THREE-DIMENSIONAL MOTION ESTIMATION AND IMAGE FORMATION WITH ACTIVE ARRAYS

Vincent R. Radzicki and Hua Lee
University of California, Santa Barbara
Department of Electrical and Computer Engineering
Santa Barbara, CA, 93106
vradzicki@ucsb.edu

Faculty Advisor:
Hua Lee

7/13/2019

ABSTRACT

For target imaging and tracking systems, a key signal processing task is motion estimation. Specifically, the trajectory of a maneuvering target undergoing rigid body motion can be described through a series of translational and rotational transformations. Estimation of these motion parameters provides the tracking system enough information to calculate the targets trajectory over time. Determining the rotational motion to a high accuracy is also very important, as the imaging system can then form an image of the target over multiple aspect angles and thus increase the resolution performance. This paper focuses on algorithm development and performance limitations for motion estimation and image formation using active sensing arrays.

INTRODUCTION

Active sensing arrays, like radar and sonar, are pivotal tools for detection and tracking of both cooperative and non-cooperative moving targets. The information acquired from these systems can be used for a variety of different applications such as target tracking, navigation, and identification. A core step for these applications is estimation of the target's motion to aid in determining its overall trajectory. Target motion estimation and tracking for sensor arrays is a research area with a rich and deep technical background. Although, position tracking can be achieved with extremely high accuracy, issues remain for estimating targets' rotational position and motion during the scan. In this paper, we present a review of array sensing and imaging techniques for motion estimation and focus on algorithm design issues that can be optimized for the different stages of data processing.

ACTIVE SENSOR ARRAYS AND MOVING TARGETS

Target detection and localization is an important function for sensor systems, such as radar and sonar, in a variety of different applications in surveillance, navigation, and security [1, 2]. Depending on the target of interest, the system will employ either active or passive sensing. Active sensing requires the system to transmit a probing waveform and measure the target's reflections to extract the tracking information. Passive sensing in contrast does not actively transmit any waveforms, and instead only receives signals either generated by the target or from other independent sources. Active sensors typically provide greater target information and provide system engineers more parameters to configure to achieve optimal performance, however, require greater system complexity and cost. The primary advantage active systems possess is the ability to accurately estimate target range and direction, whereas passive systems typically focus on estimating direction only. In this work, we will explore the design and implementation of active sensor arrays due to these advantages.

A single sensor element can perform target ranging through precise timing measurements of the reflected waveforms in comparison to the transmit waveform. The exact ranging resolution depends on the waveform parameters, specifically the frequency bandwidth. Many target structures will scatter incoming probing waveforms at several discrete spatial points. This point-scattering model is accurate for the wavelengths and target size for most active sensor systems [3, 4]. Given high enough ranging-resolution, the sensor system can not only measure the mean range to the target, but can also identify each of these point-scatterers to a precise range measurement.

To obtain further information on the direction from the sensor to the target, an array of sensors are utilized together to extract this out. Many of these techniques are described as direction-finding or angle-of-arrival algorithms [5]. The angular resolution the system is capable of achieving depends on the array parameters, and is typically proportional to the size of the array. Due to this relationship, the direction parallel to the length of the array is typically called the cross-range direction, and the cross-range resolution is also proportional to the length of the array, but will change depending on the range to the target. Processing the received signals across the array, allows for isolating the returns of the point scatterers and attributing them to precise cross-range positions. Given a sufficiently large aperture design, processing the received signals across the entire array can produce a Range & Cross-Range image of the target under investigation.

It is desirable for the sensor system to not only determine the target's position, but also estimate its motion over the observation period. The sensor measurements collected over the entire array is referred to as one data-frame, and can produce a subsequent image frame through range and cross-range processing. The probing waveforms are often comprised of a sequence of frequency pulses known as Frequency-Modulated Continuous Wave (FMCW), so the raw-data acquired is referred to as Frequency-Pulse data [6, 7]. The frames are acquired at discrete time intervals or periods T, so the estimation task is to determine the target's motion between sequential frames. A model of the discrete target motion can be seen in Fig. 3, where a simple target is displayed at two separate times. The target has multiple key scattering points (x_i, y_i, z_i), and the points are subject to a rotational R and translational (Δ_x, Δ_y) transformation between frames. Although the target is following a continuous trajectory, it's motion can be described as a sequence of discrete rigid body motion transformations applied at each frame time T. Many targets of interest are subject to rigid body motion due to their mechanical structure, and will now be examined in more detail.

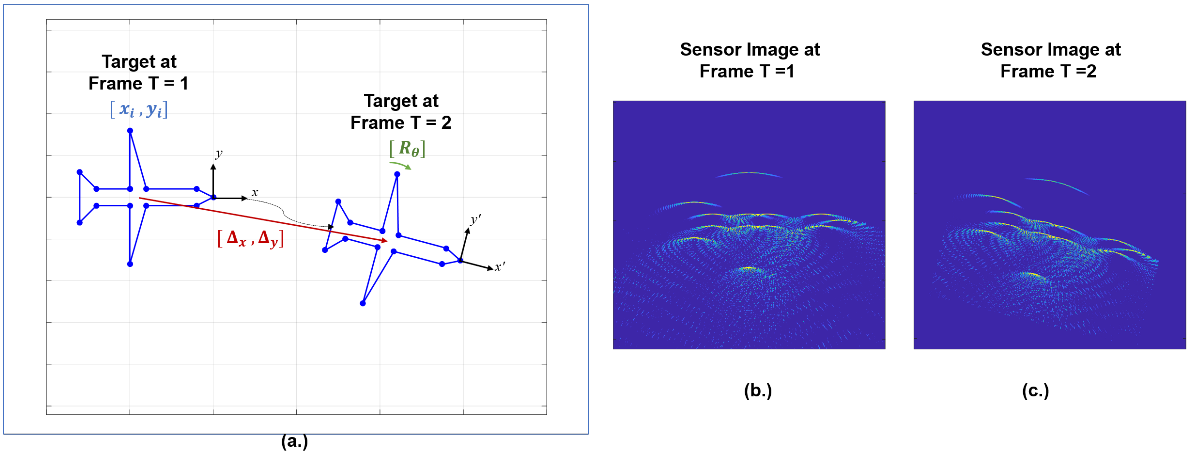

Figure 1: Moving Target and Motion Parameters at sequential frames.

For rigid body motion, the set of scattering points (x_i, y_i, z_i) at frame $T = 2$ can be related to the same scattering points for the previous at frame $T = 1$ through the matrix equation

$$\begin{bmatrix} x_i[2] \\ y_i[2] \\ z_i[2] \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\beta & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma \\ \sin\alpha\cos\beta & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma \\ -\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma \end{bmatrix} \begin{bmatrix} x_i[1] \\ y_i[1] \\ z_i[1] \end{bmatrix} + \begin{bmatrix} \Delta_x \\ \Delta_y \\ \Delta_z \end{bmatrix}$$

where α, β, and γ represent the angular motion about the yaw, pitch, and roll axis respectively, and Δ_x, Δ_y, and Δ_z is the translational shift in 3D Cartesian space. Both axis are referenced to the center of mass of the original target. This can be written more compactly in the form of

$$\mathbf{P[2]} = \mathbf{R} \cdot \mathbf{P[1]} + \mathbf{\Delta_{xyz}} \tag{1}$$

This parameterization is very useful, as the motion estimation problem can be effectively solved through estimation of the translational and rotation parameters $\mathbf{R}$ and $\mathbf{\Delta_{xyz}}$ at every time frame T. The results can be concatenated and processed further to determine the overall trajectory.

Current techniques estimate the Cartesian motion from sequential data frames, and track the target either through mechanical or electrical adjustment of the array. Oftentimes, these techniques employ long coherent integration of scans of the target to yield Doppler velocity and positional estimates, which are then inserted into a tracking filter [8]. Results are displayed showing the target trajectory via a Plan Position Indicator (PPI), or overlaid onto some other geo-mapping software. These systems usually require relatively long integration times that the scanning beam must remain on the target. Although the position of the target can be estimated to high precision, the rotational motion and relative aspect angle of the target to the array is not explicitly determined. Some systems rely on target's rotational motion to form a larger effective aperture (inverse-synthetic aperture), to form high-resolution target imagery and then deduce the rotational motion afterwards, but these algorithms are highly dependent on low-order approximations of the true target motion [9]. An alternative approach that explicitly estimates both the target's rotational and Cartesian motion at discrete time intervals could provide operators more freedom in target tracking applications and also allow for better benchmarking of system performance in terms of array design, waveform parameters, and target structure.

SENSOR ARRAY MOTION ESTIMATION TECHNIQUES

In this section we outline two approaches to motion estimation for sensor arrays. They can be defined into two broad categories depending on the processing implementation and data availability. The first approach focuses on processing the Frequency-pulse data (raw data) acquired from the array. For this, both target image localization and target motion estimation can be dealt with concurrently. The second approach is more intuitive, and presumes that the array processing has occurred and images of the target are formed at subsequent frames. Motion estimation requires applying the signal processing to the preformed target imagery.

A. Frequency-Pulse Data Domain Method

One frequency pulse data frame is comprised of the full set of measurements acquired at each sensor element in the array. The data frame at time T will be $E[k,i,T]$ where k corresponds to the transmit-illumination frequency and i is the sensor number. From this data, a coarse image of the target can be computed to map the target's reflectively over Cartesian space. Areas corresponding the main scattering centers will be of high amplitude, and areas with no reflections will be of low amplitude in the produced imagery. The coordinates that the target image is reconstructed over is a grid of 3D spatial points $(x[n], y[m], z[l])$, where the grid resolution is determined by the size and shape of the aperture along with other system parameters.

The image at frame T is formed by back-projection of the received data weighted with its inverse propagation kernel of the form

$$\begin{aligned} I[n,m,l,T] &= \sum_i \sum_k E[k,i,T] \cdot exp\left(-j\frac{4\pi}{\lambda_k}\sqrt{(x[n]-x[i])^2+(y[m]-y[i])^2+(z[l]-z[i])^2}\right) \\ &= \sum_i \sum_k E[k,i,T] \cdot h(x[n]-x[i], y[m]-y[i], z[l]-z[i]) \end{aligned} \tag{2}$$

where the $h(x[n]-x[i], y[m]-y[i], z[l]-z[i])$ is the propagation kernel shifted to the ith sensor's position in the array $(x[i], y[i], z[i])$ [10].

To perform motion estimation, a slight alteration of this procedure can be implemented. Given a sequential pair of data frames, the current image at frame T can be formed using Eq. 2, and the previous image at frame $T-1$ can be formed to a grid of the same size but where the coordinates axis are transformed by the rotational and translation shift parameters $\mathbf{R}$ and $\mathbf{\Delta_{xyz}}$. This transformed axis is $(x'[n], y'[m], z'[l])$, and the reconstruction will be

$$I'[n,m,l,T-1] = \sum_i \sum_k E[k,i,T-1] \cdot h(x'[n]-x[i], y'[m]-y[i], z'[l]-z[i]) \tag{3}$$

The two successive 3D matrices produced from Eq. 2 and Eq. 3 can be correlated against one another using a similarity metric as the motion parameters are adjusted. A simple cross-correlation would be as such:

$$C[\Delta_x, \Delta_y, \Delta_z, \alpha, \beta, \gamma] = \sum_{m,n,l} I_T^*[m,n,l] \cdot I'_{T-1}[m,n,l] \tag{4}$$

The parameters that maximize the correlation between the two grids will then be the optimal estimate [11]. Other correlation or similarity metrics may be employed here depending on the complexity of the target, but these various techniques will all need to employ some form of numerical optimization [12]. This optimization should solve the problem shown in Eq. 5, using a max search procedure or potentially a more specialized gradient-based approach. The maximum search procedure can be difficult given the number of parameters and potential lack of well-defined extrema in the correlation function. These issues may be overcome using apriori knowledge on target dynamics to help limit the search space, and better determine the accuracy needed for sufficient tracking.

$$[\hat{\Delta}_x, \hat{\Delta}_y, \hat{\Delta}_z, \hat{\alpha}, \hat{\beta}, \hat{\gamma}] = \text{argmax } C[\Delta_x, \Delta_y, \Delta_z, \alpha, \beta, \gamma] \tag{5}$$

This approach requires little pre-processing so that the amount of algorithmic artifacts to be encountered is low, but the 3D correlation and maximum search procedure can be very computationally expensive, especially for high-accuracy requirements. To overcome this a two-stage method could be implemented, in where a coarse estimate is first computed by applying a far-field, small angle-approximation of the propagation kernel. This would allow for exploitation of fast Fourier techniques and other symmetric transforms. Secondly, a fine estimate can be computed using the more exact model that would fine-tune the coarse estimate to some desired fitting criteria. A proper balance in efficiency between the coarse and fine stage would require numerical experimentation, however, these regression techniques have seen success in other difficult optimization tasks [13]. Lastly, although this procedure is of high computational cost, it can be parallelized well and done concurrently with image formation, so that if the system has large amounts of computational resources near real-time performance may be feasible. An overview of the proposed methodology is shown in Fig. 2.

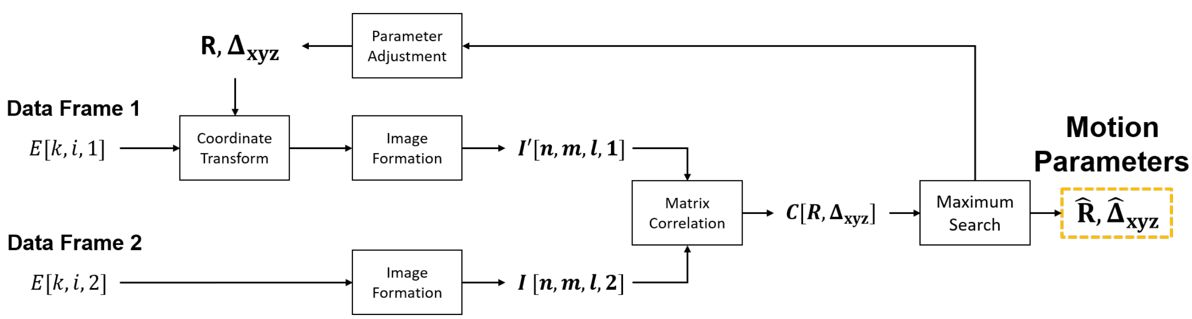

Figure 2: Frequency-Pulse Data Domain Based Motion Estimation

B. *Image Data Domain*

The approach described here is more intuitive, but also more down-stream from the raw data. Many systems perform image formation at lower software or firmware levels, so that high-level functions like motion estimation will only have the image data to process. The motion estimation problem is then to determine the motion parameters from a sequential pair of preformed images. This imagery may be interpolated, thresholded, and filtered for improved storage and visualization, but processing artifacts can be introduced and some information may be lost. There is however a deep vein of technical work on motion estimation in image processing that is drawn upon for these tasks that can overcome some of these limitations [14].

The primary methodology for image-based motion estimation is a two-step feature-extraction and registration procedure [15]. First, a set of "key-feature" points are extracted from both images and the correct correspondence between them must be determined before subsequent action. The set of N feature points from the image at time T can be stored in matrix form like

$$\mathbf{P}[T] = \begin{bmatrix} x_1[T] & x_2[T] & \dots & x_N[T] \\ y_1[T] & y_2[T] & \dots & y_N[T] \\ z_1[T] & z_2[T] & \dots & z_N[T] \end{bmatrix} = [\mathbf{p}_1[T], \mathbf{p}_2[T], \dots, \mathbf{p}_i[T], \dots, \mathbf{p}_N[T]]$$

where $\mathbf{p}_i$ is the spatial vector describing each key feature point. These feature points should be closely related to the scattering centers on the target structure described earlier. The centroid of the set of feature points should then be removed from each vector through

$$\mathbf{p}'_i[T] = \mathbf{p}_i[T] - \sum_i^N \mathbf{p}_i[T] \quad (6)$$

$$\mathbf{P}'[T] = [\mathbf{p}'_1[T], \mathbf{p}'_2[T], \dots, \mathbf{p}'_N[T]] \quad (7)$$

Centroid removal than leaves only a rotational relationship between the corresponding feature points $\mathbf{P}'[T]$ and $\mathbf{P}'[T-1]$ defined as

$$\mathbf{P}'[T] = \mathbf{R} \cdot \mathbf{P}'[T-1] \quad (8)$$

The immediate goal of feature extraction and centroid removal is estimation of the rotation parameter $\mathbf{R}$. To achieve this, a Singular Value Decomposition (SVD) can be taken of the product of the two set of feature points producing

$$\mathbf{P}'[T] \cdot \mathbf{P}'[T-1] = \mathbf{U}\boldsymbol{\Lambda}\mathbf{V} \quad (9)$$

For a set of well-defined feature points, this quantity is well-defined, and the least-squares solution for the rotation matrix $\mathbf{R}$ has been shown to be

$$\hat{\mathbf{R}} = \mathbf{U}\mathbf{V}^T \quad (10)$$

Lastly, the translation shift parameter can be found by rexaming the original feature points and applying the estimated rotation matrix. The estimate in this case will be of the form

$$\hat{\boldsymbol{\Delta}}_{\mathbf{xyz}} = \frac{1}{N}\left[\sum_i \mathbf{p}_i[T] - \mathbf{R}\sum_i \mathbf{p}_i[T-1]\right] \quad (11)$$

These procedure provides a numerically stable and computationally efficient method for calculating the full set of motion parameters from a set of sequential images formed by the array system. Although this approach has seen much success for image processing on camera systems, the approach does not always translate well to array imaging systems. The algorithm is highly dependent on image quality and having a consistent set of "good-feature" points to track. This is not given for all targets, and the performance can degrade drastically for targets that are not feature-rich. Also, the image quality and tracking points can vary significantly over the observation period as the target aspect angle changes, resulting in unexpected and significant drops in performance. If this performance variation can be accounted for, the computational efficiency of the image-based approach makes it an attractive option for target motion estimation.

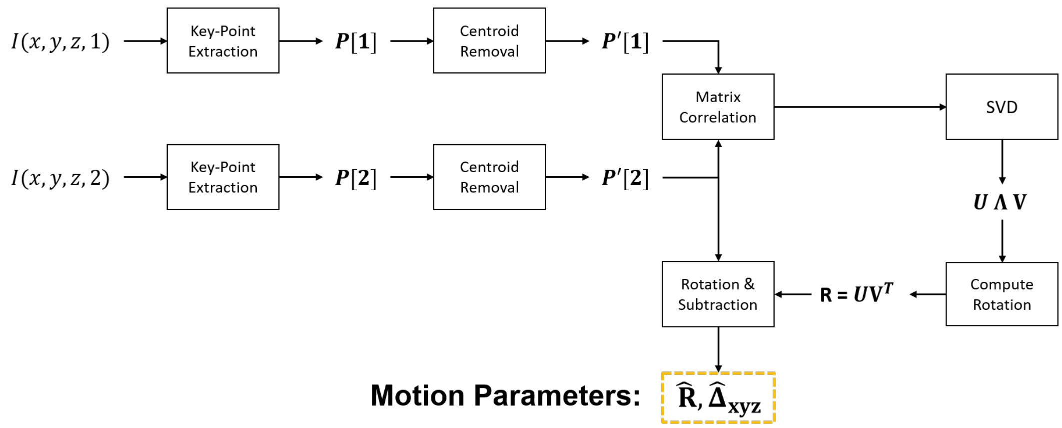

Figure 3: Image Data Domain Based Motion Estimation

CONCLUSIONS

In this paper, we presented an overview of motion estimation procedures and their limitations for active sensor array signal processing. The impact of these procedures has a wide-variety of applications and could provide another option for system designers in various fields. The review was meant to highlight the trade-offs between the methodologies and connect them to the key system parameters in terms of performance. To summarize these were:

- Frequency-Pulse Data Domain: Method that will introduce few processing artifacts and produce accurate results. Requires high-amount of computational complexity, but the accuracy is reliable, consistent, and predictable for varying target trajectories and structure.
- Image Data Domain: Method can produce high accuracy estimation results for reasonable computational complexity, but depends highly on image quality and target trajectory. Performance can vary significantly and degrade quickly as targets maneuver.

The frequency-pulse data approach provides many advantages but is costly in terms of computation, whereas the image-domain approach can be effective but performance can vary. Future work will involve numerical comparison of these proposed techniques, and exploring system parameter analysis on how the design of the arrays and waveforms affect overall accuracy.

REFERENCES

[1] D. Reid, "An algorithm for tracking multiple targets," *IEEE transactions on Automatic Control*, vol. 24, no. 6, pp. 843–854, 1979.

[2] G. Y. Delisle and H. Wu, "Moving target imaging and trajectory computation using isar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 3, pp. 887–899, 1994.

[3] M. Cheney and B. Borden, "Microlocal analysis of gtd-based sar models," in *Algorithms for Synthetic Aperture Radar Imagery XII*, vol. 5808, pp. 15–23, International Society for Optics and Photonics, 2005.

[4] N. N. Youssef, "Radar cross section of complex targets," *Proceedings of the IEEE*, vol. 77, no. 5, pp. 722–734, 1989.

[5] S. Haykin, "Radar array processing for angle of arrival estimation," *IN: Array signal processing (A85-43960 21-32). Englewood Cliffs, NJ, Prentice-Hall, Inc., 1985, p. 194-292. Research supported by the Canadian Department of Communications and NSERC.*, pp. 194–292, 1985.

[6] A. G. Stove, "Linear fmcw radar techniques," in *IEE Proceedings F (Radar and Signal Processing)*, vol. 139, pp. 343–350, IET, 1992.

[7] M. I. Skolnik, "Introduction to radar systems," *New York, McGraw Hill Book Co., 1980. 590 p.*, 1980.

[8] G. van Keuk and S. S. Blackman, "On phased-array radar tracking and parameter control," *IEEE Transactions on aerospace and electronic systems*, vol. 29, no. 1, pp. 186–194, 1993.

[9] C.-C. Chen and H. C. Andrews, "Target-motion-induced radar imaging," *IEEE Transactions on Aerospace and Electronic Systems*, no. 1, pp. 2–14, 1980.

[10] H. Lee, "Formulation of the generalized backward-projection method for acoustical imaging," *IEEE Transactions on Sonics and Ultrasonics*, vol. 31, no. 3, pp. 157–161, 1984.

[11] H. Lee and G. Wade, "A holographic approach for motion measurement," *IEEE transactions on sonics and ultrasonics*, vol. 30, no. 5, pp. 328–329, 1983.

[12] G. C. Carter, "Coherence and time delay estimation," *Proceedings of the IEEE*, vol. 75, no. 2, pp. 236–255, 1987.

[13] P. Stoica, R. L. Moses, *et al.*, "Spectral analysis of signals," 2005.

[14] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[15] H. Lee, *Acoustical Sensing and Imaging*. CRC Press, 2016.