Nonstationary blur modeling using robust eigenkernels

Mauro Luiz Brandão-Junior, Victor Carneiro Lima e Renato da Rocha Lopes

***Abstract*— In this paper we propose a robust low rank model for restoring images corrupted by a nonstationary blur. This work improves the eigenkernels model proposed by Gwak and Yang. Their method uses standard Principal Component Analysis (PCA), and is thus not well suited to data with outliers. We replace PCA by its robust version. Numerical experiments show that our proposal offers reliable blur description and restoration even in the presence of salt-and-pepper noise, in which the original framework fails.**

***Keywords*— Nonstationary blur, Robust Principal Component Analysis, Image restoration**

I. Introduction

This paper focuses on the subject of nonstationary blur modeling, which extends the idea of optical distortion modeling using space-invariant convolutions to more realistic space-variant ones. Unlike the simplistic stationary approach, nonstationary models distort each pixel in a target image using a different convolutive kernel. This extra flexibility is particularly valuable when dealing with intricate systems, such as the optics involved in imaging a large scene using a single low aperture lens [1], [2], [3], [4].

Let $I \subset \mathbb{Z}^2$ be a finite set of (i, j) indexes. Optical image acquisition provides an observed image $y : I \to \mathbb{R}$, which is a distorted version of a latent image $x : I \to \mathbb{R}$. This distortion can be modeled as a family of blur kernels $h^{(i,j)} : N(i, j) \subset I \to \mathbb{R}$, for all $(i, j) \in I$. The subset $N(i, j)$ is a neighborhood of indexes around (i, j). For every pixel indexed in I we associate a blur kernel which models the distortion affecting that single pixel. In this setting, the distorted image is calculated as a 2D linear space-variant (LSV) convolution [5] plus additive noise, which can be described using the filtering operation

$$y(i,j) = \sum_{(a,b)\in N(i,j)} x(a,b)h^{(i,j)}(i-a, j-b) + \eta(i,j) \quad (1)$$

$$= \left(x \star h^{(i,j)}\right)(i,j) + \eta(i,j), \quad (2)$$

where $\star$ represents the 2d linear space-invariant (LSI) convolution. Also, the input image x is padded at the borders when needed.

Mauro Luiz Brandão-Junior, Decom, University of Campinas, Campinas - SP, e-mail: m228001@dac.unicamp.br; Victor Carneiro Lima, Decom, University of Campinas, Campinas - SP, e-mail: v157460@dac.unicamp.br; Renato da Rocha Lopes, Decom, University of Campinas, Campinas - SP, e-mail: rlopes@unicamp.br. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior – Brasil (CAPES) – Finance Code 001, and the National Council for Scientific and Technological Development (CNPq), grant number 305480/2018-9.

The increased accuracy in nonstationary blur modeling comes with a price: In principle, we need to consider a different kernel for each position in I, the set of indices of a target image. This is a problem for large images, since kernel data is usually stored in memory. A common approach is to assume that the PSF (Point Spread Function) varies slowly for adjacent pixels and subsample it only on a subset $J \subset I$ of indices. We could take J as a regularly spaced grid of points, for example. Then, the PSF at the remaining positions can be estimated using some kind of interpolation. In this direction, the usual hypothesis made is that the blur is stationary inside regions around each point in J. This gives rise to perhaps the most well established method for handling nonstationary PSF, the Efficient Filter Flow (EFF) [1], [3], [4].

We follow the work of Gwak and Yang [2], that is also based on kernel subsampling. Gwak and Yang detail the implementation of the nonstationary blur operation using a set of eigenkernels obtained via Principal Component Analaysis (PCA) of the kernels indexed in J. To illustrate the effectiveness of their method, we blurred a test image using the LSV convolution (2) and a PCA based approximation. The results are shown in Fig. 1, with calculated Peak Signal to Noise Ratio (PSNR) with respect to the LSV convolution. This first example points out to how fast and how reliable the PCA approximation can be with respect to the full convolution.

Although the eigenkernels proposal by Gwak and Yang is impressive, standard PCA is highly susceptible to heavily corrupted observations. Thus, in a practical setting, the success of this proposal hinges on the requirement that all estimated kernels are free of outliers, which is unrealistic. To address this point, in this work we reinterpret the original eigenkernels model as a low rank representation of nonstationary blur and extend it to handle sparse anomaly corruption using a regularization approach.[1] The remainder of the paper is organized as follows: Section II reviews the original framework and presents a simple example of restoration strategy. In section III, we present our robust eigenkernel proposal which is evaluated in numerical examples in section IV.

II. Low rank representation of nonstationary blur

We consider q kernels with indices in the set J, each having m pixels in total. Let $\boldsymbol{H} \in \mathbb{R}^{m\times q}$ be the matrix whose columns

[1] All experiments and examples presented in this paper can be reproduced using blurtool, a small library containing the relevant structures and operations that are discussed here. The library is written in python and is available at `https://www.github.com/mlbj/blurtool`.

(a) (b) (c)

Fig. 1: (a) Test image. (b) Test image blurred with linear space-variant convolution $\Delta\tau = 42.306$ s). (c) Test image blurred with eigenkernels model PSNR $= 52.305$ dB, $\Delta\tau = 0.034$ s). All the computations done using the same computer setup.

are the vectorized kernels of each position in J, organized in lexicographic ordering.

The fact that kernels are usually similar to each other suggests that they lie in a low dimensional subspace. If this holds, we can efficiently describe them using Principal Component Analysis (PCA). We can evaluate PCA by obtaining a truncated Singular Value Decomposition (SVD) of $\overline{\boldsymbol{H}} = \boldsymbol{H} - \boldsymbol{\rho}\mathbf{1}^T$, where $\boldsymbol{\rho}$ is the mean sampled kernel, and $\mathbf{1}$ is a column vector of ones. From the calculated SVD, we extract $r \leq q$ principal axes, called eigenkernels in this context, and r vectors of representation weights. Ideally, if the space that contains all sampled kernels is indeed low dimensional, a small value of r should be enough to represent them with small error.

Each mean subtracted kernel indexed by J, i.e., each column of $\overline{\boldsymbol{H}}$, can be approximated using the linear combination of the eigenkernels with respective representation coefficients. This can be written in 2d signal notation as

$$h^{(i,j)} - \rho \approx \sum_{k=1}^{r} c_k(i,j) e_k, \tag{3}$$

where e_k are the eigenkernels, $c_k(i,j)$ are the corresponding representation weights and $(i,j) \in J$.

Since we assume that the original PSF varies slowly across adjacent pixels, it is reasonable to assume that the representation weights c_k, which were only estimated for the subsampled indexed set $J \subset I$, can be interpolated to the whole set I, thus approximating the complete space-varying PSF. Therefore, let $\tilde{c_k}$ be the interpolated coefficients. The rank r approximation of the PSF at position $(i,j) \in I$ is given by

$$\tilde{h}^{(i,j)} = \rho + \sum_{k=1}^{r} \tilde{c_k}(i,j) e_k = \sum_{k=0}^{r} \tilde{c_k}(i,j) e_k, \tag{4}$$

where, in the last term, we make $\rho = e_0$ and $\tilde{c_0}(i,j) = 1$ for all (i,j). This creates a more compact notation.

Now, plugging the reconstructed PSF (4) into the LSV convolution (2), we can approximate the observed image by

$$y(i,j) \approx \left(\tilde{h}^{(i,j)} \star x\right)(i,j) + \eta(i,j) \tag{5}$$

$$= \sum_{k=0}^{r} \tilde{c_k}(i,j) \left(e_k \star x\right)(i,j) + \eta(i,j). \tag{6}$$

Each convolution in (6) is with a space-invariant kernel. Thus, the eigenkernels approximate a space-varying convolution by r space-invariant convolutions, which can be implemented, assuming appropriate boundary conditions, using doubly block circulant matrices:

$$\boldsymbol{y} = \sum_{k=0}^{r} \boldsymbol{C}_k \boldsymbol{E}_k \boldsymbol{x} + \boldsymbol{n}, \tag{7}$$

where $\boldsymbol{C}_k \in \mathbb{R}^{m \times m}$ is a diagonal matrix with the weights $\tilde{c_k}(i,j)$ in its nonzero entries, and $\boldsymbol{E}_k \in \mathbb{R}^{m \times m}$ is the doubly block circulant matrix that performs the 2d convolution with the kth eigenkernel, $\boldsymbol{n}$ is random additive noise, $\boldsymbol{x}$ and $\boldsymbol{y} \in \mathbb{R}^m$ are the vectorized latent and observed images, respectively.

A. Restoration

Ignoring the noise term and assuming a periodic boundary condition, the distortion caused by the eigenkernels model to an input image $\boldsymbol{x}$ can be described by the action of the matrix $\boldsymbol{M} = \sum_{k=1}^{r} \boldsymbol{C}_k \boldsymbol{E}_k$, which is not generally doubly block circulant but has a very convenient structure. In fact, the matrix-vector products $\boldsymbol{M}\boldsymbol{x}$ and $\boldsymbol{M}^T\boldsymbol{x}$ can be efficiently calculated using the diagonalization of each $\boldsymbol{E}_k$ with the help of the Fast Fourier Transform (FFT) algorithm [6] using

$$\begin{aligned} \boldsymbol{M}\boldsymbol{x} &= \sum_{k=1}^{r} \boldsymbol{C}_k \boldsymbol{F}^{-1} \boldsymbol{D}_k \boldsymbol{F} \boldsymbol{x} \\ \boldsymbol{M}^T\boldsymbol{x} &= \sum_{k=1}^{r} \boldsymbol{F}^{-1} \boldsymbol{D}_k^* \boldsymbol{F} \boldsymbol{C}_k \boldsymbol{x}, \end{aligned} \tag{8}$$

where $\boldsymbol{F}$ is the unitary scaled 2d Discrete Fourier Transform (DFT) matrix, and $\boldsymbol{D}_k$ is the properly shifted and padded Optical Transfer Function (OTF) of the kth eigenkernel. Using (8), several deblurring algorithms can be immediately adapted for the nonstationary case. As a simple toy example, consider the l_2 regularized problem, in which, as before, $\boldsymbol{y}$ represents the observed image, $\boldsymbol{x}$ represents the latent image and $\boldsymbol{M}$ models the nonstationary distortion:

$$\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{M}\boldsymbol{x}\|_2^2 + \frac{\lambda}{2}\|\boldsymbol{x}\|_2^2. \tag{9}$$

Its solution is equivalent to the linear system

$$(\boldsymbol{M}^T\boldsymbol{M} + \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{M}^T\boldsymbol{y}. \tag{10}$$

For $\lambda > 0$, we have that $\boldsymbol{M}^T\boldsymbol{M} + \lambda\boldsymbol{I}$ is positive definite, so (10) can be solved using the conjugate gradient iterations

$$\begin{aligned} \alpha_i &:= \frac{\boldsymbol{r}_i^T\boldsymbol{r}}{\boldsymbol{p}_i^T\left(\boldsymbol{M}^T\boldsymbol{M} + \lambda\boldsymbol{I}\right)\boldsymbol{p}_i} \\ \boldsymbol{x}_{i+1} &:= \boldsymbol{x}_i + \alpha_i\boldsymbol{p}_i \\ \boldsymbol{r}_{i+1} &:= \boldsymbol{r}_i - \alpha_i\left(\boldsymbol{M}^T\boldsymbol{M} + \lambda\boldsymbol{I}\right)\boldsymbol{p}_i \\ \boldsymbol{p}_{i+1} &:= \boldsymbol{r}_{i+1} + \frac{\boldsymbol{r}_{i+1}^T\boldsymbol{r}_{i+1}}{\boldsymbol{r}_i^T\boldsymbol{r}_i}\boldsymbol{p}_i, \end{aligned} \tag{11}$$

where $\boldsymbol{p}_0 = \boldsymbol{r}_0 = \boldsymbol{M}^T\boldsymbol{y} - \left(\boldsymbol{M}^T\boldsymbol{M} + \lambda\boldsymbol{I}\right)\boldsymbol{x}_0$, where $\boldsymbol{x}_0$ is a vector of zeros. Notice that the algorithm in (11) can be implemented using only the matrix-vector products in (8), which can be implemented with low complexity. In fact, the iterations (11) only need two of such product per loop.

Several other optimization tools could be used to solve the toy problem we stated in (9), but the conjugate gradient method makes use of the expressions (8) in a very clear way. Also, other regularization priors like Total Variation (TV) [7], [8] could be adopted.

III. Robust eigenkernels

Although the eigenkernels model reviewed in section II provides a convenient way of expressing a nonstationary PSF, classical PCA notoriously loses its performance in the presence of grossly corrupted samples or outliers. Fortunately, the rich literature of Robust Principal Component Analysis (RPCA) provides several tools to deal with these adversities. Following Candès' original RPCA proposal [9], we aim to find a decomposition of the sampled kernels matrix $\boldsymbol{H}$ that is a solution of the convex optimization problem

$$\begin{aligned} &\min_{\boldsymbol{B},\boldsymbol{A}} \|\boldsymbol{B}\|_* + \lambda\|\boldsymbol{A}\|_1 \\ &\text{s.t. } \boldsymbol{B} + \boldsymbol{A} = \boldsymbol{H} \end{aligned} \tag{12}$$

where $\|\cdot\|_*$ is the nuclear norm and $\|\cdot\|_1$ is the l_1 norm of the vectorization. The simplest solution to (12) is the Principal Component Pursuit (PCP) algorithm, which is derived as an Alternating Direction Method of Multipliers (ADMM) [10] iteration that minimizes the augumented Lagrangian

$$\|\boldsymbol{B}\|_* + \lambda\|\boldsymbol{A}\|_1 + \langle\boldsymbol{V}, \boldsymbol{B} + \boldsymbol{A} - \boldsymbol{H}\rangle + \frac{\mu}{2}\|\boldsymbol{B} + \boldsymbol{A} - \boldsymbol{H}\|_F^2,$$

in each variable while fixing the others, that is

$$\begin{aligned} \boldsymbol{B}^{k+1} &= \mathcal{D}_{\mu^{-1}}\left(\boldsymbol{H} - \boldsymbol{A}^k - \mu^{-1}\boldsymbol{V}^k\right) \\ \boldsymbol{A}^{k+1} &= \mathcal{S}_{\lambda\mu^{-1}}\left(\boldsymbol{H} - \boldsymbol{B}^{k+1} - \mu^{-1}\boldsymbol{V}^k\right) \\ \boldsymbol{V}^{k+1} &= \boldsymbol{V}^i + \mu\left(\boldsymbol{B}^{k+1} + \boldsymbol{A}^{k+1} - \boldsymbol{H}\right) \end{aligned} \tag{13}$$

for arbitrary initial choices of $\boldsymbol{A}^0$ and $\boldsymbol{V}^0$. Here, $\boldsymbol{V}$ is the dual variable, and $\mu > 0$ is a penalty parameter. The operator $\mathcal{S}_{\lambda\mu}$ is the soft-thresholding function [11] and it can computed analytically in each entry using the rule $\left[\mathcal{S}_\tau\left(\boldsymbol{X}\right)\right]_{ij} = \text{sgn}(\boldsymbol{X}_{ij})\left(|\boldsymbol{X}_{ij}| - \tau\right)_+$, where sgn is the signal function, and $x_+ = \max\{x, 0\}$. Similarly, the Singular Value Thresholding (SVT) operator $\mathcal{D}$ is defined by $\mathcal{D}_\mu(\boldsymbol{X}) = \boldsymbol{U}S_\mu(\boldsymbol{\Sigma})\boldsymbol{V}^T$, where $\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ is any SVD of $\boldsymbol{X}$ [11].

Ideally, after computing some iterations of PCP, the gross anomalies of $\boldsymbol{H}$ are captured in the $\boldsymbol{A}$ component, and $\boldsymbol{B}$ is a low rank reconstruction of $\boldsymbol{H}$. Therefore, we extract the robust eigenkernels of $\boldsymbol{H}$ by computing the original eigenkernels from $\boldsymbol{B}$ as discussed in the last section.

IV. Numerical experiments

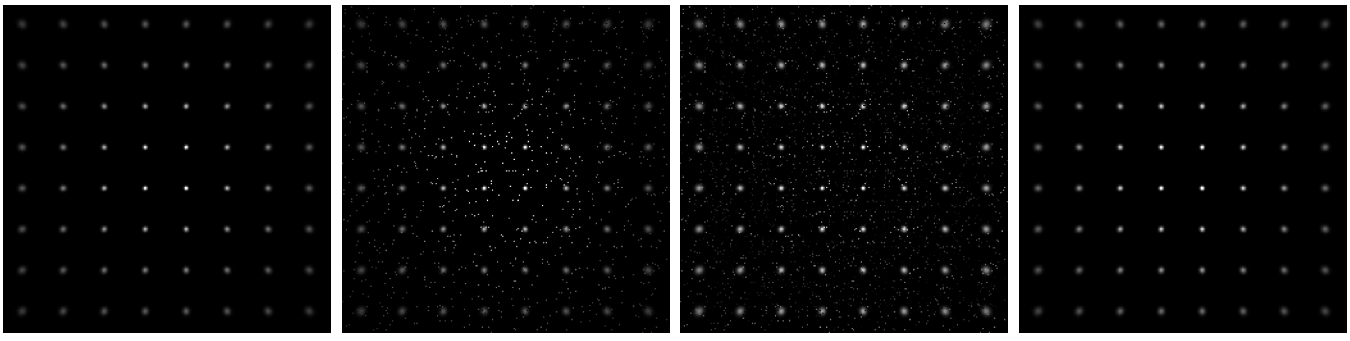
One of the simplest and most common approaches [1] to model the blur produced by an optical system is to assume that its PSF is a stationary centered 2d Gaussian kernel like

$$\omega(a, b) = \frac{1}{2\pi\sigma_a\sigma_b}\exp\left[\frac{(a - a_0)^2}{2\sigma_a^2} + \frac{(a - b_0)^2}{2\sigma_b^2}\right], \tag{14}$$

where (a, b) are rectangular indexes used to describe the kernel function and (a_0, b_0) is the center of the kernel image. A simple nonstationary model can also be described using Gaussian kernels. For instance, we are going to consider a radial nonstationary model where the kernel function at each set of centered polar indices (θ, r) is given by

$$h^{(\theta,r)}(a, b) = \mathcal{A}_\theta\left[\tilde{\omega}^r(a, b)\right], \tag{15}$$

where $\mathcal{A}_\theta$ is a rotation of θ, and $\tilde{\omega}^r$ is a reparametrized version of ω where the standard deviation in each axis is a function of r calculated using $\sigma_a(r) = \alpha_a + \beta_a r^{\gamma_a}$ and $\sigma_b(r) = \alpha_b + \beta_b r^{\gamma_b}$ with constants $\alpha_a, \alpha_b, \beta_a, \beta_b, \gamma_a, \gamma_b \geq 0$. Since each kernel is a density function, after sampling it into a 2d discrete signal, the sum of its absolute values is normalized to one.



(a) (b) (c) (d)

Fig. 2: Nonstationary PSF reconstruction discussed in Section IV-A. (a) Original noiseless PSF. (b) Original PSF corrupted with salt and pepper noise. (c) PCA reconstruction (mean PSNR= 13.799 dB). (d) PCP-RPCA reconstruction (mean PSNR= 46.301 dB).

A. *Reconstruction of corrupted PSF*

Usually, the nonstationary PSF of an optical system is estimated for a lattice of indices $J \subset I$ by combining blurred and sharp versions of an image containing a grid of the points in J. The estimated kernels are usually corrupted by noise. Thus, in this experiment we compare the performance of the original approach [2] and our robust proposal to model a nonstationary PSF in which every sampled kernel is corrupted with salt-and-pepper noise.

The low rank PSF $\boldsymbol{B}$ was generated using the rotated Gaussian model (15) with parameters $\alpha_a = 0.07, \alpha_b = 0.07, \beta_a = 0.4, \beta_b = 0.3, \gamma_a = 1, \gamma_b = 1$. The PSF was generated to match an image shape of 256×256 pixels, and sampled into a grid of 8×8 kernels, each having 32×32 pixels. Each sampled kernel was contaminated with an equaly likelly salt and pepper noise saturating $3\ \%$ of its pixels with the maximum or the minimum pixel values from that single kernel. The noiseless PSFs $\boldsymbol{B}$ are shown in Figure 2 (a) and their corrupted versions $\boldsymbol{H} = \boldsymbol{B} + \boldsymbol{A}$ are shown im Figure 2 (b).

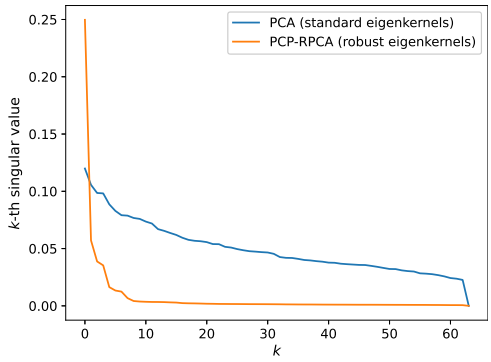
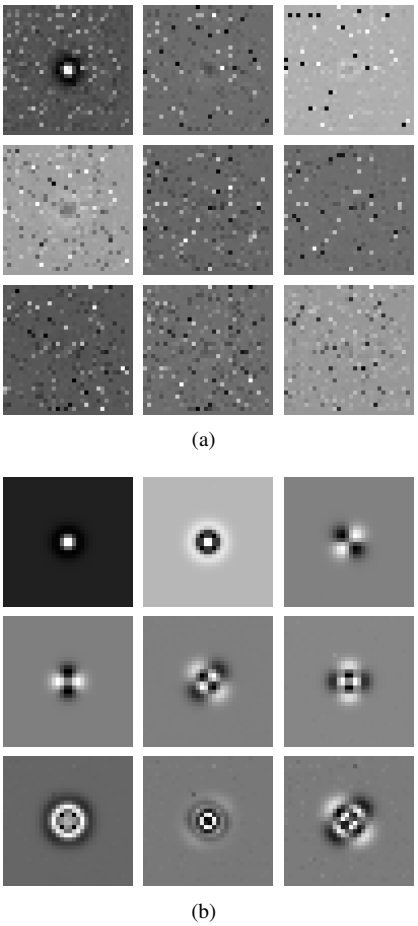


Fig. 3: Singular values obtained via standard PCA and PCP-RPCA.



(a)

(b)

Fig. 4: First 9 eigenkernels recovered from the corrupted PSF. (a) Original proposal with PCA. (b) Robust proposal with PCP-RPCA.

(a) (b)

Fig. 5: Nonstationary blur produced by (a) the original PCA model (PSNR $=$ 16.144 dB) and (b) our robust proposal (PSNR $=$ 57.189 dB). Both models estimated from a corrupted PSF, and PSNR values calculated with respect to a noiseless PCA blur model.

We evaluated PCA and PCP-RPCA in the corrupted PSF $\boldsymbol{H}$, and obtained the set of original and robust eigenkernels, each containing 64 eigenkernels. We used 500 iterations of the PCP iteration (13) and the sparsity parameter was empirically set as $\lambda = 0.149$. The singular values of each SVD run are shown in Figure 3. Also, the first 9 eigenkernels provided by each method are shown in Figure 4 (a) and (b).

At this point, we already see a glimpse of the impact due to PCP-RPCA in recovering the original PSF from the corrupted data: As displayed in the scree plot of Figure 3, the sampled PSF information is spread across all 64 singular values of the original reconstruction framework. In contrast, our robust proposal shows a fast decaying pattern, which is a desirable property that means that the first robust eigenkernels are more capable of representing the data. Also, the original PCA eigenkernels are contamined by noise, while the robust eigenkernels are much clearer and resemble circular harmonic modes, which is a also expected based on the Gaussian nature of the PSF.

The reconstruction of all 64 sampled kernels are shown in Figures 2 (c) and (d) using standard PCA and our robust proposal. In both cases we used all 64 eigenkernels for a fair comparison. We observe that the robust reconstruction looks almost the same as the original PSF, achieving a mean PSNR of 79.357 dB with respect to the original noiseless kernels. On the other hand, the reconstruction PCA resulted in a corrupted PSF with mean PSNR of 18.231 dB.

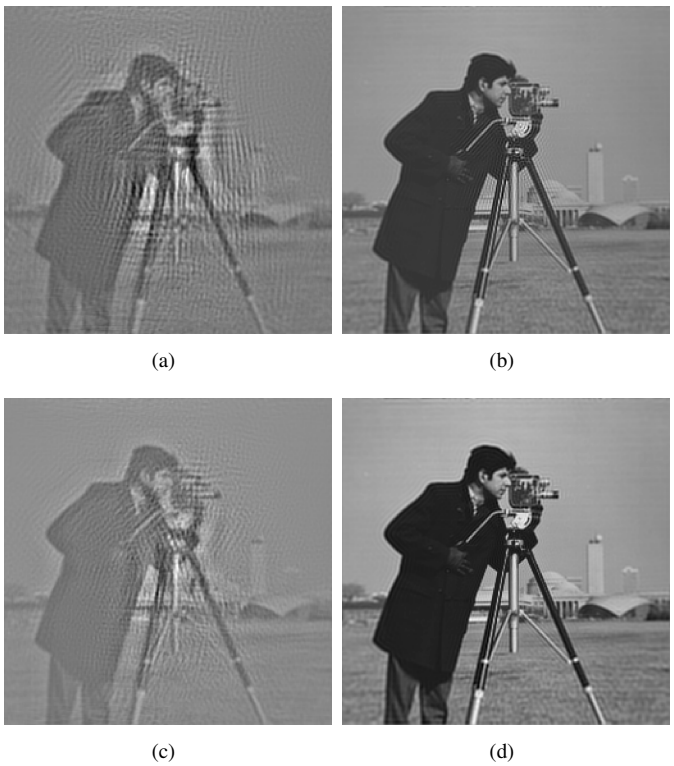
We also used both recovered PSFs to blur the cameraman image from Fig. 1. The resulting blurred images, using PCA and PCP-RPCA eigenkernels are shown in Fig. 5. To isolate the noise effect in this comparison, we calculated the PSNR of these two blurred images with respect to a ground-truth image blurred by the original eigenkernels model recovered from the

sampled PSF without salt and pepper noise $\boldsymbol{B}$ (Fig. 1 (d)).

B. Deblurring with robust eigenkernels

The main goal of improving PSF models is image restoration. As discussed earlier, the eigenkernels framework provides a simple way to adapt existing deblurring algorithms to the nonstationary case. Here, we present a restoration example using the PSFs obtained from the eigenkernels from the subsection IV-A.

As seen before, the robust eigenkernels provided a good approximation to the sampled kernels even in the presence of moderate level salt and pepper noise. In this experiment, we used the noiseless eigen PSF to generate the observed image $\boldsymbol{y}$, shown in Fig. 1(c), and applied 150 conjugate gradient iterations (11). We ran this experiment using the noisy PCA eigenkernels (Fig.4 (a)), and then using our robust proposal (Fig.4 (b)). In all cases we adopted $\lambda = 10^{-8}$.



(a) (b) (c) (d)

Fig. 6: Deblurring with robust eigenkernels exctracted from 8×8 sampled kernels (a) PCA (PSNR $= 16.146$ dB). (b) PCP-RPCA (PSNR $= 24.653$ dB). Repeated experiment with 16×16 sampled kernels. (c) PCA (PSNR $= 15.454$ dB). (d) PCP-RPCA (PSNR $= 56.632$ dB).

The deblurred images with the PCA and RPCA eigenkernels are shown in Fig. 6 (a) and 6 (b), respectively. As expected, the robust recovery is better than the original one, both from a qualitative and quantitative perspectives, with an increase of 8.507 dB in terms of PSNR. Again, to isolate the noise attenuation effects, the PSNR values were calculated from a deblurred image using the original eigenkernels model sampled from the original PSF without salt and pepper noise.

Although the robust recovery resulted in a good estimate of $\boldsymbol{x}$ even in the presence of salt and pepper noise, a careful look at it shows that its dynamic range is reduced with respect to the ground truth. To address this issue, we repeated the experiment with the same noise distribution but using 16×16 sampled kernels instead of 8×8. The PCA and PCP-RPCA recoveries are depicted in Figs. 6 (c) and (d), respectively. Comparing these to Figs. 6 (a) and (b), we see that the loss in dynamic range occured due to suboptimal generalization of PCP-RPCA in the presence of gross noise, and it can be overcome by increasing the sampling rate of the noisy PSF. Note that increasing the sampling rate does not improve the dynamic range of the PCA recovery.

V. CONCLUSIONS

In this paper we follow the work of Gwak and Yang [2], improving their eigenkernels model for nonstationary blur to handle sparse anomaly corruption in the sampled PSF. Specifically, we propose the use of PCP-RPCA instead of the standard PCA in the original framework, due to its outlier suppression properties. We compared both strategies with numerical examples regarding the eigenkernel extraction of a salt and pepper noise corrupted PSF and the indirect effects observed on a toy restoration algorithm. In both cases, we see that the standard PCA provides poor results in the presence of outliers, while PCP-RPCA is capable of mantaining the good performance of the original eigenkernels modeling, resulting in anomalies that are greatly attenuated.

REFERENCES

[1] E. Kee, S. Paris, S. Chen, and J. Wang, "Modeling and removing spatially-varying optical blur," in *2011 IEEE International Conference on Computational Photography (ICCP)*, pp. 1–8, Apr. 2011.

[2] M. Gwak and S. Yang, "Modeling nonstationary lens blur using eigen blur kernels for restoration," *Optics Express*, vol. 28, p. 39501, Dec. 2020.

[3] F. Heide, M. Rouf, M. B. Hullin, B. Labitzke, W. Heidrich, and A. Kolb, "High-quality computational imaging through simple lenses," *ACM Transactions on Graphics*, vol. 32, no. 5, pp. 149:1–149:14, 2013.

[4] H. Tang and K. N. Kutulakos, "What does an aberrated photo tell us about the lens and the scene?," in *IEEE International Conference on Computational Photography (ICCP)*, pp. 1–10, Apr. 2013.

[5] R. Bamler and J. Hofer-Alfeis, "2D linear space-variant processing by coherent optics: A sequence convolution approach," *Optics Communications*, vol. 43, pp. 97–102, Sept. 1982.

[6] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Mathematics of Computation*, vol. 19, no. 90, pp. 297–301, 1965.

[7] A. Chambolle, "An Algorithm for Total Variation Minimization and Applications," *Journal of Mathematical Imaging and Vision*, vol. 20, pp. 89–97, Jan. 2004.

[8] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, pp. 259–268, Nov. 1992.

[9] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of the ACM*, vol. 58, pp. 11:1–11:37, June 2011.

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, pp. 1–122, July 2011. Publisher: Now Publishers, Inc.

[11] N. Parikh and S. Boyd, "Proximal Algorithms," *Foundations and Trends® in Optimization*, vol. 1, pp. 127–239, Jan. 2014. Publisher: Now Publishers, Inc.