

MTI Data Clustering and Formation Recognition

MARK J. CARLOTTO, Senior Member, IEEE
Veridian Corporation

Image exploitation technology approaches have generally focused on the detection and spatial analysis of stationary groups of objects on the ground using various sensors. While spatial arrangement is clearly necessary in analyzing military formations, it is usually not sufficient. Typically the arrangement must be examined within some context in order to interpret a pattern of deployment. For moving objects the spatial arrangement of the group relative to the direction of motion is key to recognizing the formation. By examining ground moving target indicator (MTI) radar data over time, motion can be inferred and used to establish a context for interpreting the spatial arrangement of the data. New techniques that exploit the multitemporal nature of MTI data are described. The first is a space-time clustering technique that locates compact groups of objects that persist in time. The technique is an application of Marr and Hildreth's edge detection methodology to the dual problem of region segmentation, or more accurately, volumetric segmentation of space-time. The second technique is based on the use of the Hough transform for recognizing moving formations such as columns, wedges, and lines abreast by analyzing the shape of clustered MTI detections (specifically the orientation of linear arrangements within the group) with respect to their direction of motion. Preliminary results from simulated MTI data sets are presented.

Manuscript received August 18, 1999; revised September 29, 2000; released for publication October 8, 2000.

IEEE Log No. T-AES/37/2/06327.

Refereeing of this contribution was handled by L. M. Kaplan.

Author's address: 1400 Key Blvd., Suite 100, Arlington, VA 22209,
E-mail: (markc@psrw.com).

0018-9251/01/\$10.00 © 2001 IEEE

I. INTRODUCTION

Military analysts require a detailed picture of maneuvering forces on the battlefield to understand enemy force structure and disposition, and to recognize critical deployments, maneuver events, and arrangements (e.g., offensive combat formations). Image exploitation technology approaches have generally focused on the detection and spatial analysis of stationary groups of objects using synthetic aperture radar (SAR) and electro-optical (EO) sensors. Moving target indicator (MTI) radars [14] are powerful reconnaissance and intelligence resources because of their capability to detect moving objects over large areas. Together with SAR, MTI can potentially provide a more complete picture of the battlefield [6, 9, 18].

While spatial arrangement is clearly necessary in analyzing military formations, it is usually not sufficient. Typically the arrangement must be examined within some context (e.g., the direction and distance to the forward line of troops) in order to interpret a pattern of deployment. For moving objects the spatial arrangement of the group relative to the direction of motion is key to recognizing the formation. For example, a convoy and a line abreast are both linear formations. What differentiates one from the other is their arrangement with respect to their movement. The difference is very important to the analyst as the convoy is a common arrangement that a group of vehicles assume when they are moving from one location to another, while the line abreast is an offensive combat formation.

The multitemporal nature of MTI data invites pattern recognition approaches that exploit spatial arrangement and motion. By examining multiple frames of MTI over time, motion can be inferred and used to establish a context for interpreting the spatial arrangement of the data. Conventional SAR and EO sensors provide only snapshots in time of the battlefield. Typically only a subset of the vehicles in a group is observed in a given image because of terrain obscuration, camouflage, and other factors. The repetitive aspect of MTI increases the chance that most of the vehicles within a group will eventually be observed over a period of time thus improving the probability of detection and correct classification.

The present work focuses on techniques for analyzing groups of moving objects in terms of their pattern and movement. New techniques that exploit MTI are described. Section II briefly discusses techniques for analyzing and visualizing MTI data. A space-time clustering technique that locates compact groups of moving objects which persist in time is described in Section III. The technique is an application of Marr and Hildreth's edge detection methodology to the dual problem of region segmentation, or more accurately, volumetric

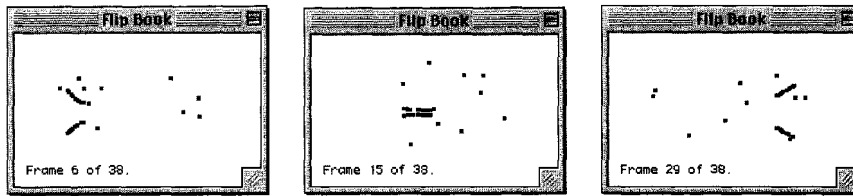


Fig. 1. Visualization of MTI data one frame at a time.

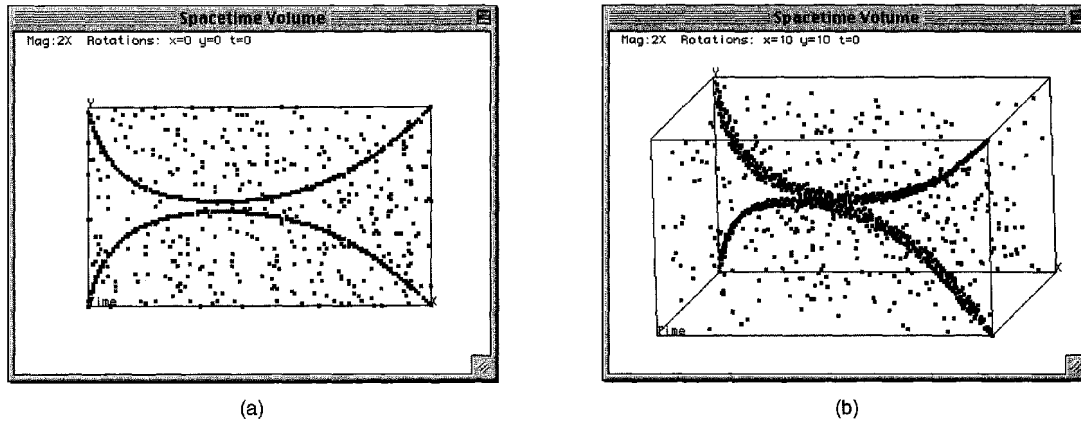


Fig. 2. Visualization of MTI data over time. (a) Accumulated over time. (b) Space-time volumetric visualization.

segmentation of space-time. Section IV presents a technique based on the use of the Hough transform for recognizing moving formations such as columns, wedges, and lines abreast by analyzing the shape of clustered MTI detections (specifically the orientation of linear arrangements within the group) with respect to their direction of motion. Formation recognition results from simulated MTI data sets are presented in Section V.

II. INTERPRETING MTI DATA

MTI reports are usually organized as a sequence of frames, each containing a set of detections (moving objects) observed at a given time. Flipbooks are a familiar metaphor for viewing MTI sequentially, one frame at a time as shown in Fig. 1.

In practice the unaided analyst often has difficulty discerning and interpreting MTI evidence of organized movement in wide area search data. Analysts sometimes find it helpful to display multiple frames of MTI accumulated over time (Fig. 2(a)) in order to better detect and associate groups of moving objects and to visualize their shape and motion. Multiple target tracking techniques [2] which use kinematic models to predict where a moving object might go from one frame to the next have been developed to assist the analyst in associating MTI reports over time. 2-D clustering techniques have also been developed to extract groups of moving targets on a frame-by-frame basis [15].

Another way to visualize MTI data is within a 3-D space-time volume consisting of two spatial dimensions plus time (Fig. 2(b)). Viewing MTI in this way suggests another approach to the association problem - one which involves clustering MTI data in 3-D space-time. Instead of using a tracking paradigm, this approach is based on the observation that most phenomena of interest in the MTI domain are spatially compact and persistent in time, i.e., they form clusters in space-time.

III. MTI DATA CLUSTERING

A variety of traditional techniques exist for clustering multivariate data. Algorithms such as *K*-means [17] involve choosing *K* initial clusters, assigning each sample to the nearest cluster (e.g., in terms of their Euclidean distance), updating the parameters of the clusters, and continuing until the process converges. When the number of clusters are not known hierarchical procedures (top-down splitting or bottom-up merging) are useful [5]. Agglomerative (bottom-up) clustering starts with one sample per cluster, computes the distances between all clusters, links the two nearest clusters, and repeats until a single cluster remains. When the distance metric between clusters is the maximum distance between the members of the clusters the algorithm encourages the formation of clusters that are compact and roughly equal in size; when the distance metric

is the minimum distance between the members of the clusters the algorithm encourages the formation of elongated clusters. The nearest neighbor algorithm is better suited to MTI data where compact and persistent groups of objects tend to form elongated clusters in space-time. A problem with the algorithm is its sensitivity to noise (background clutter) which can cause the algorithm to generate spurious links between clusters (chaining effect).

A disadvantage of clustering in general is the large number of data comparisons involved. For images and other spatial data sets image segmentation techniques provide a compute-efficient alternative to clustering. Image segmentation partition images into regions, lines, and other primitive objects [1]. For example, region growing merges neighboring pixels that are similar in value into connected regions. Because region growing involves only local operations it involves considerably less computation than clustering.

Edge detection finds the boundaries between regions and is also a local operation. Marr and Hildreth [11] developed an approach to edge detection that uses the locations of zero-crossings of the Laplacian of an image convolved with a Gaussian to define edges. Lee and Rodriguez [10] developed an automated algorithm for segmenting magnetic resonance images into different tissue types using a 3-D difference of Gaussians (DOG) filter. We apply the Marr-Hildreth approach in a similar way to the volumetric segmentation of space-time. MTI data are clustered by convolving the set of MTI detections in space-time with a Laplacian of a Gaussian (LoG) filter, labeling connected regions where the convolution is less than zero, assigning MTI detections to clusters based on their region membership, and linking labeled 2-D regions to form 3-D clusters in space-time.

The following subsections describe our algorithm, present a model for estimating the distribution of 3-D cluster lengths generated from background clutter, and show how this model can be used as the basis for detecting persistent groups of objects in MTI.

A. Space-Time Clustering

Let $D = \{\delta(x_i, y_i, t_i)\}$ be the set of MTI detections where (x_i, y_i) are the spatial coordinates and t_i the time of the i th detection. The LoG is a space-time filter of the form

$$F = \nabla^2 G * D \quad (1)$$

where

$$\nabla^2 G = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} + \frac{\partial^2 G}{\partial t^2}$$

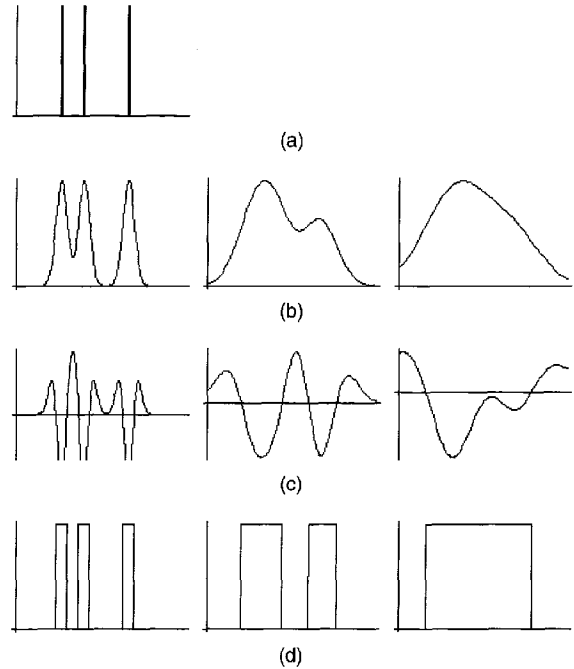


Fig. 3. 1-D clustering example showing effect of changing spatial smoothing parameter. (a) Three detections spaced 2 and 4 units apart. (b) Detections smoothed by Gaussian filter (left to right, $\sigma = 0.5, 1.5, 2.5$). (c) Second derivative of output from Gaussian filters. (d) Intervals (clusters) where second derivative in (c) is less than zero (left to right, 3 clusters each containing one detection per cluster, 2 clusters with 2 detections in one cluster and one in the other, and 1 cluster containing all 3 detections).

is the Laplacian operator. The Gaussian, factored into its space and time components, is

$$G = G_x G_y G_t \\ = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-y^2/2\sigma^2} \frac{1}{\sqrt{2\pi\tau^2}} e^{-t^2/2\tau^2}. \quad (2)$$

Connected regions in space-time where $F < 0$ are assigned unique labels and become distinct clusters. The detections assigned to the j th cluster are

$$C_j = \{\delta(x_i, y_i, t_i) \in R_j\} \quad (3)$$

where R_j is the region in space-time occupied by the cluster.

Instead of specifying the number of clusters (i.e., the number of groups of moving objects, which is in general unknown) or an objective function/algorithm for determining them, the number of clusters is controlled indirectly by the spatial and temporal smoothing factors, σ and τ , respectively. By varying these parameters clusters at different space and time scales can be extracted. Fig. 3 is a 1-D example illustrating the effect of varying the spatial parameter σ .

The LoG filter (Equation (1)) can be realized in different ways. Using 3-D arrays, the MTI detections

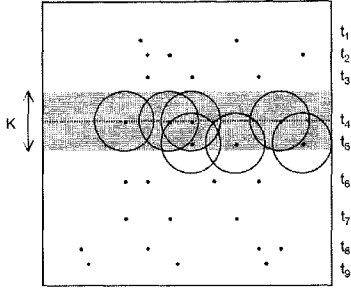


Fig. 4. MTI detections in 3-D space-time showing those contributing to a given time slice.

can be convolved with 1-D Gaussians in space and time, and the result convolved with a $3 \times 3 \times 3$ discrete approximation of the Laplacian ℓ operator:

$$\nabla^2 G * D = \ell * G_x * G_y * G_t * D. \quad (4)$$

The computational complexity of this approach is on the order of $MN^2(3K + 27)$ operations (adds and multiplies) where N is the height (and width) of the frame, M is the number of frames, and K is the size of the Gaussian operator (which should be at least 8 times the standard deviation to reduce truncation effects).

Another way to realize the space-time filter is to perform the computation sequentially using a 2-D array. The 3-D LoG impulse response is

$$H = \nabla^2 G = \left[\frac{\partial^2 G_x}{\partial x^2} G_y G_t + G_x \frac{\partial^2 G_y}{\partial y^2} G_t + G_x G_y \frac{\partial^2 G_t}{\partial t^2} \right]. \quad (5)$$

For efficiency it can be precomputed and stored in a lookup table. The space-time filter $F = H * D$ is then computed one frame at a time by superposing 2-D time slices of LoG impulse responses in a 2-D

accumulator array (Fig. 4). If the temporal extent of the filter is K , and there are on average P detections per frame, approximately PK^3 adds are required per frame. Because MTI data are relatively sparse, as the size of the frame increases this approach becomes increasingly more run-time efficient than the former.

Fig. 5 shows the steps involved in clustering a frame from the simulated MTI data shown in Fig. 1. (Recall this sequence showed two columns, each with 10 objects, which first move toward one another and then move apart. The simulated MTI probability of detection was $P_D^{MTI} = 0.9$ with $N_{FA}^{MTI} = 10$ false alarms per frame.) Detections from the current frame (a) and neighboring frames are convolved with the LoG operator (b). Regions where the convolution is less than zero are marked (c) and labeled (d). MTI detections are assigned clusters based on their region membership (e).

Multiple frames are clustered sequentially using label maps from the current and previous frames to link 2-D regions into 3-D clusters. The following labeling rules recognize the start of a new cluster, the continuation of an existing cluster, the end of a cluster, and the splitting and merging of clusters (Fig. 6).

- 1) If a region in the current label map does not intersect any region in the previous label map, assign new label (start new cluster).
- 2) If a region in the current label map intersects exactly one region from the previous label map, propagate label (continue cluster).
- 3) If a region in the previous label map does not intersect any region in the current label map, retire label (end cluster).
- 4) If a region in the previous label map intersects two or more regions in the current label map, retire

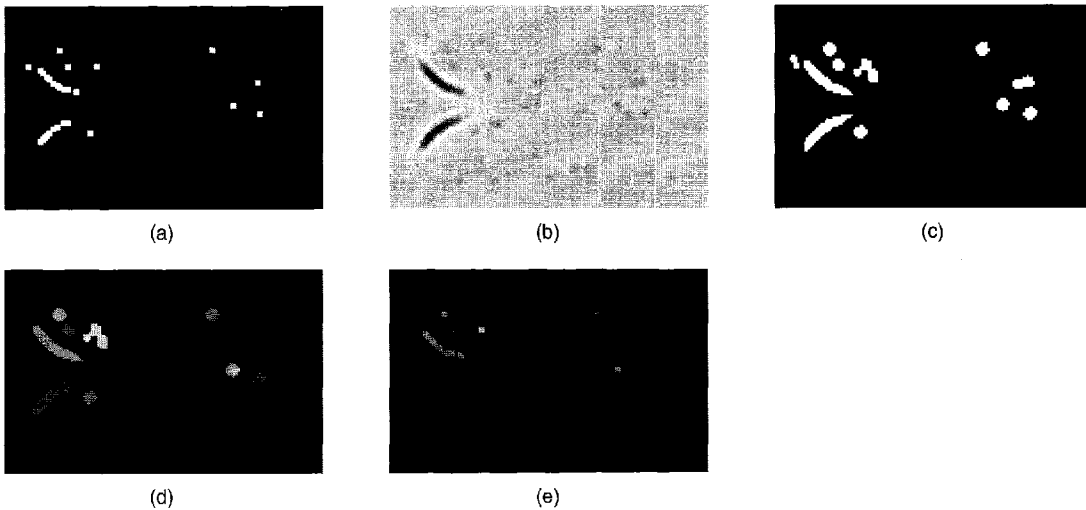


Fig. 5. Processing steps involved in clustering a single MTI frame. (a) MTI detections. (b) LoG convolution. (c) Regions where convolution is less than zero. (d) Label map (clusters). (e) Detections assigned to clusters.

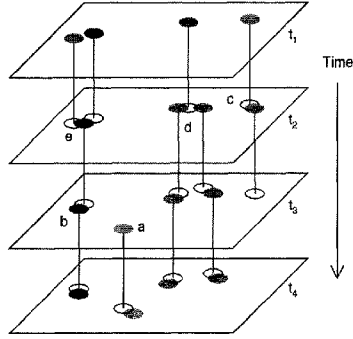


Fig. 6. Extracting 3-D clusters.

previous label and assign new labels to regions (split cluster).

5) If a region in the current label map intersects two or more regions in the previous label map, retire previous labels and start new region label (merge cluster).

In the presence of clutter groups of moving objects tend to become fragmented into a large number of short clusters (i.e., clusters that persist only for a few frames). A split/merge criterion is provided to reduce fragmentation. If the number of detections within a region is less than a threshold η , the region is not merged with, or split off from, an existing cluster.

B. Performance Modeling

That spatially compact and persistent groups of moving objects in MTI form clusters in space-time suggests the longest clusters extracted by the above algorithm are likely to be of interest. However in order to reliably differentiate persistent groups of objects from the background it is necessary to characterize the distribution of clusters generated by clutter.

It is assumed that clutter can be modeled as a Poisson process where the probability that k false alarms occur within an area s at some time t_0 is [12]

$$P[k] = e^{-\lambda s} \frac{(\lambda s)^k}{k!}. \quad (6)$$

The intensity parameter $\lambda = n/S$, where S is the area of the MTI frame and n is the MTI false alarm rate ($= N_{FA}^{MTI}$). We define

$$P = \sum_{m \geq k} e^{-\lambda s} \frac{(\lambda s)^m}{m!} \quad (7)$$

to be the probability that k false alarms or more occur within s at t_0 . For uncorrelated clutter in space-time the probability that k false alarms or more occur over the same area at two successive times t_0 and t_1 is P^2 . Over N frames the probability that this occurs K successive times, i.e., forms a cluster of length K is

$$p(K) = (N - K)P^K(1 - P)^{N-K}. \quad (8)$$

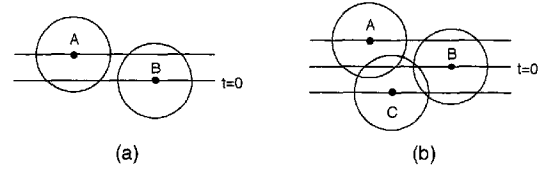


Fig. 7. Effect of neighboring detections in space-time. (a) Two detections A and B are not close enough to be grouped into the same cluster. Circles represent extent of clusters. (b) Presence of nearby detection C increases effective sizes of clusters A and B causing them to be grouped together.

Recall that 3-D clusters form when 2-D connected regions that are adjacent in time overlap (Fig. 6). For an isolated false alarm in space-time the area of its surrounding region (i.e., where the LoG response is less than zero) is $s = 2\pi\sigma^2$. As the MTI false alarm rate increases the LoG responses of neighboring false alarms in space-time begin to interact (Fig. 7). If $\tau > 0$ false alarms at other times can affect the cross-sectional area of a cluster at time t_0 and hence the probability that it will overlap adjacent 2-D regions to form 3-D clusters. As an upper bound we use

$$s = \sum_{t=-\tau}^{\tau} 2\pi\sigma^2 \left[1 - \left(\frac{t^2}{\tau^2} \right) \right] \quad (9)$$

to represent the added effect of neighboring false alarms in space-time on the area of a region at $t = 0$.

In order to validate our model we generated simulated clutter distributions over a sequence of 80 frames. Fig. 8 plots the predicted and actual distributions of cluster lengths for various parameter settings. In general longer clusters form as σ , τ , or n increases, or as η decreases. As σ or n increases, P increases (equation (7)). Increasing τ increases the area (equation (9)) which in turn increases P . Overall this model underestimates the actual distribution at lower MTI false alarm rates (Fig. 8(a)–(d)) and overestimates it at a higher MTI false alarm rates (Fig. 8(f)). The rapid decrease in the probability of long clusters forming from clutter as η increases appears to be modeled reasonably well (Fig. 8(e)).

C. Detecting Persistent Groups of Objects

We use the above clutter model as the basis for separating groups of persistent objects from the background. Let P_{FA}^{group} be a user-specified group constant false alarm rate (CFAR). The cluster threshold K_0 for a given false alarm rate satisfies

$$P_{FA}^{\text{group}} = 1 - \sum_{K=0}^{K_0} p(K). \quad (10)$$

Clusters whose length exceed K_0 are considered to represent persistent objects.

Fig. 9(a) is a space-time volumetric representation of a single moving object-embedded uncorrelated

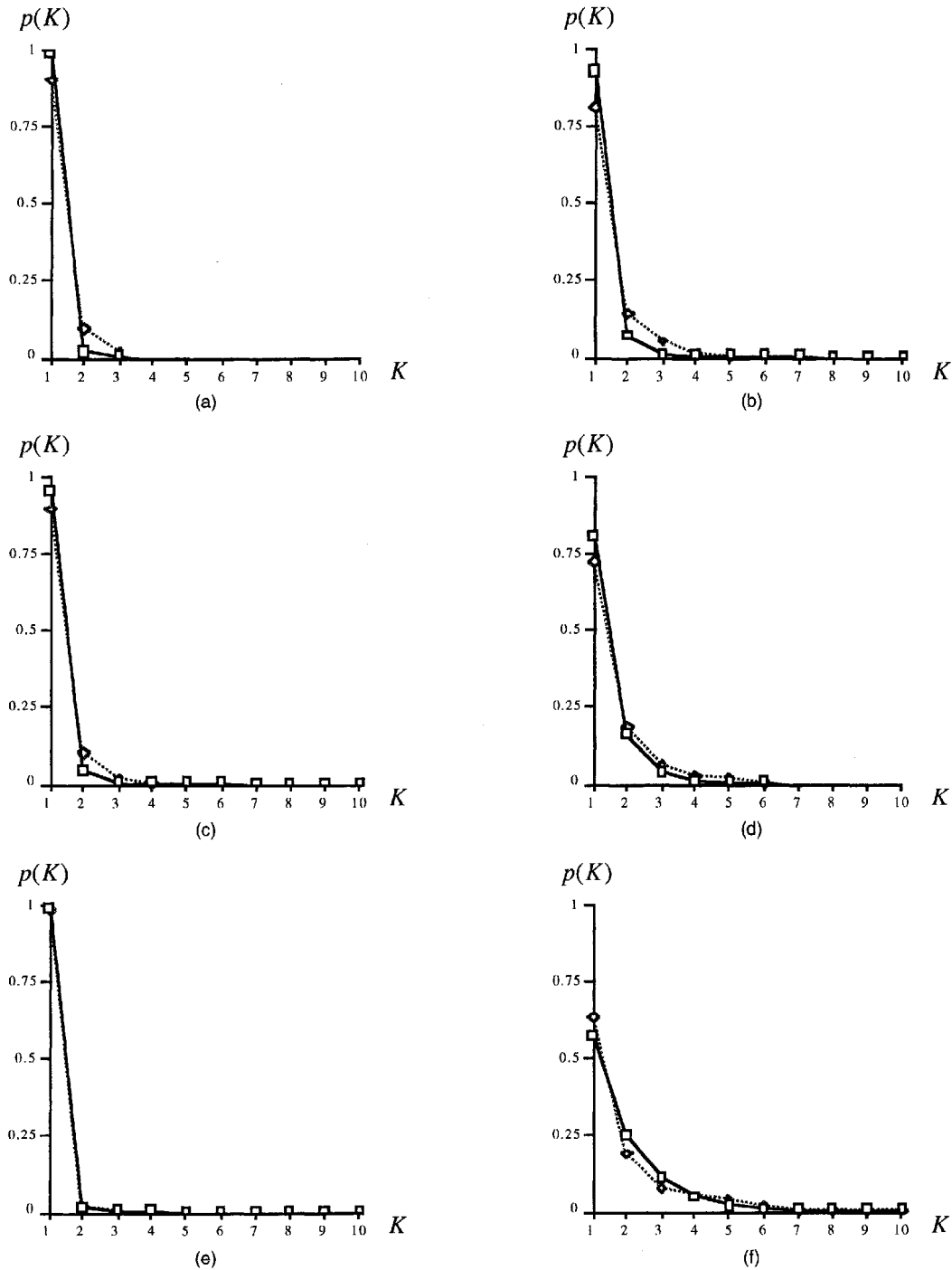


Fig. 8. Predicted (solid) and measured (dotted) background cluster length distributions. (a) $N_{FA}^{MTI} = 10$, $\sigma = 1$, $\tau = 1$, $\eta = 1$. (b) $N_{FA}^{MTI} = 10$, $\sigma = 2$, $\tau = 1$, $\eta = 1$. (c) $N_{FA}^{MTI} = 10$, $\sigma = 1$, $\tau = 2$, $\eta = 1$. (d) $N_{FA}^{MTI} = 10$, $\sigma = 2$, $\tau = 2$, $\eta = 1$. (e) $N_{FA}^{MTI} = 10$, $\sigma = 2$, $\tau = 2$, $\eta = 2$. (f) $N_{FA}^{MTI} = 20$, $\sigma = 2$, $\tau = 2$, $\eta = 1$.

background clutter ($N_{FA}^{MTI} = 10$ per frame). The measured distribution of cluster lengths are plotted against those predicted by clutter model in (b). The approximate location of the cluster length threshold for a group CFAR of 0.001 is indicated. All detections in clusters of length 3 or greater are shown in (c). The

group detection rate P_D^{group} was 1.0 with $N_{FA}^{\text{group}} = 14$ false detections. (A false detection is defined to be any detection in a cluster not associated with the moving object.) If we increase the MTI false alarm rate to $N_{FA}^{MTI} = 20$ per frame, the CFAR threshold increases to 4 (since the longer clusters will tend to form as shown

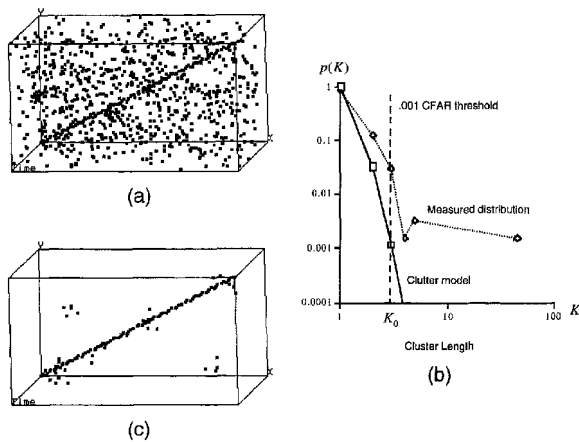


Fig. 9. Detecting single moving object embedded in clutter. (a) One object moving at constant velocity ($N_{FA}^{MTI} = 10/\text{frame}$). (b) Measured distribution of cluster lengths plotted against those predicted by clutter model. Cluster lengths greater than 3 are considered to represent groups of moving objects. (d) Detections in clusters of length > 3 ($\sigma = 2, \tau = 1, \eta = 1$).

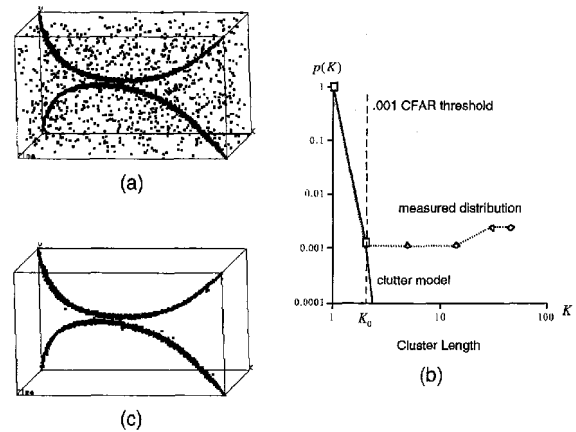


Fig. 10. Detecting groups of moving objects embedded in clutter. (a) Two groups of objects moving at constant velocity ($N_{FA}^{MTI} = 10/\text{frame}$). (b) Measured distribution of cluster lengths plotted against those predicted by clutter model. Cluster lengths greater than 2 are considered to represent groups of moving objects. (c) Detections in clusters of length > 2 ($\sigma = 2, \tau = 1, \eta = 1$).

in Fig. 8(f), and P_D^{group} drops to 0.86 with 52 false detections.

Fig. 10 is an example showing the detection of two groups of moving objects (same as in Figs. 1 and 2). At a group CFAR of 0.001, the probability of detecting the two groups of moving objects was 0.99 with 0 false detections using the same scoring criterion as above. In the process of detecting persistent groups, most of the uncorrelated radar clutter is removed. Fig. 11 shows the same 3 frames in Fig. 1 after MTI detections not associated with persistent objects have been eliminated.

It is noted that long clusters may correspond to stationary as well as moving objects. Stationary objects which appear in MTI to be moving like rotating antennas [3], trees and other reflective objects in the wind, flowing streams, etc. can be removed by eliminating clusters that do not move significantly in space.

IV. FORMATION RECOGNITION

A formation is an observable arrangement of vehicles in a group that can suggest or confirm something about the identity and near-term purpose of the group entity. Ground force units sometimes move together in formations in order to travel efficiently

and securely in groups, or to gain tactical advantage by coordinated maneuvering. The current phase of research is focused on the recognition of the following battalion-sized groupings.

Column: Vehicles driving in line ahead, with each vehicle attempting to follow the path of the one in front of it, often conforming to the lead vehicle's tread marks. Characteristic shape is a line which may be straight, curved, or kinked.

Wedge: Principal attribute is a (deformable) V-shaped forward line of vehicles. Point of the V faces the direction of travel, similar to formations of migrating birds. The point vehicle sets the direction of travel. One or more columnar tail groups may appear within or behind the open, trailing side of the V.

Line Abreast: Principal attribute is a linear trace composed of the majority of the battalion's combat vehicles in line abreast. The line forms the forward edge of the group. Its shape may be a shallow curve, or a combination of curved, wavy or fairly straight segments. One or more columnar tail groups may trail behind.

Most approaches to formation pattern recognition are based on templates. Schwartz [13] describes the use of templates for recognizing formations of moving

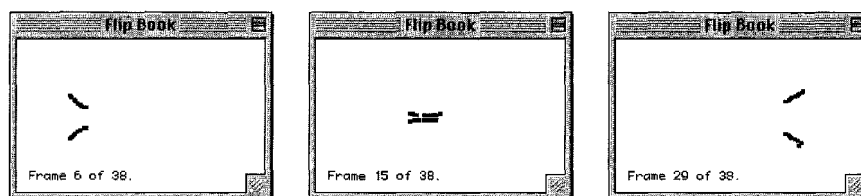


Fig. 11. Decluttered MTI frames.

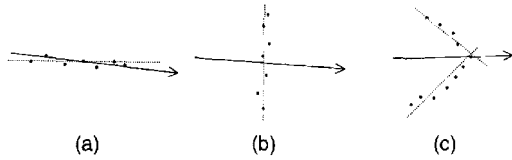


Fig. 12. Notional decomposition of formations into linear arrangements.

objects in MTI on a frame-by-frame basis. Our approach examines the shape of previously clustered MTI detections (specifically the orientation of the linear arrangements within the group) with respect to their direction of motion (Fig. 12). For example, the column is a linear arrangement oriented in the direction of motion (a). A line abreast is also a linear arrangement but moves in a perpendicular direction (b). A wedge can be decomposed into two linear arrangements oriented roughly 45° with respect to the direction of motion (c).

The following subsections describe how linear arrangements within a cluster are detected using the Hough transform, how the direction of motion is estimated from the cluster centroids over time, and how the resultant representation is used for detecting formations which may be moving in any direction. We conclude with a model to predict the false alarm rate of a formation detector as a function of number of objects in a formation and their linearity.

A. Hough Transform

The Hough transform is a method for finding geometric objects such as lines, parametric curves, and shapes in images [7]. In related applications it has been used to detect features such as road intersections in SAR [8], detecting low radar cross-section or weak infrared emitting moving targets using multiple sensors [4], and for track detection and initiation [19].

Consider the problem of detecting linear arrangements within a set of points. The equation, $y = mx + b$ represents a line by its slope m and y-intercept b . In the Hough transform, each point votes for all lines (slope-intercept combinations) that can pass through it. Assume that a single linear arrangement of points is present. Since points that lie along a line vote for the same slope-intercept combination, as votes are added those for the slope and intercept values of the line passing through the points exceed those of other lines. In the Hough space of slope and intercept combinations a peak forms at the location corresponding to the equation of the line in Cartesian space. Lines in Cartesian space thus map to points in Hough space, and vice versa. To avoid problems with vertical lines (where the slope is infinite), the polar form of the Hough transform is often used

$$\rho = x \cos \theta + y \sin \theta \quad (11)$$

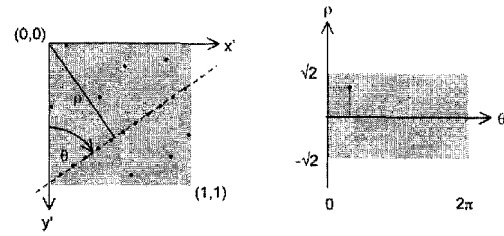


Fig. 13. Normalized Cartesian space (left) and Hough transform space (right).

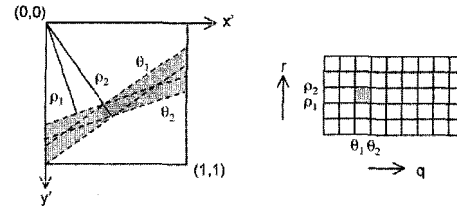


Fig. 14. Region corresponding in Cartesian space (left) to given bin in Hough accumulator array (right).

where θ is the angle between the y-axis and the line passing through the point and ρ is the distance of the line from the origin (Fig. 13).

Let us restrict our attention for the moment to the points within a cluster at a given point in time, $\{x_i, y_i\}$. For an origin we pick the upper-left point $\{x_{\min}, y_{\min}\}$. Without loss of generality the points within a cluster can be translated and scaled

$$x'_i = a(x_i - x_{\min}), \quad y'_i = a(y_i - y_{\min}) \quad (12)$$

so that they fit in the unit square, $0 \leq x'_i, y'_i < 1$. Points within this region in Cartesian space map to sinusoidal curves within the strip $-\sqrt{2} \leq \rho \leq \sqrt{2}$ in Hough space (Fig. 13). Each point in the cluster votes for Q parameter combinations $\{\theta_q, \rho_q\}$ where $\theta_q = 2\pi q/Q$ for $q = 0, 1, \dots, Q-1$ and

$$\rho_q = x'_i \cos \theta_q + y'_i \sin \theta_q. \quad (13)$$

Votes in Hough space are tallied in a $Q \times R$ accumulator array $H(q, r)$ where q and r index the parameters θ and ρ . The ρ_q are assigned the nearest integer address in the accumulator

$$r \approx \frac{R}{2\sqrt{2}}(\rho_q + \sqrt{2}). \quad (14)$$

Because of the discrete nature of the accumulator array, a bin in the accumulator array corresponds not to a line but to a bow-tie shaped region in Cartesian space (Fig. 14).

Fig. 15 shows four frames of simulated MTI data containing a group of 10 vehicles moving left to right (simulated probability of detection $P_D^{\text{MTI}} = 0.9$ and $N_{\text{FA}}^{\text{MTI}} = 10$ false alarms per frame). In each frame 9 of the ten vehicles are detected with ten false alarms. In the first frame the group is randomly distributed. Its Hough transform shows no obvious peak structure.

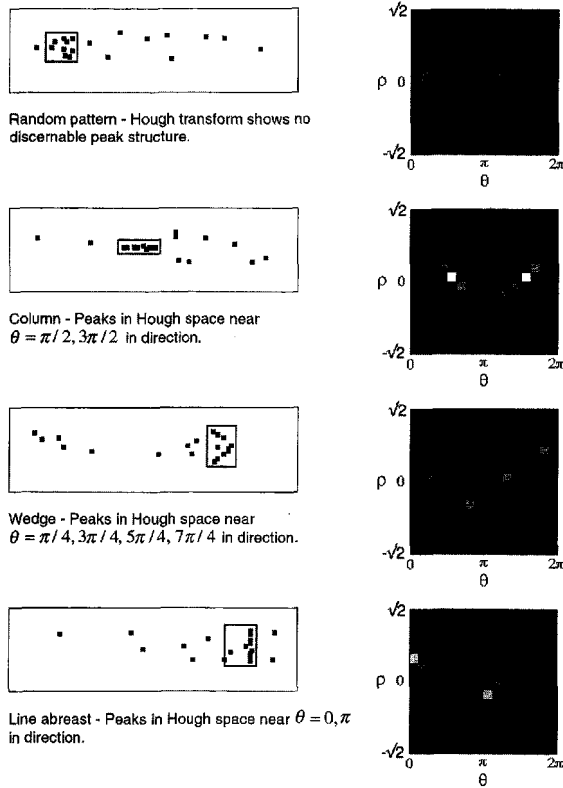


Fig. 15. Patterns (left) and corresponding Hough transforms (right).

In the second frame the group has assumed a column-like formation. Its Hough transform has peaks near $\{\pi/2, 0\}$ and $\{3\pi/2, 0\}$. In the third frame the group has changed into a wedge with peaks in Hough space near $\{\pi/4, 0\}$, $\{3\pi/4, -\sqrt{2}\}$, $\{5\pi/4, 0\}$, and $\{7\pi/4, \sqrt{2}\}$. In the last frame the group has changed into a line abreast with peaks in Hough space near $\{0, 0\}$ and $\{\pi, 0\}$.

B. Estimating the Direction of Motion

We have assumed that the above formations are moving in the same direction (left to right). In order

to recognize formations moving in any direction, we must estimate the direction of motion of the group ϕ and circularly shift Hough angles by $-\phi$. The direction of motion at time t_k is estimated by fitting a line to the cluster centroids at times t_{k-L} to t_{k+L} where $2L + 1$ is the size of the window. The optimal (least-squares) estimate of the direction of motion is

$$\phi(k) = \arctan\left(\frac{S_{XY}(k) - M_X(k)M_Y(k)}{S_{XX}(k) - M_X^2(k)}\right) \quad (15)$$

where

$$M_X(k) = \frac{1}{2L+1} \sum_{l=k-L}^{k+L} X'(l)$$

$$M_Y(k) = \frac{1}{2L+1} \sum_{l=k-L}^{k+L} Y'(l) \quad (16)$$

$$S_{XX}(k) = \frac{1}{2L+1} \sum_{l=k-L}^{k+L} X'(l)X'(l)$$

$$S_{XY}(k) = \frac{1}{2L+1} \sum_{l=k-L}^{k+L} X'(l)Y'(l)$$

are the first- and second-order moments and

$$X'(k) = X(k) - X(k-L), \quad Y'(k) = Y(k) - Y(k-L). \quad (17)$$

Fig. 16 shows a series of Hough transforms for a column traveling around in a circle. In each case, the Hough space has been shifted so that the peaks occur near $\theta = \{\pi/2, 3\pi/2\}$. It is noted that the position of the peaks vary in ρ . When the column points toward or away from the origin (a) and (d) the peaks occur near $\rho = 0$; otherwise, they lie off axis.

C. Formation Detection Filters

Instead of attempting to define a complete set of formation classes (which is at present incomplete) and constructing a classifier, we have instead developed a set of filters for detecting specific formations.

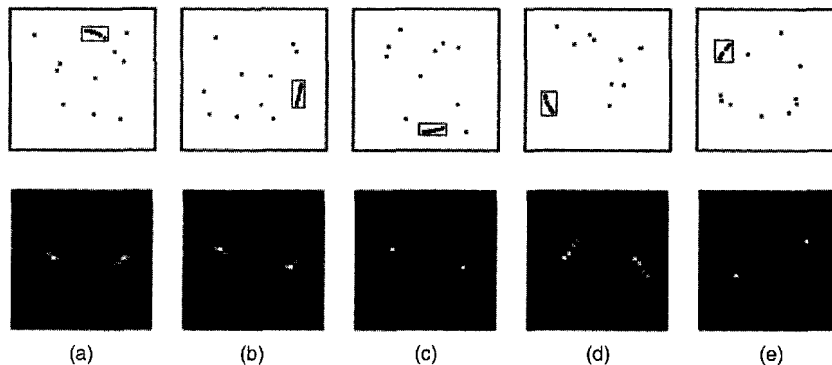


Fig. 16. Direction of motion-compensated Hough transforms of column moving in circle.

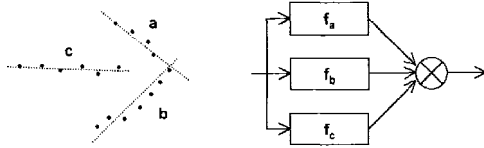


Fig. 17. Linear arrangements within a formation (left) and corresponding filters (right).

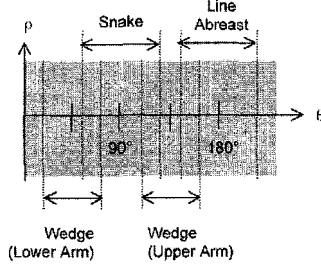


Fig. 18. Regions in Hough space.

Formations are modeled by one or more linear arrangements of detections. For example, a wedge with a single columnar tail group (Fig. 17) would be modeled by three linear arrangements. Each is represented by a separate filter.

The general form of a filter is

$$f(q_1, q_2, Q, R, T) = \text{sgn} \left[\left(\max_{0 \leq q_1 \leq q < q_2 < Q, 0 \leq r < R} H(q, r) \right) - T \right] \quad (18)$$

where $\text{sgn}(z) = 1$ for $z \geq 0$ and zero otherwise. A filter “fires” if the number of detections in any bin within a specified region in the Hough accumulator exceeds a given threshold T . The angular extent of the region is defined by two parameters q_1 and q_2 where the corresponding Hough angles are $\theta_1 = 2\pi(q_1/Q)$ and $\theta_2 = 2\pi(q_2/Q)$ and Q is the number of angle bins in the accumulator. Fig. 18 shows the regions in Hough space for a column filter ($\theta_1 = 54^\circ, \theta_2 = 126^\circ$), line abreast ($\theta_1 = -36^\circ, \theta_2 = 36^\circ$), and wedge ($\theta_1 = 18^\circ, \theta_2 = 72^\circ$ and $\theta_1 = 108^\circ, \theta_2 = 162^\circ$) where the angles are user-defined and measured relative to north (up).

To detect a formation consisting of a single linear arrangement such as a column or a line abreast, a single filter is used. Multiple filters and together are used to detect formations containing two or more linear arrangements. For example, to detect a wedge, two filters are needed, one for each arm of the wedge. The wedge is detected if both filters fire.

D. Performance Modeling

A formation detection filter is defined by the following parameters:

1) The parameters q_1 and q_2 specify the orientation of the linear arrangement with respect to the direction of motion.

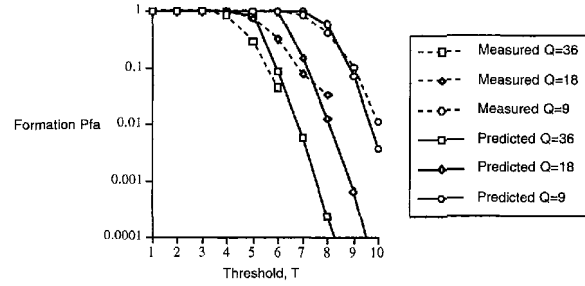


Fig. 19. Predicted and measured formation detection false alarm rates versus threshold.

2) Q is the parameter which is related to the linearity of the arrangement.

3) T is the parameter which specifies the minimum number of detections that must occur within the formation for the filter to fire.

The parameters q_1 , q_2 , and Q are domain dependent and are specified by the user. The q_1 , q_2 values plotted in Fig. 18 have been used in all experiments. The effect of varying Q is illustrated in the next section. The threshold $T \leq n$, where n is the number of objects in the formation, controls the detection and false alarm rate of the filter.

In order to determine the threshold needed to achieve a given formation recognition false alarm rate $P_{FA}^{\text{formation}}$ we must first develop a model to predict the false alarm rate of the filter as a function of q_1 , q_2 , Q , T , and n . Two detections within a group define a potential linear arrangement. For n detections there are

$$m = \binom{n}{2} \quad (19)$$

possible combinations of detections taken two at a time. For each combination the probability of k or more detections falling within the corresponding bow-tie region is

$$p(k) = \sum_{k'=k}^n \binom{n}{k'} p^{k'} (1-p)^{n-k'} \quad (20)$$

where $p = 1/s$. The size of the region is

$$s = \frac{2}{\Delta\theta} + \frac{2}{\Delta\rho} \sqrt{\frac{1}{2\pi}} \quad (21)$$

where $\Delta\theta = 1/Q$ and $\Delta\rho = 1/R$. If these regions did not overlap the probability of k or more detections falling within any bow-tie region would be $mp(k)$. Because they do overlap, the sum of the m individual probabilities can exceed unity. We therefore make the approximation

$$P_{FA}^{\text{formation}} = \min\{mp(k), 1\}. \quad (22)$$

Fig. 19 plots predicted and measured false alarm rates for any linear arrangement containing $n = 10$ detections with $q_1 = r_1 = 0$, $q_2 = r_2 = 1$, and $Q =$

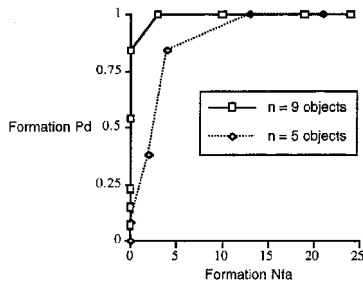


Fig. 20. ROCs for column detector as function of number of detections in column.

$R = \{9, 18, 36\}$ bins. As the number of Hough bins decreases the filters become less selective and the threshold must be increased to maintain a given CFAR. Since the above model predicts the false alarm rate for any linear arrangement of k or more objects it represents an upper bound on the false alarm rate for a particular formation detector.

V. MTI CLUSTERING AND FORMATION DETECTION RESULTS

In order to evaluate MTI clustering and formation recognition we use the simulated data shown in Fig. 15 consisting of a single group moving at a constant speed west to east (left to right) transitioning from a random arrangement into a column, then into a wedge, and finally into a line abreast. The data set consists of 80 frames total. Each frame contains a specified number of MTI detections (depending on the experiment) plus 10 false alarms. The group assumes column and wedge formations for 13 frames each, and the line abreast formation for 11 frames.

The process flow involves clustering the data and detecting persistent groups of objects as described in Section III. Each 3-D cluster is then processed as discussed in the previous section. The performance of a formation detector is defined in terms of its receiver operating characteristic (ROC) which gives the probability of detecting the correct formation versus the number of incorrect detections (false alarms) as a function of the Hough threshold.

First we consider the effect of the number of objects in a formation on detection performance, specifically on the performance of the column detector. Two simulated MTI sequences were clustered with the same parameters ($\sigma = \tau = 2, \eta = 5$) and processed by the column detector ($Q = R = 36$): one with 9 objects in a formation (i.e., per cluster) and 10 MTI false alarms per frame, the other with only 5 objects per formation and 10 MTI false alarms per frame. Fig. 20 shows the reduction in performance as the number of objects within the formation decreases. For a given detection probability, the number of false alarms increases as the number of objects in the formation decreases. Using the false alarm model

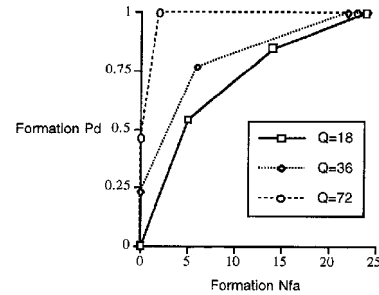


Fig. 21. ROCs for wedge detector as function of number of bins in Hough accumulator.

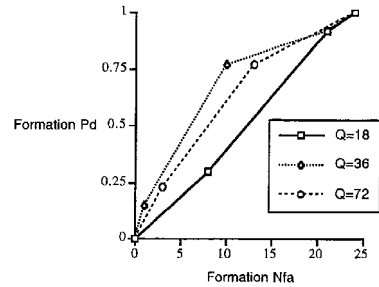


Fig. 22. ROCs for wedge detector applied to groups that are not perfectly linear in arrangement.

described in the previous section, a threshold of $T = 6$ is needed to achieve a CFAR of 0.01. Using this threshold for the case of 9 objects per formation and 10 MTI false alarms per frame, the column detection probability is 1.0 with 4 false alarms over the 80 frames.

Next we explore the relationship between the linearity of the arrangements and the number of bins in the Hough accumulator on performance, in this case, that of the wedge detector. Formations consisting of perfectly linear arrangements map into points in Hough space. Performance should thus improve as the number of Hough bins increases. The simulation was run with 9 objects in the group and 10 MTI false alarms per frame. Clustering was performed using the same parameters as above. Fig. 21 plots ROC curves for the wedge detector using $Q = R = \{18, 36, 72\}$. As the number of Hough bins increases the wedge filter becomes more selective causing the false alarm rate to decrease.

In a second sequence we simulate relative motion between objects in a group by shifting the position of detections randomly by a small amount frame to frame. The resultant groups are not perfectly linear in arrangement and spread in Hough space. (Recall the Hough transform is referenced to the upper left point in the cluster. The width of the peak in Hough space depends on the variance of the slopes of the lines between the reference and other points in the cluster.) Fig. 22 plots the ROC curves for the wedge detector applied to this second sequence using $Q = R = \{18, 36, 72\}$. Performance increases from 18 to 36

TABLE I
Summary of MTI Clustering and Formation Recognition Parameters

| Parameter | Definition | Effect |
|-----------------------|---|--|
| σ | Spatial smoothing | Controls the growth of clusters by merging detections less than 2σ apart. Longer clusters tend to form from clutter as σ increases. |
| τ | Temporal smoothing | Uses detections at other times to influence the growth of clusters (useful for clustering sparse groups). Longer clusters tend to form from clutter as τ increases. |
| η | Split/merge threshold | Reduces fragmentation of clusters and eliminates those containing less than η detections. Longer clusters tend to form from clutter as η decreases. |
| K_0 | Group detection threshold | Minimum cluster length for persistent objects. As the threshold decreases the group detection false alarm rate increases. |
| Q, R ($Q = R$) | Number of Hough bins | Controls the linearity of detections within the arrangement/formation. As the number of bins decreases the formation false alarm rate increases. |
| q_1, q_2 | Angular extent of region in Hough space | Defines the range of orientations of the arrangement/formation relative to the direction of motion. |
| T | Formation detection threshold | Minimum number of MTI detections within the arrangement/formation. As the threshold decreases the formation detection false alarm rate increases. |

bins but then decreases from 36 to 72 bins as the peak spreads into multiple Hough bins. One would expect a similar spreading in Hough space as columns move along highly curved paths and turn at intersections.

For well-aligned groups increasing the number of Hough bins will increase the detection rate while reducing the false alarm rate. However for military units operating in the field one can expect deviations in their alignment and so increasing the number of Hough bins beyond a certain point will tend to reduce the detection rate as well.

We noted earlier in Section II that analysts often accumulate multiple frames of MTI data in order to better visualize shape and motion. Accumulation is useful when groups are sparse, or due to reduced radar P_D or terrain obscuration. A disadvantage of frame accumulation is that it blurs groups along the direction of motion. While this can be useful for visualizing columns, it can make it more difficult to discriminate between wide formations. By using the group velocity estimate of a cluster, detections within clusters can be shifted and accumulated locally within the cluster. Fig. 23 show a slight improvement in the performance of the wedge detector with and without cluster-based accumulation.

VI. SUMMARY

Two new MTI exploitation techniques were described and demonstrated on simulated data. The first is a 3-D space-time MTI clustering technique that detects groups of objects (stationary or moving) that persist over time. The second technique is based on the use of a Hough transform based technique for recognizing moving formations such as columns, wedges, and lines abreast by analyzing the shape of previously clustered MTI detections (specifically the

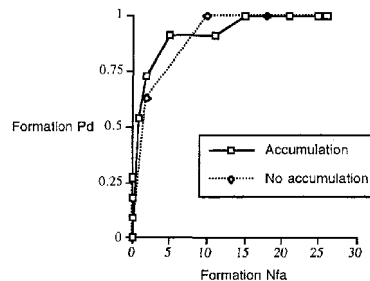


Fig. 23. Results for wedge detector with and without cluster-based accumulation of detections across frames.

orientation of linear arrangements within the group) with respect to their direction of motion.

Performance models for clustering and formation recognition were developed. Table I describes the effect of algorithm parameters on performance. We have shown that as σ , τ , n , or λ increases, or as η decreases longer clusters form from background clutter. As the spacing of objects within a group decreases σ and τ must increase to maintain a given detection rate. Similarly as the number of objects in the group decreases η must decrease. Group detection performance ($P_D^{\text{group}}/N_{\text{FA}}^{\text{group}}$) will thus decrease as the spacing between objects increases, as the number of objects within the group decreases, and as the background clutter intensity $\lambda = n/S$ increases.

Experimental results for the recognition of battalion-size formations in simulated MTI data sets were presented. Formation detection performance depends on the number of objects in the formation and the linearity of their arrangement(s). As the linearity decreases, the number of Hough bins must decrease to maintain a given detection rate. As the number of bins and/or the number of objects in the formation decreases we have shown that the false

alarm rate increases. Formation detection performance ($P_D^{\text{formation}}/N_{\text{FA}}^{\text{formation}}$) will thus decrease as the linearity decreases and/or as the number of objects in the formation decreases. We have also shown a slight improvement in performance by accumulating detections over time.

The computational complexity of our formation recognition algorithm is considerably less than template matching. Where a template matcher must be applied to the entire image, our Hough-based model is applied only over those regions in space-time that contain persistent groups of objects. Moreover, by estimating the motion of clusters and referencing our formation model to the direction of motion, we eliminate the need to search over multiple orientations.

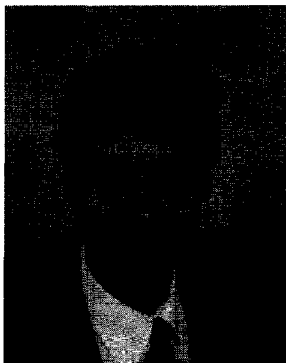
Areas for future work include the development of multiscale clustering techniques for capturing hierarchical relations between groups of moving objects, and the development of more flexible representations for formation recognition. The use of 3-D "bubbles" developed for segmenting volumetric medical data [16] may be useful for detecting groups that operate within other groups (e.g., battalions within a brigade). Such an approach is a natural extension of LoG filter. Decomposing formations into arrangements with a greater number of degrees of freedom (e.g., higher order curves) would allow for a greater amount of "flex" is one possible area for improvement. An alternative is to match formation templates such as those described by Schwartz [13] to MTI data that have been grouped into 3-D clusters where the templates are aligned to the direction of motion of the cluster.

ACKNOWLEDGMENTS

Thanks to Tom Tulenko at Veridian-PSR for providing military domain expertise, and Bruce Johnson for his support under the DARPA Moving Target Exploitation (MTE) program.

REFERENCES

- [1] Ballard, D., and Brown, C. (1982) *Computer Vision*. Englewood Cliffs, NJ: Prentice Hall, 1982.
- [2] Blackman, S. S. (1986) *Multiple-Target Tracking with Radar Applications*. Norwood, MA: Artech House, 1986.
- [3] Brown, D. B. (1996) Classification of stationary movers with simplified MTI tracking. *Proceedings of SPIE*, **2759** (1996), 596–604.
- [4] Cheng, H., Sun, Z., Liu, Q., and Chen, Y. (1997) Small-target tracking technique with data fusion of distributed sensor net. *Proceedings of SPIE*, **3163** (1997), 566–574.
- [5] Duda, R., and Hart, P. (1973) *Pattern Classification and Scene Analysis*. New York: Wiley Interscience, 1973.
- [6] Garren, D. A., Grayson, T. P., Johnson, R. O., and Strat, T. M. (1999) Theoretical analysis of a continuous tracking system. *Proceedings of SPIE*, **3709** (1999).
- [7] Hough, P. V. C. (1962) Method and means for recognizing complex patterns. U.S. Patent 3,069,654, 1962.
- [8] Iisaka, J., and Sakurai-Amano, T. (1996) Automated GCP detection for SAR imagery: Road intersections. *Proceedings of SPIE*, **2818** (1996).
- [9] Johnson, B. L., and Grayson, T. P. (1998) Moving target exploitation. *Proceedings of SPIE*, **3393** (1998), 172–183.
- [10] Lee, J. L., and Rodriguez, J. J. (1994) Volumetric segmentation of magnetic resonance images. *Proceedings of SPIE*, **2298** (1994), 652–661.
- [11] Marr, D., and Hildreth, E. (1980) Theory of edge detection. *Proceedings of the Royal Society of London B*, **207** (1980), 187–217.
- [12] Papoulis, A. (1965) *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1965.
- [13] Schwartz, S. A. (2000) Algorithm for automatic recognition of formations of moving targets. *Proceedings of SPIE*, **4051** (2000), 407–417.
- [14] Shrader, W. W. (1977) Moving target indicator radar. *Proceedings of SPIE*, **128** (1977), 86–97.
- [15] Stroud, P. D., and Gordon, R. C. (1997) Automated military unit identification in battlefield simulation. *Proceedings of SPIE*, **3069** (1997), 375–386.
- [16] Tek, H., and Kimia, B. B. (1994) Volumetric segmentation of medical images by three-dimensional bubbles. *Computer Vision and Image Understanding* (Feb. 1994).
- [17] Tou, J. T., and Ganzalez, R. C. (1974) *Pattern Recognition Principles*. Reading, MA: Addison-Wesley, 1974.
- [18] Wishner, R. P., and Fennell, M. T. (1998) Battlefield awareness via synergistic SAR and MTI exploitation. *IEEE Aerospace and Electronics Systems Magazine*, **13**, 2 (1998), 39–43.
- [19] Yankowich, S. W., and Farooq, M. (1998) Hough transform based multisensor, multitarget, track initiation technique. *Optical Engineering*, **37**, 7 (1998), 2064–2077.



Mark J. Carlotto (S'78—M'78—SM'90) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Carnegie-Mellon University, Pittsburgh, PA, in 1977, 1979, and 1981, respectively.

From 1981–1993 he was a division staff analyst at TASC in Reading, MA where he was involved in a variety of projects in the areas of multispectral image processing, object/change detection, image segmentation, scene analysis, geographical information systems, knowledge-based systems, text processing, and data visualization. Also during this period from 1981–1983 he was an Assistant Adjunct Professor in the College of Engineering at Boston University where he taught courses in computer architecture and pattern recognition. Dr. Carlotto is now a senior staff scientist with Veridian-PSR in Arlington, VA. His current research interests include moving target exploitation, terrain feature extraction, change detection and analysis, and image understanding.