Phase-coded computational imaging for depth of field extension

Erdem Sahin, Ugur Akpinar, Atanas Gotchev

Faculty of Information Technology and Communication Sciences, Tampere University, Korkeakoulunkatu 1, 33720, Tampere, Finland

erdem.sahin@tuni.fi

Abstract: We present a computational imaging approach, combining a phase-coded computational camera with a corresponding CNN-based deblurring network, that enables extended depth of field images. The simulations demonstrate promising results achieving significant depth of field extension.

OCIS codes: 110.1758, 110.7348, 100.1830

1. Introduction

Usually, the conventional imaging systems have to sacrifice from light efficieny for increased depth of field (DoF). This in turn reduces the signal to noise ratio of the captured image. The research on DoF extension has been mostly concentrated on manipulating the point spread function (PSF) of the imaging system. For instance, in [1] the DoF is increased by inserting a cubic phase mask at the aperture. More recent works [2,3] optimize the phase mask and the corresponding deblurring algorithm at the same time based on an end-to-end model, which demonstrate more promising results. In [2], the phase mask is restricted to be a single-ring pattern and the deblurring algorithm is implemented via a convolutional neural network (CNN). In [3], the phase mask is optimized within a much larger signal space and the deblurring is accomplished via Wiener deconvolution.

In this work, we also propose an end-to-end optimization framework, which is based on a phase-coded imaging network and a CNN-based deblurring network. For a given camera, we explicitly define the bandwidth of the search space for the phase mask depending on the desired depth range for DoF extension, which significantly eases the training process and is critical for the convergence of the end-to-end network.

2. Method

The proposed approach for DoF extension is illustrated in Fig. 1. The end-to-end model consisting of phase-coded imaging and CNN-based deblurring networks is fully differentiable, which ensures optimization of (i.e. learning) phase mask Φ together with the CNN weights via backpropagation (shown in red). The first five convolution layers

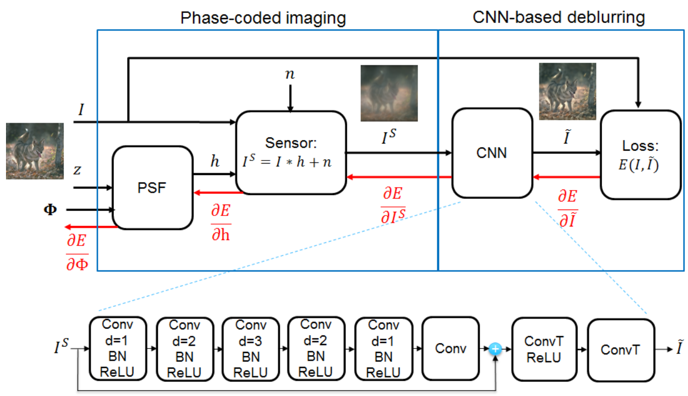

Fig. 1. End-to-end phase-coded imaging and deblurring model.

have 32 kernels with size 3×3. The sixth convolution layer has 3 kernels with size 3×3. The resulting residual image is added to the original image, and the output is driven to a simple deconvolution network, i.e., two layers of ConvT, each of which performing the operation $x^{n+1} = H^T x^n$ with H being the convolution matrix. Finally, the utilized loss function is chosen to be sum of L_1 loss and sparse image gradient [4] as

$$E(I,\tilde{I}) = |I-\tilde{I}|_1 + \alpha \sum_{x,y} \exp\left[-10\left(|\nabla I_x|^{0.8} + |\nabla I_y|^{0.8}\right)\right]\left(|\nabla \tilde{I}_x|^{0.8} + |\nabla \tilde{I}_y|^{0.8}\right) \tag{1}$$

Our phase-coded extended DoF (EDoF) camera, modeled via the imaging network in Fig. 1, includes a phase mask placed on top of a (or at the aperture position of an equivalent) thin lens that has focal length f and images the depth plane z_f onto the sensor, i.e, $1/f = 1/z_f + 1/z_s$ with z_s being the distance between the lens and the sensor. For an object point at z and nominal lens focal length f_0 at wavelength λ_0, the *generalized pupil function* is given as

$$P_{\lambda,z}(\xi,\eta) = A(\xi,\eta)\exp\left\{j\Phi_{\lambda_0}(\xi,\eta)\left[\frac{\lambda_0(n_\lambda - 1)}{\lambda(n_{\lambda_0}-1)}\right]\right\}\exp\left[j\Psi_{\lambda,z}\left(\frac{\xi^2}{r^2}+\frac{\eta^2}{r^2}\right)\right], \tag{2}$$

where $A(\xi,\eta)$ is the (binary) pupil function of the lens, $\Phi_{\lambda_0}(\xi,\eta)$ is the phase pattern that corresponds to the phase delay introduced by the mask at the nominal wavelength λ_0, n_λ is the refractive index of the material that is assumed for both thin lens and phase mask, r is the radius of the circular aperture; and $\Psi_{\lambda,z}$ is the defocus coefficient that is defined as

$$\Psi_{\lambda,z} = \frac{\pi}{\lambda}\left(\frac{1}{z}+\frac{1}{z_s}-\frac{n_\lambda - 1}{f_{\lambda_0}(n_{\lambda_0}-1)}\right)r^2. \tag{3}$$

The PSFs $h_{\lambda,z}(x,y)$ for different color channels are found by taking the magnitude-squared of the Fourier transform of the generalized pupil function, given by Eq. 2, at the corresponding wavelengths. During training, the mask is discretized so that the corresponding sampling rates, $(1/\Delta_\xi, 1/\Delta_\eta)$, are twice the one required for properly sampling the defocus term (within the extent of lens aperture). The following equation provides a good approximation for this purpose:

$$\frac{1}{\Delta_\xi} = \frac{1}{\Delta_\eta} \approx \frac{4}{\pi}\left|\frac{\partial\left(\Psi_{\lambda_{min},z}\frac{\xi^2}{r^2}\right)}{\partial\xi}\right|_{\xi=r} \tag{4}$$

3. Results

For comparison purpose, we simulate a camera with similar parameters to [3], i.e. $r = 2.5mm$ and $z_s = 35.5mm$. The (convex) thin lens is focused at $2m$ for $\lambda_0 = 628nm$, and the phase mask is optimized within $\Psi_z \approx [-49, 49]$, corresponding to the depth range of $[0.5m, \infty]$.

The network is trained by using 180×180 patches from data set [5]. The standard deviation of the sensor noise used in training is chosen from uniform distribution $U(0.001, 0.01)$, where the image intensity is normalized to 1.

Top row of Fig. 2 illustrates the optimized phase mask and the PSFs for various depths as well as wavelengths, shown in red ($628nm$), green ($537nm$) and blue ($447nm$). The mask, in a sense, minimizes the dependencies of PSF to depth and color. Bottom row demonstrates an example result of DoF extension in comparison with existing methods, which demonstrates superior performance. The planar object is assumed at $z = 0.5m$ and the sensor noise level is set to $\sigma = 0.002$. The image on the second column is taken from [3].

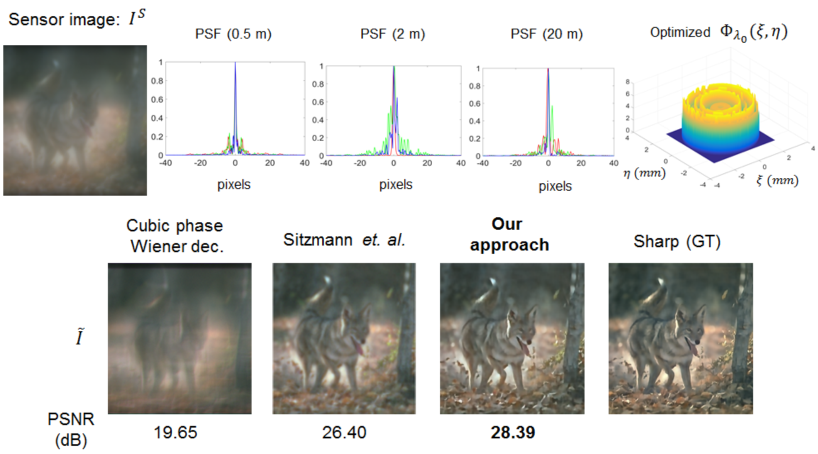

Fig. 2. The simulation results of DoF extension for an object point at $z = 0.5m$.

References

1. E. R. Dowski and W. T. Cathey, “Extended depth of field through wave-front coding,” Appl. Opt. **34**, 1859–1866 (1995).
2. S. Elmalem, R. Giryes, and E. Marom, “Learned phase coded aperture for the benefit of depth of field extension,” Optics Express **26**, 15316–15331 (2018).
3. V. Sitzmann, S. Diamond, Y. Peng, X. Dun, S. Boyd, W. Heidrich, F. Heide, and G. Wetzstein, “End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging,” ACM SIGGRAPH (2018).
4. A. Levin, R. Fergus, F. Durand, and W. T. Freeman, “Image and depth from a conventional camera with a coded aperture,” ACM transactions on graphics (TOG) **26**, 70 (2007).
5. P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, “Contour detection and hierarchical image segmentation,” IEEE Trans. Pattern Anal. Mach. Intell. **33**, 898–916 (2011).