Focussing dynamic single-channel synthetic aperture radar video with optical flow-based compressive sensing

F. M. Watson[1]
[1]*The University of Manchester, UK*
francis.watson@manchester.ac.uk

The motion of targets is well known to result in their defocussing and displacement in SAR imagery, but detection of motion and re-focussing of targets under generic conditions remains of ongoing interest. One class of methods involves forming images of sub-apertures in which motion defocussing will be reduced. In this paper, we use dynamic tomographic image formation methods utilising an optical flow constraint to form a video of SAR sub-aperture images. These retain fine resolution of the full aperture, focussing along-track motion.

Introduction: The effects of moving targets in synthetic aperture radar (SAR) imagery are well studied, resulting in both displacements and defocusing. Detection and focussing of moving targets is important both in a defence context, as well as providing potentially valuable insights of activity for commercial applications of remote sensing, and a range of techniques for focussing different types of motion exist as discussed for example by Axelsson[1]. A common class of methods involves dividing the data into sequential sub-apertures in which targets displacements are reduced, and therefore so is defocussing. Analysis of target motion between frames may be used to estimate velocity, for example forming a single re-focussed image[2], superimposing motion indication[3], or individual frames might be retained as a "SAR video". In this case, one must balance resolution of the frames and the amount of motion defocussing in determining the lengths of (possibly overlapping) sub-aperture: resolution will be coarser than that of the full aperture.

We consider the use of dynamic tomographic (4D CT) techniques to focus motion in video-SAR. For a generic 4D CT problem, limited data available for each frame results in a severely ill-conditioned image reconstruction problem, leading to loss of resolution and the potential to introduce artefacts. Such methods make use of regularisation to apply prior knowledge and stabilise the reconstructions scheme. This has often been in the form of spatio-temporal Total Variation (TV), penalising changes between frames[4]. Recent progress has instead made use of an optical flow regularisation term to jointly reconstruct both the dynamic image and velocity field[5, 6]. Optical flow describes the apparent motion between image frames based on brightness consistency between image frames[7]. Its inclusion in dynamic tomography constrains changes between image frames to be somehow "sensible" – precluding unrealistic jump changes or morphing of objects. Effectively, information from all of the sequential data subsets is incorporated into each image frame, so finer resolution can be regained.

Our contribution is to demonstrate the applicability of such 4D tomographic reconstruction techniques to single-channel video-SAR, describing some adaptations required. The effect is to regain fine cross-range resolution in sub-apertures and focus along-track motion.

Data model: Under a slow-moving target approximation, SAR data can be approximated as

$$d(f,\tau_k) = \int u(x)\tilde{S}(f)\exp\left[-4\pi \mathrm{i}(f+f_c)\Delta R(x,\tau_k)/c\right]dx, \qquad (1)$$

ignoring for simplicity (without loss of generality) the amplitude effects of antenna beam pattern and geometric spreading. Here, τ_k is the (slow) time at which the pulse reaches the scene centre, $u(x)$ is the reflectivity of the scatterers which are located at x at $\tau_k = 0$, f_c is the centre frequency of the narrow-band emitted pulse described in the (fast) time-domain as $s(\bar{t}) = \tilde{s}(\bar{t})\mathrm{e}^{-2\pi \mathrm{i} f_c \bar{t}}$, with $\tilde{S}$ the Fourier transform of the slowly-varying pulse envelope $s(\bar{t})$ (e.g. chirp) and f the baseband frequency. The (bi-static) differential range at slow-time τ_k is given by

$$\begin{aligned}\Delta R(x,\tau_k) =& R(x,\tau_k) - R(x_0,\tau_k),\\ R(x,\tau_k) =& \frac{1}{2}\big(|T_x(\tau_k) - (x+\delta x(x,\tau_k))|\\ &+ |R_x(\tau_k) - (x+\delta x(x,\tau_k))|\big),\end{aligned} \qquad (2)$$

with T_x and R_x the positions of the (possibly co-located) transmitter and receiver, respectively, x_0 is the scene reference point, and $\delta x(x,\tau)$ the displacement of a scatterer starting at x after slow-time τ. For a constant velocity model, we have simply $\delta x(x,\tau_k) = v(x)\tau_k$. The data model (1) assumes the receiver gates are set such that a scatterer at x_0 will have zero phase.

Discretisation of (1) in space x as well as frequency bin and assuming a constant velocity, the data model can be written succinctly as

$$\mathbf{d} = \mathrm{A}(\mathbf{v})\mathbf{u}, \qquad (3)$$

with $\mathbf{d} = [d(f_1,\tau_1), d(f_2,\tau_1), \ldots, d(f_{m-1},\tau_n), d(f_m,\tau_n)]^T$. Note that we do not need to form the dense matrix A, but only routines which perform the "direct" operation A$\mathbf{x}$ and the "adjoint" operation A$^*\mathbf{b}$ (equivalent to forming a matched filter SAR image). Splitting the data into T > 1 sub-apertures, the model (3) can be written as

$$\mathbf{d}_t = \mathrm{A}_t(\mathbf{v}_t)\mathbf{u}_t, \quad t = 1,\ldots,\mathrm{T}, \qquad (4)$$

where the matrix A_t simulates the data components of the t^{th} sub-aperture $\mathbf{d}_t$ (i.e. it contains the corresponding rows of A). We allow the scene $\mathbf{u}_t$ to vary between sub-apertures, as can the velocity field $\mathbf{v}_t$. Note that we assume slow-time is counted from $\tau_k = 0$ in each sub-aperture, so that $\mathbf{u}_t$ represents the scene at the start of the sub-aperture and not the start of the whole collection.

In the remainder of the paper, we use the model (4) to form video-SAR reconstructions with non-overlapping sub-apertures $\mathbf{d}_t$, for which each sub-aperture image $\mathbf{u}_y$ is able to make use of the full dataset $\mathbf{d} = [\mathbf{d}_1^T, \ldots, \mathbf{d}_\mathrm{T}^T]^T$ to retain fine resolution, automatically focussing both moving and stationary parts of the scene. Velocity is also able to vary between sub-apertures in the reconstructions, allowing for a degree of resolving manoeuvring targets.

Spatio-temporal image reconstruction: To use the model (4) to form video-SAR images, we consider the regularised least-squares reconstruction problem

$$(\tilde{\mathbf{u}},\tilde{\mathbf{v}}) = \underset{\mathbf{u},\mathbf{v}}{\operatorname{argmin}} \sum_{t=1}^{\mathrm{T}} \frac{1}{2}\|\mathrm{A}_t(\mathbf{v}_t)\mathbf{u}_t - \mathbf{d}_t\|_2^2 + \mathcal{R}(\mathbf{u},\mathbf{v}), \qquad (5)$$

where $\mathbf{u} = \left[\mathbf{u}_1^T, \ldots, \mathbf{u}_\mathrm{T}^T\right]^T \in \mathbb{C}^{\mathrm{T}N}$ and $\mathbf{v} = \left[\mathbf{v}_1^T, \ldots, \mathbf{v}_\mathrm{T}^T\right]^T \in \mathbb{R}^{2\mathrm{T}N}$ are the sequences of reconstructed complex SAR images $\mathbf{u}_i \in \mathbb{C}^N$ and of velocity $\mathbf{v}_i = [v_{1xi}, v_{1yi}, \ldots, v_{Nxi}, v_{Nyi}]^T \in \mathbb{R}^{2N}$, each on $N = N_x \times N_y$ pixels. The regularisation term $\mathcal{R}$ incorporates prior knowledge about the scene to provide a unique and stable reconstruction, for example promoting sparsity in compressive-sensing.

One common choice for $\mathcal{R}$ has been spatio-temporal TV[4, 6],

$$\mathcal{R}(\mathbf{u}) = \mathrm{TV}_{r,t}(\mathbf{u}) := \sum_{i,j,t}^{N_x,N_y,\mathrm{T}} \left(\sqrt{(D_x\mathbf{u})^2 + (D_y\mathbf{u})^2 + (D_t\mathbf{u})^2}\right)_{i,j,t}, \qquad (6)$$

where D_x, D_y and D_t are the (discrete) first derivative operators in x, y, and t. Here, one does not explicitly solve for velocity, but rather

$$\tilde{\mathbf{u}} = \underset{\mathbf{u}}{\operatorname{argmin}} \sum_{t=1}^{\tau} \frac{1}{2}\|\mathrm{A}_t(\cdot)\mathbf{u}_t - \mathbf{d}_t\|_2^2 + \lambda\mathrm{TV}(\mathbf{u}), \qquad \lambda > 0. \qquad (7)$$

which promotes solutions with only sparse changes between frames. However, without incorporating prior knowledge on "sensible" motion, (7) may be less able to improve resolution through coupling consecutive image frames where motion is present. A further challenge is that the model $\mathrm{A}(\cdot)$ requires an estimation of $\mathbf{v}$, although for short enough apertures this may be set at $\mathbf{v} = \mathbf{0}$.

More recently, improved results have been obtained by employing an optical flow constraint [5, 6], given in continuous form as

$$\partial_t u(r,t) + (\nabla_r u(r,t)) \cdot v(r,t) = 0. \qquad (8)$$

Discretisation of (8) suggests the spatio-temporal regularisation term

$$\mathcal{R}(\mathbf{u},\mathbf{v}) = \mathcal{M}(\mathbf{u},\mathbf{v}) := \sum_{t=1}^{\mathrm{T}-1} \|\mathbf{u}_{t+1} - \mathbf{u}_t + \left(\nabla^{\pm}\mathbf{u}_t\right)\cdot\mathbf{v}_t\|_2^2, \qquad (9)$$

where $\nabla^{\pm}$ denotes the discrete centred-difference Grad operator, and other choices of norm may be considered. In general, (8) is used in

conjunction with additional regularisation of $\mathbf{u}$ and $\mathbf{v}$ (such as spatial TV), and the resulting bi-convex optimisation problem in the form (5) is solved using an alternate convex search[6].

We have found that direct application of (9) to along-track focussing in SAR poses some difficulties due to the numerical differentiation. Compared to previous applications to dynamic reconstruction, in SAR we will generally have a small number of solid bodies moving over – and obscuring different parts of – a varying and cluttered background background, and this clutter may vary by observation angle. This is not well described by a diffeomorphism (i.e. smooth, differentiable, invertible map) between frames. We may also have many point-like objects (or parts of object) in the scene: employing a centred difference scheme will yield spurious results in these cases as objects are "jumped" over – yet forward or backward differences are unstable[6].

Our solution, similar to that of Dirks[8], is to use the brightness consistency condition directly,

$$u(r,t) = u(r+\delta r, t+\delta t), \tag{10}$$

for which (8) is a linearisation of. Here, δr is the displacement of the local image region at r, t after time δt. (10) suggests use of the regularisation term

$$\mathcal{R}(\mathbf{u},\mathbf{v}) = \mathcal{W}(\mathbf{u},\mathbf{v}) := \sum_{t=1}^{\mathrm{T}-1} \|\mathrm{W}(\mathbf{v}_t)\mathbf{u}_t - \mathbf{u}_{t+1}\|_2^2, \tag{11}$$

where the linear operator $\mathrm{W}(\mathbf{v})$ warps the scene by the displacement between image frames due to the velocity field $\mathbf{v}$ via a cubic interpolation scheme. (11) will be used in reconstruction of the reflectivity frames $\mathbf{u}_t$, but along-track velocity will be estimated via an alternative sparse Lucas-Kanade optical flow method[9] applied to pre-processed absolute imagery, which is more suited to the highly sparse solid-body motion.

The effect of (11) in an iterative scheme will be to coherently sum across SAR video frames with the correct warping for moving targets.

Algorithm overview: As with the previously noted spatio-temporal reconstruction methods, the image $\mathbf{u}$ and velocity field $\mathbf{v}$ will be updated in an alternating search. The problem in complex SAR image $\mathbf{u}$ reads

$$\begin{aligned}\mathbf{u}^{[i+1]} &= \operatorname*{argmin}_{\mathbf{u}} \sum_{t=1}^{\tau} \frac{1}{2}\|\mathrm{A}_t(\mathbf{v}_t^{[i]})\mathbf{u}_t - \mathbf{d}_t\|_2^2 \\ &\qquad + \lambda_1 \mathrm{TV}_r(\mathbf{u}) + \lambda_2 \mathcal{W}(\mathbf{u},\mathbf{v}^{[i]}), \\ &:= \operatorname*{argmin}_{\mathbf{u}} \mathcal{J}(\mathbf{u};\mathbf{v}^{[i]})\end{aligned} \tag{12}$$

with $\lambda_1, \lambda_2 > 0$, $\mathbf{v}^{[i]}$ the i^{th} estimate of the velocities, and where TV_r is spatial TV,

$$\mathrm{TV}_r(\mathbf{u}) := \sum_{i,j}^{N_x,N_y} \left(\sqrt{(D_x\mathbf{u})^2 + (D_y\mathbf{u})^2}\right)_{i,j}. \tag{13}$$

Equation (12) and the linear operators therein have been implemented in the open-source CCPi Core Imaging Library (CIL) framework[4, 10], providing us with a standardised suite of reconstruction algorithms and routines as well as a level of abstraction of the underlying numerical linear algebra for more rapid prototyping. The CIL implementation of the Primal Dual Hybrid Gradient algorithm (PDHG)[11] was used to solve this optimisation problem, without tailoring of the default parameters.

Given an estimate $\mathbf{u}^{[i]}$ of the image sequence, image frames are first pre-processed to filter stationary targets following the approach of Pastina *et al*[3]. Velocities are then estimated on the processed frames using the OpenCV[12] implementation of the sparse Lucas-Kanade optical flow method to determine velocities of image chips, and the estimates projected onto the along-track direction. Optionally, the extracted velocities are also fit to a linearly-varying (i.e. constant acceleration) profile, which we find helps with robustness with larger displacements or varying velocity profiles. The extracted velocities are then applied to the target locations in the velocity fields $\mathbf{v}_t^{[i+1]}$, which are zero elsewhere.

We note that this moving target extraction and tracking might be readily substituted with any other method for velocity estimation in SAR, allowing easy integration with existing processing schemes.

The resulting TV-Optical Flow (TV-OF) algorithm is summarised in Fig 1. In general we find that only a small number of outer iterations are required, say $i \leq 4$. Gaining an initial estimation of velocities in early iterations allows subsequent iterations to focus moving targets via the warping optical flow regularisation term relatively quickly, with subsequent iterations providing only marginal improvement. A prior estimation of the velocity field might reduce this further.

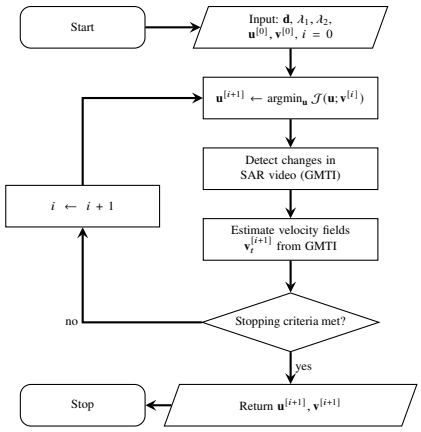


Fig 1 *Overview of TV-Optical Flow reconstruction scheme.*

Numerical results: Data is simulated for a scene containing 3 targets consisting of point scatterers arranged as a cross inside a box, with 60 cm maximum separation between point scatterers. Bandwidth and monostatic aperture length is selected for a range and cross-range resolution of 20 cm. Thus, if the data is divided into at least 3 sub-apertures they will be less than a resolution cell apart. Non-dimensionalising slow-time such that the whole aperture is considered to take $\tau = 1$ unit time, the first target has a cross-range starting speed of $1.25\,\mathrm{m}/\tau$ and is decelerating at $0.25\,\mathrm{m}/\tau^2$, the second has a starting cross-range speed of $0.95\,\mathrm{m}/\tau^2$ and accelerating at $0.1\,\mathrm{m}/\tau^2$: both travel through more than 5 resolution cells. The third is stationary. Simulated speckle clutter is added to give a signal-to-clutter ratio in the data of approximately -1 dB. The resulting normalised matched filter image of the whole aperture is shown in Fig 2.

Image formation is of 10 non-overlapping sequential sub-apertures. This number chosen to challenge the ability to regain cross-range resolution, but could be adapted to target motion. Fig 3 shows 3 of the 10 frames for each of matched filter (b), spatio-temporal TV reconstructions (c), and optical flow-spatial TV reconstructions (d). Also shown for comparison are matched filter images with double-length (i.e. 50% overlapping) sub-apertures in Fig 3 (a).

Both TV and TV-OF reconstruction techniques are able to regain cross-range resolution, as well as reduce the background clutter, which is clear by comparison to the matched filter images. However, the TV method is unable to adequately combine information from multiple sub-apertures to accurately reconstruct the targets without artefacts, whereas the TV-OF method is clearly able to reproduce these: for TV-OF, we can clearly make out the individual point scatterers, apparently regaining the resolution of the full aperture. TV-OF is also able to more closely reach the absolute value of reflectivities than TV (noting the colour scales), although still underestimates slightly: tailoring the regularisation parameters in both cases might improve this result.

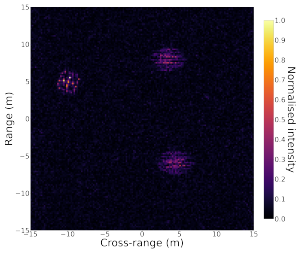
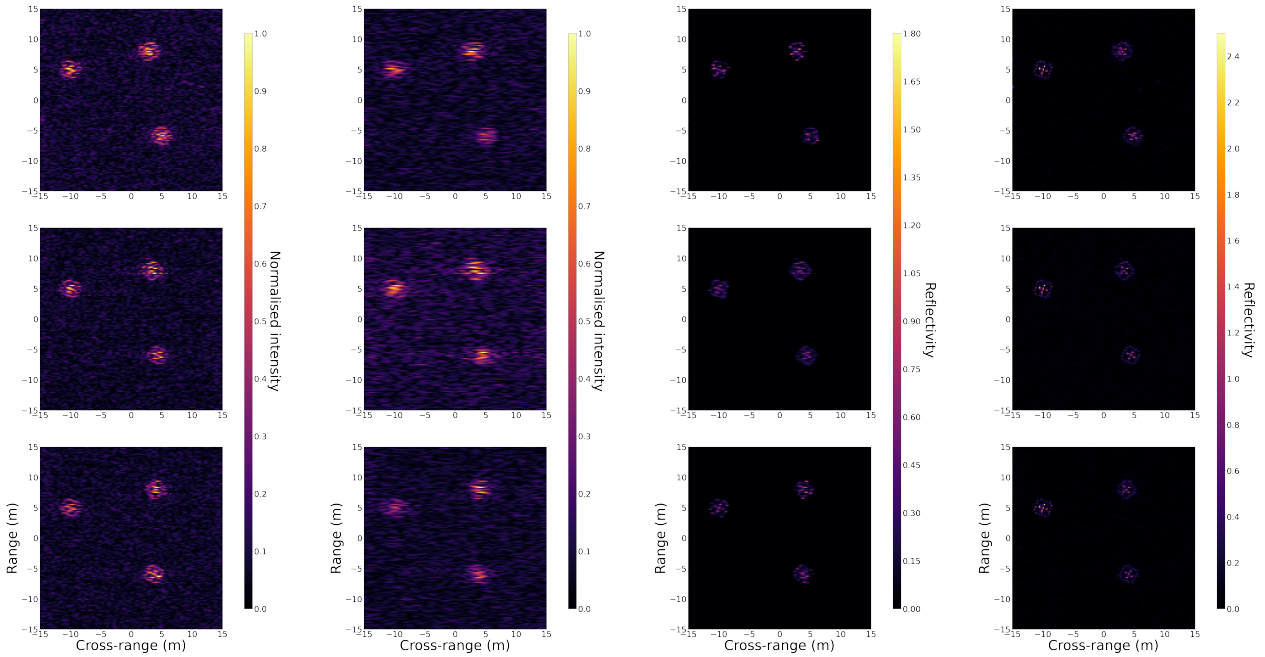


Fig 2 *Normalised matched filter SAR image of simulated dataset of 2 moving and 1 stationary target.*



(a) *Matched filter, double length sub-apertures*

(b) *Matched filter, disjoint sub-apertures*

(c) *Spatio-temporal total variation reconstruction*

(d) *Optical flow-spatial total variation reconstruction*

Fig 3 *Absolute value (detected) image frames 2, 5 and 8 (top to bottom) of 10 from the video-SAR sequence, for two slowly moving and accelerating/decelerating targets, and one stationary target. Comparing matched filter with double-length (i.e. 50% overlapping) sub-apertures (a), matched filter with non-overlapping sub-apertures (b), spatio-temporal TV reconstructions (c), and optical flow-spatial TV reconstructions (d).*

The price to pay is of course increased computational cost with the number of frames in the video (which might be increased for faster targets), the number of inner optimisation iterations (here PDHG), and the number of outer iterations. Adaptive strategies could be considered to determine these parameters for best performance, as well as alternative optimisation procedures for the inner loop which may be more efficient (e.g. Lucka *et al* prefer the ADMM algorithm for larger problems[6]).

Conclusion: We have presented a method for focussing along-track motion in video-SAR. This is based upon a joint reconstruction of both image intensity and a velocity field (constrained to the along-track direction), coupled by an optical flow-based regularisation term, similar to as has been used in other 4D tomography problems. The result of the method is to coherently combine information from each sub-aperture image frame with the correct object displacements for moving targets, regaining fine cross-range resolution of the full aperture and focussing cross-range motion whilst retaining a SAR video. Simulated results have validated the potential of the approach. This might equally apply to combining multi-look imagery containing moving targets

In future work we will apply the method to real data, in particular considering the multi-static case to combine multiple estimates of velocity components from the different along-track directions. This will also enable more straightforward use with curved or varying flight paths. We will also consider combining the dynamic reconstruction method with alternative velocity estimation techniques.

Acknowledgments: This work has been supported by the Royal Academy of Engineering and the Office of the Chief Science Adviser for National Security under the UK Intelligence Community Postdoctoral Research Fellowship programme, and has made use of computational support by CoSeC, the Computational Science Centre for Research Communities, through CCPi.

Received: 09 Aug 2022 *Accepted:* DD MMMM YYYY
doi: 10.1049/ell.10001

References

1. Axelsson, S.: Position correction of moving targets in SAR imagery. In: SAR Images Analysis, Modeling, and Techniques VI, vol. 5236, pp. 80–92. International Society for Optics and Photonics (2004)
2. Kirscht, M.: Detection and imaging of arbitrarily moving targets with single-channel SAR. In: RADAR 2002, , pp. 280–285. (2002)
3. Pastina, D., Battistello, G., Aprile, A.: Change detection based GMTI on single channel SAR images. In: 7th European Conference on Synthetic Aperture Radar, , pp. 1–4. (2008)
4. Papoutsellis, E., et al.: Core Imaging Library - Part II: multichannel reconstruction for dynamic and spectral tomography. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 379(2204), 20200193 (2021). doi:10.1098/rsta.2020.0193
5. Burger, M., Dirks, H., Schönlieb, C.B.: A variational model for joint motion estimation and image reconstruction. SIAM Journal on Imaging Sciences 11(1), 94–128 (2018). doi:10.1137/16M1084183
6. Lucka, F., et al.: Enhancing Compressed Sensing 4D Photoacoustic Tomography by Simultaneous Motion Estimation. SIAM J. Imaging Sci. 11, 2224–2253 (2018)
7. Beauchemin, S.S., Barron, J.L.: The computation of optical flow. ACM computing surveys (CSUR) 27(3), 433–466 (1995)
8. Dirks, H.: Joint large-scale motion estimation and image reconstruction. ArXiv abs/1610.09908 (2016)
9. Lucas, B.D., Kanade, T., et al.: An iterative image registration technique with an application to stereo vision. vol. 81. Vancouver (1981)
10. Jørgensen, J.S., et al.: Core Imaging Library - Part I: a versatile Python framework for tomographic imaging. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 379(2204), 20200192 (2021). doi:10.1098/rsta.2020.0192
11. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision 40, 120–145 (2010)
12. Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools (2000)