ESTIMATING NONPLANAR FLOW FROM 2D MOTION-BLURRED WIDEFIELD MICROSCOPY IMAGES VIA DEEP LEARNING

Adrian Shajkofci[1,2], *Michael Liebling*[1,3]

[1]Idiap Research Institute, CH-1920 Martigny, Switzerland
[2]École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
[3]Electrical and Computer Engineering Department, University of California, Santa Barbara, CA 93106, USA

ABSTRACT

Optical flow is a method aimed at predicting the movement velocity of any pixel in the image and is used in medicine and biology to estimate flow of particles in organs or organelles. However, a precise optical flow measurement requires images taken at high speed and low exposure time, which induces phototoxicity due to the increase in illumination power. We are looking here to estimate the three-dimensional movement vector field of moving out-of-plane particles using normal light conditions and a standard microscope camera. We present a method to predict, from a single textured wide-field microscopy image, the movement of out-of-plane particles using the local characteristics of the motion blur. We estimated the velocity vector field from the local estimation of the blur model parameters using an deep neural network and achieved a prediction with a regression coefficient of 0.92 between the ground truth simulated vector field and the output of the network. This method could enable microscopists to gain insights about the dynamic properties of samples without the need for high-speed cameras or high-intensity light exposure.

***Index Terms*—** Microscopy, optical flow, convolutional neural networks, motion blur

1. INTRODUCTION

Life is all about movement. From the microscale to the macroscale, organisms undergo growth, nutrients flow or diffuse in their environment. Quantification of the displacement in time of particles, organelles, or organisms can be done using optical flow [1]. Optical flow is a method aimed at determining the distribution of apparent velocities of any movement in image series. In medicine, and more specifically in cardiac imaging, optical flow proved to be correlated with the flow patterns measured using computational dynamics [2]. In photography, deep neural networks (DNNs) recently allowed for the prediction of 3D optical flow in a computationally-efficient way and with a good accuracy [3]–[8].

Optical flow is usually computed using two image frames at different time points. In microscopy, the physical scales are orders of magnitude smaller than in photography, especially in the axial direction due to the very small depth of field. Consequently, a small movement of the object can cause massive blurring when imaged. For optical flow to be applied successfully, the two reference images must be taken in a short interval of time. High-speed cameras are still uncommon in microscopy stations and fast movement happening during the exposure time causes motion blur. Furthermore, short exposure times require stronger illumination, accelerating phototoxicity or fluorophore depletion [9].

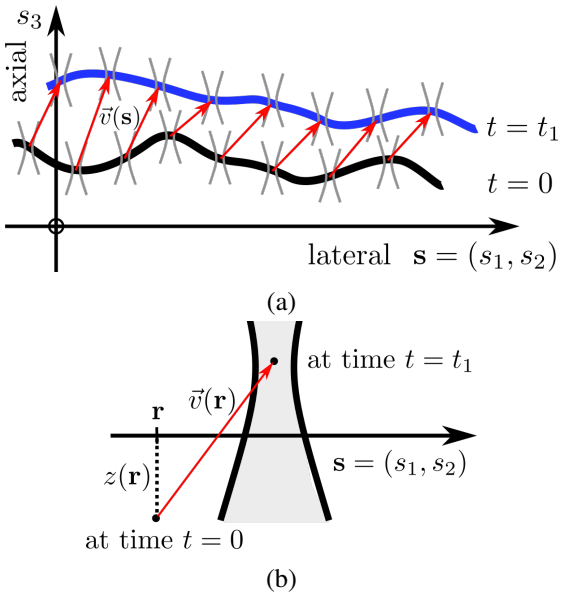


Fig. 1: Image formation model. The object is modeled as a thin moving manifold. (a) global view of the moving manifold. (b) contribution to image from one point on the manifold at time $t = t_1$; motion is encoded in image blur created by cumulated contributions over exposure interval $[0, t_1]$.

Here we present a method for estimating the movement of out-of-plane particles in a fluid, from a single optical microscopy wide-field image with a long exposure time. We take advantage of the motion blur by estimating the parameters of a spatially-variant Point Spread Function (PSF) for every point in the image. Since the PSF has been modeled to take into account the displacement in both axial and lateral directions, we are able to extract from the input image a three-dimensional

vector field of the motion.

This paper is organized as follows. In Section 2, we present the method, comprising the image formation model and the estimation of the displacement vector field. Then, in Section 3, we characterize the performance of the method by firstly simulating arbitrary motion fields in microscopy images, then by simulating a rotational flow of particles in a cylindrical pipe. We then discuss our findings and conclude in Section 4.

2. METHODS

2.1. Problem statement

We consider a single 2D widefield microscopy image $i(\boldsymbol{s})$ of a 3D object $o(\boldsymbol{s}, s_3)$, with $\boldsymbol{s} = (s_1, s_2)$ denoting lateral 2D coordinates and s_3 the axial coordinate. We model the object as a flat 2D manifold in a 3D space. We further denote the local movement of the object, measured for points in the image (focus) plane $s_3 = z_0$, by a three-dimensional displacement-vector field $\vec{v}(\boldsymbol{s}) = (v_1(\boldsymbol{s}) \quad v_2(\boldsymbol{s}) \quad v_3(\boldsymbol{s}))$. We finally define the camera shutter interval Δt, during which a point in the object initially at position $\boldsymbol{s}$ moves to a new position determined by vector $\vec{v}(\boldsymbol{s})\Delta t$. Given only the image $i(\boldsymbol{s})$ we aim to predict the field of 3D vectors $\vec{v}(\boldsymbol{s})$ in the image plane.

2.2. Image formation model

We assume that the imaged object at time $t = 0$ is a thin manifold that can be described as:

$$o(\boldsymbol{s}, s_3, 0) = o_0(\boldsymbol{s})\delta\left(s_3 - z\left(\boldsymbol{s}\right)\right). \tag{1}$$

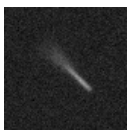
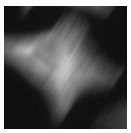
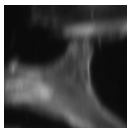
We further assume that the object undergoes a motion of velocity $\vec{v}(\boldsymbol{s}) = (v_1(\boldsymbol{s}) \quad v_2(\boldsymbol{s}) \quad v_3(\boldsymbol{s})) = (\boldsymbol{u}(\boldsymbol{s}) \quad v_3(\boldsymbol{s}))$, with $\boldsymbol{u}(\boldsymbol{s}) = (v_1(\boldsymbol{s}) \quad v_2(\boldsymbol{s}))$. For the image formation at time $t = t_1$, we start by considering each point on the manifold at time $t = 0$, with coordinates $(\boldsymbol{r}, z(\boldsymbol{r}))$ and follow its displacement to position $(\boldsymbol{r} + \boldsymbol{u}(\boldsymbol{r})t_1, z(\boldsymbol{r}) + v_3(\boldsymbol{r})t_1)$ at time $t = t_1$. This point contributes to the image via a PSF h_{3D} centered over it and weighted by the intensity $o_0(\boldsymbol{r})$ at $t = 0$ (see Fig. 1(b)). Consequently, the intensity at a given position $\boldsymbol{s}$ in the image ($s_3 = 0$) at time t_1 is then given by:

$$i(\boldsymbol{s}, t_1) = \int_{\boldsymbol{r}} o_0(\boldsymbol{r}) h_{\mathrm{3D}}(\boldsymbol{s} - \boldsymbol{r} - \boldsymbol{u}(\boldsymbol{r})t_1, -z(\boldsymbol{r}) - v_3 t_1)\mathrm{d}\boldsymbol{r}. \tag{2}$$

For a shutter exposure of duration Δt, we integrate the contributions at each time:

$$\begin{aligned} i(\boldsymbol{s}) &= \int_0^{\Delta t} i(\boldsymbol{s}, t)\mathrm{d}t \qquad (3)\\ &= \int_{\boldsymbol{r}} o_0(\boldsymbol{r}) \underbrace{\int_0^{\Delta t} h_{\mathrm{3D}}(\boldsymbol{s}\text{-}\boldsymbol{r}\text{-}\boldsymbol{u}(\boldsymbol{r})t, \text{-}z(\boldsymbol{r})\text{-}v_3 t)\mathrm{d}t}_{h_{\Delta t}(\boldsymbol{s},\boldsymbol{r})}\mathrm{d}\boldsymbol{r}\\ &= \int_{\boldsymbol{r}} o_0(\boldsymbol{r}) h_{\Delta t}(\boldsymbol{s}, \boldsymbol{r})\mathrm{d}\boldsymbol{r}. \qquad (4) \end{aligned}$$

Table 1: Example of predictions of the velocity vector $\vec{v}(\boldsymbol{s})$. a-c are simulations of beads. e-f are textured images of HeLa cells actin (Alexa Fluor 635) taken with a 10×/0.3 air objective.

	$\vec{v}(\boldsymbol{s})$	$i(\boldsymbol{s})$	$\widetilde{\vec{v}}(\boldsymbol{s})$
a	$v_1 = 0.30$ $v_2 = 0.30$ $v_3 = 0$ $z_0 = 0$		$\widetilde{v}_1 = 0.38$ $\widetilde{v}_2 = 0.31$ $\widetilde{v}_3 = 0.01$ $\widetilde{z}_0 = 0.05$
b	$v_1 = 0.30$ $v_2 = 0.30$ $v_3 = 0.80$ $z_0 = 0$		$\widetilde{v}_1 = 0.35$ $\widetilde{v}_2 = 0.29$ $\widetilde{v}_3 = 0.71$ $\widetilde{z}_0 = 0.08$
c	$v_1 = 0.30$ $v_2 = 0.30$ $v_3 = 0.80$ $z_0 = 0.30$		$\widetilde{v}_1 = 0.35$ $\widetilde{v}_2 = 0.31$ $\widetilde{v}_3 = 0.77$ $\widetilde{z}_0 = 0.41$
d	$v_1 = -0.30$ $v_2 = 0.60$ $v_3 = 0.80$ $z_0 = 0.10$		$\widetilde{v}_1 = -0.28$ $\widetilde{v}_2 = 0.51$ $\widetilde{v}_3 = 0.96$ $\widetilde{z}_0 = 0.15$
e	$v_1 = 0$ $v_2 = 0.40$ $v_3 = 1.20$ $z_0 = 0$		$\widetilde{v}_1 = 0.10$ $\widetilde{v}_2 = 0.35$ $\widetilde{v}_3 = 0.98$ $\widetilde{z}_0 = 0.06$

Eq. (4) reveals a spatially variant 2D PSF $h_{\Delta t}(\boldsymbol{s}, \boldsymbol{r})$, which captures both the local lateral and axial velocities $\boldsymbol{u}(\boldsymbol{r})$ and $v_3(\boldsymbol{r})$ at each location $\boldsymbol{r}$ in the image, as well as the local depth of the manifold $z(\boldsymbol{r})$. This suggests that if the local PSF could be estimated at every location $\boldsymbol{s}$ of an acquired image $i(\boldsymbol{s})$, the local 3D velocity field could be estimated, including out-of-plane motion.

2.3. Estimation of the displacement vector field

To estimate the displacement vector $\vec{v}(\boldsymbol{s})$ from the input image $i(\boldsymbol{s})$ we follow a procedure similar to [10], where we estimated the local PSF in every location of a still image by training a DNN that extracted the PSF parameters. Here, we again chose for h_{3D} a Zernike polynomial-based PSF model [10], [11], which we adapted to take into account linear 3D displacement starting at various depths to match Eq. (4).

We create a training set of $K = 400'000$ images taken from [12] that are blurred by spatially-variable PSFs. To do so, we first define for every k-th image, N non-overlapping 2D masks $m_k^n(\boldsymbol{s})$, with $n = 0, ..., N-1$. Then, we define, for every mask, a PSF $h_{\vec{v}_k^n}$ generated using the parameters $\vec{v}_k$ drawn from an uniform distribution ($[-1, 1]$ for (v_1, v_2) and $[0, 1]$ for v_3) and $z_k(\boldsymbol{s})$, which is the axial position where the object o is in focus. We get the final K training images by multiplying the masked input image by the PSFs in the Fourier

domain:

$$i_k(\mathbf{s}) = \beta b_p \left(\lambda = \mathcal{F}^{-1} \left[\sum_{n=0}^{N-1} \mathcal{F}(h_{\vec{v}_k^n}) \mathcal{F}(m^n * i_k) \right] (\mathbf{s}), \boldsymbol{s} \right) + b_g(\boldsymbol{s}), \quad (5)$$

with β a number between 0 and 1 reflecting the camera quantum efficiency, $b_p(\lambda, \boldsymbol{s})$ a random variable following a Poisson distribution, and $b_g(\boldsymbol{s})$ a random variable following a zero-mean half-normal distribution. Since there are cases where the PSF estimation is not possible, e.g. where the sample lacks texture, such as in uniformly black or gray areas, we added a boolean parameter $w_k(\boldsymbol{s})$ (whose values can be either 0 or 1), which indicates the "legitimacy" of the sample (i.e is this image textured enough to yield useful information?). We illustrated such input vectors with their corresponding degraded samples in Table 1.

We trained a U-Net DNN [13] with a ResNet encoder [14] pre-trained on ImageNet [15], in order to predict, with the image $i(\boldsymbol{s})$ as input, the map of parameters $(\widetilde{\vec{v}}(\boldsymbol{s}), \widetilde{z}_k(\boldsymbol{s}), \widetilde{w}_k(\boldsymbol{s}))$ converted using cylindrical coordinates. We assessed in [10] that such a network was robust to unwanted image degradations such as Poisson and Gaussian noise. We trained the network for 50 epochs in PyTorch with RAdam [16] optimizing the following loss function:

$$E^{(k)} = \gamma \left(w_k(\boldsymbol{s}) - \widetilde{w}_k(\boldsymbol{s}) \right)^2 + \frac{1 - w_k(\boldsymbol{s})}{U+1} \left[\sum_{u=1}^{U} \left(|v_u(\boldsymbol{s})| - |\widetilde{v}_u(\boldsymbol{s})| \right)^2 + \left(z_k(\boldsymbol{s}) - \widetilde{z}_k(\boldsymbol{s}) \right)^2 \right], \quad (6)$$

with $U = 3$ components in $\vec{v}$, and γ a hyperparameter regulating the importance of the validity parameter w, that we set to 1 in our further experiments.

3. EXPERIMENTS

3.1. Characterization of the displacement vector field estimation with synthetic data

We aim at defining the performance of the method using test data generated in the same way as the training data, but with a separate data set of $K_{\text{test}} = 5000$ images cropped at 224×224 pixel, preliminary acquired using a Leica DM 5500, a 10×/0.3 objective, and fixed fluorescent samples (HeLa cells actin (Alexa Fluor 635) and HeLa cells anti-α-catenin (Alexa Fluor 488)). Specifically, we took sharp images of non-moving objects and blurred them with two generated PSFs modeling different three-dimensional flow rates from a uniform distribution. We then used the DNN trained in Section 2.3 to predict the flow vector $\widetilde{\vec{v}}(\boldsymbol{s})$, the axial position $\widetilde{z}_0(\boldsymbol{s})$, and the "validity" parameter $w_k(\boldsymbol{s})$. Since it is a regression problem, our metric was set to be the squared Pearson correlation coefficient R^2 averaged over all dimensions.

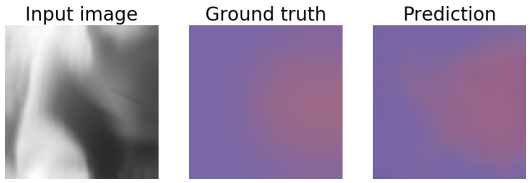


Fig. 2: Velocity vector field estimation from a single motion blurred image from the images of fluorescent HeLa cells and two PSFs drawn from a random distribution. The network output and the ground truth vectors are represented in the RGB spectrum with $\widetilde{v}_1(\boldsymbol{s})$ in the red channel, $\widetilde{v}_2(\boldsymbol{s})$ in the green channel, and $\widetilde{v}_3(\boldsymbol{s})$ in the blue channel.

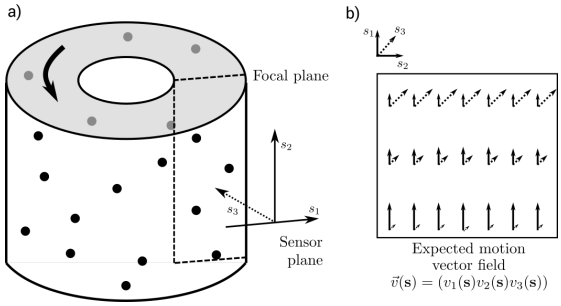


Fig. 3: (a) Simulation of a flow in a cylinder. (b) simulation of its expected flow profile as captured by the camera. The flow vector $\vec{v}(\boldsymbol{s})$ has a greater lateral component in the bottom of the image, and a larger axial component in the top of the image.

3.2. Characterization of the displacement vector field estimation with simulated flow

We then turned to a more realistic experiment and generated a synthetic testing dataset that simulates the motion of a cylinder where the camera and the focal plane are perpendicular to the motion direction. Due to the small depth-of-field (DOF) in microscopy, the effect of the cylinder curvature is negligible. The flow vector map $\vec{v}(\boldsymbol{s})$ is then similar to Fig. 3 (b). We neglected as well the effects of the non-slip condition at the walls present in Poiseuille flow.

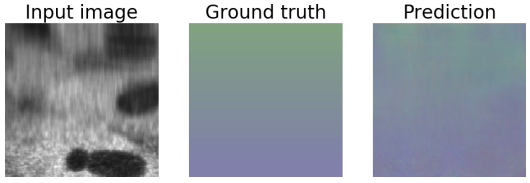


Fig. 4: Velocity vector field estimation from a single motion blurred image and a gradient of PSFs mimicking the conditions of Fig. 3. The network output and the ground truth vectors are represented in the RGB spectrum with $\widetilde{v}_1(\boldsymbol{s})$ in the red channel, $\widetilde{v}_2(\boldsymbol{s})$ in the green channel, and $\widetilde{v}_3(\boldsymbol{s})$ in the blue channel.

4. DISCUSSION AND CONCLUSION

Our experiments on simulated data confirmed the network's capability to regress a pixel-wise motion vector, not only for point-like sources, but also from a single textured image blurred with a three-dimensional motion PSF. Indeed, when it came to the task of estimating two different motion vector in two zones in an image, the network achieved a Pearson regression coefficient of $R^2 = 0.92$ averaged over all pixels of $N = 1000$ images of 224×224 pixel (see Fig. 2). Similarly, to retrieve the cylindrical flow profile in the second experiment, the regression coefficient was computed at $R^2 = 0.91$ using the same conditions as before (see Fig. 4).

In all our experiments, the axial component $\widetilde{v_3}$ was predicted with a systematically greater error than the lateral components. That could be explained by the confusion between an object with larger axial velocity that started its motion right in focus, and an object with a smaller axial velocity, but whose motion happens out-of-focus. Both situations yielded similar-looking PSF since the generation of the PSF from the parameters $\vec{v}(\boldsymbol{s})$ and $z_0(\boldsymbol{s})$ is not a perfectly bijective transformation (similar-looking PSFs can be generated from other sets of parameters). However, as illustrated in Table 1, our method for predicting the out-of-plane velocity field was relatively robust to the change in the starting position of the sample z_0.

Even though our development is done for thin manifolds, in practice, we expect microscope objectives with sufficiently shallow depths of field and sparse samples to fulfill our method's assumptions. We showed that a neural network can be used to predict a motion vector field from a single textured microscopy image degraded with motion blur with only minimal knowledge about the optical setup. This opens the possibility of retrieving flow information without the need for dedicated high-speed camera or strong light exposure.

The source code and trained model is available at the following address: `https://github.com/idiap/flowestimation`.

5. COMPLIANCE WITH ETHICAL STANDARDS

The post-mortem stained and fixed tissue slices were reused from experiments approved by the EPFL Ethics Committee.

6. ACKNOWLEDGMENTS

This work was funded by the Swiss National Science Foundation, Grant 200020_179217. The authors have no conflict of interest to disclose.

7. REFERENCES

[1] T. Corpetti, D. Heitz, G. Arroyo, E. Mémin, and A. Santa-Cruz, "Fluid experimental flow estimation based on an optical-flow scheme," *Exp Fluids*, vol. 40, no. 1, pp. 80–97, 2006.

[2] O. Brina *et al.*, "Intra-Aneurysmal Flow Patterns: Illustrative Comparison among Digital Subtraction Angiography, Optical Flow, and Computational Fluid Dynamics," *American Journal of Neuroradiology*, vol. 35, no. 12, pp. 2348–2353, 2014.

[3] A. Dosovitskiy *et al.*, "FlowNet: Learning Optical Flow with Convolutional Networks," in *IEEE ICCV*, 2015, pp. 2758–2766.

[4] J. Sun, W. Cao, Z. Xu, and J. Ponce, "Learning a Convolutional Neural Network for non-uniform motion blur removal," in *IEEE CVPR 2015*, 2015.

[5] J. Walker, A. Gupta, and M. Hebert, "Dense Optical Flow Prediction from a Static Image," in *IEEE ICCV*, Santiago, Chile: IEEE, 2015, pp. 2443–2451.

[6] D. Gong *et al.*, "From Motion Blur to Motion Flow: A Deep Learning Solution for Removing Heterogeneous Motion Blur," in *IEEE CVPR*, IEEE, 2017, pp. 3806–3815.

[7] S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *IEEE CVPR*, 2017.

[8] L. Tian, Z. Tu, D. Zhang, J. Liu, B. Li, and J. Yuan, "Unsupervised Learning of Optical Flow With CNN-Based Non-Local Filtering," *IEEE Transactions on Image Processing*, vol. 29, pp. 8429–8442, 2020.

[9] J. Icha, M. Weber, J. C. Waters, and C. Norden, "Phototoxicity in live fluorescence microscopy, and how to avoid it," *BioEssays*, vol. 39, no. 8, p. 1 700 003, 2017.

[10] A. Shajkofci and M. Liebling, "Spatially-Variant CNN-Based Point Spread Function Estimation for Blind Deconvolution and Depth Estimation in Optical Microscopy," *IEEE Transactions on Image Processing*, vol. 29, pp. 5848–5861, 2020.

[11] F. von Zernike, "Beugungstheorie des schneidenverfahrens und seiner verbesserten form, der phasenkontrastmethode," *Physica*, vol. 1, no. 7, pp. 689–704, 1934.

[12] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: A 10 million image database for scene recognition," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 40, no. 6, pp. 1452–1464, 2018.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *MICCAI 2015*, vol. 9351, Cham: Springer International Publishing, 2015, pp. 234–241.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE CVPR*, 2016, pp. 770–778.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep Convolutional Neural Networks," in *NIPS*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds., Curran Associates, Inc., 2012, pp. 1097–1105.

[16] L. Liu *et al.*, "On the Variance of the Adaptive Learning Rate and Beyond," in *ICLR 2020*, 2020.