**Source:** **China Mobile Com. Corporation**

**Title: [FS\_Beyond2D] Quality aspects of stereoscopic video content**

**Agenda Item: 9.9**

**Document for: Discussion and Agreement**

# Introduction

This document provides the quality evaluation aspects of stereoscopic video content and is applicable to Scenario 5: *UE-to-UE Beyond 2D Video Streaming*, of the FS\_Beyond2D study [1].

# Proposed Updates

### 2.1 Mean Absolute Error (MAE) Metric

Mean Absolute Error (MAE) metric is also used for quantitative evaluation of stereoscopic 3D video content [2]. The MAE can be calculated as follows:

|*y - g(x)|*

where *x* is the left view, *y* is the right view, *g(·)* is the model, and *H* and *W* are height and width of the image respectively.

### 2.2 Human-Visual-System-based 3D (HV3D) Quality Metric

A full-referenced human visual-system-based quality metric for 3D videos called HV3D had been proposed in ITU-T [3]. It takes into account the quality of individual views, the quality of the cyclopean view (fusion of the right and left view, what the viewer perceives), as well as the quality of the depth information.

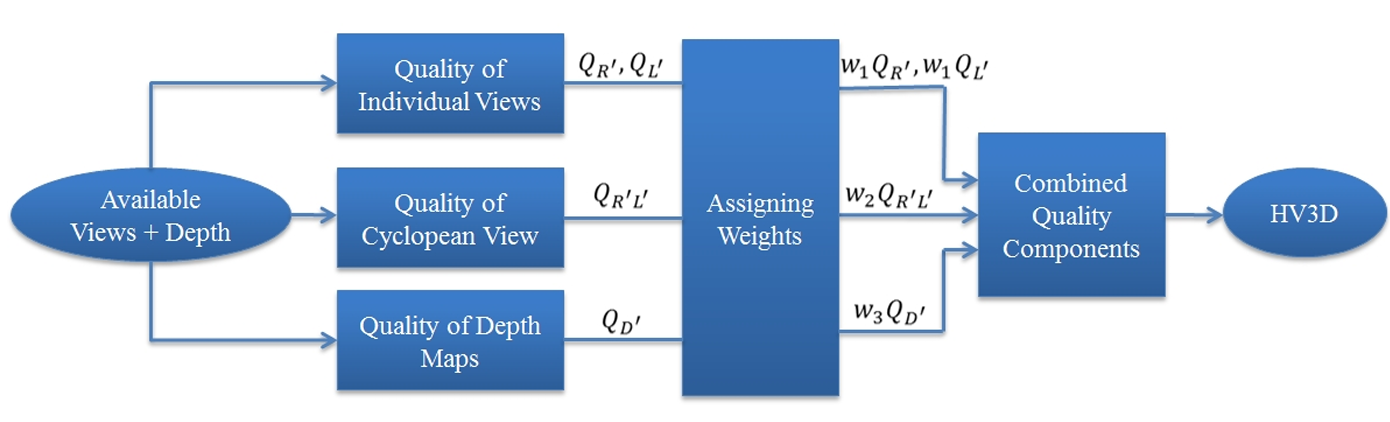


Figure 2.2-1 HV3D Flowchart

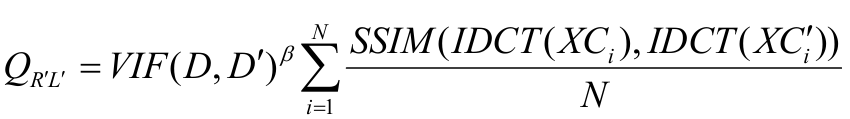
#### 2.1.1 Quality of individual views

The metrics (see below) are computed for each source view. To have one number that considers that the reconstructed view is partially synthesized, the metrics [dB] are averaged in the squared error domain and converted back to decibels.

#### 

#### 2.1.2 Quality of cyclopean view

In order to measure the quality of the cyclopean view, first the cyclopean view is constructed by combining the corresponding areas from the left and right views



#### 2.1.3 Quality of depth map

The quality of depth map is formulated as follows:



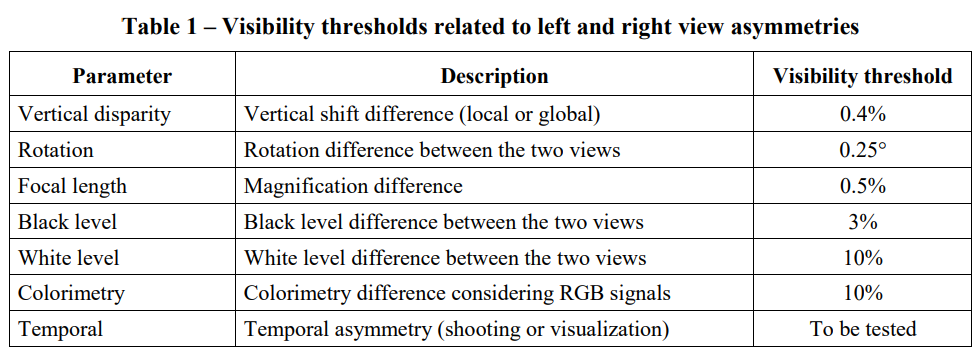
Where di is the variance of block *i* in the depth map of the 3D reference view and the local disparity variance is calculated over a block size area of 64x64.

### 2.3 Subjective evaluation

1. Subjective Assessment Methods for 3D Video Quality, document ITU-T P.3D-sam, International Telecommunication Union, Geneva,Switzerland, Jul. 2015.

*“In stereo 3D systems, a binocular 3D image is formed by presenting the left and right image to their respective eye. If discrepancies arise between these two images, they can cause psychophysical stress, and in some cases 3D viewing can fail. For example, when shooting and displaying stereoscopic 3DTV programmes, there may be geometrical, optical, electrical or temporal asymmetries, such as size inconsistency, vertical shift, rotation error, and luminance or colour levels between the left and right images. For the production of natural scene content using two independent video cameras, the main issue is to guarantee that the asymmetries of the views are under perceptual limits.”*

*Table 1 illustrates visibility thresholds obtained from subjective experiments using an impairment scale and for a viewing distance of 4.5 times the display height.*



1. Assessment Methods of Visual Fatigue and Safety Guideline for 3D Video, document ITU-T J.3D-fatigue, International Telecommunication Union, Geneva, Switzerland, 2015.

# Proposal

We propose to document section 2 to PD as the methodology for evaluating the quality stereoscopic 3D video content.

# References

[1] 3GPP TR 26.956 v0.0.2 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Evaluation and Characterization of Beyond 2D Video Formats and Codecs (Release 19), S4-240947, Jeju (Korea), May 2024.

[2] Xie, Junyuan et al. “Deep3D: Fully Automatic 2D-to-3D Video Conversion with Deep Convolutional Neural Networks.” European Conference on Computer Vision (2016).

[3] ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 - JCT3V-C0032: A human visual system based 3D video quality metric

[4] ITU-R Recommendation BT.2021, “Subjective methods for the assessment of stereoscopic 3DTV systems,” Aug. 2012.