Data Structures in Imaging

George Papanicolaou
Stanford University

ICIAM 2019, Valencia

July 18, 2019

With M. Leibovich and C. Tsogka
Thank you ICIAM for the Lagrange Prize

It has been an honor and a pleasure to work on imaging for nearly twenty years with several colleagues, young and older, junior and senior:


In different ways they have contributed enormously to imaging, and continue to contribute to its mathematical foundations, algorithm development, numerical simulations, as well as to the analysis of specific imaging systems.
Outline: Role of data structure in a particular imaging framework

Synthetic Aperture Radar (SAR) with motion estimation: The objective is to image both the stationary background reflectivity and the moving object(s).

1. The synthetic aperture needs to be divided into sub-apertures that are small enough to capture the direction of the moving target, as well as for computational, FFT, issues, but also large enough to provide a background image. Should the sub-apertures overlap and if so by how much? This is a Data Structure issue.

2. Formulate first the SAR problem. What happens if we image as if there is no motion?

3. What are some ways in which motion can be estimated?

4. The use of low-rank-plus-sparse decomposition of the data: Robust Principal Components Analysis (RPCA) and its shortcomings.

5. Tensorization of the SAR data and Tensor RPCA (TRPCA) improves very substantially the sensitivity of motion estimation.
Synthetic aperture imaging configuration: At every time $s$ the airborne platform emits a pulse $f(t)$ and records the reflections from the imaging region $D(t, s)$. There are 10 stationary point scatterers, and a single moving target. The angle $\alpha$ in which the target is moving is relative to the vertical platform-target plane at $s = 0$, $\vec{r}(0)$. 
Conventional SAR imaging with no motion

To maximize the power the probing pulses are long, of support $t_c \gg 1/B$ where $B$ denotes the bandwidth. They are linear frequency modulated chirps. To re-concentrate the energy of the reflected echoes to an interval of size $1/B$ they are convolved with the complex conjugate of the time-reversed emitted pulse. This is the pulse compression. Since the reflections relevant for imaging cover a limited area of support $\ll \Delta s$, a range compression is done, that is, we remove from the data the large phase $\omega \tau(s, \vec{p}_o)$ where $\vec{p}_o$ is a reference point.

Pulse and range compression together gives the down-ramped data,

$$D_r(s, t) = \int dt' D(s, t - t' + \tau(s, \vec{p}_o)) \hat{f}(-t') = \int \frac{d\omega}{2\pi} \hat{f}(\omega) \hat{D}(s, \omega) e^{-i\omega[t + \tau(s, \vec{p}_o)]}. \quad (1)$$

Here, $\tau(s, \vec{p}_o)$ is the round trip travel time between the platform location at slow time $s$, $\vec{r}(s)$ and the reference point location $\vec{p}_o$,

$$\tau(s, \vec{p}_o) = 2 \frac{\|\vec{r}(s) - \vec{p}_o\|}{c}, \quad (2)$$

with $c$ the speed of light.
Conventional SAR imaging continued

The SAR data matrix $\mathcal{D} \in \mathbb{R}^{(n+1) \times (m+1)}$ is defined by discrete samples of $D_r(s, t)$.

$$D_{il} = D_r \left( s_{i- \frac{n}{2} - 1}, t_{l-1} \right), \quad i = 1, \ldots, n + 1, \quad l = 1, \ldots, m + 1, \quad (3)$$

with slow times $s_j$ defined by

$$s_j = j\Delta s, \quad j = -n/2, \ldots, n/2 \quad (4)$$

and fast times $t_l$ defined as

$$t_l = l\Delta t, \quad l = 1, \ldots, m. \quad (5)$$

Here we assumed that the pulse repetition rate $\Delta s$ is an integer multiple of $\Delta t$ and set $m = \Delta s/\Delta t$. The SAR image is formed by summing coherently the down-ramped data $D_r(s_j, t)$ back-propagated to the imaging point $\vec{\rho}$ using the travel times differences $\tau(s_j, \vec{\rho}) - \tau(s_j, \vec{\rho}_o)$,

$$I_{\text{SAR}}(\vec{\rho}) = \sum_{j=-n/2}^{n/2} D_r(s_j, \tau(s_j, \vec{\rho}) - \tau(s_j, \vec{\rho}_o)). \quad (6)$$

$I_{\text{SAR}}(\vec{\rho})$ is most often computed by DFT over sub-apertures.
Conventional motion estimation

The simplest and most widely used approach is the Displaced Phase Center Antenna (DPCA) method (Review in: Muehe-Labitt 2000) in which two synchronized antennas are used. The antennas follow the same trajectory with a small time delay. By subtracting the data traces collected at the two antennas the echoes due to the stationary background are eliminated. This solution does not require any image formation as part of the motion detection stage, but requires the necessary hardware to be in place to record the extra data.

Other methods, such as using autofocus (Fineup 2001) or sparsity driven techniques (Willsky et al 2014) for moving target detection rely on forming a preliminary image first, and then detecting motion through features in the image. We can use the spectral properties of the raw-data matrix to detect motion, without imaging first, or recording more data. Specifically, we explore the use of robust principal component analysis (RPCA) for data separation in SAR. The idea was first proposed in (Borcea, Callaghan and P. 2013) and was further developed and analyzed in (Leibovich, Tsogka and P. 2018), where optimal parameters were derived for achieving robust separation in SAR.
RPCA for separating data from stationary background and moving target

RPCA, or low rank plus sparse decomposition, was originally applied in video processing (Candes et al 2009). RPCA uses the fact that the moving targets and stationary background would generate data structures with different properties. Indeed, the background data form a low rank data matrix, while the moving object’s echoes correspond to a sparse matrix. These data structures can be obtained by solving a convex optimization problem of the form

\[
\begin{align*}
\min_{L,S \in \mathbb{C}^{n_1 \times n_2}} & \quad \|L\|_* + \eta \|S\|_1 \\
\text{subject to} & \quad L + S = D.
\end{align*}
\]

(7)

Here \(\|L\|_*\) denotes the nuclear norm, that is, the sum of the singular values of \(L\), and \(\|S\|_1\) is the matrix \(\ell_1\)-norm of \(S\) (element-wise). Assuming the matrix \(D \in \mathbb{R}^{n_1 \times n_2}\) is the sum of a low rank matrix, \(L_o\), and a sparse matrix, \(S_o\), then, under some (sufficient) conditions, the solution of the optimization (7) recovers \(L_o\) and \(S_o\) exactly.
Low rank plus sparse decomposition

RPCA applied to the data matrix $D_r(s, t)$. The moving target has velocity $v_t = 15\text{m/s}$, $\alpha = 0$ and there are 10 stationary targets. The moving target’s reflectivity is 10% of the reflectivity of the stationary scatterers. Left: Original data matrix. The echoes corresponding to the moving target are weak but still visible in the data; Middle: RPCA, the low rank part for the stationary background; Right: RPCA, the sparse part for the echoes from the moving target. Good separation is achieved with RPCA when using the "optimal" value for $\eta$ in (7). This is the case for fast moving targets as in this example.
Moving at $\mathbf{v}_t = 15 \text{m/s}$ and the effect of the direction of motion. Targets with $\alpha = 0$, top, have large column support, and the phase ($\Delta \tau(s)$) is nearly linear ($|D_r(s,t)|$ left, real part right). With $\alpha = \pi/2$, bottom, the column support is much smaller, but the phase is non linear.
Slow speed, angle dependence, RPCA fails: Tensorization

Schematic of SAR data tensor representation. For motion detection purposes, the large aperture data matrix $D_r(s, t)$ is converted to a 3rd order tensor $A(s, t, \ell) = A^{(\ell)}(s, t)$. The tensor is composed of partially overlapping sub-aperture data from the full data matrix $D_r(s, t)$. This new data structure is needed because the information in the data matrix is too heterogeneous.
We define the TRPCA algorithm, using a specific extension of the nuclear norm for third-order tensors (Lu et al 2016), and solve a tensor based RPCA optimization problem for complex valued third order tensors:

\[
\min_{\mathcal{L}, \mathcal{S} \in \mathbb{C}^{n_1 \times n_2 \times n_3}} \|\mathcal{L}\|_{*,\mathcal{F}} + \eta \|\mathcal{S}\|_1,
\]

subject to \( \mathcal{L} + \mathcal{S} = A \)

\( \| \cdot \|_{*,\mathcal{F}} \), to be defined, is a specific extension of the tensor nuclear norm which involves performing a Fourier Transform with respect to the sub-aperture index \( \ell \).
How does TRPCA perform?

TRPCA applied to SAR data from a slow moving target. The moving target velocity is $v_t = 1\text{m/s}$ and $\alpha = \pi/4 - \pi/2$ has a gap. The target’s reflectivity is 10% of the reflectivity of the 10 stationary scatterers. Left: Original data matrix. The echoes of the weakly reflecting and slow moving target are barely discernible; Middle: TRPCA, the low rank part from the stationary background; Right: TRPCA, the sparse part from the moving target; We see that TRPCA achieves good data separation in a very challenging setting.

While the performance of TRPCA is not universally better than that of regular matrix RPCA, TRPCA performs significantly better in the cases where motion is hardest to detect.
The TRPCA method: The nuclear norm

A natural definition of the nuclear norm that extends to higher dimensions is

$$\|A\|_{*,\mathcal{F}} = \inf \left\{ \sum_{i=1}^{r} |\lambda_i| \mid A = \sum_{i=1}^{r} \lambda_i u_i^1 \otimes u_i^2 \otimes \cdots \otimes u_i^d, \|u_i^j\| = 1, r \in \mathbb{N} \right\}.$$  (9)

Note that the $u_i^j$ do not need to be orthogonal. As outlined in (Friedland et al 2018), this is in general an NP hard problem to compute. Hence, we look for alternative, more tractable definitions of a nuclear norm.
The simplest alternative nuclear norm

The simplest alternative would be to take the matrix panels with respect to a specific dimension and compute the matrix SVD on every panel separately, such that, for example,

\[ \mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}, \quad \mathcal{A}^{(\ell)} = \mathcal{A}(\cdot, \cdot, \ell) = U^{(\ell)} \Sigma^{(\ell)} V^{(\ell) \text{H}}, \quad \ell = 1, \ldots, n_3 \]  

\[ \|\mathcal{A}\|_{*, \mathcal{D}} = \sum_{\ell=1}^{n_3} \|\mathcal{A}^{(\ell)}\|_* = \sum_{\ell=1}^{n_3} \sum_{j=1}^{r_\ell} \sigma_j^{(\ell)}. \]  

In SAR this means that we do RPCA for each sub-aperture without allowing for any interaction between them, regardless of the overlap. This is clearly not what we want!
The Fourier nuclear norm

We can represent a third order tensor as a block-circulant matrix, with the tensor multiplication homeomorphic to the regular matrix multiplication

\[
\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3} \rightarrow \frac{1}{\sqrt{n_3}} \begin{pmatrix}
A^{(0)} & A^{(1)} & \ldots & A^{(n_3-1)} \\
A^{(n_3-1)} & A^{(0)} & \ldots & A^{(n_3-2)} \\
& \ddots & \ddots & \ddots \\
A^{(1)} & A^{(2)} & \ldots & A^{(0)}
\end{pmatrix} \in \mathbb{C}^{n_1 n_3 \times n_2 n_3}.
\] 

(12)
Block-circulant matrices can be block-diagonalized by a Discrete Fourier Transform (DFT) matrix,

\[
\frac{1}{\sqrt{n_3}} \begin{pmatrix}
A^{(0)} & A^{(1)} & \ldots & A^{(n_3-1)} \\
A^{(n_3-1)} & A^{(0)} & \ldots & A^{(n_3-2)} \\
& \ddots & \ddots & \ddots \\
A^{(1)} & A^{(2)} & \ldots & A^{(0)}
\end{pmatrix} \xrightarrow{\mathcal{F}} \begin{pmatrix}
\bar{A}^{(0)} \\
\bar{A}^{(1)} \\
\vdots \\
\bar{A}^{(n_3-1)}
\end{pmatrix},
\]

which is equivalent to

\[
\bar{A} = \mathcal{F}_3 A, \quad \bar{A}^{(k)} = \frac{1}{\sqrt{n_3}} \sum_{\ell=0}^{n_3-1} \omega_{n_3}^{\ell k} A^{(\ell)},
\]

where

\[
\omega_{n_3} = e^{\frac{i2\pi}{n_3}}, \quad \bar{A}^{(k)} \in \mathbb{C}^{n_1 \times n_2}, k = 0, \ldots, n_3 - 1.
\]
The Fourier tensor nuclear norm

Thus, another form of a nuclear norm is given by

\[ \| \mathcal{A} \|_{*,\mathcal{F}} = \sum_{k=0}^{n_3-1} \| \bar{A}^{(k)} \|_* . \] (16)

This last definition of the tensor nuclear norm proves to be well suited for the SAR motion detection problem.

Why is this so? Because the DFT captures in a particularly appropriate way motion in the sub-apertures, regardless of the overlap, as it is recorded in the data.
Theoretical study of sensitivity to $\eta$ in TRPCA

We can now consider TRPCA as

$$
\min_{\mathcal{L} + \mathcal{S} = \mathcal{A}} \| \mathcal{L} \|_{*,\mathcal{F}} + \eta \| \mathcal{S} \|_1 = \min_{\mathcal{L} + \mathcal{S} = \mathcal{A}} \sum_{\ell=1}^{n_3} \| \bar{\mathcal{L}}^{(\ell)} \|_{*,\mathcal{F}} + \eta \| \bar{\mathcal{S}}^{(\ell)} \|_1,
$$

(17)

where $\mathcal{L}^{(\ell)}$ and $\mathcal{S}^{(\ell)}$ are the sub-aperture parts of $\mathcal{L}$ and $\mathcal{S}$, respectively. We are back to matrices!

Main result of the analysis that is supported very well by simulations: The range of the parameter $\eta$ for which we have good, robust results depends on the number of sub-apertures and the overlap (other things being the same). This suggests how to choose it so as to get information about motion that is difficult to get by conventional methods, in a systematic and stable way.
**Main tool in the analysis**

For matrices $A_1, A_2, \ldots, A_k$, $A_i \in \mathbb{C}^{m \times n_i}$, the following holds for the matrix

$$A = [A_1, A_2, \ldots, A_k] \in \mathbb{C}^{m \times N}, \quad N = \sum_{i=1}^{k} n_i$$

(18)

$$\left( \sum_{i=1}^{k} \|A_i\|_2^2 \right)^{1/2} \leq \|A\|_* \leq \sum_{i=1}^{k} \|A_i\|_*.$$  \hspace{1cm} (19)

SAR interpretation: When the information in the sub-aperture data is the same (lots of overlap) the left inequality tends to an equality. When the information in the sub-aperture data is "distinct" then the right inequality tends to an equality.
Imaging: Left TRPCA, Middle decoupled, Right full

Top \( \alpha = \pi/4 \), Middle \( \alpha = 3\pi/8 \), Bottom \( \alpha = \pi/2 \).
Concluding remarks

- Tensorizing SAR data is a new and very effective way to estimate motion, and image the stationary background without loss of resolution. Finding a "good" data structure matters. The tensor structure makes a big difference.
- TRPCA systematically provides data separation with the "right" amount of sub-aperture overlap.
- Many other instances where the data structure matters in imaging.
- There are a lot of theoretical and applied (implementation) issues in this context that need to be considered in more depth and explored further.