기 값은 송신 단에서만 측정할 수 있기 때문에 앞의 2절에서 조사한 바와 같이 무기준법이 아닌 전체 기준법이다. 즉 수신 단에서 사용하기에는 불가능 하므로 MSE의 초기 값을 수신된 영상만으로 예측하는 방법이 필요하다.

동영상의 경우 움직임 정도를 나타내는 움직임 벡터를 각 블록의 좌표마다 가지고 있다. 이 움직임 벡터의 값이 각 프레임 마다 크다면 패킷 손실이 발생 하였을 경우 MSE의 초기 값은 증가 하게 된다. 즉 움직임 정도가 많은 영상은 패킷 손실에 더욱 취약 하게 된다. 움직임 정도를 나타내는 MA (Motion Activity)는 다음과 같이 정의 된다.

$$MA = \frac{\sum_{j=0}^n \sum_{i=0}^m \sqrt{(mv_i^2 + mv_j^2)}}{W \times H} \quad (5)$$

수신된 영상의 각 블록마다 있는 움직임 벡터를 좌표별로 구하여 합하면 식 (5)를 구할 수 있다.

(5)에서 i와 j는 프레임의 블록의 좌표이고 W와 H는 각 프레임의 높이와 너비이다. MA값과 초기 MSE의 관계를 알기 위해서 동영상 화질 평가의 실험을 위해 서로 다른 MA 정도를 가진 시험영상인 그림 2의 (a), (b), (c)이 3가지 시퀀스를 선정 하였다. (b) football의 경우 많은 움직임 정도를 가지고 있고 (c) foreman은 중간정도의 움직임 정도 (a) mobile의 경우 적은 움직임 정도를 가지고 있기 때문에 MA값과 초기 MSE의 관계식을 유도하기에 적합하다.

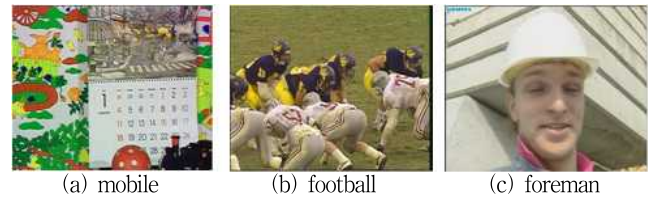


그림 2. MA와 초기 MSE 값의 관계유도에 사용한 시퀀스들

세 가지 시퀀스들을 이용해서 구해진 MA값과 초기 MSE 값을 그래프로 표현하면 다음과 같다.

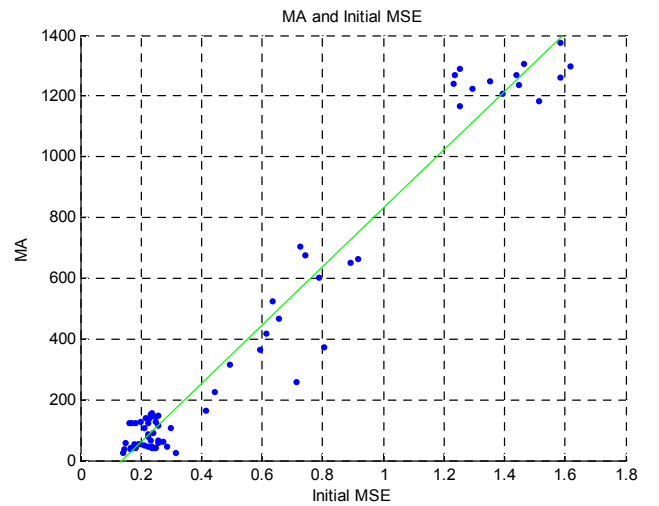


그림 3. MA 값과 MSE 초기값과의 관계

MA값과 초기 MSE값은 선형적인 특성을 가진 다음과 같은 관계식으로 나타낼 수 있다.

$$MSE_{initial} = \alpha \times MA + \beta \quad (6)$$

recursive analysis를 사용하여 식 (6)의 선형적 파라미터  $\alpha$ 와  $\beta$ 의 값을 다음과 같이 구할 수 있다.

표1. 예측한 MSE 초기값의 파라미터 값

파라미터	값
$\alpha$	963.02
$\beta$	-131.43

표 1을 이용하여 MSE 초기 값은 다음과 같이 근사 된다.

$$MSE_{initial} \approx 0.963.02 \times MA - 131.43 \quad (7)$$

최종적으로 우리가 예측 하고자 하는 우리의 예측 MSE를 식 (4),(6)을 통해서 구할 수 있다. 예측된 MSE값은 수신된 동영상의 움직임 정보와 GOP 사이즈 그리고 패킷 손실에 대한 정보만으로 표현 된다.

$$MSE_{est} = \sum_{k=0}^{NG} \sum_{j=0}^{NL} \sum_{i=0}^{EL} (1 - \frac{i}{n}) \times (963.02 \times MA - 131.43) \quad (8)$$

#### 4. 실험방법 및 결과

본 논문에서는 제안한 새로운 무기준법 알고리즘 성능을 평가하기 위해서 두 개의 테스트 시퀀스의 모션 벡터를 추출하여 비교 한다.

표2. 실험 영상 에러 패턴 세팅

GOP Size	16(IDRPPP...P)	
NG	15(GOP)	
Single Error	Pattern Number	Location of Error
	1	1 (P-frame)
	2	2 (P-frame)
	3	3 (P-frame)
	4	4 (P-frame)
	5	5 (P-frame)
	6	6 (P-frame)
	7	7 (P-frame)
	8	8 (P-frame)
	9	9 (P-frame)
	10	10(P-frame)
	11	11(P-frame)
	12	12(P-frame)
	13	13(P-frame)
Multiple Error	14	14(P-frame)
	15	1,11(P-frame)
	16	3,8(P-frame)
	17	5,6(P-frame)
	18	5,6,7(P-frame)

두 영상은 화질 평가의 실험을 위해 잘 알려진 그림 4의 (a), (b) 2개의 영상으로 선정하고 H.264/AVC 부호화기(JM 12.4 [8])를 이용해 부호화 하였다. 두 영상은 각기 다른 움직임정도와 색의 표현력을 가지고 있다. 패킷 손실을 가정하기 위해 JVT loss simulator를 사용해 총 18가지의 에러 패턴을 주고 실험을 진행 하였다 [9].



(a) Crew

(b) Soccer

그림4. 실험영상

수행된 결과를 식 (8) 을 통해 계산된 예측 값과 MSE값을 각각 상용로그를 취하여 그래프에 표현 하였다.

계산된 결과는 표 3과 그림 5, 6에 나타내었다.

표3. 실험 영상별 상관도

영상	피어슨 상관도 (Pearson Correlation)
Crew	0.953999
Soccer	0.983047

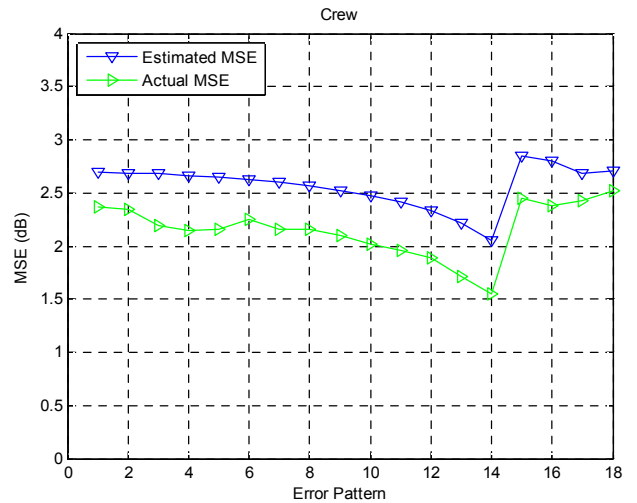


그림5. Crew 영상에 관한 MSE 값과 제안된 방법

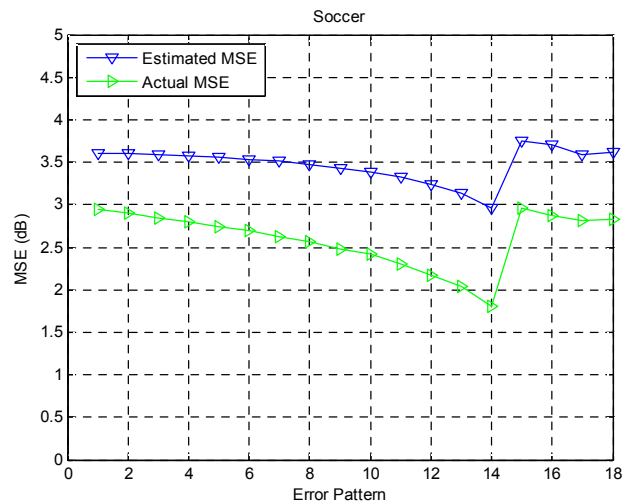


그림6. Soccer 영상에 관한 MSE 값과 제안된 방법

## 5. 결론

본 논문에서는 움직임 벡터를 사용한 새로운 무기준법을 제안하고 이 지표를 이용한 최종 평가 방법을 제시 하여 실험을 통해 성능을 평가하였다. 제안한 방법은 무기준법의 특성에 비하여 상당히 높은 피어슨 상관도(Pearson Correlation)를 보여 주었다. 하지만, 영상의 여러 지표 중 움직임 벡터만을 이용한 방법이기에 때문에 좀 더 정확한 평가를 위해서 다른 파라미터와의 상관관계 또한 앞으로 연구해야 할 것이다. 향후 연구에서는 다양한 테스트 영상을 이용해서 정확한 파라미터들을 얻은 후 패킷 손실 패턴에 따른 적응적인 무기준 화질 열화 예측 기법을 연구토록 하겠다.

## 감사의 글

“본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원 사업(NIPA-2009-C1090-0902-0011)의 연구결과로 수행되었으며, 이에 감사드립니다.”

## 참고문헌

- [1] Video Quality Experts Group, “Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment,” 2000.
- [2] Video Quality Experts Group, “Draft VQEG final report of FR-TV Phase II validation test,” 2003.
- [3] ITU-T, “User requirements for objective perceptual video quality measurements in digital cable television,” Recommendation ITU-T J.143, 2000.
- [4] Zhou Wang Bovik, A.C. Evan, B.L. “Blind measurement of blocking artifacts in images”, Image Processing, 2000. Proceedings. 2000 International Conference on Volume 3, pp. 10-13 Sept. 2000.
- [5] Marziliano, p.; Dufaux, F.; Winkler, S.; Ebragimi, T.; “A no-reference perceptual blur metric”, Image Processing. 2002. Proceedings. 2002 International Conference on Volume 3, pp. 24-28 June 2002.
- [6] Tao Liu “Subjective Quality Evaluation of Decoded Video in the Presence of Packet Losses”, ICASSP, 2007.
- [7] Yamada, Toru, Miyamoto, Yoshihiro, Serizawa, Masahiro “No-Reference Video Quality Estimation Based on Error-Concealment Effectiveness” Packet Video 2007, 2007.
- [8] H.264 / AVC JM reference software home page, <http://iphome.hhi.de/suehring/tml/>
- [9] Yi Guo, Honuqiang Li “SVC/AVC loss simulator”, JVT-Q.069.