Towards a Quality Metric for Dense Light Fields

Chaonan Song

Jun 8, 2018

Abstract

Today I read the later part of Dr. Adhikarla's thesis. Dr. Adhikarla collected some data in this section, including Scenes and Distortions. In order to obtain a subjective quality score that can simultaneously train and test different quality indicators, Dr. Adhikarla conducted a large-scale subjective experiment. Dr. Adhikarla then analyzed the results of the experiment.

1. Data collection

Dr. Adhikarla's data set consists of a light field that is parameterized using two parallel planes. Motion parallax is described by a plane and a line parallel to it.

1.1. Distortions

Dr. Adhikarla consider typical light field distortions that are specific to transmission, reconstruction, and display. For each distortion, Dr. Adhikarla generate multiple light fields by varying the severity of the distortion. The exact level was chosen to keep the difference between two consecutive levels small and similar. To do this, Dr. Adhikarla conducted a small pilot study of 10 distortion levels and then manually selected the final level.

Transmission: In order to transmit light field data, highly efficient data compression algorithms are highly required. Dr. Adhikarla consider the well-known 3D extension of the HEVC encoder Tech *et al.* [3].

Reconstruction: Light field reconstruction techniques are used to recover dense light fields from sparse view samples. They use a variety of techniques to insert missing views that change the nature and appearance of the distortion. Dr. Adhikarla chose the distortion produced by LINEAR and NN interpolation, and the image distortion using Optical Flow Estimation (OPT). Dr. Adhikarla also investigated the effect of using a quantized depth map (DQ).

Display: As an example of multi-view autostereoscopic display artifacts, Dr. Adhikarla have selected crosstalk between adjacent views, which can be modeled using Gaussian blur in the angle domain Liu *et al.* [1]. Therefore, Dr.

Adhikarla include these artifacts in our dataset (GAUSS). Upsampling to higher resolution light fields is achieved by using a display crosstalk model. An example of artifacts generated is shown in Figure 1.
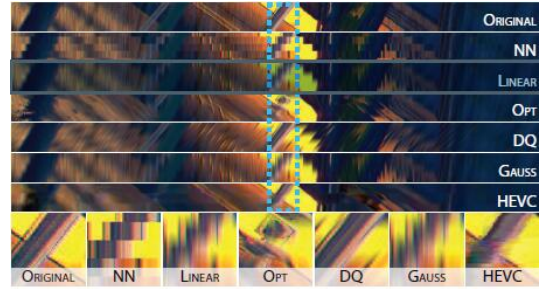

Figure 1. Examples of distortions introduced to our light field for one of our scenes (BARCELONA). The images visualize central EPI of each of the distorted light fields and the enlarged portion of it is shown in the bottom row

2. Experiment

In order to obtain a subjective quality score that can simultaneously train and test different quality indicators, Dr. Adhikarla conducted a large-scale subjective experiment. Equipment: To simulate stereoscopic viewing with high-quality motion parallax, we used on our own setup (Figure 3) that consists of ASUS VG278 27¡□□ Full HD 120 Hz LCD desktop monitor and NVIDIA 3D Vision 2 Kit for displaying stereoscopic images. Motion parallax was reproduced using a custom head tracking in which a small LED headlamp was tracked using a Logitech HD C920 Pro webcam (refer to the supplemental video).

Task: Dr. Adhikarla tried to use a direct scoring method such as ACR to measure the Mean-Opinion-Score of a distorted image. We found that these methods are insensitive to subtle but significant quality degradation. We decided to use more sensitive pairwise comparison methods and double-selection forced selection.

Participants: Forty participants participated in this test, including males (20) and females (20) aged 24-40 years, with normal or corrected vision. Each subject performs three tests within a week. In a meeting, participants see 120-180 light
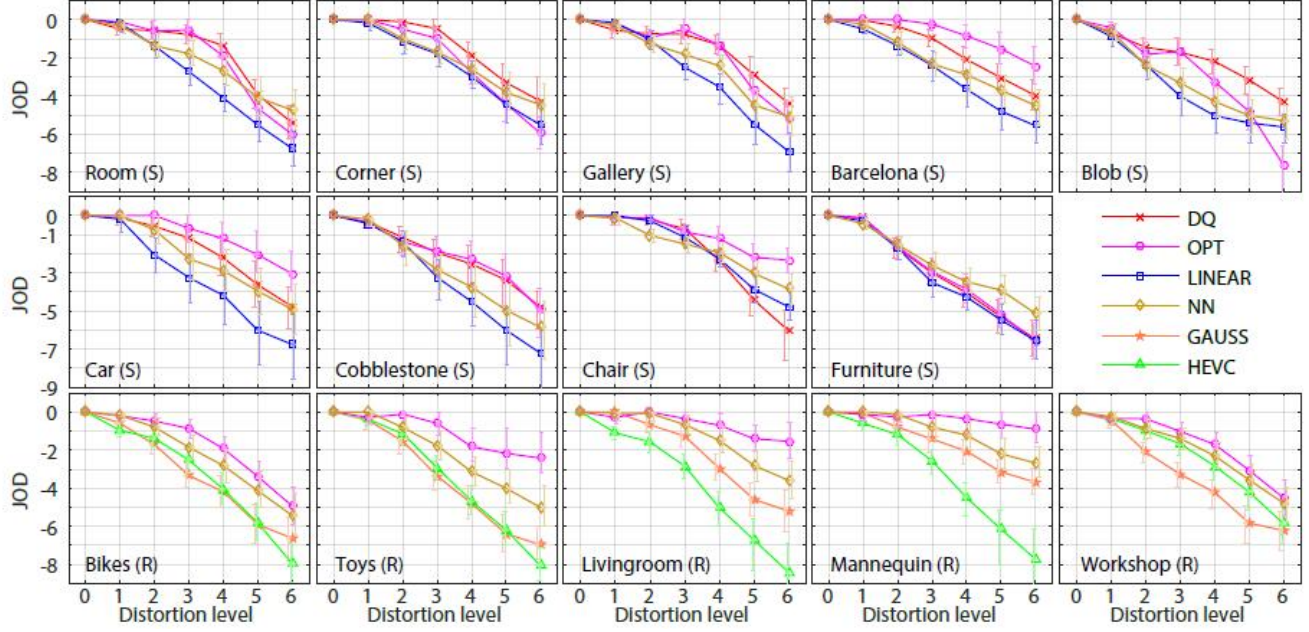
Figure 2. The results of the subjective quality assessment experiment. The distortion level indicates the distortion severity: 0¨Creference 6¨Cseverest distortion level. JOD is the scaled subjective quality value. The error bars denote 95% confidence interval. The bars are horizontally displaced to avoid overlapping. The scene names are indicated in the corner of each plot. The character in parenthesis after the scene name indicates whether the scene is synthetic (S) or real-world (R).
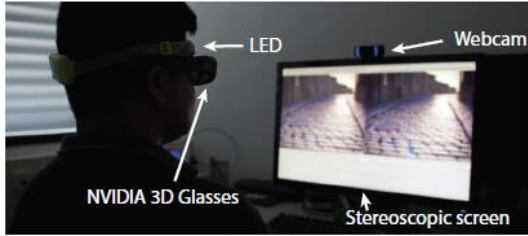

Figure 3. Experiment session: viewer¡¯s position is tracked using a head lamp and a webcam, a pair of NVIDIA 3D Vision 2 Kit active glasses provides stereoscopic viewing.

field pairs consisting of all test conditions, but for a subset of the scene. For a given subject, two tests are allowed within one day, and these tests need to rest for at least one hour.

3. Analysis of subjective data

Dr. Adhikarla observed that taking into account "obvious" measurement differences can lead to misinterpretation of experimental results. However, Dr. Adhikarla and his team experimental problem is not whether the observer can judge whether the light field is different, but which one has higher quality. As shown in Figure 4, a pair of stimuli may be significantly different from each other (JND > 1), but they seem to have the same quality. For this reason, Dr. Adhikarla expressed the measured values as just disgusting differences (JODS). These units quantify the quality differences

associated with a perfect reference image. The measure of JOD is more similar to visual equivalence Ramanarayanan *et al.* [2] or the quality is expressed as difference mean score (DMOS) rather than JND.
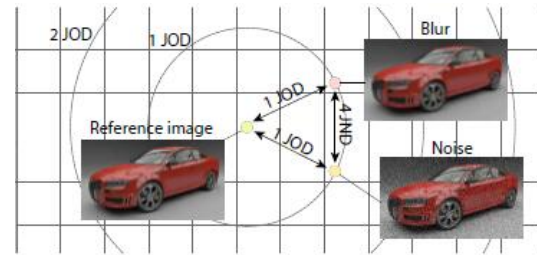

Figure 4. Illustration of the difference between justobjectionabledifferences (JODs) and just-noticeabledifferences (JNDs). The image affected by blur and noise may appear to be similarly degraded in comparison to the reference image (the same JOD), but they are noticeably different and therefore several JNDs apart. The mapping between JODs and JNDs can be very complex and the relation shown in this plot using Cartesian and polar coordinates is just for illustration purposes.

The results of subjective quality assessment experiments are shown in Figure 2. The error bars represent the 95% confidence interval relative to the reference light field and are calculated by bootstrapping using alternative samples. The result shows a distorted and disgustingly interesting model.

References

[1] J. Liu, T. Malzbender, S. Qin, B. Zhang, C.-A. Wu, and J. Davis. Dynamic mapping for multiview autostereoscopic displays. In *Proceedings of SPIE - The International Society for Optical Engineering*, 2015.

[2] G. Ramanarayanan, J. Ferwerda, B. Walter, and K. Bala. Visual equivalence: Towards a new standard for image fidelity. In *ACM Transactions on Graphics*, 2007.

[3] G. Tech, Y. Chen, K. Müller, J.-R. Ohm, A. Vetro, and Y.-K. Wang. Overview of the multiview and 3D extensions of high efficiency video coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016.