

초고속 네트워크 상의 전송 품질에 관한 주관적 평가

남동우*, 이재기*
*동아대학교 컴퓨터공학과
e-mail : andyw@hitel.net

Subjective Evaluation of Transmission Quality on Very high-speed Network

Dong-Woo Nam*, Jae-Kee Lee*
*Dept. of Computer Engineering, Dong-A University

요 약

본 논문은 기가비트 네트워크 테스트베드인 JGN 상에서 대표적인 스트림형 트래픽인 DV와 MPEG2에 대한 전송 품질의 주관적 평가를 통하여 전송지연과 패킷 손실에 따른 영향이 최종 사용자에게 어떤 영향을 미치는지를 조사한 것이다. 스트림형 데이터에 대해 음질 측정은 ITU-R BS.1116-1을 기반으로 하고 화질 측정은 ITU-R BT.500-1을 기반으로 하여 측정하였다. 또한 최종 사용자의 주관적 평가를 얻기 위한 방법으로는 DSCQS 5단계 평가 방법을 채택하였다. 그 결과 MPEG2 스트림의 경우 네트워크 트래픽 파라미터가 급증하는 54%의 부하율에서 화질과 음질의 평가 등급이 급격히 하강한다는 것을 알 수 있었고 DV의 경우 화질은 MPEG2와 마찬가지로 54%의 부하율에서 평가 등급이 감소하기 시작하였으나 그 변화도도가 MPEG2에 비하여 완만한 변화를 보였다. DV 음질의 경우 트래픽 파라미터(RTT, 패킷 손실)와는 큰 연관성을 보이지 않았고 대부분의 측정자들이 보통 수준의 평가등급을 보였다.

1. 서론

최근 인터넷 사용이 급증하면서 사용자들은 종래의 텍스트 위주의 데이터 교환에서 멀티미디어 데이터의 교환을 주로 하는 스트림형 통신에 더 많은 관심을 보이고 있으며 앞으로 이 스트림형 정보의 교환이 빈번해질 것은 자명한 사실이다. 따라서 멀티미디어 데이터의 전송을 위한 고속의 네트워크 환경이 필요하게 되었고 이러한 고속의 환경을 위해 미국의 NGI, 캐나다의 CANARIE, 일본의 JGN 등의 테스트베드들이 구축되었다[1]. 이 중 JGN 상에서 동일한 스트림형 이면서도 성질이 다른 DV와 MPEG2에 대한 트래픽에 대한 특성 비교가 시행되어 트래픽 파라미터(RTT, 패킷 손실)가 부가된 정상 부하에 따라서 어떤 영향을 미치는지에 대한 연구는 시행되었으나[2] 이런 자료들이 화질과 음질에 대한 최종 사용자의 평가에 대해서는 논의된 바 없다. 따라서 본 논문에서는 네트워

크 부하의 증가에 따라 DV와 MPEG2의 화질과 음질에 대한 최종사용자의 주관적 평가를 얻어 내므로서 네트워크 부하의 실제 가용값을 도출하고자 한다. 주관적 평가에 사용되는 스트림은 DV와 MPEG2 모두 동일한 스트림으로 평가하고 네트워크의 대역폭을 90%(45Mbps)까지 10%(5Mbps)씩 증가시킨다. 또한 디지털 영상의 경우 주어진 화상의 내용이 화질에 미치는 영향이 크므로 DV 스트림과 MPEG2 스트림간의 비교는 가능한 동일한 내용과 동일한 조건으로 행한다.

본 논문의 구성은 1장 서론과 2장에서 주관적 측정 방법과 객관적 측정 방법에 대해 다루고 4장에서는 주관적 평가에 의해 획득된 데이터에 대해 정리하며 5장에서는 얻어진 자료에 대한 결론을 맺는다.

2. 주관적 측정방법

화질의 평가 방법인 ITU-R 권고 BT.500-10 "TV 화상의 주관적 측정에 관한 방법"[6]은 주관적 평가의 목적에 따른 다양한 평가 방법을 규정함과 동시에 관찰 조건 및 평가 결과의 분석법 등도 포함한 일반적인 사항을 규정하고 있다. 화질의 평가 방법은 이 권고 사항을 바탕으로 하며 음질 평가는 ITU-R 권고 BS.1116-1 "다중 채널 시스템을 포함한 오디오 시스템의 작은 손상에 관한 주관적 평가 방법"[5]을 바탕으로 하여 측정한다. 그의 음질과 화질의 평가에 관련된 ITU-R 권고 사항은 표 1, 2 에서와 같으며 각각의 권고 사항을 참조한다.

<표 1> 화질 평가에 관한 주요 ITU-R 권고

권고 번호	내용
BT.500-10	관찰조건, 평가법, 경과 분석법
BT.1128-2	표준 TV의 관찰 조건과 평가 방법
BT.1129-2	디지털 TV의 관찰 조건과 평가 방법
BT.710-4	HDTV의 관찰 조건과 평가 방법
BT.1438	입체 TV의 관찰 조건, 평가 방법, 표준화상
BT.1382	다중 채널 서비스의 평가 방법
BT.814-1	디스플레이의 휘도, Contrast 조정법
BT.815-1	디스플레이의 Contrast 비 측정 신호, 표준화상
BT.1120-1	

<표 2> 음질 평가에 관한 주요 ITU-R 권고

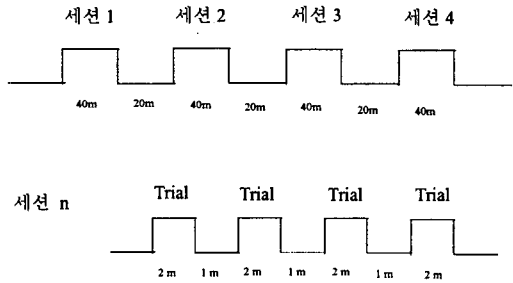
권고 번호	내용
BS.1116-1	다채널 음성 시스템을 포함하는 음질 저하가 적은 음성 시스템의 주관적 평가법
BS.1284	음질의 주관적 측정을 위한 방법 0 일반 일반 요구사항
BS.1285	오디오 시스템의 작은 손상의 주관적 측정을 위한 Pre-selection 방법
BS.1286	영상을 포함한 오디오 시스템의 주관적 측정 방법

화질 및 음질의 평가법은 DSCQS(Double Stimulus Continuous Quality Scale)법을 사용하여 <그림 1>에 나타난 5 단계 DSCQS 법의 연속 스케일로 평가한다.



<그림 1> DSCQS 법의 연속 스케일

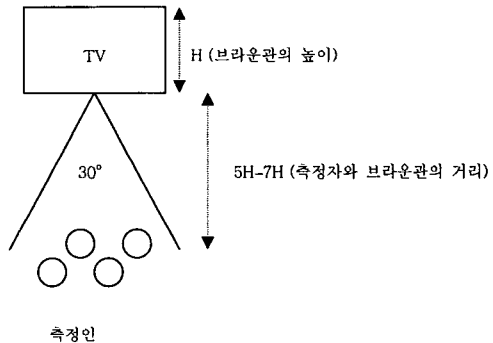
DSCQS 법은 많은 화질/음질 저하 요인을 포함하는 경우의 화질/음질을 측정하는데 적합하고 참조 영상 및 음성에 대한 평가물의 화질 및 음질을 연속 스케일 방식으로 평가한다. 측정 내용물은 4 개의 세션으로 구성되며 각각 MPEG2 화질, MPEG2 음질, DV 화질, DV 음질이다. 각 세션은 5Mbps(10%)에서 부터 45Mbps(90%)까지 네트워크 부하율에 따라 10 개의 trial로 구성된다. 즉 trial-1 은 10%부하율의 5Mbps, trial-2 는 20%부하율의 10Mbps 와 같이 된다. 각각의 trial 은 임의의 순서로 진행되어 측정인이 그 순서를 예측할 수 없도록 하였다. <그림 2>는 세션의 구성을 나타낸 것이다.



<그림 2> 세션의 구성

각각의 세션은 동일한 내용의 스트림형 데이터로 구성되고 주어진 참조 영상 및 음성과 비교하여 DSCQS 법의 5 단계 주관적 평가를 내린다.

실험실의 구성은 <그림 2>의 배치도에서 나타나는 바와 같이 브라운관의 높이를 H 라고 할때 5H-7H 사이에 배치하도록 하고 측정인이 브라운관에서 30. 이내 에 포함되도록 배치한다.



<그림 3> 배치도

앞에서 언급했던 주관적 측정 방법에 반해 영상이나 음성의 품질을 객관적으로 측정 및 감시하려고 하는 요구가 증대하고 있다. 아날로그 시스템인 경우, 화질 저하가 미미하지만 발생 빈도가 높고 많은 경우 몇 가지 테스트 화상을 사용한 평균 화질로 전체를 대표하게 할 수 있다. 한편, 디지털 압축 시스템에서는 화질이 화상 내용에 크게 의존하기 때문에 평균

화질만으로는 판단할 수 없고, 화질 저하의 발생 빈도나 최악 화질도 고려할 필요가 있다. 시간적으로 연속한 주관적 평가는 이와 같은 성질이 고려된 것이지만, 장시간 측정이나 온라인 측정은 곤란하다. 그래서 ITU에서는 화질의 객관적 평가법의 표준화가 검토되고 있고, ITU-R 과 ITU-T 를 모체로 하는 VQEG(Video Quality Expert Group)가 ITU 권고 작성을 위해 활동 중에 있다[4]. VQEG 에서의 객관적 평가법 제안 모집에 대하여 10 가지 방식이 제안되어 제안방식의 성능 평가가 행해졌으나[7] 해석 결과에 의해서 현재 단계에서는 특정 방식을 권고할 수 있는 상황이 아니라고 판단되어 본 논문에서는 객관적 방법을 배제하고 주관적 방법을 고려하여 측정하였다.

3. 측정된 데이터의 분석

2 장에서의 방법대로 초고속 네트워크 상의 전송 품질에 관한 최종 사용자의 주관적 평가를 시행한 결과 얻어진 데이터의 처리는 아래와 같다.

각 Trial 에 대한 평균값 \bar{u}_j 을 구한다. 구해진 평균값과 표준 편차를 사용하여 신뢰구간 95%를 만족하도록 유지하고 만족되지 않는 평균값은 신뢰할 수 없는 값이므로 재측정 한다.

평균값과 95%의 신뢰 구간은 반드시 주어져야 한다.

- 특정 Trial 의 평가 평균값

$$\bar{u}_j = \frac{1}{N} \sum_{i=1}^N u_{ij}$$

u_{ij} = 실험 조건 j에 대한 실험자 i의 평가 점수
 N = 실험자의 수

- 표준 편차를 이용한 95%의 신뢰 구간

아래에서 주어진 95%의 신뢰 구간을 갖는 것으로 가정한다.

$$[\bar{u}_j - \sigma_j, \bar{u}_j + \sigma_j]$$

여기서

$$\sigma_{jk} = 1.96 \frac{S_j}{\sqrt{N}}$$

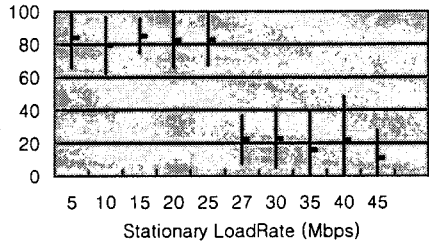
각각의 표준 편차 S_j 는 아래와 같이 구한다.

$$S_j = \sqrt{\sum_{i=1}^N \frac{(\bar{u}_j - u_{ij})^2}{(N-1)}}$$

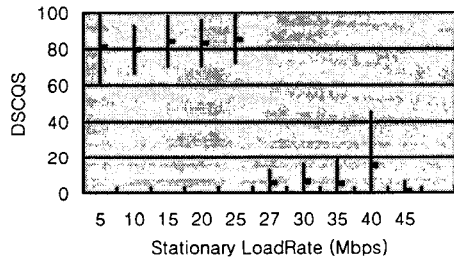
MPEG2 화질 및 음질의 경우 네트워크 부하율 27Mbps(54%)경우 부터 최종 사용자들의 평가 등급이 급속히 떨어지는 것을 알 수 있다. MPEG2 는 프레임 간 (Interframe)압축을 수행하는 스트림이므로 I 프레임이 정확하게 수신되지 못할 경우 관련 P 프레임과 B 프레임도 손상되므로 화면이 일시 정지되는 현상을

보인다. 이로 인해 27Mbps(54%) 네트워크 로드가 발생할 경우 사용자의 평가 값이 급격히 떨어진 것으로 보인다.

MPEG2 화질



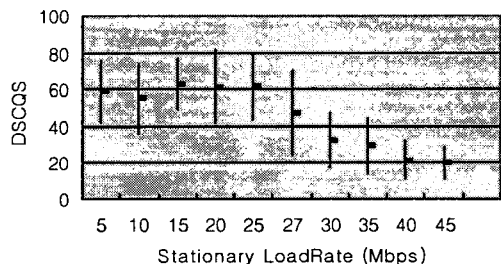
MPEG2 음질



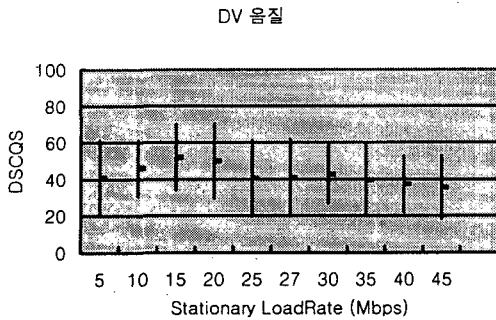
<그림 4> MPEG2 화질 및 음질 측정 결과

DV 화질의 경우 네트워크 부하율 약 27Mbps(54%)까지는 fair(보통) 수준이었으나 30Mbps(60%)부터 평가 등급이 조금씩 감소하기 시작하였다. MPEG2 화질에 비해 그 변화 정도가 완만한 것은 DV 는 프레임의 내부(Intraframe)에서 압축을 수행하는 스트림으로 트래픽 파라미터(RTT, 패킷 손실)에 의해 프레임이 손상될 경우 해당 프레임 자체에만 영향을 미치게 되어 화면에 모자이크형의 잡음이 부하율에 비례하여 많이 나타나게 된다. 그러나 MPEG2 화상처럼 화면이 일시 정지하는 현상을 일어나지 않는다. 이러한 이유로 최종 사용자의 평가 등급이 완만히 변화하는 것으로 판단된다.

DV 화질



[7] 니시다(西田), “객관적 화질 평가법의 표준화 동향”, 영상정보미디어학회지, 제 53 권 제 9 호, pp.1206-1208, 1999.



<그림 5> DV 화질 및 음질 측정 결과

DV 음질의 경우 사용자의 평가 등급이 네트워크의 부하율과는 크게 상관이 없다고 할 만큼 그 값들이 대부분 “fair(보통)” 등급에 속하는 값들이 었다.

4. 결론

연구 개발용 기가비트 네트워크인 JGN 상에서 정상 부하(512B)의 패킷 크기와 부하율을 변경하면서 DV 스트림과 MPEG2 스트림을 전송하고 이에 대한 최종 사용자의 주관적 품질평가를 측정하였다. MPEG2 스트림은 약 27Mbps (부하율 54%)부터 패킷 손실과 RTT 가 급격히 증가하는 현상을 보였다[2]. 최종사용자의 주관적 평가 역시 부하율 약 54%인 이 시점에서부터 급격히 감소하는 것을 알 수 있었다. DV 스트림의 화질평가의 경우 패킷 손실과 RTT 가 급증하는 시점에서 사용자의 평가 등급도 급격히 감소하는 현상을 보였으나 DV 스트림 음질평가의 경우는 이런 트래픽 파라미터와는 상관없이 대체적으로 보통의 평가 등급이나와 약 35Mbps(70%) 정도의 부하율에서도 사용자의 평가는 음질이 나쁘지 않다고 판단하고 있다. 위의 결과에 의하면 패킷 손실이 발생하지 않는 범위나 RTT 지연이 급증하지 않는 범위에서 트래픽 제어를 할 필요가 있을 뿐 아니라 최종 사용자의 주관적 판단에 의한 허용 범위를 고려하여 트래픽 제어를 할 필요가 있다.

참고문헌

- [1] http://www.ngi.or.kr/other_site/map21.html
- [2] 이재기 외, “JGN Traffics Measurement through Image Transmission”, 信學技報, SSE2000-222, OCS2000-103, pp.73-78, 2001.
- [3] Junji Kunmada 외, “화상 평가와 관련한 국제 표준화 동향”, 일본 영상정보 미디어 학회지, 제 54 권 제 7 호, 2000.
- [4] VQEG Official Web Site <<http://www.vqeg.org/>>
- [5] ITU-R BS.1116-1, “Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems”.
- [6] ITU-R BT.500-10, “Methodology for the subjective assessment of the quality of television pictures”.