

COMMITTEE T1
CONTRIBUTION

Document Number T1A1.5/94-154

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

TITLE: Proposed US Contribution to ITU-T SG12, Q22, Regarding
Two Criteria for Video Test Scene Selection

ISSUE ADDRESSED: Video Quality

SOURCE: Arthur Webster NTIA/ITS

DATE: October 3, 1994

DISTRIBUTION TO: T1A1.5

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

Introduction

T1A1.5 has done extensive work in developing measures of spatial information (SI) and temporal information (TI) to facilitate the selection of test scenes for T1A1.5 programs. T1A1.5 used the SI and TI measures during the process of test scene selection for the recent VTC/VT subjective and objective viewing test.

ITU-T Study Group 12 is currently seeking methods to assist in the selection of video test scenes for subjective testing of multimedia systems. The following proposed US contribution to ITU-T SG12, Q22 will provide the ITU-T with the benefits of T1A1.5's experience in the area of test scene selection.

This contribution partially fulfills an action item undertaken by NTIA/ITS at last February's meeting in Geneva and also answers a request from the Rapporteur of Q22 to provide a method for measuring spatial and temporal content of video test scenes.

The work presented herein was previously approved for inclusion in the T1A1.5 VTC/VT Draft Standard. The bulk of the contribution is taken verbatim from the T1A1.5 VTC/VT Draft Standard T1A1.5/94-107. Editing was done to remove references specific to T1A1.5.

We request that T1A1.5 approve this document as a proposed U.S. Contribution to the November 1994 meeting of ITU-T Study Group 12.

International Telecommunications Union
Telecommunications Sector
(ITU-T)
Study Period 1993-1996

Document:
Date: 28 November - 8 December, 1994
Original: English

STUDY GROUP 12 CONTRIBUTION

Question: 22/12

SOURCE: USA* (Proposed)

TITLE: TWO CRITERIA FOR VIDEO TEST SCENE SELECTION

ABSTRACT:

This contribution describes some important considerations relevant to video test scene selection and offers two measurements that can be used during such selection. Measurements of spatial information content (SI) and temporal information content (TI) are presented and their use as criteria for test scene selection is demonstrated.

We recommend that this contribution be included in Recommendation P.VQ in Section 4.3.

* Contact Person - Name: Mr. Arthur Webster

Tel: +1 303 497 3567
Fax: +1 303 497 5323
e-mail: webster@its.bldrdoc.gov

1 Introduction

This contribution presents two criteria which can be used in video test scene selection. Measurements are defined which can be used to characterize the relative coding difficulty among a group of test scenes. The spatial content and temporal content of test scenes are not the only criteria which can be useful in test scene selection. They are, however, the most important perceptually as well as being simple to calculate. While some issues remain regarding the use of these criteria in comparing scenes of different video formats, the basic measurements presented herein are a useful start.

We recommend that the following text be included in Recommendation P.VQ in Section 4.3.

2 Test Scene Selection.

The selection of test scenes is an important issue. In particular, the spatial and temporal information content of the scenes are critical parameters. These parameters play a crucial role in determining the amount of video compression that is possible, and consequently, the level of impairment that is suffered when the scene is transmitted over a fixed-rate digital transmission service channel. Figure 1 shows the relative amounts of spatial and temporal information for some representative test scenes. Fair and relevant video test scenes must be chosen such that their spatial and temporal information content is consistent with the video services that the digital transmission service channel was intended to provide. The set of test scenes should span the full range of spatial and temporal information content of interest to users of the devices under test.

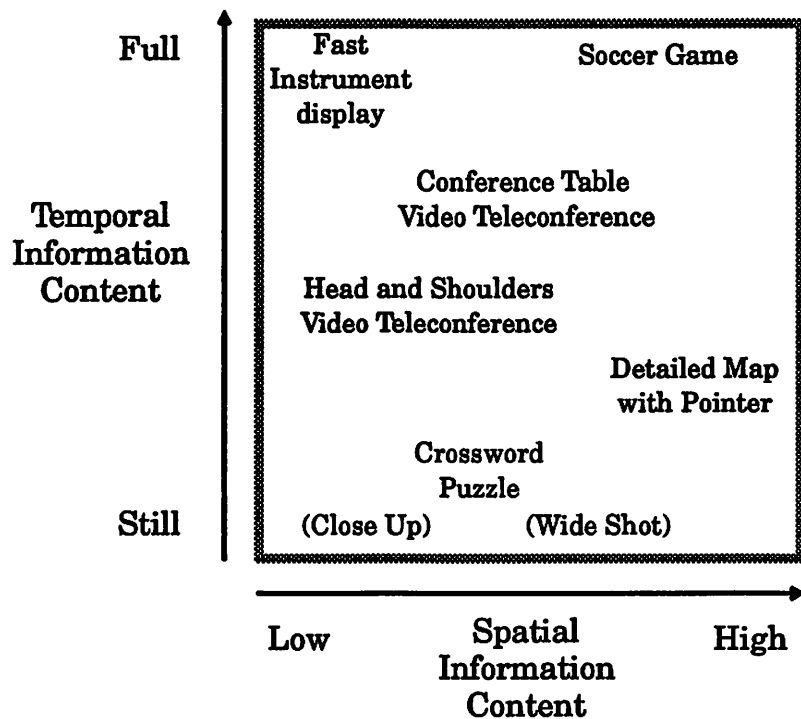


Figure 1 Information Content of Test Scenes

3 Spatial and Temporal Information Measures.

This section presents methods for quantifying the spatial and temporal information content of test scenes. These methods for evaluating the spatial and temporal information content of test scenes are applicable to video quality testing—both now and in the future. The location of the video scene within the spatial-temporal matrix is important because the quality of a transmitted video scene (especially after passing through a low bit rate codec) is often highly dependent on this location. The spatial and temporal information measures presented here can be used to assure appropriate coverage of the spatial-temporal matrix.

3.1 Spatial Information Measurement

The Spatial Information content, SI, is based on the Sobel filter. Each video frame (luminance plane) at time n (F_n) is first filtered with the Sobel filter ($Sobel(\bullet)$). The standard deviation over the pixels (std_{space}) in each Sobel-filtered frame is then computed. This operation is repeated for each frame in the video sequence and results in a time series of spatial information values. The maximum value in the time series (max_{time}) is chosen to represent the spatial information content of the scene. This process can be represented in equation form as,

$$SI = max_{time} \{ std_{space} [Sobel (F_n)] \} \quad (1)$$

3.2 Temporal Information Measurement

The Temporal Information content, TI, is based upon the motion difference feature, $M_n(i,j)$, which is the difference between the pixel values (of the luminance plane) at the same location in space but at successive times or frames. $M_n(i,j)$ as a function of time (n) is defined as,

$$M_n(i,j) = F_n(i,j) - F_{n-1}(i,j) \quad (2)$$

where $F_n(i,j)$ is the pixel at the i^{th} row and j^{th} column of the n^{th} frame in time.

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

$$TI = max_{time} \{ std_{space} [M_n(i,j)] \} \quad (3)$$

More motion in adjacent frames will result in higher values of TI. Note: For scenes that contain scene cuts, two values may be given: one where the scene cut is included in the temporal information measure, and one where it is excluded from the measurement.

4 Example

Figure 2 shows how a set of test scenes can be placed on a spatial-temporal information matrix. Along the TI=0 axis (along the bottom of the plot) are found the still scenes and those with very limited motion (such as l, f, and a). Near the top of the plot are found scenes with a lot of motion (such as p, q and i). Along the SI=0 axis (at the left edge of the plot) are found scenes with minimal spatial detail (such as l, k, x, u, and f). Near the right edge of the plot are found scenes with the most spatial detail (such as h and s). The values of SI and TI were obtained using the above equations and video which has been spatially sampled according to CCIR 601 specifications.

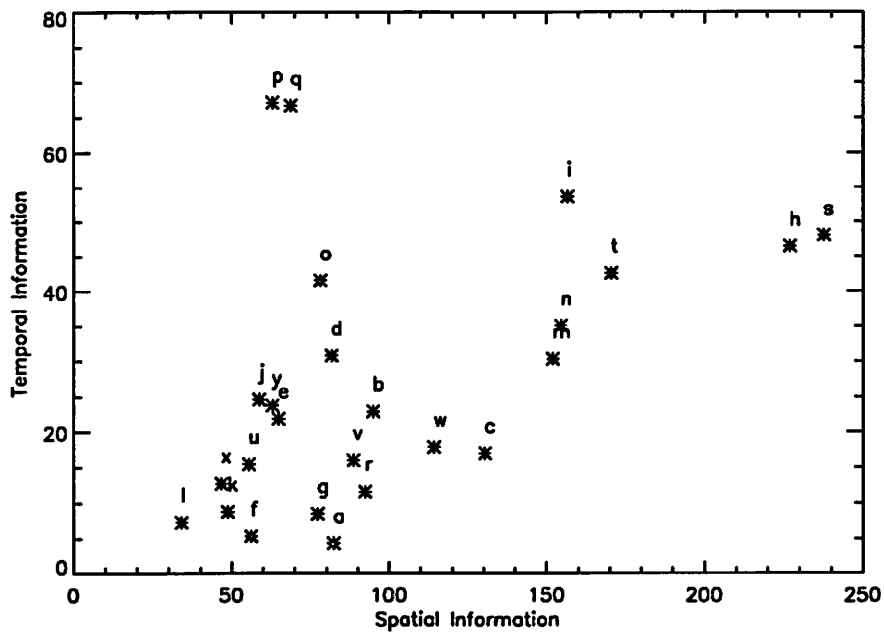


Figure 2 Spatial-Temporal Plot for Test Scene Set