ICME 2016 Grand Challenge: Light-Field Image Compression

July 11th – 15th, 2016, Seattle, USA

Call for Proposals and Evaluation Procedure

Authors: Martin Rerabek, Tim Bruylants, Touradj Ebrahimi, Fernando Pereira, Peter Schelkens

Abstract

The ICME Grand Challenge on Light-Field Image Compression calls for efficient image compression technologies, visual quality assessment methodologies and test materials for light-field (LF) images. This document provides detailed information about this call and evaluation procedure that will be used to assess the proposed image compression algorithms.

1 Introduction

LF images capture information about the intensity of light in a scene and also information about the direction of the light rays in space. They are typically acquired with a plenoptic camera enhanced by an array of micro-lenses placed in front of an otherwise conventional image sensor. In recent years, plenoptic cameras have become commercially available. However, an efficient compression format for this new modality does not exist yet. This challenge solicits contributions that demonstrate the efficient compression of LF images.

This document describes the test dataset and quality assessment methodology to assess submitted proposals. However, other evaluation methodologies and contents, as supplemental information to submitted compression technologies are welcome, and could be considered for future editions of this challenge. Therefore, evaluation methodologies and metrics as well as other publicly available LF images along with additional evidence to assess performance, including those quantifying attractive features beyond compression efficiency, can also be included in submissions of compression technologies to this call.

2 Purpose and Procedure

The main purpose of this challenge is to collect and evaluate **new light-field image compression solutions** as alternatives to existing JPEG standards. Evaluation of the proposed compression algorithms and technologies will be performed by means of conventional objective metrics as well as by formal subjective quality evaluations. The evaluation results will be presented at ICME 2016 in a special session targeted to this Grand Challenge. The proponents will be disclosed at this event. However, proponents can request to be listed anonymously. Proponents are requested to also submit a paper describing their contribution as outlined in Section 4. Attendance and registration for ICME 2016 is required.

A secondary purpose is to encourage propositions of **new methodologies to evaluate the compression of light-field images**. New assessment techniques should address the specific features of LF imaging, such as interaction with compressed media (change of focus and point of view).This effort mainly aims at gathering, enhancing, and summarizing the information about evaluation of LF images used in this call. Thus, no formal assessment of the proposed evaluation methodologies will be carried out within this challenge, but these methodologies may be used in future editions of the challenge.

Furthermore, this challenge also encourages the establishment of **additional publicly available datasets of light-field images** that can also be used in future editions of the challenge.

3 Assessment Methodology

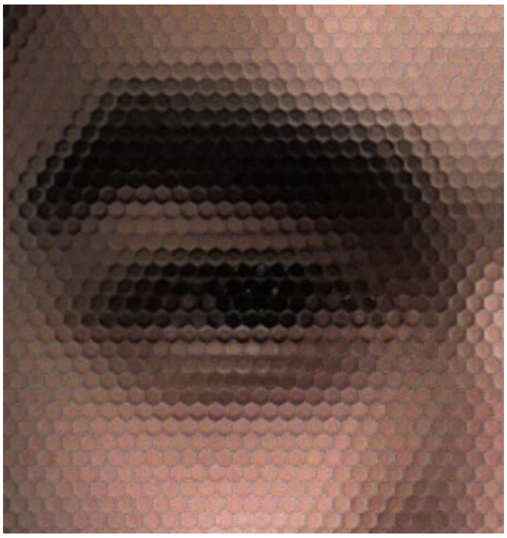
This challenge focuses on the compression of images acquired by plenoptic cameras. More specifically, it assumes compression of lenslet[1]-based LF images produced by a Lytro Illum light-field camera. The following chapter describes in more details the processing flow used to obtain the compressed and uncompressed LF images from Lytro Illum light-field camera raw sensor data. Furthermore, it presents the methodology that will be used to evaluate the compression technologies submitted to this call.

3.1 Processing Flow

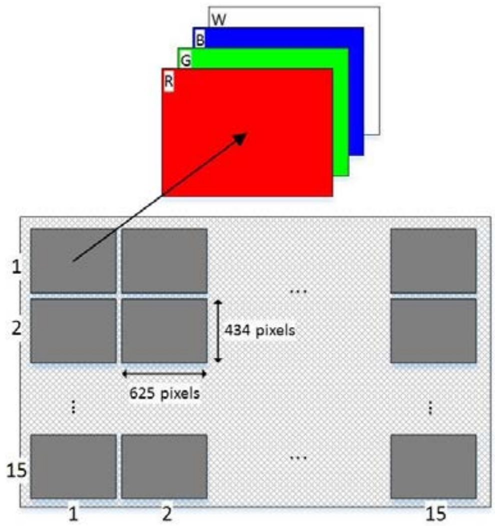
Many methods exist to convert plenoptic camera raw sensor data to what is usually called a LF image. This challenge uses a simple and practical approach by relying on the Matlab implementation of the Light Field Toolbox v0.4[2]. This toolbox converts Lytro Illum light-field camera raw sensor data to LF images following the process described in [1, 2].

First the Lytro Illum raw sensor data is preprocessed by applying demosaicing and devignetting. Subsequently, the uncompressed lenslet image is converted into an LF data structure. An example of an uncompressed lenslet image and the LF data structure as provided by above mentioned toolbox is shown in Figure 1. LF data structure is actually a stack of 2D low-resolution RGB images in addition to a weighting image. The weighting image carries the confidence associated with each pixel, which can be useful in filtering applications that accept a weighting term [3]. The resulting dimension of the LF data structure is 15×15×434×625×4, 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.

This challenge focusses on the compression of uncompressed lenslet images in YCbCr 4:2:0 color space. Moreover, various constraints are imposed on the compression and decompression. These constraints are intended to focus the challenge on the core aspects of an uncompressed lenslet image compression and its decompression and to avoid cluttering the assessment by influencing factors, which are not due to the core compression technology. The considered end-to-end processing chain for compression of a lenslet-based LF image is explained in details in the following section.



a) Crop of an uncompressed lenslet image in RGB obtained after demosaicing and devigneting of raw sensor data.



b) LF data structure as produced by LF toolbox. The dimensions of LF data structure are *15 × 15 × 434 × 625 × 4*

Figure 1: Example of a lenslet image and its corresponding LF image data.

---

[1] A lenslet is literally a small lens. The fact that distinguishes it from a small lens is that it is part of a lenslet array. A lenslet array consists of a set of lenslets in the same plane.

[2] http://www.mathworks.com/matlabcentral/fileexchange/49683-light-field-toolbox-v0-4

3.1.1 End-to-end compression and decompression chain

The end-to-end coding-decoding chain of the lenslet-based LF image is provided in Figure 2. It is composed of the following steps:

1. The 10 bit precision raw sensor data is demosaiced and devignetted.
2. The 10 bit precision data is clipped to 8 bit by dropping the two least significant bits.
3. The RGB 4:4:4, 8 bit uncompressed lenslet image (point **B** in Figure 2) is converted to YCbCr4:2:0 color space (point **A** in Figure 2) to produce uncompressed lenslet image.
4. The image information at point **A** is compressed to produce a bitstream by means of a compression algorithm, which could be either an anchor or a proposed encoder.
5. The compressed bitstream is subsequently decompressed to obtain the uncompressed lenslet image in YCbCr4:2:0 (point **A'** in Figure 2) by means of a decompression algorithm, which could be either an anchor or a proposed decoder.
6. The uncompressed lenslet image in YCbCr 4:2:0 color space is converted to an uncompressed lenslet image in RGB 4:4:4(point **B'** in Figure 2).
7. The uncompressed lenslet image in RGB is then converted to an LF data structure, i.e. a stack of low-resolution 2D images, where each image represents a different viewpoint. This representation is suitable for rendering and viewing by the provided Matlab toolbox.

3.1.2 Reference LF data structure generation

The reference LF data structure is generated directly from the uncompressed lenslet image. The produced stack of 2D uncompressed views will serve as reference data for the full reference visual quality assessment metrics and subjective quality assessment methodology used in this challenge.

The processing flow to generate the reference LF data structure is illustrated in Figure 3. The uncompressed lenslet image at point **B** is obtained in a way similar to the previously described end-to-end compression and decompression chain (i.e. steps 1 to 3). Then, the color space upsampling and conversion (step 6) are performed to obtain the reference uncompressed lenslet image (point **B'$_{\mathbf{REF}}$**). Finally, the reference uncompressed lenslet image is converted to the reference LF data structure.

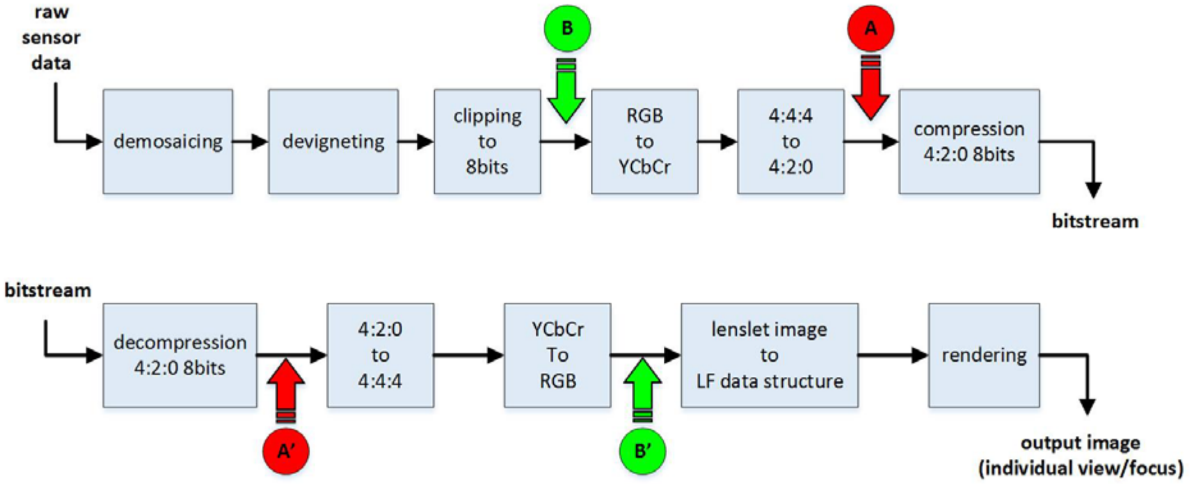


Figure 2: End-to-end processing chain showing compression and decompression of lenslet-based LF image.

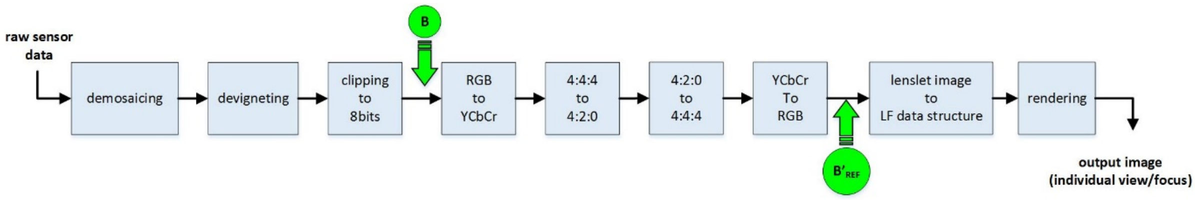


Figure 3: Processing chain for generation of reference LF data structure.

3.2 Objective Performance Evaluation

The compression distortion will be measured relative to the uncompressed image data at various bit rates corresponding to compression ratios defined in Table 1. More specifically, the objective quality metric will be computed between the uncompressed LF data structure (stack of 2D low-resolution RGB images) converted from point **B’** and the reference LF data structure converted from point **B’**$_{\mathbf{REF}}$. Both the objective quality of the individual views (2D low-resolution RGB images) and the global quality of the uncompressed LF data structure will be evaluated.

The evaluation of the proposed compression algorithms will be performed according to conventional objective metrics PSNR_Y and PSNR_YUV defined as follows:

$$PSNR_Y(k,l) = 10\log_{10}\frac{255^2}{MSE(k,l)}.$$

MSE(k,l) is defined for *k-th* and *l-th* viewpoint as:

$$MSE(k,l) = \frac{1}{m\,n}\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j) - R(i,j)]^2,$$

where *m* and *n* correspond to dimensions of one individual view (one 2D low-resolution image) in pixels, i.e., *m* = 434 pixels and *n*=625 pixels. *I(i,j)* and *R(i,j)* is the value of pixel on position *(i,j)* of corresponding channel (Y, U or V) of evaluated and reference individual view, respectively.

The PSNR_YUV value for each individual view is computed as:

$$PSNR_{YUV}(k,l) = \frac{6PSNR_Y(k,l)+PSNR_U(k,l)+PSNR_V(k,l)}{8}.$$

Average PSNR_Y_mean value is computed as a mean of PSNR_Y values for each individual view as:

$$PSNR_{Y_{mean}} = \frac{1}{(K-2)(L-2)}\sum_{k=2}^{K-1}\sum_{l=2}^{L-1} PSNR_Y\,(k,l),$$

where k and *l* correspond to *k-th* and *l-th* individual view, *K* and *L* correspond to amount of views in each horizontal and vertical direction, respectively. (*K=L=15*). Mean value for PSNR_YUV is then computed as:

$$PSNR_{YUV_{mean}} = \frac{1}{(K-2)(L-2)}\sum_{k=2}^{K-1}\sum_{l=2}^{L-1} PSNR_{YUV}\,(k,l).$$

3.3 Subjective Performance Evaluation

Formal viewing tests are performed to assess the subjective quality of the resulting LF images obtained from each proposal when compared to an anchor. The visual comparison between the individual images rendered from the LF data structure obtained from uncompressed lenslet images (point **B’)** and their corresponding images from the reference (point **B’**$_{\mathbf{REF}}$). At least 3 views and 3 focus points will be randomly selected for visual testing by the test organizers. The selected focus points will be rendered using the Matltab toolbox adopted in this challenge. Please note, that the rendered images will be corrected for colors using Matlab toolbox *(DecodeOptions.OptionalTasks = 'ColourCorrect')* and for gamma (*1/2.2*).

Subjective visual quality evaluations will be performed by means of a Double Stimuli Impairment Scale (DSIS) methodology [4], adapted to still image quality assessment. The test subject will evaluate the level of impairment between the reference and decoded images displayed simultaneously in a side-by-side configuration.

4 Test Material and Coding Conditions

This section defines the test material and coding conditions to be used by all proponents.

4.1 Anchors, Test Material, and Coding Conditions

The test material consists in a subset of publicly available light-field image dataset from EPFL[3].

The test images selected for evaluation are listed in Table 1.

The anchors are generated according to the procedure outlined in Section 2.1.1 using Legacy JPEG as the compression algorithm.

The test material can be downloaded from:

FTP SERVER:	**tremplin.epfl.ch**
USER:	**GrandChallenge@grebvm2.epfl.ch**
PASSWORD:	**LF2016**
Please use dedicated FTP clients, such as FileZilla to download the content!	

Table 1: Image test sequences and compression ratios

Image ID	Image name	Compression ratios			
		R1	R2	R3	R4
I01	Bikes	10:1	20:1	40:1	100:1
I02	Danger_de_Mort	10:1	20:1	40:1	100:1
I03	Flowers	10:1	20:1	40:1	100:1
I04	Stone_Pillars_Outside	10:1	20:1	40:1	100:1
I05	Vespa	10:1	20:1	40:1	100:1
I06	Ankylosaurus_&_Diplodocus_1	10:1	20:1	40:1	100:1
I07	Desktop	10:1	20:1	40:1	100:1
I08	Magnets_1	10:1	20:1	40:1	100:1
I09	Fountain_&_Vincent_2	10:1	20:1	40:1	100:1
I10	Friends_1	10:1	20:1	40:1	100:1
I11	Color_Chart_1	10:1	20:1	40:1	100:1
I12	ISO_Chart_12	10:1	20:1	40:1	100:1

The compression ratios are computed as ratios between the *theoretical size* of Lytro sensor raw data in 10 bit precision, i.e. 5368 × 7728 × 10b = 51854 880 Bytes, and the size of their corresponding compressed bitstream file.

[3] http://mmspg.epfl.ch/EPFL-light-field-image-dataset

4.2 Materials for Proponents

Proponents to this Challenge are provided with:

1. LFR images (*.LFR) as taken by Lytro camera including camera calibration data.
2. Uncompressed lenslet images (*.YUV) after color conversion and color subsampling (point **A** in Figure 2).
3. Matlab scripts to perform:
    a. Computation of PSNR values.
    b. Downsampling 4:4:4 to 4:2:0 and upsampling 4:2:0 to 4:4:4.
    c. Color conversion RGB to YCbCr and back YCbCr to RGB.
4. JPEG anchors
    a. Compressed JPEG bitstreams (in *.JPG file format) generated according to Figure 2.
    b. Corresponding LF data structure (in *.MAT file format) created from compressed JPEG bitstreams.
5. Reference images
    a. LF data structure (in *.MAT file format) created from uncompressed lenslet images.

5 Requirements on Proposal Submission

Proponents of light field compression algorithms should provide a detailed technical description in the form of a short paper as well as material to assess and validate the performance of their submission, e.g., compressed files and binary executable to reconstruct and measure performance. Any additional content or evaluation methodology should be submitted along with a compression algorithm. The list of materials from proponents for each submission category is detailed below. Detailed instructions for submission are provided to those proponents who will send a message to email addresses at the end of this document.

Submitted papers will be included as part of the ICME proceedings and published on IEEE Xplore if accepted after a peer review according to ICME rules.

5.1 Materials from Proponents

Proponents to this Challenge are required to submit:

1. Binary executables (for Windows) of the proposed coder and decoder along with a short description of how to run.
2. Compressed bitstreams at bit rates corresponding to the compression ratios indicated in Table 1.
3. Decoded bitstreams in YUV4:2:0 format (corresponding to point **A’** in Figure 2).
4. PSNR values (in *.mat file) for each compressed LF image (output of PSNR Matlab script).
5. Short paper including sufficiently detailed technical description of the proposed compression methodology and results. The format of the submitted papers is the same as regular papers submitted to ICME. Papers will go through the same peer review as any other ICME submissions and may be published as part of ICME 2016 proceedings.

Please note that all proponents are required to submit complete results for all test cases, this means results for test images and compression ratios. Incomplete submissions will be ineligible for consideration.

5.2 Additional Assessment Methodology Proposals

Proposals for new methodologies to evaluate compression of light field image data may demonstrate any or several of the following:

1. Objective quality measures that are well correlated with subjective assessment of image quality over a range of manipulations
2. Methods that assess the quality of interaction with the compressed media such as assessment of free-navigation, etc.
3. Methods to assess efficiency of features such as change of focal point by means of blur metrics on regions of interest, etc.

Furthermore, proposers of such additional evaluation methodologies should provide a very detailed description along with an executable or a source code (preferred option) of the methodology they submit.

Please note that none of the proposed LF image quality assessment methodology will be evaluated nor validated within this challenge. However, such proposals could be considered for use in future editions of the challenge.

5.3 Additional Content Proposal

Proposals for new LF image content are welcome along with proposals for LF image compression and evaluation methodologies. Such content should be made freely available for research and standardization purposes. Please note that none of the proposed LF image content will be evaluated nor validated within this challenge. However, such content could be considered for use in future editions of the challenge.

6 Challenge Outcome

A winner will be selected by a judging committee. Selection criteria will be based on rate distortion performance of submitted technologies measured by objective metrics explained above as well as subjective evaluations.

This challenge is aligned with the objectives of the new JPEG PLENO initiative[4].

7 Timeline

2016/02/05	Availability of detailed submission process, including JPEG anchors.
2016/04/03	Proposals submission deadline, including paper and other requested materials
2016/06/01	Notification of acceptance and feedback to proponents
2016/07/11-15	Grand Challenge session at ICME2016

8 Questions and requests

If you have any questions or requests, or need further clarifications, please contact any of the organizers below:

4 https://jpeg.org/items/20150320_pleno_summary.html

Ebrahimi Touradj touradj.ebrahimi@epfl.ch

Fernando Pereira: fp@lx.it.pt

Peter Schelkens: pschelke@etro.vub.ac.be

Martin Rerabek: martin.rerabek@epfl.ch

9 References

[1] D. Dansereau, O. Pizarro and S. Williams, “Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras,” in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013.

[2] D. G. Dansereau, O. Pizarro and S. B. Williams, “Linear Volumetric Focus for Light Field Cameras,” *ACM Trans. Graph.*, vol. 34, no. 2, pp. 15:1--15:20, 3 2015.

[3] D. G. Dansereau, “Light Field Toolbox for Matlab,” software manual, 2015.

[4] BT.500-13, ITU-R, “Methodology for the subjective assessment of the quality of television pictures,” 2012.