



HAL
open science

Analysis of public light field datasets for visual quality assessment and new challenges

Waqas Ellahi, Toinon Vigier, Patrick Le Callet

► To cite this version:

Waqas Ellahi, Toinon Vigier, Patrick Le Callet. Analysis of public light field datasets for visual quality assessment and new challenges. European Light Field Imaging Workshop, Jun 2019, Borovets, Bulgaria. hal-02504946

HAL Id: hal-02504946

<https://hal.science/hal-02504946v1>

Submitted on 11 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ANALYSIS OF PUBLIC LIGHT FIELD DATASETS FOR VISUAL QUALITY ASSESSMENT AND NEW CHALLENGES

Waqas Ellahi, Toinon Vigier, Patrick Le Callet

firstname.name@univ-nantes.fr
LS2N UMR CNRS 6003, Université de Nantes
Nantes, France

ABSTRACT

Different orientations of rays emitted by a 3D scene are captured in Light field (LF) imaging. In this, spatial along with angular information of the scene is captured. LF offers higher dimensional representation which needs advanced processing such as post-capture refocusing, depth sensing, compression, and reconstruction. These processing methods often induce artifacts as the quality is degraded. So, it is important to understand the mechanism of human visual quality perception to find the optimal performance that maximizes the quality of experience (QoE) of LF. The measurement of this data degradation is possible with the availability of a suitable dataset. In this paper, we present a comprehensive overview and discussion of publicly available LF datasets with emphasis on dataset with subject score. Furthermore, we discuss the current trends and existing challenges in LF dataset creation for QoE.

Index Terms — Light Field, Quality of experience, Quality assessment

1. INTRODUCTION

Recently, LF imaging received much popularity in research community due to its applications in computer vision and computer graphics. In LF, the distribution of light rays is recorded from free space in the LF data. On the contrary, the conventional images express the projection of the light on a 2D plane. This allows LF to capture more information from a scene as compared to a normal 2D image. The plenoptic function [1] describes light in space as a function of position, time, angle, and wavelength. The plenoptic function is high dimensional 7D mathematical model denoted as $L(x, y, z, \theta, \phi, \gamma, t)$, where (x, y, z) represent position in 3D space, (θ, ϕ) represent propagation angles, γ is wavelength and t represent time. The higher dimensional information is difficult to capture and process in real-time. Thus, the plenoptic function is simplified for LF data, and is reduced to a 4D equation [2]. The representation of a 4D LF is to parameterize the rays by the coordinates of their intersections with two planes placed in arbitrary positions. The coordinate system is denoted by (u, v) and (s, t) for first and second plane respectively as shown in Figure 1.

The content of the captured scene is degraded during transmission, reconstruction and visualization process [3]. Figure 2 clearly shows the artifacts of few processes. The figures 2-a and 2-b show the compression artifacts and 2-c and 2-d display reconstruction errors. The red rectangular mark highlights degradation in the perceived quality. The performance of LF processing schemes relies on intended application of LF and it is also influenced by visual quality assessment method.

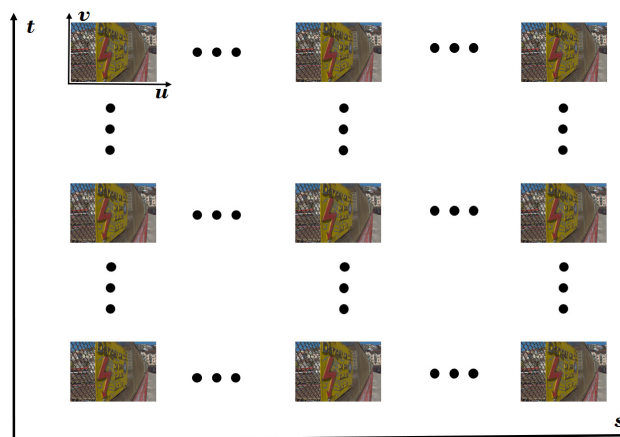


Figure 1: Parameterization of Light field

However, subjective tests are time consuming and also it poses ethical challenges as well which encourages to assess quality using objective metrics which correlate with human vision system. The scheme can be build on state-of-the-art deep learning architectures, that requires dataset with large number of images. However, such a dataset can be used to develop new compression and rendering processing.

This paper aims to provide an overview on publicly available LF imaging dataset for visual quality assessment, while simultaneously revealing the current trends and challenges. Section 2 briefly explained the general LF data-sets and datasets with subjective scores are describe in Section 3. The overview of existing challenges in the research is covered in Section 4. Finally, concluding remarks are presented in Section 5.

2. LIGHT FIELD DATASETS

An overview of the subjects, test material, viewing setup, and other details of experiments for each dataset is presented in Table 2. Stanford Light Field Archives collects the various light field datasets. The newest Stanford Multiview Light Field Datasets released in 2018 has two multi-view datasets with a total of about 7200 light fields [4]. On the contrary to earlier datasets, in new datasets is each scene is captured from a diversity of camera poses. Ziegler et al. [5] captured natural scene with wide-baseline, in the order of meters, in both vertical and horizontal direction using cantilever axes. In the dataset presented in [6] a single scene is captured from two different plenoptic cameras, i.e. plenoptic 1.0 and plenoptic 2.0 cameras from a single view point under similar conditions. Rerabek and Ebrahimi [7] provided 118 LF captured scenes with a Lytro camera for evaluation of the algorithm's per-



Figure 2: Distortion due to compression and reconstruction algorithms

formance. The large amount of LF images captured with Lytro camera enables the training of deep learning schemes.

The publicly available datasets [8, 3, 9, 10, 11] with primary focus on quality assessment of LF are explained in further details in next section.

3. PUBLIC LF DATASETS FOR QUALITY ASSESSMENT

LF is considered as a modern technology in the imaging techniques. New algorithms are to be developed for the camera capturing, over the encoding and transmission stage and rendering on the display. Each of these steps need to be re-evaluated in the LF context. Recently, few datasets have been published for quality assessment of LF data [8, 3, 9, 10, 11] on different stages in the process.

Tamboli et al. published a multi view dataset using three real objects with angular resolution on a turn table [8]. The artifacts examined are blurring, additive noise and JPEG compression in the study. The Holografika's Holovizio HV721RC LF display with an Absolute Category Rating (ACR) was used for subjective evaluation methodology. The main limitations of this study are the consideration of only horizontal parallax with scene content which is not much realistic and considers only a single object in the scene. In addition, the artifacts or distortion, considered

during experiments, are not LF specific.

A second dataset, named SMART, contains fifteen LF images compressed at different bitrates with the annotated subjective scores [3]. The author defines the criteria for the selection of source sequences (SRCs). However, the intra-based approach only considers for encoding LF scene in the proposed method, which were subpar with respect to pseudo-sequence based methods [7]. Moreover, the LF contents were presented on a conventional 2D display during subjective methodology experiment, which admittedly disregards any problem that may arise in the encoding of the depth information.

Kiran et al. [9] proposed a dataset of dense LF of total fourteen scenes with five synthetic images. Several reconstructions, compressions and display distortions are considered during subjective experiment. The stereoscopic view was established using ASUS VG278 monitor with NVIDIA 3D Vision 2 Kit. The two-alternative forced choice 2AFC with pairwise comparisons among the LF pairs of the dataset was used to extract the JOD. They considered only horizontal parallax and ignore the vertical parallax in the dataset generation.

Viola et al. [11] proposed a so-called VALID dataset composed of five uncompressed and compressed contents on various bitrates using four different compression solutions, along with subjective quality scores. Three different subjective methodologies

Table 1: Light field datasets

Datasets	Year	Purposes	Scene	Devices
Stanford Multiview Light Field Dataset [4]	2019	Scene with different camera poses	1889	Lytro Illum
Ziegler et al. [5]	2017	Dense LF of large scenes	9	Sony Alpha 7RII
W. Ahmad et al. [6]	2018	Comparison between plenoptic 1.0 and 2.0	31	Lytro Illum, Ratyrix R29
Rerabek and T. Ebrahimi [7]	2016	General	118	Lytro Illum
Tamboli et al. [8]	2016	QoE	3	Basler's ACE
SMART [3]	2016	QoE	15	Lytro Illum
Kiran et al. [9]	2017	QoE	14	Canon EOS 5D
VALID [11]	2018	QoE	5	Lytro Illum
Shi et al. [10]	2018	QoE	10	Lytro Illum

(passive, interactive, Passive and interactive) and two different visual setups (8 and 10-bit depth) were established for the experiment. A Double Stimulus Impairment Scale (DSIS) with 5-point grading scale was selected for all methodologies. However, 2D display was employed and only compression artifact was considered during experiment.

Shi et al. [10] proposed the first dataset with five degrees of freedom (5DOF) with ten scenes based on stereoscopic display. The dataset contains representative reconstruction and compression artifacts with double-stimulus continuous quality scale (DSCQS) methodology as well as both the picture quality and overall quality of LF image. The 55-inches SAMSUNG 3D television stereo display with shuttle glass were used in the experiment.

4. NEW CHALLENGES FOR QOE IN LF DATA

This section focuses on identifying the problems and challenges faced by the present studies of the researchers within experiment designing and analysis. The contribution and analysis of these publicly available datasets is shown in Table 3.

4.1. Experimental methodology

Unfortunately, there is no dataset available with all the LF artifacts present on a 3D display. The datasets [9, 10] employed the compression and reconstruction artifacts on a 3D display but missed other LF artifacts during refocusing, depth sensing process. The directional aliasing artifacts originate in a refocused image [12], and affects the quality of the image. The VALID dataset [11] considers the refocusing views, but the stimuli was presented on a 2D display. However, post-processing refocusing and depth sensing methods are important LF processing techniques and need some criteria to access quality of these methods.

The different processing methods (Hypothetical Reference Circuit, HRC) induced the artifacts that influence the perceived quality. The predefined navigation trajectories with conventional displays, carried out during evaluation studies when exploring the content. These different trajectories (Hypothetical rendering trajectories, HRT) lead to different conclusions in terms of visual quality [13]. The impact of these visual trajectories of a LF data is an open question for research.

4.2. Experience in head mounted display (HMD)

HMD is a wearable technology cable of enhancing user's sense of presence while exploring virtual environment. The main issue is that the users may experience motion sickness that will cause nausea and dizziness [14]. The other problem is to optimize the process to reduce the noticeable graphic pixilation to minimize visuals rendering. LF data are closer to genuine representation of how light exists and is perceived in the real world, which is significant to VR and AR. So a study is needed on how the 360 degree of LF content improve the user experience in the HMD. The visual interaction, navigation and immersion raises issues about depth perception and visual interfaces [15]. The LF content impact these issues and on a user experience. Quality assessment of LF rendering techniques is still an open research topic for research community as substantial research lacks in this field.

4.3. Behavior study

All publicly available datasets for LF quality assessment are with the explicit approach, which have been adopted in numerous studies. In explicit approach, a multimedia stimulus is presented to the human and asks them to fill in a questionnaire. However, the implicit approach does not require person's actions for rating stimuli or answering questions but observe the natural response passively to the given stimuli. Recently, implicit measurement using physiological signals has gained much attention due to the increasing popularity of wearable devices equipped with various physiological sensors [16]. The implicit approach can be used for real-time monitoring after training, using ground truth data, obtained from explicit methods. The bias results may be occurring due to users conceptualize, recall, and verbalize experience in explicit approach. The psychophysical methods that deal with observer bias and the uncertainty of rating on an abstract scale. In the past, the different psychophysiological signals like electroencephalography EEG [17], Eye-Tracker, galvanic skin response GSR, and Heart Rate (HR) are used for video quality assessment. The P300 wave has been shown as a good indicator of the users' experience for EEG data. Similarly, the fixation and saccadic eye movements were gathered with the Eye-Tracking which helped in evaluating user perceived visual quality [18].

5. CONCLUSION

This paper has summarized the development of LF imaging datasets for quality assessment, in terms of both the datasets published, and the challenges that lie ahead. The standardization of subjective experiment methodology, experience of participants in HMD and integration of physiological signals for LF data are the open questions for research community.

6. ACKNOWLEDGMENT

The work in this paper was funded from the European Union's Horizon 2020 research and innovation program under the Marie

Table 2: Comparison of the different datasets

Characteristics	Tamboli et al. [8]	SMART [3]	Kiran et al. [9]	VALID [11]	Shi et al. [10]
Artifacts	- Additive Noise - Compression - Blurring	Compression	- Reconstruction - Compression - Display	- Compression - Refocusing	- Compression - Reconstruction - Refocusing
Compression	JPEG	- JPEG - JPEG2000 - HEVC Intra - SSDC	3D-HEVC	- HEVC - VP9 - 3 other methods	- JPEG 2000 - HEVC
Reconstruction	-	-	- Linear - Nearest neighbour - Optical flow estimation - Quantized depth maps	-	- Linear - Nearest neighbour - 2 CNN based methods
Display	-	-	Crosstalk	-	-
Participants	20	19	40	-	29
Display	3D	2D	3D	2D	3D
Rating	SSCQE	JND	JOD	DSIS	DSCQS

Table 3: Analysis of the literatures

Datasets	Contributions	Remarks
Tamboli et al. [8]	- First dataset with high angular resolution - Dataset using three real objects in a turntable setting	- Simple scenes and artifacts are not LF specific - Ignore Vertical Parallax
SMART [3]	- Define image content selection criteria - Analysis of the features of the proposed dataset	- Intra-based approach used to encode LF images - 2D display device
Kiran et al. [9]	- Compression, display and reconstruction artifacts are considered	- Ignore vertical parallax
VALID [11]	- Lots of compression methods artifacts are considered during the experiment	- Compression artifacts are only included - 2D display device
Shi et al. [10]	- First full parallax dataset along reconstruction and compression artifacts	- This dataset covered many LF artifacts

Skłodowska-Curie grant agreement No 765911, European Training Network on Real Vision project.

7. REFERENCES

- [1] Edward H Adelson, James R Bergen, et al., “The plenoptic function and the elements of early vision,” 1991.
- [2] Steven J Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F Cohen, “The lumigraph,” in *Siggraph*, 1996, vol. 96, pp. 43–54.
- [3] Pradip Paudyal, Federica Battisti, Mårten Sjöström, Roger Olsson, and Marco Carli, “Towards the perceptual quality evaluation of compressed light field images,” *IEEE Transactions on Broadcasting*, vol. 63, no. 3, pp. 507–522, 2017.
- [4] Donald G. Dansereau, Bernd Girod, and Gordon Wetzstein, “LiFF: Light field features in scale and depth,” in *Computer Vision and Pattern Recognition (CVPR)*. IEEE, June 2019.
- [5] Matthias Ziegler, Ron op het Veld, Joachim Keinert, and Frederik Zilly, “Acquisition system for dense lightfield of large scenes,” in *2017 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*. IEEE, 2017, pp. 1–4.
- [6] Waqas Ahmad, Luca Palmieri, Reinhard Koch, and Mårten Sjöström, “Matching light field datasets from plenoptic cameras 1.0 and 2.0,” in *2018-3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*. IEEE, 2018, pp. 1–4.
- [7] Martin Rerabek and Touradj Ebrahimi, “New light field image dataset,” in *8th International Conference on Quality of Multimedia Experience (QoMEX)*, 2016, number CONF.
- [8] Roopak R Tamboli, Balasubramanyam Appina, Sumohana Channappayya, and Soumya Jana, “Super-multiview content with high angular resolution: 3d quality assessment on horizontal-parallax lightfield display,” *Signal Processing: Image Communication*, vol. 47, pp. 42–55, 2016.
- [9] Vamsi Kiran Adhikarla, Marek Vinkler, Denis Sumin, Rafal K Mantiuk, Karol Myszkowski, Hans-Peter Seidel, and Piotr Didyk, “Towards a quality metric for dense light fields,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 58–67.
- [10] Likun Shi, Shengyang Zhao, Wei Zhou, and Zhibo Chen, “Perceptual evaluation of light field image,” in *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018, pp. 41–45.
- [11] Irene Viola and Touradj Ebrahimi, “Valid: Visual quality assessment for light field images dataset,” in *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2018, pp. 1–3.
- [12] Ensun Lee, Seohee Yang, Miseon Han, and Jeongtae Kim, “Depth-based refocusing for reducing directional aliasing artifacts,” *Optics express*, vol. 24, no. 24, pp. 28065–28079, 2016.

- [13] Suiyi Ling, Jesús Gutiérrez, Ke Gu, and Patrick Le Callet, "Prediction of the influence of navigation scan-path on perceived quality of free-viewpoint videos," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2019.
- [14] Kieran Carnegie and Taehyun Rhee, "Reducing visual discomfort with hmds using dynamic depth of field," *IEEE computer graphics and applications*, vol. 35, no. 5, pp. 34–41, 2015.
- [15] Guillaume Moreau, "Visual immersion issues in virtual reality: a survey," in *2013 26th Conference on Graphics, Patterns and Images Tutorials*. IEEE, 2013, pp. 6–14.
- [16] Rainer Reisenzein, Markus Studtmann, and Gernot Horstmann, "Coherence between emotion and facial expression: Evidence from laboratory experiments," *Emotion Review*, vol. 5, no. 1, pp. 16–23, 2013.
- [17] Sebastian Bosse, Kjell Brunnström, Sebastian Arndt, Maria G Martini, Naeem Ramzan, and Ulrich Engelke, "A common framework for the evaluation of psychophysiological visual quality assessment," *Quality and User Experience*, vol. 4, no. 1, pp. 3, 2019.
- [18] Olivier Le Meur and Antoine Coutrot, "How saccadic models help predict where we look during a visual task? application to visual quality assessment.," *Electronic Imaging*, vol. 2016, no. 13, pp. 1–7, 2016.