Adaptive gate multifeature Bayesian statistical tracker

W. B. Schaming
RCA Advanced Technology Laboratories
Camden, New Jersey 08102

Abstract

A statistically based tracking algorithm is described which utilizes a powerful segmentation algorithm. Multiple features such as intensity, edge magnitude, and spatial frequency are combined to form a joint probability distribution to characterize a region containing a target and its immediate surround. These distributions are integrated over time to provide a stable estimate of the target region and background statistics. A Bayesian decision rule is implemented using these distributions to classify individual pixels as target or nontarget. An adaptive gate process is used to estimate desired changes in the tracking window size.

Introduction

This paper documents progress during the past year toward the development and demonstrations of a statistical tracking algorithm. Papers[1,2] presented in 1981 described some of the initial concepts in this development. Since that time, the statistical tracking algorithm has been expanded to incorporate (a) the simultaneous use of multiple features, (b) an adaptive gate process for control of the window size, and (c) positional dependence of the misclassification cost factor.

The tracking algorithm is based on the use of multifeature joint probability density functions for the statistical separation of targets from their background. The features currently being used are intensity, edge magnitude, and a pseudo spatial frequency feature. These features are combined to form the joint distributions which characterize a target region and its immediate surround. The distributions are integrated over time to provide a stable estimate of the target and background statistics. A Bayesian decision rule is implemented using these distributions to classify individual pixels as target or nontarget within a tracking window. An adaptive gate process is used to estimate desired changes in the tracking window size. The algorithm at present assumes manual target designation.

RCA believes this tracking process is capable of operation in all environments; insensitive to target type, signature, and orientation; applicable to a variety of sensors; and extendable to multisensor processing and readily implementable.

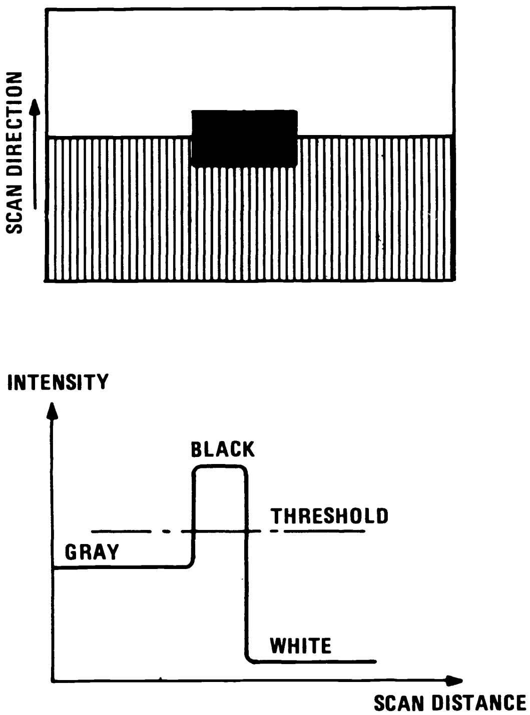
Preprocessing and A/D conversion

The video preprocessing function is an important part of any imaging sensor system, but is more critical when the sensor is an IR device which may exhibit very high dynamic range capability. In this case it is insufficient to perform a simple AGC based upon global statistics because the subsequent rescaling to reduce the dynamic range will destroy the low contrast local detail. Instead, some form of local adaptive contrast enhancement should be applied in which the gain varies with the local contrast. Lo[3] simulated and compared several such techniques.

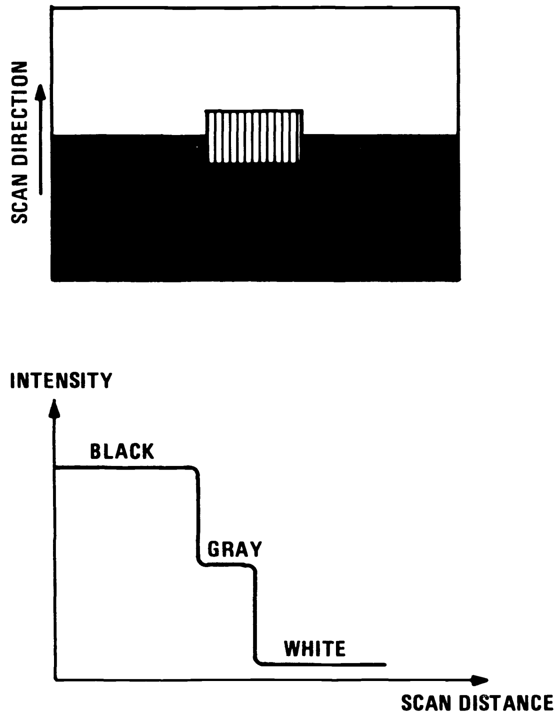
Although necessary in a hardware implementation, this function has not been included in the simulations reported here. Ten-second image sequences were digitized from video tape via an analog video disc and an image processing system. The input to the image processing system was passed through a video processing amplifier so that the levels could be properly matched to the A/D converter.

Statistical tracking algorithm

Targets are often separated from their background by a simple thresholding scheme. Sometimes the computation of the threshold is quite sophisticated and involves looking at the statistics of the video signal. However, thresholding is inherently limited in ability as can be seen by the diagrams in Fig. 1. A simple black and white target can be readily thresholded to isolate it from its background. On the other hand a gray target cannot be thresholded without using a pair of thresholds properly placed to contain the intensity levels on the target. This dual threshold in itself is not prohibitive, but rather the problem lies in the ability to place the thresholds at the appropriate levels.



(a) THIS TARGET CAN BE EASILY THRESHOLDED



(b) THIS TARGET CANNOT BE EASILY THRESHOLDED BUT REQUIRES A PAIR OF THRESHOLDS PROPERLY PLACED.

Fig. 1. Example showing two postulated targets. One is easily segmented from the background using a single threshold. The other, however, requires two thresholds which are not easily determined. The statistical process provides a separate threshold for each intensity level.

The statistical segmentation process is a technique which provides an improved method for extracting the target from its background. Figure 2 depicts this process. Shown are two histograms, one taken from a window area of the image containing the target and the other taken from the immediate surround which represents the background. A single feature, intensity, is shown in these histograms for illustrative purposes. The shape of the distribution shown is arbitrary; there are no assumptions made about their actual shape. The segmentation process makes a separate assessment of each bin in the histogram to determine if pixels whose intensity falls in the bin are more likely to be target or background. In addition to solving the threshold selection problem, the statistical tracking algorithm provides a method to both simplify the multimode tracking concept and provide added capability.

The simplification comes about in the following way. State-of-the-art multimode trackers typically operate a contrast, edge, and correlation tracker in parallel. An executive process may be defined to determine at any given time which tracking mode is providing the most reliable estimate of target position. The statistical process, as currently defined, eliminates this mode polling process by combining the available features into multidimensional statistics representing target and background. Consider the use of intensity and edge magnitude as the two candidate features. In this case the statistical approach encompasses three tracking modes in an integrated single mode without the need to poll the performance of the individual processes. When intensity is the best target background separator, the algorithm operates like a contrast tracker. When edge magnitude is predominate it operates similar to an edge centroid tracker. Because the process is searching for pixels in the current frame that are statistically similar to those pixels selected as target in previous frames, the algorithm is in a sense a correlation type process as well.

The added capability comes from the fact that there are target/background conditions which are inseparable using two features independently but are readily separable using the same two features jointly. This is illustrated quite simply in Fig. 3. In this example, neither edge magnitude nor intensity can be used independently to separate the target from background because both flat distributions cover the entire variable range for both features. On the other hand, the joint distribution clearly delineates the two areas.

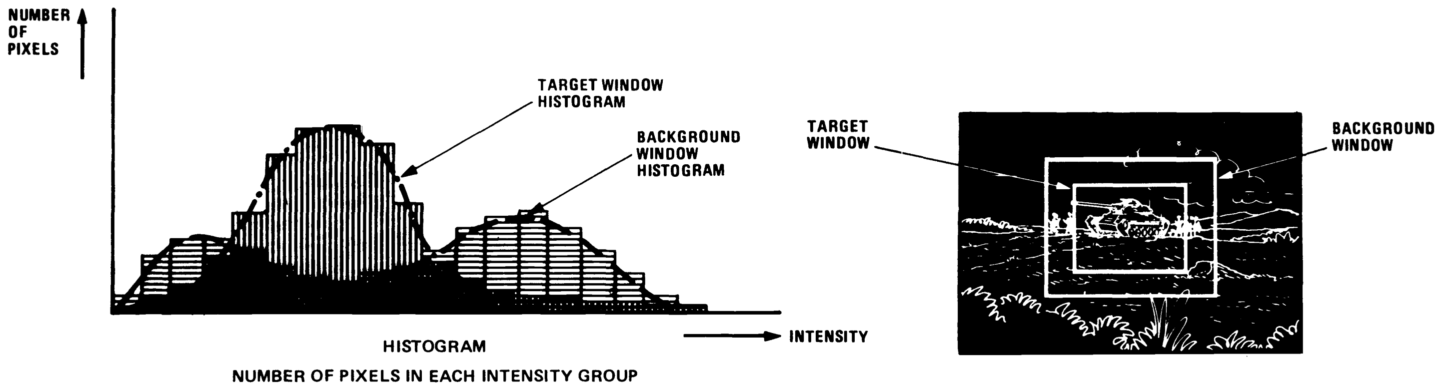


Fig. 2. Example of how histograms are used to separate a target from its background. Each bin in the histogram is examined to determine if the intensity value falling within that bin are more likely to be target or background. Although this is a single feature (intensity) example, the same process is used with multiple features in an N-dimensional histogram representing a joint probability density.

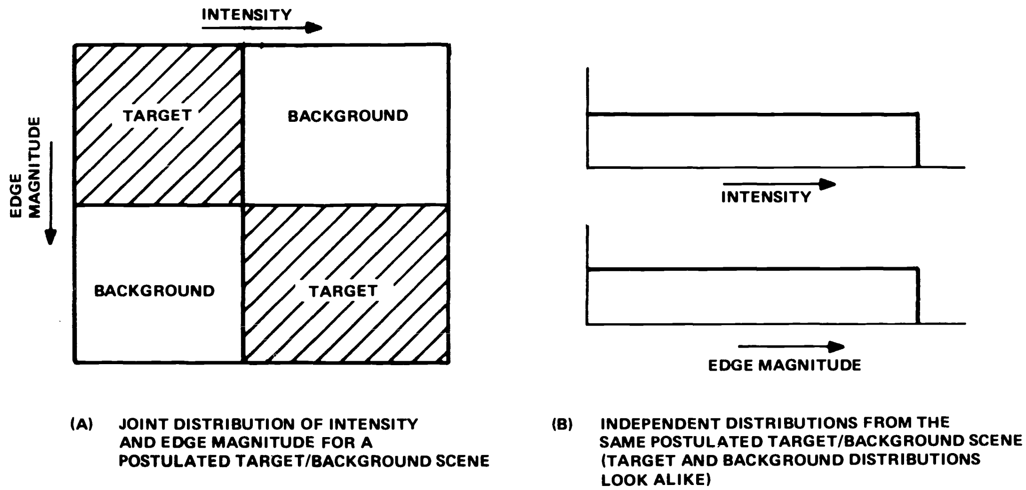


Fig. 3. Simple example showing how the use of joint statistics aids in the separation of target from background in situations where the use of the features singly fails.

Figure 4 is a flow diagram of the statistical tracking mode. The preprocessed video is used to generate multiple feature images to be used in the decision process. The features are combined into two joint probability density functions for (a) a target tracking window and (b) a background window frame. These distributions are the basis of a statistical decision process which is used to classify the image pixels inside the tracking window to separate the target from the background. In actuality the statistics from previous frames are used in the classification process for the current frame. At the same time, histograms are generated from the current image frame so that the statistics can be updated for processing subsequent frames. At the end of the classification process the segmented image is analyzed to determine the appropriate error signals as well as the window size and position for the next frame. In parallel with the pixel rate computations for the Nth frame, the statistics from the N-1st frame are integrated with past history and a decision rule is generated for the N+1st frame.

A sample output from the process is shown in Fig. 5. Only two features were used for this example, namely, intensity and edge magnitude. The total number of bits utilized for the features is seven — four for intensity and three for edge magnitude. The edge magnitude used is the absolute value approximation to the Sobel operator.

The next few paragraphs describe some of the steps in this process in more detail.

<u>Computation of features</u>

The first step in the statistical process is the generation of the features to be used. There are many potential candidates, some of which are computationally too burdensome for real-time implementation at this time. We therefore have limited our selection of features

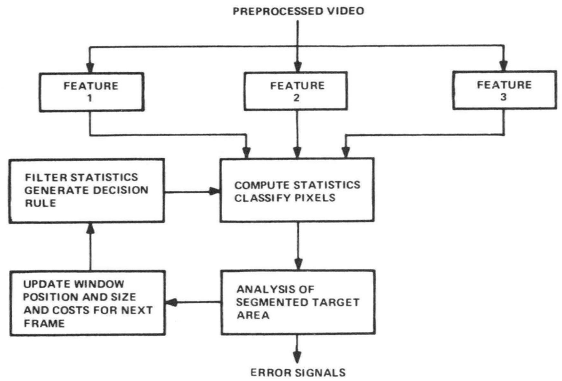


Fig. 4. Block diagram of the Bayesian statistical tracking mode. The feature computation, statistics generation, and pixel classification are performed at the pixel rate. The computation of error signals is performed during vertical sync.

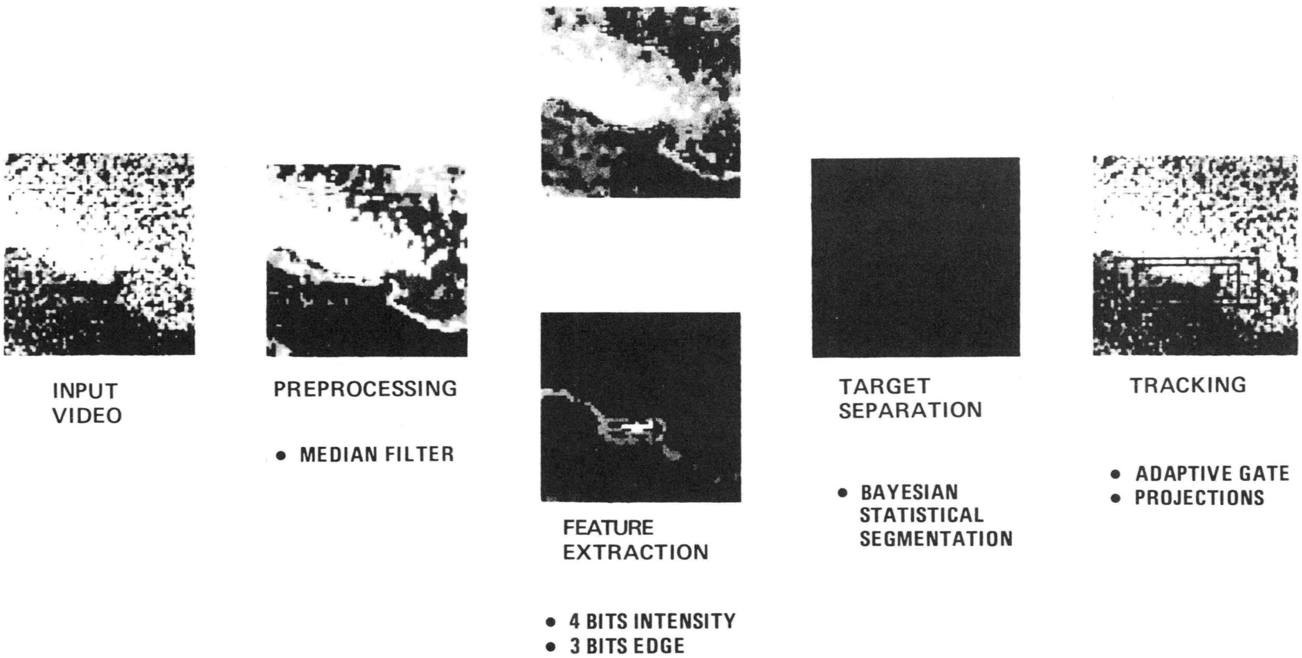


Fig. 5. Sample output from the Bayesian statistical tracker simulation using a 64-x-64 pixel image of an aircraft at a mountain boundary. Two features were used in the statistical segmentation with a total of seven bits.

to those which are readily implemented. These features are intensity, edge magnitude, and spatial frequency.

The intensity feature is simply a requantized version of the digitized video signal to obtain the desired number of bits of intensity resolution. The edge magnitude feature is the sum of absolute values approximation to the Sobel operator. The absolute sum is an acceptable and computationally more appealing approximation than the true edge magnitude.

The third feature is an approximation to spatial frequency in the horizontal direction. Because it is a measure of object size, it could also be considered a simple texture measure in a broad sense. The spatial frequency is defined as the function of the run length where a run is the number of consecutive pixels between which the pixel-to-pixel difference does not exceed a predefined threshold. The threshold used is the mean value of the absolute difference beweeen pixels in the previous frame. The feature value is then defined as:

$$SF = \text{MAXIMUM} \left[0,\ (2^N - \text{RUN LENGTH}) \right] \tag{1}$$

where 2^N is the number of levels into which the spatial frequency feature will be quantized

An example of the spatial frequency feature is shown in Fig. 6. An arbitrary function is plotted to represent the image intensity I at successive pixels in the x direction. Beneath the plotted data are shown the actual pixel intensities, absolute differences, run lengths, and feature values. The threshold used to compute run lengths in the example is 1.3 and the number of quantization levels is 8(N = 3 bits). The first sample-to-sample difference which exceeds the threshold 1.3 is the sixth sample. Samples 1 to 5 represent a run of length 5 in which the differences do not exceed threshold. The corresponding feature value is 3 which is assigned to all pixel locations in the run. The higher feature values indicate smaller distances between gradient values exceeding threshold. Note that the low amplitude variation between the pixels 6 and 14 do not exceed the threshold and therefore do not define the boundary of a run. The feature is intended to provide information about the size (in the x direction) of areas or patches which have uniform or slowly varying intensity.

Generation and integration of statistics

Histograms from two separate regions in the image must be computed to provide the probability density functions required by the decision rule. The regions from which the histograms are generated are shown in Fig. 7. The assumption in the segmentation algorithm is that the target is absent from the frame region. For both the frame and window regions a multifeature histogram is defined as

$H_{FR}^{N}\ (f_1, f_2, f_3)$ Frame Region Histogram

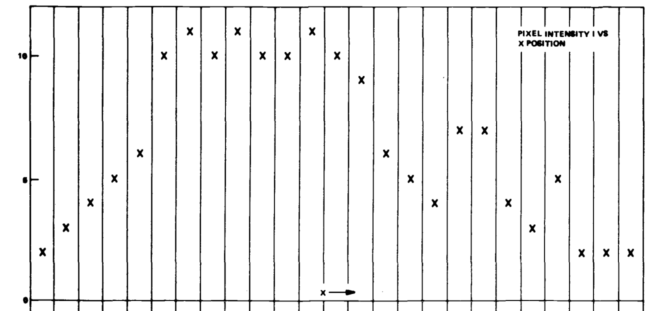
$H_{WR}^{N}\ (f_1, f_2, f_3)$ Window Region Histogram

for the Nth image in the sequence.

After normalization by the respective areas of the frame and window regions the histograms become the discrete joint probability densities

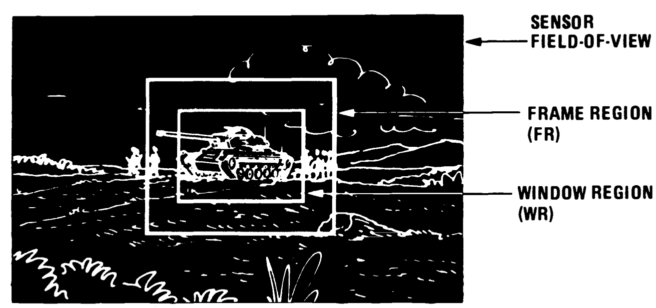
$P_{FR}^{N}\ (f_1, f_2, f_3)$

$P_{WR}^{N}\ (f_1, f_2, f_3)$.



I	2	3	4	5	5	10	11	10	11	10	10	11	10	9	6	5	4	7	7	4	3	5	2	2	2
DIFF		1	1	1	1	4	1	1	1	1	0	1	1	1	3	1	1	3	0	3	1	2	3	0	0
RUN	5	5	5	5	5	9	9	9	9	9	9	9	9	9	3	3	3	2	2	2	2	1	3	3	3
FEAT	3	3	3	3	3	0	0	0	0	0	0	0	0	0	5	5	5	6	6	6	6	7	5	5	5

Fig. 6. Sample which shows the procedure for calculating the pseudo spatial frequency feature. The absolute difference threshold used to compute run lengths in the example is 1.3, which is the average difference. The number of quantization levels for the feature is 8.



$H_{FR}\ (f_1, f_2, f_3)$ – multifeature histogram from frame region

$F_{WR}\ (f_1, f_2, f_3)$ – multifeature histogram from window region

Fig. 7. Areas of the image over which the multifeature histograms are computed. It is assumed that the target is absent from the frame region which is defined as a border around the window region containing the target.

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

Smarter legal research.