Iowa State University

From the Selected Works of Sarah A. Rajala

November 1, 1992

Tracking of unresolved targets in infrared imagery using a projection-based method

Jae-Ho Choi, *North Carolina State University*

Sarah A. Rajala, *North Carolina State University*

Available at: https://works.bepress.com/sarah_rajala/24/

Tracking of unresolved targets in infrared imagery using a projection-based method

Jae-Ho Choi
Sarah A. Rajala

Department of Electrical and Computer Engineering
North Carolina State University
Raleigh, NC 27695-7911, USA
Email: jhc@clr.ece.ncsu.edu or sar@eos.ncsu.edu

ABSTRACT

The conventional two-dimensional Hough transform technique is generalized into a projection-based transform method by using the modified Radon transform for estimating a three-dimensional target tracks embedded in a time-sequential set of image frames. The targets of concern are dim, unresolved point targets moving along straight paths across a same field of view. Since the target signal-to-noise is low and the spatial extend of the target is less than a pixel, one must rely on integration over a target track which span over many image frames. Instead of processing the entire 3-D data set, a set of projections are taken using the modified 3-D Radon transform. The projection frames are processed further to extract the track parameters using the Hough transform. This projection-based method not only lowers the data dimensionality but maintains a comparable estimation performance to that of using the entire 3-D data by successfully incorporating all available knowledge obtained from the set of projections. The simulation results are presented for the synthetic and real infrared image sequences containing synthetically generated 3-D target tracks under various signal-to-noise conditions.

1. INTRODUCTION

In 1962, Hough¹ introduced a method for detecting complex patterns of points in binary image data by determining specific values of parameters characterizing these pattern. Since then, the Hough transform has been extended to estimate the parameters of analytic curves in gray scale images². One application area in which the Hough transform has been successfully used is tracking of airborne, point targets acquired by the satellite-based optical/infrared sensor arrays. For tracking unresolved point targets, two-dimensional digital image frames of the same field of view are collected at short, periodic intervals. The target, if present, is assumed to move approximately perpendicular to the sensor line of sight and moves through the field of view. If the target is in the field of view, it appears in a succession of frames as it moves while the background stays relatively constant. In each frame in which the target appears, it shows up in one or few pixels, depending on its size, its speed, and the point spread function of the optics. The acquired set of image sequences contain noise and, if there is a target, a signal of constant or time-varying amplitude.

Cowart et al.³ applied a traditional straight line finding Hough transform to detect non-maneuvering target tracks. Later, Padgett et al.⁴ extended Cowart et al.'s algorithm to allow for the detection of maneuvering circular target tracks using a generalized Hough transform. Each of these algorithms has been evaluated using a set of two-dimensional track map sequences. A track map sequence is obtained by projecting a preprocessed time sequence of image frames along the temporal direction onto one 2-D track map

image. The track map may still contain some noise; however, it is assumed that the remaining noise is white. There is a limitation, however, in overall performance when using temporal projection to create a 2-D track map sequence from the original 3-D (spatio-temporal) image data, especially when the targets are dim. Even if the projection is performed optimally⁵, it causes a significant reduction in the effective SNR as compared to the SNR of the original 3-D data. An alternative approach would be to process the entire 3-D volume of spatio-temporal data⁶. This could potentially increase the detection and tracking performance somewhat but at the expense of significant more computation. This paper presents an extension of the Hough transform-based algorithm based on the modified Radon transform. The presented projection-based method effectively circumvent the shortcomings of the 2-D track map sequences while maintaining a much lower computation than that of processing the entire 3-D volume.

2. PROJECTION-BASED HOUGH TRANSFORM

In this section the derivation of a projection-based Hough transform algorithm is presented. Here, the Radon transform⁷ provides a method for obtaining sets of projections at arbitrary angles. Computing the Radon transform consists of computing the projections of an image along a particular pattern, e.g., a straight line. Mathematically the problem can be stated as follows. Let $T(x, y, t)$ represent a three-dimensional function. The integral of $T(x, y, t)$ along a ray is called a *ray integral*, and a set of ray integrals forms a *projection*.

The equation for an individual ray is given by

$$k_\phi = y \cos \phi - x \sin \phi \quad (1)$$

where k_ϕ is the perpendicular distance of the ray from the origin. The integral of the function $T(x, y, t)$ along this ray may be expressed as

$$P(k_\phi, t) = \int_{\text{plane}} T(x, y, t) d^3\mathbf{r} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T(x, y, t) \delta(y \cos \phi - x \sin \phi - k_\phi) dx dy dt \quad (2)$$

The two-dimensional function $P(k_\phi, t)$ as a function of k_ϕ and t is the surface integral of $T(x, y, t)$ over the plane for angle ϕ . The function $P(k_\phi, t)$ is the three-dimensional Radon transform of $T(x, y, t)$.

Instead of using surface integration as in Eq. 2, the maximum value projection method is to be applied in this implementation as follows:

$$P_m(k_\phi, t) = F_\phi[T(x, y, t)] = \text{Max}[T(x, y, t) \delta(y \cos \phi - x \sin \phi - k_\phi), \quad \forall t] \quad (3)$$

$P_m(k_\phi, t)$ denotes a set of parallel projections taken along a series of parallel planes defined by Dirac-delta function. $P_m(k_\phi, t)$ as a function of k_ϕ and t for a given value of ϕ will be referred to as the parallel projection of $T(x, y, t)$ for angle ϕ (see Figure 1).

Another unique set of parallel projections can be taken along

$$k_\psi = t \cos \psi - y \sin \psi \quad (4)$$

which leads to

$$P_m(k_\psi, x) = F_\psi[T(x, y, t)] = \text{Max}[T(x, y, t) \delta(t \cos \psi - y \sin \psi - k_\psi), \quad \forall x] \quad (5)$$

Figure 1: Radon-like forward parallel projection.

Theoretically, increasing the number of projections improves the performance of the estimation, but increases the cost of computation. Any desired number of projections can be taken uniformly or non-uniformly given $0 \leq \phi \leq \frac{\pi}{2}$ and $0 \leq \psi < \frac{\pi}{2}$. In the following, let the number of projections be M and N for Eq. 3 and Eq. 5, respectively.

Once a set of the parallel projections are obtained using the forward projection defined by Eq. 3 and/or Eq. 5, the straight-line track parameters for each of the parallel projection frames are estimated using the 2-D Hough transform. An estimate of the 3-D target trajectory in 3-D is obtained by back-projecting the $M + N$ 2-D track parameters estimates $(\rho'_i, \theta'_i)_{i=0,1,\dots,M-1}$ and $(\rho'_j, \theta'_j)_{j=0,1,\dots,N-1}$. The back-projection shown in Figure 2 is defined as follows:

$$T'(x, y, t) = \sum_i F_\phi^{-1}[T'_i(k_\phi, t)] + \sum_j F_\psi^{-1}[T'_j(k_\psi, x)] \quad (6)$$

where $T'_i(k_\phi, t)$ and $T'_j(k_\psi, x)$ are the estimated two-dimensional target track functions obtained by the inverse Hough transforms

$$T'_i(k_\phi, t) = k_\phi \cos \theta'_i + t \sin \theta'_i - \rho'_i = 0, \quad (7)$$

and

$$T'_j(k_\psi, x) = k_\psi \cos \theta'_j + x \sin \theta'_j - \rho'_j = 0, \quad (8)$$

and $T'(x, y, t)$ is the reconstructed three-dimensional target trajectory function using parallel back-projection along two sets of parallel projection planes defined by the Dirac delta functions in Eq. 1 and Eq. 4.

The accuracy of three-dimensional target track estimation solution is limited by the input image frame signal-to-noise ratio and the discretization errors (data sampling and quantization). Discretization errors occur when the data is recorded discretely and when there is a projection onto a plane that is not parallel to one of the spatial or temporal planes.

Figure 2: Reconstruction using Radon-like parallel line back-projection.

2.1 Analytical Bounds on the Hough Space Errors

The analytical bounds on the Hough space parameter errors that are introduced by image space noise contamination is summarized. The complete derivation can be found in Rajala et al.⁶.

Assuming a 3-D volume of image data is generated from a time-sequence of 2-D image frames, a set of parallel projection frames can be obtained from the 3-D volume. A target track (line) in a 2-D projection frame is modeled by:

$$\rho_i = x_i \cos \theta_i + y_i \sin \theta_i \quad (9)$$

where x_i and y_i are the image domain coordinates for an arbitrary 2-D projection and ρ_i and θ_i are the corresponding Hough space parameters. If the data are noisy, there will be errors in the estimates of the Hough space parameters. The objective of this section is to determine the bounds on the errors for ρ_i and θ_i . In general, the amount of error will depend on the geometry of the track (see Figure 3).

2.2 Error Bound on $\bar{\rho}_i$

A bound on the estimate of ρ_i can be specified as a function of θ_i as follows:

$$(x_i \mp K \Delta x_i) \cos \theta_i + (y_i - K \Delta y_i) \sin \theta_i \leq \bar{\rho}_i \leq (x_i \pm K \Delta x_i) \cos \theta_i + (y_i + K \Delta y_i) \sin \theta_i \quad (10)$$

where $\bar{\rho}_i$ is the estimate of ρ_i .

Figure 3: Geometry associated with determining error bounds on ρ_i of target track.

Consider the problem geometry shown in Figure 3. Here

$$\begin{aligned}\rho_{iU} &\equiv \text{upper bound of estimate } \bar{\rho}_i \\ \rho_{iL} &\equiv \text{lower bound of estimate } \bar{\rho}_i\end{aligned}$$

The bound on the estimate of ρ_i is as follows:

(1) For $0 \leq \theta \leq \frac{\pi}{2}$:

$$|\Delta \rho_i| \leq |\rho_i - \bar{\rho}_{iU}| = |\rho_i - \bar{\rho}_{iL}| \leq K \sqrt{\Delta x_i^2 + \Delta y_i^2} \cos |\theta_i - 45^\circ| \quad (11)$$

(2) For $\frac{\pi}{2} < \theta_i < \pi$:

$$|\Delta \rho_i| \leq K \sqrt{\Delta x_i^2 + \Delta y_i^2} \cos |\theta_i - 135^\circ| \quad (12)$$

(3) For $\frac{3\pi}{2} < \theta_i < 2\pi$:

$$|\Delta \rho_i| \leq K \sqrt{\Delta x_i^2 + \Delta y_i^2} \cos |\theta_i - 315^\circ| \quad (13)$$

Combining these three results,

$$|\Delta \rho_i| \leq \sqrt{\epsilon_1^2 + \epsilon_2^2} \cos |\delta| \quad (14)$$

where,

$$\delta = \begin{cases} \theta_i - 45^\circ, & \text{for } 0 \leq \theta_i \leq \frac{\pi}{2} \\ \theta_i - 135^\circ, & \text{for } \frac{\pi}{2} < \theta_i < \pi \\ \theta_i - 315^\circ, & \text{for } \frac{3\pi}{2} < \theta_i < 2\pi \end{cases} \quad (15)$$

If we assume $\epsilon = \epsilon_1 = \epsilon_2 \leq K\Delta x_i = K\Delta y_i = K$, (that is, $\Delta x_i = \Delta y_i = 1$) then,

$$|\Delta \rho_i| \leq K\sqrt{2} \cos |\delta| \quad (16)$$

2.2 Error Bound on $\bar{\theta}_i$

Assuming the length of the actual target track is L , the upper and lower estimation error bounds on $\bar{\theta}_i$, with δ defined as in Eq. (15). The error bound for $\bar{\theta}_i$ can be expressed as follows:

$$|\Delta \theta_i| = |\theta_i - \bar{\theta}_i| \leq \left| \frac{\pi}{2} - \tan^{-1} \left(\frac{L}{\sqrt{2\epsilon_1^2 + \epsilon_2^2} \cos |\delta|} \right) \right| \quad (17)$$

Again consider the case when $\epsilon = \epsilon_1 = \epsilon_2 \leq K\Delta x_i = K\Delta y_i = K$. Then,

$$|\Delta \theta_i| \leq \left| \frac{\pi}{2} - \tan^{-1} \left(\frac{L}{2K\sqrt{2} \cos |\delta|} \right) \right| \quad (18)$$

3. SIMULATION RESULTS

Two classes of image data sets were used in the target tracking system simulation. The first one is a synthetically generated image with a constant background. From this synthetic data, a sequence of image frames are composed to create a 3-D data set. The second one is a 3-D data set which is made of real infrared data provided by NRL (HiCamps). Two noise free images taken from the HiCamps are shown in Figure 4. The data set has a frame size of 32×32 pixels and up to 45 (diagonal) frames in the temporal direction in order to accommodate the longest possible track of a target moving at a speed of one pixel per frame. Since a 3-D data set with a real target is not available, the target points are also synthetically generated and superimposed on the sequence of image frames as a test target trajectory. Under the assumption that a target is containable to a pixel (target spatial extend less than a pixel), a target point along the 3-D line trajectory is added to each one of frames of the 3-D data as follows:

$$S_{i,j,k} = \begin{cases} \alpha C + (1 - \alpha)B_{i,j,k} & \forall i, j, k \in T(x, y, t) \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

$S_{i,j,k}$ is an intensity function that has a value of zero except at the location of the target. C is target signal intensity (set as a constant for a preliminary experiment) and $B_{i,j,k}$ is a background pixel intensity that may take a value from 0 to 255. The variable α ($0 \leq \alpha \leq 1$) controls the target intensity contribution to the pixel, i.e., for $\alpha = 1$, the target intensity C completely replaces the whole pixel. To observe the algorithm with a dim target, we set α at 0.45. An additive, white Gaussian noise with a certain variance is finally added to above created noise free 3-D data sets.

Figure 4: The noise free images taken from the HiCamp image sequence I (left) and the sequence II (right), respectively.

As a measurement of the algorithm performance, the pixel signal-to-noise ratio is used. The pixel SNR is defined as follows:

$$SNR = 10 \log_{10} \left(\frac{\frac{\alpha^2}{L} \cdot \sum_{i,j,k \in T(x,y,t)} [C - B_{i,j,k \in T(x,y,t)}]^2}{\sigma_n^2} \right) \quad (20)$$

where L is the total number of target points (frames) in the 3-D volume of data; α is the target signal contribution factor; C is the target signal intensity; $B_{i,j,k \in T(x,y,t)}$ is a set of background pixel values along the target trajectory $T(x,y,t)$; and σ_n^2 is the variance of the additive, white Gaussian noise.

First, the simulation was run on a sequence of synthetic data to evaluate the performance of the projection-based algorithm for a number of different combinations of signal-to-noise ratios, projections, and target frames. In this stage the two-dimensional track parameters (ρ, θ) were estimated in a 2-D space to test if the projection-based algorithm is justifiable. This was done by taking a projection of the 3-D reconstructed volume of the estimated 3-D track along the temporal direction onto the xy-plane. The 3-D test track has the 2-D Hough parameters at $\rho = 28$ and $\theta = 45^\circ$ on the xy-plane. The simulation results are summarized in Table 1. While the SNR changes from the relatively high value (14.1 dB) to the low value (5.6 dB) given 13 to 19 time-sequential frames of data, the algorithm performance was tested for two (2 and 6) different number of projections. When the SNR is above the 10 dB, the number of projection is not the factor if the number of projection is 2 or more. However, as the SNR sinks below 8 dB, the target tracking system performance gets comparably better with higher number of projections (5-6).

Furthermore, the simulation was run to estimate the three-dimensional linear track parameters using the sets of time-sequential frames taken from the HiCamp sequence I. The test target track has the 3-D parameter vector at (1.00, -1.00, 1.00) in x, y, and temporal direction, respectively. Figure 5 shows one of the typical projection-based 3-D target track estimation procedure using the the 3-D reconstructed volume, which is the back-projection of the set of estimated 2-D tracks from each one of projection frames. On the contrary, the simulation objective was to determine the average number of projections required for reliable

PROJ	Frame	Estimated 2-D Hough Parameters (ρ, θ)				
		14.1 dB	12.3 dB	8.10 dB	6.90 dB	5.60 dB
2	13	(28, 44°)	(28, 42°)	(30, 63°)	(30, 63°)	(27, 66°)
	15	(28, 42°)	(28, 42°)	(30, 63°)	(30, 63°)	(27, 66°)
	17	(29, 45°)	(29, 45°)	(30, 63°)	(30, 63°)	(27, 66°)
	19	(29, 47°)	(29, 47°)	(30, 63°)	(30, 63°)	(27, 66°)
6	13	(28, 45°)	(28, 51°)	(28, 50°)	(5, 130°)	(5, 130°)
	15	(28, 45°)	(28, 46°)	(28, 50°)	(28, 52°)	(4, 132°)
	17	(28, 45°)	(28, 45°)	(28, 45°)	(28, 51°)	(28, 52°)
	19	(28, 45°)	(28, 45°)	(28, 45°)	(28, 45°)	(28, 52°)

Table 1: 2-D Hough parameter estimation of a target track with true parameters (28,45°) in a sequence of a synthetic image data: The track estimation performance is tested for a number of different combinations of the pixel SNR (dB), target frames, and projections (PROJ). Better 2-D track estimation performance was observed overall given 6 projections than that of 2 projections for SNRs below 8 dB.

3-D track parameter extraction. The tracking accuracy was measured in terms of the average error distance from the true 3-D trajectory. Figure 6 shows the results for two different values of signal-to-noise ratio. From the simulation, it was determined that an average error distance below 1.0 yielded reliable 3-D parameter estimates. In other words, the projection-based algorithm for the 3-D target tracking requires minimum of 4 projections or an average of 6-7 projections given 23 frames of data. As shown in Figure 6, good 3-D target tracking performance was observed for SNRs down to 7.0 dB. The performance does depend on the variation of the image background and the formation of target track.

From the simulation, it is evident that at low SNR's the algorithm performance degrades as the number of frames decreases and/or the number of projections decreases. However, for the three-dimensional track parameter estimation using a reconstructed 3-D track volume, a reliable track parameter extraction was performed down to an average pixel SNR of 7-8 dB with 6-7 projections and 20-23 frames of data.

4. CONCLUSION

The generalized projection-based Hough transform algorithm for tracking dim, unresolved target tracks was presented. The modified 3-D Radon transform was used as a tool to project a time-sequential set of the infrared image frames into two sets of 2-D projection frames of data. The sets of 2-D projection frames were further processed using the 2-D Hough transform to obtain the estimate of the 3-D track parameters. For the dim target, the presented projection-based algorithm had a superior tracking ability compared with that of using the track map sequences taken only along the temporal direction; and the estimation performance was comparable to that of processing the entire 3-D volume of data at a highly reduced number of computation. Simulation results showed that good tracking of a 3-D target trajectories is possible down to the SNR of 7 dB given twenty-three frames although fewer number of frames were required to obtain the correct parameters at higher SNRs. Because the projections can be taken in an arbitrary angle in parallel and the Hough transformation are performed on each projection frame separately, the tracking computation can be distributed using a parallel computing architecture such that the track parameters are real-time obtainable.

5. REFERENCES

1. P.C.V. Hough, "Method and means for recognizing complex patterns," *U.S. Patent 3,069,654*, December 18, 1962.
2. Richard O. Duda and E. Hart, "Use of the Hough transform to detect lines and curves in pictures," *Communications of Association of Computing Machinery*, vol. 15, no. 1, pp. 11-15, January, 1972.
3. A.E. Cowart, Wesley E. Snyder, and W. Howard Ruedger, "The detection of unresolved targets using the Hough transform," *Computer Vision, Graphics, and Image Processing*, vol. 21, pp. 222-238, 1983.
4. Mary L. Padgett, Sarah A. Rajala, Wesley Snyder, and W. Howard Ruedger, "Detection of maneuvering target tracks," *Proceedings of SPIE*, vol. 575, pp. 145-155, August, 1985.
5. Peter L. Chu, "Optimal projection for multidimensional signal detection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 5, pp. 755-786, May, 1988.
6. Sarah A. Rajala, Loren W. Nolte, and James V. Aanstoos, "Detection of unresolved target tracks in infrared imagery," *Proceedings of SPIE*, vol. 1606, pp. 360-371, November, 1991.
7. Michael P. Ekstrom, *Digital Image Processing Techniques*, Academic Press, Orlando, Florida, 1984.

Figure 5: 3-D target track estimation (gray points) from the reconstructed 3-D track volume (small dark points) when the true 3-D trajectory (big dark points) vector is at $(1.00, -1.00, 1.00)$ in x, y , and temporal direction, respectively, given 23 frames of data for the SNR of 7.0 dB: The estimated 3-D linear target track has a 3-D vector at $(1.00, -1.02, 1.07)$; and it is on the average of less than 0.86 pixel distance from the true target trajectory.

Figure 6: The 3-D target track estimation performance in terms of average error distance between the true and the estimated 3-D linear trajectory for the different number projections with respect to the SNR values of 7.0 dB and 12.2 dB. 19 and 23 frames of the Hicamp sequence I were used with the true 3-D target trajectory vector at (1.00, -1.00, 1.00) in x, y, and temporal direction for each different SNR. For reliable tracking performance, it requires at least 6-7 projections with over 20 frames of data when the target is dim (SNR below 8 dB).