

팩시밀리 화상품질 측정에 관한 연구 (A Study on Testing Image Quality on Facsimile)

권세혁*황건*
(S. Kwon, G. Hwang)

본 연구는 아날로그 신호를 사용하는 공중교환전화망과 접속되는 그룹 3(G3) 팩시밀리의 화상품질을 측정하는 방법을 제시하였다. CCITT(현 ITU-TS) 표준 시험 도표 No.2를 이용하여 전송된 화상에 대한 평가는 설문조사를 통해 평가되었고, 그것들은 MOS 방법에 의해 계량화되었다. 설문지의 결과에 대한 상관 분석을 통해 문항을 하나의 종합 평가 문항으로 줄일 수 있음을 살펴보았다. 그리고 그 점수들의 평균들에 대한 차이를 분석함으로써 팩시밀리 화상품질에 영향을 미치는 요인들의 유의성을 검정하였다. 유의성을 검정하는 방법들로 t 검정법과 Van der Waerden Scores 방법을 제시하였다. 그리고 검정 결과 점수 평균이 유의하지 않은 그룹들을 하나의 그룹으로 하여, 그 그룹에 있어서 점수 히스토그램을 구하였다. 이 히스토그램을 하나의 정규분포 곡선으로 근사시켜 팩시밀리 화상품질 평가치를 살펴 보았다

I. 서 론

현재 그룹 3 팩시밀리를 사용하여 문서를 보내는 경우 아날로그 신호를 사용하고 있다. 한편, 보내고자 하는 문자나 그림등의 화상이 팩시밀리에 입력되면 그 정보는 디지털 신호로 읽혀진다. 그러므로 팩시밀리 안에 내장된 모뎀에 의해 디지털 신호는 아날로그 신호로 변조된 후 전화망에 의해 송신되며, 송신된 신호는 다시 디지털 신호로 복조되어 송신된 화상이 감열방식에 의해 인쇄

된다.

화상을 디지털 신호화하기 위해 주사선 밀도는 8pel(picture element)/mm를 사용하며 각 주사선은 하나의 프레임으로 인식되어 전송된다. 부주사선 밀도도 3.85로 사용하고 있지만 프레임에 포함된 정보의 내용에 따라 각 프레임은 서로 다른 수의 비트로 인식되어 전송된다. 주사선과 부주사선으로 등분된 한 펠(pel)은 하나의 정보로 인식되어 디지털 부호화하지만 같은 정보가 반복되면 송신시간 절약을 위해 독자적 부호 방식을 가지므로 같은 길이의 프레임이라도 디지털 부호화한 후의 비트 수는 상이하게 된다.

*기술기준연구실 선임연구원

일반적으로 BER(Bit Error Ratio)은 전화망에 있어서 품질을 측정하는 중요 지표가 되지만, 팩시밀리에 있어서는 비트의 에러보다는 프레임에서의 에러, 즉 FER(Frame Error Ratio)을 고려한다. 프레임 내에서 하나의 비트에 에러가 생겨도 그 프레임은 에러가 발생한 것으로 간주되지만, 비트의 에러 발생이나 프레임 에러가 발생하더라도 화상이나 문자를 판독하는 데는 문제가 없을 수 있다. 한 프레임 내의 다수의 비트 에러 발생, 혹은 다수의 프레임 에러가 발생하면 판독자는 화상의 차이를 느끼게 된다. 현재 CCITT 권고에 의하면 다음의 경우 전송된 페이지는 판독이 어려운 화상으로 판정하고 있다.

- 1) 연속적으로 4개 이상 프레임 에러들의 발생이 한 번 이상인 경우
- 2) 12개 이상의 프레임 에러가 발생한 경우
- 3) 연속적으로 2개나 3개의 프레임 에러들의 발생이 3번 이상인 경우.

그러나 전송된 화상에 있어서 FER을 측정하는데 어려움이 있으며, 프레임 에러의 형태에 따라 화상품질이 달라진다. 그러므로 본 연구에서는 팩시밀리 화상 품질을 측정하기 위해 FER보다는 주관적 평가 방법의 하나인 MOS(Mean Opinion Score) 방법을 제시하였다. CCITT No.2 시험 화상을 사용하여 실제 전송된 화상을 수집하여 설문지를 통해 응답자들의 평가를 점수화하였다. 그 점수(score)들을 여러 분석 방법을 통해 팩시밀리 화상 품질에 영향을 미치는 요인은 무엇이며, 전체 팩시밀리 화상 품질은 어느 정도인지 CCITT가 권고하는 평가치와 비교하였다.

II. 분석방법

1. 요 인

서론에서 기술한대로 BER이 FER에 영향을 미치며, FER은 화상 품질에 영향을 미치므로 BER에 영향을 주는 요인들을 팩시밀리 화상 품질에 영향을 미치는 요인들로 고려할 수 있다. 일반적으로 BER에 영향을 주는 요인들로는 다음과 같다.

가. 링크 수

송수신간에 거쳐야 하는 교환기의 수가 많을수록 BER은 커지게 된다. 시내 전화망의 경우 링크는 1이나 2이며, 시외 전화망에서는 3 혹은 4이다. 본 연구에서는 시내망과 시외망간의 차이를 살펴 보았다.

나. 통화량

통화량이 많을수록 BER은 커지므로 통화량이 많은 시간대와 적은 시간대로 나누어 통화량 효과를 분석하였다. 일반적으로 통화량이 많은 시간은 오전 9시부터 오전 12시 그리고 오후 6시부터 오후 9시까지로 간주되므로 본 연구에서도 같은 방법으로 시간대를 나누었다.

다른 요인들로는 교환기종과 전송 선로의 길이나 종류 등을 고려할 수 있으나 영향이 극히 미미하다고 판단되거나 혹은 측정상 어려움이 있어서 팩시밀리 화상 품질에 영향을 주는 요인으로서는 위의 두가지만 고려하였다. 본 연구에서 고려될 두 요인들의 결합 조건을 다음 <표 1>에 정리하였다.

〈표 1〉 요인들의 결합조건

	시내망	시외망
통화량이 많은 시간 (Busy-Hour)	(BH, LO) 그룹 1	(BH, LD) 그룹 3
통화량이 적은 시간 (NonBusy-hour)	(NB, LO) 그룹 2	(BH, LD) 그룹 4

주) LO=시내, LD=시외

2. MOS

MOS 방법은 주관적 평가방법의 하나로 5점 척도를 사용하게 된다. 각 설문 문항의 보기를 매우 좋다, 보통이다, 나쁘다, 매우 나쁘다의 순으로 배열하여 응답자의 평가를 조사한 후, 각각 5그룹 1까지의 점수로 환산하여 계량화시키는 방법이다. 앞 절에서 언급된 두 요인의 결합조건인 각 그룹 g에 대한 모집단 점수의 평균과 분산은 i 를 점수, $P_{i,g}$ 를 g 그룹이 i 점수를 얻을 확률이라고 하면 아래와 같이 정의된다.

$$MOS_g = \sum_{i=1}^5 i P_{i,g}$$

$$VARMOS_g = \sum_{i=1}^5 (i - MOS_g)^2 P_{i,g} \quad (1)$$

요인들의 결합 조건에 따라 CCITT 표준 시험 도표 No.2를 전송하여 평가될 화상을 얻고 이것을 응답자는 설문지를 통해 평가하게 된다. 이 평가된 점수들에 의해 점수의 상대도수를 나타내는 히스토그램들이 각 그룹마다 하나씩 얻어지게 된다. 각 히스토그램들에 의해 구해지는 평균과 분산은 각 그룹의 모평균(MOS_g)과 모분산($VARMOS_g$)을 추정치로 사용하게 된다. 즉, $\hat{P}_{i,g}$ 를 g 그룹의 화상에 대하여 응답자가 i 점수라고 평가했을 비율인

상대도수라 하면 각 그룹의 평균과 분산은 다음과 같이 추정된다.

$$\hat{MOS}_g = \sum_{i=1}^5 i \hat{P}_{i,g}$$

$$\hat{VARMOS}_g = \sum_{i=1}^5 (i - \hat{MOS}_g)^2 \hat{P}_{i,g} \quad (2)$$

3. CCITT 표준 시험 도표 No.2 분석

시험 화상으로 사용된 길이 297mm이고 폭 210mm인 CCITT 표준 시험 도표 No.2는 문자의 판독성에 대한 일반적인 평가용으로 권고되고 있다. 이 시험 도표는 화상품질 측정을 위해 여러가지 사항들을 고려하여 작성되었는데 이것을 정리하면 〈표 2〉와 같다.

〈표 2〉 CCITT 표준 시험 도표 No.2의 측정 항목표

도 표 항 목	측정되는 사항
1 화살표와 선의 4가지 그룹	프레임의 왜곡 정도
2 0.5mm의 교대되는 흑백선	최소 허용 해상도의 시험
3 흑·백색 대역	수평 해상도
4 대각선	시험 패턴 표면의 균일성
5 흑색원·동심원	송수신 장치간의 협동지수 차
6 수평대역	수직해상도
7 문자	문자 판독의 한계

각 도표 항목에 대한 평가를 위해 μm 단위로 응답되는 3번의 흑·백색 대역의 항목을 제외하고는 각 항목의 평가가 5점 척도로 설문지가 작성되었다. 설문지 응답자들이 이 분야의 비전문가들이면 전문적인 지식의 부족으로 인해 비록 질문이 객관식이지만 질문을 어렵게 느끼고 성의없이 응답할 가능성이 높아진다. 그리하여 응답자로 비전문

가를 활용하는 경우, 화상의 종합적인 품질을 묻는 문항만을 가지고 설문을 하여도 같은 결론을 얻을 수 있는가를 상관분석을 통해 분석하였다.

〈표 2〉에 언급된 7개의 문항에 덧붙여서 8번째 문항으로 “종합적인 화상품질은?”이라는 문항을 추가하여 얻은 40개의 설문 결과를 가지고 8번 문항의 점수들과 나머지 문항들의 점수들간의 상관 분석을 하여 〈표 3〉에 정리하였다.

〈표 3〉 CCITT 표준 시험 도표 문항들의 상관 계수

문 항	1	2	4	5	6	7
Pearson 상관계수	-.0332	.4332**	.3964*	.3682*	.3979*	.4044**
Spearman 상관계수	-.0301	.4008**	.3356*	.3493*	.3873*	.4066**

주) *:95% 유의, **:99% 유의

Pearson 상관계수는 보통 흔히 사용되는 상관계수이며, 점수들이 이산변수이므로 자료들의 순위(rank)의 상관계수를 구하는 비모수방법인 Spearman 상관계수도 구하여 분석했다. 3번 문항은 5점 척도가 아니므로 분석 대상에서 제외하였다. 1번 문항이외에는 모든 문항들은 8번 문항과 상관 관계가 유의하므로 8번 문항 하나로 대체될 수 있음을 알 수 있다. 한편, “프레임의 왜곡 정도는?”라고 질문된 1번 문항의 상관계수가 유의하지 않는 것은 문항의 질문 자체가 화살표나 선들에 대한 질문이기 보다는 다소 모호하였기 때문이다. 그러므로 비전문가가 응답자인 경우는 종합적인 품질을 묻는 하나의 문항으로 화상 품질을 측정하여도 모든 도표 문항을 포함한 설문지를 이용한 측정 결과로 대체될 수 있다고 결론지을 수 있다.

4. 유의성 검정

4 그룹들의 모집단의 점수평균(MOS_g)들의 차이에 대한 유의성 검정은 링크의 효과와 통화량의 효과를 알아보는 분산분석도 고려할 수 있으나, 점수가 이산변수이므로 일반적인 t 검정과 비모수 방법인 Van der Waerden Scores를 이용하고자 한다. S²을 두 표본 분산들의 통합분산(pooled variance)이라 하면 i그룹 모집단의 평균과 j그룹 모집단의 평균의 차이에 대한 t 검정통계량은 아래와 같다.

$$t = \frac{(\hat{MOS}_i - \hat{MOS}_j)}{S} \quad (3)$$

Van der Waerden Scores는 서로 다른 그룹들의 위치모수(location parameter)의 차이를 검정하는 비모수 방법의 하나이다. 점수의 평균들은 위치모수이므로 이 방법을 이용하고자 한다. R_j를 j번째 관측치의 순위, a(R_j)를 순위점수(rank score), n을 표본의 크기라고 정의하면, Van der Waerden Scores는 아래 식(4)와 같이 정의되고 이 순위점수들의 평균 점수들에 의해 서로 다른 그룹의 평균의 차이에 대한 유의성을 χ^2 분포를 이용해 검정을 한다. Φ 는 정규분포의 분포함수이다.

$$a(R_j) = \Phi^{-1}\left(\frac{R_j}{n+1}\right) \quad (4)$$

평균에 대한 유의성 검정 결과를 이용하여 평균의 차이가 유의하지 않는 그룹들의 점수 자료에 의해 점수를 수평축, 점수의 상대도수를 수직축으로 하는 점수 히스토그램을 작성할 수 있다. CCITT에서는 평가치를 정규분포 곡선에 근사하여 분석하는 방법을 권고하고 있으므로, 본 연구에서도 다음의 방정식에 의해 정규분포의 평균(μ_k)과

분산(σ_g^2)을 추정하여 점수 히스토그램을 정규분포 곡선으로 근사화 시켰다.

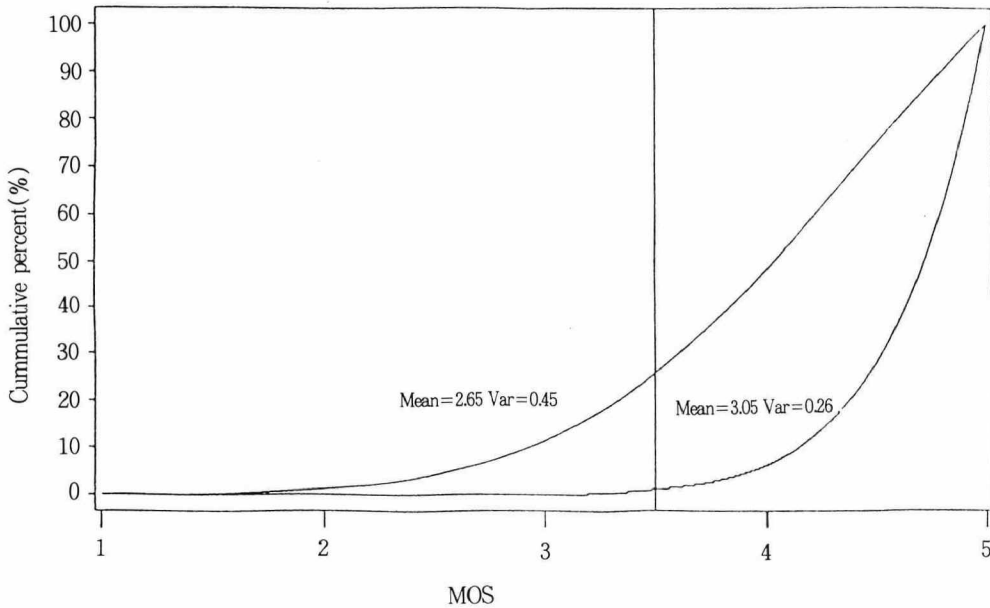
$$\hat{MOS}_g = 5 - \sum_{i=1}^4 \Phi\left(\frac{i+0.5-\mu_g}{\sigma_g}\right)$$

$$VAR\hat{MOS}_g = 25 - \sum_{i=1}^4 (2i+1) \Phi\left(\frac{i+0.5-\mu_g}{\sigma_g}\right) - (\mu_g)^2 \quad (5)$$

III. 화상품질 측정

CCITT 표준 시험 도표를 <표 1>의 결합조건에 의해 실제 선로에서 전송한 후, 10명의 비전문가 응답자를 대상으로 화상들을 설문 조사한 결과 각 그룹의 평균과 분산은 아래 <표 4>와 같이 정리되었다. 각 그룹의 평균들에 차이에 대한 유의성 검

정을 이미 언급한 t 검정과 Van der Waerden Scores 방법을 실시한 결과 그룹1과 그룹2, 그룹3과 그룹4의 평균의 유의성은 없었다. 그룹 1은 그룹 3과 그룹 4와 각각 95%, 85% 유의하였고, 그룹 2는 그룹 3과 그룹 4와 각각 85%, 75% 유의하였다. 그러므로 화상품질에 시내·외의 요인은 어느 정도 영향을 미치나, 통화량에 의한 영향은 유의하지 않으므로 그룹을 시내와 시외로만 나누어 평균과 분산을 구하였다. 시내그룹의 평균과 분산은 각각 2.65와 0.45이었고, 시외그룹의 평균과 분산은 각각 3.05와 0.26이었다. 다시 t 검정과 Van der Waerden Scores방법에 의한 검정 결과 시내의 그룹들의 평균 차이는 95% 유의하였다. 일반적으로 링크 수가 많을수록 화상품질의 평가치는 낮게 되는데 본 연구에서는 링크가 많은 시외망에 있어서 평가치의 평균 점수는 높다는 결론을 얻었다. 시외망 구성이



(그림 1) 시내·외 그룹에 있어서 근사된 정규분포 곡선

시내망보다는 더 우수하기 때문에 링크의 요인보다는 망 구성 요인이 팩시밀리 화상품질에 더 큰 영향을 미쳐 링크수에 의한 영향을 상쇄하였기 때문일 것이다.

〈표 4〉 그룹들의 평균과 분산

	그룹 1	그룹 2	그룹 3	그룹 4
평균	2.7	2.8	3.0	3.1
분산	.4558	.4888	.2222	.3222

시내·외 그룹들의 평균의 차이가 유의하므로 시내와 시외에 대한 점수 히스토그램을 정규분포 곡선에서 근사시키기 위하여 식 (5)에 의해 평균과 분산을 구하고 누적 정규분포 곡선을 그리면 (그림 1)과 같다. 시내 그룹에 있어서 근사된 정규분포의 평균과 분산은 2.65와 0.3665이고, 시외 그룹의 경우는 평균이 3.05, 분산이 0.1946이었다.

IV. 결 론

팩시밀리의 화상품질 측정을 위해 선로상에서 시험 전송하여 얻은 CCITT 표준 시험도표를 종합 평가만을 묻는 설문지를 통해 주관적 평가방법인 MCS 방법을 이용해 분석하였다. 영향을 미치는 요인으로 통화량의 차이와 시내·외망 차이를 검정한 결과 통화량은 유의하지 않았고 시내·외망간

의 차이만 유의하였다. 이 결과를 토대로 시내·외망간의 점수 히스토그램을 정규분포로 근사하여 (그림 1)을 얻었다. CCITT의 품질평가에 대한 권고는 아래와 같다.

- ① 불량 : 고객 80%가 MOS 1.0 이하로 평가
- ② 양호 : 고객 90%가 MOS 2.5 이하로 평가
- ③ 우수 : 고객 80%가 MOS 3.0 이상으로 평가
- ④ 아주 우수 : 고객 80%가 MOS 3.5 이상으로 평가

(그림 1)에 의하면 시내의 경우는 70% 이상이, 시외의 경우는 90% 이상이 MOS 3.5 이상으로 평가하고 있으므로 조사된 선로에 있어서 팩시밀리 화상품질은 아주 우수하다고 결론지을 수 있다.

참 고 문 헌

- [1] CCITT(1988), Terminal Equipment and Protocols for Telematic Services, Series T.21.
- [2] Lehmann, E.L.(1975), Nonparametrics: Statistical Methods Based on Ranks, San Francisco:Holden-Day.
- [3] CCITT(1993), Quality as corrupted by transmission-induced scan errors, Draft Recommendation E.45y.
- [4] Neter, J. and Wasserman W. (1974), Applied Linear Statistical Models, Richard D. Irwin Inc.
- [5] SAS(1992), SAS/STAT User's Guide, SAS Inc.
- [6] Quality of Service Development Group(1992), Connection Quality, QSDG.