

스마트폰용 영상품질 측정을 위한 주관적 화질평가 분석

Analysis on Subjective Image Quality Assessments for Smart Phone/Pad
Environments

김성철 (S.C. Kim) 융합서비스전략연구실 선임연구원
김병일 (B.I. Kim) (주)다바 책임연구원

- I. 영상품질 측정 방법론 개요
- II. 모바일 비디오 영상의 주관적 화질평가
- III. 모바일 비디오 영상의 주관적 화질평가 결과
- IV. 결론 및 시사점

본고에서는 ITU-T P.910 권고사항을 이용하여 영상 화질의 주관적 평가를 실시하였다. 최악의 평가 결과의 경우를 고려하더라도 스마트폰 등의 무선단말기에서는 500kbps 이상으로 부호화할 경우 원본의 영상과 뚜렷한 차이를 인지하지 못하는 것으로 분석되었으며, 원본영상 대비 테스트 영상의 주관적/객관적 화질평가 결과는 시각적으로 원본 영상과 테스트 영상 간의 품질 차이를 인지하지 못하는 수준인 것으로 나타났다. 이와 같은 점을 고려할 때 입력 영상 및 출력 영상 그리고 대역폭 관계를 고려하여 타깃 비트 레이트(target bitrate)를 적절하게 설정하는 것이 필요하다. 비트 레이트가 증가함에 따라 느끼는 이용자들의 주관적인 만족도는 점차 감소하지만, 네트워크 비용은 체증하기 때문에 적절한 수준의 비트 레이트 설정이 사회적으로 바람직할 수 있다.

I. 영상품질 측정 방법론 개요

1. 객관적 화질 측정: PSNR

객관적 화질 측정 방법에는 다양한 방법이 있으나 PSNR(Peak Signal-to-Noise Ratio)이 가장 대표적이다. PSNR은 주로 영상 또는 동영상 손실 압축에서 화질 손실정보를 평가할 때 사용되는데, 두 영상이나 음성의 차이를 숫자로 표현한 것이다. 일반적으로 동영상의 객관적 품질평가 방법은 원본 영상의 유무에 따라서 전 기준법(full reference), 감소 기준법(reduced reference), 무기준법(no reference)으로 나눌 수 있는데, PSNR은 전 기준법의 한 가지 방식이다.

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I_{orig}(i, j) - I_{imp}(i, j)]^2 \quad \text{식 (1)}$$

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad \text{식 (2)}$$

식 (1)과 식 (2)는 PSNR의 계산 방법으로, MAX_I는 해당 영상의 최대값으로서 해당 채널의 최대값에서 최소값을 빼서 구할 수 있는데, 예를들어 8bit 그레이스케일 영상의 경우는 255(255-0)가 된다. MSE란 두 개의 같은 양의 데이터에 대해 동일한 위치에 대한 분산을 계산하는 것으로, 변수가 i, j 2개인 것은 영상처리를 기준으로 식을 썼기 때문에 가로, 세로를 표현하기 위해 2개(음성처리에서는 1개)인 것이다. I_{orig}(i,j)는 좌표(i,j)의 원래영상 픽셀값이고, I_{imp}(i,j)는 좌표(i,j)의 비교 영상 픽셀값을 나타낸다.

PSNR은 신호가 가질 수 있는 최대 전력에 대한 잡음의 전력을 나타낸 것으로, 신호의 전력에 대한 고려 없이 평균 제곱 오차를 이용해서 계산할 수 있다. 두 영상이 완벽하게 동일하다면 분자인 MSE가 0이므로 PSNR은 무한대가 되며, 일반적으로 30dB가 넘으면 두 영상의 차이를 눈으로 구분하기 어려운 것으로 알려지고 있다. 로그스케일에서 측정하기 때문에, 단위는 db이며,

손실이 적을수록 높은 값을 가진다. 무손실 영상의 경우에는 MSE가 0이기 때문에 PSNR은 정의되지 않는다.

PSNR은 전 기준법 방법으로 원본 영상은 수신측에 있어야 하기에 이동통신 단말기에 사용하기에 적합하지 않으며, 또한 PSNR은 다양한 통신 오류가 발생하는 환경에서 단말 사용자가 느끼는 인지품질을 제대로 반영하지 못하는 것으로 알려져 있다는 단점이 있다.

2. 주관적 화질평가

가. ITU-R BT.500.11

객관적 품질평가 모델의 개발을 위해서는 다양한 서비스 환경을 고려한 주관적 품질평가(subjective quality assessment)가 수행되어야 한다. 주관적 품질평가는 평가자가 직접 동영상을 보고 화질을 평가하는 방법으로서 인간의 화질인지 특성을 반영할 수 있는 가장 정확한 방법으로 알려져 있다[1].

비디오 화질의 주관적 평가(subjective assessment) 방법들 중 가장 기본이 되는 것은 ITU-R BT.500.11 권고사항이다. ITU-T BT.500.11 규격은 TV 화질에 대한 주관적 평가 방법에 관한 권고사항으로 주관적 화질평가에 사용되는 실험환경 조건(디스플레이 밝기, 평가실의 조건, 평가자의 위치, 스크린과의 거리 및 각도 등)을 기술하고 있으며, 평가하고자 하는 문제에 따라 적용 가능한 여러 가지 실험 평가 방법과 실험결과 분석 방법에 대하여 기술하고 있다. 특히 위 권고사항에 여러 방법 중 DSCQS(Double Stimulus Continuous Quality-Scale Method) 방법은 overall quality를 평가하는 방법으로 ITU-R BT.710 HDTV 주관적 평가 권고사항에 포함되어 있으며, MPEG, VCEG과 VQEG 등 국제 표준화 기구에서도 DSCQS를 기반으로 한 방법을 사용하고 있다.

ITU-R BT.500.11은 평가할 문제에 따라 실험 평가 방법을 제시하고 있는데, 가장 기본적인 방법이 DSIS와

DSCQS 방법이다. 디지털 비디오 부호화 방식에서는 DSCQS를 기반으로 한 방법들이 널리 사용되고 있다. DSCQS는 전체 성능을 평가하는 데 적합한 방법으로 ITU-R BT.700 HDTV 주관적 화질평가 방법 및 MPEG의 SSMM(Single Stimulus Multi-Media)/DSMM(Double Stimulus Multi-Media)의 기본 방법으로 널리 사용되고 있다. DSCQS는 프리젠테이션에 기준 영상의 자극과 평가할 영상 자극을 2회에 걸쳐 주도록 구성되어 있고, 평가 점수를 각 자극에 대해 각각 매기는 것이 특징이다[2].

나. ITU-T P.910

ITU-T P.910 규격은 TV 화질에 대한 주관적 평가 방법에 관한 권고사항이다. 주관적 화질평가에 사용되는 실험 조건(보는 거리, 스크린의 조도, 실험실 밝기 등)을 기술하고 있으며, 평가하고자 하는 문제에 따라 적용 가능한 여러 가지 실험 평가 방법과 실험결과 분석 방법에 대하여 기술하고 있다. 실험 조건은 ITU-R BT.500.11과 거의 유사하다.

ITU-T P.910은 실험 평가 방법으로 ACR(Absolute Category Rating)과 DCR(Degradation Category Rating)을 제시하고 있다. ACR은 단일 비교방법(Single Stimulus Method: SSM)으로 한번에 한 영상씩 독립적으로 평가 점수를 매기는 것이 특징이다. 평가의 신뢰도를 높이기 위해 같은 영상 자극을 2회에 걸쳐 주도록 구성된다. DCR(Degradation Category Rating) 방법은 이중 비교방법(Double Stimulus Method: DSM)으로 영상 자극을 쌍(pair)으로 한 프리젠테이션에 기준 영상의 자극과 평가할 영상 자극을 두도록 구성되며 두 영상의 비교평가 점수를 하나로 매기는 것이 특징이다[3].

다. 주관적 화질평가의 실험수행 절차

1) 실험환경 조건

ITU-R BT.500.11 권고사항에서는 실험환경 조건에

대해 융통성있게 접근하고 있다. 즉, 디지털 비디오 부호화 분야, 특히 MPEG에서는 위 권고규격이 정한 실험 조건 중 시정거리 조건을 가장 중요시하고 있고, 그 외 사항들에는 융통성을 두고 있다. ITU-R BT.500.11규격은 TV 화질을 평가하기 위한 권고규격으로 실험환경 조건은 일반적인 디지털 비디오 부호화 방식에 적용하기에는 다소 맞지 않는 사항들이 있다. 참고로 SVC의 CfP에서 주관적 화질 실험에서 사용된 거리는 QCIF 해상도의 경우 6H(비디오 화면 세로 사이즈의 6배 거리), CIF는 4H, 4CIF는 3H를 사용하였다. VQEG(Video Quality Experts Group)에서는 CIF 6H~8H, VGA 4H~6H를 권고하고 있다. 한편 ITU-T P.910 규격은 가장 중요한 보는 거리는 1~8H(H는 화면의 높이)로 엄격히 정의되어 있지는 않은데, 이동통신 단말기이기에 상기 거리에 해당되는 한 사용자가 편한 데로 시청하면 된다.

2) 실험 참여자 모집 및 실험 전 교육




실험 참여 인원은 최소 15인 이상을 권고하고 있으며, TV 화질 분야의 비전문가이면서 주관적 화질평가 실험 경험이 없는 사람을 조건으로 한다. 실험 참여자에 대한 직업과 성, 연령 등의 정보를 실험결과와 같이 기록하는 것도 권고하고 있다. 실제로 MPEG과 VCEG과 같은 표준화 기구에서는 주로 회의 중에 주관적 화질평가를 하므로, 전문가들이 실험에 참여하여 평가를 수행하고 있다.

참여자들이 실험에 경험이 없으므로, 실험 전에 충분한 사전 지도를 수행해야 한다. 실험 방법에 대해 설명된 프린트물을 통해서 1차 교육을 수행하고, 실험에 대한 데모와 평가 점수에 대한 내용으로 2차 교육을 수행한다.

3) 세션(Session) 단위 실험 수행

실험의 한 세션은 최대 30분 이내로 제한하고 있으며, 이는 참여자의 피로와 실험 집중도를 고려한 것이다. 첫 세션에서 처음 5번의 평가 프리젠테이션(presentation)

〈표 1〉 실험대상 단말

단말	CPU	OS	디스플레이	동영상
 아이폰4	A4 Chip (1GHz)	iOS4	3.5인치 Retina LCD 640×960	해상도: 720p(1,280×720) HD 프레임: 30fps 동영상 코덱: H.264
 갤럭시S2	삼성 허밍버드 S5PC110 1GHz	안드로이드 2.3 진저브레드	4.7인치 Super AMOLED WVGA (800×480)	해상도: 720p(1,280×720) HD 프레임: 30fps 동영상 코덱: mpeg4, H.264, H.263 등
 아이패드 2	A5 듀얼코어 (900MHz)	iOS4	9.7인치 WVGA (768×1024) TFT LCD	해상도: 720p(1,280×720) 프레임: 30fps 동영상 코덱: H.264, mpeg4

은 잉여(dummy) 프리젠테이션으로 참여자의 의견을 안정화하기 위한 목적으로 실험결과에는 포함하지 않는다. 다음 세션에서는 시작 부분에 3회의 잉여 프리젠테이션만을 사용한다. 실험의 프리젠테이션은 임의의(random) 순서로 구성하여야 한다. 여러 세션을 통해서 동일한 프리젠테이션을 여러 번 테스트할 수도 있으며, 이는 일관성을 체크하기 위해 사용한다.

4) 실험결과 분석

실험결과로는 평균과 95% 신뢰구간을 제출하도록 되어 있으며, 그 외 테스트 영상 및 테스트 조건(양자화 계수, 프로파일, 프레임 수 등)에 대한 자세한 정보, 테스트 영상의 타입(예: yuv, rgb 등) 및 디스플레이 장치에 대한 정보(스크린 크기, 모델번호 등), 비교에 사용된 기준 영상에 대한 정보(비트율, 프로파일 등) 등은 선택적으로 제출된다.

II. 모바일 비디오 영상의 주관적 화질평가

1. 실험환경 및 평가단 구성

영상 화질의 주관적 평가 방법들 중 디지털 영상 품질

을 주관적으로 평가하는 ITU-T P.910 권고사항을 이용하여 실험을 진행하였다. 이 실험에서는 이중비교방법(DSM)과 단일비교방법(SSM)을 모두 진행하였다.

실험환경 조건의 경우, 평가가 이루어지는 실험실 환경조건을 실험실의 밝기, 시청각도, 시청거리 등의 항목을 상세하게 규정하고 있는 ITU-T P.910 규격에서 권고하는 조건을 충족시키도록 구성하였다. 다만 시청거리의 경우 이 규정을 따를 경우 모바일 폰에서의 시청거리와 부합하지 않는 점을 고려해, 이용자들이 가장 자연스럽게 이용할 수 있는 시청거리를 유지하였다. 평가 실험 중에는 외부와의 출입을 차단하여 외부 영향요인을 차단하였다.

대상 단말기는 총 3종으로 아이폰4, 갤럭시S2, 아이패드2를 선정하였다(〈표 1〉 참조).

ITU-T P.910에서는 최소 15명 이상의 일반적 시력을 가진 비전문가로 평가단을 구성할 것을 권고하고 있는데, 이 권고에 따라 20~40대 성인 남녀 19명이 평가단으로 참여하였다. 평가단은 영상 관련 분야에 전문지식이 없는 일반 학생, 회사원, 주부, 자영업, 취업 준비생 등으로 교정시력 0.8~1.2의 비전문가 집단으로 구성하였다.

2. 테스트 영상

본 실험에서는 KT o!leh tv에서 제공하는 프로그램 중에서 화면의 움직임, 밝기, 복잡도 등 다양한 특성이 반영되도록 원본 영상을 선택하였다. 이번 실험에서는 TV에서와는 달리, 기술적인 문제 때문에 원본 영상을 모바일에 적용할 수 없기 때문에 원본에 가까운 영상(16M)으로 압축한 영상을 오리지널 영상으로 하여 실험을 실시하였다. 원래 기준 영상(원본 영상)은 부호화가 안된 영상이어야 하나, 아이폰이나 갤럭시 등 스마트폰에서는 압축이 안된 영상의 플레이가 불가능하므로, 높은 부호화율로 부호화하여 실험을 진행하였다. 평가에 사용한 영상은 (그림 1)과 같으며, 실험 영상의 특징은 <표 2>과 같다.

실험을 위한 테스트 시퀀스의 생성 절차는 다음과 같다. 우선 원본 영상을 토대로 다양한 형태의 부호화된



(그림 1) 주관적 화질평가에 사용한 영상

영상을 생성하고, 5초/3초 길이의 회색영상을 생성하여 부호화된 영상에 붙여 만든 후, 이 영상을 부호화한다. 테스트 시퀀스의 객관적 화질 검증을 위해 PSHR을 계산한다.

3. 실험절차

피평가자들에게 각각 ACR 방법인 session 1은 같은

<표 2> 실험 영상의 특징

Sequence	영상 특성		
	텍스처 특성	움직임 특성	부호화 특성
S1(드라마)	- 복잡도: 중 - 텍스처 변화: 중 ※ 장면 자체의 전환으로 인한 배경의 변화가 있음.	- 움직임: 하 ※ 장면이 변화할 뿐 인물들의 움직임은 거의 없음.	- 300K 정도 됐을 때 인물들의 얼굴 부분에서 약간의 열화 발생
S2(애니메이션)	- 복잡도: 하 - 텍스처 변화: 중 ※ 장면마다 배경의 움직임은 아예 없지만 장면 자체의 변화가 있음.	- 움직임: 하 ※ 배경이 고정되어 있고 인물들의 움직임 또한 거의 없음.	- 500K 정도의 비트레이트에서도 화질 열화가 크게 나타나지 않으나 300K에서는 인물의 움직이는 테두리 영역에서 약간의 열화 발생
S3(영화)	- 복잡도: 중 - 텍스처 변화: 상 ※ 장면의 변화도 있고 카메라의 움직임도 어느 정도 있음	- 움직임: 중 ※ 인물들이 걸어가는 장면에서, 인물의 움직임이 많은 편임.	- 500K 정도의 비트레이트에서도 화질 열화가 크게 나타나지 않음. - 정작 인물들이 움직이는 장면은 어두운 장면이라 열화의 구분이 쉽지 않음.
S4(스포츠)	- 복잡도: 중 - 텍스처 변화: 상 ※ 텍스처가 크게 복잡하지는 않으나 축구 선수들의 빠른 움직임과 카메라 움직임으로 화면의 변화가 존재	- 움직임: 상 ※ 인물들의 움직임이 크고 배경의 움직임도 존재. 카메라의 움직임도 많은 편	- 카메라 이동 시 배경 부분에 열화 발생 - 선수들의 움직임이 많은 팔과 발 부분에 열화 발생 - 500K 이하로 떨어지면 열화가 매우 심해짐.
S5(교육)	- 복잡도: 하 - 텍스처 변화: 하 ※ 배경의 움직임이 거의 없고 단순함, 화면 속의 인물의 텍스처의 변화도 적음	- 움직임: 하 ※ 배경이 고정되어 있고 인물의 움직임도 거의 없음. 카메라의 움직임도 없음.	- 300K의 낮은 비트레이트에서도 열화가 거의 발견되지 않음.

영상을 연속해서 두 번씩 보여주고, DCR 방법인 Session 2, 3는 기준과 비교 영상을 나란히 보여준다. 또한 각 영상의 사이에는 회색 영상이 3초간 나타나고, 마지막 영상이 제시된 후 5초간 회색 영상이 나타난다.

평가자는 ACR 방법에 대해 화질평가 점수를 5단계 (5: excellent, 4: good, 3: fair, 2: poor, 1: bad)로 구분된 MOS(Mean Opinion Score)를 평가표에 직접 표시한다. DCR 방법에 대해 화질평가에 대한 MOS 점수를 5단계로 표시하는데, 품질 차이가 없는 경우(imperceptible) 5점, (품질 차이가 인지되지만 심하지 않은 경우(perceptible but not annoying) 4점, 차이가 좀 나는 경우(slightly annoying) 3점, 차이가 뚜렷한 경우(annoying) 2점, 차이가 매우 많은 경우(very annoying) 1점으로 평가한다.

III. 모바일 비디오 영상의 주관적 화질평가 결과

1. 주관적 화질평가 결과: 아이폰4

가. Session 1(ACR) 결과

움직임이 많은 S4(스포츠)를 제외한 모든 영상이 500kbps 이상에서는 MOS 평균이 4.3 이상이고 원본과의 차이도 최대 0.2 이하(원본 영상 대비 5% 이내)로 주관적 화질에 큰 차이가 없는 것으로 분석된다. 드라마, 교육 영상, 애니메이션 등과 같이 프레임 간 차이가 크지 않은 경우, 비트율 감소에 따른 품질 차이가 그리 크지 않은 것으로 나타나고 있다. 그러나 스포츠와 같이 움직임이 크고 장면 변화가 잦은 경우, 비트율 감소에 따른 품질에 대한 인지가 급격히 악화되고 있다.

S4(스포츠)를 제외한 모든 영상은 300K에서도 적정 수준의 품질 수준을 유지하는 것으로 분석된다. 300kbps에서는 모든 영상이 원본 영상 대비 5% 이상의 화질 저하가 있는 것으로 인지하고 있으며, 특히 S4(스

포츠)의 경우 품질 수준에 대한 인지가 열악한 것으로 나타난다. S4(스포츠)의 경우 인물 및 배경의 움직임이 크고 카메라의 움직임도 많아 화질열화가 눈에 띄게 증가하고 있다. 그러나 배경의 고정되어 있고 인물들의 움직임 또한 거의 없는 S2(애니메이션)이나 S5(교육)의 경우 300K 영상에서도 높은 주관적 품질 수준을 나타내고 있다.

나. Session 2(DCR) 결과

S4(스포츠)를 제외한 모든 영상이 Org-1M, Org-700K, Org-500K 간에 품질 차이를 인식하지 못하는 것으로 나타나고 있다. 특히 S2(애니메이션)의 경우 Org-300K도 비트율 감소에 따른 품질 차이가 없는 것으로 인지되고 있으며, 그 외 S1(드라마), S3(영화), S5(교육)도 Org-300K가 품질 차이가 심하지 않은 것으로 나타났다. S4(스포츠)의 경우도, 500K 영상까지도 원본 영상에 비해 품질 차이가 그리 크지 않으며, 다만 300K 영상인 경우에는 원본 영상에 비해 급격한 품질 저하가 인지되고 있는 것으로 나타났다.

다. Session 3(DCR) 결과

S4(스포츠)를 제외한 모든 영상은 1M-300K 조차도 품질 차이를 거의 인지하지 못하는 것으로 나타났다. 모든 테스트 영상을 쌍으로 비교(pairwise comparison)하는 경우 MOS 평균이 4.3~4.7점대를 유지하여 영상 간 품질 차이를 거의 인지하지 못하는 것으로 나타났다. S4(스포츠)의 경우, 300K와 비교하는 경우 MOS 평균이 2.6~2.7점대로 품질 차이가 뚜렷한 것으로 나타나고 있으나, 1M-500K 이상인 경우는 4.3 이상으로 그렇게 차이가 나지 않는 것으로 나타났다.

라. 결과의 종합 및 시사점

Session 1(ACR)에서는 움직임이 많은 S4(스포츠)를

제외한 모든 영상이 500kbps 이상에서는 MOS 평균이 4.3 이상이고 원본과의 차이도 최대 0.2 이하(원본 영상 대비 5% 이내)로 주관적 화질에 큰 차이가 없는 것으로 분석된다. 특히 S4(스포츠)를 제외한 모든 영상은 300K에서도 적정 수준의 품질 수준을 유지하는 것으로 분석된다. Session 2(DCR)에서는 S4(스포츠)를 제외한 모든 영상이 Org-1M, Org-700K, Org-500K 간에 품질 차이를 인식하지 못하는 것으로 나타났다. Session 3(DCR)에서는 S4(스포츠)를 제외한 모든 영상은 1M-300K 조차도 품질 차이를 거의 인지하지 못하는 것으로 나타나고 있으며, S4(스포츠)의 경우, 300K와 비교하는 경우 MOS 평균이 2.6~2.7점대로 품질 차이가 뚜렷한 것으로 나타나고 있으나, 1M-500K 이상인 경우는 4.3 이상으로 그렇게 차이가 나지 않는 것으로 나타났다.

대표적인 테스트 영상에 대해 ACR 방법과 DCR 방법을 적용해서 분석한 결과는 움직임이 상이한 경우를 포함해서 최악의 경우를 고려하면 500kbps 이상으로 부호화할 경우 원본의 영상과 뚜렷한 차이를 인지하지 못하는 것으로 최종 판단할 수 있다.

2. 주관적 화질평가 결과: 갤럭시S2

가. Session 1(ACR) 결과

움직임이 많은 S4(스포츠)를 제외한 모든 영상이 500kbps 이상에서는 MOS 평균이 4.3 이상이고 원본과의 차이도 최대 0.2 이하(원본 영상 대비 5% 이내)로 주관적 화질에 큰 차이가 없는 것으로 분석된다. 드라마, 애니메이션 등과 같이 프레임 간 차이가 크지 않은 경우, 비트율 감소에 따른 품질 차이가 그리 크지 않은 것으로 나타나고 있으며, 특히, S5(교육)의 경우 1M, 700K, 500K 모두 원본 영상보다 더 높은 품질로 인지되고 있다. 그러나 스포츠와 같이 움직임이 크고 장면 변화가 잦은 경우, 비트율 감소에 따른 품질에 대한 인

지가 급격히 악화되고 있다.

S2(애니메이션)를 제외한 모든 영상은 300K에서 2~3점대 평가를 받아 품질에 대한 인지가 열악한 것으로 분석된다. 300kbps에서는 모든 영상이 원본에 비해 10% 이상 품질 수준이 낮은 것으로 나타나고 있으며, 특히 S4(스포츠)의 경우 원본 화질과의 차이가 큰 것으로 나타났다. S4(스포츠)의 경우 인물 및 배경의 움직임이 크고 카메라의 움직임도 많아 화질열화가 눈에 띄게 증가하고 있다. 그러나 배경의 고정되어 있고 인물들의 움직임 또한 거의 없는 S2(애니메이션)이나 S5(교육)의 경우 300K 영상의 품질 수준은 원본 영상 대비 10% 정도 감소한 것으로 나타났다.

나. Session 2(DCR) 결과

S4(스포츠)를 제외한 모든 영상이 Org-1M, Org-700K, Org-500K 간에 품질 차이를 인식하지 못하는 것으로 나타나고 있고, Org-300K도 품질 차이가 그리 크지 않은 것으로 나타났다. 특히 S2(애니메이션)의 경우 Org-300K도 비트율 감소에 따른 품질 차이가 없는 것으로 인지되고 있으며, 그 외 S1(드라마), S3(영화), S5(교육)도 Org-300K가 품질 차이가 심하지 않은 것으로 나타났다. S4(스포츠)의 경우도, 500K 영상까지도 원본 영상에 비해 품질 차이가 그리 크지 않으며, 다만 300K 영상인 경우에는 원본 영상에 비해 급격한 품질 저하가 인지되고 있는 것으로 나타났다.

다. Session 3(DCR) 결과

S4(스포츠)를 제외한 모든 영상은 1M-300K 조차도 품질 차이를 거의 인지하지 못하는 것으로 나타났다. 모든 테스트 영상을 쌍으로 비교하는 경우 MOS 평균이 4.1~4.7점대를 유지하여 영상 간 품질 차이를 거의 인지하지 못하는 것으로 나타났다. S4(스포츠)의 경우, 300K와 비교하는 경우 MOS 평균이 2.4~2.7점대로 품

질 차이가 뚜렷한 것으로 나타나고 있으나, 1M-500K 이상인 경우는 4.4 이상으로 그렇게 차이가 나지 않는 것으로 나타났다.

라. 결과의 종합 및 시사점

Session 1(ACR)에서는 움직임이 많은 S4(스포츠)를 제외한 모든 영상이 500kbps 이상에서는 MOS 평균이 4.3 이상이고 원본과의 차이도 최대 0.2 이하(원본 영상 대비 5% 이내)로 주관적 화질에 큰 차이가 없는 것으로 분석된다. S2(애니메이션)를 제외한 모든 영상은 300K에서 2~3점대 평가를 받아 품질에 대한 인지가 열악한 것으로 분석된다. Session 2(DCR)에서는 S4(스포츠)를 제외한 모든 영상이 Org-1M, Org-700K, Org-500K 간에 품질 차이를 인식하지 못하는 것으로 나타나고 있고, Org-300K도 품질 차이가 그리 크지 않은 것으로 나타났다. Session 3(DCR)에서는 S4(스포츠)를 제외한 모든 영상은 1M-300K 초차도 품질 차이를 거의 인지하지 못하는 것으로 나타났는데, S4(스포츠)의 경우, 300K와 비교하는 경우 MOS 평균이 2.4~2.7점대로 품질 차이가 뚜렷한 것으로 나타나고 있으나, 1M-500K 이상인 경우는 4.4 이상으로 그렇게 차이가 나지 않는 것으로 나타났다.

대표적인 테스트 영상에 대해 ACR 방법과 DCR 방법을 적용해서 분석한 결과는 움직임이 상이한 경우를 포함해서 최악의 경우를 고려하면 500kbps 이상으로 부호화할 경우 원본의 영상과 뚜렷한 차이를 인지하지 못하는 것으로 최종 판단할 수 있다.

3. 주관적 화질평가 결과: 아이패드2

가. Session 1(ACR) 결과

움직임이 많은 S4(스포츠)를 제외한 모든 영상이 500kbps 이상에서는 MOS 평균이 4.0 이상이고 원본과

의 차이도 그리 크지 않아 주관적 화질에 큰 차이가 없는 것으로 분석된다. 드라마, 애니메이션, 교육 등과 같이 프레임 간 차이가 크지 않은 경우, 비트율 감소에 따른 품질 차이가 그리 크지 않은 것으로 나타난다. 특히, S2(애니메이션)과 S5(교육)의 경우 1M, 700K 영상이 원본 영상보다 더 높은 품질로 인지되고 있다. 그러나 스포츠와 같이 움직임이 크고 장면 변화가 잦은 경우, 원본 영상도 품질 수준에 대해 상대적으로 낮은 평가를 받고 있으며 비트율 감소에 따른 품질에 대한 인지가 급격히 악화되고 있다.

모든 영상은 300K에서 2~3점대 평가를 받아 품질에 대한 인지가 열악한 것으로 분석된다. 300kbps에서는 모든 영상이 원본에 비해 10% 이상 품질 수준이 낮은 것으로 나타나고 있으며, 특히 S4(스포츠)의 경우 원본 화질과의 차이가 큰 것으로 나타났다.

나. Session 2(DCR) 결과

S4(스포츠)를 제외한 모든 영상이 Org-1M, Org-700K, Org-500K 간에 품질 차이를 인식하지 못하는 것으로 나타나고 있고, Org-300K도 품질 차이가 그리 크지 않은 것으로 나타났다. 특히 S2(애니메이션)의 경우 Org-300K도 비트율 감소에 따른 품질 차이가 없는 것으로 인지되고 있으며, 그 외 S1(드라마), S2(애니메이션), S5(교육)도 Org-300K가 품질 차이가 심하지 않은 것으로 나타났다. S4(스포츠)의 경우도, Org-500K가 Org-700K보다 품질 격차가 더 작은 것으로 나타나는 등 품질 수준에 대한 정확한 인지가 곤란한 것으로 나타나고 있는데, 다만 300K 영상인 경우에는 원본 영상에 비해 급격한 품질 저하가 인지되고 있는 것으로 나타났다.

다. Session 3(DCR) 결과

S4(스포츠)를 제외하고 1M-700K, 1M-500K, 700K-

500K의 품질 차이는 없는 것으로 나타나고 있으며, 300K와 비교 시에는 그 차이가 심하지는 않은 것으로 나타났다. S4(스포츠)의 경우, 1M-500K와 700K-500K의 품질 차이가 미미한 것으로 나타났다. 다만, 300K와 비교 시에는 차이가 뚜렷할 정도로 품질 격차가 심한 것으로 나타났다.

라. 결과의 종합 및 시사점

Session 1(ACR)에서는 움직임이 많은 S4(스포츠)를 제외한 모든 영상이 500kbps 이상에서는 MOS 평균이 4.0 이상이고 원본과의 차이도 그리 크지 않아 주관적 화질에 큰 차이가 없는 것으로 분석된다. 다만, 300K에서는 2~3점대 평가를 받아 품질에 대한 인지가 열악한 것으로 분석된다. Session 2(DCR)에서는 S4(스포츠)를 제외한 모든 영상이 Org-1M, Org-700K, Org-500K 간에 품질 차이를 인식하지 못하는 것으로 나타나고 있고, Org-300K도 품질 차이가 그리 크지 않은 것으로

나타났다. Session 3(DCR)에서는 S4(스포츠)를 제외하고 1M-700K, 1M-500K, 700K-500K의 품질 차이는 없는 것으로 나타나고 있으며, 300K와 비교 시에는 그 차이가 심하지는 않은 것으로 나타났다. S4(스포츠)의 경우, 1M-500K와 700K-500K의 품질 차이가 미미한 것으로 나타나고 있으나, 300K와 비교 시에는 차이가 뚜렷할 정도로 품질 격차가 심한 것으로 나타났다.

대표적인 테스트 영상에 대해 ACR 방법과 DCR 방법을 적용해서 분석한 결과는 움직임이 상이한 경우를 포함해서 최악의 경우를 고려하면 500kbps 이상으로 부호화할 경우 원본의 영상과 뚜렷한 차이를 인지하지 못하는 것으로 최종 판단할 수 있다. 다만, S4(스포츠)의 경우 주관적인 품질 수준에 대해서는 높은 평가를 받고 있지 못하지만, 쌍으로 비교 시 영상 간 품질 차이가 그리 크지 않은 것으로 나타났다. S4(스포츠)의 경우 텍스처가 크게 복잡하지는 않으나 축구 선수들의 빠른 움직임과 카메라 움직임으로 화면의 변화가 존재하여 카메라 이동 시 배경 부분에 열화가 발생하고 있고, 선수들

〈표 3〉 영상별 화질평가 결과: Session 1

S1(드라마)	Org	1M	700K	500K	300K
아이폰 4	4.51	4.49	4.43	4.33	4.05
갤럭시 S2	4.44	4.49	4.29	4.24	3.85
아이패드 2	4.39	4.25	4.12	3.94	3.68
S2(애니메이션)	Org	1M	700K	500K	300K
아이폰 4	4.67	4.68	4.56	4.56	4.42
갤럭시 S2	4.71	4.67	4.58	4.49	4.24
아이패드 2	4.44	4.55	4.45	4.27	3.80
S3(영화)	Org	1M	700K	500K	300K
아이폰 4	4.55	4.36	4.45	4.39	4.12
갤럭시 S2	4.71	4.67	4.58	4.49	4.24
아이패드 2	4.44	4.55	4.45	4.27	3.80
S4(스포츠)	Org	1M	700K	500K	300K
아이폰 4	4.43	4.26	4.15	3.94	2.73
갤럭시 S2	4.22	4.02	3.87	3.66	2.48
아이패드 2	3.71	3.78	3.47	3.13	2.18
S5(교육)	Org	1M	700K	500K	300K
아이폰 4	4.58	4.57	4.52	4.45	4.19
갤럭시 S2	4.33	4.45	4.42	4.45	3.86
아이패드 2	4.27	4.36	4.40	4.12	3.69

〈표 4〉 영상별 화질평가 결과: Session 2

S1(드라마)	Org-1M	Org-700K	Org-500K	Org-300K
아이폰 4	4.66	4.62	4.51	4.31
갤럭시 S2	4.67	4.63	4.27	4.16
아이패드 2	4.57	4.61	4.41	4.05
S2(애니메이션)	Org-1M	Org-700K	Org-500K	Org-300K
아이폰 4	4.66	4.62	4.51	4.31
갤럭시 S2	4.67	4.63	4.27	4.16
아이패드 2	4.57	4.61	4.41	4.05
S3(영화)	Org-1M	Org-700K	Org-500K	Org-300K
아이폰 4	4.69	4.52	4.52	4.41
갤럭시 S2	4.72	4.66	4.55	4.24
아이패드 2	4.56	4.52	4.54	3.99
S4(스포츠)	Org-1M	Org-700K	Org-500K	Org-300K
아이폰 4	4.57	4.27	4.13	2.67
갤럭시 S2	4.29	4.32	4.17	2.79
아이패드 2	4.25	3.75	4.02	2.6
S5(교육)	Org-1M	Org-700K	Org-500K	Org-300K
아이폰 4	4.57	4.44	4.6	4.34
갤럭시 S2	4.63	4.56	4.53	4.17
아이패드 2	4.52	4.52	4.25	4.01

〈표 5〉 영상별 화질평가 결과: Session 3

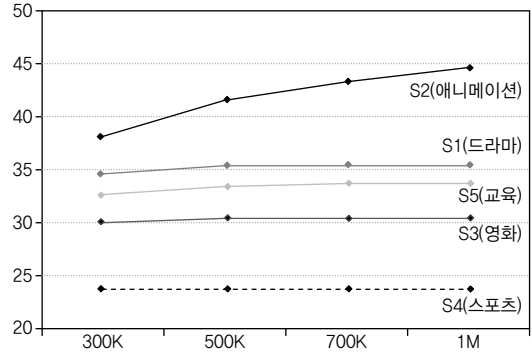
S1(드라마)	1M-700K	1M-500K	1M-300K	700K-500K	700K-300K	500K-300K
아이폰 4	4.59	4.46	4.46	4.58	4.48	4.37
갤럭시 S2	4.62	4.56	4.13	4.63	4.2	4.39
아이패드 2	4.62	4.4	4.11	4.49	4.2	4.12
S2(애니메이션)	1M-700K	1M-500K	1M-300K	700K-500K	700K-300K	500K-300K
아이폰 4	4.71	4.69	4.59	4.61	4.43	4.39
갤럭시 S2	4.66	4.66	4.4	4.62	4.34	4.4
아이패드 2	4.57	4.54	4.09	4.59	4.09	4.13
S3(영화)	1M-700K	1M-500K	1M-300K	700K-500K	700K-300K	500K-300K
아이폰 4	4.62	4.48	4.48	4.6	4.51	4.45
갤럭시 S2	4.64	4.59	4.17	4.59	4.21	4.26
아이패드 2	4.56	4.48	4.13	4.47	4.09	3.88
S4(스포츠)	1M-700K	1M-500K	1M-300K	700K-500K	700K-300K	500K-300K
아이폰 4	4.55	4.27	2.77	4.47	2.63	2.78
갤럭시 S2	4.44	4.43	2.48	4.24	2.62	2.67
아이패드 2	3.8	4	2.41	4.14	2.38	2.64
S5(교육)	1M-700K	1M-500K	1M-300K	700K-500K	700K-300K	500K-300K
아이폰 4	4.63	4.66	4.46	4.67	4.35	4.39
갤럭시 S2	4.62	4.59	4.04	4.59	4.2	4.44
아이패드 2	4.55	4.41	3.87	4.49	3.8	3.99

의 움직임이 많은 팔과 발 부분에 열화가 발생하고 있다.

이상 아이폰4, 갤럭시S2, 아이패드2의 session 1~3의 화질평가 결과를 정리하면, 〈표 3〉, 〈표 4〉, 〈표 5〉와 같다.

4. 테스트 영상에 대한 객관적 화질평가

S4(스포츠)를 제외한 모든 테스트 영상은 품질 수준에 관계없이 PSNR 값이 30dB 이상으로 나타나고 있어, 스마트폰에서 300K 이상의 영상조차도 높은 수준의 영상 품질을 제공할 수 있다고 평가할 수 있다. 통상적으로 30db 이상의 영상은 인간의 시각 특성상 화질 저하로 판단하기 쉽지 않은 영상으로, 원본 영상과 테스트 영상간 품질 격차가 거의 없는 것으로 판단할 수 있다. 일반적으로 30dB가 넘으면 두 영상의 차이를 눈으



(그림 2) PSNR 값 비교

로 구분하기 어려운 것으로 알려져 있고, 특히 무선전송에서 JPEG 이미지의 경우 받아들여질 수 있는 수준은 약 20~25dB 정도로 간주되고 있다[4],[5].

S3(영화)와 S4(교육)의 경우 PSNR이 30~33dB 수준이지만 TV와 스마트폰의 화면 크기를 고려할 때, 이용자들이 시각적으로 느끼는 수준은 TV보다 오히려 상대적으로 우위에 있다고 할 수 있다. 특히, 30인치 TV를 기준으로 하는 경우(HD급 36~37dB, VGA급 32~33dB)와 상대적으로 비교해 보면, 아이폰4(3.5인치) 및 갤럭시 S2(4. 인치)의 화면 크기가 TV에 비해 거의 10배 정도 작다는 점을 감안하면 상당히 높은 수준의 화질을 제공할 수 있다고 평가할 수 있다(그림 2 참조).

다만, 일반적으로 비트 레이트가 높을수록 PSNR이 높게 나오지만 S4(스포츠)의 경우에는 그렇지 않은데, 그 이유는 2가지 정도로 추정된다. 우선 PSNR-U와 PSNR-V에 상대적으로 높은 비트(bit)가 할당된 경우이며, 두 번째 이유로는 아주 복잡한 시퀀스의 경우 적절한 레이트(rate)가 할당되지 않을 경우 일정 비트 레이트 이하에서는 레이트 컨트롤(rate control)이 제대로 되지 않을 수 있다.

IV. 결론 및 시사점

PSNR을 이용한 객관적 화질평가 결과는 주관적 화질

평가에 참여한 피실험자들이 눈으로 인식하는 결과 값의 차이에서와 유사한 패턴을 보이고 있다. 즉, 객관적 화질평가 결과(PSNR 값)의 차이의 정도가 주관적 화질 평가에서도 피실험자들이 느끼는 화질의 수준과 유사한 패턴을 보이고 있어, 주관적 화질평가 결과의 신뢰도가 높은 것으로 판단할 수 있다.

따라서 최악의 경우를 고려하더라도 스마트폰 등의 무선단말기에서는 500kbps 이상으로 부호화할 경우 원본의 영상과 뚜렷한 차이를 인지하지 못하는 것으로 최종 판단할 수 있다. 원본 영상 대비 테스트 영상의 주관적/객관적 화질평가 결과는 시각적으로 원본 영상과 테스트 영상 간의 품질 차이를 인지하는 못하는 수준인 것으로 나타났다. 움직임이 많은 S4(스포츠)를 제외한 모든 영상이 500kbps 이상에서는 MOS 평균이 4.3 이상이고 원본과의 차이도 최대 0.2 이하(원본 영상 대비 5% 이내)로 주관적 화질에 큰 차이가 없는 것으로 나타나고, 이는 객관적 화질평가도 30dB 이상의 PSNR 결과를 나타냈다.

Florence Agboma & Antonio Liotta(2012)는 실증분석을 통해 화면 전환이 많은 콘텐츠의 경우 비트 레이트 보다는 프레임 레이트(frame rate)가 QoE에 더 큰 영향을 미친다는 결과를 제시한 바 있다[6]. Florence Agboma & Antonio Liotta는 액션 영화의 경우 최저 수준은 128bps/12FPS, 코미디 영상의 경우 128bps/10FPS 제시하였다. 한편 Quan Huynh-Thu & Mohammed Ghanbari(2008)은 실증분석을 통해 frame rate에 따른 MOS(Mean Opinion Score) 수준을 평가하였는데, 일반적으로 4 이상을 고품질로 분류하였다[7]. Alcatel Lucent(2010)도, 모바일 기기는 일반적으로 화면 크기가 유선에 비해 더 작으므로 비디오 스트림당 무선 네트워크에서 필요한 대역폭은 유선보다 7~14배 더 적다고 지적하며, HD 비디오 스트리밍에 578kbps면 충분하다고 밝힌바 있다[8].

이와 같은 점을 고려할 때 입력 영상 및 출력 영상 그

리고 대역폭 관계를 고려하여 타겟(target) 비트 레이트를 적절하게 설정하는 것이 필요하다. 비트레이트 증가에 따른 사용자들의 한계효용은 체감하지만, 비트레이트가 증가할수록 그에 따른 네트워크 비용은 체증하기 때문에 적절한 수준의 타겟 비트레이트 설정이 사회적으로 바람직할 수 있다.

용어해설

PSNR 주로 영상 또는 동영상 손실 압축에서 화질 손실정보를 평가할 때 사용되는데, 두 영상이나 음성의 차이를 숫자로 표현한 것

주관적 화질평가 주평가자가 직접 동영상을 보고 화질을 평가하는 방법으로서 인간의 화질인지 특성을 반영할 수 있는 가장 정확한 방법으로 알려져 있음.

ITU-T BT,500.11 규격 TV 화질에 대한 주관적 평가 방법에 관한 권고사항으로 주관적 화질평가에 사용되는 실험 환경 조건(디스플레이 밝기, 평가실의 조건, 평가자의 위치, 스크린과의 거리 및 각도 등)을 기술.

약어 정리

ACR	Absolute Category Rating
DCR	Degradation Category Rating
DSCQS	Double Stimulus Continuous Quality-Scale Method
DSIS	Double Stimulus Impairment Scale
DSM	Double Stimulus Method
DSMM	Double Stimulus Multi-Media
PSNR	Peak Signal-to-Noise Ratio
SSM	Single Stimulus Method
SSMM	Single Stimulus Multi-Media

참고문헌

- [1] 박인경 외 “4K-UHD 비디오 시청환경 특성분석을 위한 주관적 화질평가 분석,” 방송공학회논문지, vol. 15, no. 4, 2010, pp. 563-581.
- [2] ITU-R BT-500.11, “Methodology for the Subjective Assessment of the Quality of Television Pictures,” International Telecommunications Union, 2002.
- [3] ITU-T P.910, “Subjective Video Quality Assessment Methods for Multimedia Applications,” International

Telecommunications Union, 1999.

- [4] N. Thomos and N.V. Boulgouris, and M.G. Strintzis, "Optimized Transmission of JPEG2000 Streams over Wireless Channels," *IEEE Trans. Image Process.*, vol. 15, no. 1, 2006.
- [5] X. Li, and J. C, "Robust Transmission of JPEG2000 Encoded Images over Packet Loss Channels," *ICME*, 2007, pp. 947-950
- [6] F. Agboma and A. Liotta, "Quality of Experience Management in Mobile Content Delivery Systems," *Telecommun. Syst.*, vol. 49, no. 1, 2012, pp. 85-98.
- [7] Q. Huynh-Thu and M. Ghanbari, "Temporal Aspect of Perceived Quality in Mobile Video Broadcasting," *IEEE Tran. Broadcast.*, vol. 54, no. 3, Sept. 2008.
- [8] Alcatel Lucent, "Analysis of the Impact of Traffic Growth on the Evolution of Internet Access," 2010.