A Generalized Additive Convolution Model for Efficient Deblurring of Camera Shaken Image

Hong Deng, Dongwei Ren, Hongzhi Zhang,
Kuanquan Wang, and Wangmeng Zuo^(✉)

School of Computer Science and Technology,
Harbin Institute of Technology, Harbin, China
cswmzuo@gmail.com

Abstract. Image blur caused by camera shake is often spatially variant, which makes it more challenging to recover the latent sharp image. Geometrical camera shake model based non-uniform deblurring methods, modeling the blurry image as a weighted summation of the homographically transformed images of the latent sharp image, although can achieve satisfactory deblurring results, still suffer from the problems of heavy computation or extensive memory cost. In this paper, we propose a generalized additive convolution (GAC) model for efficient non-uniform deblurring. A camera motion trajectory can be decomposed into a compact set of in-plane translations (slice) and roll rotations (fiber), which with an insightful analysis can both be formulated as convolution. The GAC model provides a promising way to overcome the difficulties of both computational load and memory burden. The experimental results show that GAC can obtain satisfactory deblurring results, and is much more efficient than state-of-the-arts.

Keywords: Image deblurring · Camera shake · Non-uniform deblurring

1 Introduction

Camera shake blur, as one of the major causes that ruin a photograph taken in low-light condition, is often inevitable, thus having caught research interests for several decades. To simplify the problem, earlier researches usually assume that the blur is spatially invariant, and that blur procedure is modeled as convolution of blur kernel and sharp image. Thus a large number of deblurring methods formulated as blind deconvolution have come forward [2, 4, 13, 14, 16, 17, 22].

However, image blur caused by camera shake is generally spatially variant [9, 12, 23], which is far from simple convolution, and much more difficult to model, making the removal of camera shake blur much more challenging than uniform blind deconvolution. Much ahead of designing non-uniform deblurring algorithms, it is firstly in dire need to properly model non-uniform blur. On one hand, a direct strategy is to divide blurry image into several regions, within which is assumed to take uniform blur, and thus local blind deconvolution can be

Fig. 1. The results of different deblurring approaches on a real camera shaken image. (a) The input image with size 512×768 , (b) Close-ups of (a). Deblurring results by (c) Whyte et al. [20], (96833s, 0.98GB), (d) Gupta et al. [5] (3614s, 13.75GB), (e) Hu et al. [10] (3583s, 10.09GB), (f) Xu et al. [21] (1582s, 2.95GB), and (g) GAC (659s, 2.05GB). The proposed GAC method can achieve satisfactory results and is much more efficient.

performed to estimate the kernels [11]. However, uniform blur assumption within small regions does not often hold exactly, producing visual deblurring artifacts especially at region boundaries. Most importantly, this approach neglects the global constraints on local kernels based on camera motion during exposure. On the other hand, based on the projective motion path blur model (PMPB) [19], a physically plausible model was developed to globally model non-uniform blur, where blurry image is modeled as a weighted summation of the homographically transformed versions of the latent sharp image in the 6D camera motion space, including x -, y -, z -axes rotation, and x -, y -, z -axes translation.

However, the full 6D model involves large amount of variables to estimate, restricting it as a reasonable choice in practice. Thus two simplified 3D homography based models have been developed to approximate the 6D camera motion space, and been validated that both can well approximate the original 6D model [12]. Specifically, Whyte et al. [20] suggested to adopt x -(pitch), y -(yaw), z -(roll) axes rotation, while x - and y -axes (in-plane) translation and z -axis rotation are used by Gupta et al. [5]. Unfortunately, even with promising motion space reduction, the simplified 3D homography-based methods [5, 20] still suffer from the computational inefficiency and huge memory burden, restricting the wide applications of the camera shake removal methods. For example, the method proposed by Whyte et al. [20] usually requires hours to process a 300×400 image. To achieve high computational efficiency, Gupta et al. [5] pre-computed a sparse matrix for each homographical transform, while Hu and Yang [10], with elimination of inactive poses over iterations, further constrained the possible camera pose in a low dimensional subspace. Actually, these methods [5, 10] relaxed the computational inefficiency at the cost of increasing memory burden to store the sparse matrices. By incorporating camera shake with the efficient filter flow (EFF) framework [8], Hirsch suggested an interesting forward model [7] to construct local uniform blur models guided by the global constraints on camera motions, and gained considerable speed-ups, however it actually adopted a locally uniform approximation of the geometrical models.

In this paper, we propose an efficient generalized additive convolution (GAC) model for direct computation of the global geometrical camera shake model.

With a insightful analysis of in-plane translations (slice) and z -axis rotation (fiber), we prove that any camera motion trajectory can be decomposed into a set of slices and fibers, which actually can be formulated as convolution. Given a specific roll angle, the in-plane translations of the sharp image can be equivalently formulated as the rotation of a convolution image, while given a specific in-plane translation, roll rotations of the sharp image can be equivalently formulated as the inverse polar transform of the convolution of a kernel with the transformed sharp image. Thus, given a camera motion trajectory, we propose a greedy algorithm to generate GAC, which only requires several fast Fourier transform (FFT) and a number of pixel-wise operations. Moreover, GAC also concurrently reduces the memory complexity, since only the basis kernels and a number of rotation matrices are required to be pre-stored. Experimental results show that the proposed GAC-based deblurring method can achieve comparable or better deblurring results than the competing methods (See Fig. 1), has a much lower peak memory usage than [5, 10], and is much more efficient than the state-of-the-art camera shake removal methods [5, 7, 10, 20, 21].

The contributions of this paper can be summarized from three aspects:

- We develop a generalized additive convolution (GAC) framework together with two GAC models, slice-based and fiber-based, for forward modeling of camera shake blur, concurrently reducing the computational and memory complexity.
- In order to further reduce the number of basis convolution kernels, a greedy algorithm is proposed to generate hybrid (slice and fiber mixed) GAC from any camera motion trajectory, resulting in more compact and efficient GAC for deblurring.
- By incorporating the GAC with the existing non-uniform deblurring model, a GAC-based method is provided to remove camera shake from a single image. Experimental results validate that the GAC-based method can obtain comparable deblurring results with improved efficiency and significantly reduced memory usage compared with state-of-the-arts [5, 7, 10, 20, 21].

The rest of the paper is organized as follows. Section 2 briefly reviews the geometrical model of camera shake blur. Section 3 presents the proposed GAC model, together with a greedy algorithm for hybrid GAC parameterization, and the GAC-based non-uniform deblurring method. Section 4 demonstrates experimental results, and finally Sect. 5 concludes this paper.

2 Prerequisites

With the projective motion path blur (PMPB) model, blurry image is a weighted summation of all possible homographically transformed versions of the latent sharp image, which is faithful to real world camera motion. In [5, 10], the full 6D camera pose space is approximated by the 3D subspace of roll rotation and in-plane translations and roll rotation,

$$\boldsymbol{\theta} = (\theta_z, t_x, t_y) \quad (1)$$

where θ_z is the roll angle, t_x and t_y are the translations along x - and y - axes, respectively. For each pose in the 3D camera subspace $\theta_j = (\theta_{z,j}, t_{x,j}, t_{y,j})$, By defining,

$$\begin{aligned} \mathbf{M}_{\theta_j} &= \left(\mathbf{R}_{\theta_{z,j}} + \frac{1}{d} * \mathbf{t}_j \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \right) \\ &= \begin{bmatrix} \cos(\theta_{z,j}) - \sin(\theta_{z,j}) \frac{1}{d} * t_{x,j} & \sin(\theta_{z,j}) \frac{1}{d} * t_{x,j} & 0 \\ \sin(\theta_{z,j}) \frac{1}{d} * t_{y,j} & \cos(\theta_{z,j}) \frac{1}{d} * t_{y,j} & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned} \quad (2)$$

the homography [18] can be defined as

$$\mathbf{H}_{\theta_j} = \mathbf{C} \mathbf{M}_{\theta_j} \mathbf{C}^{-1} \quad (3)$$

where $\mathbf{R}_{\theta_{z,j}}$ is the rotation matrix, $\mathbf{t}_j = [t_{x,j}, t_{y,j}, 1]^T$ is the translation vector, d is the depth of the scene, and $\mathbf{C}$ is the camera calibration matrix.

Given $\mathbf{H}_{\theta_j}$, the corresponding warping matrix $\mathbf{K}_{\theta_j}$ can be constructed (Please refer to [3, 5, 10] for details of construction of $\mathbf{K}_{\theta_j}$ and definition of calibration matrix $\mathbf{C}$). Now the forward blur model can be formulated as

$$\mathbf{y} = \sum_{\theta_j} w_{\theta_j} \mathbf{K}_{\theta_j} \mathbf{x} + \mathbf{v} \quad (4)$$

where w_{θ_j} is the fraction of time the camera spent at pose θ_j , $\mathbf{y} \in \mathbb{R}^{n \times n}$ is the blurry image, $\mathbf{x} \in \mathbb{R}^{n \times n}$ is the latent sharp image, and $\mathbf{v}$ denotes the additive Gaussian white noise.

Deng et al. [3] further suggested an additive convolution model, where each homography is formulated as convolution operation,

$$\mathbf{y} = \sum_{\theta_j} \alpha_{\theta_j} \circ (\mathbf{x} \otimes \mathbf{k}_{\theta_j}) + \mathbf{v} \quad (5)$$

where $\circ$ and $\otimes$ denote entry-wise multiplication and convolution, $\mathbf{k}_{\theta_j}$ is basis convolution kernel corresponding to pose θ_j , and α_{θ_j} is weighting matrix. The computational complexity of additive convolution based methods are only $O(Cn^2 \log_2 n)$, where n^2 is the number of pixels of the image $\mathbf{x}$ and C is the number of the basis convolution kernels. For non-blind deconvolution problem, principal component analysis (PCA) is adopted to learn the basis convolution kernels in advance [3]. However, when applied to blind non-uniform deblurring, both the basis kernels and the weighted maps should be updated in each iteration, making it unpractical in blind deblurring of camera shaken images.

In this paper, we propose a novel generalized additive convolution (GAC) model for efficient non-uniform deblurring, where with specific design, basis convolution kernels and the weighted maps can be efficiently constructed from the camera motion trajectory.

3 Generalized Additive Convolution Model for Removal of Camera Shake

The generalized additive convolution (GAC) model is defined as

$$\mathbf{y} = \sum_{\theta_j} \mathbf{f}_{\theta_j} \left(\mathbf{g}_{\theta_j}(\mathbf{x}) \otimes \mathbf{k}_{\theta_j} \right) \quad (6)$$

where $\mathbf{f}_{\theta_j}$ and $\mathbf{g}_{\theta_j}$ are two pixel-wise operators. GAC is obviously the generalization of the additive convolution model (5) by defining $\mathbf{f}_{\theta_j}(\mathbf{x}) = \boldsymbol{\alpha}_{\theta_j} \circ \mathbf{x}$ and $\mathbf{g}_{\theta_j}(\mathbf{x}) = \mathbf{x}$. Moreover, since both $\mathbf{f}_{\theta_j}$ and $\mathbf{g}_{\theta_j}$ are pixel-wise operators, GAC still has the computational complexity $O(Cn^2 \log_2 n)$. Thus, the key issue of GAC is to properly set $\mathbf{k}_{\theta_j}$, $\mathbf{f}_{\theta_j}$ and $\mathbf{g}_{\theta_j}$ to reduce the C value. In the following, we will show that, with carefully designed $\mathbf{k}_{\theta_j}$, $\mathbf{f}_{\theta_j}$ and $\mathbf{g}_{\theta_j}$, we can derive an explicit and efficient scheme for the forward computation of the blurry image caused by camera shake.

3.1 Slice and Fiber

To give the definition of slice and fiber, we first define $\mathbf{K}_{\theta_{j,r}}$, $\mathbf{K}_{\theta_{j,t}}$ as warping matrix of rotation and translation, respectively,

$$\mathbf{K}_{\theta_{j,r}} \mathbf{x} = \mathbf{R}_{\theta_{z,j}}(\mathbf{x}) \quad (7)$$

$$\mathbf{K}_{\theta_{j,t}} \mathbf{x} = \mathbf{k}_{t_j} \otimes \mathbf{x} \quad (8)$$

where $\mathbf{R}_{\theta_{z,j}}$ denotes the pixel-wise image rotation operation, and the translation convolution kernel is defined as:

$$\mathbf{k}_{t_j}(x, y) = \begin{cases} 1, & \text{if } x = t_{x,j} \text{ and } y = t_{y,j} \\ 0, & \text{else} \end{cases} \quad (9)$$

Given pose θ_j , $\mathbf{K}_{\theta_j} \mathbf{x}$ actually first translates the image $\mathbf{x}$ by $t_{x,j}$ and $t_{y,j}$ along x -axis and y -axis respectively, and then rotates the translated image by roll angle $\theta_{z,j}$, i.e., $\mathbf{K}_{\theta_j} = \mathbf{K}_{\theta_{j,r}} \mathbf{K}_{\theta_{j,t}}$. Thus, the non-uniform blur caused by camera shake can be formulated as:

$$\mathbf{y} = \mathbf{K} \mathbf{x} = \sum_{\theta_j} w_{\theta_j} \mathbf{K}_{\theta_j} \mathbf{x} = \sum_{\theta_j} w_{\theta_j} \mathbf{K}_{\theta_{j,r}} \mathbf{K}_{\theta_{j,t}} \mathbf{x} \quad (10)$$

It should be noted that, the poses along a real camera motion trajectory form a connected 1D path, and thus the weights of most poses are zeros. So in Eq. (10) we can only consider the subset of poses with positive weights, $\mathcal{P} = \{\theta_j : w_{\theta_j} > 0\}$.

Finally, we define two special classes of subsets of $\mathcal{P}$: slice and fiber. A slice $\mathcal{S}_\theta$ is defined as $\mathcal{S}_\theta = \{\theta_j = \{\theta_{z,j}, t_{x,j}, t_{y,j}\} : \theta_j \in \mathcal{P} \text{ and } \theta_{z,j} = \theta\}$, while a fiber $\mathcal{F}_t$ is defined as $\mathcal{F}_t = \{\theta_j = \{\theta_{z,j}, t_{x,j}, t_{y,j}\} : \theta_j \in \mathcal{P} \text{ and } (t_{x,j}, t_{y,j}) = t\}$. Actually, $\mathcal{P}$ can be decomposed into a number of non-intersected slices and fibers. In the following, we will introduce how to construct slices and fibers from $\mathcal{P}$ and how to reformulate camera shake as the GAC model.

3.2 Slice Based GAC

Given a slice $\mathcal{S}_\theta$, the blur caused by camera motion within $\mathcal{S}_\theta$ can be formulated as

$$\begin{aligned} \mathbf{K}_\theta \mathbf{x} &= \sum_{\theta_j \in \mathcal{S}_\theta} w_{\theta_j} \mathbf{K}_{\theta_j, r} \mathbf{K}_{\theta_j, t} \mathbf{x} \\ &= \sum_{\theta_j \in \mathcal{S}_\theta} w_{\theta_j} \mathbf{R}_\theta(\mathbf{k}_{t_j} \otimes \mathbf{x}) \\ &= \mathbf{R}_\theta \left(\left(\sum_{\theta_j \in \mathcal{S}_\theta} w_{\theta_j} \mathbf{k}_{t_j} \right) \otimes \mathbf{x} \right) \\ &= \mathbf{R}_\theta(\mathbf{k}_{\mathcal{S}_\theta} \otimes \mathbf{x}) \end{aligned} \quad (11)$$

where $\mathbf{k}_{\mathcal{S}_\theta}$ denotes the slice kernel with respect to the roll angle θ .

For a general camera motion trajectory $\mathcal{P}$, we first classify the poses in $\mathcal{P}$ into a number of non-intersected slices with $\mathcal{P} = \cup_\theta \{\mathcal{S}_\theta\}$, and then the non-uniform blur in Eq. (10) can be equivalently reformulated as

$$\mathbf{K} \mathbf{x} = \sum_\theta \mathbf{R}_\theta(\mathbf{k}_{\mathcal{S}_\theta} \otimes \mathbf{x}) \quad (12)$$

It is obvious that this is a GAC model with $\mathbf{f}_\theta(\mathbf{x}) = \mathbf{R}_\theta(\mathbf{x})$ and $\mathbf{g}_\theta(\mathbf{x}) = \mathbf{x}$. If the range of the roll angles is discretized into n_z intervals, we can see that the number of basis filters should be not higher than n_z .

Similarly, we can define the adjoint operator $\mathbf{K}^T \mathbf{y}$ as

$$\mathbf{K}^T \mathbf{y} = \sum_\theta \tilde{\mathbf{k}}_{\mathcal{S}_\theta} \otimes \mathbf{R}_\theta^T(\mathbf{y}) \quad (13)$$

where $\tilde{\mathbf{k}}_{\mathcal{S}_\theta}$ is the adjoint operator of $\mathbf{k}_{\mathcal{S}_\theta}$ constructed by flipping the $\mathbf{k}_{\mathcal{S}_\theta}$ upside-down and left-to-right, and $\mathbf{R}_\theta^T$ is the strict adjoint operator of the discrete version of $\mathbf{R}_\theta$. It is obvious that the adjoint operator in Eq. (13) is also a GAC model and can be efficiently computed.

3.3 Fiber Based GAC

Given a fiber $\mathcal{F}_t$ with the same translation $\mathbf{t} = [t_x, t_y]^T$, the non-uniform blur caused by the camera motion along $\mathcal{F}_t$ can be formulated as

$$\begin{aligned} \mathbf{K}_t \mathbf{x} &= \sum_{\theta_j \in \mathcal{F}_t} w_{\theta_j} \mathbf{K}_{\theta_j, r} \mathbf{K}_{\theta_j, t} \mathbf{x} \\ &= \sum_{\theta_j \in \mathcal{F}_t} w_{\theta_j} \mathbf{R}_{\theta_{z,j}}(\mathbf{K}_{\theta_{j,t}} \mathbf{x}) \\ &= IPT(\mathbf{w}_t \otimes PT(\mathbf{K}_{\theta_t} \mathbf{x})) \end{aligned} \quad (14)$$

where $\theta_t = [0, t_x, t_y]^T$, $\mathbf{K}_{\theta_t} \mathbf{x}$ denotes the in-plane translation operation, $PT(\cdot)$ and $IPT(\cdot)$ stand for the polar transform and inverse polar transform [15], respectively. In the polar transform, we use the same interval to discretize angular and roll angles, and thus the basis filter $\mathbf{w}_t$ can be defined as,

$$\mathbf{w}_t = [w_{t, \theta_1}, w_{t, \theta_1}, \dots, w_{t, \theta_{n_z}}] \quad (15)$$

where θ_1 is the minimal roll angle and θ_{n_z} is the maximal roll angle.

Fig. 2. Decomposing a camera motion trajectory into a slice set and a fiber set.

For a general camera motion trajectory $\mathcal{P}$, we first classify the poses in $\mathcal{P}$ into a number of non-intersected fibers with $\mathcal{P} = \cup_t \{\mathcal{F}_t\}$, and then the non-uniform blur in Eq. (10) can be equivalently reformulated as

$$\mathbf{K}\mathbf{x} = \sum_t \mathbf{K}_t \mathbf{x} = \sum_t IPT(\mathbf{w}_t \otimes PT(\mathbf{K}_{\theta_t} \mathbf{x})) \quad (16)$$

It is obvious that this is a GAC model with $\mathbf{f}_t(\mathbf{x}) = IPT(\mathbf{x})$ and $\mathbf{g}_t(\mathbf{x}) = PT(\mathbf{K}_{\theta_t} \mathbf{x})$. If the range of the in-plane translations is discretized into $n_x \times n_y$ intervals, then the number of basis filters should be not higher than $n_x n_y$.

We can then define the adjoint operator $\mathbf{K}^T \mathbf{y}$ as

$$\mathbf{K}^T \mathbf{y} = \sum_t \mathbf{K}_{-\theta_t} (PT^T(\tilde{\mathbf{w}}_t \otimes IPT^T(\mathbf{x}))) \quad (17)$$

where $\tilde{\mathbf{w}}_t$ is the adjoint operator of $\mathbf{w}_t$, PT^T and IPT^T are the adjoint operators of PT and IPT , respectively.

3.4 Hybrid GAC for Modeling

For GAC, the key to save computational cost is to reduce C , the number of basis filters. Given a general camera motion trajectory as shown in Fig. 2, neither pure slice-based nor pure fiber-based GAC can guarantee sufficiently small value of C , while in a hybrid (slice and fiber mixed) decomposition, only 2 slices and 2 fibers are required to model the camera motion trajectory, so that the computational complexity is significantly reduced. Thus, we propose a greedy method to decompose a camera motion trajectory into a hybrid set of slices and fibers to reduce the C value. The experimental results also validates that the hybrid GAC performs better than either slice based GAC or fiber based GAC by providing both high efficiency and better visual performance.

Given the pose subset $\mathcal{P}$ and the 3D weight matrix $\mathbf{W}$ with $\mathbf{W}(\theta_z, t_x, t_y)$ the weight of the pose $\theta = (\theta_z, t_x, t_y)$, in each iteration, the proposed method

first finds a candidate slice $\mathcal{S}_{\hat{\theta}_z}$ and a candidate fiber $\mathcal{F}_{\hat{t}_x, \hat{t}_y}$, compare their relative contributions, and then choose the slice or fiber with higher weights. By this way, we can obtain a slice set $\{(\theta_j, \mathcal{S}_j, \mathbf{k}_j) : j = 1, \dots, n_s\}$ and a fiber set $\{(t_i, \mathcal{F}_i, \mathbf{w}_j) : i = 1, \dots, n_f\}$. As shown in Fig. 2, the proposed greedy algorithm can provide the correct compact decomposition of the camera motion trajectory. The detailed algorithm is summarized in Algorithm 1.

Based on the slice set and the fiber set, the non-uniform blur caused by camera shake can be reformulated as:

$$\mathbf{K}\mathbf{x} = \sum_{i=1}^{n_f} IPT(\mathbf{w}_i \otimes PT(\mathbf{K}_{\theta_{t_i}} \mathbf{x})) + \sum_{j=1}^{n_s} \mathbf{R}_{\theta_j}(\mathbf{k}_j \otimes \mathbf{x}) \quad (18)$$

The adjoint operator $\mathbf{K}^T \mathbf{y}$ is then defined as,

$$\mathbf{K}^T \mathbf{y} = \sum_{i=1}^{n_f} \mathbf{K}_{-\theta_i} PT^T(\tilde{\mathbf{w}}_i \otimes IPT^T(\mathbf{y})) + \sum_{j=1}^{n_s} \tilde{\mathbf{k}}_j \otimes \mathbf{R}_{\theta_j}^T(\mathbf{y}) \quad (19)$$

Algorithm 1. Greedy algorithm for hybrid GAC

Input: $\mathcal{P}$ and $\mathbf{W}$

Output: The fiber set $\{(t_i, \mathcal{F}_i, \mathbf{w}_j) : i = 1, \dots, n_f\}$ and the slice set $\{(\theta_j, \mathcal{S}_j, \mathbf{k}_j) : j = 1, \dots, n_s\}$.

- 1: $\mathbf{WS}(\theta_z) = \sum_{t_x, t_y} \mathbf{W}(\theta_z, t_x, t_y)$, $e = e_0 = \sum_{\theta_z} \mathbf{WS}(\theta_z)$, $\mathbf{WF}(t_x, t_y) = \sum_{\theta_z} \mathbf{W}(\theta_z, t_x, t_y)$, $n_s = 0$, $n_f = 0$, ε , $0 < \alpha$, null fiber set and null slice set.
 - 2: **while** $e/e_0 > \varepsilon$ **do**
 - 3: $(\hat{t}_x, \hat{t}_y) = \arg \max_{t_x, t_y} \mathbf{WF}(t_x, t_y)$ and $w_f = \mathbf{WF}(\hat{t}_x, \hat{t}_y)$.
 - 4: $\hat{\theta}_z = \arg \max_{\theta_z} \mathbf{WS}(\theta_z)$, $w_s = \mathbf{WS}(\hat{\theta}_z)$.
 - 5: Update the slice set: If $w_s \geq \alpha w_f$, add the $\hat{\theta}_z^{th}$ slice to the slice set, and modify $\mathbf{WS}(\hat{\theta}_z) = 0$, $\mathbf{WF}(t_x, t_y) = \mathbf{WF}(t_x, t_y) - \mathbf{W}(\hat{\theta}_z, t_x, t_y)$, and $n_s = n_s + 1$.
 - 6: Update the fiber set: If $w_s < \alpha w_f$, add the $(\hat{t}_x, \hat{t}_y)^{th}$ fiber to the fiber set, modify $\mathbf{WF}(\hat{t}_x, \hat{t}_y) = 0$, $\mathbf{WS}(\hat{\theta}_z) = \mathbf{WS}(\hat{\theta}_z) - \mathbf{W}(\hat{\theta}_z, \hat{t}_x, \hat{t}_y)$, and $n_f = n_f + 1$.
 - 7: Update $e = \sum_{\theta_z} \mathbf{WS}(\theta_z)$.
 - 8: $k \leftarrow k + 1$
 - 9: **end while**
-

With FFT, the computational complexity of the model in Eq. (18) is $O((n_s + n_f)n^2 \log_2 n)$. If n_s and n_f are small, GAC would be more efficient than the other methods. Let n_z be the number of intervals for the roll angle. It is reasonable to assume that $(n_s + n_f) < n_z$, otherwise, we can use the pure slice-based GAC model. To further improve the computational efficiency, the look up table (LUT) method can be adopted for fast image rotation, polar and inverse polar transform, and $n_z + 2$ LUTs should be pre-computed and stored in memory.

3.5 GAC-based Non-uniform Deblurring

GAC is a computational model of non-uniform blur caused by camera shake, and can be incorporated with the existing image priors and optimization algorithms

for deblurring. Typical non-uniform deblurring model can be formulated as:

$$\min_{\mathbf{W}, \mathbf{x}} \left\| \sum_{\theta \in \mathcal{P}} w_{\theta} \mathbf{K}_{\theta} \mathbf{x} - \mathbf{y} \right\|^2 + \Phi_1(\mathbf{x}) + \Phi_2(\mathbf{W}) \quad (20)$$

where Φ_1 and Φ_2 stands for the regularizers on the latent image $\mathbf{x}$ and the pose weights $\mathbf{W}$, respectively. By substituting the summation of homographies as the proposed hybrid GAC (18), we obtain the GAC-based non-uniform deblurring model. In this paper, we use the l_2 -norm regularizer [10] for $\mathbf{W}$, and the regularizer in [10, 17] for $\mathbf{x}$.

To solve the model in Eq. (20), we adopt an iterative method by updating $\mathbf{W}$ and $\mathbf{x}$ alternatively. Given $\mathbf{x}$, we use the method in [10] to update $\mathbf{W}$, and given $\mathbf{W}$, the accelerated proximal gradient algorithm [1] is adopted to update $\mathbf{x}$.

4 Experimental Results

In this section, we first show the superiority of hybrid GAC method over both slice and fiber-based GAC methods. Then hybrid GAC method is compared with the state-of-the-art geometrical methods, i.e., Whyte et al. [20], Gupta et al. [5], and Hu et al. [10], and the state-of-the-art EFF-based methods, i.e., Hirsch et al. [7] and Xu et al. [21]. All the experiments are conducted on a PC with Intel Core i5-2400 3.10 GHz CPU and 16G RAM memory, and the proposed method is implemented in Matlab.

4.1 Comparison of GAC Variants

On three of the real world camera shaken images shown as Fig. 5, we report the deblurring results of the three GAC methods. Table 1 shows that the three GAC methods have comparable peak memory cost to each other. Since hybrid GAC method usually can decompose the camera motion trajectory into a more compact set of slices and fibers, it is more efficient than slice based GAC and fiber based GAC as shown in Table 1. As shown in Fig. 3, where 3 deblurring examples are presented, hybrid GAC and slice based GAC achieve better deblurring quality than fiber based GAC. As a conclusion, hybrid GAC method outperforms pure fiber based GAC and slice based GAC in items of computational efficiency and memory burden along with satisfactory deblurring quality, and thus is adopted in the following comparison experiments.

4.2 Comparison with Geometrical Methods

We used five real world camera shaken images, shown in the top row of Fig. 4, to compare GAC with three geometrical methods [5, 10, 20]. For the close-ups in Fig. 4, from top to down demonstrate the blurry images, and deblurring results by Whyte et al. [20], Gupta et al. [5], Hu and Yang [10], and the proposed GAC, respectively. GAC can achieve satisfactory deblurring result, and preserve visually plausible edges, e.g., for the image *Six-Books*, the restored text by GAC

Table 1. Running time (s) and Peak memory usage (GB) of the three GAC models.

Image ID	Image size	Time			Memory		
		Fiber	Slice	Hybrid	Fiber	Slice	Hybrid
<i>Books</i>	512×768	1056	837	667	1.74	1.77	1.63
<i>Butcher shop</i>	401×601	567	528	398	0.95	0.96	0.96
<i>Statue</i>	710×523	1014	840	794	1.38	1.50	1.53

Fig. 3. The results of different GAC models on three real camera shaken image. For the close-ups, from top to bottom rows are the input real camera shake blurred images, deblurring results by fiber based GAC model, deblurring results by slice based GAC model, deblurring results by hybrid GAC model.

is more clear than Whyte et al. [20]. Meanwhile, Gupta et al. [5], Hu and Yang [10] adopted the same 3D model to approximate the full 6D camera motion space, and can obtain the similar deblurring results with GAC. In terms of computational efficiency and memory usage, although GAC costs more memory than Whyte et al. [20] shown in Table 2, is at least 100x faster, shown in Table 3. Compared with the accelerated 3D models i.e., Gupta et al. [5], Hu and Yang [10], GAC not only is at least 2.5x faster, but also has much less memory burden, shown as Tables 2 and 3.

Fig. 4. Visual comparison of GAC model with three geometrical methods on five images. For the close-ups, from top to bottom rows are blurry images, deblurring results of Whyte et al. [20], Gupta et al. [5], Hu and Yang [10], and GAC, respectively.

Table 2. Running time (s) of the four geometrical methods.

Image ID	Image size	Whyte et al.	Gupta et al.	Hu &Yang	GAC
<i>Cabin</i>	512×768	98872	7789	2113	550
<i>Car-in-yard</i>	512×768	87527	8829	7495	637
<i>Dim-station</i>	512×768	88767	8512	2512	639
<i>Station</i>	406×679	89673	7982	1402	348
<i>Six-books</i>	512×768	96732	7356	2601	634

Table 3. Peak memory usage (GB) of the four geometrical methods.

Image ID	Image size	Whyte et al.	Gupta et al.	Hu &Yang	GAC
<i>Cabin</i>	512×768	1.00	14.00	10.06	2.06
<i>Car-in-yard</i>	512×768	0.98	14.04	10.20	2.05
<i>Dim-station</i>	512×768	0.91	13.98	10.20	0.98
<i>Station</i>	406×679	0.92	14.01	8.25	1.43
<i>Six-books</i>	512×768	0.91	14.05	11.4	1.96

4.3 Comparison with Non-geometrical Methods Based on EFF

We further compare GAC with three non-geometrical methods based on EFF, i.e., Harmeling et al. [6], Hirsch et al. [7] and Xu et al. [21]. Since source codes of Harmeling et al. [6] and Hirsch et al. [7] are not available, and Xu et al. [21] provides an executable program, we only report the CPU running time com-

Table 4. Running time of the non-geometrical methods and hybrid GAC.

Image ID	Image size	Xu et al. [21]	GAC
<i>Book</i>	512×768	1698	667
<i>Butcher shop</i>	401×601	922	398
<i>Pantheon</i>	273×366	260	87
<i>Boy statue</i>	751×451	1017	767
<i>Coke bottles</i>	680×731	1335	1008
<i>Statue</i>	710×523	1333	794

parison with Xu et al. [21] on several collected camera shaken blurry images shown in Fig. 5. Table 4 shows that GAC is also much faster than Xu et al. Since EFF-based non-geometrical methods actually neglect the global constraints on camera motion trajectory, the deblurring results greatly rely on the rationality of regions division and often introduce artifacts at region boundaries. From Fig. 6(c), one can see that GAC can achieve comparable or better deblurring results than Xu et al. To compare the visual effect with Harmeling et al. [6] and Hirsch et al. [7], we collected deblurring results from their websites. As shown in Fig. 6(a) and Fig. 6(b), GAC can achieve more visually plausible deblurring results.

5 Conclusions

In this paper, we propose a generalized additive convolution (GAC) model for efficient non-uniform deblurring. By decomposing the camera motion trajectory into a compact set of fibers and slices, which can be formulated as convolution operations, GAC directly models the 3D geometrical camera motion, and thanks to the efficient convolution computation via FFT, the computational complexity of GAC is only $O((n_s + n_f)n^2 \log_2 n)$. As to memory usage burden, GAC only requires pre-store several basis filters and mapping matrices. Finally, GAC model is embedded into non-uniform deblurring framework, and can be efficiently solved by an alternating strategy. Since GAC globally models the geometric 3D camera shake model, it can achieve comparable or better deblurring results, compared to both state-of-the-art geometrical and non-geometrical deblurring methods.

Fig. 5. Real camera shake blurred images for comparing with non-geometrical methods.

Fig. 6. Deblurring effect comparison with non-geometrical methods. The top row from left to right are by Harmeling et al. [6], Hirsch et al. [7] and Xu et al. [21], respectively. The bottom row are deblurring results by GAC.

Besides, GAC is much more efficient as well as providing a lower memory cost than most methods.

References

1. Beck, A., Teboulle, M.: Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE TIP* **18**(11), 2419–2434 (2009)
2. Cho, S., Lee, S.: Fast motion deblurring. *ACM Trans. Graph.* **28**(5), 145 (2009)
3. Deng, H., Zuo, W., Zhang, H., Zhang, D.: An additive convolution model for fast restoration of nonuniform blurred images. *Int. J. Comput. Math.* **91**(11), 2446–2466 (2014)
4. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.: Removing camera shake from a single photograph. *ACM Trans. Graph.* **25**(3), 787–794 (2006)
5. Gupta, A., Joshi, N., Lawrence Zitnick, C., Cohen, M., Curless, B.: Single image deblurring using motion density functions. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) *ECCV 2010, Part I. LNCS*, vol. 6311, pp. 171–184. Springer, Heidelberg (2010)
6. Harmeling, S., Hirsch, M., Schölkopf, B.: Space-variant single-image blind deconvolution for removing camera shake. In: *NIPS*, pp. 829–837 (2010)
7. Hirsch, M., Schuler, C.J., Harmeling, S., Schölkopf, B.: Fast removal of non-uniform camera shake. In: *ICCV* (2011)
8. Hirsch, M., Sra, S., Schölkopf, B., Harmeling, S.: Efficient filter flow for space-variant multiframe blind deconvolution. In: *CVPR*, pp. 607–614, June 2010

9. Hu, Z., Xu, L., Yang, M.H.: Joint depth estimation and camera shake removal from single blurry image. In: CVPR (2014)
10. Hu, Z., Yang, M.H.: Fast non-uniform deblurring using constrained camera pose subspace. In: BMVC (2012)
11. Ji, H., Wang, K.: A two-stage approach to blind spatially-varying motion deblurring. In: CVPR (2012)
12. Köhler, R., Hirsch, M., Mohler, B., Schölkopf, B., Harmeling, S.: Recording and playback of camera shake: benchmarking blind deconvolution with a real-world database. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part VII. LNCS, vol. 7578, pp. 27–40. Springer, Heidelberg (2012)
13. Levin, A., Weiss, Y., Durand, F., Freeman, W.: Efficient marginal likelihood optimization in blind deconvolution. In: CVPR (2011)
14. Liu, G., Chang, S., Ma, Y.: Blind image deblurring using spectral properties of convolution operators. IEEE TIP **23**, 5047–5056 (2014)
15. Matungka, R., Zheng, Y.F., Ewing, R.L.: Image registration using adaptive polar transform. IEEE TIP **18**(10), 2340–2354 (2009)
16. Perrone, D., Favaro, P.: Total variation blind deconvolution - the devil is in the details. In: CVPR (2014)
17. Shan, Q., Jia, J., Agarwala, A.: High-quality motion deblurring from a single image. ACM Trans. Graph. **27**(3), 73 (2008)
18. Stančin, S., Tomažič, S.: Angle estimation of simultaneous orthogonal rotations from 3d gyroscope measurements. Sensors **11**(9), 8536–8549 (2011)
19. Tai, Y.W., Tan, P., Brown, M.S.: Richardson-lucy deblurring for scenes under a projective motion path. IEEE TPAMI **33**(8), 1603–1618 (2011)
20. Whyte, O., Sivic, J., Zisserman, A., Ponce, J.: Non-uniform deblurring for shaken images. Int. J. Comp. Vis. **98**(2), 168–186 (2012)
21. Xu, L., Zheng, S., Jia, J.: Unnatural l0 sparse representation for natural image deblurring. In: CVPR (2013)
22. Yue, T., Cho, S., Wang, J., Dai, Q.: Hybrid image deblurring by fusing edge and power spectrum information. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part VII. LNCS, vol. 8695, pp. 79–93. Springer, Heidelberg (2014)
23. Zhang, H., Wipf, D.: Non-uniform camera shake removal using a spatially-adaptive sparse penalty. In: NIPS (2013)