

# 인간의 시각적 특성을 이용한 동영상 품질 측정 방식

\*박진철, 이광현, 이상훈  
연세대학교 전기전자공학부

e-mail : dewofdawn@yonsei.ac.kr, kwangsabu@yonsei.ac.kr, slee.yonsei.ac.kr

## Video Quality Assessment based on Human Visual Characteristics

\*Jincheol Park, Kwanghyun Lee, Sanghoon Lee  
School of Electrical and Electronic Engineering  
Yonsei University

### Abstract

In the present paper, a new framework, which is called visual SSIM (VSSIM), is proposed by incorporating crucial human factors into the SSIM. The human factors are foveation, luminance, frequency and motion information. Subjective quality test compliant with the Video Quality Expert Group (VQEG) multimedia group test plan shows that the visual SSIM is more correlated with the subjective quality result than the conventional SSIM.

### I. 서론

최근에 기존의 영상품질 측정을 위한 접근들과는 다르게 HVS는 영상을 인지할 때 신호의 구조적 정보를 추출한다는 새로운 가정이 제기 되었다. 따라서 품질 손실은 특정 타입의 에러에 의해서라기 보다는 신호 자체의 구조적 왜곡에 의해 발생한다는 가정 하에 구조적 왜곡을 측정하는 structural similarity index measure (SSIM)이 제안 되었다 [1]. 이는 더 정확한 품질 손실을 예측한다. 하지만 SSIM은 여전히 추가적인 성능 개선의 여지를 contain하고 있다. SSIM은 HVS이 영상에서 구조적 정보를 인지한다는 HVS의 특성은 고려하지만 그 신호의 특성에 따라서 HVS에 중요한정도가 다르다는 특성은 고려하지 않았다. 하지만 영상 안에서 각 신호의 특성에 따라 HVS가 갖는 중요도가 다르다는 것은 이미 많은 논문의 연구 결과를 통해서 증명된 사실이다[2]. 이를 바탕으로 본 논문에서는 SSIM의 가정에 HVS의 시각적 중요도에 대한 특성을 추가하여 개선된 방법을 제안한다. 구체적으로,

포비에이션의 모델로 spatial weight를, 휘도와 주파수를 이용하는 모델인 information allocation function (IAF)로는 IAF weight를, 모션정보를 이용해서 motion weight를 정의 하였다. 이러한 시각적 중요도들은 결합되어 8x8의 모든 블록을 샘플윈도우로 하는 구조적 유사성의 결과에 적용되어 각 블록마다의 구조적 유사성의 결과를 조절한다. 본 논문에서는 SSIM의 강점과 HVS의 sensitivity 특성을 자연스럽게 결합하는 프레임 워크를 제시함으로써 성능을 향상시킨 것에 기여점이 있다.

### II. 본론

본 논문에서는 신호 자체가 시각적으로 얼마나 중요한 정보를 담고있느냐에 따라 동일한 정도의 구조적 왜곡에서도 시각적으로 미치는 영향의 정도는 다르다는 가정하에 신호 자체의 시각적 중요도를 정의한다. 따라서 각 신호 자체가 지니고 있는 특성들이 HVS에 얼마나 큰 영향력을 가지고 있는지를 중요도로 정의하고 적용함으로써 각 신호마다의 구조적 왜곡의 전체 스코어에 대한 영향력을 중요도에 따라 다르게 반영시킨다. 물론 본 논문에서의 HVS 모델도 이상적일 수는 없지만 이전에 다른 인지적 영상 품질 측정 방식들의 모델들보다 신뢰성 있는 성능을 보이며 이는 기존 SSIM의 분명한 성능향상을 유도함을 본 과제의 실험 결과를 통해서 알 수 있다. 이는 실제 주관적 화질 평가가 주어진 이미지 자체의 수학적 관계를 통해서 직접적으로 유도되기 보다는 HVS를 통해서 어떻게 인지되는가에 의해 정해지기 때문이다. 따라서 본 논문은 이러한 시각적 가중치를 사용하는 영상품질 메트릭들의 기본 원리를 SSIM의 우수한 두 영상간의 구조적 유사성 측정 능력에 적용함으로써 하향식 접근과 상향식 접근방식들의 장점을 동시에 취한것이라 할 수 있다. 따라서 추후에 훌륭한 시각적 모델들이 나온다면 추가적인 성능 향상도 기대할 수 있다.

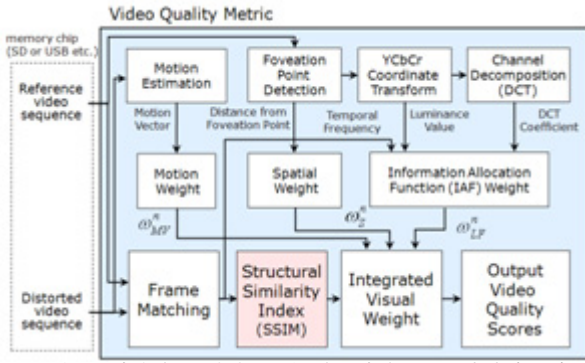


그림 1. 제안하는 영상 품질 매트릭의 블록 다이어그램

그림 1.는 제안하는 영상품질 매트릭의 블록 다이어그램을 나타낸다. 여기에서의 모든 과정은 8x8 의 블록 단위로 진행된다. 이는 HVS의 특성과 SSIM을 자연스럽게 결합하는것을 가능하게 한다. 구조적인 정보를 추출하기 위해서는 8x8의 샘플 윈도우가 필요한데 이것이 매 블록이 된다. 시각적 중요도 또한 매 블록 단위로 정의될 수 있다. 따라서 매 블록마다 구해진 SSIM의 결과는 매 블록마다의 시각적 중요도에 의해서 전체 score에 반영되는 정도가 달라지게 된다. 여기에서의 SSIM은 두 영상간의 구조적 유사도를 비교하는 기본 측정 장치 처럼 사용되어 기존의 픽셀단위로 단순히 두 영상간의 거리를 측정했던 방식을 대체 한다. 어찌보면 SSIM이 error sensitivity metric으로부터 직접적으로 개선된 점은 HVS model의 한계점을 극복한것이 아니라 두 영상간의 비교방식이 우수하다는 것이기 때문이다.

### III. 시뮬레이션

VQEG Multimedia (MM) Group에서는 영상 품질 매트릭의 성능을 검증하기 위해 주관적 화질 평가의 결과와 비교를 하기위한 테스트 플랜을 제공한다. 본 과제에서는 이 테스트 플랜을 바탕으로 주관적 화질 평가를 실시했으며 제안하는 영상 품질 매트릭의 성능을 평가하였다. 테스트를 위해 총 18개의 테스트 시퀀스가 사용되었으며 28명의 관찰자가 동원되었다. 각 테스트 시퀀스는 H.263 코덱을 이용하여 여러 가지 다양한 비트율과 프레임 율을 적용하여 인코딩 하였다. 주관적 화질 평가를 위해서는 adjectival categorical rating (ACR) 방식을 이용하였는데 이는 테스트 영상을 한번씩만 보여주고 Excellent, Good, Fair, Poor, Bad 중에서 하나에 관찰자가 보팅하는 방식으로써 VQEG Multimedia (MM) Group에서 주관적 화질 평가를 위해 사용하는 방식이다. 그림 2는 주관적 화질 평가를 통해 구한 스캐터 플롯이다.

주관적 테스트 결과를 이용하여 객관적 결과와 비교하는 매트릭으로는 pearson linear correlation coefficient, spearman's rank-order correlation coefficient, outlier ratio를 이용하였다. Pearson linear correlation coefficient은 주관적 화질결과와 객관적 결과와의 상관도를 구하는 기준이며 spearman's rank-order correlation coefficient는 두 결과의 monotonicity 혹은 consistency를 측정하기 위한 기준이다. Outlier ratio는 주관적 테스트 결과가 기준 값으로부터 일정 범위 이상 떨어진 결과들의 비율을 구하는 기준이다. 이러한 3가지 기준을 이용하여 비교한 결과는 그림 3과 같다.

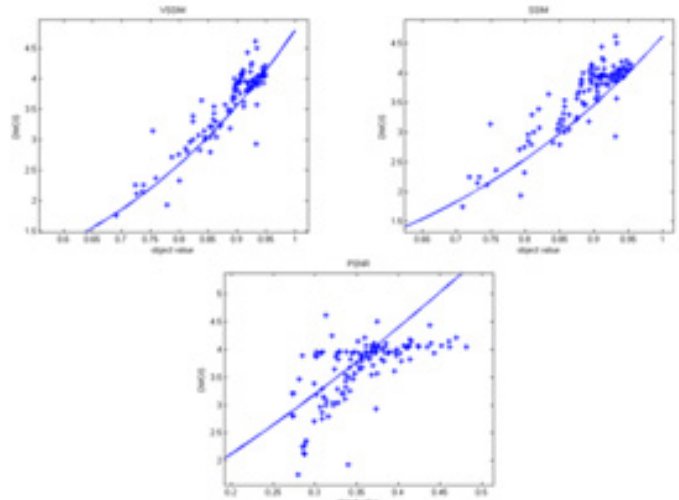


그림 2. 스캐터 플롯그림



3. 주관적 화질 평가 결과

이 그래프에서 볼 수 있듯이 단순한 상관도만을 비교하는 pearson correlation의 경우 기존의 SSIM과 비교해 봤을 때 약간의 성능 개선이 있음을 알 수 있다. 하지만 monotonicity 혹은 consistency 를 측정하는 spearman's rank-order correlation의 경우 높은 성능 개선을 얻었다.

### 감사의 글

이 논문은 2008년도 정부(과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임(No. R01-2007-000-11708-0). 또한 본 연구는 서울시 산학연 협력사업 중 (과제번호:11136) 보유기술 사업화 지원 사업의 연구 결과로 수행되었음.

### 참고문헌

[1] Zhou Wang, Ligang Lu and A. C. Bovik "Video quality assessment based on structural distortion measurement," *Signal Processing Image Communication*, vol.19, no.2, pp.121-132, Feb. 2004.  
 [2] S. Winkler, "Issues in vision modeling for perceptual video quality assessment," *Signal Processing*, vol. 78, pp. 231-252, Oct. 1999.