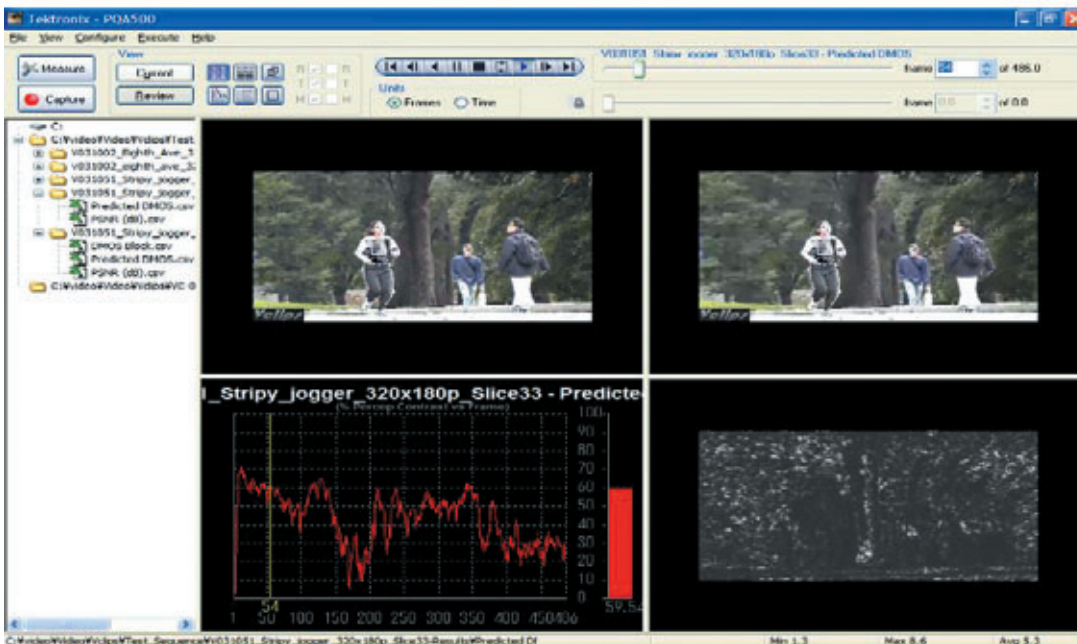
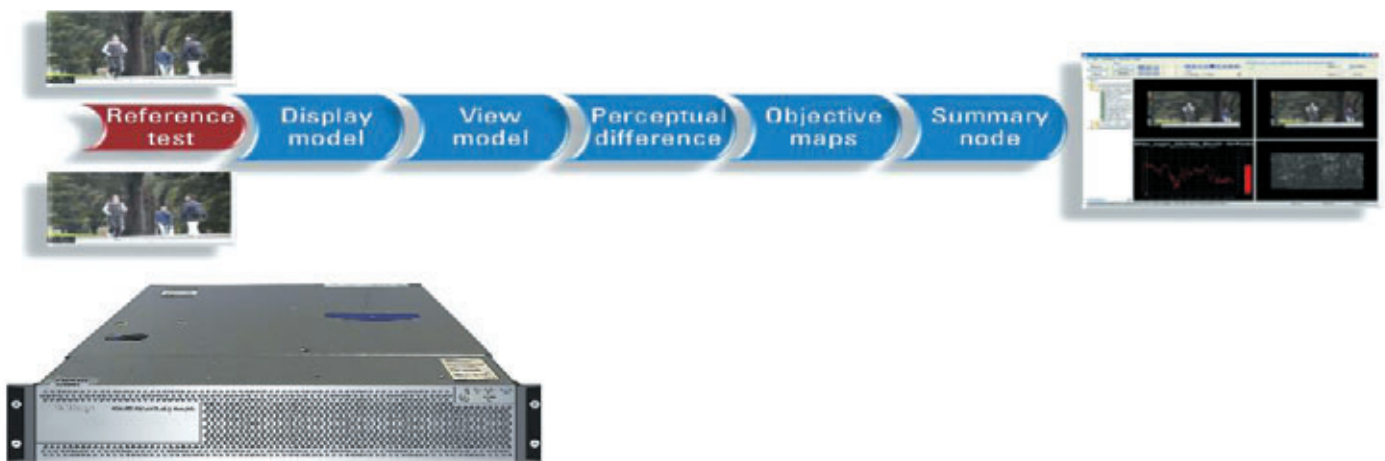


+ 정동호 · 한국텍트로닉스 Application Engineer

Tektronix PQA500 – Picture Quality Analysis System



Tektronix

최근의 디지털 방송환경에서 방송사의 제작환경과 전송과정이 아날로그에서 SD, 그리고 HD로 빠르게 전환되고 있으며, 다양한 이미지 압축기술을 적용하여 시청자들에게 콘텐츠를 전달하고 있다. 이러한 변화의 궁극적인 목적은 더 나은 해상도, 더 좋은 품질의 영상을 전달하고자 함일 것이다. 이를 위해서 많은 엔코더/디코더를 포함한 방송장비 제조업체들은 최적의 알고리즘, 코덱을 찾기 위해 노력하고 있고, 방송사 및 네트워크 사업자는 정해진 대역폭 내에서 최상품질의 비디오를 전송할 수 있는 방법, 콘텐츠 제작 업체들은 편집, 포맷변환, 특수효과 등의 작업 중에 화질의 열화가 생기지 않도록 하는 방법을 찾아내고자 노력한다.

본 글에서는 이러한 노력에 도움이 될 수 있는 Tektronix의 PQA500이라는 화질 분석 장비에 대한 소개와 더불어 화질분석과 관련된 몇 가지 고려사항들에 대하여 살펴보도록 한다.

주관적인 화질 평가 방법

시각적인 화질이란 본래 주관적이고 많은 요소들에 의해 영향을 받기 때문에 정확하게 화질을 측정한다는 것은 쉬운 일이 아니다. 시각적인 화질에 대한 견해는 공간적인 사실성(화면이 얼마나 깨끗하게 보이는가 또는 눈에 띄는 왜곡 현상이 있는가)과 시간적인 사실성(움직임이 자연스럽게 부드러운가)에 영향을 받게 된다. 또한, 화질에 대한 보는 사람의 견해는 주위의 환경, 보는 사람의 마음 자세, 시각적인 주의력(관찰자는 화면의 모든 것을 동시에 관찰하기 보다는 일부분만을 응시함으로써 장면을 인지한다) 등이 복합적으로 작용하여 결정되기 때문에 이를 정확하게 수학적으로 측정하는 데에는 많은 어려움이 있다.

주관적인 화질 평가에 관한 여러 가지 테스트 방법들이 ITU-R BT.500-11에 정의되어 있다. 표준에서 일반적으로 사용되는 방법은 DSCQS(Double Stimulus Continuous Quality Scale) 방법으로, 평가자에게 A와 B라는 한 쌍의 이미지 또는 짧은 동영상 을 순차적으로 보여주고 각 이미지 또는 동영상의 화질에 대한 점수를 “매우 좋음(Excellent)”, “좋음(Good)”, “보통(Fair)”, “나쁨(Poor)”, “매우 나쁨(Bad)”까지의 5단계로 평가하게 하는 방법이다.

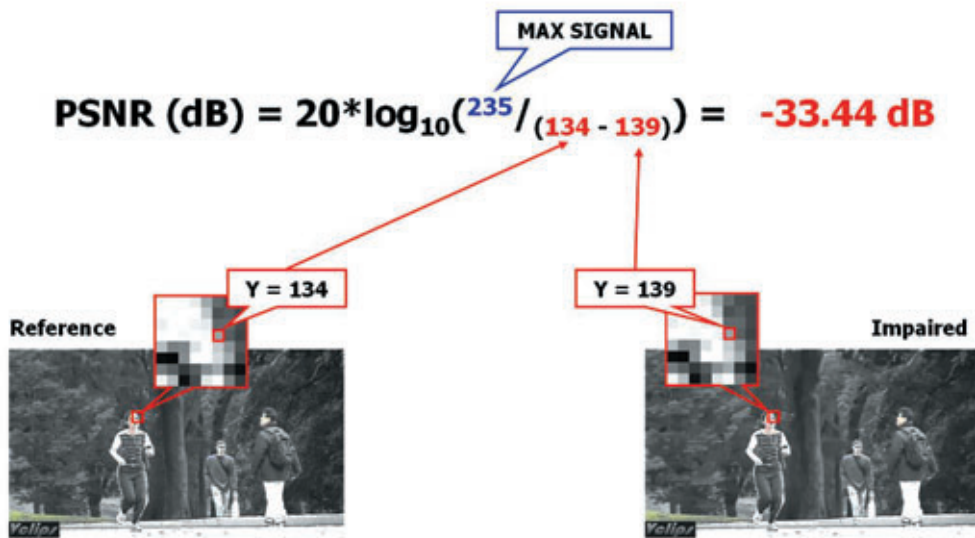
A와 B 중 하나는 원본 영상이고, 다른 하나는 테스트를 위해 가공된 영상이다. 평가자가 어떤 것이 원본 영상이고 어떤 것이 손상된 영상인지 알 수 없게 하기 위해 원본 영상과 손상된 영상의 순서는 테스트 과정에서 무작위로 정해진다. DSCQS와 같은 테스트들은 주관적인 시각 화질 측정방법을 현실화한 것으로 받아들여진다.

그러나, 이러한 형태의 테스트는 실질적인 문제를 가지고 있는데 평가자와 테스트되는 동영상에 따라서 그 결과가 현격한 차이를 나타낼 수 있다. 이는 많은 수의 평가자에 의해서 보완될 수 있지만 이러한 요인들로 인하여 많은 비용과 시간이 소요된다.

객관적인 화질 평가 방법

주관적인 화질 측정방법의 복잡성과 비용으로 인해, 알고리즘을 이용하여 자동으로 화질을 측정하는 방법이 관심을 끌게 됐다.

가장 널리 사용되는 방법은 MSE(Mean Squared Error)를 기반으로 계산된 PSNR(Peak Signal to Noise Ratio)이다.



$$\text{MSE}_{\text{seq}} = \sqrt{\frac{1}{MN_v N_h} \sum_{n=0}^{M-1} \sum_{j=0}^{N_v-1} \sum_{i=0}^{N_h-1} [Y_{\text{ref}}(i, j, f_n) - Y_{\text{test}}(i, j, f_n)]^2}$$

$$\text{PSNR}_{\text{seq}} = 20 \log_{10} \left[\frac{Y_{\text{peak}} = 235}{\sqrt{\frac{1}{MN_v N_h} \sum_{n=0}^{M-1} \sum_{j=0}^{N_v-1} \sum_{i=0}^{N_h-1} [Y_{\text{ref}}(i, j, f_n) - Y_{\text{test}}(i, j, f_n)]^2}} \right]$$

[MSE 및 PSNR 측정]

PSNR은 로그 단위로 측정되며 이미지에 존재할 수 있는 최대 샘플 개수의 제곱[일반적으로는 8비트 비디오의 경우 255의 숫자가 사용되지만, ITU Committee T1, Video Quality Expert Group, T1.TR.74-2001 Objective Video Quality Measurement Using a Peak-Signal-to-Noise Ratio(PSNR) Full Reference Technique에서는 Peak White를 235로 정함]과 원본 이미지와 손상된 이미지 또는 비디오 프레임 사이의 평균제곱오차(MSE)의 비율에 의해 결정된다. 따라서, 높은 MSE 값은 원본 영상과의 차이가 많다는 것을 의미하고 이는 낮은 PSNR로 측정되기 때문에 낮은 품질의 영상으로 판단한다. 한편, 낮은 MSE 값은 원본 영상과의 적은 차이로 간주되고 이는 높은 PSNR로 측정되기 때문에 좋은 품질의 영상으로 판단할 수 있다. 하지만, 다음 영상을 보면 PSNR에 따른 평가가 실제 주관적인 화질과 꼭 연관이 있지 않다는 것을 알 수 있다.



MSE=27.10



MSE=21.26

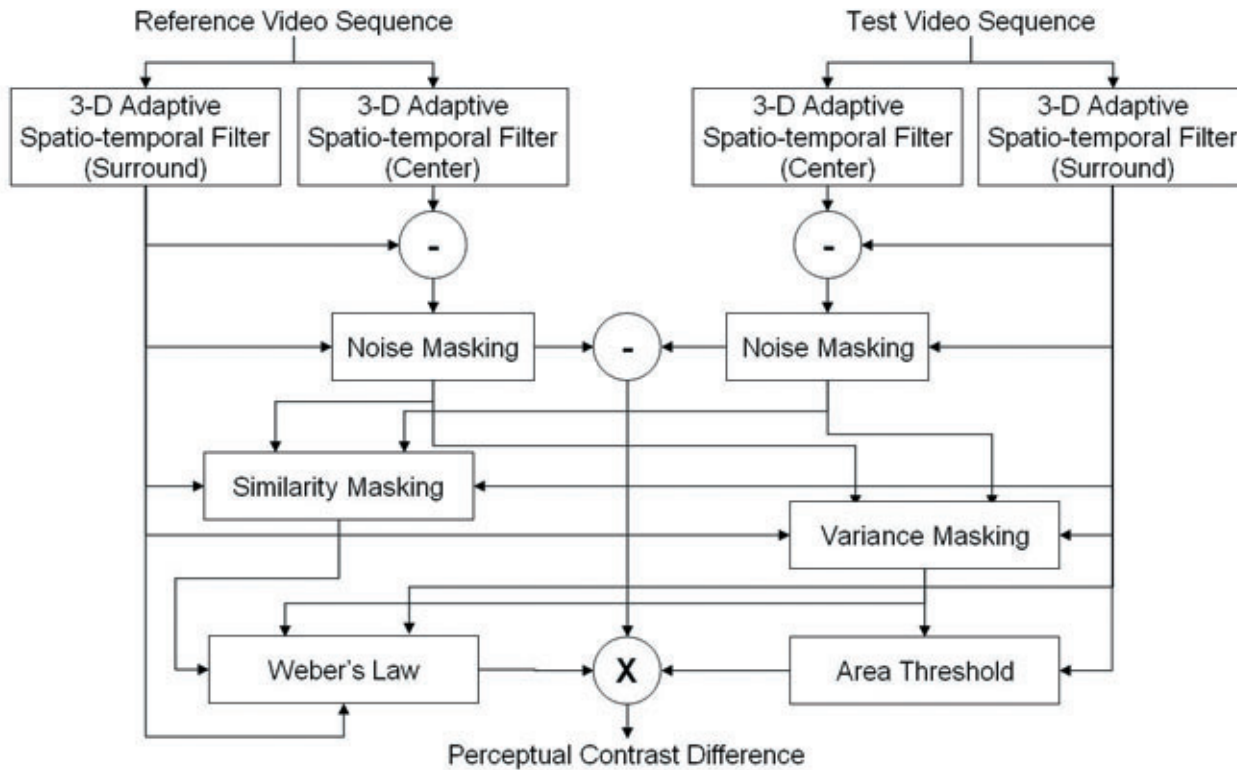
[MSE/PSNR ≠ 실제 화질]

그림에서 실제로 사람의 눈에는 왼쪽 그림의 화질이 오른쪽 것에 비해서 좋아 보이지만, 오른쪽 그림의 MSE 값이 낮게 측정되었기 때문에 MSE로 측정된 값으로 판단할 때는 오른쪽의 그림을 좋은 화질로 판단할 수 있는 오해가 있다. 이는 실제 인간의 시각 시스템이 높은 공간적 주파수(high spatial frequency)보다 낮은 공간적 주파수(low spatial frequency)에 더욱 민감하게 반응하기 때문이다.

Human Vision Model

앞서 언급한 것처럼 PSNR과 같은 객관적인 화질 평가 방법의 한계 때문에 실제 인간의 시각 모델을 적용한 새로운 방법이 적용된 화질 평가 방법이 필요하게 됐다. Human Vision Model에서는 단순히 원본 영상과 테스트 영상 간의 픽셀 값의 차이를 구하는 것이 아니라 실제 인간이 인지할 수 있는 차이 Perceptual Contrast Difference(PCD)를 구하여 이를 객관적인 화질을 평가할 수 있는 수치로 이용한다.

다음은 PCD를 구하기 위해서 PQA500에 적용된 방법을 블록 다이어그램으로 나타낸 것이다.



[Human Vision Model Block Diagram]

인간 시각 시스템에서는 공간적, 시간적 요소에 대한 Contrast Sensitivity가 다르다. 일반적인 Contrast란 배경과 물체를 구분할 수 있는 시각적 요소의 차이를 퍼센트로 표시한 것이지만, 실제 인간이 느끼는 Contrast Sensitivity는 화면을 보는 곳이 밝은 곳인지 어두운 곳인지에 따라 달라진다. 따라서, 인간이 인지할 수 있는 Perceptual Contrast를 구하기 위한 특별한 Filter가 적용된다.

또한, Weber's Law에 의하면 물체 주변의 Luminance가 증가하게 되면 물체를 구분하기 위해서는 더 큰 Luminance의 차이가 필요하기 때문에 이런 요소들도 Human Vision Model에 반영이 돼야 한다.



[Luminance가 높아지면 물체를 구분하기 위해 더 많은 Luminance의 차이가 필요함]

Variance Masking의 경우는 다음의 그림을 보면 쉽게 이해할 수 있다. 물이 흘러가는 그림에서는 쉽게 에러가 있는 블록을 찾아 낼 수 있지만, 나뭇잎 그림에서는 쉽게 블록을 인지하기 어려운 것을 알 수 있다.



Reference



Test

[적은 공간적 변화 영상(Low Spatial Variance Video)]



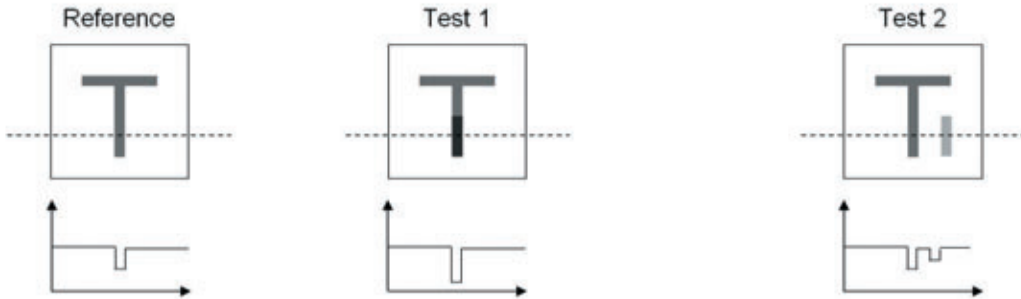
Reference



Test

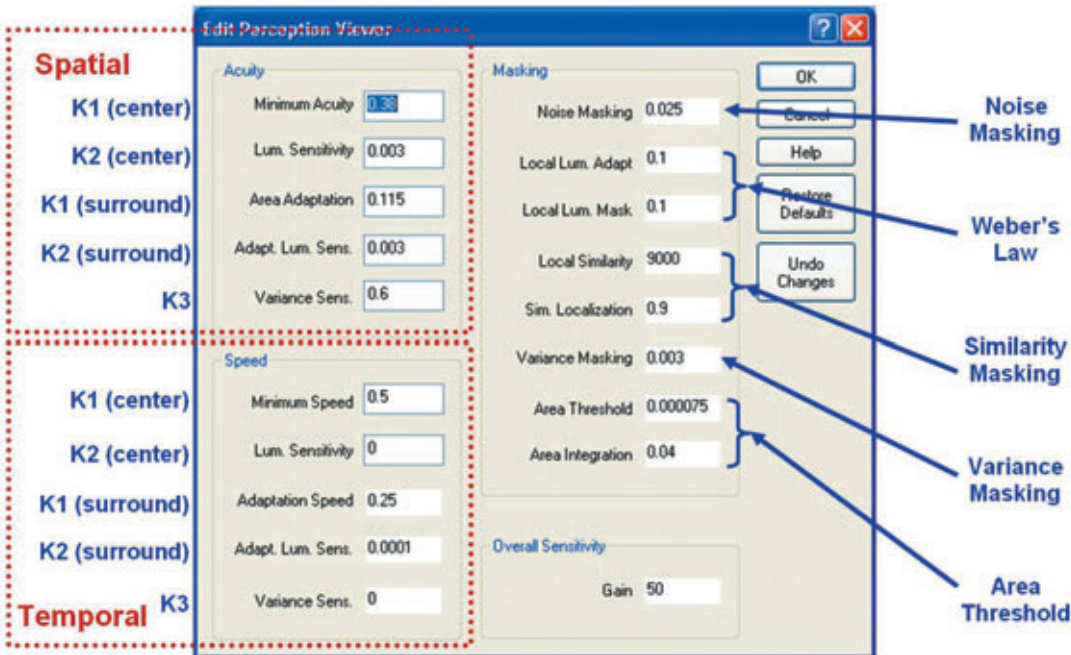
[많은 공간적 변화 영상(High Spatial Variance Video)]

또한, 다음 두개의 Test 1과 Test 2의 그림은 동일한 MSE 값을 가지고 있지만, 실제 사람은 이미지와 연관성이 없는 오점을 가지고 있는 Test 2 영상에서 보다 더 쉽게 오점을 찾아 낼 수 있다. 이러한 특징을 반영한 것을 Similarity Masking이라고 부른다.



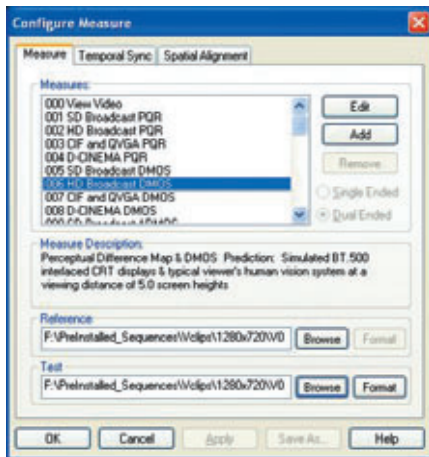
[Similarity Masking]

이러한 여러 가지 요소를 모두 반영하여 개발된 Tektronix의 PQA500에서는 실제 사람이 인지할 수 있는 차이를 계산하기 위한 요소들을 손쉽게 세팅할 수 있다.



[PQA500의 Viewer 특성 편집]

일반 사용자들은 이러한 항목에 대한 항목들을 일일이 세팅하는 일이 매우 번거롭고 어려운 일이다. 그렇기 때문에 PQA500에서는 34가지의 Pre-Config 측정항목이 제공되어 일반 사용자도 손쉽게 Human Vision Model이 적용된 메뉴를 선택할 수 있다.



[347개의 측정항목]

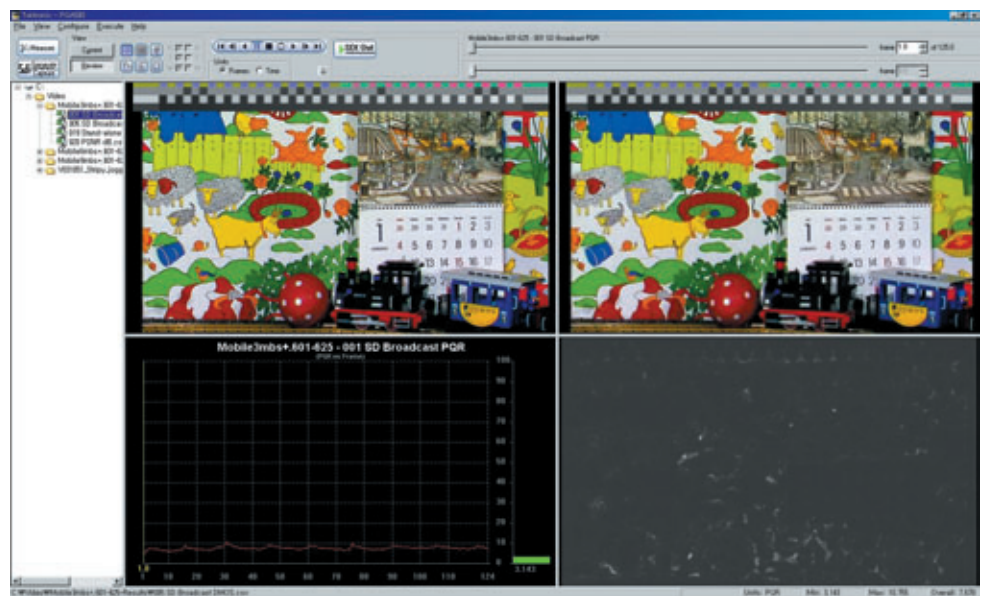
Tektronix PQA500 – Picture Quality Analysis System

Tektronix의 PQA500은 일반적인 객관적인 화질 측정방법인 PSNR뿐만 아니라 앞에서 언급한 Human Vision Model을 적용하여 원본과 테스트 영상 간의 사람이 인지할 수 있는 차이, 즉, Perceptual Contrast Difference를 계산하고 이를 이용하여 PQR(Picture Quality Rating)과 DMOS(Difference Mean Opinion Score) 수치를 제공한다.

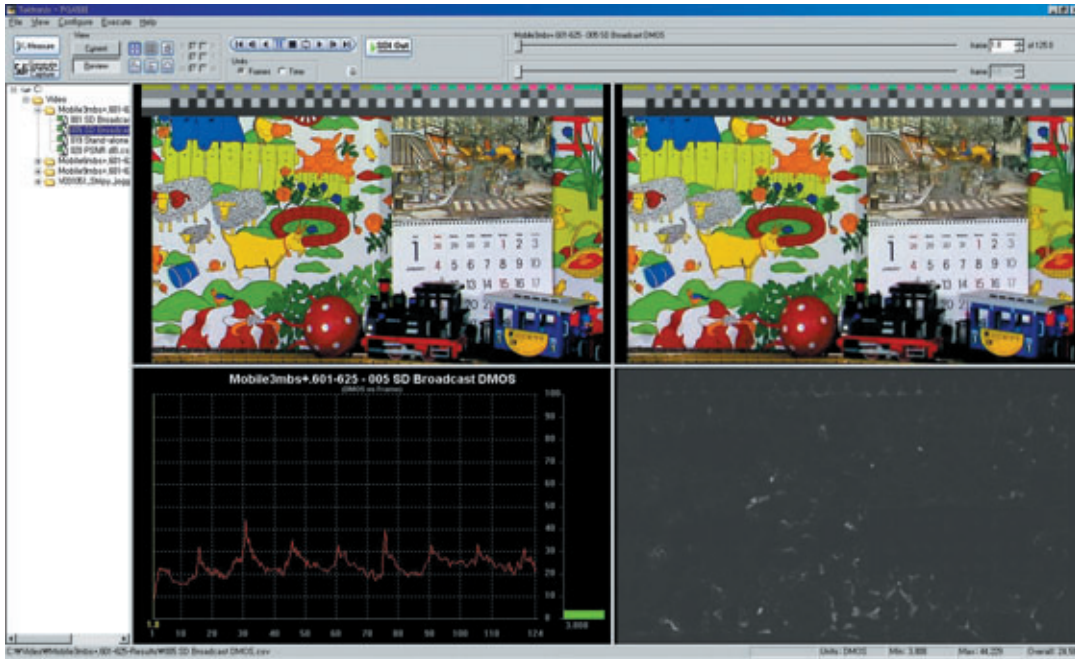
PQR은 JNDs(Just Noticeable Differences) 방식으로 결정하는 화질과 매치되는 화질 수치이며, JNDs 방식으로 보통 100명 중 75명이 원본 영상과 테스트 영상 간의 차이를 느끼고 나머지 25명이 차이를 느끼지 못한다면 그 때의 JND 값은 10이 되며, 1 JND 값은 PQR 1의 값과 같다고 볼 수 있다. 1 PQR의 값은 원본 영상과 테스트 영상 간의 약 0.1%의 PCD 차이를 가지고 있다고 판단한다. 보통 PQR 3~4의 값은 화질의 열화가 있지만 거의 느끼지 못하는 정도로 판단하고, PQR 5는 MPEG-2 SD 10Mb/s의 화질, PQR 6~8은 MPEG-2 SD 4Mb/s 정도로 판단할 수 있다.

한편, DMOS는 ITU-R BT.500에 정의되어 있는 프로시저를 통하여 얻어진다. DMOS 값은 0~100까지의 값을 가질 수 있으며, 이는 ITU Scale의 "Excellent"에서 "Good"까지의 값과 매치될 수 있다. 일반적으로 MOS 혹은 DMOS 값은 테스트들 간의 평가할 수 있는 기준이 다를 수 있기 때문에 트레이닝이 필수적이다. 이를 통하여 Excellent(원본 혹은 원본과 견줄 수 있는 테스트 영상)와 Bad(최악의 테스트 영상)의 범위를 규정할 수 있도록 하기 위함이다. 이와 마찬가지로 PQA500의 DMOS 값 역시 최악의 테스트 영상(worst case training sequence)에 따라 결정된다. 즉, 최악의 테스트 시퀀스에 따라서 실제 DMOS 값이 달라질 수 있다.

다음은 PQA500으로 측정한 PQR 및 DMOS 화면이며, 왼쪽 위부터 원본 영상, 테스트 영상, 프레임별 PQR/DMOS, PCD(Perceptual Contrast Difference Map)를 나타낸다.

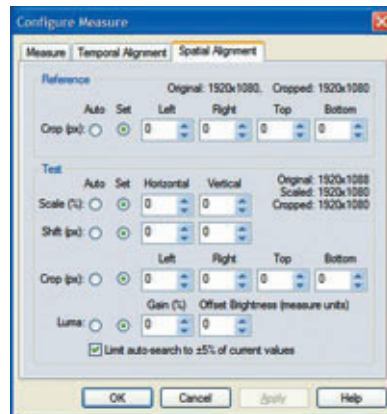
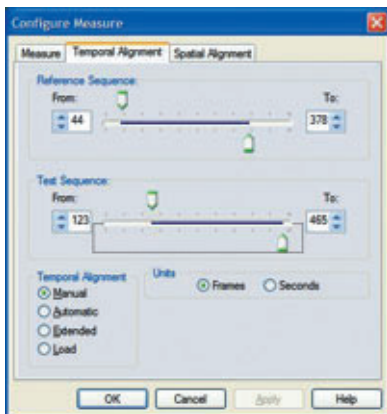


[PQR 측정화면]



[DMOS 측정화면]

한편, 원본 영상과 테스트 영상 간의 화질을 비교하기 위해서는 테스트를 하는 프레임을 정확하게 맞춰주는 것이 필요하다. 이런 과정이 없다면 전혀 다른 프레임을 비교하게 되므로 정확한 화질의 열화를 측정하기 힘들며, 같은 프레임을 비교한다고 하더라도 픽셀이 밀려 (Shift) 있거나 라인이 밀려 있으면 마찬가지로 정확한 측정이 어려울 것이다. 이를 위하여 PQA500에서는 자동 혹은 수동으로 시간적 공간적 정렬(alignment)을 지원한다. 사용자는 이를 이용하여 테스트하는 프레임을 맞추는 시간적 정렬과 프레임 내에서의 크롭(Crop), 픽셀이동(Shift) 등의 공간적 정렬을 수행할 수 있다.



[Temporal/Spatial Alignment 화면]

앞서 관찰자는 화면의 모든 것을 동시에 관찰하기 보다는 일부분만을 응시함으로써 장면을 인지한다고 언급했다. 이에 따라 사람이 집중하고 있는 물체에 관련된 화질의 열화는 전체적인 화질 평가에 많은 영향을 미치지만, 사람이 응시하고 있지 않은 곳에서의 화질의 열화는 전체적인 화질에 큰 영향을 미치지 않을 수 있다. 이를 응용하여 PQA500에서는 화면 상에 사람이 집중하게 되는 부분을 물체의 크기, 위치, 움직임 등의 항목을 자동으로 인식하는 기능이 포함되어 있다.



[사람이 집중하는 부분을 자동으로 인식한 화면]

따라서, PNSR, PQR, DMOS 등의 화질 측정항목에서 이러한 가중치를 적용하여 사람이 집중하는 부분에서 화질의 열화를 보다 더 정확히 살펴 볼 수 있도록 제공한다. 또한, Blockiness, Blurring, Ringing의 정도를 자동으로 측정하여 사람이 집중하는 부분을 가중치에 반영하듯이 실제 화질을 측정할 때 반영하여 측정할 수 있다.



[Blockiness, Blurring, Ringing 측정]

PQA500은 현재 출시되고 있는 화질 평가 시스템 중에서 가장 정확하게 인간의 화질 평가를 대체할 수 있는 장비로 여겨지고 있다. 실제 삼성, LG, Sony, Philips 등의 DVD 혹은 BlueRay Player의 화질을 테스트하고 있는 프랑스 FNAC에서도 PQA200/300과 함께 PQA500을 사용하고 있다.

또한, 최근에 많이 늘어가고 있는 NPS 제작 시스템 상에서는 송출 전에서도 여러 번의 엔코딩/디코딩 과정을 거치게 되며, 이에 따라 화질의 열화가 생길 수 있는 가능성이 많아지므로 방송사에서의 화질 검증 및 엔코더/디코더 제조업체에서의 화질 검증 등에 많이 이용될 수 있을 것으로 생각된다.

※ Human Vision Model 및 PQA500의 화질 평가와 관련된 추가 자료는
http://www.tek.com/products/video_test/pqa500/의 Application Note를 참조