Compact Multi-frame Blind Deconvolution

James Nagy
Mathematics and Computer Science
Emory University
Atlanta, GA, USA

Michael Hart, University of Arizona
Douglas Hope, HartSCI Company
Stuart Jefferies, University of Hawaii and Georgia State University

Motivation

Big Data Imaging Problems

- Multi-Frame Blind Deconvolution (MFBD) where multi = very large
- MFBD combined with 3D (shape) or 4D (shape and color) reconstructions.

Requirements

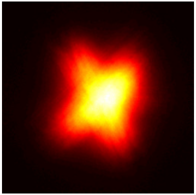
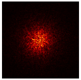
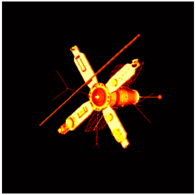
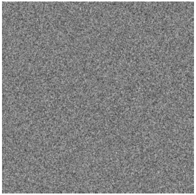
- Powerful computers
- More efficient algorithms
    - More efficient use of current algorithms
    - Smarter ways to process massive data sets
- Collaborative/synergistic teams

    e.g., physics, math, computer science, engineering

Outline

Convolution

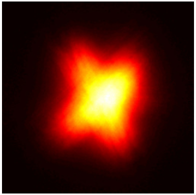
Consider the convolution image formation model:

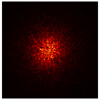
 =  *  + 

Deconvolution

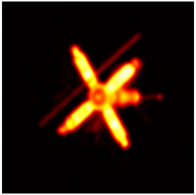
Deconvolution: Given

- Blurred image, and
- Point spread function (convolution kernel).





$\Rightarrow$ compute $\Rightarrow$

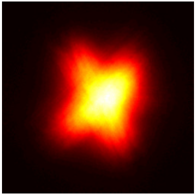


- Compute estimate of true image.

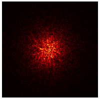
Blind Deconvolution

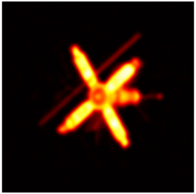
Blind Deconvolution: Given

- Blurred image.



$\Rightarrow$ compute $\Rightarrow$



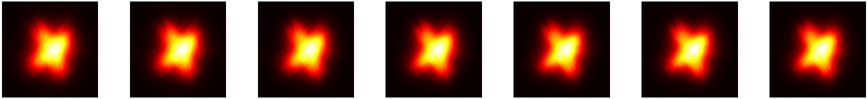


- Compute estimate of true image, and
- Compute estimate of PSF.

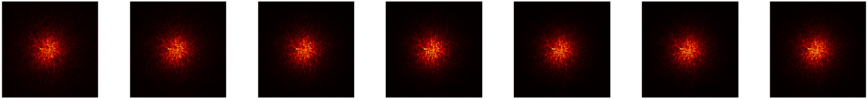
Multi-Frame Blind Deconvolution

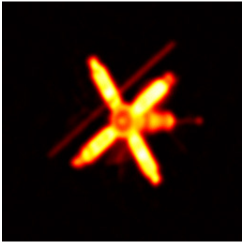
Multi-Frame Blind Deconvolution (MFBD):

- Given multiple frames of blurred images:



- Reconstruct PSFs and object:



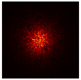
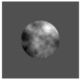


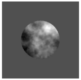
Single Frame Blind Deconvolution (SFBD) Model

Parameterize point spread function

- Using convolution model: $\mathbf{b} = \mathbf{psf}(\mathbf{y}) * \mathbf{x} + \boldsymbol{\eta}$
- Or, using matrix notation: $\mathbf{b} = \mathbf{A}(\mathbf{y})\mathbf{x} + \boldsymbol{\eta}$

Example parameterizations:

- PSF pixels **psf** ()
- Wavefront phase: **psf** () $= |\mathcal{F}^{-1}(Pe^{i\mathbf{y}})|^2$
- Wavefront phase with Zernikes:

psf () $= |\mathcal{F}^{-1}(Pe^{i(y_1\mathbf{z}_1 + \cdots y_m\mathbf{z}_m})|^2$

General Mathematical Model

General mathematical model for image formation:

$$\mathbf{b} = \mathbf{A}(\mathbf{y})\,\mathbf{x} + \boldsymbol{\eta}$$

where

- $\mathbf{b}$ = vector representing observed image
- $\mathbf{x}$ = vector representing true image
- $\mathbf{A}(\mathbf{y})$ = matrix defining blurring operation

  For example,

  - Convolution with imposed boundary conditions
  - Spatially variant blurs
- $\mathbf{y}$ = vector of parameters defining blurring operation

Goal: Given $\mathbf{b}$, jointly compute approximations of $\mathbf{y}$ and $\mathbf{x}$.

Multi-Frame Blind Deconvolution (MFBD)

The MFBD problem is:

$$\begin{aligned} \mathbf{b}_1 &= \mathbf{A}(\mathbf{y}_1)\mathbf{x} + \boldsymbol{\eta}_1 \\ \mathbf{b}_2 &= \mathbf{A}(\mathbf{y}_2)\mathbf{x} + \boldsymbol{\eta}_2 \\ &\vdots \\ \mathbf{b}_m &= \mathbf{A}(\mathbf{y}_m)\mathbf{x} + \boldsymbol{\eta}_m \end{aligned}$$

To solve, could consider least squares best fit objective:

$$\left\| \begin{bmatrix} \mathbf{b}_1 - \mathbf{A}(\mathbf{y}_1)\mathbf{x} \\ \vdots \\ \mathbf{b}_m - \mathbf{A}(\mathbf{y}_m)\mathbf{x} \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_m \end{bmatrix} - \begin{bmatrix} \mathbf{A}(\mathbf{y}_1) \\ \vdots \\ \mathbf{A}(\mathbf{y}_m) \end{bmatrix} \mathbf{x} \right\|_2^2 = \|\mathbf{b} - \mathbf{A}(\mathbf{y})\mathbf{x}\|_2^2$$

Also need regularization, but we omit that complication for now.

Compact Multi-Frame Blind Deconvolution

Processing a large number of frames is computationally intensive.

Compact MFBD (CMFBD)

D. Hope, S. Jefferies, Optics Letters, 36 (2011), pp. 867–869.

- Identify a small set of *control frames* that contain most independent information.
- Reduce full set of data to small set of control frames, without losing any important information.

CMFBD: Identifying Control Frames

- Suppose $\mathbf{A}_j \equiv \mathbf{A}(\mathbf{y}_j)$ are simultaneously diagonalizable (e.g. Fourier transforms for circulant matrices)

$$\mathbf{A}_j = \mathbf{F}^* \mathbf{\Lambda}_j \mathbf{F}$$

- Consider noise free data, and the j-th frame:

$$\begin{aligned} \mathbf{A}_j \mathbf{x} = \mathbf{b}_j \quad &\Rightarrow \quad \mathbf{\Lambda}_j \hat{\mathbf{x}} = \hat{\mathbf{b}}_j \\ &\Rightarrow \quad \mathbf{\Lambda}_j \operatorname{diag}(\hat{\mathbf{x}}) = \operatorname{diag}(\hat{\mathbf{b}}_j) \\ &\Rightarrow \quad \operatorname{diag}(\hat{\mathbf{b}}_j)^\dagger = \operatorname{diag}(\hat{\mathbf{x}})^\dagger \mathbf{\Lambda}_j^\dagger \end{aligned}$$

where $\hat{\mathbf{x}} = \mathbf{F}\mathbf{x}$ and $\hat{\mathbf{b}}_j = \mathbf{F}\mathbf{b}_j$ are unitary Fourier transforms.

CMFBD: Identifying Control Frames

- Assume there is a uniformly "best" conditioned matrix $\mathbf{A}_k$. That is, there is a $\boldsymbol{\Lambda}_k$ such that

$$[\,|\boldsymbol{\Lambda}_k|\,]_{ii} \geq \tau \quad \text{if there exists } j \text{ with} \quad [\,|\boldsymbol{\Lambda}_j\,|]_{ii} \geq \tau$$

  where $\tau > 0$ is a tolerance.
- In this case, where there is a single control frame, observe:

$$\text{diag}(\hat{\mathbf{b}}_j) = \boldsymbol{\Lambda}_j \text{diag}(\hat{\mathbf{x}}) \quad \text{and} \quad \text{diag}(\hat{\mathbf{b}}_k)^\dagger = \text{diag}(\hat{\mathbf{x}})^\dagger \boldsymbol{\Lambda}_k^\dagger$$

- This allows to compute *spectral ratios*

$$\underbrace{\text{diag}(\hat{\mathbf{b}}_j)\,\text{diag}(\hat{\mathbf{b}}_k)^\dagger}_{\text{known}} = \underbrace{\boldsymbol{\Lambda}_j\,\boldsymbol{\Lambda}_k^\dagger}_{\text{unknown}}$$

CMFBD: Exploiting Control Frames

- WLOG, assume the control frame is $k = 1$, and observe:

$$\left\| \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_m \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \mathbf{\Lambda}_1 \\ \mathbf{\Lambda}_2 \\ \vdots \\ \mathbf{\Lambda}_m \end{bmatrix} \hat{\mathbf{x}} - \begin{bmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \\ \vdots \\ \hat{\mathbf{b}}_m \end{bmatrix} \right\|_2^2$$

$$= \left\| \begin{bmatrix} \mathbf{\Lambda}_1 \mathbf{\Lambda}_1^\dagger \\ \mathbf{\Lambda}_2 \mathbf{\Lambda}_1^\dagger \\ \vdots \\ \mathbf{\Lambda}_m \mathbf{\Lambda}_1^\dagger \end{bmatrix} \mathbf{\Lambda}_1 \hat{\mathbf{x}} - \begin{bmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \\ \vdots \\ \hat{\mathbf{b}}_m \end{bmatrix} \right\|_2^2$$

CMFBD: Exploiting Control Frames

- WLOG, assume the control frame is $k = 1$, and observe:

$$\left\| \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_m \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \boldsymbol{\Lambda}_1 \\ \boldsymbol{\Lambda}_2 \\ \vdots \\ \boldsymbol{\Lambda}_m \end{bmatrix} \hat{\mathbf{x}} - \begin{bmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \\ \vdots \\ \hat{\mathbf{b}}_m \end{bmatrix} \right\|_2^2$$

$$= \left\| \begin{bmatrix} \mathbf{I} \\ \mathbf{D}_2 \\ \vdots \\ \mathbf{D}_m \end{bmatrix} \boldsymbol{\Lambda}_1 \hat{\mathbf{x}} - \begin{bmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \\ \vdots \\ \hat{\mathbf{b}}_m \end{bmatrix} \right\|_2^2$$

where $\mathbf{D}_j = \mathrm{diag}(\hat{\mathbf{b}}_j)\,\mathrm{diag}(\hat{\mathbf{b}}_1)^\dagger$

CMFBD Observations

- The initial MFBD problem has unknowns:

$$\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_m, \mathbf{x} \quad \text{or, equivalently} \quad \mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \ldots, \mathbf{\Lambda}_m, \hat{\mathbf{x}}$$

- After identifying a control frame, significantly fewer unknowns:

$$\mathbf{A}_1, \mathbf{x} \quad \text{or, equivalently} \quad \mathbf{\Lambda}_1, \hat{\mathbf{x}}$$

- More control frames may be needed to capture all $[\,|\mathbf{\Lambda}_j\,|]_{ii} \geq \tau$.
- For noisy data, algebra relating known and unknown information holds only approximately.
- Frame Selection: Based on heuristics
  - "Best" conditioned $\mathbf{A}_k \Leftrightarrow$ least blurred image
  - Many techniques can be used – we use a Fourier based power spectrum approach.

CMFBD Practical Details

- **Reduction of Single Frame Problem:** Use Givens rotations

$$\left\| Q^* \left(\begin{bmatrix} \mathbf{I} \\ \mathbf{D}_2 \\ \vdots \\ \mathbf{D}_m \end{bmatrix} \boldsymbol{\Lambda}_1 \hat{\mathbf{x}} - \begin{bmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \\ \vdots \\ \hat{\mathbf{b}}_m \end{bmatrix} \right) \right\|_2^2 = \left\| \begin{bmatrix} \mathbf{D} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \boldsymbol{\Lambda}_1 \hat{\mathbf{x}} - \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_m \end{bmatrix} \right\|_2^2$$

Therefore, we need only consider

$$\| \mathbf{D} \boldsymbol{\Lambda}_1 \hat{\mathbf{x}} - \mathbf{d}_1 \|_2^2 = \| \mathbf{D} \mathbf{F} \mathbf{A}_1 \mathbf{x} - \mathbf{d}_1 \|_2^2$$

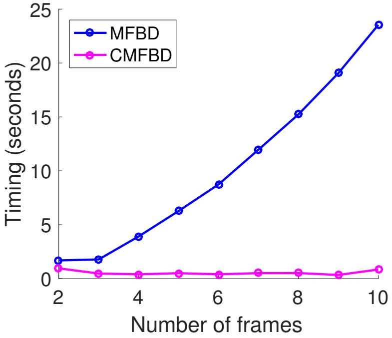
Thus, the MFBD problem

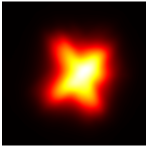
$$\min_{\mathbf{y}_j,\, \mathbf{x}} \sum_{j=1}^{m} \| \mathbf{A}(\mathbf{y}_j)\, \mathbf{x} - \mathbf{b}_j \|_2^2$$

reduces to the CSFBD problem

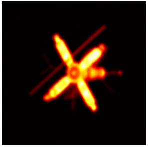
$$\min_{\mathbf{y}_1,\, \mathbf{x}} \| \mathbf{W} \mathbf{A}(\mathbf{y}_1)\, \mathbf{x} - \mathbf{d}_1 \|_2^2 , \quad \mathbf{W} = \mathbf{D}\mathbf{F}$$

Numerical Illustration of Time Savings

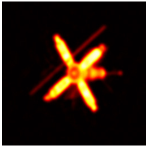




Mean data frame



MFBD



CMFBD

Global Variable Consensus

The MFBD problem can be written as:

$$\min_{\mathbf{y}_i, \mathbf{x}} \sum_{i=1}^{m} \|\mathbf{b}_i - \mathbf{A}(\mathbf{y}_i)\mathbf{x}\|_2^2 + g(\mathbf{x})$$

where here we include an object regularization term, $g(\mathbf{x})$.

Remarks:

- Regularization $g(\mathbf{x})$ can be used to enforce nonnegativity, sparsity, etc.
- The unknown $\mathbf{x}$ couples the objective terms $i = 1, \ldots, m$
- We can get a partial decoupling by reformulating as:

$$\min_{\mathbf{y}_i, \mathbf{x}_i} \sum_{i=1}^{m} \|\mathbf{b}_i - \mathbf{A}(\mathbf{y}_i)\mathbf{x}_i\|_2^2 + g(\mathbf{z}) \quad \text{subject to } \mathbf{x}_i = \mathbf{z}, \; i = 1, \ldots, m$$

Global Variable Consensus

Using an augmented Lagrangian approach, and the Alternating Direction Method of Multipliers (ADMM)[1], the optimization decouples:

for k = 1, 2, ...

$$\left[\mathbf{y}_i^{(k+1)}, \mathbf{x}_i^{(k+1)}\right] = \operatorname*{argmin}_{\mathbf{y}_i, \mathbf{x}_i} \|\mathbf{b}_i - \mathbf{A}(\mathbf{y}_i)\mathbf{x}_i\|_2^2 + \frac{\beta}{2}\|\mathbf{x}_i - \mathbf{z}^{(k)} + \mathbf{u}_i^{(k)}\|_2^2$$

$$\bar{\mathbf{x}}^{(k+1)} = \frac{1}{m}\sum_{i=1}^{m} \mathbf{x}_i^{(k+1)}$$

$$\bar{\mathbf{u}}^{(k)} = \frac{1}{m}\sum_{i=1}^{m} \mathbf{u}_i^{(k)}$$

$$\mathbf{z}^{(k+1)} = \operatorname*{argmin}_{\mathbf{z}} \left\{ g(\mathbf{z}) + \frac{m\beta}{2}\|\mathbf{z} - \bar{\mathbf{x}}^{(k+1)} - \bar{\mathbf{u}}^{(k)}\|_2^2 \right\}$$

$$\mathbf{u}_i^{(k+1)} = \mathbf{u}_i^{(k)} + \mathbf{x}_i^{(k+1)} - \mathbf{z}^{(k+1)}$$

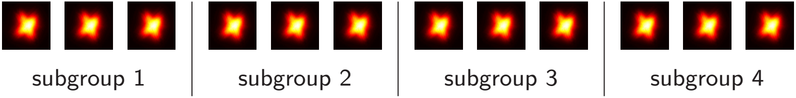
end

[1] Good ADMM references: Wahlberg, Boyd, Annergren, Wang, *Proc. 16th IFAC Symposium on System Identification*, 2012, and Boyd, et. al., *Foundations and Trends in Machine Learning*, 2010.

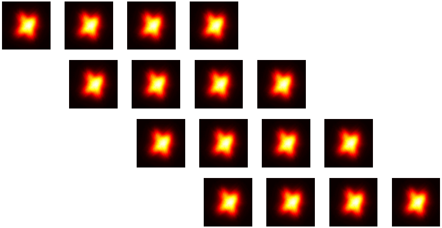
Global Variable Consensus

Advantages:

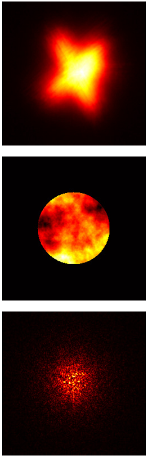
- Decoupling allows for easy parallel processing of groups of frames



  - Can either use standard MFBD on subgroups of frames, or
  - Use CMFBD on subgroups of frames
- Regularization term is also decoupled, allowing users to plug in many options, and it simplifies the computation.
- Sliding window approach might be possible:

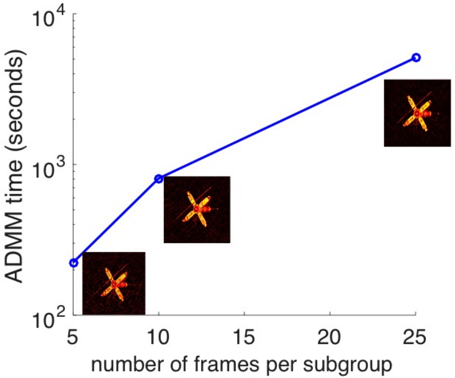


Global Variable Consensus: Numerical Illustration[2]



50 total frames, split in three different ways

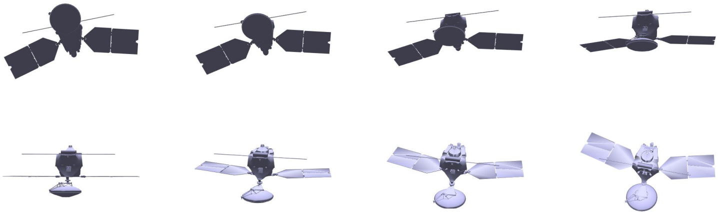
10 subgroups (5 frames each)	5 subgroups (10 frames each)	2 subgroups (25 frames each)

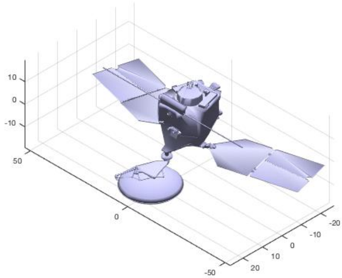


[2] J. D. Schmidt, Numerical Simulation of Optical Wave Propagation, SPIE Press Monograph Vol. PM199, 2010

Higher Dimensional Image Reconstruction

Three dimensional reconstruction from two dimensional measurements:





Higher Dimensional Image Reconstruction

Some computational challenges:

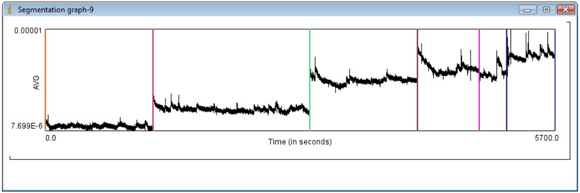
- Requires processing many, many frames of data.
- Mathematical model is similar to MFBD, but
  - Number of unknowns for object significantly increases.
  - Additional unknowns associated with parameters defining object orientation.
- Some related work has been done for molecular structure determination, e.g. in Cryo-EM and x-ray crystallography[3]

[3] J. Chung, P. Sternberg and C. Yang, *High Performance 3-D Image Reconstruction for Molecular Structure Determination*, International Journal of High Performance Computing Applications, 24 (2010), pp. 117–135.

Higher Dimensional Image Reconstruction

What further information can be used?

- Possibly assume blocks of data have constant orientation parameters
  Idea like this was used in PET brain image reconstruction[4]



- Can use consensus ADMM type approach on blocks of data.
- Use other information (e.g., a frozen flow assumption), or technologies (e.g., laser guide stars).

[4]P. Wendykier, J. Nagy, *Parallel Colt: High Preformance Java Library for Scientific Computing and Image Processing*, ACM Transactions on Mathematical Software, 37 (2010), pp. 31:1–31:22

Summary

- Big data, multi-frame image processing requires not only powerful computers, but also new approaches to process massive data sets.

  This is especially true for 3D/4D image reconstructions.

- Goal should be to extract as much information as possible from collected data, but to also do it quickly.

- Important to have synergistic collaborations with various expertise.