The IST-EURECOM Light Field Face Database

Alireza Sepas-Moghaddam¹, Valeria Chiesa², Paulo Lobato Correia¹, Fernando Pereira¹, Jean-Luc Dugelay²

¹Instituto de Telecomunicações, Instituto Superior Técnico – Universidade de Lisboa, Lisbon, Portugal
²EURECOM, Campus SophiaTech, Sophia Antipolis, France
alireza@lx.it.pt, chiesa@eurecom.fr, plc@lx.it.pt, fp@lx.it.pt, dugelay@eurecom.fr

Abstract— Light field cameras are emerging as powerful devices to capture rich scene representations that provide unique advantages for analysis and representation purposes. Some recent works have shown the power and usefulness of the richer information carried out by light field imaging, notably for face recognition. However, it is still difficult to fully assess how face recognition technology can benefit from these novel imaging sensors, notably due to the lack of appropriate test material. To support face recognition research exploiting light field images, the IST-EURECOM Light Field Face Database (IST-EURECOM LFFD) is presented in this paper. The purpose is to report the public availability of a light field face database which should be instrumental for designing, testing and validating light field imaging based recognition systems. The proposed face database includes data from 100 subjects, captured by a Lytro ILLUM camera in two 1-6 months separated sessions, with 20 samples per each person per session. To simulate multiple scenarios, the images are captured with several facial variations, covering a range of emotions, actions, poses, illuminations, and occlusions. The database includes the raw light field images, 2D rendered images and associated depth maps, along with a rich set of metadata. The IST-EURECOM LFFD is expected to become a valuable addition to existing face database repositories.

Keywords— Face Recognition; Light Field Imaging; Face Database.

I. INTRODUCTION

Nowadays, automatically recognizing the identity of a person is of paramount importance. Face recognition is one biometric modality that is widely used in personal devices, but also in legal and commercial applications, with high acceptability, collectability and universality [1], [2]. Advances in imaging technologies have been leading to a revolution in face recognition systems over the last decade. The recent availability of new imaging sensors opens a new range of possibilities also for image-based personal recognition systems, notably allowing a boost in the performance of many face recognition systems and applications [3]. Light field imaging technologies have recently come into prominence as they record not only the intensity of light on a specific 2D plane position but rather the intensity of the light rays for any direction in space [4], [5]. Preliminary works have shown the effectiveness of the supplementary information captured by light field cameras for face recognition applications [6], [7], even when considering one single shot and without using additional optics.

Recent works explored the possibility of creating multiple focus images, rendered from the same light field image acquisition, for instance using super-resolution and fusion schemes for the multi-face recognition problem [6], [7], [8], [9]. The results demonstrate the added value of light field imaging in terms of post-capture refocusing capability, and improved accuracy when compared with conventional images. In order to assess the light field imaging based face recognition methods, the so-called Light Field Face and Iris Database (LiFFID) [6] has been recently proposed. It includes a number of 2D images focused at different depths for each person, rendered from light field images acquired by the first generation Lytro camera. While the preliminary works processed multiple 2D images at different focus or depth, face recognition systems based on light field imaging can be further extended in other directions. More precisely, by processing the rich information associated to a light field in its native form, there is potential to improve the performance of current face recognition systems. However, since the existing LiFFID database only includes 2D rendered images, it cannot be used as a benchmark database for testing and validating those recognition systems which more deeply exploit the whole light field not to speak about the fact that they are dependent on the specific used rendering solution. This context, highlights the immediate, strong need for publicly available imaging resources in this domain.

In this paper, a new database, the so-called IST-EURECOM Light Field Face Database (IST-EURECOM LFFD), is introduced to address the lack of publicly available light field face resources to be used as the basis for the design, testing and validation of novel light field imaging based face recognition systems. The proposed database can be used not only in the context of face recognition research but also for other research areas such as emotion recognition, gender classification, age estimation, ethnicity classification and face modeling. The proposed face database consists of 100 subjects, with images captured by a Lytro ILLUM light field camera [10], at the Multimedia Signal Processing-Lx (MSP-Lx) at Instituto de Telecomunicações, Instituto Superior Técnico, Lisbon, Portugal, and the Imaging Security Lab at EURECOM, at the SophiaTech Campus, Nice, France. The image acquisition process has been repeated in the two labs with a very similar setup. Two separate sessions were performed for each subject with a temporal separation between 1 and 6 months. The database includes 20 image shots per person in each of the two sessions, with several facial variations including emotions, actions, poses, illuminations, and occlusions. The IST-
EURECOM LFFD is the first database of its kind as it includes the raw light field images, sample 2D rendered images and the corresponding depth maps along with a rich collection of metadata, including the location of a set of facial landmarks.

The rest of the paper is organized as follows. Section II provides an overview of available light field databases and also available face databases. Section III presents the fundamentals and potential of light field imaging technology. Section IV describes the acquisition setup and provides some statistics of the database. Sections V and VI report the considered face variations and the structure of the database, respectively. Finally, Section VII lists the database access and usage conditions and section VIII presents some final remarks.

II. AVAILABLE LIGHT FIELD AND FACE DATABASES

Light field imaging is a relatively new topic and thus only a few databases have been made available. At the date of writing, the following databases containing the full light field data were identified, none of them focused on biometric traits. The MMSPG dataset, created at EPFL, contains 118 light field images captured by a Lytro ILLUM camera [11]. The dataset is organized into ten different categories, covering a wide range of potential uses. The IRISA Lytro First Generation Dataset is a collection of 30 images including indoor and outdoor scenarios, some of them taken with motion blur or long exposure time [12]. In [13], a light field database for material recognition composed by 1200 images acquired using a Lytro ILLUM camera is presented. The Light Field Saliency Dataset (LFSD) consists of 100 images taken with a first generation Lytro camera, targeting saliency detection [14].

Some of the light field image databases are created with the purpose of studying the acquisition process or developing 2D image rendering algorithms; for this reason, the size, resolution, content and even the provided metadata are very different. The Computer Graphics Laboratory at Stanford University has created a 22 images light field archive using a multi-camera acquisition methodology [15]. The Northwestern University database is composed of 30 images of varied scenes, captured with a first generation Lytro camera, with the main target to evaluate a dictionary learning based color demosaicing algorithm [16]. The Heidelberg Collaboratory for Image Processing (HCI) project database contains two categories of images: seven artificial light fields rendered with a Blender, some of them with segmentation information, and six real-world light field acquired with a gantry [17]. Finally, a synthetic light field database with 18 images, including transparencies, occlusions and reflections, is presented in [18].

Currently, there are over 100 publicly available face databases. Table I overviews the main characteristics of a set of selected prominent existing face databases (a more complete list can be found in [19]). This information is complemented in Table II, which tabulates the variations addressed in these databases, sorted according to their release date. For comparison, also the characteristics of the IST-EURECOM LFFD being proposed in this paper are included in the tables.

TABLE I. OVERVIEW OF A FEW PROMINENT FACE DATABASES WITH DIFFERENT CHARACTERISTICS

<table>
<thead>
<tr>
<th>Database Name</th>
<th>Year</th>
<th>No. of subjects</th>
<th>Image type</th>
<th>Image modality</th>
<th>Spatial Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>FERET [25]</td>
<td>2003</td>
<td>1199</td>
<td>Grayscale 2D</td>
<td>2D</td>
<td>256×384</td>
</tr>
<tr>
<td>MIT-CBCL [26]</td>
<td>2004</td>
<td>10</td>
<td>Color       2D</td>
<td>2D</td>
<td>768×576</td>
</tr>
<tr>
<td>FEI [27]</td>
<td>2006</td>
<td>200</td>
<td>Color       3D</td>
<td>2D</td>
<td>640×480</td>
</tr>
<tr>
<td>FRAV3D [28]</td>
<td>2007</td>
<td>106</td>
<td>Color       2D</td>
<td>3D</td>
<td>N/A</td>
</tr>
<tr>
<td>Bosphorus [29]</td>
<td>2008</td>
<td>105</td>
<td>Color       2D</td>
<td>3D</td>
<td>1600×1200</td>
</tr>
<tr>
<td>MOBIO [20]</td>
<td>2010</td>
<td>150</td>
<td>Color       2D</td>
<td>3D</td>
<td>Different resolutions up to 2048×1356</td>
</tr>
<tr>
<td>3DFRD [31]</td>
<td>2010</td>
<td>118</td>
<td>Color       2D</td>
<td>3D</td>
<td>751×501</td>
</tr>
<tr>
<td>SCface [21]</td>
<td>2011</td>
<td>130</td>
<td>Color/infrared 2D</td>
<td>3D</td>
<td>100×75, 144×108, 224×168, 426×320</td>
</tr>
<tr>
<td>BU-3DFE [22]</td>
<td>2013</td>
<td>100</td>
<td>Color       3D</td>
<td>3D</td>
<td>1040×1329</td>
</tr>
<tr>
<td>Kinect Face DB [23]</td>
<td>2014</td>
<td>52</td>
<td>Color       2D</td>
<td>3D</td>
<td>640×480</td>
</tr>
<tr>
<td>LiFFID [6]</td>
<td>2016</td>
<td>112</td>
<td>Grayscale 2D</td>
<td>2D</td>
<td>1054×1054</td>
</tr>
<tr>
<td>IST-EURECOM LFFD</td>
<td>2016</td>
<td>100</td>
<td>Color       4D</td>
<td>2D</td>
<td>1515×434×625, 2022×1404, 2022×1404</td>
</tr>
</tbody>
</table>

Among the selected databases, several consider the usage of sensors that had not been considered previously, thus motivating their creation. For instance, the MOBILE Biometrics (MOBIO) dataset [20] was recorded using two mobile devices, a mobile phone and a laptop computer, to boost the research on face recognition technology for mobile devices. The Surveillance Cameras face (SCface) database [21] was collected to provide visible and infrared spectrum images in an uncontrolled indoor environment. The Binghamton University 3D Facial Expression (BU-3DFE) [22] database was developed for analyzing facial expressions in dynamic 3D spaces. The Kinect Face database [23] provides RGB-D face images, captured by a Kinect sensor, to evaluate how face recognition technology can benefit from this imaging sensor. As the emergence of novel imaging sensors motivates the research community to work with associated new and richer imaging formats, gathering a powerful light field face database is becoming a pressing need.

The Light Field Face and Iris Database (LiFFID) [6] is the first and currently the only available face database where the importance of light field imaging sensors for facial recognition tasks has been acknowledged. The major novelties of the proposed IST-EURECOM LFFD over LiFFID are:
1. IST-EURECOM LFFD provides the light field raw imaging information (in LFR format).

2. IST-EURECOM LFFD provides a vast set of metadata, notably to validate automated facial recognition systems.

3. IST-EURECOM LFFD images are captured with the Lytro ILLUM light field camera [10], which is a major improvement regarding the first generation Lytro camera, e.g. in spatial resolution.

4. IST-EURECOM LFFD provides samples of 2D RGB rendered face images, instead of grayscale; this is an advantage as color features may play a positive role in face recognition systems [32], [33].

### TABLE II. OVERVIEW OF VARIATIONS IN A FEW PROMINENT FACE DATABASES

<table>
<thead>
<tr>
<th>Database Name</th>
<th>Facial occlusions</th>
<th>Facial expressions</th>
<th>Views/poses</th>
<th>Different dates</th>
<th>Illumination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yale [24]</td>
<td>N</td>
<td>N</td>
<td>9 poses per subject</td>
<td>N 64 levels</td>
<td></td>
</tr>
<tr>
<td>FERET [25]</td>
<td>Glasses; Hair</td>
<td>2 expressions per subject</td>
<td>10 poses per subject</td>
<td>1 year interval</td>
<td>Standard Low</td>
</tr>
<tr>
<td>MIT-CBCL [26]</td>
<td>N</td>
<td>N</td>
<td>Frontal; +30°; -30°</td>
<td>N Y</td>
<td></td>
</tr>
<tr>
<td>FEI [27]</td>
<td>N</td>
<td>Neutral; Smile</td>
<td>10 poses per subject</td>
<td>N Standard Low</td>
<td></td>
</tr>
<tr>
<td>FRAV3D [28]</td>
<td>N</td>
<td>Max 2 gestures per subject</td>
<td>N</td>
<td>N N</td>
<td></td>
</tr>
<tr>
<td>Bosphorus [29]</td>
<td>Glasses; Hand; Hair</td>
<td>35 expression per subject</td>
<td>14 poses per subject</td>
<td>N N</td>
<td></td>
</tr>
<tr>
<td>Multi-PIE [30]</td>
<td>N</td>
<td>Neutral; Surprise; Squint; Smile; Disgust; Scream</td>
<td>15 poses per subject</td>
<td>N 19 levels</td>
<td></td>
</tr>
<tr>
<td>MOBIO [20]</td>
<td>N</td>
<td>Different expressions (uncontrolled)</td>
<td>Uncontrolled</td>
<td>N Uncontrolled</td>
<td></td>
</tr>
<tr>
<td>3DFRD [31]</td>
<td>N</td>
<td>Smile; Talk; Mouth and/or Eyes opened and closed</td>
<td>N</td>
<td>N N N</td>
<td></td>
</tr>
<tr>
<td>SCface [21]</td>
<td>N</td>
<td>N</td>
<td>Uncontrolled</td>
<td>N Uncontrolled</td>
<td></td>
</tr>
<tr>
<td>BU-3DFE [22]</td>
<td>N</td>
<td>25 expressions per subject</td>
<td>Frontal; +45°; -45°</td>
<td>N N</td>
<td></td>
</tr>
<tr>
<td>Kinect Face DB [23]</td>
<td>Eye, Nose; Face occlusions</td>
<td>Neutral; Smile; Open mouth</td>
<td>Frontal; Right profile; Left profile</td>
<td>Up to 1 year interval</td>
<td>Y</td>
</tr>
<tr>
<td>LiFFID [6]</td>
<td>N</td>
<td>Neutral; Smile</td>
<td>6 poses per subject</td>
<td>N Uncontrolled</td>
<td></td>
</tr>
<tr>
<td>IST-EURECOM LFFD</td>
<td>Glasses; Sunglasses; Mask; Hat; Eye, Mouth occlusions</td>
<td>Neutral; Smile; Surprise; Angry</td>
<td>7 poses per subject</td>
<td>Up to 6 months interval</td>
<td>Standard Low; Low; High</td>
</tr>
</tbody>
</table>

### III. LENSET LIGHT FIELDS: BASICS AND POTENTIAL

The (static) light field representation can be interpreted as a 4D function that models the light in a region of the 3D space by recording the light rays at every point \((x,y)\), toward every possible direction \((u,v)\), i.e., \(P(x,y,u,v)\) [4]. To capture the light field, lenslet light field cameras typically place a micro-lens array on the focal plane of the main lens to split the incoming light based on the direction of the rays into the sensor area of the corresponding micro-lens [5]. Light field cameras can provide richer scene representations for processing and post-processing tasks, e.g., depth map computation, \textit{a posteriori} refocusing, and smooth motion parallax in the viewing range in multiple directions [34]. Examples of commercial light field cameras are the Raytrix [35] and Lytro [10] cameras. The use of light field sensors allows capturing additional data over conventional 2D sensors. For instance, previous works proved that face recognition performance improves when RGB images are extended with the corresponding depth information [23], [36].

Images acquired by a light field camera include information about depth as well as about different viewpoints, supporting functionalities such as \textit{a posteriori} focusing on a given depth plane, slightly changing the observation point, and enabling to perform face recognition with less cooperation from subject. A light field may also be used to render high-resolution or all-in-focus 2D images. For lenslet light field cameras also the error-prone and computationally expensive process of establishing multi-camera correspondences may be eliminated.

### IV. IST-EURECOM LIGHT FIELD FACE DATABASE: ACQUISITION SETUP AND STATISTICS

The proposed IST-EURECOM Light Field Face Database (IST-EURECOM-LFFD) was acquired in the context of a cooperation between the Multimedia Signal Processing-Lx laboratory at Instituto de Telecomunicaciones, Instituto Superior Técnico and the Imaging Security Lab at EURECOM, SophiaTech campus. Image acquisition was performed in an indoor environment, using a lenslet-based light field camera: Lytro ILLUM [10]. The Lytro ILLUM camera has a 40 Megapixel sensor and a 30-250 mm lens with 8.3x optical zoom and f/2.0 aperture. The acquisition setup, illustrated in Figure 1, included a white backdrop background behind a chair at a fixed distance of 1.25 m to the camera. The scene was illuminated with a three-point lighting kit, including a key light, a fill light and a back light, placed to limit shadows and allow ease segmentation of the subject from the background. The image acquisition process has been repeated in the two labs with the same predefined setup. Each volunteer participated in two separate acquisition sessions, with a time interval between 1 and 6 months. The database includes 20 shots per person in each session, with different facial variations including facial emotions, actions, poses, illuminations and occlusions. Before the acquisition process, volunteers were asked to fill and sign consent and metadata forms.

Figure 1: Acquisition setup at (a) IST; (b) and EURECOM.

The IST-EURECOM LFFD includes data from 100 volunteers, with 66 males and 34 females, with a total number of 4000 light field face images in the database, corresponding to a total disk space of about 270 GB. The participants were born between 1957 and 1998, and are from 19 different countries. Figure 2 illustrates the distribution of subjects by age.
V. IST-EURECOM-LFFD: FACE VARIATIONS

To fully benefit from the non-intrusive nature of face recognition, a face recognition system may be required to recognize a face in an arbitrary situation without the explicit cooperation of the subject. This flexibility is of great interest in many face recognition applications, notably many video surveillance environments.

The IST-EURECOM LFFD includes a total of 20 face variations per person, categorized into 6 dimensions:

1. **Neutral image** (one image): image captured with standard illumination, frontal pose, neutral emotion, no action, and no occlusion;
2. **Emotions** (3 images): images with three different emotions, notably happy, angry and surprise;
3. **Actions** (2 images): images with two different actions, notably closed eyes and open mouth;
4. **Poses** (6 images): images with different poses, notably looking up, looking down, right half-profile, right profile, left half-profile, left profile;
5. **Illumination** (2 images): images with different illumination intensities, notably low and high illumination levels;
6. **Occlusions** (6 images): images with occlusions, notably eye occluded by hand, mouth occluded by hand, with glasses, with sunglasses, with surgical mask and with hat.

Examples of the various face variations considered in the IST-EURECOM LFFD are illustrated in Figure 3. All images were taken under controlled conditions, but there were no restrictions imposed on clothing, make-up and hair style.

VI. IST-EURECOM-LFFD: ELEMENTS AND STRUCTURE

The IST-EURECOM LFFD is the first biometric database to include raw light field imaging files. It also includes additional information that can be useful for developing and testing face recognition systems. The database is composed by the following elements:

1. **Raw Light Field Images**: Raw light field image stored in the Lytro Illum native file format, so-called Light Field Raw (LFR) files, with a size of about 50 MB/image.
2. **2D Rendered Images**: Central view 2D rendered image for each light field image; the rendering has been performed using the Lytro Desktop Software [37].
3. **Depth Map**: Depth map for each 2D rendered image, also generated with the Lytro Desktop Software [37].
4. **Landmark Information**: Information about the location of several facial regions, notably face, left eye, right eye, nose and mouth bounding boxes, for the selected variations.
5. **Metadata Information**: Supplementary data used to describe the subject, e.g. age, sex, nationality, and imaging conditions at the time of acquisition.
6. **Calibration Information**: Information about the Lytro ILLUM sensor manufacturing properties which may be used for calibration purposes.

A. Light Field Images

Light field images are the most important component of the database; they are stored in the Lytro Illum native format, using the so-called Light Field Raw (LFR) files. LFR files can be used as initial input for both the Lytro camera software i.e., Lytro Desktop Software [37], or to any other processing library/toolbox, such as the Matlab Light Field Toolbox V0.4 [38]. Face recognition systems working on light field images will typically require to perform some conversion of the LFR files in order to further process them. As an example, the Matlab Light Field Toolbox supports converting the LFR raw data into a multi-view sub-aperture array, with dimension $U \times V \times X \times Y \times 4$, where $U \times V$ represents the number of views, $X \times Y$ represents the resolution of each sub-aperture image, and 4 corresponds to the number of components, notably R, G, B and a confidence level matrix associated to each pixel.

B. 2D Rendered Images

Since light field images are not directly viewable in conventional 2D displays, the proposed database also includes 2D rendered images for the central view of each image variation, generated using the Lytro Desktop Software. It is worth noting that this software automatically performs a number of processing steps, including up-sampling and color correction, to enhance the quality of the output images. As the raw light field images are made available, any other rendering solutions may also be used. The 2D rendered face images can
be viewed using conventional 2D displays or be further processed.

C. Depth Maps

A depth map can be used to bridge the gap between 2D and 3D face recognition. Depth maps (see example in Figure 4.a) can provide geometric information about the position and shape of objects, to be explored by recognition systems. The supplied depth maps are generated using the Lytro Desktop Software.

D. Landmark Information

Facial landmarks are relevant for facial region extraction and normalization in face recognition systems. In the IST-EURECOM LFFD, the facial landmarks information includes the location of the face, left eye, right eye, nose and mouth bounding boxes, as illustrated in Figure 4-b. To obtain these data, the detection solution proposed in [39] has been used to identify the selected facial landmarks. The landmark information is extracted for the central view 2D rendered images.

E. Metadata Information

Metadata information can be used for the evaluation face recognition, facial expression recognition, gender classification, and age estimation automated results. The IST-EURECOM LFFD rich metadata includes the image acquisition date, as well as the subject gender, age, facial hair, makeup, haircut and usage of accessories; the range of values for each of these metadata fields is listed in Figure 5.

F. Calibration Information

Calibration data is essential to compensate for the specific properties of each camera’s sensor. For example, it is a required input for some light field image processing software products, such as the Lytro Desktop Software [37] and the Matlab Light Field Toolbox [38].

G. Database File Structure

The files composing the database are organized according to a hierarchical structure, as illustrated in Figure 6. The root level of the hierarchy includes the metadata information and facial landmarks for all the subjects and the camera calibration files. The root level also includes a folder for each of the N subjects in the database, named using a 3 digit identifier, xxx. Each of these folders contains 3 sub-folders: “LFR files”, “2D rendered images” and “Depth map images”.

H. Naming Convention

The naming convention for the database light field images is type_xxx_s_yy_variation where:

- “type” refers to the type of image, notably “LF” (Light Field), “2D” (2 Dimensional) and “DM” (Depth Map);
- “xxx” is a three digit integer uniquely identifying the subject, starting from 001; the first 50 subjects have been recorded at IST and the second 50 subjects at EURECOM;
- “s” is a digit indicating the acquisition session number, notably “1” or “2”;
- “yy” is a two digit integer indicating the variation number, ranging from 01 to 20, corresponding to different variations illustrated in Figure 6.
- “variation” is a three letter acronym identifying the variation in a format more suitable for human reading, e.g., EHF (Emotion Happy Face) for the face image with happy emotion.

VII. IST-EURECOM-LFFD: ACCESS AND USAGE CONDITIONS

The database proposed in this paper will be freely distributed for standardization and academic research purposes. The first part of the database, captured at Instituto de Telecomunicações – Instituto Superior Técnico, Lisbon, Portugal can be accessed at http://www.img.lx.it.pt/LFFD/. The second part, captured at EURECOM, SophiaTech Campus, Nice, France can be accessed at http://lffd.eurecom.fr/.

VIII. FINAL REMARKS

This paper presents the novel IST-EURECOM Light Field Face Database. The need for this type of database is highlighted and its composition is discussed. In particular, for the first time a public biometric database will include raw light field images, along with sample 2D rendered images, depth maps and extensive metadata. The included face variations are discussed...
and the database organization and naming convention is presented.

The IST- EURERCOM LFFD is proposed to fill the gap between traditional face recognition and the emerging light field face recognition technologies. In particular, the proposed database can be used to develop and test face recognition methods exploiting all the information available in a full light field image. The proposed database can also be used for other research domains, such as emotion recognition, gender classification, age estimation, ethnicity classification and face modeling. As future work, the authors plan to create a database extension including images captured in unconstrained environments and capturing groups of people at different depth planes.

ACKNOWLEDGMENTS
This work has been partially supported by Instituto de Telecomunicações under Fundação para a Ciência e Tecnologia, Grant UID/EEA/50008/2013, and also partially supported by the ITEA3 IDEA4SWIFT project and from the European Union’s Horizon 2020 research and innovation program under grant agreement No 700259.

REFERENCES