

Light Field Completion Using Focal Stack Propagation

Terence Broad and Mick Grierson
Goldsmiths, University of London*

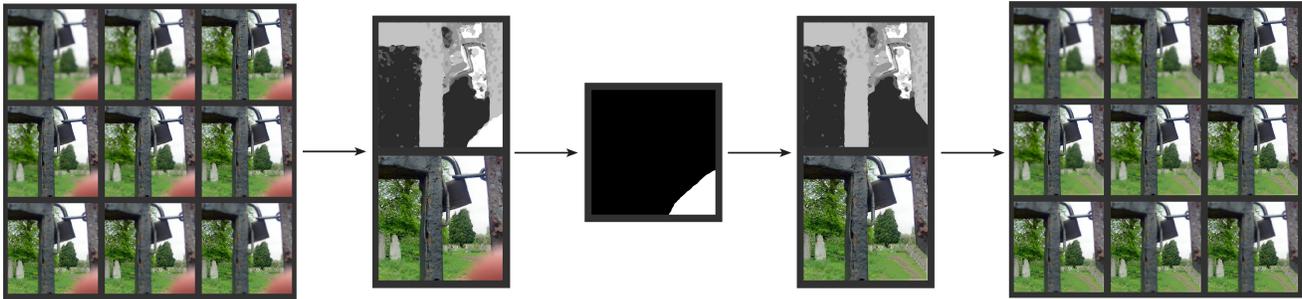


Figure 1: Key steps in algorithm: We extract a focal stack from the light field, create a depth map and all-in-focus image, given a user defined mask (center) we first perform image completion on the depth map then the all-in-focus image, finally we propagate the synthesised image segment through the focal stack which can then be used to re-sample the light field.

Keywords: light field, image completion, texture synthesis, focal stack photography, computational photography, relative depth maps

Concepts: •Computing methodologies → Image manipulation; Computational photography;

1 Introduction

Both light field photography and focal stack photography are rapidly becoming more accessible with Lytro’s commercial light field cameras and the ever increasing processing power of mobile devices. Light field photography offers the ability of post capturing perspective changes and digital refocusing, but little is available in the way of post-production editing of light field images. We present a first approach for interactive content aware completion of light fields and focal stacks, allowing for the removal of foreground or background elements from a scene.

2 Technical Approach

While some occluded information can be recovered when multiple views of the scene are available [Yatziv et al. 2004], this alone is not sufficient for the completion of large areas in light fields. This method is also not applicable to light fields captured by plenoptic cameras, as they do not capture enough lateral variation in potential viewpoints. Large areas must be synthesised in a way that is visually plausible and coherent within the rest of the scene, applying image completion algorithms independently to images captured from separate viewpoints does not result in light fields that maintain spatial or angular consistency. Therefore we pursued an approach where the target segment is synthesised only once.

*e-mail: {ma201tb, m.greirson}@gold.ac.uk

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). © 2016 Copyright held by the owner/author(s). SIGGRAPH ’16, July 24-28, 2016, Anaheim, CA, ISBN: 978-1-4503-4371-8/16/07 DOI: <http://dx.doi.org/10.1145/2945078.2945132>

Levin and Durand [2010] demonstrated that novel viewpoints from the 4D light field can be rendered from a 3D focal stack sequence (all rendered from the central view point) by spatially shifting the focal stack sequence, averaging the focal stack slices into one image, and de-convolving the images with perspective dependent point-spread function. We thus have chosen to synthesise an image segment and propagate it through the focal stack, as maintaining spatial and angular consistency of a synthesised area is much simpler in the focal stack domain than in the domain of images taken from multiple viewpoints.

The algorithm is as follows: 1. Render the focal stack sequence from the central perspective of the light field. 2. From the focal stack make a depth map by performing the laplacian transformation of all the focal stack images and calculate approximate depth from regions of highest frequency in focal stack images. 3. Create an all-in-focus image by extracting and composing the highest-frequency image content throughout the focal stack slices. 4. The user interactively defines a mask on the all-in-focus image, defining the area of the image to be completed. 5. The defined region is completed in the depth map using the Efros and Leung texture synthesis algorithm [1999]. 6. The all-in-focus image is completed using an adapted version of the Efros and Leung texture synthesis algorithm, adding a heuristic to weight against pixels that are more distant in depth from the target pixel on the completed depth map. 7. The synthesised segment is separated into smaller segments corresponding to the regions the segment covers on the depth map. 8. A Gaussian convolution kernel for each in-focus region corresponding to the same region that is out of focus in the other focal stack layers is approximated; taking the highest frequency image region of the laplacian, iteratively convolving the region with varying isotropic Gaussian kernels and selecting the kernel that produces the smallest mean difference between convolved image region and the target out of focus region. 9. Each segment is then propagated into the focal stack images (in order of most distant to nearest) being convolved the corresponding approximated kernels, then superimposed on the the original focal stack images. 10. Once the all of the images in the focal stack have had the synthesised segments propagated through them, novel viewpoints from the completed 4D light field can be rendered from the focal stack using the linear view synthesis algorithm described by Levin and Durand [2010].

Relative Depth Maps

In order to solve some of the unique challenges faced in the implementation of this method, we have developed a novel feature map to spatially manipulate and weight pixel information. We create relative depth maps for each focal stack slice that are then subtracted from the focal stack images and laplacian transformation of the focal stack images in order to enhance feature extraction.

This technique is used in step 8 when selecting and comparing the in-focus and out of focus image regions, greatly improving the results of our kernel estimation algorithm. We also use them in step 9 to spatially restrict segment masks once they have been blurred in order to prevent haloing artefacts during the propagation of synthesised segments throughout the focal stack.

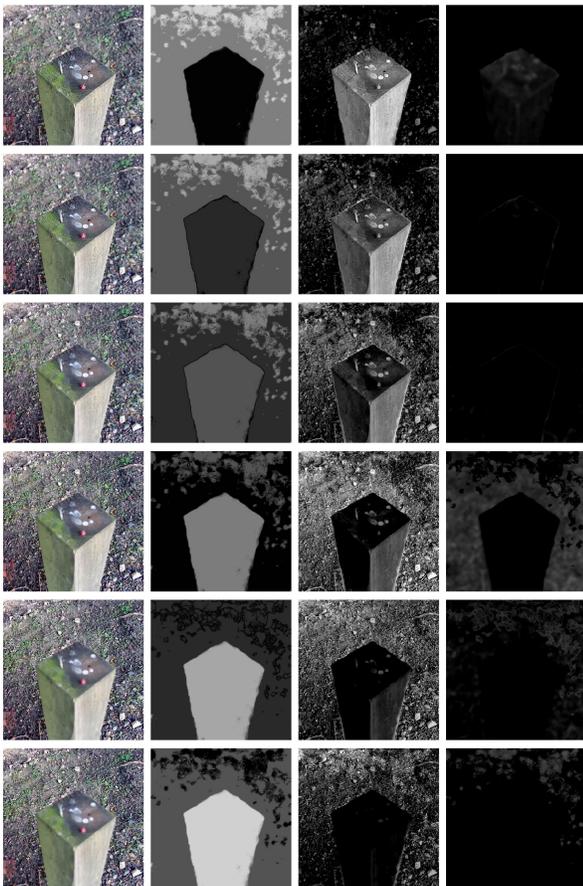


Figure 2: Relative depth maps and their application in pre-processing images to enhance feature extraction. (left to right) Focal stack images, relative depth maps, grayscale focal stack image subtracted by relative depth map, laplacian transformation of focal stack images subtracted by relative depth map.

3 Limitations and Future Work

The texture synthesis algorithm [Efros and Leung 1999] used for synthesising the image segment has serious limitations; the algorithm does not take into account larger shapes and structures in the image. Performing edge completion to find the most likely smooth curves when completing the depth map and all-in-focus image could rectify this.

Although using an isotropic Gaussian blur kernel to artificially de-focus synthesised image segments does not cause any perceptible artefacts, a more accurate method would be to estimate the spatially varying point-spread function and use that to de-focus the synthesised segments.

The focal stack sequence only covers a 3D subset of the full 4D light field, converting focal stacks into light fields leads to imperfect reconstruction at occlusion boundaries and reduced spatial resolution [Levin and Durand 2010]. In future we want to explore light field completion algorithms that are performed directly in the 4D ray-space.

References

- EFROS, A., LEUNG, T. K., ET AL. 1999. Texture synthesis by non-parametric sampling. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 2, IEEE, 1033–1038.
- LEVIN, A., AND DURAND, F. 2010. Linear view synthesis using a dimensionality gap light field prior. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, IEEE, 1831–1838.
- YATZIV, L., SAPIRO, G., AND LEVOY, M. 2004. Lightfield completion. In *Image Processing, 2004. ICIP'04. 2004 International Conference on*, vol. 3, IEEE, 1787–1790.