Depth from focus: an application for fabrics captured at microscale

David Pascual[1] Elena Garces[1,2] Carlos Aliaga[1] Jorge Lopez-Moreno[1]

Seddi Labs[1] Universidad Rey Juan Carlos[2]

Abstract

Estimating depth from a single image is a very hard problem and naturally ill-posed as geometry and illumination are integrated into the single rgb pixels. RGB-D cameras have proved to be suitable for this task in coarse environments, where the depth resolution does not need to be extremely high. However, when the size of the scene is micrometric, the available techniques to infer depth with sensory data reduce considerably and become much more expensive. In this paper, we propose an optical system suitable for estimating depth at yarn-level resolution from focal stack pictures. We evaluate a graph-based depth from defocus algorithm, and compare the depth precision with a high-end depth camera.

1. Introduction

Depth estimation is by itself a very hard problem due to the complexity of the real world. A number of methods exist that estimate material and normal maps from pictures of planar materials at millimeter scale. They commonly rely on multiple views and/or illumination directions to estimate normal values and use inverse rendering techniques to further estimate the material parameters that allow the material to be reproduced.

Depth estimation at microscale is an even more difficult problem as current technologies, like RGB-D cameras, do not reach such a high precision and, even if so, they are highly expensive (e.g. Gocator 3504). Very few methods exist that have developed techniques for depth estimation at such scale. Nam et al. [6] propose a camera system with a macro lens which obtains very short optical working distances reducing the field of view to 1.5×1.5 cm. Their system uses multiple LEDs and a camera with fixed viewing direction. Our optical system [1], similar to theirs, allow us to take captures up to 1.8 μm per pixel.

Yarn fibers, like hair, are very complex structures where scattering effects dominate the final appearance [4], thus, for properly reproducing them, a really high resolution in depth is necessary [2]. Our low cost optical system is able to capture a field of view of 8.9×6.7 mm with a pixel resolution of 4912x3684 px and a depth of field of approximately 0.5 mm with our maximum aperture (f/2.8). As fabrics have typically between 0.2-1.5mm of thickness, we found that depth from focus techniques turn out to be suitable for this task. In particular, we compare raw depth data from a high-resolution depth scanner, and a graph-based depth from focus technique similar to Suwajanakorn et al. [8].

2. Related Work

The amount of methods that estimate depth from focus images are numerous. One of the most recent one [3] leverages deep learning techniques to train a neural network with focal stacks obtained with a light field camera. The approach, although with huge potential given the success in neural architectures, is difficult to train with real data at our depth scale. Other approaches rely on total variation regularization. The work of Moeller et al. [5] combines smooth with non-smooth terms within a ADMM optimization framework. The results with this method tend to be oversmooth and some details are lost at the regularization step. On the contrary, it is a very fast method which can be executed in parallel architectures. The work of Suwajanakorn et al. [8] presents a different approach where the use of optical flow enables hand-held mobile scenarios where the optical system parameters are unknown. We rely on a simple implementation of this method and evaluate the two representations of depth that it provides. First, a fine depth estimation is given per pixel, then, a posterior graph-based optimization procedure provides a smoother depth map. The estimation of a sharpness metric is also an open-problem. We refer the reader to the overview provided in Pertuz et al. [7]. All these methods have been evaluated in regular scale images. Our optical system, however, captures data at microscale, making the system highly sensible to small changes in parameters or setup. A closely related paper to ours is the work of Nam et al. [6], which propose a capture system that allows the estimation of both normal maps and material parameters of images at microscale.

3. Optical Capture System

Our optical capture system, as illustrated in Figure 1, is composed of an RGB camera of 4912x3684 pixel resolution, and a 50 mmm focal length with an extension tube of 25 mm which provide 0.5 magnification at a distance of 100 mm. Additionally, focus is controlled by manual a rail with micrometric precision to move away/closer the lens from the holder where the fabric is placed. This setup allows us to obtain a field of view of 8.9×6.7mm and 1.8μm of pixel size.

As our setup is not standard, the depth of field (DoF), defined as the distance between the nearest and the furthest objects that are in acceptably sharp focus in an image, needs to be measured separately. To this end, we use an empirical approach using a grid of circular dots of known size and a measure of sharpness (described below). We obtain an approximately value of 0.5 mm for aperture of $f/2.8$.

Sharpness measure A critical step in depth from focus techniques is the choice of the sharpness metric. Following previous work [7, 8], we define it as the sum of $\exp|\nabla I(x,y)|$ over a Gaussian patch with zero standard deviation and three pixels of window size around the pixel $I(x,y)$

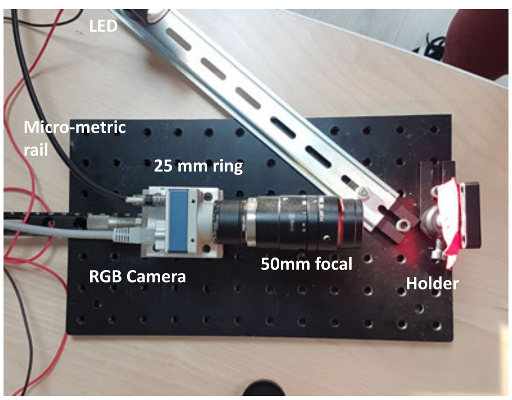

Figure 1. Optical camera setup.

4. Depth from Focus Method

We follow an approach similar to the work of Suwajanakorn et al. [8], but simplified for our controlled setup [1]. First, we build an *all-in-focus image* by means of graph cut optimization. Then, the *camera and scene parameters* are optimized. Finally, the *final depth* is obtained by another graph-based optimization.

Alignment and All-in-focus image In the first step the images need to be aligned. To this end, we estimate rigid transformations and warp all the images to the first image of the stack. Then, we compute a sharpness value for each pixel of each image of the focal stack after alignment. Taking these values as unary term, we build a graph cut optimization to find the all-in-focus image. The pairwise term will be a L1 smoothness term between depths of adjacent pixels. This all-in-focus image can additionally be visualized in the form of a coarse theoretical depth map s, which can be computed by taking the working distance and the focus step of the optical system.

[1] https://github.com/cgearhart/DepthFromDefocus

Camera parameters Before the final smoothing operation, the values for camera parameters (aperture A, focal length F, and focal increment between images f_i) are optimized taking the initial depth map s obtained in the previous step. Assuming a disk shaped point spread function (PSF), the relationship between these parameters and the PSF radius r_i is defined as:

$$r_i = A\frac{|f_i - s|}{s}\frac{F}{f_i - F} \tag{1}$$

We generate a stack of blurred versions $\{\hat{I}_i\}$ of the all-in-focus image using gaussian kernels of increasing radius, simulating the effect of the PSF in the images. We then optimize for the camera parameters by minimizing the difference between the real images and the blurred versions of the all-in-focus image.

$$\min_{A,F,f_i}\sum_{i=1}^{n}||I_i - \hat{I}_i(A,F,f_i)||^2 \tag{2}$$

Final depth image In a final step, we build another graph cut optimization problem where the goal is to interpolate depth values to obtain a smooth image with real depths. We use as unary term the difference between the PSF radius estimation for each pixel and the theoretical one, given by our optimized camera parameters. Once again, the pairwise term is a measure of L1 smoothness between adjacent pixels.

5. Results

Figure 3 shows the results of the main steps of the process. First column shows the image from the stack with the best focus value, second column shows the initial depth values after the all-in-focus step, and third column shows the final depth after the final regularization step. We can see from the images that the quality of the reconstruction is fairly reasonable for thick fabrics whereas for thin fabrics the main effect that we can observe is their misalignment with respect to the camera plane. This is quite interesting and can be used to detect wrong fabric placement in industrial setups.

In Figure 2 we provide a depth map obtained with the high-end depth camera Gocator 3504. In comparison, our optical system provides much more depth resolution than such camera.

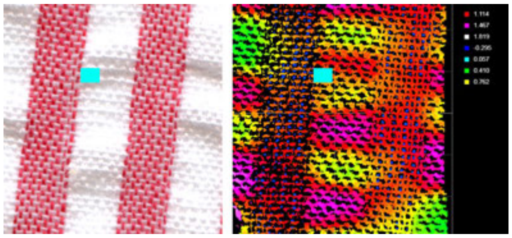

Figure 2. Depth Estimated with the camera Gocator 3504, a high-end camera that captures depth at high resolution. The rectangle in blue corresponds to the same field of view that our capture system is able to obtain.

6. Discussion

In this paper we have seen a practical implementation of depth from defocus techniques for capturing depth of fabrics at microscale. We have shown that our optical system allows us to estimate a depth map of fabrics of a range of thickness between 0.2mm and 1.5mm. As expected, the thicker the fabric, the better the estimated depth. However, even if the fabric is very thin we can still make use of the estimated depth to detect fabric misalignment.

Acknowledgments Elena Garces was supported by a Juan de la Cierva Fellowship from the Spanish Ministry of Science and Technology

References

[1] Raul Alcain, Carlos Heras, Iñigo Salinas, Jorge Lopez-Moreno, and Carlos Aliaga. Microscale optical capture system for digital fabric recreation. In *7th International Conference on Photonics, Optics and Laser Technology*, pages 114–119, 01 2019.

[2] Carlos Aliaga, Carlos Castillo, Diego Gutierrez, Miguel A. Otaduy, Jorge Lopez-Moreno, and Adrian Jarabo. An appearance model for textile fibers. *Computer Graphics Forum (Proc. of the Eurographics Symposium on Rendering)*, 36(4), 2017.

[3] C. Hazirbas, S. G. Soyer, M. C. Staab, L. Leal-Taix, and D. Cremers. Deep depth from focus. In *Asian Conference on Computer Vision (ACCV)*, December 2018.

[4] Stephen R Marschner, Henrik Wann Jensen, Mike Cammarano, Steve Worley, and Pat Hanrahan. Light scattering from human hair fibers. In *ACM Transactions on Graphics (TOG)*, volume 22, pages 780–791. ACM, 2003.

[5] Michael Moeller, Martin Benning, Carola Schönlieb, and Daniel Cremers. Variational depth from focus reconstruction. *IEEE Transactions on Image Processing*, 24(12):5369–5378, 2015.

[6] Giljoo Nam, Joo Ho Lee, Hongzhi Wu, Diego Gutierrez, and Min H. Kim. Simultaneous acquisition of microscale reflectance and normals. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia 2016)*, 35(6):185:1–185:11, 2016.

[7] Said Pertuz, Domenec Puig, and Miguel Angel Garcia. Analysis of focus measure operators for shape-from-focus. *Pattern Recogn.*, 46(5):1415–1432, May 2013.

[8] Supasorn Suwajanakorn, Carlos Hernandez, and Steven M Seitz. Depth from focus with your mobile phone. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3497–3506, 2015.

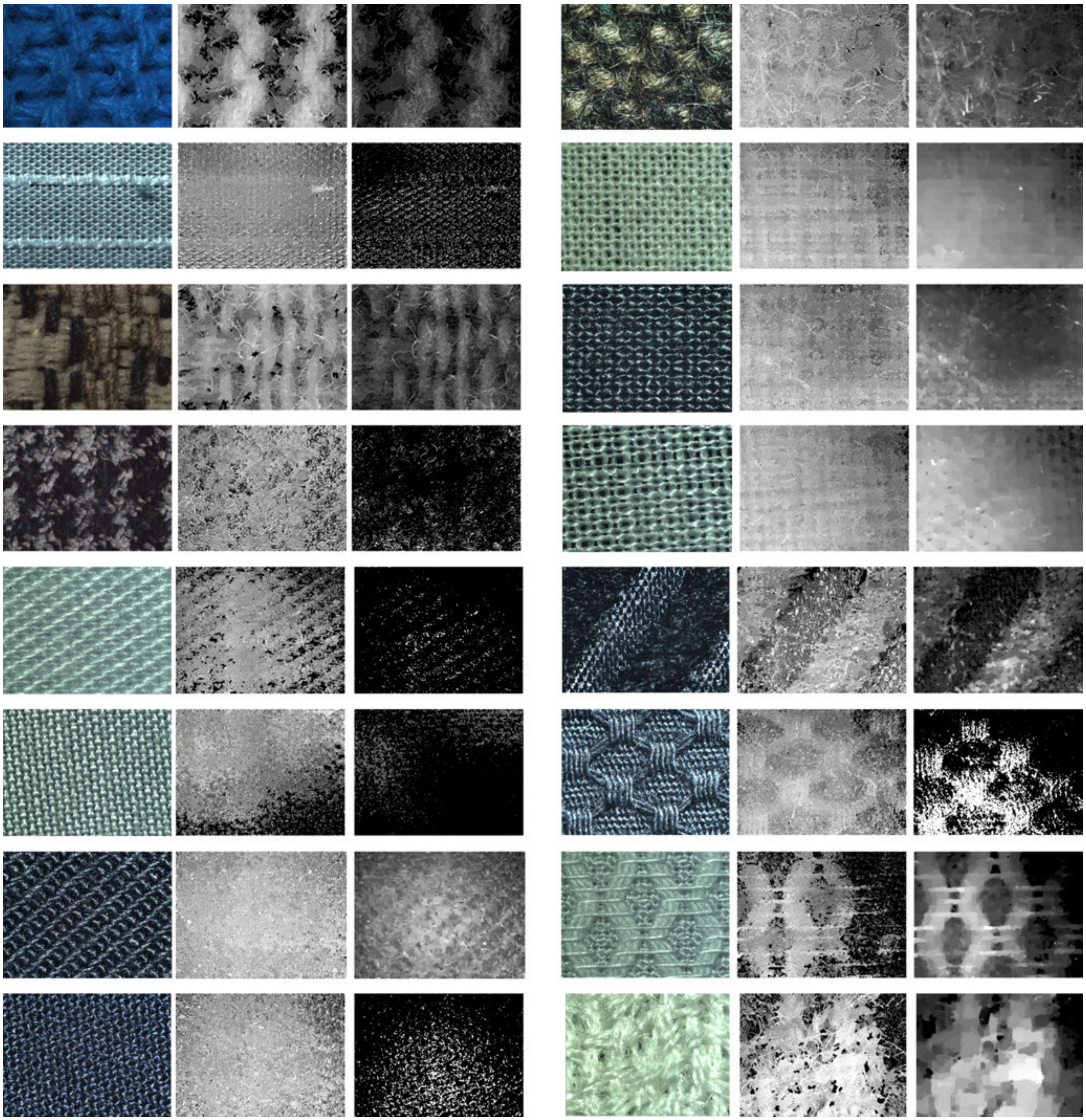

Figure 3. Results of depth values for a set of captured fabrics of different thickness. For each fabric, we show the best focused image from the stack, the initial raw depth estimation after all-in-focus, and the final depth after regularization. All the images cover the same field-of-view.