COMMITTEE T1
CONTRIBUTION

Document Number T1A1.5/92-113

****

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

****

TITLE: Spatial and Temporal Information Measures - Test Scene Evaluation

****

ISSUE ADDRESSED: Video Quality

****

SOURCE: NTIA/ITS - Arthur Webster

****

DATE: March 30, 1992

****

DISTRIBUTION TO: T1Q1.5

****

KEYWORDS: Video Teleconferencing, Video Telephony, Video Quality, Subjective Quality, Objective Quality

****

DISCLAIMER:

****

Spatial and Temporal Information Measures - Test Scene Evaluation

1. Introduction

Given that the perceived quality at the output of a video system is a function of the input to that system, it becomes necessary to evaluate the relative information content of video test scenes. This contribution will present a method that can be used to determine two measures of information content of a video test scene: one is a spatial information measure and the other is a temporal information measure. See Reference [1] for background material concerning spatial and temporal information measures.

First, the measures for spatial and temporal information will be presented. Then it will be shown how these measures can be used to determine the location of scenes in the spatial-temporal information matrix given in Table 1, T1Q1.5/92-107 R1.

2. A Spatial Information Measure

The spatial information measure (SI) used is based on the Sobel filter [2]. Each video frame (luminance plane) is filtered with the Sobel operators. The standard deviation over the pixels in each Sobel-filtered frame is then computed. Since the Sobel operator enhances edges, those images with more and/or sharper edges (i.e. those images that contain more spatial information) will produce higher values of the metric. This operation is repeated for each frame in a video sequence and results in a time series of spatial information values. For the NTIA set of test scenes, each of which is 9 seconds long (or 270 frames), we chose the maximum value in the time series to represent the spatial information content of the scene.

$$SI = MAX_{time}\ of\ STD_{space}\ of\ Sobel(VideoSequence)$$

3. A Temporal Information Measure

This class of temporal measures is based upon the motion difference feature, $M_{ij}(n)$, which is the difference between the pixel values (of the luminance plane) at the same location in space but at successive times or frames. $M_{ij}(n)$ as a function of time (n) is defined as,

$$M_{ij}(n) = I_{ij}(n) - I_{ij}(n-1)$$

where $I_{ij}(n)$ is the pixel of image I at the ith row, jth column, and nth frame in time.

The measure of temporal information content, TI, is computed as the maximum over n (time) of the standard deviation of $M_{ij}(n)$ over all i and j (space).

$$TI = MAX_{time}\ of\ STD_{space}\ [M_{ij}(n)]$$

More motion in adjacent frames will result in higher values of TI.

4. Spatial-Temporal Matrix

Figure 1 shows how the set of 36 test scenes used in the NTIA video quality experiments can be placed on a spatial-temporal information matrix. Along the TI=0 axis (x-axis) are found the still scenes and those with very limited motion. Along the SI=0 axis (y-axis) are found scenes with minimal spatial detail. Note that values are normalized for this data set such that the highest and lowest values are placed at 0 and 1 respectively.

As the distance from the origin, $R=(SI^2 + TI^2)^{1/2}$, increases, the total information content increases. This results in increased coding difficulty and may result in increased distortion for a fixed-rate digital system. See Table 1.

Figure 2 shows how the set of 36 test scenes depicted above might be grouped into the spatial-temporal matrix on page 18 of T1Q1.5/92-107, i.e. a matrix divided into 4 levels of spatial performance and 3 levels of temporal performance. Presumably, for a video system to qualify at a given level, it must be tested with scenes that exercise that level of performance.

5. Conclusions

We have discussed some of the information content measures (SI and TI) that have been used in determining the location of a video scene in the spatial-temporal information matrix (as given in Table 1, T1Q1.5/92-107-R1). The location of the video scene within the spatial-temporal matrix is important because the quality of a transmitted video scene may be highly dependent on this location.

The SI and TI measures may be used to assist the T1A1.5 VTC/VT sub-working group with scene selection for the upcoming H.261 tests. SI and TI can be used to assure uniform coverage of the spatial-temporal matrix, as well as assuring that each performance block in the spatial-temporal grid on page 18 of T1Q1.5/92-107 is adequately stressed.

The work presented in this contribution is the result of the efforts of the Video Quality Project team (which is part of the System Performance Standards Group at the Institute for Telecommunication Sciences) consisting of Steve Wolf, Arthur Webster, Steve Voran, Coleen Jones, Margaret Pinson, and Paul King. The measures presented here complement work on other video quality measures that have been presented to the T1Q1.5 VTC/VT sub-working group.

6. References

[1] T1Q1.5/92-13, Webster, A. & Wolf, S., Spatial and Temporal Information Measures for Video Quality, January 22, 1992.

[2] Gonzalez, R. & Wintz, P., Digital Image Processing, 2nd Edition, Addison Wesley, 1987, p.336-8.

Figure 1. Scene Placement - Spatial-Temporal Matrix

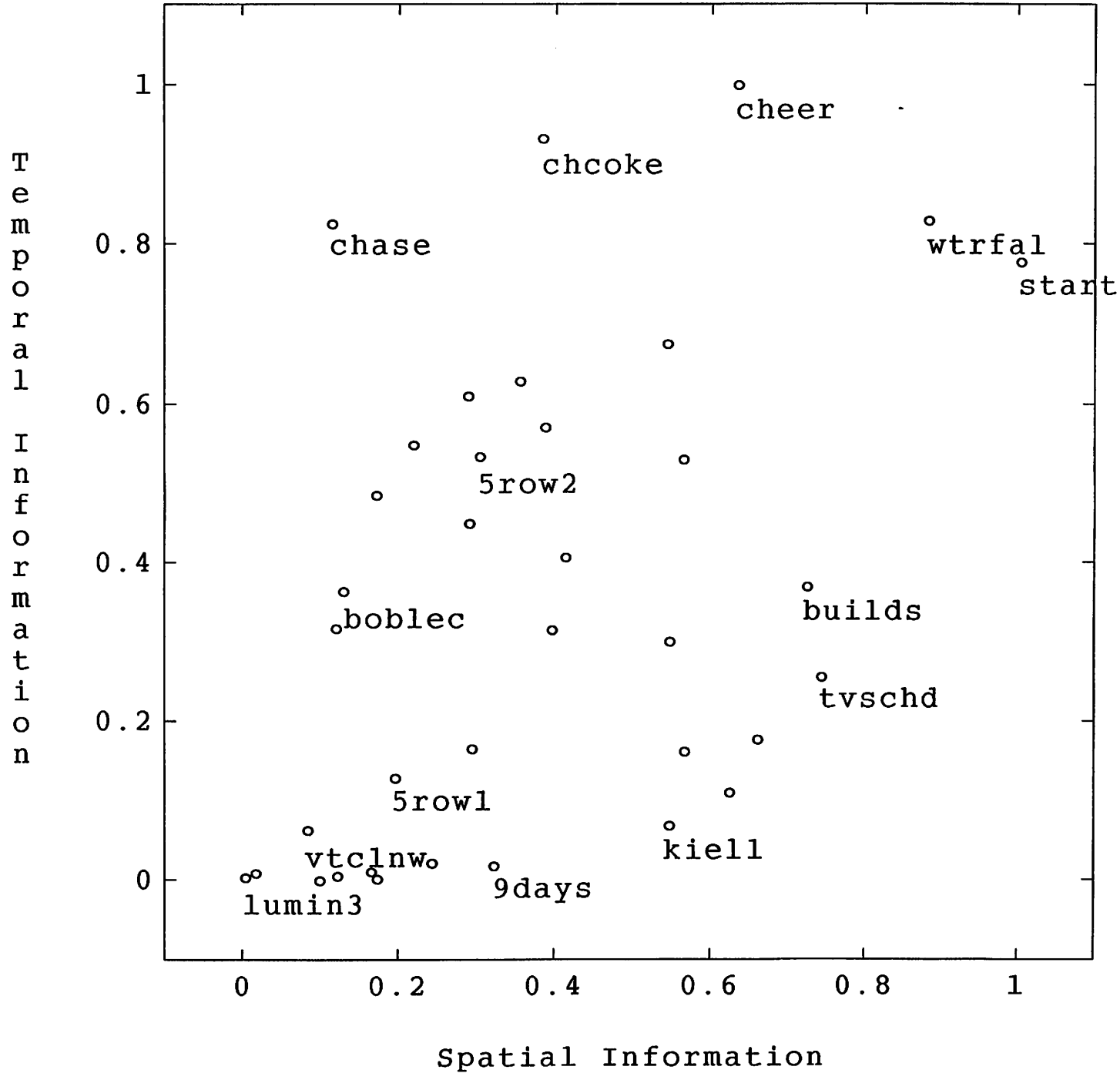

TABLE 1.

Spatial Information & Temporal Information

Scene numbers correspond to Figure 2.

#	name	SI	TI	R	Description
1	lumin3	.00	.00	.00	Luminarios III
2	moontw	.01	.01	.02	Moon and Twigs
3	acehrt	.10	.00	.10	Ace of Hearts
4	vtc1nw	.08	.06	.10	VTC1 Take 2 News Story
5	schlst	.12	.01	.12	School of Athens Still
6	peolak	.16	.01	.16	People at the Lake
7	shutle	.17	.00	.17	Space Shutle
8	5row1	.19	.13	.23	5 People in a Row I
9	vtc2mp	.24	.02	.24	VTC2 - MP + Map
10	9days	.32	.02	.32	Calendar - 9 Days
11	iceck1	.29	.17	.33	Icecicle
12	schlzm	.12	.32	.34	School of Athens Zoom
13	boblec	.13	.36	.39	Bob's Lecture
14	mounts	.39	.32	.50	Two Mountains
15	truck	.17	.48	.51	Truck
16	lanite	.29	.45	.53	Los Angeles at Night
17	kiel1	.54	.07	.55	Kiel Harbor I
18	vtc2zm	.41	.41	.58	VTC2 Zoom w/Pointer
19	3inrow	.21	.55	.59	3 People in a Row - Two Pans
20	lake1	.56	.16	.59	Lake I
21	5row2	.30	.53	.61	5 People in a Row II
22	washdc	.54	.30	.62	Washington DC Map w/Pencil
23	mntbik	.62	.11	.63	Mountain Bike Shop I
24	steve1	.28	.61	.67	Steve1 - The Box
25	crswrd	.66	.18	.68	Crossword Puzzle
26	1duck	.38	.57	.69	One Duck (Close-up)
27	wtrbub	.35	.63	.72	Water Bubbling
28	2ducks	.56	.53	.77	Two Ducks
29	tvschd	.74	.26	.78	TV Schedule
30	builds	.72	.37	.81	Grand Prix I (Buildings)
31	chase	.11	.83	.83	Chase (Dragnet)
32	flowr2	.54	.67	.86	Flowers2
33	chcoke	.38	.93	1.01	Cherry Coke on Mountain
34	cheer	.63	1.00	1.18	Cheerleaders
35	wtrfal	.88	.83	1.21	Waterfall
36	start	1.00	.78	1.27	Grand Prix III (Start)

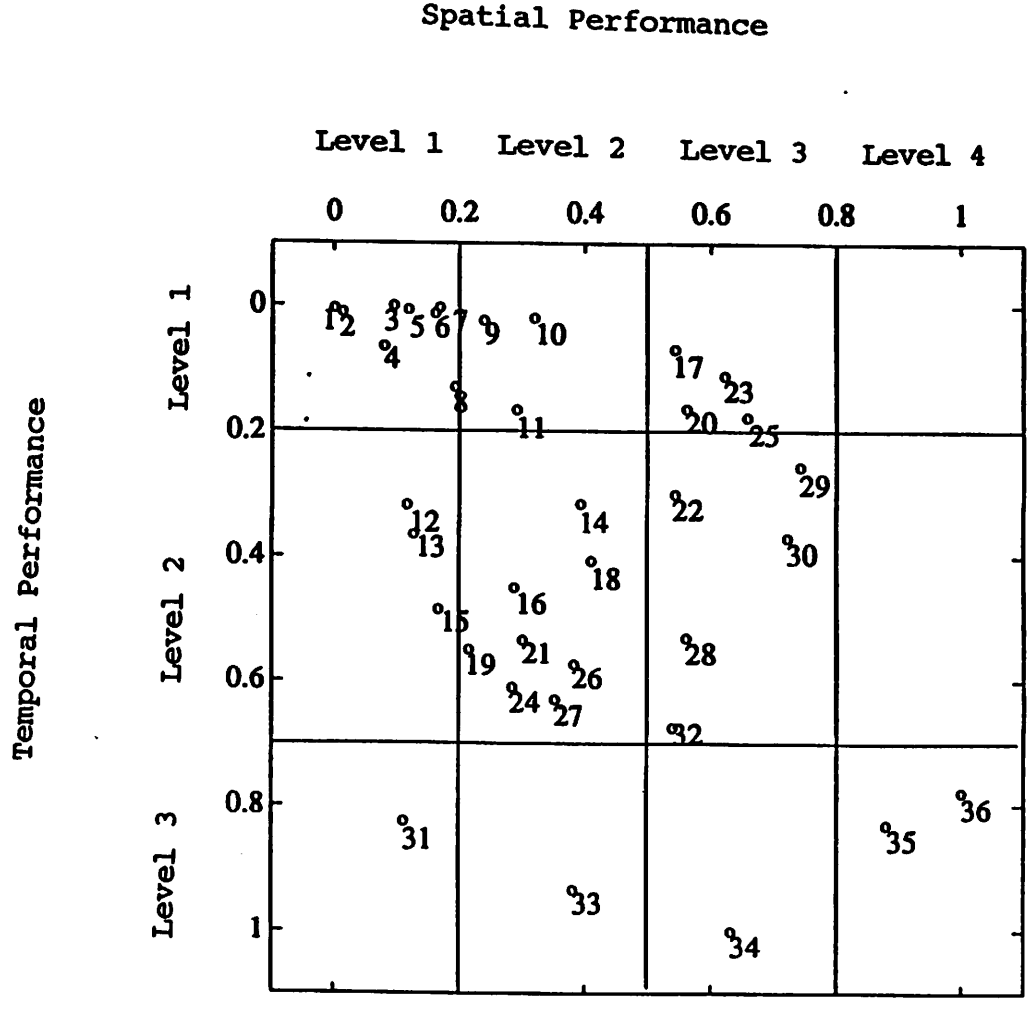

Figure 2. Scene Placement - Spatial-Temporal Performance Grid