Lenslet Light Field Panorama Creation: a Sub-Aperture Image Stitching Approach

André Oliveira, Catarina Brites, João Ascenso, Fernando Pereira
Instituto Superior Técnico, Universidade de Lisboa - Instituto de Telecomunicações
Av. Rovisco Pais, 1049-001, Lisboa, Portugal

***Abstract*— Typically, a single light field camera only captures a limited portion of the visual scene and the rendered views offer low spatial resolution since the full sensor resolution also needs to capture the light arriving to the sensor from multiple directions. However, many applications require a large field of view and higher spatial resolution, which cannot be offered by a single lenslet light field image. In this paper, a lenslet light field panorama creation solution is proposed based on the stitching of the sub-aperture images associated to several lenslet light fields. The proposed approach consists in capturing multiple light fields with a single lenslet camera, which is rotated to capture different scene angles; corresponding sub-aperture images associated with the multiple light field images are then stitched while avoiding misalignments. The created lenslet light field panorama should preserve the directional light information, thus allowing to change *a posteriori* the perspective and the objects in focus on a panorama basis.**

Keywords*— lenslet light field; panorama creation; multi-perspective; stiching

I. INTRODUCTION

With the desire to capture, in a single image, a wide field of view (FOV), panoramic photography has emerged as a way to combine a set of partially overlapping elementary images of a visual scene acquired from different viewpoints. Panoramic images bring more intense and immersive experiences to users since they support a free and intuitive navigation in the 3D visual world. Recently, this type of content has been increasingly used in many application domains as it stimulates user interaction, e.g. in virtual reality, surveillance and photography. Panorama creation is usually performed using a stitching procedure that combines multiple images of the same scene with overlapping fields of view.

Sensors in conventional cameras merely capture the total light intensity landing on each position and thus directional information about the light rays is lost. This is clearly a limited representation of the real scene, which corresponds to a fixed point of view and focus. Recently, new sensors have emerged with the capability to capture higher dimensional visual representations; for example, using a micro-lens (i.e. lenslet) array in the optical path, it is possible to capture the light for each spatial position (x, y) and coming from any angular direction (θ, Φ). This richer imaging representation model is known as *lenslet light field* and offers additional imaging functionalities such as refocusing on any part of the image after the capture, slightly changing the point of view, relighting and recoloring, and selecting objects automatically based on their depth [1]. With these new sensors, images became volumes, changing the conventional imaging representation model based on (2D) flat planes. However, in practice, since the lenslet light fields are acquired with a single camera (and not an array of cameras), they tend to have a small field of view and spatial resolution, which limits the immersion and navigation functionalities that can be offered to the users. For the sake of brevity, lenslet light fields will be simply referred to as light fields from now on.

Naturally, the creation of panoramas using light field images and not conventional images is an exciting path to pursue considering the potential additional functionalities and the pressing need to offer the users more intense and immersive experiences. In this context, the main objective of this paper is to propose a lenslet light field panorama creation solution able to exploit the potential of light field cameras for panorama production and consumption. With the proposed solution, a high-resolution panorama is created that allows *a posteriori* refocus as well as perspective changes. This representation is suitable for many applications, from stereoscopic omnidirectional imaging to augmented reality with suitable focus cues.

The remaining of this paper is organized as follows: First, a brief review of some representative light field panorama creation solutions in the literature is made in Section II. After, Section III proposes the novel light field panorama creation solution whose performance will be presented and discussed in Section IV through the presentation and analysis of some representative light field panoramas. Finally, Section V concludes with final remarks and suggestions for future work.

II. LIGHT FIELD PANORAMA CREATION: BRIEF REVIEW

Nowadays, there are two main approaches to create a panoramic light field: with multiple cameras and with a single camera that moves/rotates during the acquisition process. The first approach usually requires specialized hardware, such as the Jaunt Neo camera [2] capable of high-quality, high-resolution, full 360-degree capture with custom optics to capture light field data. Also, the Lytro Immerge camera [3] is able to capture the visual scene simultaneously from many viewpoints using an array of sensors, which effectively acts as one giant light field camera with a wider field of view. In [4], a circular array of conventional wide field-of-view cameras are used and a dense omnidirectional light field is created by view interpolation. In [5][6], multiple spherically arranged cameras are used to record light field information from different directions. Using this approach, omnidirectional panoramas are created from the perspective of a virtual observer anywhere inside the spherical camera array by fusing information in the light ray space domain.

A panoramic light field can be also acquired with a single light field camera. In [7], a light field camera is rotated to capture the scene visual information in multiple light field views that are after stitched into a panoramic light field. In this solution, the acquired light rays are directly processed and this solution does not require depth reconstruction or matching of image features. However, a special panoramic tripod head is necessary to

ISBN 978-90-827970-1-5

capture panorama light fields of arbitrarily complex scenes. In a similar work, multiple light fields are stitched using the so-called ray-space motion matrix, which establishes a relationship between neighboring light fields [8]. Finally, in [9], a light field camera prototype is proposed which combines a spherical lens with a planar sensor, with the sensor rotating relative to a fixed main lens to capture a wide field of view panoramic image. The solution proposed here follows a different approach by performing the stitching of sub-aperture images.

III. PROPOSED LIGHT FIELD PANORAMA CREATION

In this section, the proposed light field panorama creation solution is presented, namely the system architecture and walkthrough, followed by a detailed description of the key modules. As the light field panorama created by the proposed solution preserves the directional ray information, it allows to change *a posteriori* the perspective and the objects in focus.

A. Architecture and Walkthrough

For the Lytro Illum light field camera, each light field image can be represented as a 4D array of 15×15 sub-aperture images with a 625×434 resolution each, which is directly obtained from the raw light field image organized in micro-images. In this paper, the key idea is to create the light field panorama by using a large set of 2D panoramas obtained by stitching corresponding sub-aperture images (corresponding to a specific angular direction) of the multiple light fields; naturally, it is assumed that these sub-aperture images have an appropriate overlapping. Fig. 1 illustrates three different light field images and the association between sub-aperture images of different light field captures, which will be used in the sub-aperture based stitching process.

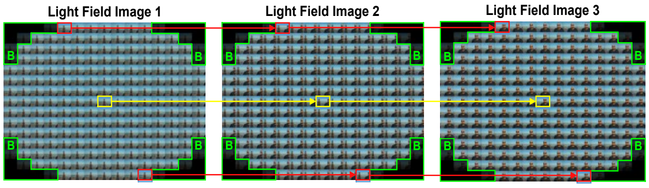

Figure 1. *Illustration of the stitching process for the sub-aperture images.*

First, the central images (yellow) located at position (8,8) of each sub-aperture structured light field image (the top-left sub-aperture image corresponds to the index (1,1)) are registered and stitched. Then, registration parameters are used to make the stitching of the remaining sub-aperture images (red rectangles and arrows). This will allow to obtain 2D panoramas whose disparity between them is similar to the disparity between the sub-aperture images used to create the respective 2D panoramas. In addition, by using the same registration parameters computed from the central sub-aperture image, it is possible to make the stitching process more coherent. That is, any stitching errors that may occur in the central sub-aperture image will also occur in the remaining sub-apertures images and the sub-aperture image content will only change due to occlusions or new content and not due to a different deformation or blending between adjacent sub-aperture images. From the stitching process, a set of 2D sub-aperture panoramic images are obtained, which are regarded as the final panoramic light field. The sub-aperture images located in the corners of the sub-aperture structured light field image (labeled with a green letter B in Fig. 1) are not used to create the final light field panorama, as they are black or very dark images due to the vignetting effect. These sub-aperture images may lead directly to black panoramas in the final light field panorama. Fig. 2 shows the system architecture of the proposed light field based panorama creation solution, which uses a feature-based registration approach for the sets of sub-aperture images.

Figure 2. *Architecture of the proposed light field panorama creation solution.*

The proposed solution walkthrough is as follows:

- **Light Field Acquisition:** In this step, the light fields of a visual scene are acquired from different perspectives using a Lytro Illum light field camera, a Nodal Ninja 4 panoramic tripod head and a Manfrotto 190CXPRO tripod. The Lytro Illum camera is mounted on the tripod head, with a constant rotation angle between each acquisition, to perform the acquisition of all parts (in the horizontal plane) of the visual scene. Due to the acquisition procedure, the light field panorama obtained has a 360º FOV in the horizontal direction and approximately 62º FOV in the vertical plane (no vertical rotation was performed). Regarding the Lytro Illum camera, the light rays are collected by a CMOS sensor (with 7728×5368 samples) containing an array of pixel sensors organized in a Bayer-pattern filter mosaic; this sensor produces GRBG RAW samples with 10 bit/sample. A lenslet array on the optical path allows to capture the different light directions.
- **Light Field Data Pre-Processing:** In this step, the raw light field data is pre-processed to obtain a 4D array with two ray directions and two spatial indices of RGB pixel data. At this stage, several operations are performed, namely demosaicing, devignetting, transforming and slicing, and finally color correction. The pre-processing is performed with the Light Field Toolbox developed by D. Dansereau [10]. At the end, a 4D array of sub-aperture images with a 625×434 resolution is obtained.
- **Central Sub-Aperture Image Registration:** In this step, the central sub-aperture images located at position (8,8) of all acquired light field images are registered. The goal is to obtain a set of registration parameters to perform the composition procedure described below. The main processes involved in this step are feature detection and extraction, image matching, an initial (rough) pairwise camera parameters estimation, global camera parameters refinement, wave correction and final panorama scale estimation. This processing module is described in detail in Section III.B. The outcome of this process are the registration parameters, notably the camera intrinsic and extrinsic parameters.
- **Composition:** In this step, all corresponding sub-aperture images in the multiple light field images are aligned and composed to produce a 4D array of panoramic images (i.e. the panoramic light field). The main processes involved in this step are image warping, exposure compensation, seam detection and blending. The central sub-aperture panorama is created first since the creation of the remaining panoramas use blending information from this central panorama. This step is described in detail in Section III.C.
- **Light Field Panorama Creation:** In this step, all 2D sub-aperture panoramic images are rearranged into a 4D light field in the same way as the input light field is represented. Thus, it is possible to perform rendering, e.g. extract a panoramic

image from a certain perspective, and *a posteriori* refocus to a specific depth plane in the visual scene in the same way as for the multiple input light field images.

In the next two sections, the most relevant modules in the proposed light field panorama creation framework are described.

B. Central Sub-Aperture Images Registration

The architecture for the central sub-aperture images registration process is shown in Fig. 3. The main goal is to compute a set of registration parameters from all central sub-aperture images of the several light field images, which cover different areas of the visual scene. These central sub-aperture images registration parameters will be used to compose all different sub-aperture panoramas in the composition process.

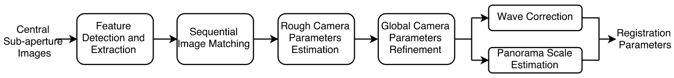

Figure 3. *Central sub-aperture images registration architecture.*

The registration process is performed with the following steps:

1. **Feature Detection and Extraction:** Local features are detected and extracted from all central sub-aperture images (one for each light field) using the SURF feature detector and extractor [11].
2. **Sequential Image Matching:** The set of features detected and extracted in the previous step are pairwise matched according to the acquisition order. This order reflects the position of each acquired light field image in the final light field panorama. The feature matcher acts as follows: 1) for each feature in one image, the two best descriptors in the other image are identified and thus two matches are obtained; 2) then, the two corresponding distances are obtained, which express how similar the two descriptors involved in each match are; 3) the ratio between the two distances (for the two matches) is computed and the best match is preserved only if this ratio is lower than a given threshold; 4) after, the RANSAC algorithm [12] is applied, estimating the transformation model between each pair of central sub-aperture images. The features that are not coherent with the estimated transformation model (i.e. outliers) are removed and only the inlier matches are kept. Fig. 4 shows an example of inlier matches between two overlapping central sub-aperture images after applying the RANSAC algorithm.

Figure 4. *Inlier matches between two central sub-aperture images.*

3. **Rough Camera Parameters Estimation:** Camera intrinsic (focal length) and extrinsic parameters (camera rotation) are roughly estimated. For each pair of overlapping central sub-aperture images, the camera intrinsic parameters (focal length) are estimated from the corresponding homography under the assumption that the camera undergoes a pure rotation to capture the different perspectives. All homographies used to estimate the camera intrinsic parameters are generated from the previously estimated sequential pairwise matches. Then, the final focal length value is computed as the median value of all focal length values estimated for each pair of overlapping central sub-aperture images.
4. **Global Camera Parameters Refinement:** The camera intrinsic (focal length) and extrinsic parameters (rotation) roughly estimated in the previous step are globally refined over each pair of matching images with a global alignment procedure. A bundle adjustment technique [13] is used to reduce accumulated registration errors resulting from the sequential pairwise image registration. This algorithm is used to update the camera parameters by minimizing the sum of squared errors associated to the projections of each feature into overlapping images with the corresponding features.
5. **Wave Correction:** A panorama straightening technique is applied to reduce the wavy effect that may occur in the final light field panoramic images. This technique straightens the final panorama by correcting the camera extrinsic parameters (e.g. rotation) to keep the ground level. This effect is due to unknown motion of the camera rotation central point (i.e. the nodal or no-parallax point) relative to a chosen world coordinates frame. Camera parameters are updated according to a global rotation that is applied in a way such that the vector normal to the horizontal plane containing both the horizon and camera centers is vertical to the projection plane. Fig. 5 illustrates the result of applying the panorama straightening technique on a sub-aperture panoramic image.

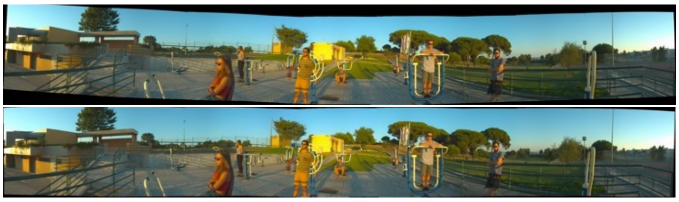

Figure 5. *Final panorama after all composition steps: (top) without and (bottom) with the Wave Correction step.*

6. **Panorama Scale Estimation:** In this step, the image scale of all the sub-aperture panoramas is estimated according to a specific focal length value. This is done by sorting in ascending order all the focal length values previously refined (i.e. updated in the global camera parameters refinement step) and selecting the middle value of this set. This module is performed in parallel with the previous step since the focal length values available after Step 4 are used. This scale value will be used later in the image warping process of all sub-aperture panoramic images.

C. Composition and Light Field Panorama Creation

Fig. 6 shows the architecture for the composition process. The dashed module and arrows represent a processing step (seam detection) that is only performed for the central sub-aperture images. The first panorama created is always the central sub-aperture panorama since the creation of the remaining panoramas requires some seam detection process information, notably warped masks corners relating the position of each sub-aperture image in the final light field panorama.

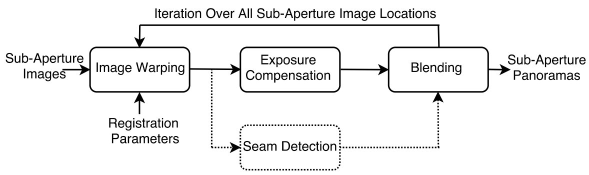

Figure 6. *Light field panorama composition architecture.*

ISBN 978-90-827970-1-5 © EURASIP 2018

The composition process aiming to create the sub-aperture panoramic images works as follows:

1. **Sub-aperture Image Warping:** In this step, image warping is performed using the sub-aperture images of the several light fields and the central sub-aperture images registration parameters (i.e. camera intrinsic and extrinsic parameters) previously estimated. All sub-aperture images are projected/warped using a spherical rotation warper according to the final sub-aperture panorama scale value and the camera parameters previously estimated in the central sub-aperture images registration process. Besides the warped images, the output of this step is also a collection of top-left corners (one corner for each warped image), which will be used in the blending process; the top-left corner corresponds to the location of the warped image top-left pixel on the sub-aperture panoramic image.
2. **Exposure Compensation:** In this step, an exposure compensation technique [14] is used with the goal to attenuate the intensity differences between the warped images composing the final panorama. This technique removes exposure differences between overlapping sub-aperture images by dividing each warped image into blocks and adjusting pixel intensities to achieve soft transitions in the overlapping regions.
3. **Seam Detection:** In this step, a graph-cut seam detection technique [15] is used with the goal of estimating seams, i.e. contours defining how the overlap areas in warped images contribute to the creation of the final sub-aperture panoramic image. The graph-cut seam detection technique determines the optimal position of each seam between all warped central sub-aperture images.
4. **Blending:** In this step, a multi-band blending technique [16] is applied to the regions where images are overlapping. The goal of this technique is to attenuate some undesired effects in each final sub-aperture panorama, such as visible seams due to exposure differences, blurring due to misregistration, ghosting due to objects moving in the scene, radial distortion, vignetting, and parallax effects, among others. This step uses the image masks obtained in the seam detection step (which is only performed for the set of central sub-aperture images), the corresponding top-left corners associated to the central sub-aperture images and the warped sub-aperture images.

After the blending process is finished and a sub-aperture panoramic image is obtained, the proposed solution starts the composition of the next sub-aperture panorama (goes to the first step of the composition process). This iterative loop runs over all sub-aperture images in the same position of the multiple acquired light fields, see Fig. 6. The final outcome of this process is a set of sub-aperture panoramic images, which all together, rearranged in a 4D array, represent a (lenslet) light field panorama.

IV. Performance Evaluation

In this section, the performance of the proposed light field panorama creation solution is assessed. This assessment involves perspective shift and refocus capabilities associated to the created light field panorama and, therefore, can only be made in a qualitative way, which makes it a challenging task when intending to show results on paper.

A. Test Conditions

A set of test scenarios was designed to evaluate the light field panorama creation solution where each scenario targets reproducing relevant acquisition conditions and corresponds to a single final light field panorama. In this context, the camera refocus range refers to the range of depth planes for which *a posteriori* refocusing is possible. The following scenarios were proposed:

- ***Short camera refocus range (indoor):*** Test case designed to evaluate the performance of the proposed solution when all objects are close to the camera, thus with large disparity between the partly overlapping light field images used in the panorama creation process.
- ***Large camera refocus range (outdoor):*** Test case designed to evaluate the proposed solution when the objects are near and far away from the camera, thus with very different disparities (from large to small) while still using a large refocus range.

For each test scenario, the rotation angle around the camera's optical center between each acquisition, the camera zoom, focus and refocus range remained constant; also, no zoom was used and the manual exposure mode was selected. The rotation angle between acquisitions was 15º for the first scenario and 30º for the second. The camera refocus range was 22cm to 6m for the first scenario and 30cm to ∞ for the second scenario.

B. Experimental Results and Analysis

To evaluate the proposed solution, the perspective shift and the refocus capabilities are assessed in a qualitative way and potential registration errors and blending artifacts (as in conventional 2D panoramas) are identified. To assess the perspective shift capability, five sub-aperture images located with some angular distance from the central perspective were extracted and evaluated. Since it is not easy to recognize the horizontal and vertical perspective shifts by visual inspection, especially due to the wide field of view (horizontal 360º) of the panorama, they are not shown here, but are available in [17]. To assess the refocus capability, the Light Field Toolbox software [10] was used. In this case, refocus is performed by shifting all the available sub-aperture images of each light field image to the same depth, and adding after all the sub-aperture images together to produce a single 2D depth plane. A parameter called *slope* allows to control the optical focal plane, and thus the object to be focused. For each light field panorama shown, some different focal planes are extracted and presented, as well as some close-ups corresponding to the presented depth planes; the red rectangles in Fig. 7 highlight the close-ups that are used to visualize the focus in specific parts of the light field panorama. For each close-up, the red circles in Fig. 8 highlight the interesting objects in focus.

As can be observed from Fig. 7 and Fig. 8, the light field panoramas obtained with the proposed solution do not have any clear registration errors or blending artifacts but still retain the desired refocus capability. In Fig. 7(a) and in the two corresponding close-ups in Fig. 8(a)(b), the two toy cars are in focus and the remaining acquired scene is not, as expected. On the other hand, in Fig. 7(b) and the two associated close-ups in Fig. 8(c)(d), show the background scene in focus and the rest of the scene blurred. Regarding the outdoor scenario with a large camera refocus range, the visual inspection of Fig. 7(c) and Fig. 8(e)(f) allows to notice that the persons close to the camera are the only object in focus, while the background is blurred.

From the results shown, it can be concluded that: 1) the light

Figure 7. *Three light field panoramas: depth plane extracted with (top) slope = - 0.05, (middle) slope = 0.6 and (bottom) slope = 0.15.*

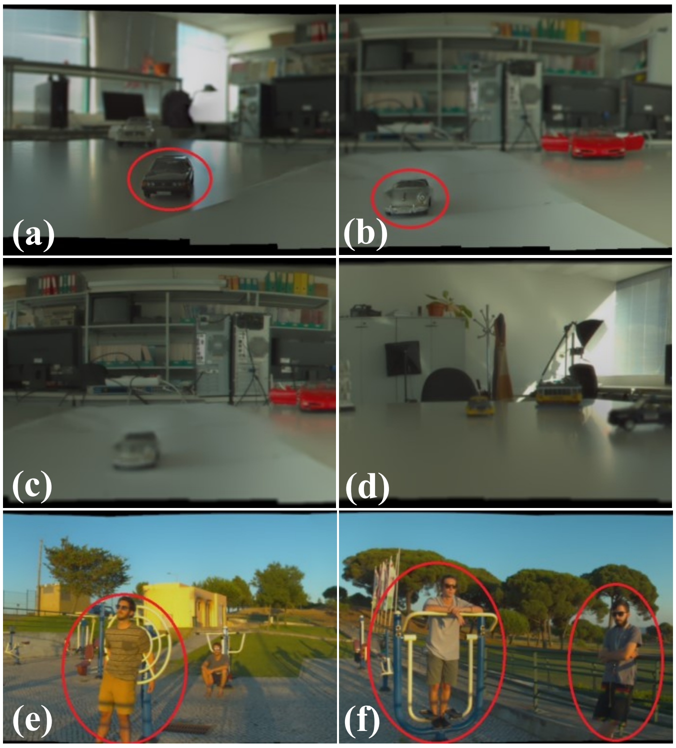

Figure 8. *Three close ups of the light field panoramas of Fig. 7*

field panoramas created can be refocused on different objects of the visual scene at the user's choice; 2) if the objects in the acquired visual scene are very distant from the camera, they will present very small disparities which can compromise the light field refocus capability (since, in this case, the depth of all scene objects is the same). Moreover, the light field panoramas with objects close to the camera have larger perspectives shifts in both horizontal and vertical directions since they have higher disparity that the objects far away from the camera. Finally, it is important to highlight that only a few selected examples (and test scenarios) are shown here due to space constraints. A more extensive evaluation is available in [17].

V. CONCLUSIONS

A light field panorama creation solution is proposed based on the stitching of the sub-aperture images for the same angular direction from partially overlapping light fields. For the stitching to be coherent among the sub-aperture images, it was necessary to determine key registration and composition parameters for the central sub-aperture image, which were after applied to the other sub-aperture images. The stitching process can be improved by explicitly estimating multiple homographies for regions of the image that are in different depth planes, thus enabling a more accurate registration between overlapping light fields.

REFERENCES

[1] G. Wu *et al.*, "Light field image processing: An overview," IEEE J. Sel. Topics Signal Process., vol. 11, no. 7, pp. 926-954, Oct. 2017.
[2] Jaunt, "First look at neo, " 2016, https://www.jauntvr.com/neo-first-look/.
[3] Lytro, "Built for the next generation of immersive storytelling," 2016, https://www.lytro.com/immerge.
[4] S. Maesen, P. Goorts, and P. Bekaert, "Omnidirectional free viewpoint video using panoramic light fields," 3DTV-CON, Hamburg, Germany, Jul. 2016.
[5] A. Akin *et al.*, "Hemispherical multiple camera system for high resolution omni-directional light field imaging," IEEE J. Emerg. Sel. Topics Circuits Syst., vol. 3, no. 2, pp. 137-144, Jun. 2013.
[6] V. Popovic, H. Afshari, A. Schmid, and Y. Leblebici, "Real-time implementation of gaussian image blending in a spherical light field camera," IEEE Int. Conf. on Industrial Technology, Cape Town, South Africa, Feb. 2013.
[7] C. Birklbauer and O. Bimber, "Panorama ligh-field imaging," Computer Graphics Forum, vol. 33, no. 2, pp. 43–52, May 2014.
[8] X. Guo, Z. Yu, S. B. Kang, H. Lin, and J. Yu, "Enhancing light fields through ray-space stitching," IEEE Trans. Vis. Comput. Graphics, vol. 22, no. 7, pp. 1852–1861, Jul. 2016.
[9] D. G. Dansereau, G. Schuster, J. Ford, and G. Wetzstein, "A wide-field-of-view monocentric light field camera," IEEE Conf. on Computer Vision and Pattern Recognition, Honolulu, HI, USA, Jul. 2017.
[10] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," IEEE Conf. on Computer Vision and Pattern Recognition, Portland, OR, USA, Jul. 2013.
[11] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: speeded up robust features," European Conf. on Computer Vision, Graz, Austria, May 2006.
[12] M. A. Fischer and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," Commun. ACM, vol. 24, no. 6, pp. 381-395, Jun. 1981.
[13] B. Triggs, P. Mclauchlan, R. Hartley, and A. Fitzgibbon, "Bundle adjustment – a modern synthesis," Int. Work. on Vision Algorithms, Corfu, Greece, Sept. 1999.
[14] M. Uyttendaele, A. Eden, and R. Szeliski, "Eliminating ghosting and exposure artifacts in image mosaics," IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, HI, USA, Dec. 2001.
[15] V. Kwatra, A. Schödl, I. Essa, G. Turk and A. Bobick, "Graphcut textures: image and video synthesis using graph cuts," ACM SIGGRAPH, San Diego, CA, USA, Jul. 2003.
[16] P. Burt and E. Adelson, "A multiresolution spline with application to image mosaics," ACM Trans. Graphics, vol. 2, no. 4, pp. 217-236, Oct. 1983.
[17] A. Oliveira, "Light Field based 360º Panoramas," M.Sc. Thesis, Insitituto Superior Técnico, Nov. 2016, https://fenix.tecnico.ulisboa.pt/downloadFile/1407770020544822