|  |  |  |  |
| --- | --- | --- | --- |
| **Question(s):** | 12/9 | **Meeting, date:** | VQEG, 8-12 July 2013  |
| **Study Group:** | JRG-MMQA | **Working Party:** | 1/9 | **Intended type of document** (R-C-TD): TD |  |
| **Source:** | NTIA/ITS |
| **Title:**  | Proposal to insert the VQEG subject screening method into draft new Recommendation J.av-dist |
| **Contact:** | Margaret H. PinsonNTIA/ITSUSA | Tel: +1 303-497-3579Fax: +1 303-497-5969Email: mpinson@its.bldrdoc.gov |
| Please don’t change the structure of this table, just insert the necessary information. |

**Introduction**

Draft new Recommendation J.av-dist is missing the subject screening method used by the Video Quality Experts Groups (VQEG).  This method has been used successfully in three VQEG validation tests.

**Proposal**

Add this subject screening method to J.av-dist. Text is supplied below. This insertion will require the insertion of related text elsewhere in J.av-dist. The related text and clause locations are included below.

New definitions for clause 3.2:

**Processed Video Sequence:** The processed video Sequence (PVS) is the impaired version of a video sequence.

**Hypothetical Reference Circuit:** A hypothetical reference circuit (HRC)is a fixed combination of a video encoder operating at a given bitrate, network condition, and video decoder.The term HRC is preferred when vendor names should not be identified.

New acronyms for Clause 4:

PVS Processed Video Sequence

HRC Hypothetical Reference Circuit: ?

Modified text of clause 11.3, Post-screening of subjects:

Post-screening of subjects is recommended. The following subject screening methods are appropriate: Annex A, [ITU-R BT.500] clause 2.3, or [ITU-R BT.1788] clause 3.

New Annex A:

Annex A

Method for Post-Experiment Screening of Subjects Using Pearson Linear Correlation

(This annex forms an integral part of this Recommendation.)

The rejection criterion verifies the level of consistency of the raw scores of one subject according to the corresponding average raw scores over all subjects. Decision is made using correlation coefficient.

Linear Pearson correlation coefficient (LPCC) for one subject versus all subjects is calculated as:

 $LPCC\left(x,y\right)=\frac{\sum\_{i=1}^{n}x\_{i}y\_{i}-\frac{\sum\_{i=1}^{n}x\_{i}\sum\_{i=1}^{n}y\_{i}}{n}}{\sqrt{\left(\sum\_{i=1}^{n}x\_{i}^{2}-\frac{\left(\sum\_{i=1}^{n}x\_{i}\right)^{2}}{n}\right)\left(\sum\_{i=1}^{n}y\_{i}^{2}-\frac{\left(\sum\_{i=1}^{n}y\_{i}\right)^{2}}{n}\right)}}$ (A1)

Where *x* and *y* are arrays of data and *n* is the number of data points.

To calculate LPCC on individual stimuli (i.e., per PVS), compute

 $r1\left(x,y\right)=LPCC\left(x,y\right)$ (A2)

Where

*xi* = MOS of all subjects per PVS

*yi* = individual score of one subject for the corresponding PVS

*n* = number of PVSs

*i* = PVS index

To calculate LPCC on systems (i.e., per HRC), compute

 $r2\left(x,y\right)=LPCC\left(x,y\right)$ (A3)

Where

*xi* = condition MOS of all subjects per HRC (i.e. condition MOS is the average value across all PVSs from the same HRC)

*yi* = individual condition MOS of one subject for the corresponding HRC

*n* = number of HRCs

*i* = HRC index

One of the following two rejection criteria may be used.

**A.1 Screen by PVS**

Screening analysis is performed per PVS only, using Equation (A2). Subjects are rejected if *r1* falls below a set threshold. A discard threshold of (*r1* < 0.75) is recommended for ACR and ACR-HR tests of entertainment video.

Different thresholds may be needed depending upon the method, technology or application.

**A.2 Screen by PVS and HRC**

Screening analysis is performed per PVS and per HRC, using Equations (A2) and (A3). Subjects are rejected if *r1* or *r2* fall below set thresholds. For ACR and ACR-HR tests of entertainment video, a subject should be discarded if (*r1* < 0.75 and *r2* < 0.8). Both *r1* and *r2* must fall below separate thresholds before a subject is discarded.

Different thresholds may be needed depending upon the method, technology or application.

The reason for using analysis per HRC using *r2* is that a subject can have an individual content preference that is different from other subjects. This preference will cause *r1* to decrease, although this subject may have voted consistently. Analysis per HRC averages out individual’s content preference and check consistency across error conditions.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_