

# 멀티미디어 서비스 품질의 양방향 주관평가법

Interactive Test Methods for Multimedia Service Quality

김정환(J. H. Kim) 정보분석실 선임연구원

본 고에서는 최근 비약적으로 발전하고 있는 멀티미디어 서비스의 품질설계와 품질관리에 기초자료로서 활용할 주관 품질평가법의 필요성과 목적을 알아보고, ITU-T Study Group 12 Question 22/12(Audiovisual Quality in Multimedia Services)에서 추진하고 있는 멀티미디어 서비스의 주관 품질평가법 권고화 동향중, AV 통신에 대한 양방향 주관 품질평가법에 관한 권고(Draft Rec. P.920)를 분석하였다.

## I. 서 론

최근 통신과 컴퓨터 기술의 비약적인 발전에 힘입어, 우리는 사무실, 학교, 공공기관은 물론 일반 가정에서도 손쉽게 멀티미디어의 혜택을 누리게 되었다. 이렇게 다양한 멀티미디어 서비스의 개발과 함께, 미국을 비롯한 일본 및 유럽공동체 등 통신 선진국에서는 21세기를 향한 경제, 사회, 문화적 측면에서의 변화와 급속한 기술혁신의 발전 등을 고려하여, 정보통신 서비스 사용자의 요구에 부응하기 위한 구체적인 서비스 계획을 마련하고 있으며, 일본 NTT의 VI & P(Visual, Intelligent and Personal Communications Service)와 같은 것이 한 예이다[1].

멀티미디어 서비스는 영상, 음성/음향, 문자, 도형 및 수치 등 표현 미디어가 다른 일련의 정보들을 디지털화하여 공통의 전송계에 실어 사용자에게 보내는 서비스를 말한다. 기존의 통신품질 주관평가법(subjective tests method on communication quality)

들은 음성/음향과 영상이라는 개별 미디어에 대해 서만 그 품질을 평가하여 왔고, 음성/음향과 영상 그리고 그외의 미디어가 복합되었을 때의 품질평가법 연구는 아직 시작단계에 머물러 있다.

한편, “End-to-End Performance of Telephone Networks and Terminals”라는 과제명으로, 음성, 문자 및 영상 서비스를 제공하는 망(network)과 단말(terminal)의 주관적인 품질과 단말간 전송성능에 대한 연구를 수행하고 있는 ITU-T Study Group 12에서는 1993~1996년의 연구 회기기간중 Question 22/12(Audiovisual Quality in Multimedia Services)를 통해 멀티미디어 서비스의 주관 품질평가법 권고화를 추진하고 있다[2]. 1996년초 현재 Question 22/12에서 권고화를 추진하고 있는 주된 내용을 <표 1>에 나타낸다[3-7].

본 고에서는 제II장에서 멀티미디어 서비스에 대한 주관 품질평가의 목적을 간략히 기술하고, 제III

〈표 1〉 ITU-T SG 12 Question 22/12에서 추진중인 권고들

권고	제목 및 내용
ITU-T Rec. P.VQ	<ul style="list-style-type: none"> <li>제목: Subjective video Quality assessment methods for multimedia applications</li> <li>내용: 저비트율 또는 중간 비트율로 부호화된 디지털 영상의 일방향적 주관품질 평가법에 대한 권고</li> </ul>
ITU-T Rec. P.RISV	<ul style="list-style-type: none"> <li>제목: Reference impairment system for video</li> <li>내용: 디지털 영상압축시스템이 생성한 영상의 주관적인 품질을 규정하는데 필요한 기준 조건들을 발생시키는 조절형 영상기준시스템에 대한 권고</li> </ul>
ITU-T Rec. P.ITU	<ul style="list-style-type: none"> <li>제목: Interactive test methods for audiovisual communications</li> <li>내용: 두지점간 또는 다지점간 AV(audiovisual) 통신에 있어서 부호화 오류(coding artefacts) 및 전송지연(transmission delay)의 효과를 수량화할 양방향 평가법에 대한 권고</li> </ul>
ITU-T Rec. P.AVQ	<ul style="list-style-type: none"> <li>제목: Subjective assessment methods for global audiovisual quality evaluation in multimedia applications</li> <li>내용: 음향이 수반된 디지털 영상의 전체적 품질평가를 위한 일방향적 주관품질 평가법에 대한 권고</li> </ul>

\* 1996년 5월 6일~15일에 Geneva에서 개최된 ITU-T SG12 회의에서 P.VQ, P.ITU 및 P.RISV는 회의 참가자들의 만장일치로 각각 권고 P.910, P.920 및 P.930(Principles Reference impairment system for video)으로 확정되었음[8, 9]

장에서는 ITU-T SG 12 Question 22/12에서 권고화를 추진하고 있는 AV(audiovisual) 통신에 대한 양방향 주관 품질평가법(Rec. P.ITU이 Draft Rec. P.920[10]으로 명칭 변경)에 대해 고찰하였다.

되도록 서비스를 설계하려면, 사용자가 느끼는 품질요인의 영향을 정확하게 평가할 수 있어야 하고, 이러한 품질평가의 결과를 망과 단말의 품질설계와 품질관리에 활용해야 한다.

통신서비스의 좋고 나쁨은 사람이 음성이나 영상에 대해 느낀 심리적인 판정에 의해 결정된다. 그러나, 사람을 대상으로 실시하는 주관평가법은 시간과 환경에 가변적인 특성을 가지고 있으므로, 재현성(reproducibility)이 좋은 주관 품질평가법을 구성하기 위해서는 측정환경의 표준화가 반드시 필요하다. 이렇게 표준화된 주관평가법에 의해 얻어진 평가결과는 계측 가능한 물리량을 추출해 내는 객관측정법(objective test)의 기초자료가 되므로, 결국 품질평가법의 최종목표는 시간과 금전적인 측면에서 경제적인 객관측정법을 개발하는 것이라 볼 수 있다. 따라서, 멀티미디어 서비스의 품질평가를 위해 우리가 해야 할 일은 먼저 표준화된 주관 품질평가법을 개발하고 궁극적인 객관 품질평가법 개발로 나아가는 것이다.

## II. 멀티미디어 서비스에 대한 주관 품질평가의 목적

멀티미디어 서비스를 경제적이면서도 좋은 품질로 제공하기 위해서는 멀티미디어의 품질요인(quality factor)을 파악하고, 사용자들의 실제 감각에 맞는 주관 품질평가법을 개발하여야 한다. 주관 품질평가는 사용자의 요구사항을 파악하고, 품질설계와 품질관리에 필요한 진단적 평가, 서비스간 상호 품질비교 및 규격 적합성 평가에 활용될 수 있다. 그러나, 실제로 망이나 단말의 경제성과 품질은 상충하는 면을 가지고 있는데, 즉, 경제성을 추구하면 품질이 나빠지고, 품질이 좋아지면 일반적으로 가격이 상승한다. 따라서, 경제성과 품질이 잘 조화

### III. AV 통신에 대한 양방향 주관 품질 평가법(Draft Rec. P.920)

#### 1. 개요

이 AV 양방향 평가법은 양방향 AV 통신에 영향을 미칠 수 있는 부호화 오류(coding artefacts)나 전송지연(transmission delay)과 같은 요인들의 효과를 수량화하기 위한 것으로, 이 평가법의 효율성은 실험실 환경이 실제 상황을 얼마나 잘 재생하는가에 전적으로 달려 있다. 따라서 피험자(subject)에게 주어지는 과제(task)의 선정에 특별한 주의가 필요하다. 일반적으로, 전화에 대한 대화실험(conversation tests)에서 사용되는 과제들은 종종 피험자의 주의(attention)를 영상화면에서 분산시키기 때문에 AV 품질평가는 적절하지 못하다. 따라서, 본 권고에서 기술하고 있는 기준에 따라 새로운 과제가 개발되어야 한다. 비록, AV 품질에 대한 모든 측면들이 아직까지 완전히 이해되지는 않았지만, 이 분야에서의 실질적인 활동이 이루어져야 하며, 본 권고는 양방향 AV 실험에 대한 현재의 초기 연구상태를 반영하고 있다. 이러한 작업이 지속적으로 이루어지면, 이 양방향 평가법에 대한 이해의 폭이 반드시 넓어질 것이며, 새로운 연구결과에 따라 권고는 개정될 것이다.

본 권고는 두지점(point-to-point) 또는 다지점(multipoint)간 AV 통신에서의 부호화 오류와 전송지연의 영향을 수량화하는 양방향 평가법을 규정하기 위한 것으로, 채택한 방법론은 대화형 오피니언 평가법(conversation opinion tests)에 기초하고 있는데, ITU-T 권고 P.80 부록 A[11]에서 규정하고 있는 방법의 확장으로 볼 수 있다. 그리고 이미 다른

P-series 권고에서 규정하고 있는, 링크 품질(link quality)에 대한 객관평가와 같은 주제는 다루지 않고 있다.

#### 2. 실험설계(Experimental Design)

##### 가. 연구의 기본적 접근과 요인들

양방향 통신을 수행하는데 영향을 미칠 수 있는 전송지연과 같은 요인들의 효과를 수량화하기 위해, 본 권고에서 제안하고 있는 접근법은 통화자들의 실제 대화평가에 기초하고 있다. 또한, 평정체계를 이용해 이러한 오피니언들을 표현할 필요가 있어 몇가지 단일-자극 평정척도(single-stimulus rating scale)를 제안하고 있다.

##### 나. 대화과제에 사용되는 자극들

일반적으로, 대화형 오피니언 평가에서는 환경의 인위성을 최소화하여야 한다. 그러나, 동시에 평가되고 있는 조건들을 이용해 양방향통신을 유도하기 위한 몇가지 방법들을 시행할 필요도 있다. 전화에 대한 품질평가에서는 이러한 목적을 위해, 일련의 사진들이나 다른 인쇄물들을 이용하는 경우가 일반적이지만, AV 단말기의 성능평가에서는 그러한 방법이 참여자의 주의를 영상화면에서 분산시킬 수 있어, AV 단말기에 대한 부자연스러운 통신환경을 유발할 수 있다.

과제형 평가(task-based test)를 설계하는데 다음과 같은 지침(guidelines)이 제공되고 있다[12].

- 피험자들이 대화하는 동안 AV 단말기에 지속적인 주의를 두도록 과제가 설계되어야 한다.
- 과제는 실제 AV 통신과 매우 유사해야 한다. 특

히, 과제는 두명의 피험자가 수행해야 하며, 피험자와 실험자가 수행해서는 안된다.

- 과제는 통신의 효율성(communication efficiency)에 대한 적합한 측정치를 나타내는 재생 가능한 수량적 결과를 산출해야만 한다.

이 대화과제에는 고령자와 청각 장애인을 포함해, 다양한 피험자가 과제를 수행할 수 있어야 하며, 과제 자체가 피험자에게 충분한 보상을 줌으로써 피험자는 더 빨리 과제를 학습하게 되고 피로와 동기의 상실을 덜 느끼게 된다.

피험자들이 사전에 서로 알고 있는 경우, 실제 AV 대화가 촉진될 수 있음이 이전 실험에서 밝혀졌다. 또한, 원래의 자극원(source of stimulation)과는 별도로 손으로 작성된 자극이 대안적으로 사용될 수 있다. 따라서, 전화와는 달리, 대화하는 참여자간의 친숙성(familiarity)이 요구될 수 있으나 반드시 그러한 것은 아니다.

그러나, 특정한 목적을 위해, 평가 시스템이 제공하고자 하는 서비스를 고려하여 대화과제는 수정될 수 있다. 연구대상 요인에 의한 의미있는 측정치를 도출하기 위해서는, 다음 사항을 고려해 대화형 과제가 구성되어야 한다.

- 정보 변화율(the rate of information exchange)
- 음향 및 영상신호 이용도(the degree of audio and video signal utilization)

예를 들어, 정보변화율의 속성들을 설명하기 위해, 과제는 일방향 통신(one-way communication)부터 자유대화(free-conversation)까지, 영상, 음향 또는 2개의 신호를 통한 빠른 정보교환까지의 범위를 가져야 한다. 그리고 음향 및 영상신호 이용도의 속성들을 평가하기 위해, 과제는 앞에 놓인 하드카피 문서

로 작업하는 피험자부터(영상 정보의 최소 사용) 영상 링크를 통해 표현어를 읽는 피험자까지(영상 정보의 최대 사용)의 범위를 가져야 한다. 그리고 실제 과제는 이 2가지 속성을 결합해야 하며, 이러한 지침들은 부록 I에서 기술된 과제의 개발과, 부록 II에서 기술된 과제를 위한 프로토콜 개발에 적용되었다.

#### 다. 평가조건과 실험설계

일반적으로, 하나의 평가에서는 적어도 하나의 전송손실 요인(transmission impairment factor)이나 평가조건(test conditions)이 평가되어야 하며, 그러한 요인의 효과가 최소가 되는 기준조건(reference condition)이 있어야 한다. 그러나, 대화실험은 시간이 걸리므로 피험자의 피로를 최소로 하고 실험의 정확성을 최대로 하기 위해 전체 실험조건수는 적절히 제한되어야 한다. 이러한 요구사항은 각 대화 조건의 지속시간을 적어도 5분 정도로 해야하는 점과 적절히 균형을 맞추어야 한다. 대화형 과제에서는 Latin이나 Graeco-Latin square design이 적절한 실험설계로 확인되었는데[13], 이 경우 방형(square)의 횡축은 피험자, 종축은 실험조건의 제시순서에 해당할 수 있다.

다른 실험조건들은 연구대상이 되는 요인들에 따라 다를 수 있는데, 예를 들어, 과거의 실험에서 AV 통신의 path 품질과 전송지연 효과의 지각 사이에 상호작용이 있을 수 있음이 밝혀졌다. 결과적으로, Graeco-Latin square design으로 2개의 조건을 이용하는 것이 바람직하며, 첫번째 알파벳은 전송지연의 다른 값들을, 두번째 알파벳은 이미지/음성 부호화율의 다른 값을 각각 나타낸다.

물론, 반복적인 블록 설계(replicated block design)와 Youden square design을 포함한 다른 실험설계가 이용될 수 있는데, 실험자는 실험목적에 따라 적절한 실험설계를 선택할 수 있다. 또한, 과제 수행의 순서와 관련되어 있을 수 있는 어떠한 효과도 고려해야만 한다.

#### 라. 피험자

적어도 16명의 피험자가 실험에 참여해야 하며, 정확한 피험자의 수는 실험설계와 결과도출에 요구되는 정확성에 의해 결정될 것이다. 피험자들은 비전문가로 그들의 일상적 업무의 일부가 음향이나 영상기술과 직접적으로 무관해야 한다. 그럼에도 불구하고, AV 통신시스템 개발의 초기단계와 광범위한 실험전에 수행되는 예비실험에서는 소규모의 전문가집단(4~8명)이나 다른 특정 피험자가 참여함으로서 신뢰성이 높은 결과를 얻을 수도 있다.

#### 마. 피험자 훈련과 기준접속

실험을 실시하기 전에, 실험대상이 된 시스템이 어떻게 사용될지에 대한 시나리오를 피험자에게 제시하여야 한다. 손실(impairments)의 범위와 유형은 연습단계에서 보여주어야 하는데, 이 단계에서 피험자가 최악(또는 최상)의 실험조건에서 최상(또는 최악)의 실험조건을 경험하도록 해야 한다. 본 실험에서는 실험조건의 세부사항을 피험자에게 알려주어서는 안된다.

#### 바. 방음실과 장비특성

실험실의 암소음 특성(ambient noise characteristics)은 일반 사무실 환경을 대표하여야 하고, 두 환경간

의 음향방진(acoustice isolation)은 60dB가 좋다. 대화 평가에 사용되는 실험실의 음향특성은 ITU-T 권고 P.80에서 기술되어 있다. 발광체 및 피험자와 관련된 실험실의 시각적/조도 특성(visual/luminance characteristics), 화면 디스플레이의 밝기 및 장비특성은 다음과 같다.

- 스크린과의 거리: 4H~8H(H는 스크린 높이)
- 최대 조도:  $70\text{cd}/\text{m}^2$ ~ $200\text{cd}/\text{m}^2$
- 배경 발광체가 없는 경우, 스크린의 명암 대비율: 30~50
- 배경조도와 최대 스크린 조도와의 비: ~0.25
- 휘도: 약 400lux
- 일반적 색채: 흰색

스크린과의 거리는 스크린 크기 뿐 아니라, 스크린의 유형, 어플리케이션의 유형 및 실험의 목적을 고려해야 한다. 주어진 스크린 높이에 대해, 화질이 떨어질 때까지 피험자가 선호하는 거리가 되도록 증가시키는 것이 바람직하다. 이러한 점을 고려하여, 스크린과의 거리는 확인실험(qualification tests)으로 미리 결정해 두어야 하며, 보다 확실한 거리는 최적실험(optimization tests)에서 사용되어야 한다.

영상을 디스플레이 하는데는 전체 스크린을 사용하는 것이 좋지만, 몇 가지 이유로 영상을 스크린의 한 원도에 디스플레이 해야 할 때는, 스크린의 배경색을  $Y=U=V=128$ 과 일치하도록 50% 회색화 시켜야 한다. 핸드프리 동작과 관련된 다른 가능한 요인들과 같이, 통제되지 않은 실험변수들은 최소화하여야 한다.

### 3. 오피니언 도출

전화기에 대한 대화평가와 같이, 각 피험자는 각

실험조건이 끝난 후에 그들의 오피니언을 산출해야 한다. 질문의 수를 최소화하고, 통제되지 않은 변수들을 최소화하기 위해 미리 질문의 유형을 구성하는 것이 필요하며, 이러한 목적에 사용될 수 있는 질문들은 부록 III에 나타내었다.

몇 가지 범주판단 척도(category judgement scales)가 AV 단말기의 성능평가에 사용될 수 있으나, 다양한 전송손상 요인들에 대한 이러한 척도들의 효율성은 다를 수 있다. 이러한 목적에 사용될 수 있는 척도들의 예는 다음과 같다.

- 전체 AV 품질, 영상품질 및 음향품질은 일반적으로 다음 범주들을 사용해 평가된다: 매우 좋다, 좋다, 보통이다, 나쁘다, 아주 나쁘다
- 상대방의 말을 끊기(interrupt)에 필요한 노력은 일반적으로 다음 범주들을 사용해 평가된다: 노력없음, 적은 노력, 보통 노력, 상당한 노력, 매우 노력
- 통신난이도(communications difficulty)와 통신의 수용성(acceptability)은 일반적으로 양극 선택(binary choice)에 의해 평가된다: 예, 아니오

비록 개별적인 실험의 목적에 따라 특정 척도가 선정될 수 있다고 해도, 이러한 척도들은 다른 실험 실에서도 지속적으로 사용되어야 한다는 점이 중요하다. 만일 AV 단말기가 대부분의 피험자에게 생소하다면, 마지막 실험조건 제시후 마무리 질문(exit question)을 받아두는 것이 좋다. 그러한 질문들은 실험에서 부주의해서 빠질 수 있는 어떤 다른 요인들을 찾아낼 수 있기 때문이다. 그러한 질문의 예를 부록 IV에 나타내었다.

## 부록 I. 대화를 위한 과제와 자극의 예

### I.1 대화를 위한 자극

다음 과제는 자연스런 대화의 정도에 있어서 각각 차이가 있으며[12], 과제를 위한 프로토콜은 부록 II에 보다 자세히 기술되어 있다.

- 과제 1: 이름-사고 과제(name-guessing task)는 고정된 프로토콜에 따라 수행하는 질문-대답 게임으로, 매우 제한된 대화를 하게 된다.
- 과제 2: 이야기-비교 과제(story-comparison task)에서 피험자는 동일한 이야기에 대한 2개 다른 버전간의 차이수를 발견해야 한다. 피험자들은 무제한적인 대화를 할 수 있으며, 실험전에 두 피험자는 단편적인 이야기를 읽고 기억해야 한다.
- 과제 3: 그림-비교 과제(picture-comparison task)에서 피험자들은 하나의 그림을 기억하고, 이어서 주어진 그림이 같은지 또는 다른지를 결정해야 하는데, 대화는 무제한이다.

### I.2 통신품질에 대한 음성지연의 효과를 평가하기 위한 과제들

다음 과제에서는 발화속도를 과제 1부터 6까지 증가시키면서, 반면에 대화율은 감소시킨다[14].

- 카운팅을 교대로 한다.
- 가능한 한 빠르고 크게 교대로 무작위 수를 읽는다.
- 가능한 한 빠르고 크게 교대로 무작위 수를 확인한다.
- 빠진 철자가 있는 단어는 다른 발성자가 불러준

철자로 메꾼다.

- 가능한 빨리 도시 이름을 교대로 확인한다.
- 언어적으로 기술된 윤곽을 결정한다.
- 자유대화

위의 과제(과제 1과 7은 제외)는 피험자에게 화면이 아닌 질문지에 주의를 집중하도록 요구하기 때문에 AV 품질평가에 사용할 수 없다.

### I.3 통신품질에 대한 AV 지역의 효과를 평가하기 위한 과제들

다음의 과제들은 영상신호에 대한 평가자들의 주의를 기술할 수 있다.

- 피험자중 한명이 블록이 쌓여 있는 모양을 말해 주면, 다른 한명이 그것을 만들어야 한다.
- 피험자중 한명이 목 통증을 완화시키는 몇가지 운동을 보여 주면, 다른 한명이 그것을 따라 한다.

### I.4 음향과 영상 신호의 동기화를 평가하기 위한 과제

다음 과제는 음향과 영상신호간 동기화에 대한 평가자들의 주의를 이끌기 위한 것이다.

- 1) 동작과 소리의 동기화에 대해 피험자 한명은 손뼉을 치고, 다른 한명은 그것을 체크한다.

과제 1은 입술 동기화(lip synchronization)를 평가하기에는 적합하지 않으며, 2개의 신호간 동기화를 평가하기 위한 다른 과제들은 연구중이다.

## 부록 II. 대화용 자극을 위한 프로토콜

### II.1 이름-사고 과제를 위한 프로토콜

[12]에서 이름-사고 과제를 위해 다음 프로토콜이 사용되었다.

“첫번째 과제인 이름-사고 과제는 고정된 프로토콜에 따라 수행되는 질문-대답 게임이다. 이 과제에서 피험자중 한명은 3가지 정보를 받는다: 첫째 생각할 것이 상표인지 (유명한) 사람의 이름인지를 나타내는 상표명(brand) 또는 사람(person)이라는 단어; 둘째, 상표인 경우는 제품에 대한 기술, 사람인 경우는 직업; 셋째 생각해야 할 이름. 따라서, 예를 들면, 피험자는 다음을 제시받는다.

- 상표명
- 담배
- Camel

상표명이나 성명을 생각할 때, 두번째 피험자는 다음 프로토콜에 주목해야 한다: 첫번째 질문은, 그 것은 상표명인가 사람인가?; 두번째, 상표명일 경우 무슨 제품인가?, 그리고 사람인 경우 그 사람의 직업은 무엇인가? 이때, 이름의 연속적인 철자를 요청할 수 있다. 이것은 이름을 정확히 생각할 때까지 또는 전체 이름의 철자가 완성될 때까지 계속된다. 질문받은 피험자는 대화하는 동안 문장을 볼 수 있지만, 대부분의 피험자는 문장에 담겨진 제한된 양의 정보를 기억하는데 문제가 없기 때문에 그럴 필요성을 못느낀다. 이름이 생각난 후에, 전체 소요 시간과 요청된 철자수가 기록된다. 길이와 난이도가 가변적인 몇가지 이름이 실험조건에 있었다. 데이터를 선형회귀분석(linear regression analysis)하여 직접 사고(요구한 철자가 없는 경우)에 걸린 시간과 철자당 추가 소요시간에 대한 추정치가 얻어진다.”

### II.2 이야기-비교 과제에 대한 프로토콜

[12]에서 이야기-비교 과제를 위해 다음 프로토콜이 사용되었다.

“두번째 과제에서, 피험자들은 한가지 이야기의 두가지 다른 버전간에 차이수를 탐색해야 한다. 그들에게는 무제한적인 대화가 주어지며, 실제 실험에 앞서 두명의 피험자는 단편적인 이야기를 읽고 기억해야 한다. 그들에게는 기본적으로는 같지만 몇가지 확실한 차이가 있는 이야기들이 주어진다 (예를 들어, 약 200 단어로 구성된 이야기에는 6개의 차이가 있음). 또한, 이야기에 대한 기억을 향상시키기 위한 질문목록을 받고 스스로 대답해야 한다. 암기시간 후에, 피험자들은 두 이야기간 모든 차이를 가능한 한 빨리 찾기 위해 대화를 개시해야 하며, 대화중에 피험자들은 문장에 관해 상의할 수는 없다. 피험자들은 얼마나 많은 차이가 있는지를 알아야 하고, 차이가 발견되고 얼마만큼의 차이를 그들이 찾았는지 서로 피드백 한다. 대화는 모든 차이를 찾을 때까지 또는 특정시간내에 더 이상의 차이를 찾을 수 없을 때까지 지속된다.

대화 개시시간과 차이발견에 걸린 시간이 기록된다. 분석에서 개시시간과 첫번째 차이 발견시간 사이의 시간간격, 그리고 후속 발견과의 시간간격들이 도출된다. 이들 시간간격들은 지속시간에 따라 순서화 된다. 시간간격, 즉, 다음 차이를 탐지하는데 걸리는 시간은 피험자쌍이 사용한 책략과, 차이가 문장의 어디에 위치하고 있었는지에 의존하기 때문에 이러한 데이터의 측정이 필요하다. 시간간격을 재순서화하는 것은 이러한 요인들에 의해 일어나는 분산(variation)을 감소시킨다. 또한, 시간간격의 지속시간에 대한 함수로서 연구하고자 하는 효과에 대한 평가가 가능하다.

### II.3 그림-비교 과제에 대한 프로토콜

[12]에서 그림-비교 과제를 위해 다음 프로토콜

이 사용되었다.

“이 과제에서 피험자는 그림을 기억하고 그들에게 주어진 그림이 동일한지 아닌지를 결정해야 한다(풍경, 건물 및 도심 등 다양한 주제의 그림들이 사용될 수 있다). 대화는 제한이 없으나, 피험자들은 대화중 그림을 볼 수 없다. 채점되는 결과들은 피험자들이 결정을 내리고, 결정이 맞는지 또는 틀리는지에 대해 걸린 전체 시간이다. 피험자들은 일반적으로 그림이 다를 때보다 같을 때 좀 더 시간을 소비하기 때문에, 두경우에 대한 결과는 분리해야 한다. 비록 이 과제가 다소 이야기-비교 과제와 유사하다고 해도, 자유 대화가 일어나는 정도의 차이가 있다. 이야기를 비교할 때 피험자는 문장내의 문구를 인용하는 경향이 있는 반면, 그림을 비교할 때는 그들 자신의 언어만을 이용하기 때문이다.”

### 부록 III. 실험조건 질문

각 실험조건 후에 사용되는 질문:

실험자: 당신들이 방금 나누었던 대화중 접속(connection)에 대한 당신들의 오피니언과 관련된 몇 가지 질문을 할 것입니다. 해당되는 것에 원을 표시해 주시기 바랍니다. 준비되었습니까?

1) 전체 AV 품질은 어떻습니까?

매우 좋다	좋다	보통이다
나쁘다	매우 나쁘다	

2) 접속의 영상품질은 어떻습니까?

매우 좋다	좋다	보통이다
나쁘다	매우 나쁘다	

3) 접속의 음향품질은 어떻습니까?

매우 좋다	좋다	보통이다
나쁘다	매우 나쁘다	

4) 상대방의 말을 끊는데 얼마만큼의 노력이 들었다고 생각하십니까?

노력없음      적은 노력      보통 노력

상당한 노력      매우 노력

5) 접속동안 어떤 어려움은 없었습니까?

예      아니오

6) 접속은 수용할만 합니까?

예      아니오

실험자: 감사합니다. 이제 몇분 동안 실험을 재정비하고 실험을 계속할 예정입니다. 접속이 되는 즉시, 대화할 수 있음을 알려드리겠으니, 상대방과 대화하시기 바랍니다.

#### 부록 IV. 마무리 질문

실험 끝에 사용될 부가적인 질문:

실험자: 실험이 끝났기에, 몇가지 부가적인 질문을 드리겠습니다. 준비되었습니까?

당신이 생각하기에 다음의 요인들중 통신품질의 향상에 중요하다고 생각되는 순서로 요인의 번호를 말씀해 주시기 바랍니다(피험자는 여기에 있는 요인들의 번호를 읽을 것이다).

7a    7b    7c    7d

전체 실험에 관해 저희에게 하고 싶으신 말씀이 있으면 해주시기 바랍니다.

---



---



---



---

실험에 참가해 주셔서 대단히 감사합니다. 실험실

을 나가셔도 좋다고 알려드릴 때까지 잠시만 기다려 주십시오.

## IV. 결 론

본 고에서는 멀티미디어 서비스에 대한 주관 품질평가의 목적과 ITU-T SG 12 Question 22/12에서 권고화를 추진하고 있는 AV 통신에 대한 양방향 주관 품질평가법에 관해 고찰하였다. 미래의 21세기에는 많은 미디어가 출현함과 동시에 질적인 면에서의 변화가 일어남으로서, 멀티미디어의 서비스 품질 측면에서 이제까지 염두에 두지 못했거나 해결하지 못했던 문제들이 제기될 것으로 생각한다. 따라서, 그러한 멀티미디어에 대한 서비스 품질요인들의 정확한 변별과 그 요인이 사용자에게 미치는 영향을 수량화할 표준화된 주관 품질평가법 연구는 반드시 필요하다. 본 고에서 살펴본 AV 통신에 대한 양방향 주관 품질평가법은 미래의 멀티미디어 서비스 품질평가를 위한 기초가 되며, 서비스 품질요인들이 확인될 때마다 적절한 권고의 수정이 이루어져야 할 것이다.

## 참고문헌

- [1] 北脇信彦, “マルチメディアの品質評価を目指して,” NTT R&D, vol. 42, no. 10, 1993.
- [2] 김정환, “국제회의참가 종합보고서 - ITU-T SG12 회의,” 한국통신기술협회, '95-ITU-T-01호, 1995.
- [3] ITU-T COM 12-R 20-E, Study Group 12 - Report R 20, Geneva, Sep. 1995.
- [4] ITU-T TD 57, Draft ITU-T Recommendation P.VQ: Subjective video quality assessment methods for multimedia applications, Geneva, Sep. 1995.
- [5] ITU-T TD 58, Draft ITU-T Recommendation P.RISV: Reference impairment system for video, Geneva, Sep. 1995.
- [6] ITU-T TD 56, Draft ITU-T Recommendation P.ITM: Interactive

- test methods for audiovisual communications, Geneva, Sep. 1995.
- [7] ITU-T TD 55, Draft ITU-T Recommendation P.AVQ: Subjective assessment methods for global audiovisual quality evaluation in multimedia applications, Geneva, Sep. 1995.
- [8] ITU-T COM 12-R 28-E, Report of Working Party 2/12 Meeting, Geneva, June 1996.
- [9] ITU-T COM 12-R 31-E, Report of The Fifth Meeting of Study Group 12 - Part A - General, Geneva, June 1996.
- [10] ITU-T COM 12-R 26-E, Draft Recommendation P.920: Interactive test methods for audiovisual communications, Geneva, Jan. 1996.
- [11] ITU-T Recommendation P.80, Methods for subjective determination of transmission quality, 1993.
- [12] A. W. Bronkhorst and J. A. Verhave, "The Effect of audio-video desynchronization on communication efficiency in videotelephony," Study for PTT Telecom Netherlands, TNO Institute for Perception, Report IZF 1992 C-35, 1992.
- [13] R. E. Kirk, Experimental Design, Procedures for the Behavioural Sciences. 2nd Edition, Brooks/Cole Publishing Co., 1982.
- [14] ITU-T COM XII-85, Proposals of Quality Assessment Method Due to Delay and Acceptable Delay Time, NTT, July 1991.