Intra-Frame Deblurring by Leveraging Inter-Frame Camera Motion

Haichao Zhang[1], Jianchao Yang[2]
[1]Duke University, [2]Adobe Research.

Camera motion introduces motion blur, degrading the quality of video. A video deblurring method is proposed based on two observations: (*i*) camera motion *within* capture of each individual frame leads to motion blur; (*ii*) camera motion *between* frames yields inter-frame mis-alignment that can be exploited for blur removal. The proposed method effectively leverages the information distributed across multiple video frames due to camera motion, jointly estimating the motion between consecutive frames and blur within each frame (Figure 1). This joint analysis is crucial for achieving effective restoration by leveraging temporal information.

We argue that camera motion is both an enemy and a friend for restoration, particularly in the context of video. On the one hand, camera motion occurring during the exposure period (within capture of a single frame) introduces blur, which smears detailed structures. On the other hand, due to inter-frame motion, different observations (frames) typically have different blurs, thus containing complementary information. This therefore provides more constraints for the ill-posed inverse problem, and hence is beneficial both in blur-kernel estimation and sharp-video-frame recovery.

The task of video deblurring is challenging mainly due to the *coupled nature of motion and blur*. On the one hand, accurate estimation of the blur using multiple frames requires reliable motion estimation, to compensate for the inter-frame camera motion as well as object motion. On the other hand, reliable motion estimation requires robust blur removal for avoiding matching ambiguities caused by smeared structures due to blur [3]. We aim to tackle the problem of motion and blur estimation *jointly*, for removing camera-motion blur in video, in the presence of non-rigid inter-frame motion.

To exploit this complementary information for improved estimation, we connect two latent sharp frames as $\mathbf{x}_l \approx \mathbf{F}_l\mathbf{x}$, where $\mathbf{F}_l$ is the non-rigid spatial transformation operator induced by the flow field that relates two video frames. While it is possible to establish this connection for each pair of frames, and recover all the latent sharp frames jointly, we concentrate on a specified frame $\mathbf{x}$ without loss of generality. We solve the following problem incorporating the temporally relevant frames

$$\min_{\mathbf{x},\{\mathbf{w}_l,\lambda_l\}\geq 0,\{\mathbf{F}_l\}} \underbrace{\frac{1}{\lambda}\|\mathbf{y}-\sum_j w_j\mathbf{P}_j\mathbf{x}\|_2^2}_{\text{data term}} + \underbrace{\sum_{l\neq 0}\frac{1}{\lambda_l}\|\mathbf{y}_l-\sum_j w_{jl}\mathbf{P}_j\mathbf{F}_l\mathbf{x}\|_2^2}_{\text{temporal term}} + R(\mathbf{x},\{\mathbf{w}_l,\lambda_l,\mathbf{F}_l\}) \quad (1)$$

where $\mathbf{w}_l = [w_{1l}, w_{2l}, \cdots]^T$. Each term in (1) is as follows:

- The first term in (1) is a standard data-fidelity term, based directly on the observation model, and is referred to as *data* term in the sequel.
- The second term is a *motion-aware* regularizer, which exploits the motion-induced complementary information across multiple temporal frames, by relating $\mathbf{x}$ to $\mathbf{x}_l$ via $\mathbf{F}_l$, for further improving the estimation; it is referred to as a *temporal* term.
- The temporal term is helpful only if $\mathbf{y}_l$ contains additional information compared to $\mathbf{y}$ in the data term, meaning $\sum_j w_{jl}\mathbf{P}_j\mathbf{F}_l$ is different from $\mathbf{H} \triangleq \sum_j w_j\mathbf{P}_j$, which is satisfied in general due to the random nature of camera shake.
- The choice of the generic regularization term $R(\cdot)$ is detailed later.

The proposed joint approach (1) has several advantages:

- This method can estimate non-uniform blur, non-rigid motion, and sharp video frames from the blurry video itself, without introducing simplified assumptions (uniform blur, rigid motion etc.) or hardware assistance [1, 4, 6, 7].
- The proposed method can effectively fuse information from multiple frames, and does not rely on the existence of sharp video frames to achieve enhancement. It is applicable to the scenario where all the video frames are blurred, which is a significant difference from most previous methods [2]. However, the existence of sharp frames will help to further improve the restoration quality.
- While using video deblurring as a motivational example, the proposed method is generic and applicable to other scenarios involving multiple *non-uniformly* blurred observations with possibly *non-rigid* inter-frame motions.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

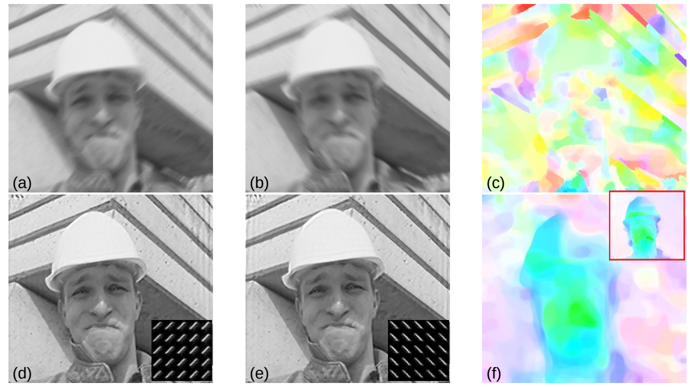

Figure 1: **Illustration of the proposed approach**. (a-b) two blurry video frames (c) optical flow estimated directly on the blurry images (d-f) the output of the proposed method: (d-e) deblurred video frames and blur kernels (f) optical flow (with color encoding [5]) jointly estimated with blur using the proposed method (for comparison, the ground-truth flow estimated on the original sharp images is shown on the top-right of (f)).

Problem (1) is difficult to solve. To reduce the complexity of the problem and increase the robustness of the algorithm, we approach the problem in two phases: in the *blind phase*, we recover the spatially-variant blur kernels and motion fields jointly, using an image penalty that has a strong ability to promote sparsity, for enhancing kernel estimation and reduce the ambiguity in flow estimation. In the second, *non-blind phase*, sharpened video frames are recovered using the estimated blur kernels and motion fields from the first phase, using a robust deblurring method with a natural-image prior. Extensive experiments are carried out on synthetic data as well as real-world blurry videos. Comparisons with several state-of-the-art methods verify the effectiveness of the proposed method.

[1] M. Ben-Ezra and S.K. Nayar. Motion-based Motion Deblurring. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(6):689–698, Jun 2004.

[2] Sunghyun Cho, Jue Wang, and Seungyong Lee. Video deblurring for hand-held cameras using patch-based synthesis. *ACM Trans. Graph.*, 31(4):64, 2012.

[3] A. Rav-Acha and S. Peleg. Two motion blurred images are better than one. *Pattern Recognition Letters*, 26:311–317, 2005.

[4] Filip Šroubek and Peyman Milanfar. Robust multichannel blind deconvolution via fast alternating minimization. *IEEE Trans. on Image Process.*, 21(4):1687–1700, 2012.

[5] Deqing Sun, Stefan Roth, and Michael J. Black. Secrets of optical flow estimation and their principles. In *CVPR*, 2010.

[6] Haichao Zhang and L. Carin. Multi-shot imaging: Joint alignment, deblurring and resolution enhancement. In *CVPR*, 2014.

[7] Haichao Zhang, David P. Wipf, and Yanning Zhang. Multi-image blind deblurring using a coupled adaptive sparse prior. In *CVPR*, 2013.