

영상 화질평가 기술동향

Trends in Visual Quality Assessment

이상광 (S.K. Lee) 콘텐츠보호관리연구팀 선임연구원
유원영 (W.Y. Yoo) 콘텐츠보호관리연구팀 선임연구원
서영호 (Y.H. Suh) 콘텐츠보호관리연구팀 팀장

* 본 연구는 문화체육관광부의 클린 클라우드를 위한 내용기반 이용제한 기술 개발 사업의 연구 결과로 수행되었음(2012-cloud-9500).

디지털 영상은 캡처 디바이스를 통해 획득된 후, 처리, 압축, 저장, 전송, 재생산 등의 과정을 거치면서 반드시 왜곡이 수반된다. 이렇게 왜곡된 영상은 인간 지각이 허용할 수 있는 범위 내에서 정확히 재생되어야 한다. 따라서, 영상이 재생되기 전 영상의 화질을 평가할 필요가 있다. 실제 응용에 있어서, 인간의 견해로 영상 화질의 점수를 정하는 주관적 화질평가는 일반적으로 불편하고 시간도 많이 소요되며, 비용이 많이 든다. 예를 들어, 사진, 동영상 등의 영상 콘텐츠를 대규모로 저장하고 있는 클라우드 서비스의 경우, 영상 콘텐츠가 상용으로 서비스되기 전 모든 영상의 화질을 인간이 직접 평가하는 것은 불가능하다. 영상 화질평가(visual quality assessment)에 대한 연구목적은 지각된 영상의 화질을 자동으로 예측하는 측도를 개발하는 것이다. 본고에서는 기존에 제안된 영상 화질평가 기술들을 설명하고, 앞으로의 연구에서 해결해야 할 이슈들을 살펴본다.

사용자 중심
차세대콘텐츠기술 특집

- I. 서론
- II. 이미지 화질평가
기술동향 및 이슈
- III. 비디오 화질평가
기술동향 및 이슈
- IV. 결론

1. 서론

기술 발전으로 인해 우리는 쉽고 효율적으로 이미지 및 비디오 영상을 캡처하고, 전송, 저장, 공유하며, 휴대형 단말기로부터 대형 HD 스크린에 이르기까지 다양한 기기들로 그 내용을 감상할 수 있다. 인간이 시각에 매우 예민한 창조물이라는 사실과 맞물려 이러한 일련의 과정 중에 재생산되는 이미지 및 비디오 영상은 인간 시각이 허용할 수 있는 범위 내에서 정확히 재생되어야 한다. 이러한 시각 신호를 캡처하는 데 사용하는 기기의 물리적 특성의 한계와 제한된 대역폭으로 정보손실과 손상이 반드시 발생하게 된다. 영상 화질평가(Quality Assessment: QA) 연구는 이러한 왜곡이 인간이 지각하는 화질에 어떠한 영향을 미치는지 이해하여 화질을 측정하는 데 그 목적을 두고 있다[1].

객관적인 화질평가(objective quality assessment)는 인간의 시신경으로 지각된 화질(perceived quality)을 측정하는 알고리즘을 구현하는 것이다. 이러한 정의를 위해 ‘지각된’ 화질의 정의가 필요하다. 어떠한 시각 자극을 최종적으로 수신하는 것은 인간 관찰자이기 때문에, 화질에 대한 견해가 그 시각 자극에 대한 지각된 화질로 간주된다. 그러한 인간의 지각된 화질은 상당량의 시각 자극에 대해 특정 등급을 매긴 결과를 분석하여 측정된다. 이러한 시각 자극의 평균 점수를 평균 평가점(Mean Opinion Score: MOS)이라고 하며 대표적인 지각된 화질 측정 방식이다. 이러한 인간의 화질평가를 주관적인 화질평가(subjective quality assessment)라고 한다[2]. 현재, 이미지 화질평가[3]와 비디오 화질평가[4]를 위한 대규모 데이터베이스가 공개되어 일반인이 접근하여 이용할 수 있다.

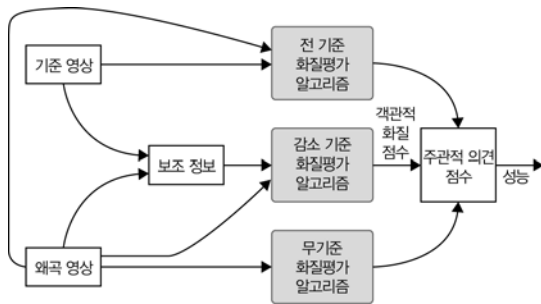
객관적인 화질평가 알고리즘의 성능은 인간 지각과 어느 정도 연관성이 있는가로 측정한다. 즉, Spear-

man의 순위 상관계수(Spearman's Rank Ordered Correlation Coefficient: SROCC), 선형 상관계수(Linear Correlation Coefficient: LCC), 평균 제곱근 오차(Root-Mean-Squared Error: RMSE), 가외치 비율(Outlier Ratio: OR) 등과 같은 다양한 통계 측도를 이용하여, 화질평가 알고리즘에 의해 도출된 객관적 화질 점수와 주관적 견해 점수(MOS) 간의 연관성을 측정하여 알고리즘의 성능을 평가한다. SROCC와 LCC는 1에 가까울수록, RMSE, OR은 0에 가까울수록 알고리즘 성능이 좋다는 것을 나타낸다. 이러한 알고리즘 성능 평가를 위한 측도에 관한 연구가 여전히 진행되고 있다.

객관적인 화질평가 방법은 일반적으로 다음과 같이 전 기준(Full-Reference: FR), 감소 기준(Reduced-Reference: RR), 무기준(No-Reference: NR) 방식으로 분류될 수 있다[5].

전 기준 화질평가(FR QA) 방식은 ‘왜곡된 영상’과 비교할 수 있는 ‘기준 영상(reference image)’을 이용할 수 있다. ‘기준 영상’은 일반적으로 영상이 압축이나 전송오류에 의해 왜곡되기 전에 고화질 캡처 기기로 얻어진 영상이다. 하지만, 기준 영상은 왜곡 영상보다 훨씬 더 자원비용이 발생하여, FR QA는 영상처리 알고리즘 설계 시 연구실 수준의 실험용으로만 사용되며 실제 응용에는 쓰이지 않는다.

반면, 무기준 화질평가(NR QA) 방식은 기준 영상의 어떠한 정보도 이용하지 않고 왜곡된 영상만을 접근하여 화질 추정을 한다. NR QA 방식은 기준 영상 정보를 필요로 하지 않기 때문에, 화질 측정이 요구되는 어떠한 응용에도 이용될 수 있다. 하지만, 이러한 가용성으로 인해 화질 예측 정확도에 어느 정도 희생이 수반되며, NR QA 알고리즘의 적용범위도 JPEG, JPEG2000, 번짐 현상(blur) 등과 같이 특정 왜곡으로 제한된다.



(그림 1) 객관적 화질평가 접근 방식과 성능 평가

마지막으로, 감소 기준 화질평가(RR QA) 방식은 위 두 방식의 중간 위치에 있으며, 기준 영상 자체가 아닌 기준 영상에 관한 일부 정보만(예를 들면, 워터마킹(watermarking), 보조 채널(auxiliary channel) 등)을 이용하여 화질평가를 한다(그림 1) 참조).

본고에서는 기존의 이미지 화질평가(Image Quality Assessment: IQA) 기술과 비디오 화질평가(VQA: Video Quality Assessment) 기술을 전 기준/무기준 알고리즘으로 나누어 살피고 앞으로의 연구에서 해결해야 할 기술 이슈들을 살펴보고자 한다.

II. 이미지 화질평가 기술동향 및 이슈

본 장에서는 IQA 기술동향을 FR 방식과 NR 방식으로 나누어 살피고 각 방식에 대해 앞으로 진행되어야 할 연구 분야에 대해 알아본다. 또한, IQA 알고리즘의 성능을 평가하기 위한 IQA 데이터베이스에 관한 이슈에 대해서도 살펴본다.

1. FR IQA 기술

간단하면서도 가장 보편적으로 이용되었던 화질 측도로는 왜곡 영상과 기준 영상의 제곱근 화소값 차이를 평균한 제곱근 오차(Mean-Square Error: MSE)가 있으며, 최대 신호대 잡음비(Peak Signal-to-

Noise Ratio: PSNR)는 이렇게 계산된 MSE를 이용한 화질 측도이다. 이러한 측도들은 계산이 간단하고, 물리적 의미도 직관적이며 수학적으로 최적화하기에 용이하지만, 지각된 화질과 만족할 만하게 일치되는 편은 아니다. 최근 객관적인 FR IQA에 관한 연구는 인간 시각 시스템(Human Visual System: HVS) 모델 기반 방식과 자연 장면 통계(natural scene statistics) 기반 방식에 초점이 맞춰지고 있다. 그 중 구조적 유사도(Structural SIMilarity: SSIM)[6] 인덱스와 시각 정보 정확도(Visual Information Fidelity: VIF)[7] 인덱스는 PSNR에 비해 통계학적으로 중요한 성능을 나타낼 뿐만 아니라 다른 IQA 방식들과 비교해도 인간 시각 관점에서 좋은 결과를 보여준다.

SSIM 인덱스의 간소함과 뛰어난 성능 때문에 그보다 좀 더 나은 방식에 대한 연구가 활발히 진행되었다. 비록 이러한 연구 결과들 중 일부가 SSIM보다 소폭 개선된 성능을 보여줬지만, 성능 개선에 의한 이점은 매우 미비하거나 통계학상으로 무의미했다. 따라서, FR IQA 분야는 현재 안정기에 이르렀다. 즉, 성능 개선에 의한 이점은 매우 적고 이러한 성능 개선을 위한 비용은 그에 비해 적지 않다.

앞으로 진행되어야 할 FR IQA 연구 분야 중 하나는 효율적인 화질 계산에 관한 연구이다. 이러한 연구는 IQA에서 뿐만 아니라, 화질 측정을 실시간으로 구현해야 하는 VQA에 적용될 수 있다.

수많은 성공적인 IQA 방식이 제안되었지만, 이러한 알고리즘들은 단지 벤치마킹으로만 이용되어 왔다. 게다가 객관적인 화질 측정 함수에 이용되는 영상 처리 알고리즘의 최적화는 상대적으로 연구가 미비한 상태이다. 일부 학자는 최적화를 위한 기준으로 널리 쓰이는 MSE 대신 화질평가 측도를 잡음제거 및 선형 필터링에 이용하고 있다. 하지만, 이 분야에 대한 연구 내용이 여전히 남아있다. 인간 지각에 잘 연관된

IQA 알고리즘은 압축 또는 필터링 최적화에도 이용될 수 있다. VIF와 같은 복잡한 모델은 그러한 연산을 쉽게 만드는데 적합하지 않다. 앞으로의 연구에서는 압축/필터링을 최적화할 수 있도록 화질 측정자 분석이 포함될 것으로 예상된다.

위에서 언급했듯이 FR IQA 알고리즘은 성능에 관한 한 포화 상태에 도달했다. 하지만, HVS를 좀 더 잘 이해한다면 의미심장한 성능 개선이 있을 것으로 본다. 더욱이 화질평가를 위한 주시(fixation)와 안구운동(eye-movement)에 대한 연구는 앞으로의 연구에 있어 또 다른 중요한 분야다[2]. 요약하면, HVS 기반의 FR IQA 알고리즘 개발이 활발히 진행될 것으로 기대되며, 객관적인 IQA 측도는 다양한 영상 처리에 응용될 것이다.

2. NR IQA 기술

대부분 최근의 NR IQA 알고리즘은 특정 왜곡을 기준으로 개발되었다. 예를 들면, JPEG, JPEG2000 등과 같은 압축 영상이나 번짐 영상의 화질을 평가하기 위한 수많은 알고리즘이 제안되었다. 이러한 NR IQA 알고리즘에서는, 왜곡에 특화된 화질 지표가 계산되고 주관적인 화질평가 점수를 이용하여 화질 측도로 함수화시켰다. 일반적으로, 이러한 알고리즘들은 대부분 윤곽선 정보를 추출한 후 번짐이나 물결 현상(ringing)에 대한 화질평가를 위해 윤곽선 확산 또는 윤곽선 분포를 찾는다. 왜곡에 따라 측정된 이러한 지표들은 매개변수를 이용한 함수 형태로 조합되며, 매개변수는 주관적 데이터베이스의 화질 점수를 이용하여 추정된다.

일반적인 NR JPEG IQA 방식은, 블록 경계에서 윤곽선의 세기를 측정하고 그 세기와 이미지의 활동성을 나타내는 특징들을 지각된 화질과 연결시킨다. 예를 들면, 이미지의 1차 차분과 영교차율(zero-crossing

rate)을 이용하여 블록 현상과 이미지 활동성을 측정 한 후, 블록 현상과 활동성 특징들을 조합하여 지각된 화질을 추정한다[8].

NR JPEG2000 IQA 알고리즘은 일반적으로 윤곽선 검출 기반의 방식을 이용하여 윤곽선 확산을 측정함으로써 이미지 내 물결 현상을 모델링한다. 이 윤곽선 확산이 화질과 연결된다. 다른 방식으로는 자연 장면의 통계 특성을 이용하는 방식[9]이 있다. 자연 영상은 전체 영상의 아주 작은 부분에서 도출되고, 왜곡 과정이 그 작은 부분들의 통계를 흐트러뜨리기 때문에, 자연 영상 통계로부터 전체 영상의 편차가 계산될 수 있으며 화질과 연결시킬 수 있다. 참고문헌 [9]에는 왜곡이 존재할 때, 하나의 웨이블릿 계수와 이웃하는 계수 사이의 의존성이 변경된다는 사실을 이용하였다. 이 알고리즘은 JPEG2000 압축뿐만 아니라 웨이블릿 기반의 압축에 응용될 수 있다. 번짐 현상에 대한 NR IQA 알고리즘은 NR JPEG2000 IQA와 유사하게, 윤곽선 확산을 모델링한 후, 이렇게 얻어진 윤곽선 확산을 지각된 화질과 연결시킨다.

위에서 언급한 JPEG, JPEG2000, 번짐 현상 등의 왜곡에 특화된 NR IQA 알고리즘은 특정 왜곡에 대한 화질평가 방식이기 때문에, 왜곡을 알 수 없는 일반적인 시나리오에서는 적합하지 않다.

NR IQA에 관한 많은 연구가 진행되고 있지만, 대부분의 연구는 위에서 언급한 왜곡에 특화된 알고리즘에 집중되어 있다. 최근 왜곡에 독립적인 접근 방식이 참고문헌 [10], [11]에 제안되고 있으며, 이 방식들은 모두 자연 영상의 통계적 특성에 기반하고 있다.

참고문헌 [10]에서는 DCT 영역 기반 방식이 제안되었다. 이 방식은 영상의 매 화소를 중심으로 일정한 크기의 영상패치에 2D DCT를 적용한 후, DCT 계수를 통해 화질과 관련 있는 특징을 두 스케일에서 계산한다. 계산된 특징은 매개변수를 이용한 접근 방식

을 이용하여 조합되어 합수화된다. 이 NR IQA 방식은 HVS에 잘 부합되고, DCT 프레임워크에 기반이기 때문에 계산이 편리하며, FR IQA 방식인 PSNR보다 좋은 성능을 보여준다.

또한, 참고문헌 [11]에서는 NR IQA를 위한 두 단계 프레임워크가 제안되었다. 왜곡된 영상의 통계를 이용하여 왜곡 범주를 결정한 후, 왜곡 범주를 인식한 IQA는 화질평가 측정을 수행한다. 이 NR IQA 방식은 모든 왜곡에 대해 PSNR 성능과 비슷하거나 많은 경우 앞선 성능을 보여주었다. 하지만, 위의 두 방식 모두 훈련/시험(train-test) 조합에 기반하기 때문에, 이러한 접근 방식들은 여전히 왜곡에 완전히 독립되지 않는다.

대부분의 왜곡에 특화된 QA 알고리즘들이 좋은 성능을 보여주기 때문에, 어떠한 왜곡을 포함한 영상에 대해서도 화질평가가 가능한 왜곡에 독립된 NR IQA 알고리즘이 앞으로의 연구의 주요 목표가 될 것이다. 더욱이 현재 NR IQA 성능은 FR IQA 알고리즘의 성능에 미치지 못하고 있다. FR IQA 경우와 마찬가지로, NR IQA에 인간의 지각 메커니즘을 도입하는 것이 이러한 알고리즘에 유용할 것이다.

3. IQA 데이터베이스

현재 알고리즘 성능은 일반인에게 공개된 IQA 데이터베이스들[3],[4]을 통해 측정된다. 비록 이러한 데이터베이스들이 다양한 왜곡 범주를 포함하고 있지만, 대부분의 범주가 압축, 잡음, 번짐 현상에 대한 단일 왜곡만을 포함하고 있다. 현재 특별히 다중 왜곡 영상과 그에 대한 주관적 화질 점수를 도입한 데이터베이스는 존재하지 않는다. 즉, 심하게 압축된 잡음 영상이 있다면, 압축과 잡음이 서로 상호작용하여 화질평가에 어떻게 영향을 미치는지, 한 범주의 왜곡이 다른 범주의 왜곡을 감출 수 있는지, 상대적인 왜곡

감도를 어떻게 화질평가에 포함할 것인지를 시험할 수 있는 다중 왜곡 영상 데이터베이스가 존재하지 않는다. 실제 영상들은 두 가지 왜곡을 넘어 더욱더 복합적으로 왜곡된다. 예를 들면, 잡음 영상이 압축된 후 손실 채널을 통해 전송된다.

복합적으로 왜곡된 영상에 관한 화질평가 이슈가 관심받기 시작했음에도 불구하고, 이러한 방향으로 진행되고 있는 연구는 많지 않다. 다중 왜곡 영상의 경우, 심지어 선두적인 FR IQA 알고리즘조차 인간 지각에 잘 연계되었는지 불분명하다. 따라서, 인간의 주관적인 화질 점수와 함께 복합적으로 왜곡된 영상 데이터베이스를 구축하는 것은 앞으로 관심을 가져야 할 연구 분야이다. 요약하면, FR IQA 성능을 뛰어넘을 뿐만 아니라 다중 왜곡 영상에 대해서도 화질평가가 가능한 NR IQA 알고리즘에 관한 연구가 진행되어야 한다.

III. 비디오 화질평가 기술동향 및 이슈

VQA 방식은 간단히 IQA의 확장으로 생각할 수 있지만, IQA를 통한 공간상에서의 화질로만으로는 비디오 영상에서 발생하는 왜곡을 모두 표현할 수 없다. VQA 알고리즘은 시간상에서의 화질도 함께 평가해야 한다. 본 장에서는 이러한 관점에서 기존에 제안된 VQA 기술들을 설명하고, 각 방식에 대해 앞으로 해결해야 할 이슈들에 대해 살펴보고자 한다.

1. FR VQA 기술

VQA은 IQA 알고리즘을 프레임 단위로 적용하여 간단히 구현해 볼 수 있다. 예를 들면, MSE나 SSIM을 각 프레임에서 계산하여 비디오의 화질을 측정한다. 비디오의 평균 화질은 프레임레벨 점수의 Min-kowski 합이 된다. 하지만, 많은 학자들이 지적한 바

와같이, VQA에서 움직임(motion) 정보가 상당히 중요하다. 최근 움직임 정보를 이용한 알고리즘들 [12],[13]이 제안되고 있으며 좋은 성능을 보여주고 있다.

참고문헌 [12]에서는 HVS 안에서 움직임 처리 메커니즘을 분석하여 움직임 기반 비디오 무결성 평가(Motion-Based Video Integrity Evaluation: MOVIE) 인덱스를 제안하였다. MOVIE 인덱스는 시공간 필터뱅크(spatio-temporal filter bank)를 이용하여 기준 비디오와 왜곡 비디오를 부대역으로 분해한 후, 각 부대역별로 화질평가를 수행한다. MOVIE 인덱스의 실험결과는 다시 한 번 PSNR이 지각된 화질을 나타내는 좋지 못한 인덱스 중의 하나라는 것을 보여주었다. 비록 MOVIE가 HVS와의 연관성 관점에서는 좋은 성능을 보여주고 있지만, 인덱스 연산의 복잡성으로 실제 응용에서 그 구현이 쉽지 않다. 또한, HVS에서 영감을 받아 개발되었기 때문에 진보된 HVS 모델을 이용하지 않는 한 의미심장한 성능 개선은 어려울 것이다.

참고문헌 [13]에서는 움직임이 보상된 구조적 유사도(Motion Compensated Structural SIMilarity: MC-SSIM) 인덱스가 제안되었다. MC-SSIM 인덱스는 공간상에서의 화질평가를 위해 간단한 단일 스케일 구조적 유사도(Single Scale Structural SIMilarity: SS-SSIM) 인덱스를 이용하고, 시간상에서의 화질평가는 움직임 보상과 SS-SSIM을 조합하여 이용한다. MC-SSIM 인덱스는 참고문헌 [12]에서 제안된 MOVIE에 비해 계산이 간단하며 인간 지각과의 연관성 관점에서도 좋은 성능을 보여준다.

FR IQA와 마찬가지로, FR VQA 알고리즘을 실제 적용하는데 있어서 효율적인 연산이 앞으로 주요 이슈가 될 것이다. 즉, 실시간성을 유지하면서 인간 지각과의 연관성 관점에서 만족할 만한 성능을 유지하

는 것이 도전 과제이다. 일부 학자들은 일반적으로 많이 이용하고 있는 SSIM 인덱스에 대해 실시간성에 관한 문제를 제기해왔다. 따라서, VQA에 대해, 기존에 제안된 접근 방식을 확장하는 것 외에 실시간 화질 예측이 가능한 새로운 알고리즘 연구가 필요하다.

인간의 주의 시각(human attention)은 IQA보다 VQA에서 그 연관성이 더욱 크다. 인간의 주시점(human fixation location)에 대한 정확한 예측은 효율적인 압축뿐만 아니라 지각된 화질을 예측하는 알고리즘을 개발하는 데 큰 도움을 준다. FR VQA의 경우, 기준 비디오 영상으로부터 로우레벨 특징을 추출한 후, 주시점을 효과적으로 예측하여 화질평가 성능을 개선할 수 있다. 하지만, 왜곡이 인간의 주의 시각에 영향을 미칠 수 있다는 증거를 보여주는 것[14]과, 어떻게 왜곡이 인간의 주의 시각에 영향을 미치는가를 분석하는 것은 VQA와 관련된 서로 다른 분야이다. 즉, 인간의 왜곡이 어떻게 인간의 주시와 안구운동에 영향을 미치는 지에 대한 메커니즘 연구가 필요하다.

2. NR VQA 기술

상대적으로 관심받지 못했던 NR VQA 알고리즘은 대부분 압축 비디오에 특화되어 설계되었다. 참고문헌 [15]에서는 프레임 손실량, 객체 움직임, 그리고 국부적인 시간 화질 대비(contrast)로 구성되는 시간상에서의 지각 화질 매트릭(Perceptual Temporal Quality Metric: PTQM)을 제안하였다. PTQM은 규칙적이거나 불규칙한 프레임 손실로 발생하는 시간상 화질 왜곡을 정확히 측정할 수 있으며, 시간상에서의 불연속성으로 인한 인간 시각의 불편함을 정확히 예측할 수 있다. 하지만, PTQM은 시간상에서의 화질평가만 가능하며, 공간상에서의 화질 매트릭과 결합하여 시공간상에서의 화질평가로 발전할 필요가 있다.

VQA에 대한 다른 접근 방식으로, 손실 네트워크를 통해 전송된 비디오에 대해 손실된 패킷 수를 계산하고 패킷 손실의 형태를 분석한 후, 바이트 스트림으로부터 화질에 연결시킬 수 있는 기타 다른 특징을 추출하는 알고리즘도 있다[16]. 왜곡된 비디오 영상의 화질을 예측하는 NR VQA 알고리즘은 여전히 연구 중에 있는 분야이다.

NR VQA의 목표를 달성하기 위해 성공적인 FR VQA 접근 방식으로부터 그 방안을 찾는 것도 의미가 있다. 지각 모델과 움직임 정보를 도입한 NR VQA 알고리즘이 화질평가에 좀 더 성공할 공산이 크다. 실제 최근 움직임 정보 통계 연구를 통해 왜곡된 비디오로부터 움직임을 분석하는 것이 화질평가의 좋은 인덱스라는 것이 밝혀지고 있다[17]. 더 나아가, 왜곡된 비디오를 볼 때 인간의 주의 시각을 예측할 수 있도록 설계된 인간 주의 시각 메커니즘도 앞으로의 연구 분야 중 하나이다.

NR VQA 알고리즘들에 대해, FR VQA에 대해 수행되었던 것[18]과 유사한 체계적인 평가가 없다. 공정한 방식으로 알고리즘들을 비교하기 위해서는, 성능 평가를 위해 어느 정도 합의된 통계분석을 이용하여 공개된 데이터베이스상에서 그 성능이 평가되어야 한다. 앞으로의 연구에서는 NR VQA 알고리즘의 성능을 정확히 측정하는데 이용될 수 있는 대규모의 비교가 필요하다.

VQA의 경우, 시간적 양상이 특히 중요하다. 대부분 IQA를 위한 공간 관점에서의 연구[19]가 수행되고 있으며, VQA에 대한 시간 관점의 연구는 극히 일부이다. 게다가, 재생시간이 긴 비디오일수록 시간 관점의 접근 방식이 더 요구된다.

IQA의 경우와 마찬가지로, FR/NR VQA 알고리즘은 화질 기반의 압축이나 잡음제거[20]와 같은 실질적인 응용에 좀처럼 이용되지 않고 있다. VQA 알고

리즘에 대한 앞으로의 연구에서는 알고리즘적인 화질 평가의 응용에 초점이 맞춰질 것으로 본다. VQA 알고리즘 기반 비트율-왜곡(rate-distortion) 최적화나 움직임 예측이 그 실제적인 응용에 포함될 수 있다.

IV. 결론

지금까지 본고에서는 이미지와 비디오 영상에 대한 화질평가 관련 기술동향을 살펴보았다. 영상 화질평가 기술에 대해 앞으로 해결해야 할 이슈들이 곳곳에 산재해 있지만, 한편으로는 기술 성장 관점에서 무궁무진한 가능성을 가지고 있다. 특히, 음원, 사진, 동영상 등의 멀티미디어 콘텐츠가 기하급수적으로 증가하면서 콘텐츠 품질을 객관적으로 평가하는 일은 매우 중요한 이슈가 되었다. 최근 몇 년간 화두가 되고 있는 클라우드 서비스를 보면, 클라우드 내 대규모 콘텐츠를 서비스할 시점이 다가오면서 콘텐츠 품질을 보장해 줘야 할 필요성이 대두되고 있다. 본문에서 언급했듯이 인간이 직접 품질을 평가하는 일은 시간이 많이 걸리고 비용도 적지 않게 든다. 따라서, 객관적 품질 평가 방식이, 서비스 형태에 맞게 개발되고 적용되어야 할 것이다.

본고에서는 이미지와 비디오 영상에 대한 화질평가 기술만을 다루었지만, 3D 영화로 잘 알려진 3D 비디오 영상에 대한 화질평가도 앞으로 많은 관심을 받으리라 예상된다.

용어해설

영상 화질평가(visual quality assessment) 전송, 저장, 재생산 등을 통해 발생된 영상 왜곡이 인간이 지각하는 화질에 어떠한 영향을 미치는지 분석하여 화질을 측정

지각된 화질(perceived quality) 시각 자극을 최종적 수신하는 인간의 화질에 대한 견해

기준 영상(reference image) 영상이 압축이나 전송오류에 의해 왜곡되기 전에 고화질 캡처 기기로 얻어진 영상

약어 정리

FR	Full-Reference
HVS	Human Visual System
IQA	Image Quality Assessment
LCC	Linear Correlation Coefficient
MC-SSIM	Motion Compensated Structural SIMilarity
MOS	Mean Opinion Score
MOVIE	MOTION-based Video Integrity Evaluation
MSE	Mean-Square Error
NR	No-Reference
OR	Outlier Ratio
PSNR	Peak Signal-to-Noise Ratio
PTQM	Perceptual Temporal Quality Metric
QA	Quality Assessment
RMSE	Root-Mean-Squared Error
RR	Reduced-Reference
SROCC	Spearman's Rank Ordered Correlation Coefficient
SSIM	Structural SIMilarity
SS-SSIM	Single Scale Structural SIMilarity
VIF	Visual Information Fidelity
VQA	Video Quality Assessment

참고문헌

- [1] A.K. Moorthy and A.C. Bovik, "Blind Image Quality Assessment: From Natural Scene Statistics to Perceptual Quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, Dec. 2011, pp. 3350-3364.
- [2] A.K. Moorthy and A.C. Bovik, "Visual Quality Assessment Algorithms : What Does the Future Hold," *Int. J. Multimedia Tools Appl.*, Special Issue on Survey Papers in Multimedia by World Experts, vol. 51, no. 2, Feb. 2011, pp. 675-696.
- [3] H.R. Sheikh, et al., "LIVE Image Quality Assessment Database Release 2," <http://live.ece.utexas.edu/research/quality>
- [4] K. Seshadrinathan et al., "LIVE Video Quality Database," <http://live.ece.utexas.edu/research/quality>
- [5] A.K. Moody, "Image and Video Quality Assessment at LIVE," <http://live.ece.utexas.edu/research/quality>
- [6] Z. Wang et al., "Image Quality Assessment: from

- Error Visibility to Structural Similarity," *IEEE Trans. Image Process.*, vol. 13, no.4, Apr. 2004, pp. 600-612.
- [7] H.R. Sheikh and A.C. Bovik, "Image Information and Visual Quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, Feb. 2006, pp. 430-444.
- [8] Z. Wang, H.R. Sheikh, and A.C. Bovik, "No-Reference Perceptual Quality Assessment of JPEG Compressed Images," *IEEE Int. Conf. Image Process.*, vol. 1, 2002, pp. 477-480.
- [9] H.R. Sheikh, A.C. Bovik, and L.K. Cormack, "No-Reference Quality Assessment Using Natural Scene Statistics: JPEG2000," *IEEE Trans. Image Process.*, vol. 14, no. 12, Dec. 2005, pp. 1918-1927.
- [10] M.A. Saad, A.C. Bovik, and C. Charrier, "A DCT Statistics-Based Blind Image Quality Index," *IEEE Signal Process. Lett.*, vol. 17, no. 6, June 2010, pp. 583-586.
- [11] A.K. Moorthy and A.C. Bovik, "A Two-step Framework for Constructing Blind Image Quality Indices," *IEEE Signal Process. Lett.*, vol. 17, no. 5, May 2010, pp. 513-516.
- [12] K. Seshadrinathan, *Video Quality Assessment Based on Motion Models*, PhD Thesis, University of Texas at Austin, 2008.
- [13] A.K. Moorthy and A.C. Bovik, "A Motion Compensated Approach to Video Quality Assessment," *IEEE Asilomar Conf. Signals, Syst. Comput.*, 2009, pp. 872-875.
- [14] A.K. Moorthy, W.S. Geisler, and A.C. Bovik, "Evaluating the Task Dependence of Eye Movements for Compressed Videos," *5th Int. Workshop Video Process. Qual. Metrics Consum. Electron. (VPQM)*, 2010.
- [15] K.C. Yang et al., "Perceptual Temporal Quality Metric for Compressed Video," *IEEE Trans. Multimedia*, vol. 9, no. 7, 2007, pp. 1528-1535.
- [16] M. Naccari et al., "No-Reference Modeling of the Channel Induced Distortion at the Decoder for H.264/AVC Video Coding," *IEEE Int. Conf. Image Process.*, 2008, pp. 2324-2327.
- [17] M.A. Saad and A.C. Bovik, "Natural Motion Statistics for No-Reference Video Quality Assessment," *Int. Workshop Qual. Multimedia Experience*, 2009, pp. 163-167.

- [18] K. Seshadrinathan et al., "Study of Subjective and Objective Quality Assessment of Video," *IEEE Trans. Image Process.*, vol. 19, no. 6, June 2010, pp. 1427–1441.
- [19] Z. Wang and X. Shang, "Spatial Pooling Strategies for Perceptual Image Quality Assessment," *IEEE Int. Conf. Image Process.*, 2006, pp. 2945–2948.
- [20] G. Varghese and Z. Wang, "Video Denoising Based on a Spatiotemporal Gaussian Scale Mixture Model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 7, July 2010, pp. 1032–1040.