

## 주목도 지도가 화질 평가에 미치는 영향

권보준, \*윤일동, 이상욱  
서울대학교, \*한국외국어대학교

baukwon@spl.snu.ac.kr yun@hufs.ac.kr sanguk@spl.snu.ac.kr

### The Effect of Saliency Map on Image Quality Assessment

Bojun Kwon \*Il Dong Yun Sang Uk Lee

Seoul National University \*Hankuk University of Foreign Studies

#### 요 약

영상에서의 화질 평가는 학술적으로나 산업적으로나 중요한 문제이지만 지금까지 주로 쓰여온 PSNR 과 같은 방법은 실제 사람이 인식하는 평가에 잘 부합하지 못하는 문제점이 있었다. 본 논문에서는 모든 밝기 영역에서 영상을 평가할 수 있는 HDR-VDP2 와 주목도 지도를 결합하여 범용적으로 사용이 가능하며 성능도 뛰어난 화질 측정 방법을 제시하고, 다양한 주목도 지도에 대하여 그 성능과 화질 평가 성능 사이의 관계를 살펴본다. 구현에는 주목도 지도를 가중치로 사용함으로써 간단하게 더 좋은 성능의 화질 평가 시스템을 만들었고 이를 실험으로 보였다. 또한 주목도 지도의 성능과 화질 평가 시스템의 성능 사이에는 약한 양의 상관관계가 있는 것으로 나타났는데 주목도 지도와 함께 구조적 특징점들의 정보를 성능 평가 시스템에 포함시키면 더 좋은 결과를 얻을 것으로 기대된다.

#### 1. 서론

영상에서의 화질 평가는 영상 압축이나 초해상도화, 합성 등의 방법에서 결과를 서로 비교할 수 있어 학술적으로나 산업적으로나 중요한 문제이다. 이를 위한 대표적인 방법의 하나인 영상을 하나의 신호로 취급하는 PSNR(Peak Signal-to-noise Ratio)이 아직도 많이 쓰이고 있다. 하지만 PSNR 은 실제로 사람들이 인식하는 화질 평가와 잘 부합하지 못하기에 이를 대체할 방법의 필요성이 대두되었다.[1]

그러한 방법의 하나로 등장한 SSIM(Structural Similarity)[2]는 단순한 신호의 차이 대신 원본 영상과의 구조적인 차이를 표현하였다. 이 방법은 간단하면서도 사람의 인식과 비슷한 성능을 보여주어서, 디노이징이나 인페인팅 등의 알고리즘을 평가하는 데 쓰이기도 했다. 또한 그 단순함으로 인해 MS-SSIM(Multi-scale SSIM), P-SSIM(Percentile SSIM), F-SSIM(Fixation SSIM)[3]처럼 다양한 방법으로 확장되기도 했다.

앞선 방법들이 모두 이미지 도메인에서 화질을 평가하는 반면 HDR-VDP2[4]는 실제 물리적 공간에서 화질을 평가한다. 이 방법은 모니터의 방출 스펙트럼이나 눈 내부의 산란, 가리움 효과 등을 효율적으로 모사하여 눈으로 인식하는 차이를 계산하여 영상을 평가한다. 따라서 영상 표시 장치에 따른 변화나 앞으로 발전할 HDR 영상의 평가에도 사용할 수 있는 범용적인 방법이라 할 수 있다.

한편 사람들에게 의한 영상의 화질 평가는 주관적인 부분에 큰 영향을 받는다. 그러므로 사람들이 주목하는 부분에 왜곡이

생기면 그렇지 않은 부분에 왜곡이 생기는 것에 비해 더 나은 점수를 주게 된다. 이를 이용하는 것이 주목도 지도를 이용한 화질 평가로, 이미 효과가 있음이 보고되었다.[5]

본 논문에서는 위의 개념들을 이용하여 모든 밝기에서 적용 가능한 사람의 인식에 가까운 화질 평가 방법을 제안한다. 그리고 이를 아이 트래킹을 통해 구한 주목도 지도와 다른 알고리즘들로 얻어진 주목도 지도를 사용하여 성능을 비교해 보고 사용하는 주목도 지도와 화질 평가 성능과의 관계를 살펴본다.

#### 2. 주목도를 이용한 화질 평가

모든 밝기에서 적용이 가능하도록 기본이 되는 방법으로 HDR-VDP2 를 이용하였다. 이 방법에서는 원본 영상과 왜곡 영상에 대해 각각 시각적으로 어떻게 받아들여지는지를 구하게 된다. 이 둘의 차이를 통해 각 픽셀 별로 원본 영상과 왜곡된 영상의 차이를 관측자가 얼마나 눈치챌 수 있는지를 구하게 된다. 이렇게 구해진 시각적 차이를 모아 하나의 수치로 나타냄으로써 화질을 평가한다.

제안하는 방법에서는 아래 식과 같이 앞서 픽셀 별로 구해진 시각적 차이에 0 에서 1 로 정규화 시킨 주목도 지도를 가중치로 줌으로써 화질을 측정한다.

$$D'(x, y) = S(x, y) \times D(x, y)$$

여기서 S 와 D 는 각 픽셀 별로 구해진 주목도와 시각적 차이이고 D'은 가중치가 적용된 시각적 차이이다.

### 3. 실험 결과

실험에는 LIVE Image Quality Assessment [6] 데이터베이스가 사용되었다. 이 데이터베이스는 29 장의 고해상도 영상(주로 768x512)에 대하여 JPEG2000 압축, JPEG 압축, 백색 잡음, 가우시안 블러, JPEG2000 전송 오류(fast fading)의 5 가지 왜곡에 대해 다른 강도로 구성되어있다. 20~29 명의 사람들에게 화질을 평가하게 한 후 982 장의 영상에 대해 DMOS(difference mean opinion score)를 기록해 놓았다.

우선 제안하는 방법을 검증하기 위해 아이 트래킹 데이터를 주목도 지도로 사용하여 실험을 하였다. 아이 트래킹 데이터는 주목도 지도의 정확도를 측정할 때 정답으로 이용되고, 눈에 띄는 부분이라는 제안하는 알고리즘의 요구 조건에도 부합한다. 사용한 데이터베이스는 TUD LIVE Eye Tracking [7]으로 LIVE IQA 데이터베이스의 원본 영상들에 대해 아이 트래킹을 한 결과들을 포함하고 있다.

실험은 주어진 데이터베이스에 대해 화질 평가 알고리즘을 평가하는 데 사용되는 스피어만 상관계수를 구하는 것으로 진행되었다. 결과는 아래 표 1 과 같다.

왜곡 형태	기존 방법	제안한 방법
JPEG2000	0.9774	0.9765
JPEG	0.9597	0.9578
백색 잡음	0.9800	0.9860
가우시안 블러	0.9534	0.9735
JPEG2000 전송 오류	0.9763	0.9800
모두	0.9567	0.9594

표 1 기존 방법과 아이 트래킹 데이터를 주목도 지도로 하여 구한 제안된 방법에서의 왜곡 형태에 따른 화질 평가 성능 비교

JPEG2000 과 JPEG 의 경우에는 기존의 기법에 비해 조금 성능이 떨어졌지만 다른 경우에 좋은 성능을 보여줌으로써 전체적인 성능이 향상되었다.

다음으로 컴퓨터를 이용해 추출한 주목도 지도의 성능과 이를 이용한 화질 평가 기법의 성능의 관계를 살펴보았다. 실험에는 Itti [8], GBVS [9], Judd [10] 의 세가지 다른 주목도 추출 알고리즘을 사용하였다. 각 알고리즘을 이용하여 29 장의 원본 영상에 대해 주목도 지도를 구한 뒤 TUD LIVE Eye Tracking 데이터를 기준으로 하여 성능을 평가하였다. 성능 평가는 영상마다 ROC 곡선을 그려 AUC 값을 구함으로써 측정되었다.

왜곡 형태	기존	Itti	GBVS	Judd
JPEG2000	0.9774	0.9791	0.9804	0.9790
JPEG	0.9597	0.9597	0.9599	0.9591
백색 잡음	0.9800	0.9859	0.9847	0.9822
가우시안 블러	0.9534	0.9775	0.9794	0.9639
JPEG2000 전송 오류	0.9763	0.9755	0.9753	0.9769
모두	0.9567	0.9610	0.9613	0.9594

표 2 기존 방법과 여러 기법으로 얻은 주목도를 이용한 제안된 방법에서의 왜곡 형태에 따른 화질 평가 성능 비교

컴퓨터로 구한 주목도를 사용한 결과는 표 2 와 같다. 이 경우에도 기존의 방법보다 전체적으로 향상된 결과를 보였다. 평균 AUC 는 Itti, GBVS, Judd 가 각각 0.7899, 0.8445, 0.8674 로 AUC 가 낮은 방법에서 오히려 더 크게 성능이 향상되었다. 주목도 지도의 성능과 화질 평가 성능의 관계를 알기 위해, 기존 방법과 주목도를 사용한 방법에서 각 영상 별로 스피어만 상관계수를 구하였다. 구해진 스피어만 상관계수의 비와 AUC 사이의 피어슨 상관계수는 0.2445 로 약한 양의 상관관계를 보였다.

### 4. 결론

주목도를 사용하는 제안하는 방법이 기존의 방법에 비해 영상에서의 화질 평가에 더 좋은 성능을 보이는 것을 확인하였다. 하지만 주목도 지도의 성능과 화질 평가의 성능 사이에는 약한 상관관계가 있는 것으로 나타났다. 이는 주목도 지도에서 주로 사용하는 영상의 구조적 특징점들이 왜곡되면, 원래는 사람의 주목을 받지 않았더라도 성능 평가에 큰 영향을 끼치게 되는 것으로 해석할 수 있다.

향후에는 주목도 지도와 구조적 특징점들을 동시에 고려하여 화질 평가의 성능을 개선시킬 수 있을 것이다. 또한 주목도 지도를 단순 가중치로 주는 것 이상으로 활용할 수 있는 방법을 찾아내는 것도 중요하다.

### 5. 참고문헌

- [1] H. R. Sheikh et al., A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms, IEEE Transactions on Image Processing, 2006.
- [2] Z. Wang et al., Image Quality Assessment: From Error Visibility to Structural Similarity, IEEE Transactions on Image Processing, 2004.
- [3] A. K. Moorthy and A. C. Bovik, Visual Importance Pooling for Image Quality Assessment, IEEE Journal of Selected Topics in Signal Processing, 2009.
- [4] R. Mantiuk et al., HDR-VDP-2: A Calibrated Visual Metric for Visibility and Quality Predictions in All Luminance Conditions, ACM Transactions on Graphics, 2010.
- [5] A. Guo et al., Visual Attention Based Image Quality Assessment, IEEE International Conference on Image Processing, 2011.
- [6] H.R. Sheikh et al., LIVE Image Quality Assessment Database Release 2, <http://live.ece.utexas.edu/research/quality>
- [7] H. Liu and I. Heynderickx, TUD Image Quality Database: Eye-Tracking Release 1, [http://mmi.tudelft.nl/iqlab/eye\\_tracking\\_1.html](http://mmi.tudelft.nl/iqlab/eye_tracking_1.html).
- [8] L. Itti et al., A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, IEEE Transaction on Pattern Analysis and Machine Intelligence, 1998.
- [9] J. Harel et al., Graph-based Visual Saliency, Advances in Neural Information Processing System, 2006.
- [10] T. Judd et al., Learning to Predict Where Humans Look, IEEE International Conference on Computer Vision, 2009.