

COMMITTEE T1
CONTRIBUTION

Document Number: T1A1.5/94-147

STANDARDS PROJECT: Analog Interface Performance Specifications for Digital Video
Teleconferencing/Video Telephony Service

TITLE: Brief Statistics for 25 Scenes Rated for Video Quality

ISSUE ADDRESSED: Standard Scenes for Testing Video Quality

SOURCE: GTE Laboratories, Incorporated
(Gregory W. Cermak)

DATE: September 20, 1994

DISTRIBUTION TO: T1A1.5

KEYWORDS Video Quality, Test Scenes, Subjective Quality, Statistical
Measures

Notice

This contribution contains information that was prepared to assist Committee T1 and specifically Technical Subcommittee T1A1 in their work program. This document is submitted for discussion only, and is not to be construed as binding on GTE. Subsequent study may lead to revision of the details in the document, both in numerical value and/or form, and after continuing study and analysis GTE Telephone Operations specifically reserves the right to change the contents of this contribution.

Brief Statistics for 25 Scenes Rated for Video Quality

This contribution is intended as a brief guide to the use of the 25 scenes that had been tested and calibrated as part of the T1A1.5 program for developing measures of video quality. The scenes themselves are not remarkable; the data collected about these scenes, however, make them a unique resource. For further information about the testing program and its results, consult Morton (Subjective Test Plan, T1A1.5/94-118) and Cermak & Fay (T1A1.5 Video Quality Project: Preliminary Results, T1A1.5/94-131).

Scenes. Briefly, 25 scenes had been chosen to represent five categories that are typical of video telephony, video teleconferencing, and entertainment video. These categories cover a wide range of movement and detail. Each scene was edited to a length of nine seconds (plus three seconds of leader to allow the coding system to stabilize). Members of the committee had donated the video footage. The categories:

- A. One person, mainly head and shoulders (scenes f,j,k,l)
- B. One person with graphics and/or more detail (scenes a,b,e,m,n,w)
- C. More than one person (scenes d,g,o,p,q,r)
- D. Graphics with pointing (scenes c,s,t,u,v,x)
- E. High object and/or camera motion (scenes h,i,y)

HRC's. Each scene was processed through each of 25 actual video systems (Hypothetical Reference Circuits, or HRC's) chosen to span a broad range of video quality. The 25 HRC's represented the following categories (see Morton, 1994, for a fuller description of the individual HRC's). These systems cover a range of (nominal) bit rates from 112 Kb/s to 70 Mb/s. Categories:

- High quality
- Vector quantization, medium rate
- Proprietary, low to medium rate
- Proprietary, medium to high rate
- QCIF, low rate
- QCIF, medium rate
- CIF, low rate
- CIF, medium rate
- CIF, high rate

Ratings. 114 judges rated each scene;¹ the judges were from three geographically distinct areas, Boston, Boulder, and Washington, D.C. Roughly half of the sample of judges was experienced with video teleconferencing.² Each presentation of a scene consisted first of the scene in its original form, then the scene as processed through one of the 25 HRC's. The judge rated the processed scene on the following five-point scale taken from CCIR Recommendation 500-5:

5. Imperceptible (difference from the original)
4. Perceptible, but not annoying
3. Slightly annoying
2. Annoying
1. Very annoying

Results. For each scene the table shows the mean rating on the five-point scale above and the standard deviation of the ratings. The more important statistic is the standard deviation because it measures the ability of the particular scene to discriminate among video systems -- the larger the standard deviation, the greater the discrimination. The mean rating is of some interest: Scenes rated near 5 or 1 show especially poor discrimination in the sense that they make nearly all HRC's look equally good or bad.

The table is broken down into 10 columns. The scene is designated by lower case letter (see Morton, 1994, for the "names" of the scenes). The category, denoted by uppercase letter, corresponds to those listed above. The mean (Mean) and standard deviation (Std) are given for each of four classes of HRC's, so the scenes are calibrated with respect to classes of HRC's and not just in general. The first class is all 25 HRC's (All). The second class is 10 HRC's with nominal bit rates below 300 Kb/s (<300K). The third class is nine HRC's with nominal bit rates between 300 Kb/s and 800 Kb/s (300-800K). The fourth class is six HRC's with nominal bit rates above 800 Kb/s (>800K).

By using the results below, one can choose scenes better able to discriminate among the systems being tested. Also, one might decide not to test two scenes with nearly identical results in the table (e.g., f and k, m and n, w and x) if resources for testing are constrained.

¹ Each judge was exposed to all combinations of the 25 scenes with 10 of the 25 HRC's. The 25 HRC's were all covered by each testing laboratory.

² The results were not significantly affected by judges' experience or by the particular laboratories performing the testing, so the results presented here for the 25 scenes should be representative for quite a general population.

T1A1.5 Scenes Guide

Scene	Category	All		<300K		300-800K		>800K	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std
a	B	3.65	1.13	3.01	0.98	3.90	1.04	4.60	0.62
b	B	3.15	1.09	2.55	0.91	3.27	0.89	4.23	0.83
c	<i>washdc</i> D	2.67	1.17	2.06	0.94	2.74	0.91	3.89	1.04
d	<i>SINGTON</i> C	3.01	1.15	2.30	0.91	3.22	0.90	4.18	0.87
e	B	2.80	1.29	1.84	0.85	3.13	0.97	4.28	0.84
f	<i>vtc/nw</i> A	3.67	1.05	3.15	0.94	3.81	1.00	4.54	0.65
g	C	3.34	1.19	2.55	0.97	3.65	0.98	4.49	0.66
h	E	2.65	1.16	2.05	0.88	2.65	0.94	3.96	0.98
i	E	2.09	1.25	1.40	0.70	2.00	0.90	3.76	1.19
j	<i>ssie</i> A	3.25	1.16	2.38	0.89	3.62	0.84	4.46	0.61
k	A	3.65	1.04	2.99	0.88	3.90	0.86	4.60	0.66
l	A	3.59	1.07	2.95	1.00	3.85	0.83	4.53	0.64
m	B	2.54	1.28	1.65	0.74	2.67	0.91	4.25	0.88
n	<i>swing 2</i> B	2.50	1.29	1.64	0.78	2.57	0.92	4.22	0.92
o	C	2.76	1.22	1.99	0.86	2.96	0.95	4.05	1.07
p	C	3.12	1.15	2.37	0.91	3.38	0.90	4.26	0.79
q	C	2.31	1.31	1.38	0.66	2.48	0.97	4.00	1.08
r	C	2.72	1.36	1.62	0.74	3.12	0.97	4.36	0.82
s	D	2.22	1.25	1.52	0.75	2.25	1.03	3.70	1.18
t	D	2.57	1.26	1.72	0.83	2.75	0.93	4.06	0.97
u	D	3.51	1.14	2.81	0.98	3.76	1.01	4.55	0.59
v	D	2.69	1.33	1.71	0.85	3.01	1.02	4.24	0.84
w	B	3.08	1.14	2.31	0.88	3.33	0.88	4.30	0.73
x	D	3.07	1.19	2.26	0.96	3.37	0.90	4.30	0.72
y	E	2.66	1.33	1.66	0.83	2.98	1.01	4.23	0.83