Towards the Perceptual Quality Evaluation of Compressed Light Field Images

Pradip Paudyal, *Student Member, IEEE,* Federica Battisti, *Senior Member, IEEE,*
Mårten Sjöström, *Senior Member, IEEE,* Roger Olsson, *Member, IEEE,* and Marco Carli, *Senior Member, IEEE*

***Abstract*—Evaluation of perceived quality of light field images, as well as testing new processing tools, or even assessing the effectiveness of objective quality metrics, relies on the availability of test dataset and corresponding quality ratings. This article presents SMART light field image quality dataset. The dataset consists of source images (raw data without optical corrections), compressed images, and annotated subjective quality scores. Furthermore, analysis of perceptual effects of compression on SMART dataset is presented. Next, the impact of image content on the perceived quality is studied with the help of image quality attributes. Finally, the performances of 2D image quality metrics when applied to light field images are analyzed.**

***Index Terms*—Quality of Experience, Perceptuaal Quality, Light Field Image, Quality Metrics, Dataset.**

I. Introduction

LIGHT Field (LF) imaging, also known as holoscopic imaging, integral imaging, or plenoptic imaging [1], derives from the fundamentals of light field sampling, where the spatial information about a scene can be captured with angular information. That is, light field imaging is based on a camera recording information about the intensity of light of the scene and about the direction of light rays.

The basic concept of light field imaging was first introduced by Lippmann [2] [3] and it has been improved by many researchers throughout the years. The idea was to use an array of thin lenses that can record multiple views of a scene that in turn is projected onto a single sheet of film [2]. Exploiting light field for generating parallax images using a single lens is explored in [4]. In [5] a plenoptic camera using a lenticular array is designed. In [6] and [7] light field camera with an additional microlens array inserted between the camera sensor and the main lens is used. The camera can be considered as a relay system, where the main lens creates a main image in the air, then this main image is re-mapped to the sensor by microlens array, and thus it is able to provide multiple views of the scene in a single shot [8].

The most appealing feature of light field camera is that even a single light field snapshot can provide pictures where focus, exposure, and depth of field can be adjusted after the picture is taken. Light field camera offers novel opportunities in many applications such as photography, astronomy, robotics, medical imaging, and microscopy [9]. The wide range of possible applications and the rapidly developing light field technology pulled the attention of the consumer, industry, and academics. In this scenario, the Joint Photographic Experts Group (JPEG) committee has launched a new activity, called JPEG-PLENO [10], aiming at developing a standard framework for the representation and exchange of new imaging modalities such as light field imaging.

Light field images are subject to a wide variety of distortions during acquisition, processing, compression, storage, transmission, and rendering; any of these steps may result in a visual quality degradation. The rapidly developing light field technology and consumer interest towards this technology is pushing the need for perceptual quality evaluation of such contents. In this article, the perceptual quality is evaluated through subjective quality assessment.

A. Research questions

In the following, the motivations and the contributions of this work are presented.

R1) Light field image quality dataset: Image processing community devoted many efforts to image quality assessment. The issues: How to evaluate the quality of images generated from the algorithms? How to compare the performance of algorithms? How to determine the quality criteria to optimize the system or algorithm? etc., repeatedly come up when dealing with image processing [11] [12]. Since, in most applications the final user of light field image content is human, the goal of quality assessment is to predict the quality as it is perceived by the average human observer. To train, test, and benchmark the light field Image Quality Metrics (IQMs) a comprehensive and well defined light field image quality dataset, test light field images with corresponding subjective quality scores, is important. To the best of our knowledge, none of existing light field image datasets include processed images and corresponding subjective quality scores.

R2) Subjective quality of compressed light field images: It is useful to notice that light field camera provides a grid of elementary images, and the contents of the elementary images are quite similar to their neighbors. Therefore, compression of light field content is very important for storage and transmission. Since lossy compression may introduce visible artifacts, the analysis of the overall subjective quality is needed. Currently, many efforts have been devoted to develop light field image compression methods and the performance

P. Paudyal, F. Battisti, and M. Carli are with the Department of Engineering, Università degli Studi Roma TRE, via Vito Volterra, 62 00146 Rome, Italy, e-mail: pradip.paudyal@uniroma3.it; federica.battisti@uniroma3.it; marco.carli@uniroma3.it (see http://www.comlab.uniroma3.it/people.html).

R. Olsson and M. Sjöström are with the Department of Information Systems and Technologies (IST), Mid Sweden University, Holmgatan 10, 851 70 Sundsvall, e-mail: marten.sjostrom@miun.se; roger.olsson@miun.se (see https://www.miun.se/stc/realistic3d)

analysis of these methods is important. To this aim, the impact of compression artifacts on the subjective quality has been analyzed and the performance of image compression methods has been studied.

Moreover, the subjective quality of images and the performance of compression methods is significantly influenced by the image content characteristics and Human Visual System (HVS) [13]. Therefore, the influence of the image content with respect to HVS on the subjective quality of light field image has been studied by exploiting key image-quality attributes: spatial information, colorfulness, hue, saturation, brightness, contrast, etc.

R3) Performance evaluation of existing 2D image quality metrics, when applied to light field image: Currently, there are no quality metrics specifically designed for light field content. However, in literature many IQMs have been proposed, and some of them are very common in the image processing field.

In this contribution, the performances of existing 2D IQMs, when applied to light field images, are evaluated.

The rest of the paper is organized as follows: Section II describes a survey of the related works, Section III details the followed steps for dataset formation, Section IV summarizes the performed experiment, Section V includes the results and analysis, the issues and challenges arisen buy our work are discussed in Section VI, and in Section VII the conclusions are drawn.

II. Related works

With respect to R1, in literature, few light field datasets have been proposed; the main features are reported in Table I. Light field images in Stanford light field archive [14] are captured by using a traditional light field image acquisition framework, the multi-camera system. Synthetic light field archive [15], light field analysis [16], LCAV-31 [17], Lytro dataset [18], light field saliency dataset (LFSD) [19] and GUC light field face and iris dataset [20] are designed for specific purposes and light field images captured by using multi-camera system and Lytro. The Lytro Illum camera is used in [21], however, as detailed in [22], the scene selection has not been properly addressed thus resulting in redundancy in the dataset images.

Advances in light field imaging technology and the availability of commercial light field cameras (Lytro Illum and Raytrix) allow the consumer to exploit such a technology. The state-of-the-art datasets are not sufficient to deal with new challenges (quality evaluation, performance testing for processing algorithms, etc.) arising with the advancement of the light field technology. Moreover, none of the datasets contain the information about processed light field images and annotated subjective quality scores.

The novelty of the media has an impact on the subjective evaluation too. To collect the subjective opinion scores for the test light field images, a subjective quality assessment experiment need to be designed. In literature, many standard guidelines [23] [24] [25], have been recommended to design the subjective experiment for images and videos. However, currently there is no guideline defined for light field images.

With respect to R2, many ongoing efforts have been given to encode the light field image. In [26], a wavelet packet based hierarchical disparity compensated coding scheme is presented. In [27], an adaptive prediction method based on micro images is proposed to reduce the data correlation before the entropy coding. To compress lenslet images, a sub-aperture images streaming scheme is presented in [28]. Where, sub-aperture images are firstly extracted and video stream is produced from it, and it is compressed as a video. In [1], the self-similarity prediction concept is exploited to explore the inherent correlation of the contents. A displacement intra prediction scheme is introduced for light field image compression in [29] [30]. However, the perceived quality of the compressed image has not been analyzed.

Moreover, at semantic level, the image content may be characterized by high level features, such as indoor, outdoor, sports, movie, and even different types of movies such as action or documentaries etc. Though there is no direct mapping of extractable features (such as color, texture, shape, structure, etc.) into semantic concept, the effective strategy in content analysis is to use attributes extractable from the source image [31]. In literature, few efforts have been devoted to analyze content features by exploiting the extractable attributes. The spatial perceptual information and temporal perceptual information is recommended to analyse the scene characteristics of video in [24]. The spatial information and colorfulness is used to analyse image scene in [32]. In [33] and [34] the parameters, spatial activity, temporal activity, spatial-temporal interaction, contrast, and colorfulness are considered to characterize video content. To the best of our knowledge, the impact of image content on the subjective quality has not been studied for light field image.

With respect to R3, there has been a growing interest and many efforts [35]–[38] have been given to the development of IQM. However, there are many challenges and issues regarding the applicability and performances of the metrics [11], [39]. The performances of the most widely used metrics have been investigated for 2D images [40] [41]. The image quality metric, Peak Signal-to-Noise Ratio (PSNR), has been used to compare the performance of light field image compression methods [42] [29]. However, the performances of the 2D metrics, including PSNR, have not been studied for light field images.

III. Dataset Formation

In this section, the research question R1 is addressed while in the following subsections, the steps performed for creating the light field image quality dataset are detailed.The dataset is a result of collaboration between Mid Sweden University and Università degli Studi Roma TRE. The term SMART reflects the aim to have a compact dataset as complete as possible, that is a set of SRCs with a wide range of key quality attributes affecting the subjective quality.

A. Source sequences

In the following, procedures employed for selecting the Source Sequences (SRCs), and SRCs description are presented.

Datasets	Date	Purpose	Features	Acquisition Devices	Depth Map (DM)
GUC Light Field Face and Iris Database [20]	2016	Face and Iris Recognition	It provides two biometric image databases collected by using a Lytro camera on multiple faces and visible iris (112 subjects for faces and 55 subjects for eye pattern).	Lytro	No
Lytro dataset [18]	2015	Light field Reconstruction	It provides 30 images, with indoor and outdoor, motion blur, long exposure time, and flat image.	Lytro	No
EPFL Light-Field Image Dataset [21]	2015	General	It provides 118 Lytro images with different categories: buildings, landscapes, people, etc.	Lytro Illum	No
LCAV-31 [17]	2014	Object Recognition	It provides light field images of 31 object categories captured from ordinary household objects and designed for object recognition purpose.	Lytro	No
Light Field Saliency Dataset (LFSD) [19]	2014	Saliency Map Estimation	It provides 100 light field images with 60 indoor scenes and 40 outdoor scenes.	Lytro	Yes (Estimated)
Synthetic Light Field Archive [15]	2013	General	It provides many artificial light fields including images with transparencies, occlusions, and reflections.	Camera (Artificial light field)	No
Light Field Analysis [16]	2013	Depth Map Estimation	It provides 7 Blender and 6 Gantry images; however, images do not cover the wide range of natural scenes.	Blender Software and Gantry device	Yes
Stanford Light Field Archive [14]	2008	General	It provides more than 20 light fields sampled using a camera array, a gantry and a light field microscope.	Gantry, Light field microscope, and Camera Array	No

TABLE I
LIGHT FIELD IMAGE DATASETS WITH CORRESPONDING BASIC FEATURES. THE GENERAL PURPOSE DATASET CAN BE USED FOR A WIDE RANGE APPLICATIONS SUCH AS TO BENCHMARK THE PROCESSING ALGORITHMS (ENCODING METHODS AND QUALITY METRICS) AND QUALITY ASSESSMENT.

1) SRCs content selection: A careful selection of SRCs is important to create a dataset. In particular, image features such as spatial information, color information, and brightness are important parameters. By using these parameters, it is possible to quantify the distortions suffered by data compression or transmission over a bandwidth-limited channel. Therefore, selected SRCs should span a wide range of content features.

In this article, based on the key quality attributes: Spatial distribution Information (SI), colorfulness (CF), contrast, correlation, homogeneity, brightness, hue, and saturation, the SRCs have been selected. The features were considered based on the image quality attributes and HVS characteristics [43]. SI is a perceptual indicator of spatial information of the scene [25]. Colorfulness is the main perceptual attribute underlying image quality and naturalness of the images [44]. Contrast is the most important parameter used to evaluate image quality [36], because the meaningful visual information is conveyed by contrast. As an example, a largely uniform picture carries little or no information [45]. Hue, saturation, and brightness are major attributes that are used to express the HVS characteristics of images [34]. The achieved results, presented in Section V-B3, show that the subjective quality of the light field image is significantly influenced by the image content, and that more than one content attribute should be considered for characterizing the scene.

A light field camera provides information about depth dependence and Lambertian lighting. The depth dependence implies multiple depth of semitransparent objects and the Lambertian surface reflects light with equal intensity in all directions [46]. The depth dependence information can be exploited during coding, and the variation in depth of field information could give different compression levels at a same quality level. Reflections and transparency are prevalent in natural images, that is, reflected and transmitted lights are super-imposed on each other. The image can be modeled as a linear combination of transmitted layer, which contains the scene of interest, and a secondary layer, which contains the reflection or transparency [47] [48]. The decomposition of the images into two layers is a ill-posed problem in the absence of additional information about the scene [49]. The light field camera recorded information, particularly multiple views of a single scene, can be exploited to solve the problem. Therefore, in a test dataset images with transparency, reflections, and wide Depth of Field (DoF) variation are needed.

2) Number of SRCs: Having a large number of SRCs is preferable in order to obtain a complete understanding of the phenomena under investigation. However, in some applications, in particular subjective quality assessment and pilot test of processing algorithms, the number of SRCs is limited by processing time, duration of experiment, and available number of the subjects for the experiment.

As specified in ISO 20462 standard [23], to get relative quality values in Just Noticeable Differences (JNDs), the selected attributes should appear in at least three images. A single image can cover more than one attribute; in this work two to three significant attributes per image have been considered. As consequence, we have captured 16 images.

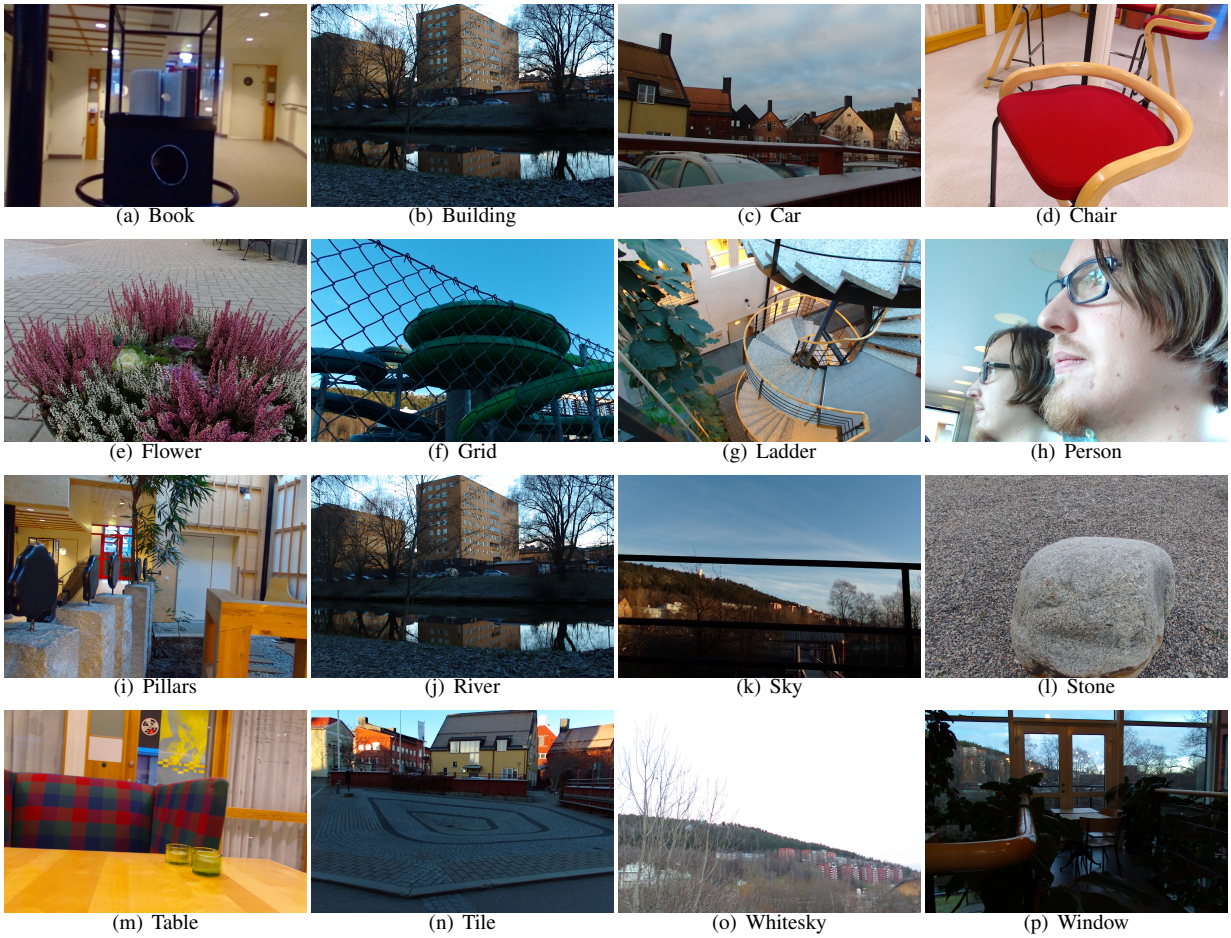


Fig. 1. All-in-focused 2D views of SRCs in SMART dataset.

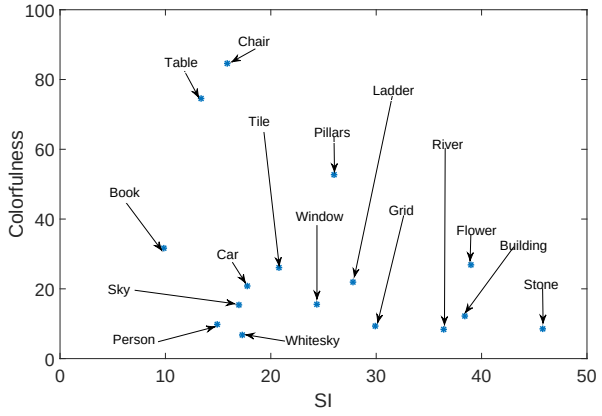


Fig. 2. Spatial and colorfulness information distribution of SRCs.

3) SRC description: Thumbnails of the selected SRCs are shown in Figure 1 and the corresponding key features are summarized in Table II. As can be noticed, the considered images cover a large number of quality attributes and content variations. The analysis of the SRCs in Figure 2, shows that they cover a wide range of key quality attributes [22].

B. Hypothetical reference circuits

To create Processed light field Image Sequences (PISs) from SRCs, test conditions (Hypothetical Reference Circuits (HRCs)) need to be defined [50]. Figure 3 shows that before the test image showing to the subject, the light field image goes through encoding and rendering steps. To the best of our knowledge, there is no standard encoding and rendering method. As can be noticed, HRCs can be generated at the encoder and/or at the renderer block.

In this work, to evaluate the perceptual quality of compressed light field images, rendering method is fixed while compression methods are varied. As HRCs, encoding techniques: JPEG, JPEG2000, HEVC intra [51] [52] and an ad hoc designed plenoptic image compression system (Sparse Set and Disparity Coding-SSDC) [30] have been considered

As mentioned in [51] and [40], four to five levels of distortion are sufficient to create the PISs. The distortion strengths have been selected after initial tests: a large set of test images has been generated. In more details, for JPEG the following quality levels (QLs): 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100; for JPEG2000 the compression ratios (CRs): 1, 25, 50, 75, 100, 125, 150, 175, 200, 250, and 300; for HEVC intra and SSDC the quantization parameters (QPs): 22, 23, 24,..., 50 have been used. Based on the following considerations, a subset of these images that spanned a wide

Index	Name	Description	Key Features	Remarks
(a)	Book	Book inside a transparent box	Homogeneity, Transparency	Indoor
(b)	Building	Building and its reflection on the river	SI, contrast, reflection	Outdoor
(c)	Car	Car roof and building with sky	Homogeneity, DoF	Outdoor
(d)	Chair	Chair on the floor	Colorfulness, DoF	Indoor
(e)	Flower	Flower with tile on the floor	SI, hue	Outdoor
(f)	Grid	Grid with natural scenes	DoF, hue	Outdoor
(g)	Ladder	Ladder top view	DoF, SI	Outdoor
(h)	Person	Close-up picture of a person with reflection	Reflection, contrast	Indoor
(i)	Pillars	Pillars	Colorfulness, DoF	Outdoor
(j)	River	Flower and river with reflection of the building	Contrast, DoF	Outdoor
(k)	Sky	Sky with natural scenes	Homogeneous, correlation	Outdoor
(l)	Stone	Stone on the concrete ground	SI, contrast	Outdoor
(m)	Table	Table with sofa	Colorfulness, Correlation	Indoor
(n)	Tile	Tile with background building	brightness , hue	Outdoor
(o)	Whitesky	Natural scene with white sky	brightness, correlation	Outdoor
(p)	Window	Natural outdoor scene with indoor objects	Transparency, DoF	Outdoor/Indoor

TABLE II
THE BRIEF DESCRIPTION OF SRCS. THE SRCS COVER A WIDE RANGE OF KEY IMAGE QUALITY ATTRIBUTES AND IMAGE CONTENT VARIATIONS.

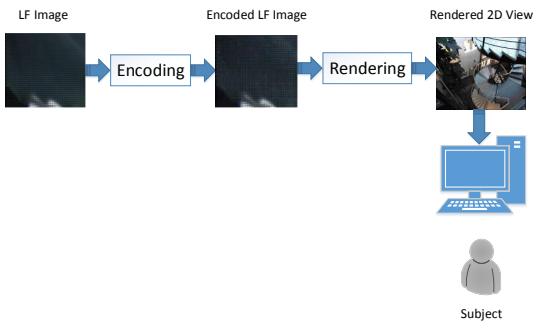


Fig. 3. Signal processing steps and experimental design.

Encoding methods	Level 1	Level 2	Level 3	Level 4
JPEG (QLs))	30	50	70	90
JPEG2000 (CRs)	25	50	100	200
HEVC intra (QPs)	32	37	42	47
SSDC (QPs)	32	37	42	47

TABLE III
COMPRESSION METHODS AND CORRESPONDING COMPRESSION LEVELS

.

range of visual quality scores (minimum to maximum quality) have been selected for the dataset, as detailed in [53]. The QPs used in literature for comparing performance of H.265/HEVC, VP9, and H. 264/AVC encoders are 22, 27, 32, and 37 [54]. From a preliminary test, the difference in subjective quality of rendered images with original image was very small for HEVC intra and SSDC for QPs 22 and 27. Therefore, higher values of QPs (42 and 47) have been used. The selected compression level for each compression method are shown in Table III.

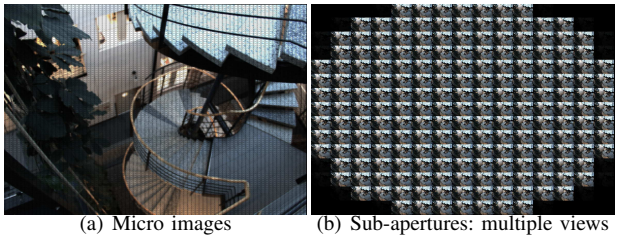
C. Light field image processing

The following steps were taken for preparing the test images:

1) Raw sensor data decoding: To decode the Lytro Illum recorded raw sensor data, Light Field Toolbox v0.4 [55] was used. Demosaicing, devignetting, and colour and gamma corrections were applied during the decoding [56]. Next, the content was converted into a light field data structure: a stack of 2D low-resolution RGB images in addition to the weighting image. The weighting image conveys the confidence associated with each pixel, which can be useful in filtering applications that accept a weighting term [55]. The resulting dimension of the light field data structure is $15 \times 15 \times 434 \times 625 \times 4$, where 15×15 represents the number of views, 434×625 represents the resolution of each view and 4 corresponds to the RGB and weighting image components. For the processing, the weighting component is discarded, i.e. only RGB color channels are considered. In Figure 4 (b), views around the corers are black, this is due to the hexagonal geometry of the micro lens.

2) Light field data arrangement: Light field data structure is depending on the targeted encoding method. In general, light field image encoding method can be divided in two categories: lenslet image (micro images) encoding and sub-aperture image encoding, based on the input (data structure) given to the encoder. In lenslet image encoding, the light field data need to be arranged as lenslet image (shown in Figure 4 (a)). In sub-aperture image encoding, sub-aperture images (multiple views) are given as input to the encoder. For an example, in [57], the sub-aperture images are used as the frames of the pseudo-video: coding order of the views accounts the similarities between adjacent views, and it is encoded as a video.

In literature, most of the light field image encoding methods use the HEVC intra profile [1], [30], [58]. Since, in this study, the HEVC intra profile together with standard image encoding methods are considered, and thus, the light field data is arranged in a lenslet image (2D image) structure. In more details, HEVC intra related encoders are used for encoding, and thus, 16 bit precision data is clipped to 8 bits by dropping the least significant bits, and RGB 8 bit uncompressed image is converted to YCbCr4:2:0 color space. The resulted light image is encoded and decoded. After de-



(a) Micro images (b) Sub-apertures: multiple views

Fig. 4. Light field image content arrangement.

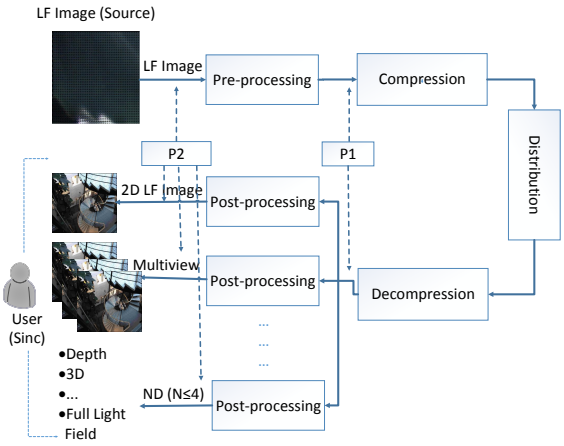


Fig. 5. Signal processing chain of light field image processing. At quality evaluation point, P2, the image HVS properties are more relevant than point P1. The final deliverable to the sink (user) could be 2D view, multiple views, depth map, stereo views, light field content, etc.

coding, the image is converted back to the RGB color channel for display. For HEVC compression, *HM software* (Encoder Version [11.0][Windows][VS 1700][64 bit]) [52] has been used, and test parameter configurations have been selected as detailed in [51] and [59].

D. Light field image visualization

The basic steps followed in the light field image processing system are shown in Figure 5. In addition to conventional 2D/3D processing steps (i.e., acquisition, processing, compression, storage, transmission, and reproduction) light field imaging is affected by increased computation complexity as described in the following. In pre-processing (raw data decoding and representation) step, information (such as color) may be lost and/or distorted. Due to the amount of light field camera recorded information, aggressive compression of the light field content is needed. Next, in most of the applications, since the end user is a human subject, the light field content is shown to the final user after post-processing, such as rendering. The selection of rendering method and the resulting introduced artifacts are depending on the targeted visualization techniques: refocused image, extended depth of field, 3D image, parallax, depth, 360 degree light field display [60], etc.

The quality of compressed light field image can be assessed at two points:

- P1: The effect of compression on light field content may be evaluated before the post processing such as rendering. As an example, the quality of compressed image can be evaluated with the help of objective quality metric such as PSNR.
- P2: The decoded light field content is usually rendered before being shown to the final user. The light field content can be shown in different ways, such as 2D view (refocused at different point/depth or all-in-focused view), multiple views, 3D views, depth information, full light field, etc.

In general, light field content is shown to the final user after the post-processing, and thus image HVS properties are more pertinent at point P2 than P1. In this work, the subjective quality is evaluated at point P2. The subjective quality evaluation of light field content is even more difficult to properly address the concept of quality. This is also due to the fact that light field image can be shown to the user in different ways, and there are many possibilities for its visualization. Based on the easily availability of 2D display devices, possible approaches are:

- *Random selection of refocused single view:* A single refocused view can be shown to the observer, so that only a subset of light field data would be evaluated. In the refocused view, out of focus area of the scene will be blurred, and thus the complete scene cannot be evaluated equally.
- *Random selection of multiple refocused views:* In this technique, bigger amount of light field data can be evaluated than for a single view. There will be many views and refocus points, and it is not possible to consider all views. As a result, for the subjective evaluation randomly selected 3 views and 3 focus points are recommended in [42]. The major limitations of this technique are: 1) as a result of refocusing, the out of focus area will be blurred and entire image scene can not be evaluated uniformly, 2) the consideration of 3 refocused views can not cover the complete light field information, 3) the refocused images, refocusing at foreground object or in background scene, could have a different subjective quality, and 4) there is no validated protocol to run the subjective experiment for the refocused views.
- *Pseudo-video sequence:* The multiple views of light field image can be used to create a video sequence. The video sequence could be circular animated video sequence, such as bobbling; light field image is displayed with a circular animation [61] [56]. However, this technique may not be a good option for refocused views. Subjective quality of refocused views could be different based on the point of focus and refocusing algorithm. In particular, handling inter-perspective aliasing ("jumps" between the views), as a result of large depth variation, is challenging. Moreover, there is no validated subjective assessment protocol for this technique.
- *3D visualization:* Light field data can be rendered for 3D visualization by estimating depth information from the light field data. A light-field 3D rendering using Depth-

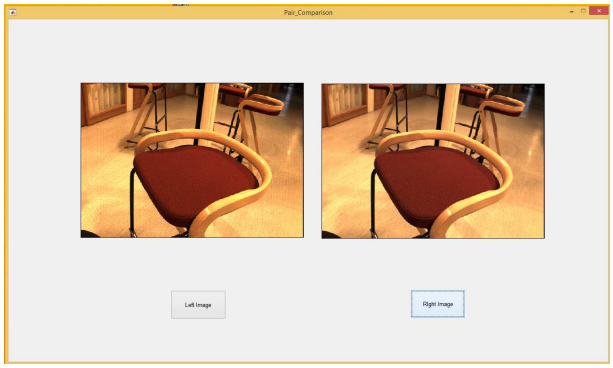


Fig. 6. Subjective experiment setup.

Image-Based Rendering (DIBR) is presented in [62]. The 3D visualization technique could be a good option to evaluate the quality of depth map; however, there is no standard 3D rendering technique.

- *All-in-focused view:* In this work, all-in-focused view of the light field image is used for subjective quality evaluation. Motivation behind the selection of this method is that the foreground as well as the background information can be shown to the subject without blurred regions. Moreover, validated image quality assessment protocols can be adopted. The major issues of this technique are: subject can evaluate only few part of the light field content and the level of distortion on the depth map may not be evaluated. To create the view, a basic or full resolution rendering method has been exploited [63]. The resulted image is equivalent to the central view of the light field image.

IV. Subjective Evaluation

In the following, the protocol adopted during the subjective experiment is described.

A. Subjective experiment methodology

Given the novelty of this technology, during the selection of subjective quality assessment method, the following issues:

- subjects are not familiar with the light field content;
- light filed camera recorded raw data decoding process introduce the artifacts (such as color distortions, blurring, Gaussian noise, and Salt and Pepper noise), and thus, the quality of reference image is already (before compression) reduced;

need to be considered. For reducing the impact of the pre-introduced artifacts, the double stimulus paradigm is an appropriate choice for the subjective quality assessment. Therefore, to collect the opinion scores for the rendered 2D views, the Pairwise Comparison (PC) method [24], is selected. The main advantage of PC is its high discriminatory power, which is of particular value when several test items are nearly equal in quality [24].

In PC, a pair of images is presented to the observer who selects the one that, according to his/her opinion, has best image quality. The Graphical User Interface (GUI) used in this experiment is shown in Figure 6. At the end of each paired presentation, the subject expressed his/her preference by ticking the boxes. The number of pairs is $n \times (n-1)$ for n number of images. In our case, there were 17 test images (272 pairs) for each SRC, and we had 16 SRCs. The result of the PC experiment is a PC Matrix (PCM), that contains the number of times that each option was preferred over the other option [64].

B. Subjects

To collect reliable results, 19 subjects have been selected [23] [25]. The subjects were drawn from a pool of undergraduate to post-doctorate students from Università degli Studi Roma TRE, Rome, Italy. The subjects were naive concerning image impairments and the associated terminology. They were asked to wear any vision correcting devices (glasses or contact lens) that they normally wear. The light field image is displayed as a classical 2D image (all-in-focused view), and thus specific visual activity tests such as color and vision tests, are not performed as for HD/3D image.

C. Experiment length

To minimize the effect of viewers' fatigue on quality assessment, four experimental sessions have been scheduled. The experiment length is maintained shorter than 30 minutes by dividing each session into two sub-sessions to retain the attention of subjects [25]. Each sub-session lasted 12 to 15 minutes including evaluation and training time of 2 minutes, and at least 5 minutes gap between each sub-session.

D. Stimuli arrangement

In the experiment, PISs from 2 SRCs have been evaluated in each sub-session. The PISs from the same SRC are displayed at a time. To compensate the effect of a potential bias based on order or position of stimuli in the averaged results [65], stimuli are shown in random order for each subject according to their compression and intensity.

E. Training

Before the experiment, the subject is verbally given instructions, followed by written instructions as specified in ITU-T Recommendation. P.910 (Appendix II) [24].

In the training stage, the subject is shown pairs of images having different levels of impairment, from the lowest to the highest found in the experiment. In this phase, each subject gets familiar with the assessment procedure and establish the annoyance values range. The images used in the training session are different from the test images.

F. Apparatus and environment

The experiment is conducted in a controlled environment in order to produce reliable and reproducible results by avoiding involuntary influence of external factors [25]. The characteristics of the display device and system are used in the experiment are briefly described in Table IV. The viewing conditions and viewing distance (1–8H, H = picture height) are maintained as specified in [24].

TABLE IV
SYSTEM AND DISPLAY PARAMETERS.

(a) System parameters.

Parameters	Values
Processor	Intel(R)Core(TM)i7-4770
Processor Speed	@3.40GHZ
RAM	8GB
System type	64-bit OS
Operating System	Windows 8.1
GUI	MATLAB R2015a

(b) Display parameters.

Parameters	Values
Display Device	DELL U2413f
Screen Refresh Rate	60Hz
Screen Resolution	1920x1200 pixels
Brightness and Contrast	50
Sharpness	50
Aspect Ratio	Wide 16:10

V. RESULTS ANALYSIS AND DISCUSSIONS

In this section, the research questions R2, and R3 have been addressed.

A. Overview of the SMART dataset

In the following the contents provided in the dataset (available at http://www.comlab.uniroma3.it/SMART/) are briefly reported.

1) Light field content: The dataset includes the following test materials.

- *Raw light field images:* The SRCs in raw format (light field content and camera calibration data) that can be used for a wide range of applications such as benchmarking processing algorithms (raw data decoding, compression, rendering, etc.) or quality evaluation.
- *Decoded light field images:* Decoded light field data are available in a light field data structure: matrix of the multiple views. The decoded light field images can be used directly (without raw data decoding and processing) to train, test, and benchmark the processing algorithms and for quality assessment.
- *Encoded light field image:* The encoded or distorted light field images can be used to study the response of HVS to the light field content.

2) Collected quality scores: To widen the possible application of the SMART dataset, raw and processed opinion scores of the test (reference and distorted) light field images are provided.

As mentioned in Section IV-A, the result of the subjective experiment is the PCM matrix. The matrix contains the number of times that each option (HRC) was preferred. Therefore, there is a matrix corresponding to each SRC, since distorted images of the same SRC were compared. The matrix (subjective opinions) can be analyzed and interpreted in different ways.

To analyze the subjective quality, a preliminary step is to convert the results in a continuous rating scale [66]. In this article, the conversion of the preference frequencies to continuous-scale quality scores is performed by using the Bradley-Terry (BT) model [67] [68]. A detailed description of the BT score and Confidence Interval (CI) estimation procedure is available in [69]. Higher BT score indicates higher preference rate. Therefore, the BT value can be used as a quality score (mean opinion score). The distribution of the BT scores for the test images is shown by means of the histogram in Figure 7. The histogram shows that HRCs are well selected for creating a set of data spanning a wide range of subjective quality. The quality scores are in the range 1 (excellent quality) to -12 (worst quality). In general, the range of the quality scores is depending on the difference between the preference scores of the image pair under test and employed mapping model.

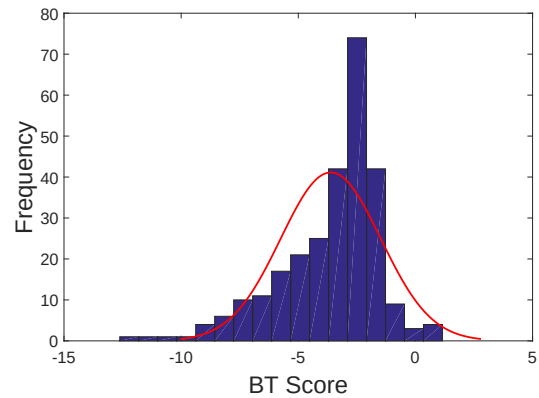


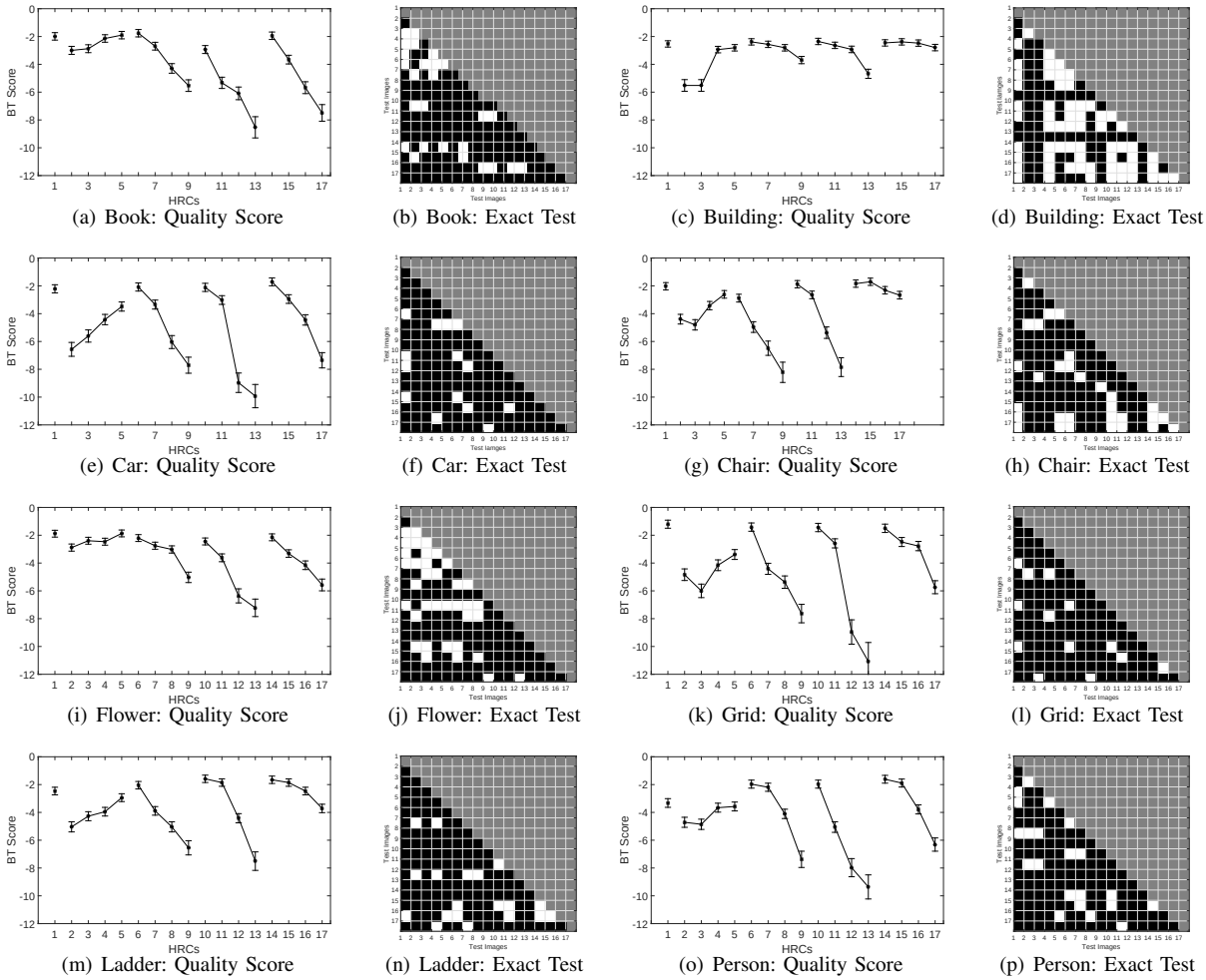
Fig. 7. Distribution of BT scores. The test light field images have a wide range of subjective quality scores.

In PC method, the subjects are only requested to indicate the binary opinion (better or worse) between two images under test. This kind of comparison is less confusing than discrete and/or continuous quality rating system [70], [71], and thus scores given by all the subjects are considered for the study [24].

B. Subjective quality of compressed light field images

To address the issue arisen by R2, the analysis of the subjective quality of compressed light field image is presented. For the analysis, the numbers 1 to 17 have been used as indexes to represent 17 HRCs. The index numbers and corresponding HRC description are shown in Table V. The impact of compression on the quality is analyzed in two steps. First, the overall impact of the HRCs on the subjective quality is presented and then, the same analysis is performed for each SRC.

1) Overall impact of HRC: To analyze the subjective quality of light field images of different HRCs, the opinion scores given by all the subjects and for all SRCs are used. The results reported in Figure 8, show that the subjective quality of processed light field images varies with the selected HRCs. This result is further confirmed by the Analysis Of Variance (ANOVA) [72]. During the testing, test statistic is measured with the help of the *Fisher-Snedecor distribution*, indicated as a F_{value}. If the probability, p_{value}, for the F-statistic is smaller than the significance level, then the test rejects the null



(a) Book: Quality Score (b) Book: Exact Test (c) Building: Quality Score (d) Building: Exact Test
(e) Car: Quality Score (f) Car: Exact Test (g) Chair: Quality Score (h) Chair: Exact Test
(i) Flower: Quality Score (j) Flower: Exact Test (k) Grid: Quality Score (l) Grid: Exact Test
(m) Ladder: Quality Score (n) Ladder: Exact Test (o) Person: Quality Score (p) Person: Exact Test

Fig. 10. BT Score with 95% confidence interval and result of exact test for eight images: Book, Building, Car, Chair, Flower, Grid, Ladder, and Person. The Quality Score figure shows that the BT score with 95% confidence interval for 17 HRCs; the index number and corresponding description are shown in Table V. In Exact Test sub-figures, white box indicates quality scores for the test points (pairs) are not significantly different and black box indicates the quality scores are significantly different, where as gray box is used to express the test is not necessary for the pairs.

Index	Artifacts	Index	Artifacts
1	Reference		
2	JPEG (q=30)	10	HEVC intra (QP=32)
3	JPEG (q=50)	11	HEVC intra(QP=37)
4	JPEG (q=70)	12	HEVC intra (QP=42)
5	JPEG (q=90)	13	HEVC intra (QP=47)
6	JPEG2000 (CR=25)	14	SSDC (QP=32)
7	JPEG2000 (CR=50)	15	SSDC (QP=37)
8	JPEG2000 (CR=100)	16	SSDC (QP=42)
9	JPEG2000 (CR=200)	17	SSDC (QP=47)

TABLE V
INDEXES AND CORRESPONDING ARTIFACTS (Q=QUALITY LEVEL, CR=COMPRESSION RATIO, QP=QUANTIZATION PARAMETER LEVEL).

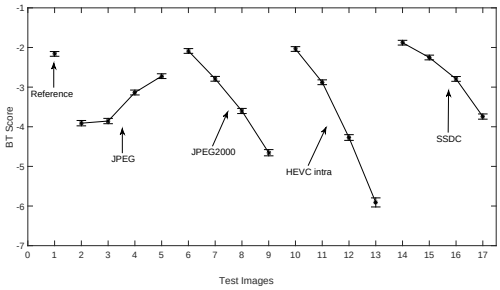
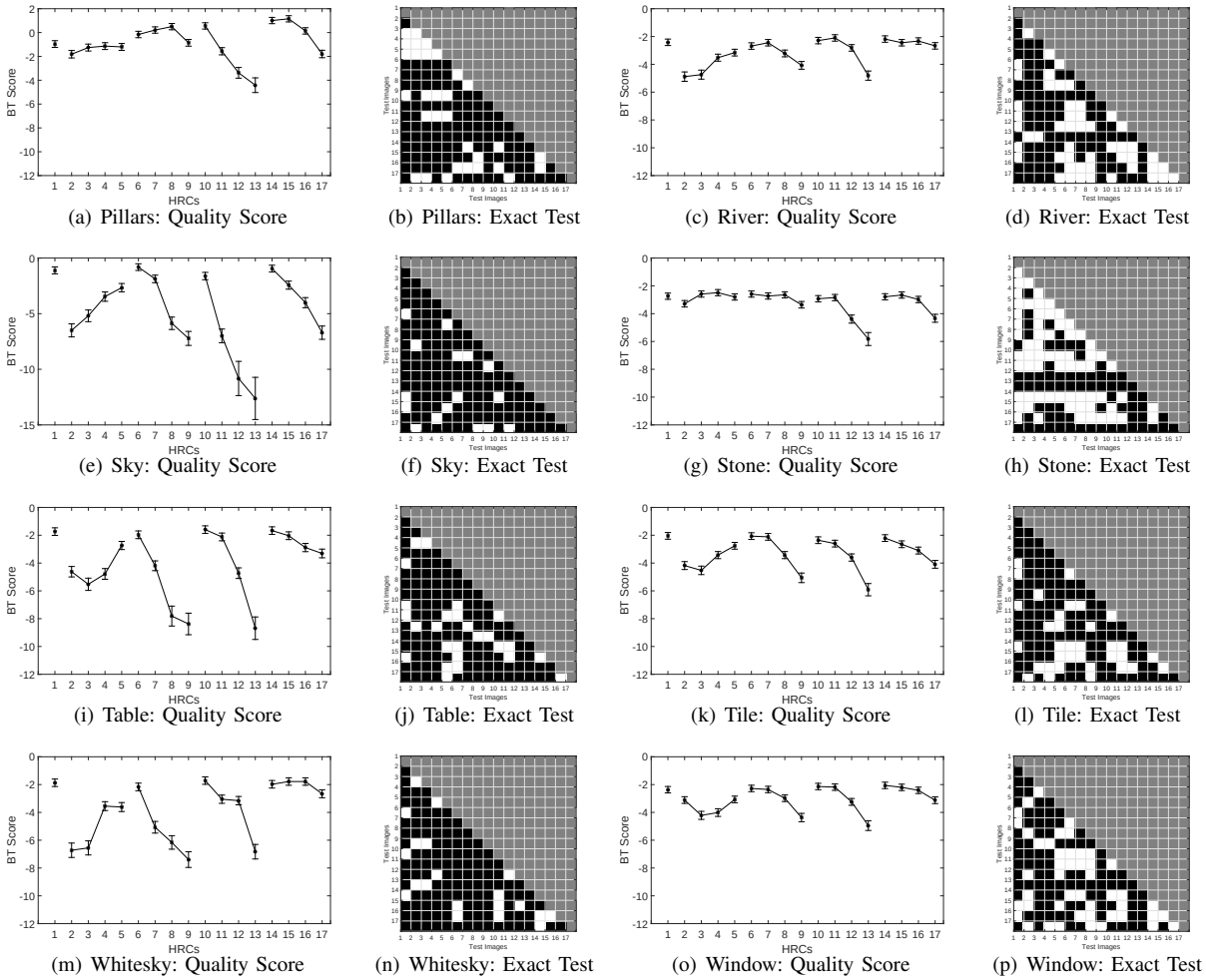


Fig. 8. Overall impact of HRCs the subjective quality: BT score with 95% of Confidence Interval (CI).

hypothesis, i.e. accepts the alternative hypothesis (at least one of the group means is significantly different from the others). In this work, a significance level of 0.05 is considered. The result $F(16, 255) = 18.1$ and p-value $\simeq 0 < 0.05$, indicates that the quality scores for the HRC are significantly different.

2) Impact of HRC on individual SRC: To study the influence of HRCs for individual SRCs, the result has been further analyzed by using the box plot. The box plot (box and whisker diagram) is used to display the distribution of data based on the five number summary: minimum, first quartile, median, third quartile, and maximum [73]. From Figure 9 it can be noticed that some of the points are outliers and that the range of the first and third quartile is noticeable. This result indicates that the BT scores are different for the SRCs even for the same level of compression. Therefore, it is worthwhile to analyze the



(a) Pillars: Quality Score (b) Pillars: Exact Test (c) River: Quality Score (d) River: Exact Test
(e) Sky: Quality Score (f) Sky: Exact Test (g) Stone: Quality Score (h) Stone: Exact Test
(i) Table: Quality Score (j) Table: Exact Test (k) Tile: Quality Score (l) Tile: Exact Test
(m) Whitesky: Quality Score (n) Whitesky: Exact Test (o) Window: Quality Score (p) Window: Exact Test

Fig. 11. BT Score with 95% confidence interval and result of exact test for eight images: Pillars, River, Sky, Stone, Table, Tile, Whitesky, and Window. The Quality Score figure shows that the BT score with 95% confidence interval for 17 HRCs; the index number and corresponding description are shown in Table V. In Exact Test sub-figures, white box indicates quality scores for the test points (pairs) are not significantly different and black box indicates the quality scores are significantly different, where as gray box is used to express the test is not necessary for the pairs.

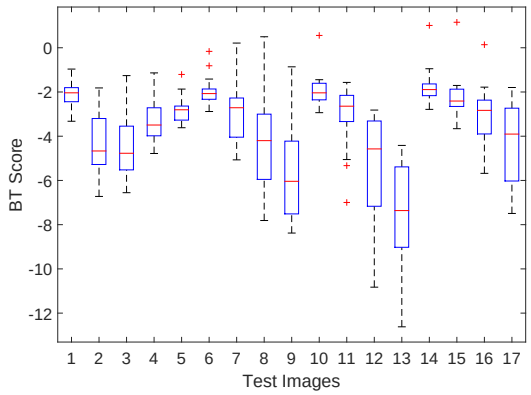


Fig. 9. Box plot of the BT scores at different level of HRCs for 16 SRCs. On each box, the central mark is the median score, the edges of the box are the 25th and 75th percentiles, and the outliers are plotted individually.

subjective quality of individual scene at different compression levels.

The subjective quality of the SRCs for different HRCs is shown in Figures 10 and 11. As can be noticed, the patterns of quality scores are noticeably different for the SRCs even at the same level of compression. As an example, variation on the BT scores is smaller for *Building* compared to *Car* for the same HRCs or at a same level of compression.

To confirm the results, Barnard's test [74] has been considered to verify whether the probability P_{ij} (scores given for a test sequence/HRCs) is significantly different from a probability of 0.5 (i.e., whether the observers are undecided) or not [75]. For example, at a 95% confidence, p-value < 0.05 means there is significant difference between the probabilities that observers chose image I_i over I_j.

In Figures 10 and 11, the difference of the subjective quality of test images is expressed for each possible pair. It is assumed that each pair of images follows the commutative law. Therefore, the region of interest is the left side of the matrix plot. As can be noticed, if the opinion scores for the pairs of images are not significantly different, the corresponding square is filled with white box otherwise it is black, where as gray box is used to express the test is not necessary. The Figures 10 and 11 show that different set of SRCs have different numbers

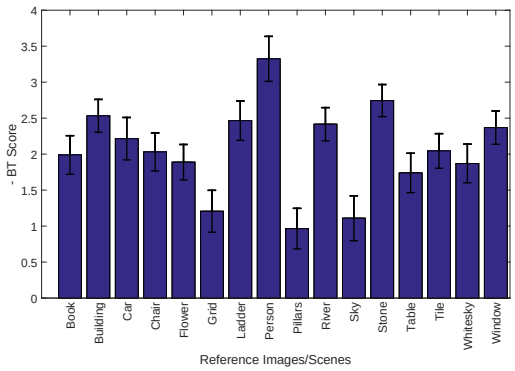


Fig. 12. BT score with 95% CI for reference light field images. The subjective quality is different for reference images.

Descriptors	SI	CF	Contrast	Homogeneity
PLCC	-0.184	0.317	-0.344	0.347
SROCC	-0.267	0.352	-0.423	0.452
Descriptors	Hue	Saturation	ColorValue	Brightness
PLCC	-0.047	**0.520**	0.018	-0.150
SROCC	-0.229	0.394	0.191	-0.050

TABLE VI
CORRELATION COEFFICIENT: RELATIONSHIP BETWEEN QUALITY SCORES (BT SCORE) AND IMAGE ATTRIBUTES.

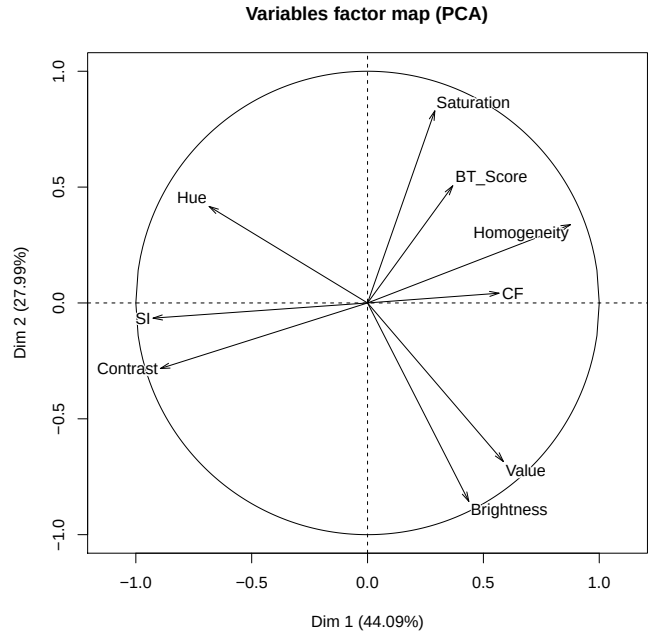


Fig. 13. Result of PCA. The wide angles between the content descriptors and BT score indicate there is no strong correlation between the perceived quality and the attributes.

of white squares. This result indicates the subjective quality of light field images is also influenced by the SRC or image content.

Moreover, a one-way-ANOVA test is performed for a group of SRCs at different HRCs to confirm the result. The result, $F(15, 256) = 3.71$ and p-value $\simeq 0 < 0.05$, indicates that the subjective quality of the test images are significantly different for the SRCs. This is because the image content has a significant impact on the quality [13].

3) Effect of image content on subjective light field image quality: The performed analysis, reported in Figure 12, shows that the subjective quality is significantly different even for the SRCs. This result strengthens the need to investigate the impact of image content on the subjective quality.

At the moment, no standard content descriptors have been defined. In this work, key image quality attributes defined in Section III-A1, have been used. Based on the descriptors scores and corresponding image quality scores, correlation analysis has been performed.

For the analysis, Pearson Linear Correlation Coefficient (PLCC) and Spearman Rank Order Correlation Coefficient (SROCC) have been used. The correlation between the content descriptors scores and corresponding subjective quality scores is shown in Table VI, and it shows that there is no significant correlation between the considered descriptor with the subjective quality of the light field image.

The result is further confirmed with the help of Principle Component Analysis (PCA). In the PCA the correlation between the variables is measured in terms of the angle between the variables, where a small angle corresponds to a higher correlation [76] [77]. The PCA has been performed by using *FactoMineR* Software [78] and the result, Figure 13, shows that the descriptor saturation has a small angle with BT score compared to other descriptors. However, the angle between the BT score and saturation is noticeable.

The results show that there is no strong correlation between the SRC descriptors and subjective quality. Though, the quality of the SRCs is significantly different. This result could be due to the fact that together with system factors (such as transmission impairments and compression artifacts) the subjective quality is influenced by human and context factors [79] too. However, the considered descriptors mostly explain the system factors and the subjective quality is more about the subjects, and hence physiological, psychological, social, and role-related aspect of the subject also influence the quality score [80].

C. Subjective quality analysis of the compression methods for light field image

From Figure 8, it can be noticed that the subjective quality of SSDC compressed light field image is high when compared to other compression methods, and the same result is replicated in Figures 10 and 11 for different SRCs. In those plots the test points for the compression methods are different: quality level for JPEG, compression ratio for JPEG2000, and QPs for HEVC intra and SSDC. To compare the performance the bit per pixel (bpp) for each test image is computed and compared with the corresponding subjective quality scores.

From Figure 14, we can observe that the subjective quality of JPEG2000 is significantly high compared to JPEG at same level of bpp and a similar trend is shown for SSDC and HEVC intra. These results indicate that the SSDC outperforms the other methods for the light field images. A detailed description

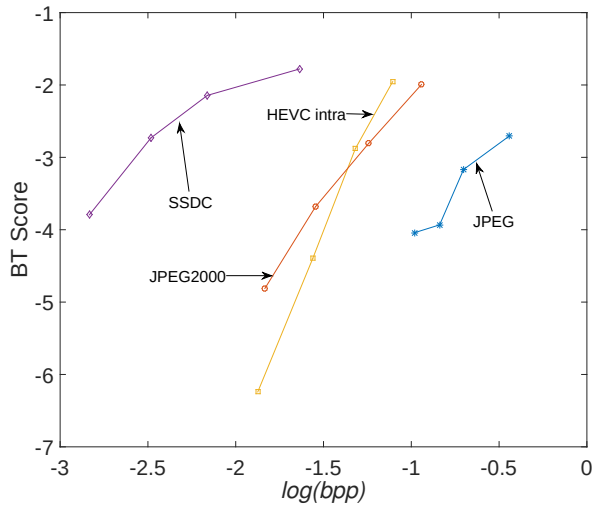
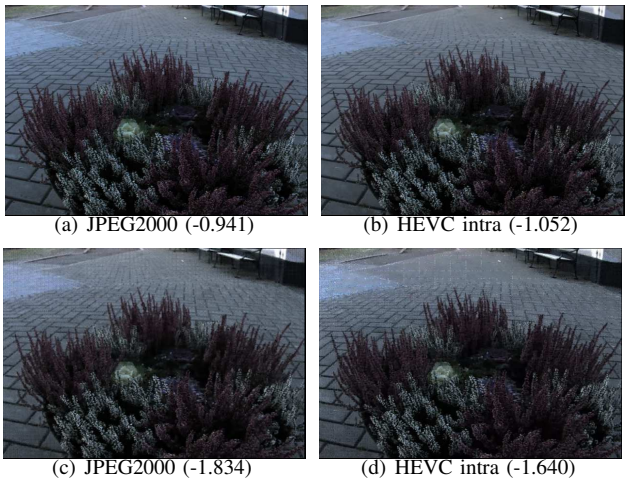


Fig. 14. Performance of the considered compression methods: subjective quality of light field images at different level of compression by the different compression methods. The recently proposed light field image compression method, SSDC, has a significantly high level of subjective quality compared to other methods.



(a) JPEG2000 (-0.941) (b) HEVC intra (-1.052)

(c) JPEG2000 (-1.834) (d) HEVC intra (-1.640)

Fig. 15. Quality of light field images at low and high level of compression for JPEG2000 and HEVC intra, and level of compression is expressed in terms of *log(bpp)*.

of SSDC is available in [30], where the objective performance of this compression method on light field images from focused plenoptic cameras, such as Raytrix, is described. The results of our work are different in two ways: 1) comparison of the compression methods is performed on light field images from a light field camera, Lytro Illum, and 2) the subjective quality of the compressed image has been evaluated by using a subjective assessment, and results show that the SSDC performs better than other considered methods.

Moreover, it can be noticed that there is a crossover between the curves for JPEG2000 and HEVC intra. This crossover indicates that, at a high level of compression, JPEG2000 produced blurring and ringing artifacts having a small impact on the subjective light field image quality compared to the block artifact produced by the HEVC compression [81]. This

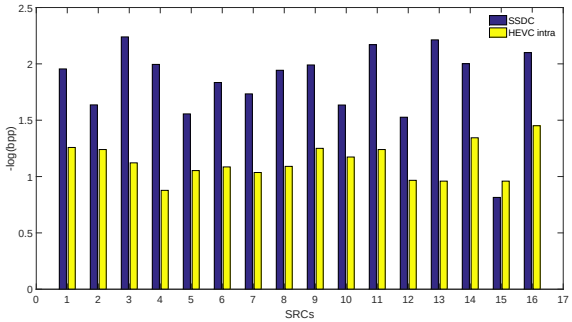


Fig. 16. The compression level of SRCs at QP = 32 for top performing compression methods, SSDC and HEVC intra.

Compressions	SI	CF	Contrast	Homogeneity
SSDC	0.013	-0.302	0.036	-0.046
HEVC intra	-0.008	0.399	0.089	-0.153
	Hue	Saturation	Value	Brightness
SSDC	-0.404	-0.536	0.451	0.617
HEVC intra	-0.290	-0.243	0.844	0.780

TABLE VII
PLCC: CORRELATION COEFFICIENT BETWEEN IMAGE ATTRIBUTES AND IMAGE COMPRESSION LEVELS MEASURED IN TERMS OF BPP (COMPUTED AT QP = 32).

result is further confirmed by Figure 15. Figure 15 (a) and (b) show that at low level of compression the HEVC intra compressed light field image has a high quality even the level of compression is higher than for JPEG2000 but at high level of compression (Figure 15 (c) and (d)) the JPEG2000 compressed image has high perceptual quality. It is useful to notice that the employed values of QP (42 and 47) are significantly higher than the commonly used QPs (22, 27, 32, and 37) [54].

Influence of image content: The result shown in Figure 16 indicates that even at a same QP level the SRCs have a different level of compression. Moreover, the effect of the image content on the subjective quality of light field image is analysed with the help of content descriptors for top performing light field image compression methods, HEVC intra and SSDC. The result, Table VII, shows that the level of compression significantly depends on the image color value and brightness of the light field image. Therefore, inclusion of a wide range of content variations is important in the dataset.

D. Performance analysis of image quality metrics

As stated in R3, in the following the performances of the 2D IQMs when applied for light field image have been analyzed.

As mentioned in Section I, in literature [1], [30], to evaluate the performance of encoding methods objective quality metric (PSNR) is used. However, several 2D metrics, such as MSE and SSIM, could be used for evaluating the light field image quality. To the best of our knowledge, performance of these metrics has not been yet evaluated on the light field image. Therefore, in this work, the performance of the well accepted and validated metrics when applied to light field image is evaluated.

In general, based on the availability of the reference signal, the level of computational complexity, and the provided accuracy, the objective metrics are categorised into three groups: Full Reference (FR), Reduced Reference (RR), and No Reference (NR). The FR metrics exploits complete information about reference image, while NR do not require any reference information. In RR only partial information is needed. The availability of FR information allows better prediction of quality. In particular, FR metrics are easy to compute and provide accurate results. Therefore, the FR metrics are widely used. In literature, many FR metrics [11], [40], [82], [83] have been proposed. However, there are no quality metrics specifically designed for light field images, and thus, in this work, we analyze the performances of the following 2D metrics:

- Mean Squared Errors (MSE) [84]
- Peak Signal-to-Noise Ratio (PSNR) [84]
- Signal-to-Noise Ratio (SNR) [84]
- Weighted SNR (WSNR) [85]
- PSNR based on Human Visual System (PSNR-HVS) [86]
- PSNR-HVS-M [87]
- Visual Signal-to-Noise Ratio (VSNR) [88]
- Structural Similarity Index (SSIM) [36]
- Multi-scale Structural Similarity Index (MSSIM) [89]
- Image information and Visual quality (VIFP) [90]
- Universal image Quality Index (UQI) [91]

For encoding, the light field data is arranged in a 2D image (lenslet) structure. The distorted image resulted from the encoder is also in 2D image structure. Therefore, quality score of the distorted light field image is estimated by using the FR metrics with the help of reference (before encoding) light field image.

Correlation between the subjective opinion scores and objective quality scores has been computed by using PLCC, SROCC, and Kendall Tau rank Correlation Coefficient (KTCC). PLCC evaluates the linear relationship between two continuous variables: subjective and estimated quality score. SROCC evaluates the monotonic relationship between two continuous or ordinal variables. Finally, KTCC rank correlation is a non-parametric test that measures the strength of dependence between two variables [92]. KTCC is considered since it is less sensitive to errors and discrepancies in data with respect to SROCC. High correlation values indicate that best prediction capability of the metrics with respect to subjective scores.

Figure 17 illustrates that the objective IQMs perform differently, and the PSNR-HVS-M has higher values of PLCC and SROCC compared to other metrics. This result could be due to the fact that the PSNR-HVS-M takes into account the HVS and the contrast sensitivity of the image [87]. A similar result is yielded from PCA, as shown in Figure 18, the PSNR, PSNR-HVS, and PSNR-HVS-M have a small angle with subjective quality score (BTScore) with respect to other metrics.

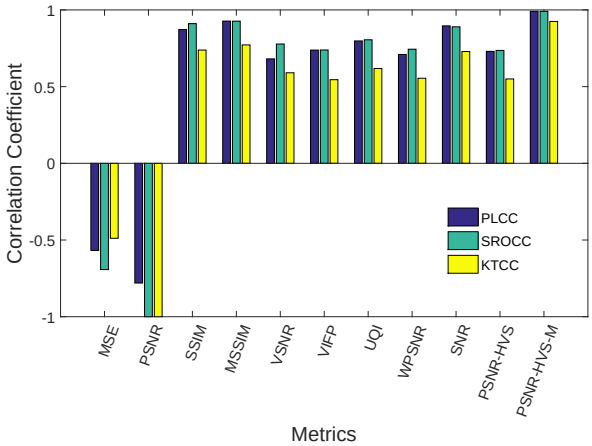


Fig. 17. Objective metric performance: correlation coefficient between the estimated quality scores and subjective opinion scores.

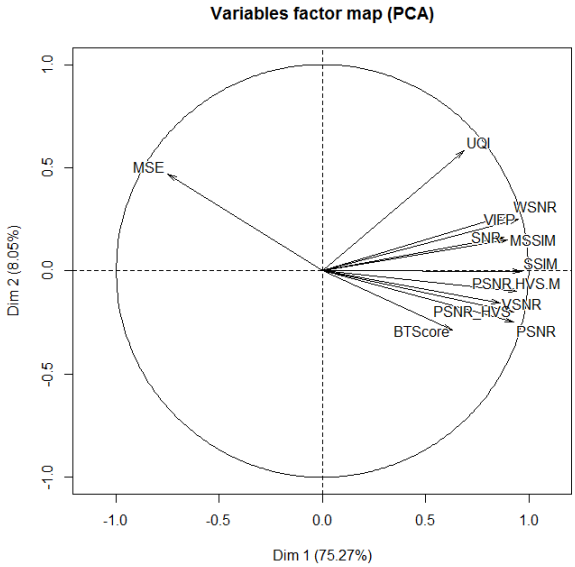


Fig. 18. PCA results: bipolar plot for image quality metrics performance analysis.

VI. Challenges and possible future directions

Several challenges remain to be tackled on the subjective and objective quality evaluation of the light field images. To the best of our knowledge, no works have been performed in this direction. Some of the issues arisen with our work are reported in the following and they could be the motivations and directions for future work.

- In this work, to decode the Lytro Illum recorded light field information, MATLAB Light Field Toolbox [55] has been used. During the decoding process some of the color and depth informations are distorted, though the color correction function available in the toolbox was considered. The distortion produced by the toolbox may have an influence on the perceptual quality of light field image.
- Visualization technique: As mentioned before, the light field image can be shown to the user in many ways. Some of the possible ways and their features are mentioned in Section III-D. There are still open questions such as

which is the appropriate visualization method considering the complexity of the light field contents? How to evaluate the multiple views or the complete light field content? How to evaluate the influence of depth distribution and distortion on the perceived quality? How to handle the inter-perspective aliasing, as a result of large depth?, etc.
- Rendering method: There is no standard rendering method for light field image. Many ongoing efforts have been given to develop the rendering methods. However, the selection of the method is also depending on the possible application of the light field image content. In our study, the basic full resolution rendering method has been adopted. The results that we got in this experiment could be different for other rendering methods since the distortion introducing by the rendering methods would be different.
- Depth information: The depth information is a valuable information that we get from the light field image. When we process the light field image, even during the compression, the depth information is also distorted. Therefore, for the objective metric a measure of depth distortion is also important. Moreover, the compression algorithm may work differently well with different depth distributions. On the other hand, we have only assessed the central 2D view, the adopted visualization technique may not be a good choice to measure the depth distortion. In this situation, we might have to employ different kind of visualization techniques such as 3D visualization or animated GIF (by considering all possible 2D views).
- In this article, we adopted a subjective quality assessment method (PC) but it needs to be developed further in order to cover the issues such as content complexity and several application areas.
- Finally, it is also worthwhile to have a similar study on light field images from a focused plenoptic camers such as the Raytrix.

VII. Conclusion

In this paper, a study on the perceived quality of compressed light field image is presented. Major contributions of the article are described in the following.

- A new SMART light field image quality dataset has been created. For dataset population, based on the key image quality attributes (colorfulness, spatial perceptual information, contrast, etc.) and light field camera specific capabilities (depth dependence and Lambertian lighting) the SRCs are selected and captured by Lytro Illum Camera. The captured SRCs are processed and compressed. An experiment was scheduled to collect the subjective quality rating for the processed image sequences. Source sequences, test sequences, subjective quality scores, and adopted experimental setup procedures are made freely available for the research community.
- The impact of the compression artifacts on the subjective light field image quality is studied. The results show that the compression methods significantly degrade the subjective quality of light field image, and the level of the quality degradation varies for the images with different content.
- The impact of the image content on the subjective quality is studied with the help of key image quality attributes and subjective quality scores of the light field images. The results show that there was no strong correlation of the descriptors with the corresponding subjective quality. However, the results presented in Section V-C show that the level of compression of light field images varies for different content even at a same level of QPs. Therefore, the inclusion of SRCs with a wide range of content variation is important.
- The performances of compression methods are evaluated for light field images. The results show that the recently proposed plenoptic image compression method, SSDC [30], has a better subjective quality at a same level of compression compared to other considered compression methods. The results demonstrate the importance of specific compression algorithms for light field data in order to reach a best possible quality.
- The performances of widely used 2D image quality metrics are evaluated for the light field images. The results show that, among the considered metrics, the quality scores predicted by PSNR-HVS-M has a better correlation with corresponding subjective quality scores.

VIII. Acknowledgement

Most of this work has been carried out in the Short Term Scientific Missions (STSM) under the scope of EU ICT COST Action IC1105, 3D content creation, coding and transmission over future media networks (3D-ConTourNet).

Moreover, we are very grateful to the people who gave their time for the subjective experiment.

References

[1] C. Conti, L. Ducla Soares, and P. Nunes, "HEVC-based 3D holoscopic video coding using self-similarity compensated prediction," *J. Signal Processing: Image Communication*, vol. 42, pp. 59–78, 2016.
[2] G. Lippmann, "Sur la théorie de la photographie des couleurs simples et composées par la méthode interférentielle," *J. Phys. Theor. Appl.*, vol. 3, no. 1, pp. 97–107, 1984.
[3] L. Gabriel, "Epreuves reversibles photographies integrals," *J. Comptes-Rendus Academie des Sciences*, vol. 146, pp. 446–451, 1908.
[4] H. E. Ives, "Parallax panoramagram," Jul. 18 1933, US Patent 1,918,705.
[5] E. H. Adelson and J. Y. Wang, "Single lens stereo with a plenoptic camera," *IEEE Trans. on pattern analysis and machine intelligence*, vol. 14, no. 2, pp. 99–106, 1992.
[6] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Computer Science Technical Report CSTR*, vol. 2, no. 11, pp. 1–11, 2005.
[7] A. Lumsdaine and T. Georgiev, "The focused plenoptic camera," in *IEEE Int. Conf. on Computational Photography (ICCP)*, 2009, pp. 1–8.
[8] T. Georgiev and A. Lumsdaine, "Depth of field in plenoptic cameras," in *Proc. Eurographics*, 2009, pp. 1–4.
[9] I. Montilla, M. Puga, J. P. Luke, J. G. Marichal-Hernandez, and J. M. Rodriguez-Ramos, "Design and laboratory results of a plenoptic objective: From 2D to 3D with a standard camera," *J. of Display Technology*, vol. 11, no. 1, pp. 73–78, 2015.
[10] T. Ebrahimi, S. Foessel, F. Pereira, and P. Schelkens, "Jpeg pleno: Toward an efficient representation of visual reality," *IEEE MultiMedia*, no. 4, pp. 14–20, 2016.
[11] Z. Wang, "Objective image quality assessment: Facing the real-world challenges," in *Image Quality and System Performance (keynote speech paper)*, 2016, pp. 1–6.

[12] K. Gu, S. Wang, G. Zhai, W. Lin, X. Yang, and W. Zhang, "Analysis of distortion distribution for pooling in image quality prediction," *IEEE Trans. on Broadcasting*, vol. 62, no. 2, pp. 446–456, 2016.

[13] P. Paudyal, F. Battisti, and M. Carli, "Impact of video content and transmission impairments on quality of experience," *Multimedia Tools and Applications*, pp. 1–25, 2016.

[14] V. Vaish and A. Adams, "The (New) Stanford Light Field Archive," 2008. [Online]. Available: http://lightfield.stanford.edu/

[15] "Synthetic Light Field Archive," 2016. [Online]. Available: http://web.media.mit.edu/~gordonw/SyntheticLightFields/

[16] S. Wanner, S. Meister, and B. Goldluecke, "Datasets and benchmarks for densely sampled 4D light fields," in *Annual Workshop on Vision, Modeling and Visualization: VMV*, 2013, pp. 225–226.

[17] A. Ghasemi, N. Afonso, and M. Vetterli, "LCAV-31: a dataset for light field object recognition," in *IS&T/SPIE Electronic Imaging*, 2014, pp. 902 014–902 014.

[18] A. Mousnier, E. Vural, and C. Guillemot, "Partial light field tomographic reconstruction from a fixed-camera focal stack," *arXiv preprint arXiv:1503.01903*, 2015.

[19] N. Li, J. Ye, Y. Ji, H. Ling, and J. Yu, "Saliency detection on light field," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 2806–2813.

[20] R. Raghavendra, K. Raja, and C. Busch, "Exploring the usefulness of light field camera for biometrics: An empirical study on face and iris recognition," *IEEE Trans. on Information Forensics and Security (TIFS)*, vol. 11, no. 5, pp. 922 – 936, 2016.

[21] M. Rerabek, L. Yuan, L. A. Authier, and T. Ebrahimi, "[ISO/IEC JTC 1/SC 29/WG1 contribution] EPFL Light-Field Image Dataset," ISO/IEC JTC 1/SC 29/WG1, Tech. Rep., 2015.

[22] P. Paudyal, R. Olsson, M. Sjöström, F. Battisti, and M. Carli, "SMART: a light field image quality dataset," in *ACM Multimedia Systems Conference*, 2016, pp. 1–6.

[23] B. W. Keelan and H. Urabe, "ISO 20462: a psychophysical image quality measurement standard," in *Electronic Imaging 2004*, 2003, pp. 181–189.

[24] ITU-T Rec. P.910, "Subjective video quality assessment methods for multimedia applications," 2008.

[25] ITU-R Rec. BT.500, "Methodology for the subjective assessment of the quality of television pictures," 2012.

[26] X. Dong, D. Qionghan, and X. Wenli, "Data compression of light field using wavelet packet," in *IEEE Int. Conf. on Multimedia and Expo (ICME)*, vol. 2, 2004, pp. 1071–1074.

[27] C. Perra, "Lossless plenoptic image compression using adaptive block differential prediction," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2015, pp. 1231–1234.

[28] F. Dai, J. Zhang, Y. Ma, and Y. Zhang, "Lenselet image compression scheme based on subaperture images streaming," in *IEEE Int. Conf. on Image Processing (ICIP)*, 2015, pp. 4733–4737.

[29] Y. Li, M. Sjöström, R. Olsson, and U. Jennehag, "Coding of focused plenoptic contents by displacement intra prediction," *IEEE Trans. on Circuits and Systems for Video Technology*, p. 1, 2015.

[30] ——, "Scalable coding of plenoptic images by using a sparse set and disparities," *IEEE Trans. on Image Processing*, vol. 25, no. 1, pp. 80–91, 2016.

[31] N. Dimitrova, H.-J. Zhang, B. Shahraray, I. Sezan, T. Huang, and A. Zakhor, "Applications of video-content analysis and retrieval," *IEEE multimedia*, no. 3, pp. 42–55, 2002.

[32] S. Winkler, "Analysis of public image and video databases for quality assessment," *IEEE J. of Selected Topics in Signal Processing*, vol. 6, no. 6, pp. 616–625, 2012.

[33] O. Hadar, R. Shmueli, R. Huber, and M. Huber, "Effects of compression parameters on the perceived quality of video stream over a lossy internet protocol network," *Optical Engineering*, vol. 45, no. 8, pp. 087 003–087 003, 2006.

[34] R. Shmueli, O. Hadar, R. Huber, M. Maltz, and M. Huber, "Effects of an encoding scheme on perceived video quality transmitted over lossy internet protocol networks," *IEEE Trans. on Broadcasting*, vol. 54, no. 3, pp. 628–640, 2008.

[35] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Hybrid no-reference quality metric for singly and multiply distorted images," *IEEE Trans. on Broadcasting*, vol. 60, no. 3, pp. 555–567, 2014.

[36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[37] F. Shao, W. Lin, S. Wang, G. Jiang, and M. Yu, "Blind image quality assessment for stereoscopic images using binocular guided quality lookup and visual codebook," *IEEE Trans. on Broadcasting*, vol. 61, no. 2, pp. 154–165, 2015.

[38] T. Zhao, J. Wang, Z. Wang, and C. W. Chen, "SSIM-based coarse-grain scalable video coding," *IEEE Trans. on Broadcasting*, vol. 61, no. 2, pp. 210–221, 2015.

[39] P. Paudyal, F. Battisti, and M. Carli, "Evaluation of the effects of transmission impairments on perceived video quality by exploiting retrieved dataset," *Journal of Electronic Imaging*, vol. 26, no. 2, pp. 023 003–023 003, 2017.

[40] N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti *et al.*, "Image database TID2013: Peculiarities, results and perspectives," *Signal Processing: Image Communication*, vol. 30, pp. 57–77, 2015.

[41] L. Liu, B. Liu, H. Huang, and A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," *Signal Processing: Image Communication*, vol. 29, no. 8, pp. 856–863, 2014.

[42] M. Rerabek, T. Bruylants, T. Ebrahimi, F. Pereira, and P. Schelkens, "Grand challenge: Light-field image compression," Tech. Rep., 2016.

[43] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz, "Efficient and effective querying by image content," *J. of intelligent information systems*, vol. 3, no. 3-4, pp. 231–262, 1994.

[44] D. Hasler and S. E. Suesstrunk, "Measuring colorfulness in natural images," in *Electronic Imaging*, 2003, pp. 87–95.

[45] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *J. of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.

[46] T. E. Bishop, S. Zanetti, and P. Favaro, "Light field superresolution," in *IEEE Int. Conf. on Computational Photography (ICCP)*, 2009, pp. 1–9.

[47] R. Szeliski, S. Avidan, and P. Anandan, "Layer extraction from multiple images containing reflections and transparency," in *IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, 2000, pp. 246–253.

[48] Q. Wang, H. Lin, Y. Ma, S. B. Kang, and J. Yu, "Automatic layer separation using light field imaging," *arXiv preprint arXiv:1506.04721*, 2015.

[49] A. Levin and Y. Weiss, "User assisted separation of reflections from a single image using a sparsity prior," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 9, p. 1647, 2007.

[50] ITU-T Tutorial, "Objective perceptual assessment of video quality: full reference television," *ITU-T Telecommunication Standardization Bureau*, 2004.

[51] F. Bossen, "Common test conditions and software reference configurations," *Joint Collaborative Team on Video Coding (JCT-VC), JCTVC-F900*, 2011.

[52] "High Efficiency Video Coding (HEVC): HM Software," 2016. [Online]. Available: https://hevc.hhi.fraunhofer.de/,

[53] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video," *IEEE Trans. on Image Processing*, vol. 19, no. 6, pp. 1427–1441, 2010.

[54] D. Grois, D. Marpe, A. Mulayoff, B. Itzhaky, and O. Hadar, "Performance comparison of H. 265/MPEG-HEVC, VP9, and H. 264/MPEG-AVC encoders," in *Picture Coding Symposium (PCS), 2013*, 2013, pp. 394–397.

[55] G. D. Donald, "Light Field Toolbox for Matlab (v 0.4)," pp. 1–26, 2016. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/49683-light-field-toolbox-v0-4

[56] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Linear volumetric focus for light field cameras," *ACM Trans. on Graphics (TOG)*, vol. 34, no. 2, pp. 1–15, 2015.

[57] D. Liu, L. Wang, L. Li, Z. Xiong, F. Wu, and W. Zeng, "Pseudo-sequence-based light field image compression," in *IEEE Int. Conf. on Multimedia & Expo Workshops (ICMEW)*, 2016, pp. 1–4.

[58] R. Monteiro, L. Lucas, C. Conti, P. Nunes, N. Rodrigues, S. Faria, C. Pagliari, E. da Silva, and L. Soares, "Light field HEVC-based image coding using locally linear embedding and self-similarity compensated prediction," in *IEEE Int. Conf. on Multimedia & Expo Workshops (ICMEW)*, 2016, pp. 1–4.

[59] H. Common, "Test conditions and software reference configurations, JCTVC-L1100," ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Tech. Rep., 2013.

[60] A. Jones, I. McDowall, H. Yamada, M. Bolas, and P. Debevec, "Rendering for an interactive 360 light field display," in *ACM Trans. on Graphics (TOG)*, vol. 26, no. 3, 2007, p. 40.

[61] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2013, pp. 1027–1034.

[62] Y. J. Jeong, J.-H. Lee, Y. H. Cho, D. Nam, D.-S. Park, and C.-C. J. Kuo, "Efficient light-field rendering using depth maps for 100-mpixel

multi-projection 3d display," *IEEE J. of Display Technology*, vol. 11, no. 10, pp. 792–799, 2015.
[63] T. Georgiev and A. Lumsdaine, "Focused plenoptic camera and rendering," *J. of Electronic Imaging*, vol. 19, no. 2, pp. 021 106–021 106, 2010.
[64] K. Tsukida and M. R. Gupta, "How to analyze paired comparison data," Tech. Rep., 2011.
[65] J. Korhonen, N. Burini, J. You, and E. Nadernejad, "How to evaluate objective video quality metrics reliably," in *Fourth Int. Workshop on Quality of Multimedia Experience (QoMEX)*, 2012, pp. 57–62.
[66] L. Krasula, K. Fliegel, P. Le Callet, and M. Klíma, "Using full-reference image quality metrics for automatic image sharpening," in *SPIE Photonics Europe*, 2014.
[67] R. A. Bradley and M. E. Terry, "Rank analysis of incomplete block designs: I. the method of paired comparisons," *Biometrika*, vol. 39, no. 3/4, pp. 324–345, 1952.
[68] J. C. Handley, "Comparative analysis of bradley-terry and thurstone-mosteller paired comparison models for image quality assessment," in *PICS*, 2001, pp. 108–112.
[69] F. Wickelmaier and C. Schmid, "A matlab function to estimate choice model parameters from paired-comparison data," *Behavior Research Methods, Instruments, & Computers*, vol. 36, no. 1, pp. 29–40, 2004.
[70] L. Xu, J. Li, W. Lin, Y. Zhang, Y. Zhang, and Y. Yan, "Pairwise comparison and rank learning for image quality assessment," *Displays*, vol. 44, pp. 21–26, 2016.
[71] R. K. Mantiuk, A. Tomaszewska, and R. Mantiuk, "Comparison of four subjective methods for image quality assessment," in *Computer Graphics Forum*, 2012, pp. 2478–2491.
[72] M. Piegel, *Schaum's Outline of Theory and Problems of Statistics*. Schaum Publishing, 1961.
[73] R. McGill, J. W. Tukey, and W. A. Larsen, "Variations of box plots," *The American Statistician*, vol. 32, no. 1, pp. 12–16, 1978.
[74] G. Barnard, "A new test for 2× 2 tables," *Nature*, vol. 156, 1945.
[75] J. Li, M. Barkowsky, and P. Le Callet, "Subjective assessment methodology for preference of experience in 3DTV," in *IEEE 11th IVMSP Workshop*, 2013, pp. 1–4.
[76] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.
[77] P. Hanhart, E. Bosc, P. Le Callet, and T. Ebrahimi, "Free-viewpoint video sequences: a new challenge for objective quality metrics," in *IEEE Int. Workshop on Multimedia Signal Processing (MMSP)*, 2014, pp. 1–6.
[78] S. Lê, J. Josse, F. Husson *et al.*, "FactoMineR: an R package for multivariate analysis," *J. of statistical software*, vol. 25, no. 1, pp. 1–18, 2008.
[79] P. Le Callet, S. Möller, A. Perkis *et al.*, "QUALINET white paper on definitions of quality of experience," *European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)*, 2012.
[80] M. Varela, L. Skorin-Kapov, K. De Moor, P. Reichl, C. W. Chen, P. Chatzimisios, T. Dagiuklas, and L. Atzori, "QoE- defining a user-centric concept for service quality," *Multimedia Quality of Experience (QoE): Current Status and Future Requirements*, pp. 1–5, 2015.
[81] A. Norkin, G. Bjontegaard, A. Fuldseth, M. Narroschke, M. Ikeda, K. Andersson, M. Zhou, and G. Van der Auwera, "Hevc deblocking filter," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1746–1754, 2012.
[82] S. Saha, M. Tahtali, A. Lambert, and M. Pickering, "Perceptual dissimilarity: A measure to quantify the degradation of medical images," in *Int.l Conf. on Digital Image Computing Techniques and Applications (DICTA)*, 2012, pp. 1–6.
[83] ——, "Perceptual dissimilarity metric: A full reference objective image quality measure to quantify the degradation of perceptual image quality," in *IEEE Int. Symposium on Signal Processing and Information Technology*, 2013, pp. 327–332.
[84] M. Gaubatz and S. Hemami, "MeTriX MuX visual quality assessment package," 2016. [Online]. Available: http://foulard,ece.cornell.Edu/gaubatz/metrix_mux
[85] T. Mitsa and K. L. Varkur, "Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, 1993, pp. 301–304.
[86] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli, "New full-reference quality metrics based on HVS," in *Second Int. Workshop on Video Processing and Quality Metrics*, vol. 4, 2006.
[87] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin, "On between-coefficient contrast masking of DCT basis functions," in *Third Int. Workshop on Video Processing and Quality Metrics*, vol. 4, 2007.
[88] D. M. Chandler and S. S. Hemami, "VSNR: a wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. on Image Processing*, vol. 16, no. 9, pp. 2284–2298, 2007.
[89] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *ASILOMAR Conf. on Signals, Systems and Computers*, 2003, pp. 1398–1402.
[90] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. on Image Processing*, vol. 15, no. 2, pp. 430–444, 2006.
[91] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, 2002.
[92] P. Y. Chen and P. M. Popovich, *Correlation: Parametric and nonparametric measures*. Sage, 2002, no. 137-139.

Pradip Paudyal received his M.Sc. in Information and Communication Engineering from Tribhuvan University, Kathmandu, Nepal in 2010. Since 2011, he is working as Assistant Director at Nepal Telecommunications Authority (NTA). Since January 2014, he is PhD student at Roma TRE University, Rome Italy. His current research fields include multimedia communication and quality assessment.

Federica Battisti received her PhD from Universit degli Studi Roma TRE, Italy, in 2010. She is currently a non-tenured assistant professor at the same university. Since November 2006, she has been with the Telecommunication Group of the Department of Engineering at the Universit degli Studi Roma TRE. Her current research interests include 2-D and 3-D image/video quality, image processing, and watermarking. She is an IEEE senior member.

Mårten Sjöström received the M.Sc. degree in Electrical Engineering and Applied Physics from Linköping University, Sweden, in 1992, the Licentiate of Technology degree in Signal Processing from KTH, Stockholm, Sweden, in 1998, and the Ph.D. degree in Modeling of nonlinear systems from EPFL, Lausanne, Switzerland, in 2001. He worked as an Electrical Engineer at ABB, Sweden, from 1993-1994, was a fellow at CERN from 1994-1996, and a Ph.D.-student at EPFL, Lausanne, Switzerland during 1997-2001. He joined Mid Sweden University in 2001, and was appointed Associate Professor and Full Professor in Signal Processing in 2008 and 2013, respectively. He is the head of the subject Computer and System Sciences at Mid Sweden University since 2013. He founded the Realistic 3D research group in 2007. His current research interests are within multidimensional signal processing and imaging, as well as system modelling and identification.

Roger Olsson received the M.Sc. degree in electrical engineering and the Ph.D. degree in telecommunication from Mid Sweden University, Sweden, in 1998 and 2010, respectively. He was with the video compression and distribution industry from 1997 to 2000. He was a Junior Lecturer with Mid Sweden University from 2000 to 2004, where he taught courses in telecommunication, signals and systems, and signal and image processing. Since 2010, he has been a Researcher with Mid Sweden University. His research interest includes plenoptic image capture, processing, and compression; plenoptic system modeling; and depth map capture and processing

Marco Carli is Assistant Professor with the Department of Engineering at the Universit degli Studi 'Roma TRE', Roma, Italy. He received the Laurea degree in Telecommunication Engineering from the Universit degli Studi di Roma 'La Sapienza', Roma, Italy and the Ph.D. degree from Tampere University of Technology, Tampere, Finland. He was a Visiting Researcher with the Image Processing Laboratory, directed by prof. S.Mitra, UCSB, University of California, Santa Barbara, California, USA (2000-2004). His research interests are in the area of digital signal and image processing with applications to multimedia communications. Specifically, he has been working on digital watermarking, multimedia quality evaluation, information security. He is an Associate Editor of EURASIP Journal on Image and Video Processing (2011 - present) and Area Editor of Elsevier Signal Processing: Image Communication