

Using Centroid Covariance in Target Recognition

Gang Liu and Robert M. Haralick
Department of Electrical Engineering
University of Washington
Seattle, WA 98195-2500, U.S.A.
{gliu,haralick}@isl.ee.washington.edu

Abstract

An ATR algorithm for low quality imagery is reported. Compact shaped targets are represented by their 2D silhouettes. Associated with each point on the silhouette, there is a direction roughly perpendicular to the local segment of the silhouette. The location of each silhouette point is assumed to be perturbed along that direction. A statistical technique is used to estimate the variance of that perturbation for the silhouette points of the hypothesized target. This variance is then used to estimate the location covariance of the target centroid. Target detection and recognition is based on this covariance. Target scaling, aspect, and rotation are not considered. Experiments on 31 FLIR images give a correct recognition of target identity and target location for 29 of the 31 images.

1. Introduction

Automatic target recognition (ATR) applications are characterized by their great variability in target shape and appearance, scale and orientation, lighting and imaging condition, natural background clutter, camouflage, etc. This has resulted in the fact that generic target recognition problem is far beyond the capability of the state-of-the-art ATR systems. Currently, the research in this area is mainly focused on algorithms with various restrictions on the application.

In this work, we study the situation where interested targets, which are 3D objects in general, always appear at the same distance (hence with the same scale) and with the same aspect. This way, although the target may look different in different images due to the varying imaging conditions, its silhouette, or its boundary in the 2D image, remains the same and there is no rotation involved. The information contained in the geometric shape of the target boundary is all we have to discriminate one target from another. [1][4]

Due to the great variability in the image forming process and the usual low quality of the images, image segmentation and edge detection based algorithms give very low performance. The low level information has to be gathered with the help of higher level knowledge. We adopt a hypothesis and test paradigm to use target shapes in making measurements at each location in the image for each interested target. Target detection and recognition is carried out based on those measurements.

The measurement made here for each target hypothesis is the location covariance of its centroid. This is obtained by taking the target boundary as a polygon, estimating the location covariance of the vertices, and analytically propagating the covariance of the polygon's vertices to that of its centroid. Correct hypotheses are associated with small measurement values, and large location covariance values are usually observed for incorrect hypotheses.

Experiments on 31 FLIR (forward looking infrared) images give a correct recognition of target identity and target location for 29 of the 31 images.

2. Framework of the algorithm

There are a finite number of interested targets, each being represented by a digitized 2D contour. Targets may appear at any location in the image except the boundary regions. The task is to decide how many, if any, targets are in the image, their identities, and the locations of their centroids.

In the kernel of the algorithm, a hypothesis is made that some target t is present in the image with its centroid located at some location (r, c) . The contour of target t is imagined in the image with its centroid at (r, c) . The pixel locations on the contour form the set of vertices of a polygon.

The pixel values of the input image is assumed to be corrupted by some additive *iid* normal perturbation. This perturbation on the pixel values causes the uncertainty in the vertices' locations if they were to be estimated. This uncertainty is modeled as a perturbation. We are mainly interested in the part of the perturbation that is perpendicular

to the local segment of the target contour. Next section describes how the variance of this perturbation is estimated, and then analytically propagated to the location of the polygon's centroid, resulting in an estimate of the covariance matrix of the centroid location. The trace (sum of diagonal elements) of this matrix is the major measurement, from which the inference about the presence of target t at location (r, c) will be made. Bigger values of the trace show greater variability in the centroid location, indicating less likelihood of the target's presence at that location.

Multiple hypotheses are made by the algorithm for all targets at all possible locations, and the associated measurements are made. These hypothesis are compared by these measurements, and the choice is made in favor of smaller measurement values.

Right now we make use of the domain knowledge that each image in our test image set contains one and only one target. With this knowledge, the decision rule for target detection and recognition is simply to claim the presence of the target at the location whose corresponding hypothesis has the smallest trace value. We are still working on properly setting up thresholds for the trace values for different targets. This will allow multiple target presence in the image even with overlap between them, and will avoid claiming at least one target per image when there is none.

The geometric shape of the polygon plays an important part in deciding the contribution from each boundary point to the trace. The importance of the boundary points depends highly on each other, and in general differs from each other. Our study shows that using all boundary points to model the targets is often less as effective as using only a subset which is optimal in certain sense. In choosing the subset, we want *a)* the subset retains enough discriminatory power to distinguish the target from others; and *b)* the trace of the centroid location covariance matrix is the smallest among all subsets of the same size. The details of the criteria and the algorithm for choosing the subset is omitted here.

3. Centroid covariance matrix estimation

3.1. Boundary point location variance

Figure 1 illustrates the hypothesis that target t is present at (r, c) . Target t happens to be a side view of a car. The black dots show its boundary points, and the small cross shows the centroid. The dotted lines parallel to the row and column axes show the pixel grid of the image.

Let P be one of the boundary points, and d be the direction assigned for it, which is roughly perpendicular to the local segment of the hypothesized target boundary. Through P and along the direction d we have a straight line l which is digitized by the pixel grid of the image. The pixel locations through which l passes form the *support* of l . This

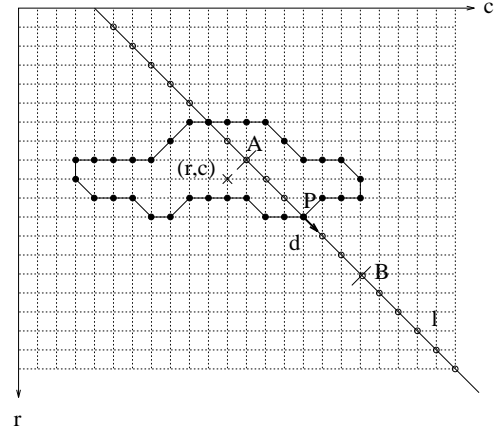


Figure 1. Illustration of the estimation of the location variance of hypothesized target boundary points.

support is denoted by $\Gamma = (\gamma_1, \gamma_2, \dots, \gamma_M)$ where M is the total number of points on l . Γ is shown in the figure by the unfilled circles lying on the line l . Let P be the θ -th point on l . For integer valued $\tau > 0$ ($\tau = 3$ in the figure) we define the τ -segment for P to be the segment of Γ centered on P and with length $2\tau + 1$. Let it be denoted by

$$\Pi_{P,\tau} = (\gamma_{\theta-\tau}, \gamma_{\theta-\tau+1}, \dots, \gamma_{\theta+\tau})$$

In the figure, it is represented by the unfilled circles on l between A and B inclusively. Let $S = (s_1, s_2, \dots, s_M)$ and $H = (h_1, h_2, \dots, h_N)$ where $N = 2\tau + 1$ be the profiles of the unperturbed image pixel values on Γ and $\Pi_{P,\tau}$, respectively. H is just one segment of S and $h_n = s_{(\theta-\tau-1)+n}$ for $n = 1, 2, \dots, N$.

Since Γ and $\Pi_{P,\tau}$ are essentially along a straight line, by translating and rotating the axes of the coordinate system, they can be converted into sets of 1D positions. Hence S and H are taken as 1D digital signals.

When the image pixel values are perturbed by additive *iid* normal noise, the observed pixel values on Γ form $\hat{S} = (\hat{s}_1, \dots, \hat{s}_M)$ where $\hat{s}_m = s_m + \nu_m$ for $m = 1, 2, \dots, M$. $\nu_1, \nu_2, \dots, \nu_M$ are *iid* $n(0, \sigma^2)$ random variables.

The location of the boundary point P in Γ can now be estimated as $\hat{\theta}$ satisfying

$$\sum_{n=1}^N (\hat{s}_{(\hat{\theta}-\tau-1)+n} - h_n)^2 = \min_{\theta} \sum_{n=1}^N (\hat{s}_{(\theta-\tau-1)+n} - h_n)^2$$

The noise perturbation in \hat{S} propagates to $\hat{\theta}$. Then the perturbation $\Delta\theta = \hat{\theta} - \theta$ causes uncertainty on the location of the target boundary point P .

$$\mathbf{E}(\Delta\theta) = 0$$

$$\begin{aligned} \text{Var}(\Delta\theta) &= \left\{ \sum_{n=1}^{N-1} \left(h_{n+1} - h_n - \frac{1}{6} [h''_{n+1} + 2h''_n] \right)^2 \right. \\ &\quad \left. + \left(h_N - h_{N-1} + \frac{1}{6} [2h''_N + h''_{N-1}] \right)^2 \right\}^{-1} \cdot \sigma^2 \end{aligned}$$

where h''_m is the second order derivative value at γ_m of the cubic spline constructed from H .

The last equation gives the location variance of the boundary point P along the direction d . In our application, we choose $\tau = 3$, and since we do not have the unperturbed H , we substituted \hat{H} for H .

3.2. Polygon vertices and centroid

A simple polygon is represented by an ordered (either clockwise or counter-clockwise) set of N 2D locations P_1, \dots, P_N of its vertices.

Let P_i have coordinates (x_i, y_i) for $i = 1, 2, \dots, N$. Let $\mathbf{x} = (x_1, \dots, x_N)^T$ and $\mathbf{y} = (y_1, \dots, y_N)^T$. The centroid of the polygon is denoted by $Q(\mathbf{x}, \mathbf{y}) = (Q_x(\mathbf{x}, \mathbf{y}), Q_y(\mathbf{x}, \mathbf{y}))^T$, and can be expressed as

$$Q_x(\mathbf{x}, \mathbf{y}) = \frac{K_x}{3S} \quad Q_y(\mathbf{x}, \mathbf{y}) = \frac{K_y}{3S}$$

where

$$\begin{aligned} K_x &= \sum_{i=1}^N (x_i y_{i+1} - x_{i+1} y_i) (x_i + x_{i+1}) \\ K_y &= \sum_{i=1}^N (x_i y_{i+1} - x_{i+1} y_i) (y_i + y_{i+1}) \\ S &= \sum_{i=1}^N (x_i y_{i+1} - x_{i+1} y_i) \end{aligned}$$

where the notation of $x_{N+1} = x_1$ and $y_{N+1} = y_1$ is used.

3.3. Perturbation on the vertices

Suppose $P_i = (x_i, y_i)^T$ is affected by an additive perturbation $\Delta P_i = (\Delta x_i, \Delta y_i)^T$, which is called the input perturbation. We assume that this perturbation is along a certain known direction denoted by a constant unit vector $d_i = (d_{i,x}, d_{i,y})^T$. We model the perturbation along this direction by a normal random variable $v_i \sim N(0, \sigma_i^2)$, where the estimation of σ_i^2 is described in Section 3.1. We further assume that v_i and v_j are independent if $i \neq j$. Then $\mathbf{v} = (v_1, \dots, v_N)^T$ has the mean vector $\mathbf{0}$ and covariance matrix $\text{Diag}(\sigma_1^2, \dots, \sigma_N^2)$.

The perturbed location of the i -th vertex is given by

$$\begin{pmatrix} \hat{x}_i \\ \hat{y}_i \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \begin{pmatrix} v_i d_{i,x} \\ v_i d_{i,y} \end{pmatrix} \quad (1)$$

3.4. Perturbation on the centroid

Due to the perturbation on the vertices, the location of the centroid is perturbed. This perturbation is called the output perturbation and is expressed as

$$\begin{aligned} \Delta Q_x &= Q_x(\mathbf{x} + \Delta \mathbf{x}, \mathbf{y} + \Delta \mathbf{y}) - Q_x(\mathbf{x}, \mathbf{y}) \\ \Delta Q_y &= Q_y(\mathbf{x} + \Delta \mathbf{x}, \mathbf{y} + \Delta \mathbf{y}) - Q_y(\mathbf{x}, \mathbf{y}) \end{aligned}$$

Using linear approximation, we have

$$\begin{aligned} Q_x(\mathbf{x} + \Delta \mathbf{x}, \mathbf{y} + \Delta \mathbf{y}) &\approx Q_x(\mathbf{x}, \mathbf{y}) + A^T \cdot \mathbf{v} \\ Q_y(\mathbf{x} + \Delta \mathbf{x}, \mathbf{y} + \Delta \mathbf{y}) &\approx Q_y(\mathbf{x}, \mathbf{y}) + C^T \cdot \mathbf{v} \end{aligned}$$

where $A = \frac{\partial Q_x(\mathbf{x}, \mathbf{y})}{\partial \mathbf{v}}$ and $C = \frac{\partial Q_y(\mathbf{x}, \mathbf{y})}{\partial \mathbf{v}}$ are $N \times 1$ real vectors. For $i = 1, 2, \dots, N$, let $a_i = \frac{\partial Q_x}{\partial v_i}$ and $c_i = \frac{\partial Q_y}{\partial v_i}$.

$$a_i = \frac{1}{3S} \cdot \left\{ \frac{\partial K_x}{\partial v_i} - \frac{K_x}{S} \cdot \frac{\partial S}{\partial v_i} \right\} \quad (2)$$

$$c_i = \frac{1}{3S} \cdot \left\{ \frac{\partial K_y}{\partial v_i} - \frac{K_y}{S} \cdot \frac{\partial S}{\partial v_i} \right\} \quad (3)$$

where K_x , K_y , and S are given in Section 3.2, and

$$\begin{aligned} \frac{\partial S}{\partial v_i} &= (y_{i+1} d_{i,x} - x_{i+1} d_{i,y}) + (x_{i-1} d_{i,y} - y_{i-1} d_{i,x}) \\ \frac{\partial K_x}{\partial v_i} &= (x_i y_{i+1} - x_{i+1} y_i) d_{i,x} \\ &\quad + (x_i + x_{i+1})(y_{i+1} d_{i,x} - x_{i+1} d_{i,y}) \\ &\quad + (x_{i-1} y_i - x_i y_{i-1}) d_{i,x} \\ &\quad + (x_{i-1} + x_i)(x_{i-1} d_{i,y} - y_{i-1} d_{i,x}) \\ \frac{\partial K_y}{\partial v_i} &= (x_i y_{i+1} - x_{i+1} y_i) d_{i,y} \\ &\quad + (y_i + y_{i+1})(y_{i+1} d_{i,x} - x_{i+1} d_{i,y}) \\ &\quad + (x_{i-1} y_i - x_i y_{i-1}) d_{i,y} \\ &\quad + (y_{i-1} + y_i)(x_{i-1} d_{i,y} - y_{i-1} d_{i,x}) \end{aligned}$$

where the notation of $x_0 = x_N$, $y_0 = y_N$ is used. The mean and covariance of ΔQ are then given by

$$\mu_{\Delta Q} \approx \begin{pmatrix} A^T \cdot \mathbf{E}(\mathbf{v}) \\ C^T \cdot \mathbf{E}(\mathbf{v}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (4)$$

$$\Sigma_{\Delta Q} \approx \begin{pmatrix} \sum_{i=1}^N a_i^2 \sigma_i^2 & \sum_{i=1}^N a_i c_i \sigma_i^2 \\ \sum_{i=1}^N c_i a_i \sigma_i^2 & \sum_{i=1}^N c_i^2 \sigma_i^2 \end{pmatrix} \quad (5)$$

The above result was validated using a statistical test in a controlled experiment where a quadrilateral is used for the polygon, and noise perturbations with known parameters were generated and used to perturb the locations of the vertices. The centroid of the vertex-perturbed polygon is calculated, and its sample mean and covariance matrix are compared with the theoretically predicted values given by Equations (4) and (5). The empirical result agreed with the theoretical prediction, although the details are omitted here.

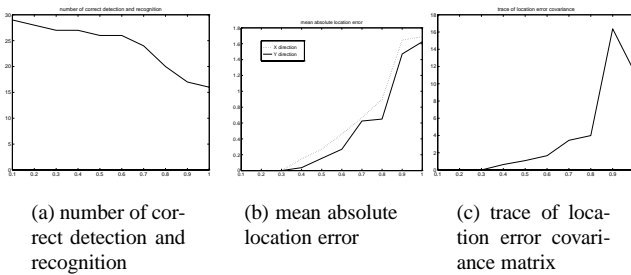


Figure 2. Performance of the algorithm using different target model sets in the experiment.

4. Experiment on FLIR imagery

Our test image set consists of 31 FLIR images, in each of which there is a distinct target. The images are 8-bit images of size 512×410 . For each of the targets, we have the locations of all the boundary points relative to the centroid of that target. The sizes of the target boundaries range from 270 to 728, and the area inside them range from 2609 to 15085, all in number of pixels. For each boundary point, we define a direction that is roughly perpendicular to the local segment of the boundary, as needed in estimating the covariance of the centroid's location.

For each target, we choose 9 different subsets of the boundary points. The sizes of the subsets are 10%, 20%, ..., 90% of that of the original boundary. These percentage values are called the percentage indices of the subsets. Together with the original boundary, with percentage index 100%, there are 10 models for each target. For each of the percentage indices, we construct a target model set by including the target models of that percentage index from all 31 targets. For example, the so-called "10% model set" includes 31 models, each of which is a 10% target model constructed from the respective original target boundary. Each of the 10 model sets is used in a separate experiment to detect and recognize the targets in the 31 images.

The performance of the algorithm, in terms of number of correct detection and recognition, and the location accuracy for the correctly detected and recognized targets, varies with the different target model sets in use, as shown in Figure 2.

The best performance occurred for the 10% model set, where the detection and recognition rate is satisfactorily high, and the location performance is perfect. This shows that enough discriminatory power is retained by only the 10% of the target boundary points which are properly chosen. When the percentage index of the model sets increases, more and more boundary points are used which include

points that not only are marginally discriminatory but also enlarge the difference between the assumed model for the noise and the actual noise itself.

We also ran the algorithm with model sets of lower percentage indices. When the percentage index is smaller than 5%, the performance drops rapidly. This is because too few boundary points are used which do not bear enough feature information to detect targets and to discriminate between them.

We see that the model sets with percentage indices between 10% and 50% yield quite good performance in both correct detection and recognition rate, and location accuracy for correctly detected targets, although we should report that the best result on this test image set of 30 correct detection and recognitions has been achieved by the maximum likelihood ratio algorithm [4] with a 60% model set constructed in a different way.

5. Conclusion

We described a technique for estimating the location covariance matrix of the target's centroid in grey level images, and using it for target detection and recognition. The estimation procedure was validated by a statistical test. Experiments on FLIR imagery yielded quite satisfactory results. Thus the use of a target's location covariance is a promising alternative as a criterion for detecting and locating targets.

Research is still under way for taking out the assumption of one and only one target per image. This will involve a proper way of setting up thresholds for the trace measurements associated with the hypotheses.

References

- [1] S. Z. Der. Automatic target recognition using passive infrared and laser radar sensor. Technical Report CAR-TR-762, Comp. Vis. Lab, Center for Automation Research, Univ. of Maryland, College Park, March 1995.
- [2] S. Z. Der and R. Chellappa. Probe-based automatic target recognition in infrared imagery. *IEEE Trans. on Image Proc.*, 6(1):92–102, Jan 1997.
- [3] R. M. Haralick. Propagating covariance in computer vision. *Int. J. of PR and AI*, 10(5):561–572, 1996.
- [4] G. Liu and R. M. Haralick. FLIR ATR algorithm using maximum likelihood ratio criterion. Technical report, Int. Sys. Lab, Dept. of EE, Univ. of Washington, May 1996.
- [5] G. Liu and R. M. Haralick. Automated construction of templates for matching. In *1997 Image Understanding Workshop Proceedings*, volume II, pages 1247–1254, New Orleans, LA, May 1997.
- [6] G. Liu and R. M. Haralick. Performance evaluation of template matching approaches to location estimation (1-d case). Technical report, Int. Sys. Lab, Dept. of EE, Univ. of Washington, Mar 1997.