Linear Blur Parameters Estimation using a Convolutional Neural Network

Nasonov Andrey Vladimirovich [a,*], Nasonova Alexandra Andreevna [a]

[a] *Lomonosov Moscow State University,*

Faculty of Computational Mathematics and Cybernetics

Laboratory of Mathematical Methods of Image Processing

GSP-1, Leninskie Gory, Moscow, 119991, Russian Federation

* e-mail: nasonov@cs.msu.ru

Abstract— Motion blur is visible whenever the shutter speed of a camera is slow compared to the speed of unintended camera motion. General motion blur is a very complex type of blurring, and state-of-the-art blind image deconvolution methods rarely produce adequate results due to the ill-posed nature of the problem. Even modern deep-learning algorithms sometimes fail at the task.

Modern deblurring approaches typically use a series of noisy images with shorter exposure time for the reconstruction of a high quality image. However, even with a shorter exposure time some blurring still remains. The good news is that, with little time for the motion vector to change direction significantly, this particular type of motion blur is much easier to model.

The crucial stage in any deblurring process is the estimation of blur parameters. In this article we present a patch-based linear approximation to motion blur with the focus on effective estimation of the direction of linear blur. We use a CNN model for estimating the parameters of a linear blur kernel for each 32x32-pixel patch of an image and calculating a confidence value for each patch.

Keywords: linear blur estimation, convolutional neural network, motion blur

INTRODUCTION

Photographs obtained with handheld cameras represent a significant category of digital images. The resolution of an image is measured in megapixels, which influences the amount of details that can be captured in an image, but a higher megapixel count doesn’t always equate to a better picture. The amount of noise is directly related to the overall amount of light captured in an image, and the factor that contributes a lot to image quality is the size of the camera's sensor.

A sure way to increase the amount of light the sensor receives is prolonging the exposure time, but that inevitably leads to blurring from even the slightest motions of a camera. Especially in low-light conditions, it is impossible to acquire both sharp and noise-free images using hand-held cameras.

Reconstructing a sharp image from a blurry one is an ill-posed problem, with various additional constraints used to regularize the solution. While numerous blind deconvolution algorithms have shown decent performance in certain cases [1, 2], they typically do not perform well in more complex yet common scenarios such as images with strong motion blur.

Some modern approaches for image enhancement are based on reconstructing a high-quality image from a series of images. For example, the algorithm [3] utilizes a pair of images that can be easily acquired in low-light conditions: a blurred image taken with low shutter speed and low ISO value, and a noisy image captured with high shutter speed and high ISO value. Both images are sliced into patches, and the authors extend the Gaussian mixture model to model the underlying intensity distribution of each patch using the corresponding patches in the noisy image.

The algorithm [4] makes use of natural hand tremor, which is typical in hand-held photography, to acquire a burst of raw frames. These frames are then aligned and merged to form a single image.

The increase of the resolution of modern hand-held cameras makes the blur more prominent, and even with a shorter exposure time some blurring still remains. This supports the demand for high-quality image deblurring algorithms. During a short exposure, there is little time for the motion vector to change direction, so the motion blur can be approximated with linear blur, which is much easier to model.

Many state-of-the-art deblurring algorithms are based on the deep learning approach [5]. In [6], a neural network is trained to estimate a set of image-adaptive basis motion kernels with weight coefficients for each pixel, which produces a per-pixel motion blur field.

Gong et al. [7] use a Fully Convolutional Network (FCN) for the estimation of a dense linear motion flow parameterized by the horizontal and vertical components. For FCN training they generate synthetic pairs of blurred images and corresponding motion flow.

Sun et al. [8] consider a set of pre-defined linear motion kernels parameterized by their lengths and orientations. They split the image into patches and use a CNN to predict probabilistic distribution of the kernel parameters for each patch. The sparse patch-level distribution is then converted to a dense motion field using a Markov random field that ensures its smoothness.

The existing deep learning solutions to image deblurring usually present a pipeline with an image at the input and an enhanced image at the output, yet there are cases when some parts of an image remain blurry. This commonly happens due to inaccurate estimation of the blur parameters as the neural network solves the problem as a whole and does not provide the capability to control the parameters of the blur.

It is our belief that refining the parameters of the deblurring process warrants improvement of the overall performance of existing algorithms. We dedicate our research to the assessment of the non-uniform linear motion blur instead of developing yet another deblurring pipeline. In this article we focus on the estimation of the direction of linear blur.

LINEAR BLUR MODEL

We use the model of a linear blur kernel with direction θ and length l in the following form, with the examples shown in Fig.1:

$$h[\theta, l](x, y) = h[l](x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta),$$

$$h[l](x, y) = \frac{1}{l}\int_{-l/2}^{l/2} G_\sigma\,(x - p, y)dp,$$

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Here $h[l]$ is the linear blur kernel along x-axis with the length l and G_σ is Gaussian filter kernel which is used to prevent aliasing, we use $\sigma = 0.3$.

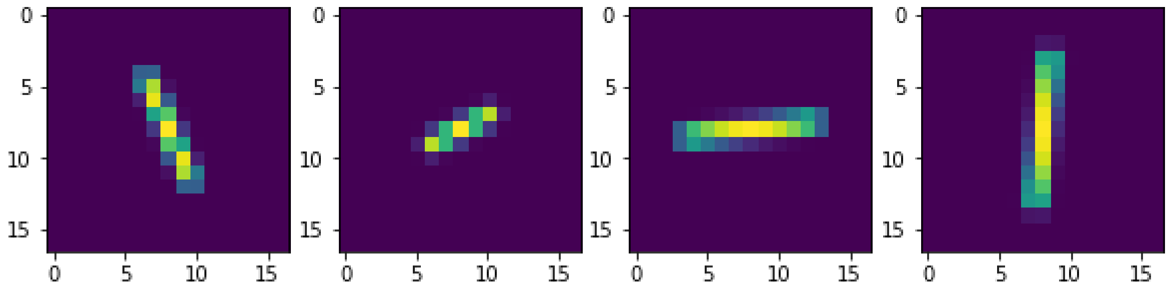


Fig. 1. Examples of linear blur kernels with different parameters.

CNN MODEL

In our work we split the images into patches and develop an algorithm that infers the parameters of the linear motion blur for each patch. The resulting sparse motion vector field can be interpolated to a dense motion vector field using various methods: simple averaging [9], fine-tuning [10] or more sophisticated methods like Markov random field for ensuring the motion smoothness [8].

We use a convolution neural network (CNN) to solve the problem of assessing the parameters of linear blur (the same structure, illustrated in Fig.2, applies to both direction and length of the linear blur kernel).

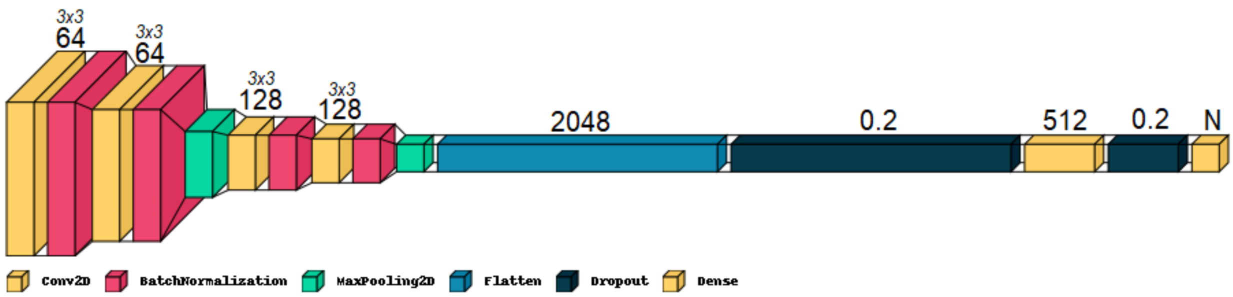


Fig. 2. Structure of the proposed CNN for the estimation of the parameters of a linear blur kernel.

We have explored several options for the output of CNN:

1. Indicator vector. Consider a discrete set of parameters $\{\theta_i, l_j\}, i = 1, \dots, N, j = 1, \dots, M$ of linear motion blur, where θ is the direction of the kernel and l represents its length. In this case, CNN output is an indicator vector which characterizes the probabilistic distribution of motion kernels [8]. The disadvantage of this approach is that different blur kernels may produce similar blurred patches, which would impair the learning process. In this case constructing an adequate training dataset becomes an overly complicated problem.
2. Pairs of values $\{\theta, l\}$. The main problem here is that the direction wraps over π, which cannot be handled by a common CNN model.
3. A vector $\{\sin^2\theta, \cos^2\theta, l\}$. Here we calculate $\sin^2\theta$ and $\cos^2\theta$ values instead of the direction θ itself. The values belong to the interval $[0, 1]$ and change smoothly.

We have observed that synchronous estimation of both direction and length of the linear blur kernel fails to produce accurate estimates for the length value since the direction has greater impact on the blurred image.

Further investigation has proven that the most accurate results can be obtained for independent estimation of the direction of a linear blur kernel, while the problem of accurate estimation of the length still remains open.

During the follow-up investigation we have observed a higher error rate with directions close to $\pi/4$ and $3\pi/4$, when the values of $(\sin^2\theta, \cos^2\theta)$ are farthest from both 0 and 1. In order to overcome this problem, we have increased the number of values in the output vector $\{v_0, \dots, v_{N-1}\}$:

$$v_n = \sin^2\left(\theta + \frac{\pi n}{N}\right).$$

We have compared the distribution of absolute errors for inferred and ground truth blur directions from the test part of the training dataset. The histograms are shown in Fig. 3. It can be seen that adding two more values (*N*=4) drastically decreases the amount of patches with absolute error greater than 15 degrees. Further increasing *N* leads to better accuracy.

We have chosen *N*=6.

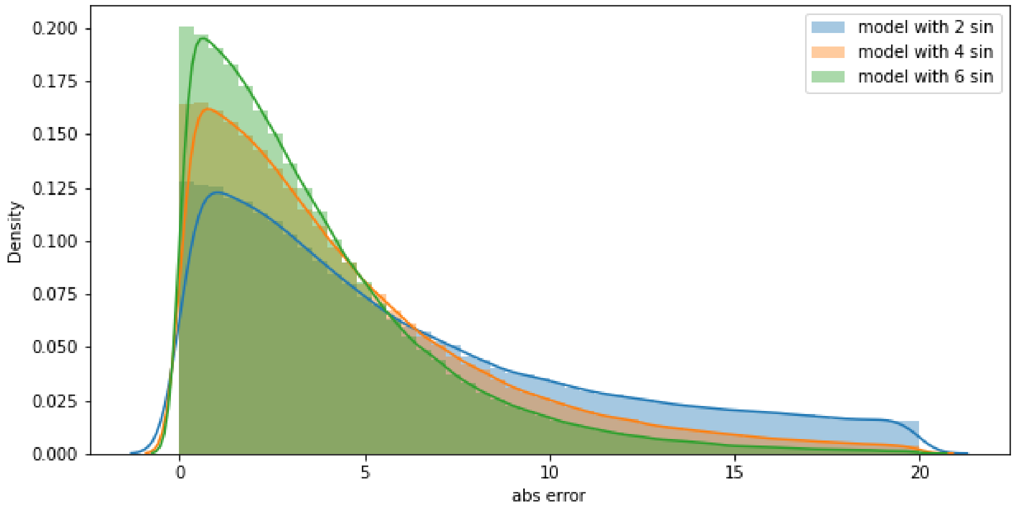


Fig. 3. Histograms of erroneously predicted blur directions using CNN models with different output vector length *N*.

In order to find the angle θ from the vector output $\mathbf{v} = \{\mathrm{v_n}\}$ of the CNN, we find θ that minimizes

$$F(\mathbf{v},\theta) = \sum_{n=0}^{N-1}\left(v_n - \sin^2\left(\theta + \frac{\pi n}{N}\right)\right)^2.$$

The value $F(\mathbf{v},\theta)$ for a given vector **v** can be used as a confidence level: low values $F(\mathbf{v},\theta)$ corresponds to blocks that likely contains pronounced motion blur. We eliminate the patches with $F(\mathbf{v},\theta) > 0.02$. The threshold has been set experimentally.

DATASET PREPARATION

For the creation of our dataset we use images from the KonIQ-10k dataset [11] that contains about 10 thousand images of diverse content. Each image is split into 32x32-pixel patches with 16-pixel overlap. From each image we take the top 10% of the patches based on the standard deviation in order to exclude flat areas. We apply patch-wise centering, subtracting the mean intensity value from each patch.

We add random impairments to each patch: first, we apply linear blur kernel with random parameters (θ, l), $\theta \in [0, \pi]$, $l \in [0, 10]$; then we add Gaussian noise with random standard deviation $\sigma \in [0, 8]$.

The dataset is split randomly into train/test sets with 80%/20% ratio.

EXPERIMENTS AND RESULTS

We evaluate the proposed method using images with simulated blur corresponding to camera movement in the hands. The blur is modeled by translation (dx, dy) and rotation α around image center. The parameters (dx, dy, α) are chosen randomly such that the maximal length of motion blur does not exceed $l_{max} = 10$.

The results are presented at Fig. 4 and Fig. 5.

In can be seen that the proposed algorithm produces sparse vector field with reliable motion information. Incorporation of the proposed algorithm into the deblurring pipeline in is progress so we cannot provide numerical comparison with state-of-the-art motion deblurring algorithms yet.

Source code and data are available at

https://imaging.cs.msu.ru/en/research/motiondeblur

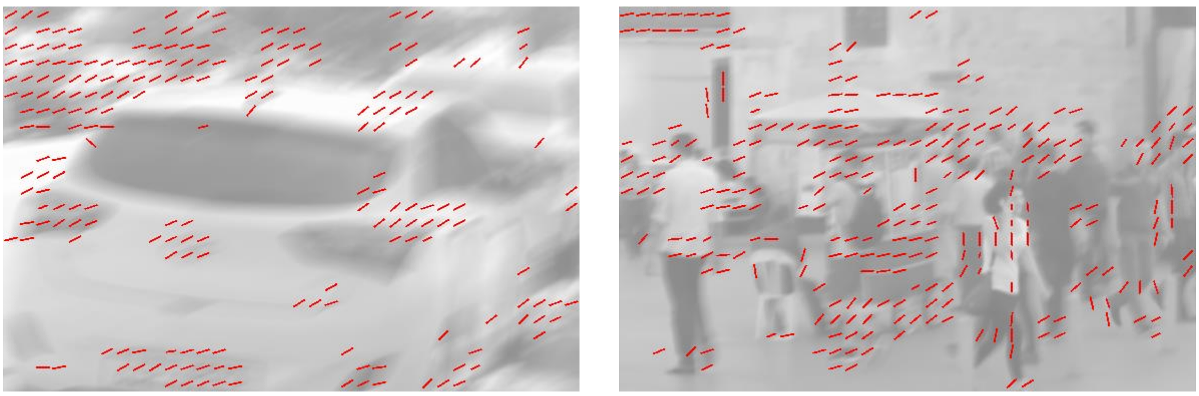
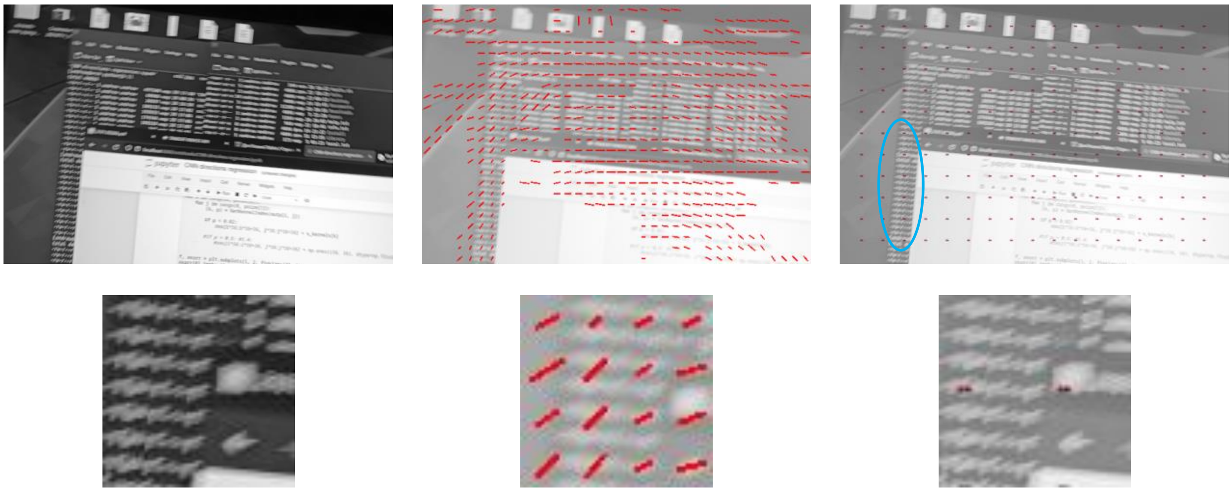


Fig. 4. An example of sparse motion vector fields produced by the proposed algorithm. Images are takes from GOPRO dataset [12].



a) Blurred image b) Proposed method c) Carbajal et al. [6]

Fig. 5. A visual comparison of the proposed method and state-of-the-art algorithm.

CONCLUSIONS

An algorithm for linear blur parameters estimation using a convolution neural network has been proposed. It produces sparse direction field for the motion blur. Interpolation of the sparse field into dense field will be part of the future work.

FUNDING

The work was supported by Russian Science Foundation Grant 22-41-02002.

CONFLICT OF INTERESTS

The process of writing and the content of the article does not give grounds for raising the issue of a conflict of interest.

COMPLIANCE WITH ETHICAL STANDARDS

This article is a completely original work of its authors; it has not been published before and will not be sent to other publications until the PRIA editorial board decides not to accept it for publication.

REFERENCES

1. Z. Xu, H. Chen, Z. Li, “Fast blind deconvolution using a deeper sparse patch-wise maximum gradient prior,” Signal Processing: Image Communication, 90, 116050 (2020).
2. Y. Bai, G. Cheung, X. Liu, X., W. Gao, “Graph-based blind image deblurring from a single photograph,” IEEE Transactions on Image Processing, 28(3), 1404–1418 (2018).
3. C. Gu, X. Lu, Y. He, C. Zhang, “Blur removal via blurred-noisy image pair,” IEEE Transactions on Image Processing, 30, 345–359 (2020).
4. B. Wronski, I. Garcia-Dorado, M. Ernst, D. Kelly, M. Krainin, C.K. Liang, P. Milanfar, “Handheld multi-frame super-resolution,” ACM Transactions on Graphics (TOG), 38(4), 1–18 (2019).
5. K. Zhang, W. Ren, W. Luo, W.S. Lai, B. Stenger, M.H. Yang, H. Li, “Deep Image Deblurring: A Survey,” arXiv preprint arXiv:2201.10700 (2022).
6. G. Carbajal, P. Vitoria, M. Delbracio, P. Musé, J. Lezama, “Non-uniform blur kernel estimation via adaptive basis decomposition,” arXiv preprint arXiv:2102.01026 (2021).
7. D. Gong, J. Yang, L. Liu, Y. Zhang, I. Reid, C. Shen, Q. Shi, “From motion blur to motion flow: A deep learning solution for removing heterogeneous motion blur,” IEEE Conference on Computer Vision and Pattern Recognition, 2319–2328 (2017).
8. J. Sun, W. Cao, Z. Xu, Z., J. Ponce, “Learning a convolutional neural network for non-uniform motion blur removal,” IEEE Conference on Computer Vision and Pattern Recognition 769–777 (2015).
9. A. Chakrabarti, “A neural approach to blind motion deblurring,“ European Conference on Computer Vision, 221–235 (2016).
10. X. Xu, J. Pan, Y. J. Zhang, M. H. Yang, “Motion blur kernel estimation via deep learning,” IEEE Transactions on Image Processing, 27(1), 194–205 (2017).
11. V. Hosu, H. Lin, T. Sziranyi, D. Saupe, “KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment,” IEEE Transactions on Image Processing, 29, 4041–4056 (2020).

12. S. Nah, T.H. Kim, K.M. Lee, “Deep multi-scale convolutional neural network for dynamic scene deblurring,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3883–3891 (2017).

Authors profile

Nasonov Andrey Vladimirovich, PhD, is a senior researcher at Laboratory of Mathematical Methods of Image Processing, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University. E-mail: nasonov@cs.msu.ru

Nasonova Alexandra Andreenva, PhD, is a mathematician at Laboratory of Mathematical Methods of Image Processing, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University. E-mail: nasonova.alexandra@gmail.com