

Liberté Égalité Fraternité



End-to-end sensor design: from Co-design to Deep-Codesign



P. Trouvé-Peloux¹, J-B. Volatier², A. Halé¹, R. Leroy¹, F. Champagnat¹, G. Le Besnerais¹ and G.Druart² 1. ONERA/DTIS-IVA 2. ONERA/DOTA-ERIO pauline.trouve@onera.fr

1 Introduction

- 2 Co-design examples using performance models
- 3 "Deep co-design"
- 4 On going work at ONERA



1 Introduction

- 2 Co-design examples using performance models
- 3 "Deep co-design"
- 4 On going work at ONERA









Joint design



Introduction Main applications in photography



 \Longrightarrow A "nice image" is no longer required at the lens output !







1 Introduction

2 Co-design examples using performance models

3 "Deep co-design"

4 On going work at ONERA



2 Co-design examples using performance modelsDepth of field extension

- Monocular depth estimation
- Monocular RGB-D camera
- Conclusion





 \Longrightarrow EDOF: extension of the DoF region



Principle A phase mask to reduce PSF variation vs depth [1]



Without the phase mask:

$$h_{\psi}(x_{P}, y_{P}) \propto \left| \mathsf{FT}\left(A(x, y) \exp^{i\psi \frac{(x^{2}+y^{2})}{R^{2}}}\right) \right|_{\left(\frac{x_{P}}{\lambda z_{i}}, \frac{y_{P}}{\lambda z_{i}}\right)}^{2} \text{ with } \psi = \frac{\pi R^{2}}{\lambda} \left(\frac{1}{z} + \frac{1}{z_{i}} - \frac{1}{f}\right)$$

With the phase mask :

$$h_{\psi}(x_P, y_P) \propto \left| \mathsf{FT}\left(A(x, y) \exp^{i\psi \frac{(x^2+y^2)}{R^2}} \exp^{i\phi_{\mathsf{mask}}(x, y)} \right) \right|_{\left(\frac{X_P}{\lambda z_i}, \frac{Y_P}{\lambda z_i}\right)}^2$$

[1] Dowski, E. R. and Cathey, W. T. Appl. Opt., 1995, 34



Example of phase mask

- Cubic phase plate [1, 2]
 - $\phi_{mask}(x,y) = \alpha(x^3 + y^3)$

PSF at various depths (W = $\frac{\psi}{2\pi}$):





 \implies Image restoration is required !

[1] Dowski, E. R. and Cathey, W. T. Appl. Opt., 1995, **34** [2] F. Diaz et al., Opt. Lett., (2009) **34**



Image formation model : y = h * x + n

Scene and noise PSD : PSD_x and PSD_n

Restoration filter for a discrete set of depths [2]

A unique deconvolution filter for all the depths (mean Wiener filter) :

$$\tilde{d}^{\alpha}(\mu,\nu) = \frac{\frac{1}{K}\sum_{k=1}^{K}\tilde{h}^{\alpha}_{\psi_{k}}(\mu,\nu)\mathsf{PSD}_{x}(\mu,\nu)}{\frac{1}{K}\sum_{k=1}^{K}|\tilde{h}^{\alpha}_{\psi_{k}}(\mu,\nu)|^{2}\mathsf{PSD}_{x}(\mu,\nu) + \mathsf{PSD}_{n}(\mu,\nu)}$$

Performance model of image restoration at depth k

$$\mathsf{MSE}(\alpha,\psi_k) = \int_{\mu} |\tilde{d}^{\alpha}\tilde{h}^{\alpha}_{\psi_k} - 1|^2\mathsf{PSD}_x(\mu)d\mu + \int_{\mu}\mathsf{PSD}_n|\tilde{d}^{\alpha}|^2(\mu)d\mu$$

Expression in terms of image quality

$$IQ(\alpha, \psi_k) = \frac{E_x}{MSE(\alpha, \psi_k)}$$
 with E_x the energy of the scene



Design criterion for a discrete set of depths $\{k\}$



 \implies Trade off between strictly invariant PSF and image quality



Co-design of a phase mask for EDOF

Deconvolution results







Co-design of phase mask for EDOF

Other phase masks

Annular phase mask - further works at IOGS





- Optimization of the number of rings [3]
- Application to 3D particle microscopy [4]
- Experimental validation and application to panchromatic illumination [5]
- [3] R. Falcon et al., Imaging and Applied Optics 2016
- [4] O. Leveque et al., Opt. Express, 2020, 28
- [5] A. Fontbonne et al., SPIE, 2021, 60.

2 Co-design examples using performance models

- Depth of field extension
- Monocular depth estimation
- Monocular RGB-D camera
- Conclusion



Monocular 3D camera using Depth from Defocus



- Depth can be estimated from local defocus blur estimation
- Compact 3D camera but tricky processing
- How to modify the optics to improve defocus blur estimation ?



Coded aperture (2007)[6] for a better defocus blur discriminiation





CLIM Workshop P.Trouvé-Peloux 30th September 2021

Image formation model y = h * x + n

Scene gaussian model

$$p(\mathbf{x}) \propto \exp{-\frac{\alpha}{2} \left(|F_v \mathbf{x}|^2 + |F_h \mathbf{x}|^2 \right)}$$

 F_h and F_v convolution matrices relative to image gradients Noise gaussian model : White gaussian noise of variance σ_n^2 Image likelihood in Fourier domain for depth k

$$p_k(\tilde{y}) \propto \exp{-\frac{1}{2} \sum_{\mu,\nu} \frac{|\tilde{y}(\mu,\nu)|^2}{\sigma_k(\mu,\nu)}}$$
$$\sigma_k(\mu,\nu) = |\tilde{h}_k(\mu,\nu)|^2 \left(\alpha |\tilde{F}_1(\mu,\nu)|^2 + \alpha |\tilde{F}_2(\mu,\nu)|^2\right)^{-1} + \sigma_n^2$$



Principle : Maximisation of Kullback-Leibler divergence between depths

$$D_{\mathcal{K}L}(p_{k_1}||p_{k_2}) = \sum_{\mu,
u} \left(rac{\sigma_{k_1}(\mu,
u)}{\sigma_{k_2}(\mu,
u)} - \lograc{\sigma_{k_1}(\mu,
u)}{\sigma_{k_2}(\mu,
u)}
ight)$$





Coded aperture Depth estimation results







2 Co-design examples using performance models

- Depth of field extension
- Monocular depth estimation
- Monocular RGB-D camera
- Conclusion





Defocus blur variation for a conventional lens

Depth



Defocus blur variation for a conventional lens

Depth Same blur level for two depths







Depth

Same blur level for two depths No blur variation within the camera depth of field







Depth

Same blur level for two depths No blur variation within the camera depth of field

But chromatic aberration can overcome these issues !



Variation of the focal length with the wavelength



Extracted from Guichard et al., SPIE (2009) [7]





Defocus blur variation for a lens with chromatic aberration



Depth

- No dead zone
- No depth ambiguity
- Reduces image quality







Two outputs lead to two performance models for co-design

[8] P. Trouvé et al., IEEE CVPR Workshops on Comp. Cameras and Displays, (2013)
[9] P. Trouvé-Peloux et al., J. Opt. Soc. Am. A 38(10) (2021)



Co-design of a monocular RGB-D camera *Performance model for DFD for a conventional lens*

Image formation model Gaussian scene model

ONERA

$$\mathbf{y} = H(z)\mathbf{x} + \mathbf{n}$$

 $p(\mathbf{x}) \propto \exp{-\frac{||D\mathbf{x}|}{z}}$

D concatenation of convolution matrices relative to image gradients (h and v) **Noise gaussian model** : White gaussian noise of variance σ_n^2 **Image likelihood**

$$p(\mathbf{y}; z) = \left| \frac{Q(z)}{2\pi} \right|_{+}^{1/2} \exp{-\frac{1}{2} \mathbf{y}^t Q(z) \mathbf{y}}$$

With $Q(z) = \frac{1}{\sigma_b^2} (I - H(z)(H^t(z)H(z) + \alpha D^t D)^{-1}H^t(z))$ and $\alpha = \sigma_b^2/\sigma_x^2$

Performance model Cramér Rao Lower Bound

$$\sigma_{CRB}^{2}(z) = \left(\frac{1}{2} \operatorname{tr}\left(Q^{+}(z)\frac{\partial Q}{\partial z}Q^{+}(z)\frac{\partial Q}{\partial z}\right)\right)^{-1}$$

 $|A|_+$ product of non zeros eigenvalues of matrix A and A^+ pseudo-inverse of A

Co-design of a monocular RGB-D camera

Performance model for DFD for a conventional lens



- Theoretical depth estimation accuracy for a given depth
- With respect to SNR value
- With respect to a given lens

Principle: use a Gaussian scene model in the luminance/chrominance decomposition





Co-design of a monocular RGB-D camera *Framework*



Two outputs lead to two performance model for co-design

```
    [8] P. Trouvé et al., IEEE CVPR Workshops, (2013)
    [9] P. Trouvé-Peloux et al., J. Opt. Soc. Am. A 38(10) (2021)
```



Principle

- With chromatic aberration there is a "sharpest" channel
- High frequencies transfer to the blurred channels [7]
- Use of the depth map to guide the transfer [9]

$$y_c = \alpha(z) HF_R + \beta(z) HF_G + \gamma(z) HF_B$$

Definition of the Generalized Depth of Field (GDOF) :





Application to small UAV navigation

- Depth range 1 to 5 m
- Field of view 25°
- Spatial depth map resolution 2cm at 3 m
- Accuracy of few cm








Several parameters are fixed from the settings





Principle

- Simulation of optical systems with geometrical optics (Gaussian PSF)
- Performance comparison of various systems







Architecture design on Zemax



- Design of a first architecture using Zemax
- Fine tuning using performance models and PSF from Zemax





Co-design of a monocular RGB-D camera *Experimental results*



Depth estimation results

Image restoration results





2 Co-design examples using performance models

- Depth of field extension
- Monocular depth estimation
- Monocular RGB-D camera
- Conclusion



States of the art of co-design using performance models

Common approach

- Choice of an (un)conventional optical model (Fourier optics, geometrical optics, optical design software...)
- Choice of a processing model
- Definition of a performance model (CRB, RMSE, KL divergence...)
- Joint optimization of the parameters
- Experimental validation

Performance model enables the analysis of the interaction between lens properties and processing



States of the art of ("old school") co-design using performance models



How to make co-design with a neural network ?



1 Introduction

2 Co-design examples using performance models

3 "Deep co-design"

4 On going work at ONERA





- State of the art (not exhaustive)
- Principle







Several fields of application :

- Depth of field extension [10, 11, 12]
- Lateral field of view extension [13]
- **3D** [14, 15, 16, 17, 18]
- Object classification [19, 15]
- HDR [20]
- Lensless imaging [21]

Unconventional lens :

- Phase mask [10, 11, 12, 14, 16]
- Coded aperture [21]
- Freeform lens [15]
- \Rightarrow How to conduct joint lens and network optimization ?

3 "Deep co-design"

- State of the art (not exhaustive)
- Principle



Image formation model $y = h(\phi) * x + \eta$

 \Rightarrow A lens can be modeled using a convolutional layer of a neural network



Figure extracted from Metzler et al, CVPR 2020 [20]

End-to-end optimization can be conducted as long as the gradient of h (i.e. the PSF) can be defined with respect to ϕ

A more detailed example : the presentation of E. Sahim yesterday !



1 Introduction

- 2 Co-design examples using performance models
- 3 "Deep co-design"
- 4 On going work at ONERA



4 On going work at ONERA

- Our approach
- Application to EDOF
- Conclusions and perspectives



State of the art papers

- Optical model based on Fourier Optics (thin lens and paraxial rays approx.)
- Optimization of a single optical element (phase mask, single lens...)

– Our approach

Optical model based on differential ray tracing (DRT) [22]

- No thin lens nor paraxial approximation
- Optimization possible of the full set of lens parameters θ_{opt}
- Valid for any field of view and point source depth (α, z)

[22] Halé et al., Optics. Express (to appear), 2021







4 On going work at ONERA

- Our approach
- Application to EDOF
- Conclusions and perspectives



Application to EDOF *Context*

Add-on to add chromatic aberration to an existing conventional lens [23]





Settings

- Add-on in front of a Double Gauss lens
- A state of the art restoration neural network [24]
- Texture database [25] (patch 64x64)
- **Single** degree of freedom : the sensor position (RGB focus)
- [23] Trouvé-Peloux et al. Applied Optics, 57, 2018.
- [24] Mao et al. Advances in neural information processing systems, 2016.
- [25] Cimpoi et al., IEEE CVPR, 2014.



EDOF problem : Joint optimization of the sensor position and the restoration network within a depth range [20 to 50 m]



 \Rightarrow Validation of the DRT model and the network joint optimization \Rightarrow The focus is automatically adapted to the depth range of interest



	Without	sensor position optimization	With sens	or position optimization
Object position	RMSE	MAE	RMSE	MAE
10m	0.1032	0.0667	0.1169	0.0774
15m	0.0503	0.0317	0.0841	0.0530
20m	0.0503	0.0320	0.0468	0.0298
25m	0.0678	0.0430	0.0490	0.0310
30m	0.0786	0.0499	0.0488	0.0312
35m	0.0833	0.0528	0.0478	0.0298
40m	0.0876	0.0555	0.0482	0.0309
45m	0.0919	0.0586	0.0563	0.0356
50m	0.0955	0.0614	0.0697	0.0392
Average	0.0787	0.0502	0.0631	0.0398

Table: RMSE stands for root mean square error and MAE for mean absolute error.



Application Restoration qualitative results



 \Rightarrow Restoration quality is visible on image patch



4 On going work at ONERA

- Our approach
- Application to EDOF
- Conclusions and perspectives



Conclusion

- Validation the DRT model and neural network joint optimization
- Application in a single parameter optimisation for EDOF

Parallel work

Use of differential ray tracing in co-design of complex optical system and neural network [26]

Perspectives

- Optimization of several optical parameters
- Variation of the PSF off axis
- Application to 3D : see R. Leroy's talk !

[26] Sun et al, ACM SIGGRAPH, 2021.



Open questions

- How do optical and processing parameters interact during the joint optimization ?
- What image quality (i.e. type of aberration) will be acceptable by the network for a given task ?
- How to ensure a physically realistic solution ?
- How important is the optical system starting point of the optimization ?
- Can this approach generate unconventional and original system ?
- How to model the performance ?





Liberté Égalité Fraternité



Thank you for your attention Do you have any question ?

www.onera.fr

References

- E. R. Dowski and W. T. Cathey, Extended depth of field through wave-front coding, Applied optics, 34, pp. 1859–1866 (1995).
- F. Diaz, F. Goudail, B. Loiseaux and J.-P. Huignard, Increase in depth of field taking into account deconvolution by optimization of pupil mask, Opt. Lett., 34 (19), pp. 2970–2972 (Oct 2009), http://ol.osa.org/abstract.cfm?URI=ol-34-19-2970.
- [3] R. Falcón, F. Goudail and C. Kulcsár, How many rings for binary phase masks co-optimized for depth of field extension?, In Imaging and Applied Optics 2016, p. CTh1D.5. Optical Society of America (2016), http:

//www.osapublishing.org/abstract.cfm?URI=COSI-2016-CTh1D.5.

 [4] O. Lévêque, C. Kulcsár, A. Lee, H. Sauer, A. Aleksanyan, P. Bon, L. Cognet and F. Goudail, *Co-designed annular binary phase masks for depth-of-field extension in single-molecule localization microscopy*, Opt. Express, 28 (22), pp. 32426–32446 (Oct 2020), http: //www.opticsexpress.org/abstract.cfm?URI=oe-28-22-32426.

 [5] A. Fontbonne, H. Sauer and F. Goudail, Theoretical and experimental analysis of co-designed binary phase masks for enhancing the depth of field of panchromatic cameras, Optical Engineering, 60 (3), pp. 1 – 20 (2021), https://doi.org/10.1117/1.0E.60.3.033101.

[6] A. Levin, R. Fergus, F. Durand and W. T. Freeman, Image and Depth from a Conventional Camera with a Coded Aperture, In ACM SIGGRAPH 2007 Papers, SIGGRAPH '07, New York, NY, USA, ACM (2007), http://doi.acm.org/10.1145/1275808.1276464. [7] F. Guichard, H. Nguyen Phi, R. Tessières, M. Pyanet, I. Tarchouna and F. Cao,

Extended Depth-of-Field using sharpness transport across color channels,

Proceedings of SPIE - The International Society for Optical Engineering, 7250, p. 72500 (01 2009).

- [8] P. Trouvé, F. Champagnat, G. Le Besnerais, G. Druart and J. Idier, Design of a Chromatic 3D Camera with an End-to-End Performance Model Approach,
 In IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops (June 2013).
- [9] P. Trouvé-Peloux, F. Champagnat, G. L. Besnerais, G. Druart and J. Idier, *Performance model of depth from defocus with an unconventional camera*.
 - J. Opt. Soc. Am. A, 38 (2021).
- S. Elmalem, R. Giryes and E. Marom, Learned phase coded aperture for the benefit of depth of field extension, Opt. Express, 26 (12), pp. 15316–15331 (Jun 2018).

[11] 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 Trans. Graph. (SIGGRAPH) (2018).

- U. Akpinar, E. Sahin and A. Gotchev, Learning Optimal Phase-Coded Aperture for Depth of Field Extension, In 2019 IEEE International Conference on Image Processing (ICIP), pp. 4315–4319 (2019).
- [13] Y. Peng, Q. Sun, X. Dun, G. Wetzstein, W. Heidrich and F. Heide, Learned Large Field-of-View Imaging With Thin-Plate Optics, ACM Trans. Graph. (SIGGRAPH Asia), (6) (2019).
- [14] H. Haim, S. Elmalem, R. Giryes, A. M. Bronstein and E. Marom, Depth Estimation From a Single Image Using Deep Learned Phase Coded Mask, IEEE Transactions on Computational Imaging, 4 (3), pp. 298–310 (2018).

[15] J. Chang and G. Wetzstein,

Deep optics for monocular depth estimation and 3d object detection, In Proceedings of the IEEE International Conference on Computer Vision, pp. 10193–10202 (2019).

[16] Y. Wu, V. Boominathan, H. Chen, A. Sankaranarayanan and A. Veeraraghavan, *PhaseCam3D — Learning Phase Masks for Passive Single View Depth Estimation*, In 2019 IEEE International Conference on Computational Photography (ICCP), pp. 1–12 (2019).

[17] E. Nehme, D. Freedman, R. Gordon, B. Ferdman, L. E. Weiss, O. Alalouf, T. Naor, R. Orange, T. Michaeli and Y. Shechtman, *DeepSTORM3D: dense 3D localization microscopy and PSF design by deep learning*, Nature Methods, 17 (7), p. 734–740 (Jun 2020).

 H. Ikoma, C. M. Nguyen, C. A. Metzler, Y. Peng and G. Wetzstein, Depth from Defocus with Learned Optics for Imaging and Occlusion-aware Depth Estimation, IEEE International Conference on Computational Photography (ICCP) (2021).

- [19] J. Chang, V. Sitzmann, X. Dun, W. Heidrich and G. Wetzstein, *Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification*, Scientific Reports (2018).
- [20] C. Metzler, H. Ikoma, Y. Peng and G. Wetzstein, Deep Optics for Single-shot High-dynamic-range Imaging, In Proc. CVPR (2020).
- [21] R. Horisaki, Y. Okamoto and J. Tanida, *Deeply coded aperture for lensless imaging*, Optics Letters, 45 (11), pp. 3131–3134 (2020).
- [22] A. Halé, P.Trouvé and J.-B. Volatier, *End-to-end design of lens and a neural network using differential ray tracing*, Opt. Express (2021).

 [23] P. Trouvé-Peloux, J. Sabater, A. Bernard-Brunel, F. Champagnat, G. L. Besnerais and T. Avignon, *Turning a conventional camera into a 3D camera with an add-on*, Appl. Opt., 57 (10), pp. 2553–2563 (Apr 2018), http://ao.osa.org/abstract.cfm?URI=ao-57-10-2553.

- [24] X. Mao, C. Shen and Y.-B. Yang, Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections, In Advances in neural information processing systems, pp. 2802–2810 (2016).
- [25] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed and A. Vedaldi, Describing Textures in the Wild, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 3606–3613 (IEEE, 2014).
- [26] Q. Sun, C. Wang, Q. Fu, X. Dun and W. Heidrich, End-to-End Complex Lens Design with Differentiate Ray Tracing, ACM Trans. Graph., 40 (4) (July 2021), https://doi.org/10.1145/3450626.3459674

Evaluate the derivative of the ray propagation with respect to "system" parameters



 $\label{eq:phi} \Psi \mbox{ contains curvatures} \\ \mbox{and positions of the} \\ \mbox{mirrors for instance.} \\$



Sum of optical paths :

$$L = \sum_{i} n_{i,i+1} ||P_i P_{i+1}||$$
(1)

Fermat principle : $\frac{\partial L}{\partial u_i}=0$ and $\frac{\partial L}{\partial v_i}=0$

$$n_{i-1,i}\frac{\partial ||P_{i-1}P_i||}{\partial u_i} + n_{i,i+1}\frac{\partial ||P_iP_{i+1}||}{\partial u_i} = f_{u,i} = 0$$
(2)

(idem for ν_i). For any system parameter Ψ_j :

$$\frac{df_{u,i}}{d\Psi_j} = \frac{\partial f_{u,i}}{\partial \Psi_j} + \sum_{k=0}^N \frac{\partial f_{u,i}}{\partial u_k} \frac{\partial u_k}{\partial \Psi_j} + \sum_{k=0}^N \frac{\partial f_{u,i}}{\partial v_k} \frac{\partial v_k}{\partial \Psi_j} = 0$$
(3)

The derivatives $\frac{\partial f_{u,i}}{\partial u_k}$ and $\frac{\partial f_{u,i}}{\partial \Psi_j}$ can be calculated by automatic differntiation (Julia). Solving this linear system yields : $\frac{\partial u_k}{\partial \Psi_i}$

