|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extension Development**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  2nd Meeting: Shanghai, CN, 13–19 Oct. 2012 | Document: JCT3V-B0140 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Comments on 3D Video Quality Assessment in JCT-3V** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Information | | |
| *Author(s) or Contact(s):* | Sehoon Yea (LG)  Jincheol Park (Yonsei University)  Sanghoon Lee (Yonsei University) | Email: | [sehoon.yea@lge.com](mailto:sehoon.yea@lge.com)  [dewofdawn@yonsei.ac.kr](mailto:dewofdawn@yonsei.ac.kr)  [slee@yonsei.ac.kr](mailto:slee@yonsei.ac.kr) |
| *Source:* | LG Electronics, Yonsei University | | |

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

# Abstract

This contribution discusses some of the issues related to 3D video quality assessments. The intent is to bring to the attention of the AhG on 3D Quality Assessment a few technical considerations to be made when recommending improvements to evaluation methodology as well as to provide pointers to relevant activities and literatures.

# Introduction

In JCT2-A0150, it was reported that quality metrics such as VIFp, VQM, MS-SSIM, and SSIM are more correlated with perceived quality than PSNR for the 3D quality evaluation scenario relevant to the current JCT-3V activity. Also an AdHoc group was established on 3D Quality Assessment in the 101st MPEG meeting in Stockholm. The mandates of the AdHoc group included to study objective quality assessment metrics that would provide an accurate evaluation of synthesized views. This contribution briefly discusses some of the technical aspects to be considered as well as relevant references to help facilitate aforementioned activities.

# Quality Assessment of Synthesized View

## Sources of 3D perceptual distortion

Stereoscopic 3D introduces new types of perceptual distortions such as the keystone effect, ghosting effect, and cardboard effect that could be objectionable to viewers. In addition, use of view-synthesis for baseline adjustment of stereo or multiview generation for autostereo often introduces artifacts especially around rendered object-boundaries due to the presence of disocclusion or inaccuracy of the estimated depth. Furthermore, compression of depth or texture video adds an additional source of 3D distortion to the viewer’s eye. In general, these three sources will combine and one might dominate over the others depending on the relative severities. Important to consider under the context of JCT-3V is obviously how to separate and measure the effect of compression on 3D distortion from those of other sources.

## Alternative metrics to PSNR

One of the previous contributions, i.e., JCT-V3 A0150 [1] as well as [2][3] discuss alternative ways to quantify 3D distortion of synthesized views that better correlate with human perception. For example, [2] tests seven alternative measures and concludes WSNR, pixel-based PSNR, NOM are better than PSNR while [3] proposes two measures based upon the analysis of the orientation of the edge shifts and the analysis of disoccluded areas, respectively. As for the reference (aka anchor) selection for quality evaluation, the current approach of using synthesized views based upon uncompressed texture/depth as anchor seems reasonable as it will not unduly penalize only the coded cases. In other words, the adverse effects of disocclusion and inaccurate estimated depth on the quality of synthesized views will be roughly identical whatever the capabilities and limitations of the reference view synthesis algorithm.

Assuming a reasonable success in securing acceptable estimated depth quality and a reference view synthesis algorithm, on which opinions might split in the current JCT-3V, the ensuing question would be back to how to measure the compression-induced distortion in an objective manner that correlates better with 3D human perception. Further collaborative investigations in a prompt manner seem necessary in order for this standard committee to have a higher-level of confidence in the technical soundness of the developed (depth) coding tools as well as the marketability of the resulting standard contingent on it.

## Penalizing local distortions (aka alternative pooling)

One important question to ask is whether current typical ways of pooling the quality scores of every local region in a synthesized picture by *averaging* is adequate. By averaging *with equal weights* for local distortions in coming up with an overall figure of (de-)merit will often lead to a quantitative measure that does not correlate well with human perception. References [5][6] discuss this issue and propose alternative pooling methods of local distortions.

## Relevant Databases

* EPFL-PoliMI Video Quality Assessment Database : <http://vqa.como.polimi.it/>
* VQEG Video Quality Assessment Database : <ftp://vqeg.its.bldrdoc.gov/>
* LIVE Image and Video Quality Assessment Database : <http://live.ece.utexas.edu/research/Quality/index.htm>

## Relevant (Standard) Activities

There is an on-going WG activity within IEEE-SA called “Standard for the Quality Assessment of Three Dimensional (3D) Displays, 3D Contents and 3D Devices based on Human Factors” that focuses on 3D quality assessment. Those interested might want to check out the following link:

<http://grouper.ieee.org/groups/3dhf/>

# Conclusion

This contribution discussed some of the issues related to 3D video quality assessments. Further collaborative investigations in a prompt manner to find alternative objective metric that better correlates with 3D perception seem necessary.

# References

1. ISO/IEC JTC1/SC29/WG11, "3DV: Quality assessment of stereo pairs formed form two synthesized views," Doc. JCT2-A0150, Stockholm, SE, July 2012.

2. E. Bosc, M. Köppel, R. Pépion, M. Pressigout, L. Morin, P. Ndjiki-Nya, and P. Le Callet, "Can 3D synthesized views be reliably assessed through usual subjective and objective evaluation protocols?," in International Conference on Image Processing, 2011, pp. 2597-2600.

3. E. Bosc, R. Pepion, P. Le Callet, M. Koppel, P. Ndjiki-Nya, M. Pressigout, and L. Morin, “Towards a new quality metric for 3D synthesized view assessment,” IEEE Journal of Selected Topics, 2011.

4. N. A. El-Yamany, K. Ugur,M.M. Hannuksela, andM.Gabbouj, “Evaluation of depth compression and view synthesis distortions in multiview- video-plus-depth coding systems,” in Proc. 3DTV-Conf.: The True Vision—Capture, Transmission, Display of 3-D Video, Tampere, Finland, 2010, pp. 1–4.

5. E. Ekmekcioglu, S. T. Worrall, D. De Silva, W. A. C. Fernando, and A. M. Kondoz, “Depth based perceptual quality assessment for synthesized camera viewpoints,” in Proc. of UCMedia, Palma de Mallorca, Spain, Sep. 2010.

6. J. Park, K. Seshadrinathan, S. Lee and A. C. Bovik, “Spatio temporal quality pooling accounting for transient severe impairment and egomotion,” IEEE Int. Conf. Image Processing, Brussels, Sept. 2011.