Lucky DCT Aggregation for Camera Shake Removal

Sanjay Ghosh, Satyajit Naik, and Kunal N. Chaudhury

Department of Electrical Engineering, Indian Institute of Science.

Email: {sghosh, satyajit, kunal} @ee.iisc.ernet.in

Objectives

- To remove the effect of camera shake during a long exposure in hand-held photography.
- A simple and cheap algorithm that can effectively recover the original sharp image from multiple burst images.

Introduction

- Modern cameras come with a *burst* mode for capturing a series of images in quick succession.
- *Multiple-image blind deconvolution* [1, 2]: recovering a sharp image from such a burst.

The mathematical model is that we have blurred versions $y_1, y_2, ..., y_N$ of a sharp image x:

$$y_i = k_i * x + \sigma n_i \qquad (i = 1, \ldots, N),$$

where (k_i) are the blurring kernels, (n_i) are i.i.d. $\mathcal{N}(0,1)$, and σ is the noise level.

- The problem is to recover the unknown image x from the *burst* images $y_1, y_2, ..., y_N$.

Proposed Algorithm

Input: Images $y_1, y_2, \ldots, y_N$ of size $M_1 \times M_2 \times c$.
Parameters: Integers $p \geq 0$ and $N_0 \leq N$.
Output: Output image $\hat{x}$.
Initialize: Null images w, Y, Z of size $M_1 \times M_2 \times c$;

1. **for** $i = 1, 2, \ldots, N$ **do**
   Compute $\mathcal{E}_i$ using (1);
   end
2. Rank images according to decreasing $\mathcal{E}_i$ values
3. Select first N_0 images **for** $i = 1, 2, \ldots, N_0$ **do**
   $\bar{Y}_i = \mathcal{D}(y_i)$; % DCT
   $Y_i = \mathcal{G}(\bar{Y}_i)$; % Smoothing
   $Z = Z \oplus |Y_i|^p$
   end
4. **for** $i = 1, 2, \ldots, N_0$ **do**
   $\bar{w}_i = |Y_i|^p \oslash Z \quad w_i = \mathcal{G}(\bar{w}_i)$; % Smoothing
   $Y = Y \oplus (w_i \otimes Y_i)$ % Aggregation
   $w = w \oplus w_i$
   end
5. $\hat{X} = Y \oslash w$; % Normalization
6. $\hat{x} = \mathcal{D}^{-1}(\hat{X})$. % Inverse DCT

Lucky DCT Aggregation

- The kernels [a] are used to model the camera shake which occurs due to hand tremor.
- The proposed method is built upon – *lucky* imaging using Dirichlet energy [3] and Fourier Burst Accumulation (FBA) [4].
- The images (total N) are sorted according to decreasing Dirichlet energy. Then the top N_0 images with larges energies aggregated.

  The Dirichlet energy for an image y is defined as [3]:

$$\mathcal{E} = \sum_{\ell \in \text{support}(y)} \sum_{\Omega_\ell} \|\nabla y(\ell)\|^2. \qquad (1)$$

  where ∇y is gradient of y and Ω_ℓ is an $n \times n$ window around pixel ℓ.

- Let Y_i denote the Gaussian-smoothed version of the DCT of burst image y_i, that is, $Y_i = \mathcal{G}(\mathcal{D}(y_i))$. For some non-negative integer p, we define the weights:

$$\bar{w}_i(\nu) = \frac{|Y_i(\nu)|^p}{\sum_{j=1}^{N_0} |Y_i(\nu)|^p}, \qquad (2)$$

  where ν is the frequency index for the DCT. The integer p controls the nature of the aggregation.

- The DCT of the aggregated image is:

$$\hat{X}(\nu) = \frac{\sum_{i=1}^{N_0} w_i(\nu) Y_i(\nu)}{\sum_{i=1}^{N_0} w_i(\nu)}. \qquad (3)$$

  The aggregated image is: $\hat{x} = \mathcal{D}^{-1}(\hat{X})$, where $\mathcal{D}^{-1}$ stands for the inverse DCT.

The complete algorithm is summarized at the left. The symbols $\oplus, \otimes$, and $\oslash$ denote pixelwise addition, multiplication, and division, performed on each of the c channels.

[a] The kernels in our experiments were obtained from: http://dev.ipol.im/~mdelbra/fba/

Results

- For simplicity, we have assumed that the images in the burst are perfectly aligned [4].

(a) Input image (512×768). (b) FBA [4], **23.32 / 83.95**, 1.76 s. (c) Proposed, **24.22 / 84.20**, 0.62 s.

Fig. 1. Deblurring results for the blured images (also with additive Gaussian $\sigma = 5$) using FBA [4] and the proposed method. We used $N_0 = 2$ for our method, from the dataset of total $N = 14$ blurred images.

(a) Zoomed of Fig 1 (b). (b) Zoomed of Fig 1 (c).

Fig. 2. Zoomed versions of the boxed portions of the deblurred images in Fig. 1.

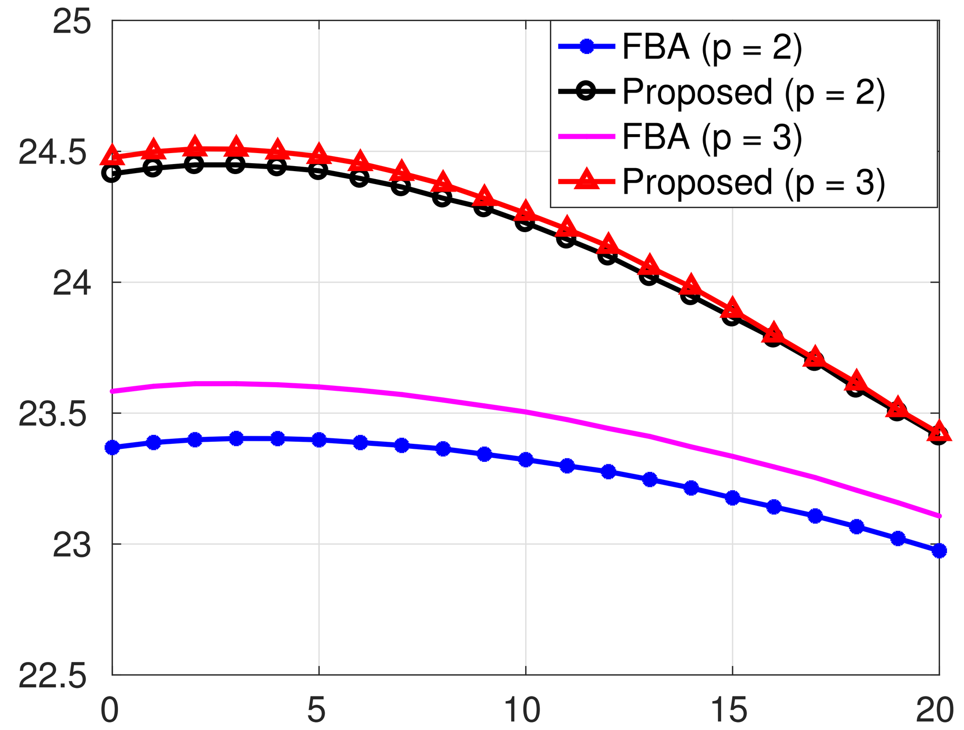

Fig. 3. PSNR *vs* σ, where x is in Fig. 1 (a).

- The proposed method is faster than [4] by a factor of about N_0/N, neglecting the overhead of computing the Dirichlet energies and ranking them.

Conclusions

- Better and faster camera shake correction is achieved through *outlier rejection*.
- Similar idea can be extended to remove camera shakes from videos.

Acknowledgements

- Centenary Conference Travel Fund, Department of Electrical Engineering, IISc.
- SPCOM Student Travel Grant.

References

[1] J.-F. Cai, H. Ji, C. Liu, and Z. Shen, "Blind motion deblurring using multiple images," *Journal of Comput. Phys.*, vol. 228, no. 14, 2009.

[2] F. Sroubek and P. Milanfar, "Robust multichannel blind deconvolution via fast alternating minimization," *IEEE Transactions in Image Processing*, vol. 21, no. 4, Apr. 2012.

[3] G. Haro, A. Buades, and J.-M. Morel, "Photographing paintings by image fusion," *SIAM Journal in Imaging Science*, vol. 5, no. 3, 2012.

[4] M. Delbracio and G. Sapiro, "Removing camera shake via weighted fourier burst accumulation," *IEEE Transactions in Image Processing*, vol. 24, no. 11, Nov. 2015.