Light Field based 360º Panoramas

André Alexandre Rodrigues Oliveira

Abstract — This paper describes in detail the developed light field based 360º panorama creation solution, named as multi-perspective image stitching. The idea followed here is to create a light-field panorama from a set of 2D panoramas obtained by stitching all the perspective image stacks (i.e. the sub-aperture images corresponding to different sub-aperture light field images). This method stitch together a set of perspective images at the same location of each sub-aperture light field image stack, using classical 2D panorama creation techniques. The performance assessment of the proposed multi-perspective image stitching solution is made with relevant test scenarios proposed in this Thesis, since they attempt to reproduce relevant acquisition conditions of a common user in a given real scenario. The experimental results obtained show both light field refocus and multi-perspective capabilities in light field panoramic images created with the proposed multi-perspective solution.

Index Terms — Digital Photography, 360º Panorama Creation, Stitching, Plenoptic Function, Light Field

I. INTRODUCTION

This paper will focus on the development of an light field based 360º panorama creation solution presenting both perspective shift and refocus light field capabilities.

Photography is the process of recording visual information by capturing light rays on a light-sensitive recording medium, e.g. film or digital sensors. The result of this process, the images, is one of the most important communication medium for human beings, largely employed in a variety of application areas, from art and science to all types of businesses. However, the most common photography cameras developed up today present an important limitation: whether analog or digital, they have a limited field of view. Thus, it is not an easy task to encompass wide fields of view in a single camera shot.

With the desire to capture, in a single image, wide fields of view, panoramic photography has emerged as a technique that combines a set of partially overlapping elementary images of a visual scene acquired from a single camera location to obtain an image with a wide field of view. Conventionally, 360º panoramas creation is a process involving a sequence of different steps, starting with the acquisition of several images representing different parts of the scene and ending with a stitching process that combines the multiple overlapping images, resulting in the desired panorama. This procedure suffers from several limitations associated to the conventional imaging representation paradigm used where the images are just a collection of rectangular 2D projections for some specific wavelength components. Also, conventional cameras can merely capture the total sum of the light rays that reach a certain point in the lens using the two available spatial dimensions at the camera sensor, thus leading to loss of valuable visual information. This valuable information is the directional distribution of the light of the scene, which can be used in a number of different ways to improve the creation of panoramas and also to provide several new functionalities to the users. Without the use of this precious visual information, the user experience becomes greatly restricted and limited. The full real scene light field can be fully expressed by a well-known function characterizing the amount of light traveling through every point in space, in every direction, for any wavelength, along time, the so-called (7D) plenoptic function. For this reason, the search for novel, more complete representations of the world visual information (i.e. higher dimensional representations) has become a hot research field, in a demand to offer the users more immersive, intense and faithful experiences.

Recently, new sensors and cameras, e.g. the plenoptic or light field cameras has emerged, which capture the angular light information. These new cameras have an innovative design where a micro-lens array allows to capture the light for each position and from each angular direction, i.e. they capture and record a 4D light field. Each of these micro-lenses captures a slightly different perspective/view from the scene, allowing these cameras to record not only the location and intensity of light, but also to differentiate and record the light intensity for each incident direction.

This important characteristic allows capturing a much richer visual information representation of the scene, which can be used to overcome the limitations related to the conventional imaging representation and capture. For instance, this richer representation of the visual information brings additional interesting capabilities like the ability to a posteriori refocus any part of the image, relighting and recoloring, slightly changing the user viewpoint, among others. All these new capabilities will inevitably lead to the reinvention of concepts and functionalities associated with digital photography and 360º panoramic image creation solutions.

With this in mind, the objective of this work was to develop a light field based 360º panorama creation solution, named as multi-perspective image stitching. The proposed multi-perspective image stitching solution is inspired by the work of Brown and Lowe in [1]. The light field 360º panorama creation developed is able to still maintain the desired refocus and perspectives shift capability on the light field panoramas.
created.

In the following section an overview over the State-of-the-art on convention 360º panoramas creation solution will be presented. Section II will present the light field based 360º panorama creation solution proposed, notably its architecture and main tools. Section III will conclude with a summary and the future work plan.

II. MULTI-PERSPECTIVE IMAGE STITCHING SOLUTION

This section describes the proposed light field based 360º panorama creation solution. In this context, it starts by describing the global system architecture and walkthrough, followed by a detailed description of the main parts, namely the light field data pre-processing module (which corresponds to a solution already available in the literature [2]) and the key modules used to create the panorama light field image.

A. Global System Architecture and Walkthrough

The main goal of this section is to describe the global system architecture and walkthrough of the proposed light field based 360º panorama creation solution. Each light field image is acquired with the Lytro Illum camera [46] and is represented as a 4D matrix of 15x15 sub-aperture images (called here as a perspective image stack). The idea followed here is to create a light-field panorama from a set of 2D panoramas obtained by stitching all the perspective image stacks (light field images). This method, which is named here as multi-perspective image stitching, stitches together a set of perspective images at the same location of the image stack, using classical 2D panorama creation techniques. Figure 1 illustrates 3 different light field images (as perspective images stacks) and the association between perspective images of different light field captures which will be used as input for the stitching process.

The idea is to first perform stitching in the central image (yellow rectangles and arrows) located at position (8,8) of the sub-aperture light field image and, then, derive some (registration) parameters which are used to make the stitching of the remaining perspective images (red rectangles and arrows). This will allow to maintain the disparity between the different perspectives of the light field panoramas similar to the disparity between the perspectives of each of the light field image (before stitching). In addition, by using the same (registration) parameters obtained from the central perspective image is possible to make this process more coherent, i.e. any stitching errors will occur in all perspective images and the perspective image content will only change due to occlusions or new content and not due to a different deformation or blending between adjacent perspective images. From the stitching process it is obtained a set of 2D perspective panoramic images which are regarded as the final panoramic light field. The perspective images that are within the matrix 4 regions (for each light field image presented) highlighted with green and labeled with a green letter “B” are not used to create the final light field 360º panorama because they are black or very dark images and thus not useful. These perspectives are replaced in the final light field 360º panorama by black panoramas. All the perspective panoramas used (i.e. the 2D perspective panoramas and the black panoramas) have the same resolution.

In the following, a brief walkthrough of the proposed solution depicted in Figure 2 is presented:

1. Light Field Acquisition: In this step, the light field of a visual scene from different perspectives is acquired using a Lytro Illum light field camera [46]. The Lytro Illum camera is mounted on the panoramic tripod head which rotates around the camera’s optical center (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 may have a FOV of 360° in the horizontal direction and approximately 62° of FOV in the vertical plane (corresponds to the vertical FOV of the Lytro Illum camera since no vertical rotation is performed in this acquisition step). Regarding the Lytro Illum camera, the light rays are collected by a CMOS sensor (with 7728 x 5368 samples) containing an array of pixel sensors organized in a Bayer-pattern filter mosaic as illustrated in Figure 25(a); this sensor produces GRBG RAW samples with 10 bit/sample. Naturally, a lenslet array on the optical (illustrated in Figure 25(b)) path allows to capture the different light directions. The Lytro camera stores the acquired information in the so-called .LFR files; this is a container format that stores various types of data, notably the Raw Bayer pattern GRBG image, associated metadata, a thumbnail in PNG format and system settings, among others. The remaining acquisition conditions are described in Section 5.1.

2. Light Field Data Pre-Processing: In this step, the RAW light field data produced by the Lytro Illum camera is pre-processed to obtain a 4D light field, i.e. a 4D dimensional array with two ray directions and two spatial indices of pixel RGB data. In this case, several operations are performed, namely demosaicing, devignetting, transforming and slicing, and finally color correction (rectification light field processing is not performed). The pre-processing applied over the RAW light field data is done using the Light Field Toolbox developed by D. Dansereau [51]. After, all the 2D perspective images (of the 4D light field) are stored, i.e. it was extracted from the 4D Light Field (LF) array, 193 perspective images, which are stored in the bitmap format (the images obtained from one light field image correspond to a perspective image stack). Also, a directory file which lists all extracted perspective images obtained from the set of light field images that compose the final panorama was written. This file indicates the order (i.e. sequential order, thus the images need to be processed.
According to their position in the final light field panorama) of the input light field images and associated perspective image stacks which will be used in the creation of the final light field 360° panorama; note that the order needs to be sequential. All perspective images have a resolution of 623 x 434 pixels. Ideally, the resolution should be 625 x 434 pixels but due to the presence of black pixels in the first and last columns of each perspective image it was necessary to remove the first and the last column of each perspective image.

3. Central Perspective Images Registration: Central Perspective Images Registration: In this step, the central perspective images located at position (8,8) of the sub-aperture light field image (one for each different perspective image stack created in the previous step) are registered. The goal of this step is to obtain a set of registration parameters that will be used to perform the composition (next step) of all perspective images of each different light field. The main processes involved in this step are feature detection and extraction, image matching, an initial (rough) pairwise camera parameters estimation (intrinsic and extrinsic parameters), global camera parameters refinement, wave correction and final perspective panorama scale estimation. The outcome of this process is the registration parameters, which are the camera intrinsic and extrinsic parameters.

4. Composition: In this step, all corresponding perspective images in each different perspective image stack will be composed to produce the required perspective panoramic images to create the final light field panorama. Thus, the goal of this step is to align and compose all 2D panoramas, this means one 2D panorama for each different perspective of the 4D light field. These 2D panoramas will use the central perspective images registration parameters previously estimated (camera parameters). The main processes involved in this step are image warping (where is applied the spherical projection), exposure compensation, seam detection and blending. The registration module is described in detail in Section 4.3. The first perspective panorama created is the central perspective panorama since the creation of the remaining panoramas use information of the blending process (image warped masks and corners) obtained from the composition of the central perspective. The perspective images that correspond to the corners of the 2D array of light field images that are too dark (as described in Step 2) will be replaced by black panoramic images with the same resolution of the created 2D panoramas. The outcome of this process will be a set of perspective 2D panoramas (193 perspective panoramic images and 32 black panoramas), all with the same resolution.

5. Light Field 360° Panorama Creation: In this step, all perspective 2D panoramic images are rearranged into a 4D light field in the same way as the input light field is represented after the pre-processing module previously described in Step 2. This means that by storing the final light field 360° in this 4D format it is possible to perform some rendering, e.g. extract a single perspective panoramic image or refocus a posteriori in a specific depth plane of the acquired visual scene in the same way as a usual light field image.

B. Main Tool: Detailed Description

This section describes in detail the main tools of the proposed light field based 360° panorama creation solution. In the implementation of the proposed solution, it was used the OpenCV library [54] where some of the processing modules described in the following are implemented. The main tools are the central perspective images registration and composition processes of the global system architecture presented in the following.

The central perspective images registration process architecture of the proposed light field based 360° panorama creation solution is shown in Figure 31. The main goal of the central perspective images registration is to compute a set of registration parameters from all central perspective images of the different perspective image stacks (obtained from the several LF 4D images covering different areas of the visual scene). This central perspective images registration parameters will be used to compose all different perspective panoramas in the composing process described in detail in Section 4.3.2. Figure 3 depicts the central perspective images registration process architecture.

Figure 3 - Central perspective images registration architecture of the proposed light field based 360° panorama creation solution.

A walkthrough of the registration process architecture illustrated in Figure 31 is presented, where the main tools deserve more detail:

1. Feature Detection and Extraction: In this step, local features [11] are detected and extracted from all central perspective images (one for each perspective image stack) using the SURF feature detector and extractor [10]. The SURF detector is a blob detector which is based on the Hessian matrix to find points of interest. The SURF descriptors characterize how pixel intensities are distributed within a neighborhood of each detected point of interest (keypoint). SURF descriptors are robust to rotation, scale and perspective changes in a similar way to the SIFT descriptors. Figure 4 shows the features detected from 2 overlapping central perspective images (no feature scale or orientation is shown to allow the visualization of the content and keypoint descriptor location).

Figure 4 - Features detected and extracted from 2 overlapping central perspective images.

2. Sequential Image Matching: In this step, the set of features detected and extracted (from all central perspective images of each perspective image stack) in the previous step is pairwise matched according to the order presented by the directory file
Global Camera Parameters Refinement

Global Camera Parameters Refinement: In this step, the camera intrinsic (focal length) and extrinsic parameters (camera rotation) are roughly estimated. For each pair of overlapping central perspective 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 different areas of the visual scene. All transformations (i.e., homographies) used to estimate the camera intrinsic parameters are generated from the previously estimated sequential pairwise matches (Step 2). Then, the median value from all estimated focal length values (one for each pair of overlapping central perspective images) will be considered the focal length value to be used in the next step. Camera translation is assumed to be zero during the whole light field 360º panorama creation pipeline.

4. Global Camera Parameters Refinement: In this step, the camera intrinsic (focal length) and extrinsic parameters (rotation) roughly estimated in the previous step are globally refined with a global alignment procedure over each pair of matching images thus reducing accumulated registration errors resulting from the sequential pairwise image registration. This is achieved using a bundle adjustment technique [27] which simultaneously refines the camera intrinsic (focal length) and extrinsic (camera rotation) parameters. The bundle adjustment technique only considers the overlapping pair of images that have a confidence value (expresses the reliability of estimated homography for each pair) above a given threshold. In this case, the bundle adjustment technique used minimizes the sum of the distances between the rays passing through the camera center and the SURF features and the matches estimated in Step 2. The Levenberg-Marquardt algorithm [28] is used to update the camera parameters by minimizing the sum of squared projection errors associated to the projections of each feature into overlapping images with corresponding features.

5. Wave Correction: In this step, a panorama straightening technique is used with the goal of reducing the wavy effect that may occur in each final 2D perspective panoramic images. This technique is able to straighten 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 since it is rather hard to maintain the nodal point perfectly static and stable during the camera acquisition of all light field images that compose the final panorama. Since only camera horizontal rotations are considered during the whole light field 360º panorama creation pipeline, the unknown motion of the nodal point is not considered in previous registration steps. Camera parameters are updated according to a global rotation which is applied such that the vector normal to the horizontal plane containing both the horizon and camera centers is vertical to the projection plane. Figure 6 illustrated the result of applying the panorama straightening technique described over perspective panoramic image.

Figure 6 - Wave correction examples: (a) without and (b) with applying the panorama straightening technique. Both examples presented are the final panorama that was obtained after all composition steps.
performed to the central perspective images. The creation of the remaining panoramas requires the use of some information of the seam detection process of the central perspective panorama, namely image warped masks corners relating the position of each image in the final light field 360° panorama. The orange arrow coming from the blending process to the image warping process symbolize the iteration loop over all perspective images of different perspective images stacks, i.e. the proposed solution iterates over all perspective stacks to create a set of perspective panoramas.

Figure 7 – Composition architecture of the proposed light field based 360° panorama creation solution.

In the following, the walkthrough of the composition process architecture illustrated in Figure 7 is presented while describing in detail the main tools:

1. **Image Warping:** In this step, image warping is performed using all perspective image stacks and the central perspective images registration parameters (i.e. camera intrinsic and rotation parameters) previously estimated in the registration process. The goal of this process is to apply a deformation of all input images according to the selected projection and to obtain a set of top-left corners that will be used in the blending process later described. Thus, all perspective images are projected/warped using a spherical rotation warper according to the final perspective panorama scale value and the camera parameters (i.e. intrinsic parameters and rotation) previously estimated in the central perspective 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). The top-left corners obtained from the image warping of the central perspective image are used later in the blending process of all remaining perspectives of the final light field 360° panorama. All warped images will be used later in the exposure compensation and blending processes.

2. **Exposure Compensation:** In this step, an exposure compensation technique [55] is used with the goal of attenuate the intensity differences between the warped images that compose the final panorama. The technique used tries to remove exposure differences between overlapping perspective images by adjusting image block intensities. By dividing each warped image in blocks and making use of the overlapping and non-overlapping information for each pixel, soft transitions are achieved within a perspective panorama containing various overlapping regions and also between different overlapping perspective images.

3. **Seam Detection:** In this step, a graph-cut seam detection technique [56] is used with the goal of estimating seams, i.e. lines which define how the overlap areas in warped images will contribute in the creation of the final perspective panoramic image. With this goal in mind, image masks the seams contribute in the creation of the final perspective panoramic images, which is the reason why this module is dashed. The graph-cut seam detection technique determines the optimal position of each seam between all warped central perspective images, enabling the composition process of all perspective panoramas. This technique creates the images masks which defines the seam to compose all images of the panorama using the top-left corners obtained from the image warping process previously described. Figure 8 illustrates an image mask resulting of the seam detection process over all central perspective images. The white region defines the position of the central perspective image corresponding to the image mask presented and the lines that separate the white region from the black one are detected seams.

4. **Blending:** In this step, a multi-band blending technique [29] is applied to the region where images are overlapping. The goal of this technique is to attenuate some undesired effects that may exist in each final perspective panorama, such as visible seams due to exposure differences, blurring due to misregistration, ghosting due to objects moving in the scene, radial distortion, vignetting, parallax effects, among others. To solve this problem, each image, expressed in spherical coordinates, is iteratively filtered, using a Gaussian filter with a different standard deviation value in each iteration, and a high pass version of each image is created by subtracting the original image from its filtered version (in each iteration); the high pass version of the image represents the spatial frequencies in the range established by the Gaussian filter standard deviation value. Blending weight maps are also created for each image in each (image filtering) iteration. The blending weight map is initialized for each image by finding the set of samples in the previously created panoramic image for which that image is the most responsible (for the sample values) and then is iteratively filtered (at the same time the image filtering process takes place) using the same filter applied to the image. The final panorama results from a weighted sum of all high pass filtered versions of each overlapping image where the blending weights were obtained as previously described. Therefore, low frequencies are blended over a large spatial range, while high frequencies use a short range, thus allowing smooth transitions between images even with illumination changes (while at the same time preserving high frequency details). This step uses the image masks obtained in the seam detection step (which is only performed to the set of central perspective images) and top-left corners associated to the central perspective images and the warped perspective images needed to create a perspective 2D panorama.

After finish the blending process over some perspective panoramic image, the proposed solution starts the composition of the next perspective panorama (i.e. it comes back to Step 1 of the composition process) that corresponds to the neighboring perspective of the right side. The final outcome of this process are a set of perspective 2D panoramic images, which all
The motivation of this work was the enhancement of 360° panoramic photography with additional features such as refocusing and the major objective was the development of a light field based 360° panorama image creation solution that was named multi-perspective image stitching and is inspired by the work developed by Brown and Lowe. The main concept behind this solution was to create light field 360° panoramas from a collection of 2D perspective panoramas. The conventional 360 panorama creation architecture was adapted to deal with light field input and thus, for the stitching to be coherent among sub-aperture images it was necessary to calculate key registration and composition parameters only for the central view which are applied to other views of the collection of light field images that compose the final panorama.

The performance assessment of the proposed multi-perspective image stitching solution was made with relevant test scenarios proposed for the first time in this Thesis, which are critical for the adequate assessment of the proposed solution. These test scenarios attempt to reproduce relevant acquisition conditions of a common user in a given real scenario. The experimental results obtained show both light field refocus and multi-perspective capabilities in panoramic images created with the multi-perspective solution, however the results are not presented in this article because of their size and format, which would lead to a difficult understanding due to the lack of quality, therefore it is suggested the full reading of the thesis document.

In this context, it is possible to conclude that the proposed multi-perspective image stitching solution allows the creation of light field 360° panoramas under different types of realistic scenarios. Also, both light field refocus and multi-perspective capabilities are possible in all light field panoramas created. By observing the perspective shift capability assessment, it is possible to conclude that: 1) the light field panoramas that were acquired in visual scenes where the objects are close to the camera present higher perspectives shifts in both horizontal and vertical directions. This is justified with the fact the objects that are close to the camera present higher disparity that the objects that are far away from the camera; 2) the design of the light field camera used (i.e. the Lytro Illum camera) does not allow to capture great levels of disparity between different perspectives. By observing the refocus capability assessment, it can be deduced that: 1) the light field panoramas created can be refocused on different objects present in the acquired 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 compromising the light field refocus capability because, in this case, the refocus technique considers that the depth of all scene objects is the same; 3) the resolution of the sub-aperture created using the Light Field Toolbox [45] (and thus, the resolution of the final light field 360° panorama) is a considerable limitation when finding very different depth planes to refocus the scene objects since it is not very easy to accurate distinguish focus in different objects if these objects are beyond a certain distance from the camera. Considering all results obtained one of the major conclusions of this Thesis is that the creation of light field panoramas excels for visual scenes containing objects close to the camera.

The light field 360° panorama creation developed is able to still maintain the desired refocus and perspectives shift capability on the light field panoramas created. However, there are important limitations that may be addressed to improve the proposed multi-perspective solution. Since the light field imaging representation is a relatively new topic there are not many panorama creation solutions based on light field images and thus, it is expected that new and innovative light field 360° panorama creation techniques will be proposed in the future.

Regarding the proposed solution, some improvements are possible to improve the quality of the light field 360° panoramas created. Some suggestions aiming to improve the developed solution are listed:

- **Depth-based Light Field Panorama Creation:** To minimize the stitching errors and allow to capture well the disparity to objects close to the camera and of objects that are far away it is possible to improve the stitching process by: 1) estimate the depth of the acquired visual scene in each light field image used and 2) use this information in the registration process by estimating multiple homographies for regions of the image which are in different depth planes, thus, enabling a more accurate multi-perspective stitching process [40].

- **Light Field Panorama Rendering:** Another topic, is the development of a rendering tool appropriate for the light field panoramas, giving the user the possibility to interact with light field 360° panorama content, e.g. using his/her mouse by rotating it in all directions or navigate through the whole acquired visual scene, making zoom-ins and zoom-outs, etc. to enjoy a more immersive user experience. In addition, to the usual interactions with conventional panoramas the visual scene could be rendered with certain depth of field and allow minor perspective adjustments. This type of rendering tool could be also relevant for the visualization of light field panoramas while giving the user the depth impression, e.g. rendering the content in stereoscopic or virtual reality head mounted displays.

- **Unrestricted Light Field Panorama Creation:** Another topic that could be interesting to improve the proposed solution is the creation of light field 360° panoramas in an unrestricted way, i.e. moving the camera handheld and thus with some unrestricted rotation and translation camera motion. The tripod-base scenario used here assumes that camera undergoes a pure rotation around its no-parallax point and is very common among professional photographers. However, there are many solutions which do not have this constraint (e.g. using smartphone cameras) and thus it is important to also target these cases.

IV. **BIBLIOGRAPHY**


