Computers & Graphics 53 (2015) 127–135

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Technical section

Active guidance for light-field photography on smartphones

C. Birklbauer*, O. Bimber

Institute of Computer Graphics, Johannes Kepler University Linz, Austria

ARTICLE INFO

Article history:
Received 17 December 2014
Received in revised form
16 September 2015
Accepted 17 September 2015
Available online 3 October 2015

Keywords:
Light field
Augmented reality
Mobile
Guidance

ABSTRACT

We present a novel approach for guided light-field photography using off-the-shelf smartphones. In contrast to previous work that requires the user to decide where next to position a mobile camera, we actively compute and visualize during runtime a recommendation for the next sampling position and orientation taking into account the current camera pose and required camera alignments. This supports efficient capture of various types of large-field-of-view light fields in just a matter of minutes and without specialized camera equipment. To further reduce the overall capture time, we describe an extension of our guidance algorithm to collaborative light-field photography by small groups of users.

1. Introduction and contributions

Light fields have the potential to revolutionize photography and radically change everything we relate to digital imaging today. By providing not only spatial but also directional information, they support synthetic refocusing, multi-perspective recording, depth-variant filtering, and much more. First compact plenoptic cameras have already been introduced to the market (e.g., Lytro, Raytrix, Pelican Imaging). Although these cameras offer quick and easy recording of light fields, they are specialized devices that are clearly constrained by limited resolution, small aperture, and narrow field of view.

Modern smartphones represent an alternative with the potential to bring light-field photography to the mass-market. Rather than recording light fields in one shot, moving a mobile camera enables sequential sampling of the ray data. A key to success, however, is a capture procedure that is efficient, fast, and easy to use. Thus far, first attempts at sequential light-field sampling with a manually moved camera have only visualized the current sampling status of spherical light fields [1], leaving the decision of what to capture next entirely to the user. Furthermore, the scenes captured with such passive guidance methods have to fit into the field of view of the capture camera due to the only two-dimensional visualization of the sampling status.

In this paper, we present a first active guidance algorithm that allows capturing different types of light fields, such as two-plane, spherical, and panoramic light fields with large fields of view (cf. Fig. 1). We compute and display a recommendation for the next sampling position and orientation taking into account the current camera pose and required camera alignments to enable fast and easy recording of high-quality light fields without the need for in-depth knowledge of the underlying parameterization. Our recommended sampling poses ensure a fast and uniformly distributed reduction of undersampling artifacts and consequently allow for an early abortion of the capture process by the user. Our guidance approach is efficient enough to run in real-time on performance limited mobile platforms. Furthermore, we explain how our algorithm is extended to collaborative light-field photography to reduce recording times by simultaneously guiding small groups of users to each capture parts of the same light field.

Our key contributions are:(1) A novel guidance algorithm for capturing a variety of light-field types that proposes and actively directs users to a next sampling position and orientation; (2) an extension of this guidance algorithm to support a first realization of collaborative light-field photography; (3) a method for automatically computing the parameterizations of two-plane, spherical and panoramic light-field sampling based on scene and camera properties.

We implemented our algorithm on Android smartphones and evaluated its effectiveness for single and multiple collaborating users.

2. Related work

In many cases, special hardware, such as microlens arrays [2,3], camera gantrys [4], or camera arrays [5], is required for light-field imaging. While compact microlens-based light-field cameras are becoming increasingly popular, large-field-of-view light fields can only be captured with the help of spherical mirrors [6], rotating

* Corresponding author.
E-mail addresses: clemens.birklbauer@jku.at (C. Birklbauer), oliver.bimber@jku.at (O. Bimber).

http://dx.doi.org/10.1016/j.cag.2015.09.008
0097-8493/

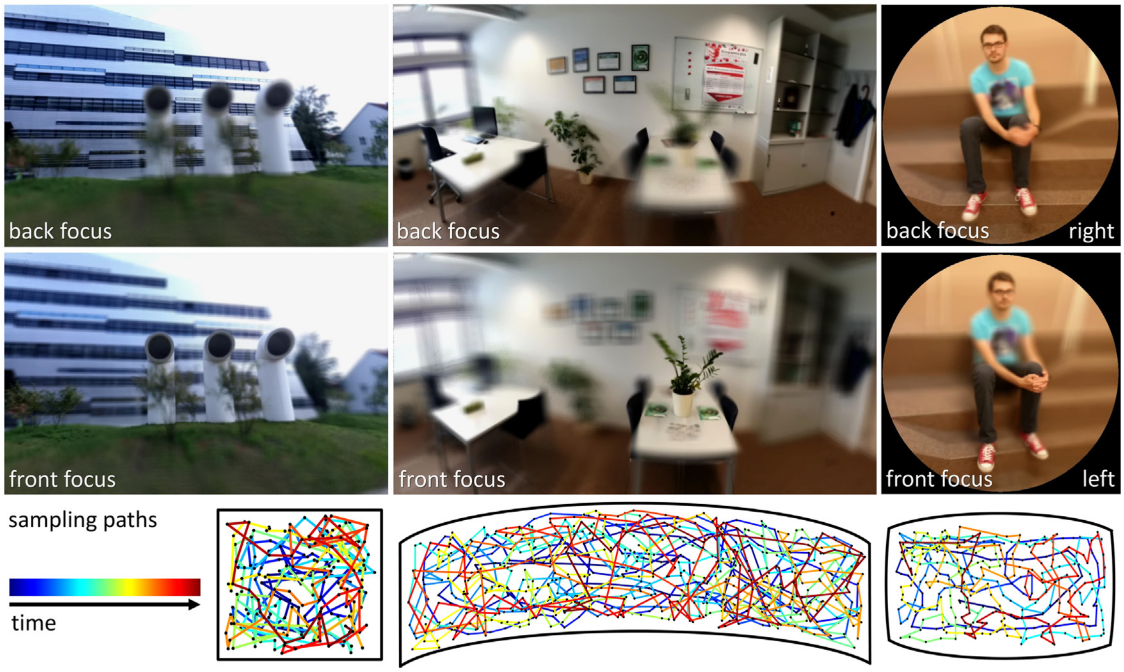

Fig. 1. Three different types of light fields, each captured with a smartphone (640 × 480, FOV: 62°H, 48°V) within 5–10 min. From left to right: two-plane parameterized light field rendered with a wide aperture (FOV: 83°H, 58°V; # images: 225), panoramic light field rendered with a wide aperture (FOV: 130°H, 60°V; # images: 675) and spherical light field rendered with a medium aperture (FOV: 40°H, 40°V; # images: 332). The bottom row shows a visualization of the sampling path for all three light fields over time.

camera arms [7], or rotating light-field cameras [8]. Early approaches that apply hand-held cameras for light-field recording do not provide any online guidance to the user [9,10], but determine camera poses using structure from motion during post-processing after recording. This can lead to inefficient under- or oversampling. Techniques such as that in [1,11] guide the user passively during recording by visualizing only previously captured camera positions on a sphere around the scene. Ensuring adequate sampling density and distribution, however, is left to the user. Thus, an active guidance through a recommended sampling is not supported in these cases. Furthermore, a simple visualization of the (two-dimensional) sampling status only supports guidance for those scene portions that are inside the field of view of the camera. An extension to four dimensions is also not trivial.

Collaborative imaging enabled by asynchronous crowd sourcing and offline processing, as in [12], has become a common research topic. Simultaneous crowd imaging, as in [13], supports smooth view-point transitions within a captured scene. Such crowd-sourcing techniques do not offer users an ideal sampling pattern to follow. Only for panoramic video recording active guidance for collaborative recording was presented in [14]. However, the two-dimensional nature of video recording is too simple to be extended to the recording of four-dimensional light fields.

Optimal light-field sampling was analyzed in [15] for two-plane parameterizations and in [1] for unstructured (mainly spherical) light fields. Others (e.g. [16]) also explored the possibility to reduce the required number of samples by taking depth reconstruction into account. By reconstructing depth, a light field can even be created from a single stereo pair as in [17]. We apply similar ideas to identify ideal sampling patterns based on the scene's depth range and present a practical way for automatically computing two-plane, spherical and panoramic light-field parameterizations from a sparse point cloud of the scene. However, we currently do not consider dense depth reconstruction as this may reduce the applicability of our method for scene parts where it completely fails (e.g. for highly specular or semi-transparent surfaces or occluded scene parts). In contrast to previous approaches, we provide active online guidance to the user during runtime to capture wide-aperture and wide-field-of-view light fields with off-the-shelf smartphones aiming for a minimal recording effort. We also extend our guidance approach to collaborative light-field photography, and compare both single-user and collaborative sampling in the context of a user study.

Active user guidance requires some sort of view planning to suggest ideal sampling positions to the user. Content based view planing for (non-uniform) synthetic light-field acquisition was explored in [18]. Potential new views are first interpolated from existing views (with known depth) and are only rendered if the estimated quality of the interpolated view is low. A similar metric is used in [19] for virtual scenes where depth is not known but reconstructed. However, a high quality and dense depth reconstruction, as required in these algorithms, is unrealistic to achieve for complex real-world scenes (e.g. scenes containing transparency or specularity). Different to our method these approaches are also not designed to run in real-time on performance limited mobile platforms and do not plan novel views based on the current user position.

3. Guided light-field photography

Our goal is to capture light fields while moving a continuously recording smartphone successively to suggested next sampling poses that uniformly minimize undersampling and avoid oversampling while taking into account that the recording can be aborted at any time. During capturing, we provide visual and acoustic guidance for the user to follow it as quickly and as easily as possible. While Section 3.1 explains how the recommendations for next sampling poses are computed, Section 3.2 presents our guidance interface. All of the above requires real-time pose estimation, which is achieved with SLAM tracking, as implemented in [20].