# Image Cover Sheet

| | |
|---|---|
| **CLASSIFICATION**<br>UNCLASSIFIED | **SYSTEM NUMBER** 35539 |

**TITLE**

TARGET TRACKING BY IMAGE SUBTRACTION

**System Number:**

**Patron Number:**

**Requester:**

**Notes:**

**DSIS Use only:**

**Deliver to:** JR

UNCLASSIFIED
UNLIMITED DISTRIBUTION

CRDV RAPPORT 4226/82
DOSSIER: 3632D-008
JANVIER 1982

DREV REPORT 4226/82
FILE: 3632D-008
JANUARY 1982

UNLIMITED
DISTRIBUTION
ILLIMITÉE

# TARGET TRACKING BY IMAGE SUBTRACTION

J.F. Boulter

Centre de Recherches pour la Défense
Defence Research Establishment
Valcartier, Québec

BUREAU - RECHERCHE ET DÉVELOPPEMENT
MINISTÈRE DE LA DÉFENSE NATIONALE
CANADA

NON CLASSIFIÉ
DIFFUSION ILLIMITÉE

RESEARCH AND DEVELOPMENT BRANCH
DEPARTMENT OF NATIONAL DEFENCE
CANADA

# TARGET TRACKING BY IMAGE SUBTRACTION

by

J. F. Boulter

CENTRE DE RECHERCHES POUR LA DEFENSE
DEFENCE RESEARCH ESTABLISHMENT
VALCARTIER
Tel.: (418) 844-4271

Quebec, Canada

January/janvier 1982

## RESUME

On peut rehausser et poursuivre des cibles qui se déplacent devant un arrière-plan stationnaire en soustrayant les images obtenues de la scène à des temps différents. On étudie, par la simulation numérique, les conditions nécessaires pour qu'un système de cette sorte donne une performance adéquate, et on montre deux exemples de la poursuite automatique de cibles. Le premier démontre la poursuite d'un véhicule blindé de transport de personnel (VBTP) enregistrée sur un film pris dans la lumière visible, et le deuxième, la poursuite de chars d'assaut et de VBTP dans une séquence d'images infrarouges. On décrit comment on peut construire un système fondé sur la soustraction des images et capable de poursuivre, en temps réel, des cibles multiples dans des images prises avec une caméra de télévision. (NC)

## ABSTRACT

50

//Targets moving against a stationary background can be enhanced and tracked by subtracting images of the scene obtained at different times. We examine by digital simulation the conditions required for such a system to give adequate tracking performance, and we give examples of the automatic tracking of an APC target in a visible-light cinefilm and of tank and APC targets in a sequence of FLIR images. We propose a possible hardware implementation of a system based on image subtraction for tracking multiple targets in real time in television-line-scan imagery.// (U)

## TABLE OF CONTENTS

## 1.0 INTRODUCTION

Targets of interest moving against a fixed background may be extracted and tracked by subtracting registered images of the scene obtained at different times. The stationary background cancels leaving the objects that have changed position. A limitation of the method is that targets which do not move laterally in the field of view of the imaging sensor, but move only along its line of sight, cannot be tracked. Such processing can now be readily performed at standard television-video rates. We examine by means of digital simulation target tracking based on image subtraction, and present a possible hardware implementation of it.

Section 2.0 defines and examines the conditions required for acceptable tracking performance. In Sect. 3.0, we give examples which show the tracking of an APC target in a 16-mm color cinefilm taken in visible light, and the tracking of multiple tank and APC targets in a sequence of IR images obtained with a FLIR system. Section 3.0 also examines experimentally some of the factors that limit the practical performance of a target-tracking system based on image subtraction. In Sect. 4.0, we describe a possible system based on image subtraction for tracking multiple targets in television-line-scan imagery, and we compare the analog and the digital implementations of it.

This work was performed at DREV during the period December 1980 to March 1981 under PCN 32D08, Target Enhancement Using Spectral and Textural Information.

# 2.0 THEORY

## 2.1 Requirements

Assume a fixed imaging system views a scene consisting of a target moving against a stationary background. We will subtract the images of the scene obtained at different times to enhance and track the moving target. If the requirements described in the present section are satisfied, the subtraction will cancel the stationary background and only the areas corresponding to the moving target will have intensities significantly different from zero. This permits us to determine the target position and to track it by using simple thresholding and centroid-estimation operations.

Three general conditions must be satisfied to permit successful target tracking by image subtraction. First, the target motion resolution must be adequate, i.e., the distance the target moves between successive images must be larger than the spatial resolution of the target image. Second, the background regions of the two images must be sufficiently well registered, both in space and in intensity, to insure that it cancels adequately relative to the intensities of the moving target regions. Third, the target must contrast in intensity with its immediate background in both images.

## 2.2 Motion Resolution

Suppose we acquire every T seconds the image of a target moving with a velocity v. Each image is obtained by integrating over the time interval t. Assume the spatial resolution r of the target images is composed of a part p caused by imperfections and noise in the imaging system, and a part q caused by motion blurring of the target occurring over the time interval t. We approximate r by adding the two contributions in quadrature:

$$r = \sqrt{p^2 + q^2} = \sqrt{p^2 + v^2t^2} \qquad (1)$$

The ratio of the distance the target moves between successive images to the spatial resolution of the target image in the direction of its motion is:

$$R = vT/\sqrt{p^2 + v^2t^2} \qquad (2)$$

Good target motion resolution, i.e., the ability to easily detect the target motion when successive images are subtracted, is achieved when R is sufficiently greater than one.

For fast targets, where motion blurring is dominant, the ratio can be approximated by:

$$R = vT/vt = T/t \qquad (3)$$

In this case, the ratio is independent of the target velocity, and we obtain good target motion resolution by making the integration time t shorter than the interval T between frames. For slow targets, or short exposure times, where motion blurring is negligible:

$$R = vT/p \qquad (4)$$

and a large R value is obtained when the distance the target moves between successive images is much greater than the spatial resolution of the imaging system.

In practice, the value for R that will be acceptable depends on how well the background cancels when the images are subtracted. A background that cancels poorly due, for example, to the motion of trees or clouds present in it will require a larger R value than a background of stationary ground terrain which cancels well. On the other hand, vT cannot be too large or the target may pass out of the field of view of the imaging system between successive frames.

Note that some imaging systems may not satisfy the requirement that R be much greater than one when successive frames are subtracted. Consider a cinefilm camera operating at 24 frames/s (T=0.042 s) with an exposure time t=0.004 s, and a vidicon television line-scan camera with a 0.03-s decay constant (t=0.03 s) operating at 30 frames/s (T=0.033 s). Suppose both view a well-resolved fast target such as an aircraft. Target resolution is limited by motion blurring in both cases, so we use eq. 4 to calculate R. In the former case, R=T/t=0.042/0.004=10.5, i.e., the target moves 10.5 times the blur length between frames. In the latter case, R=0.033/0.03=1.1, i.e., the target moves only 1.1 blur length between frames. The target motion is more apparent and easier to detect when successive frames from the cinefilm are subtracted than when successive frames from the vidicon camera are subtracted.

In some cases, where the subtraction of successive frames does not give a sufficiently high R value to permit good target detection, the frame rate may be fixed and not subject to change. Here, we can increase the effective R value by storing in a buffer memory a number of successive frames, and then subtracting images separated by more than one frame interval, e.g., subtract Frame 1 from Frame 5, Frame 2 from Frame 6, etc.

## 2.3 Spatial Registration

The stationary background region in the two images to be subtracted will not cancel unless they are spatially registered. We can use standard spatial correlation techniques (Refs. 1-2) to compare the background regions of successive images to determine the coordinate transformation required to achieve spatial registration. For simple types of sensor motion, this may require only a horizontal or a vertical translation of one image relative to the other one. Optical correlation techniques, fast digital correlation algorithms such as the SSDA (Ref. 1), or those based on the fast Fourier transform can yield the magnitude and the direction of the required displacement.

In more complex situations, such as when the imaging system is on board a missile closing on a target, registration may require, in addition to the preceding translation, a rotation, a scaling or even a description of the three-dimensional nature of the scene. A variety of approaches have been proposed to determine the parameters required to specify the coordinate transformation. These may be based on the scale invariant Mellin transform (Refs. 3-4), the rotation invariant polar Fourier transform (Refs. 3-4) or on the correlation of selected subregions containing common features or "control points" (Refs. 5-6). In general, the image-registration techniques being developed for correlation-tracking applications (Refs. 1-2 and 5-6) are directly applicable to the problem of registering images for subtraction.

## 2.4 Intensity Registration

If the two images to be subtracted are degraded by different gray-scale nonlinearities, their background regions will not be registered in intensity and will not cancel. Provided the nonlinearities are invertible, intensity registration can be restored without actually determining the nonlinearities by using histogram modification techniques (Ref. 7). An invertible nonlinearity is one that transforms each input gray level into a unique output gray level. A shift in the DC level or in the gain of a television video amplifier are examples of invertible nonlinearities. On the other hand, hard limiting is not invertible.

Histogram modification is performed by measuring the intensity distribution of an image, and then applying the gray-scale transformation required to modify the distribution to a specified shape. For example, histogram equalization produces an image with a uniform distribution of gray levels, i.e., all gray levels appear with equal frequency independent of their probabilities in the original image. The gray-scale transformation required to perform histogram equalization is the integral of the intensity distribution of the original image.

We can restore the intensity registration of the background regions of two images distorted by different gray-scale nonlinearities by histogram equalizing them, or modifying them to some other shape. This assumes that the scene content in the two images is the same. While the intensities in the two histogram-modified images will equal one another, they will not equal the intensities in the original scene.

## 2.5 Target Segmentation

Suppose we obtain at different times two images of a uniform intensity target moving against a uniform intensity stationary background. Assume the spatial and the intensity resolutions of the imaging system are infinite, and that the integration time t=0, so there is no motion blurring. The target will appear in the two images as an extended object with a uniform intensity j against a background with a uniform intensity k. When we subtract the first image from the second one, the region at the leading edge of the target, i.e., in the direction of its motion, will have an intensity j-k and the region at the trailing edge of it an intensity k-j. Except for these two areas, all other regions of the difference image will be zero.

In practice, system noise, the motion of objects in the background, and spatial and intensity misregistrations (Sect. 2.3 and 2.4) will prevent the complete cancellation of the background, and an RMS residue b will remain after the subtraction is performed. We use the ratio $|j-k|/b$ as a measure of the target contrast in the difference image. High values of this ratio are desirable and values close to one indicate that the target will be difficult to distinguish from the background.

To extract the target from the background, we classify as belonging to the target region all elements of the difference image with an absolute value that exceeds a specified threshold value. All other elements are classified as belonging to the background. The threshold value must be sufficiently greater than the RMS background b to limit the false-alarm rate. The average position of the target during the time between the acquisition of the two images is assumed to equal the centroid of the elements classified as belonging to the target. If the intensity of the target is known, a priori, to be either less than or greater than that of the background, the polarity of the two regions corresponding to the leading and the trailing edges of the target can be used to determine the direction the target is moving. In IR imagery, for example, military targets are usually hotter than their backgrounds, i.e., j is greater than k. Therefore, the region with positive polarity represents the leading edge of the target and the one with negative polarity the trailing edge of it.

Spatial filtering the difference image after taking its absolute value, but before thresholding it, may improve our ability to extract the target from the background if the target and the background regions have different spectral contents. Background noise often has a uniform spatial-frequency content, for example, whereas an extended target has more power at the lower frequencies. In this case, low-pass filtering would increase the amplitudes of the target regions relative to those of the background.

Bear in mind that the preceding discussion of target extraction is highly simplified, and is only intended to illustrate some of the basic factors that are involved. First, real targets and backgrounds will not have uniform intensities, and some areas of the target may have a higher gray level than the background, whereas other areas of it may have a lower gray level. The extracted target will then consist of a number of regions of different polarity, rather than the two regions considered above. Second, as discussed in Sect. 2.3, motion blurring and the limited resolution of the imaging system will prevent the target edges from being sharp. Third, the background may contain extended moving objects, such as trees or clouds, that resemble the moving targets of interest.

## 2.6 Spectral Ratioing

Spectral ratioing consists in obtaining two images of a scene in different spectral ranges, e.g., in red and blue light, and dividing one by the other. As described in Refs. 8 and 9, the spectral ratioing of images formed by the reflection of radiation removes the contrast caused by changes in the viewing conditions, such as the orientation and the illumination of the surfaces, but retains that caused by differences in the spectral reflectivites of the surfaces. In a visible-light image, the intensity reflected from a target with a homogeneous surface may vary because of shadows cast on it, or because of changes in the viewing angles of the target surfaces. In this case, it may be easier to track the target in the ratio image where this contrast is absent or reduced, than in the intensity image. For example, the contrast reversals that frequently occur when we view an aircraft moving against a sky or cloud background can hamper correct tracking. The frequency of such contrast reversals may be reduced by using the spectral-ratio images in place of the intensity ones.

The spectral ratioing of images formed by the emission of radiation can reduce the contrast caused by differences in the surface emissivity relative to that due to differences in temperature (Ref. 10). In some situations, the ratioing of imagery obtained with a "multicolor"

IR imaging system may improve the target tracking performance. Consider, for example, a hot target moving against a natural background that has a large variation in surface emissivity but a relatively constant temperature. The target may be more easily tracked by subtracting the ratio images where the spectral ratio of the background appears more uniform than it does in intensity.

## 3.0 EXAMPLES

### 3.1 Visible-Light Cinefilm

We recorded at 24 frames/s on a 16-mm color cinefilm an APC target moving from right to left against a background of vegetation and ground terrain. An interactive digital image-processing system (Ref. 11) was used to digitize, to a resolution of 256 by 256 elements, square areas measuring 0.5 cm on a side on 16 frames of a first-generation copy of the original cinefilm. We only digitized every twelfth frame, so successive images are separated by 0.5 s. The APC target measures 0.6 mm wide on the film and 32 elements in the digital image. The target moves approximately 13 elements between successive digital images.

Since the digitized images represent the enlargement of a small area on the film, the resolution is severely limited by film granularity. The original images are given in Fig. 1, and one can see that the APC target is poorly defined and does not contrast well with its background because of the limited film resolution. As shown in Fig. 2, high-pass filtering improves, somewhat, the visibility of the target, but at the expense of an increase in the noise level.

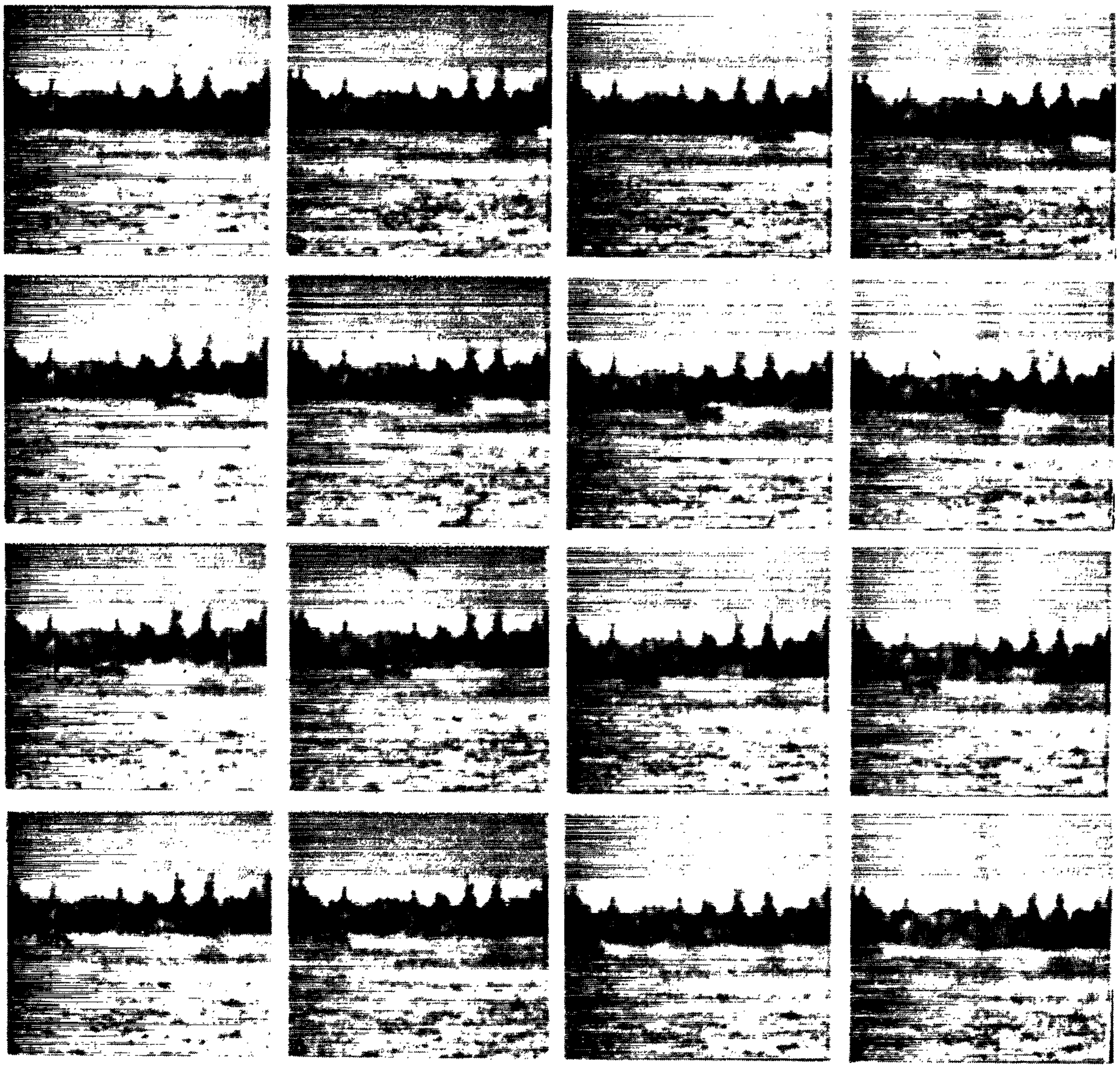


FIGURE 1 - Digitized cinefilm sequence showing an APC crossing the field of view from right to left. Each image represents the enlargement of a small area on the film, and the target resolution and contrast are limited by the film granularity.

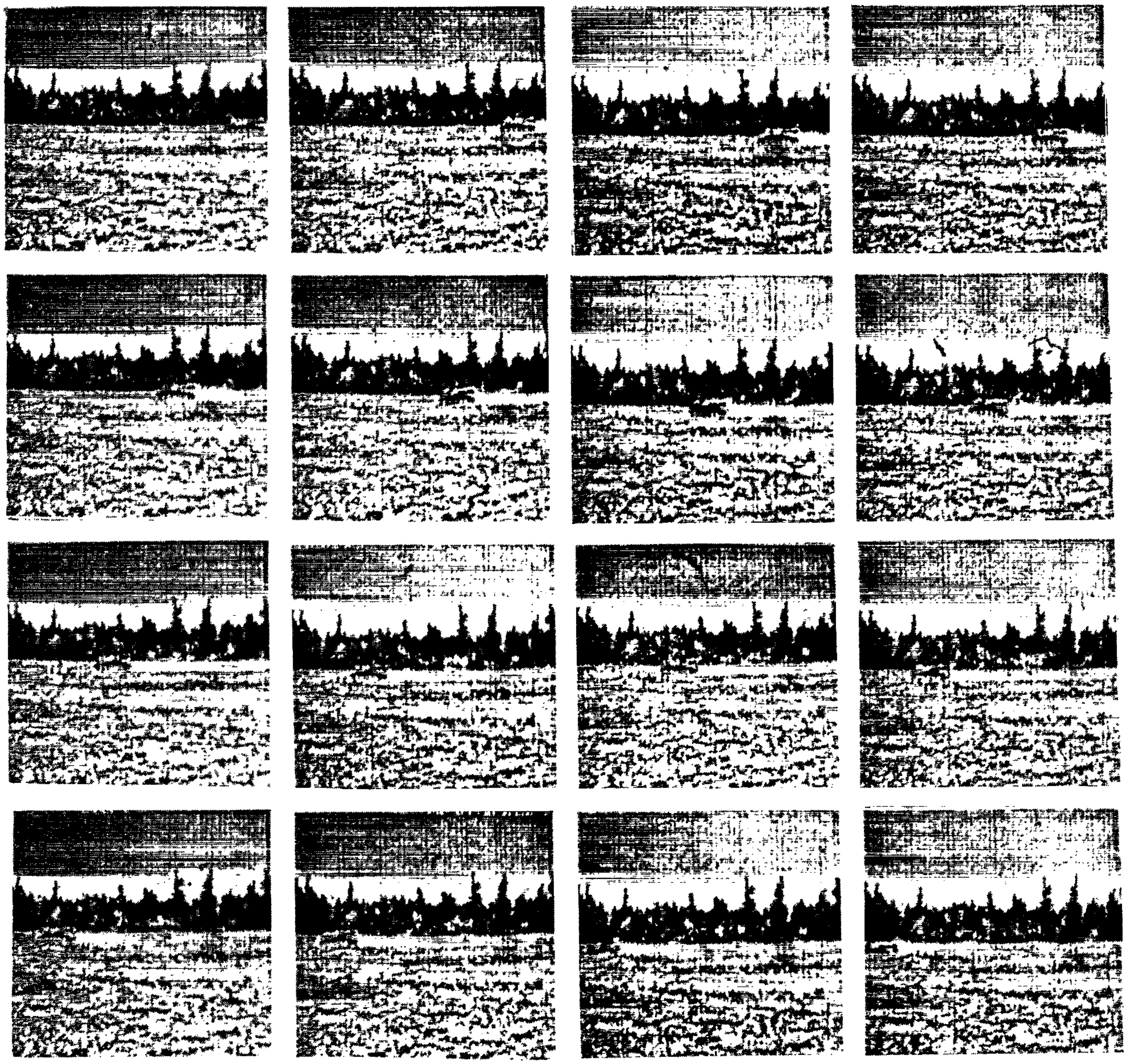


FIGURE 2 - The cinefilm sequence given in Fig. 1 after high-pass filtering to enhance the visibility of the target

Figure 3 gives the 15 difference images corresponding to the 16 original images shown in Fig. 1. The upper left photograph in Fig. 3 was obtained by subtracting Image 1 from Image 2, the next photograph to the right of it by subtracting Image 2 from Image 3, etc. We added a bias level of 0.5 before displaying each difference image so that a correctly cancelled region appears with a mid-gray intensity, a negative difference appears darker, and a positive difference lighter.

Despite the poor quality of the images given in Fig. 1, the moving target is clearly apparent in the difference images shown in Fig. 3. Dust particles present on several frames are also evident in the difference images. These particles first appear with a negative polarity and then, in the next difference image, with positive polarity. The low-contrast dust cloud thrown up by the moving APC can be seen in the difference images and, in several cases, motion of trees in the background region is apparent.

We see in Fig. 1 that the upper part of the APC target is superimposed upon a background of trees, whereas the lower part of it has a background of ground terrain. On the average, the trees are darker than the target, whereas the ground is lighter than it. The target area in the difference images should, therefore, consist of four regions rather than the two regions predicted by the simple model given in Sect. 2.5. In particular, the leading edge of the target should consist of an area of positive polarity with a negative polarity area below it, whereas at the trailing edge of the target, the polarities of the upper and the lower areas should be reversed. These four regions are clearly evident in the difference images given in Fig. 3.

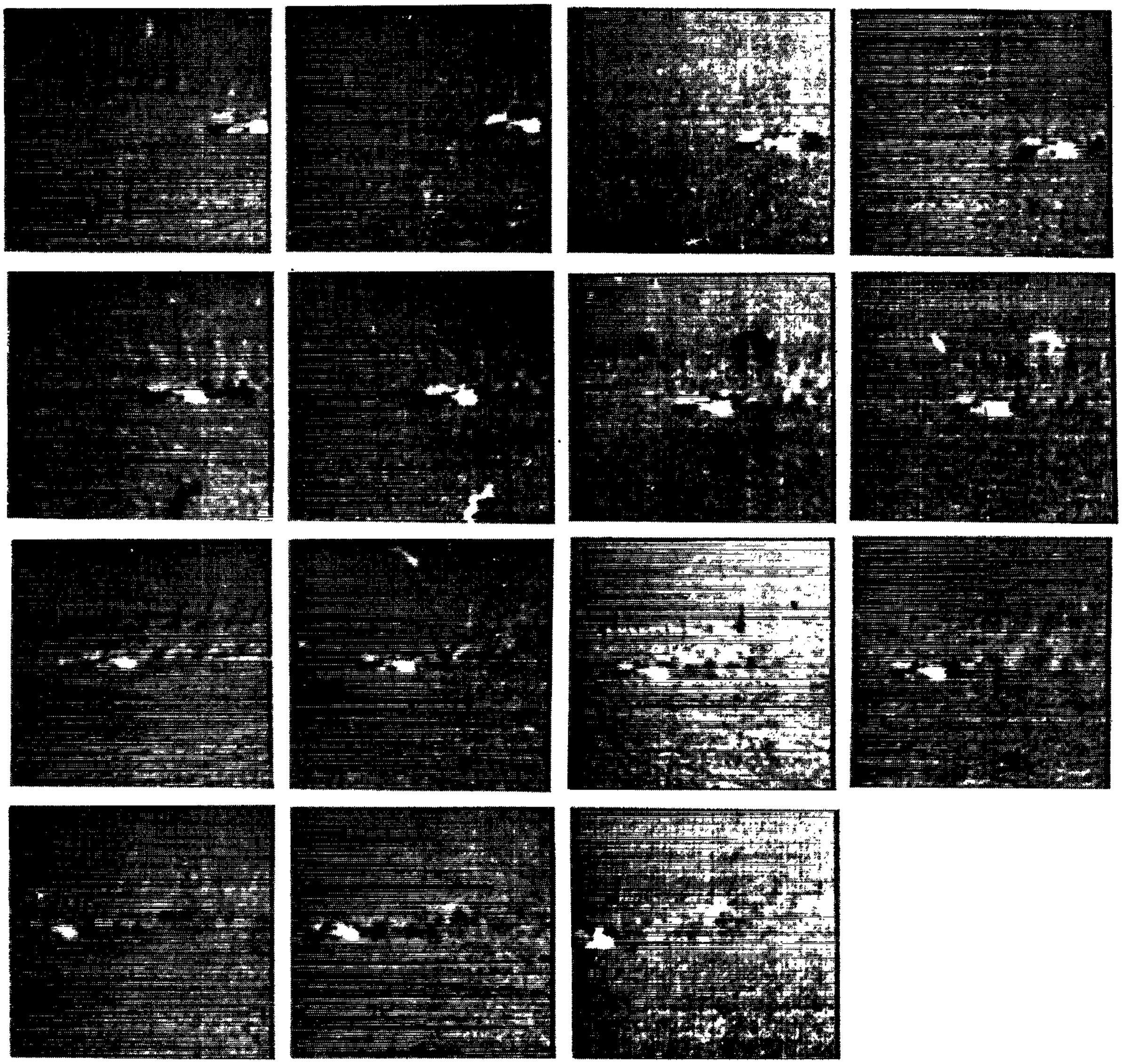


FIGURE 3 - The difference images obtained by subtracting successive images of the cinefilm sequence given in Fig. 1. The upper left photograph is Image 2 minus Image 1, the next one to the right Image 3 minus Image 2, etc. The display has been biased so that correctly cancelled regions appear gray, negative differences are dark and positive differences are light. The dark and light areas corresponding to the moving target are clearly evident.

As discussed in Sect. 2.5, we first low-pass filter the absolute value of the difference images to enhance the extended target areas relative to the film noise. Then we detect the target areas by setting to one all elements that exceed a specified threshold (here equal to 0.1 of the maximum intensity in the original images) and setting all others to zero. The resulting binary images are shown in Fig. 4. Note that the areas corresponding to the moving target have been cleanly extracted.

Next, we define a rectangular window to select the target region in the first thresholded difference image, and calculate the centroid of all elements classified as belonging to the target. This centroid is assumed to represent the average position of the target in the time between the acquisition of the first two images. The centroids of the other 14 thresholded difference images were similarly calculated. The size of the tracking window was held fixed, but its position was changed so that it followed the target. In particular, we set the window position for Difference Image N equal to its position for Difference Image N-1 plus the distance that the target moved between Difference Images N-2 and N-1.

Figure 5 shows the result of superimposing the tracking windows and the target centroids on the original images. The centroid and the window corresponding to Image 2 minus Image 1 is superimposed on Image 1, etc., so the designated target positions appear slightly ahead of the true target positions in all displays. Despite the poor target contrast and the relatively large noise level in the images, the target is correctly tracked in the 15 frames.

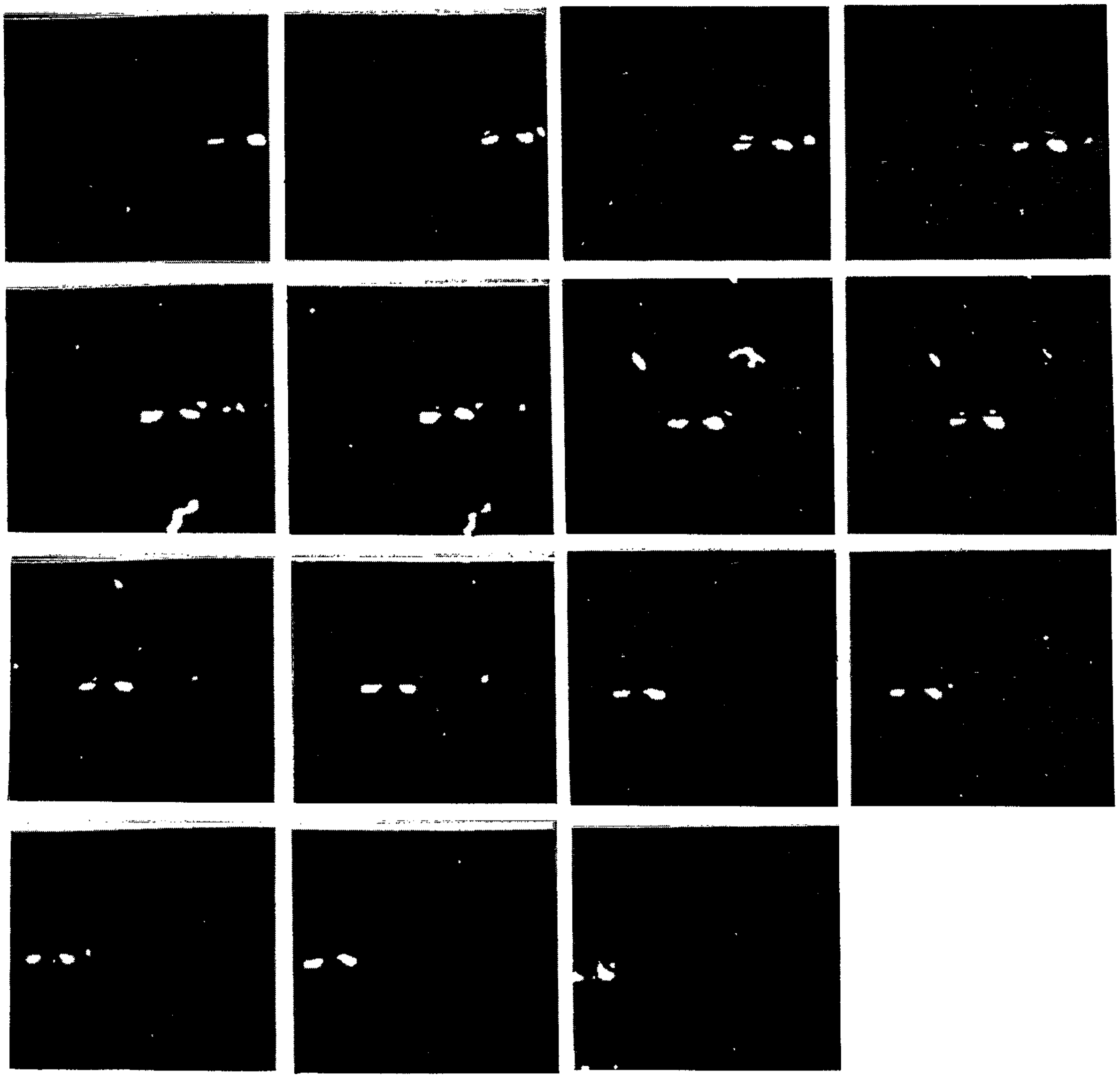


FIGURE 4 - The difference images given in Fig. 3 after thresholding and smoothing to extract the regions corresponding to the moving target

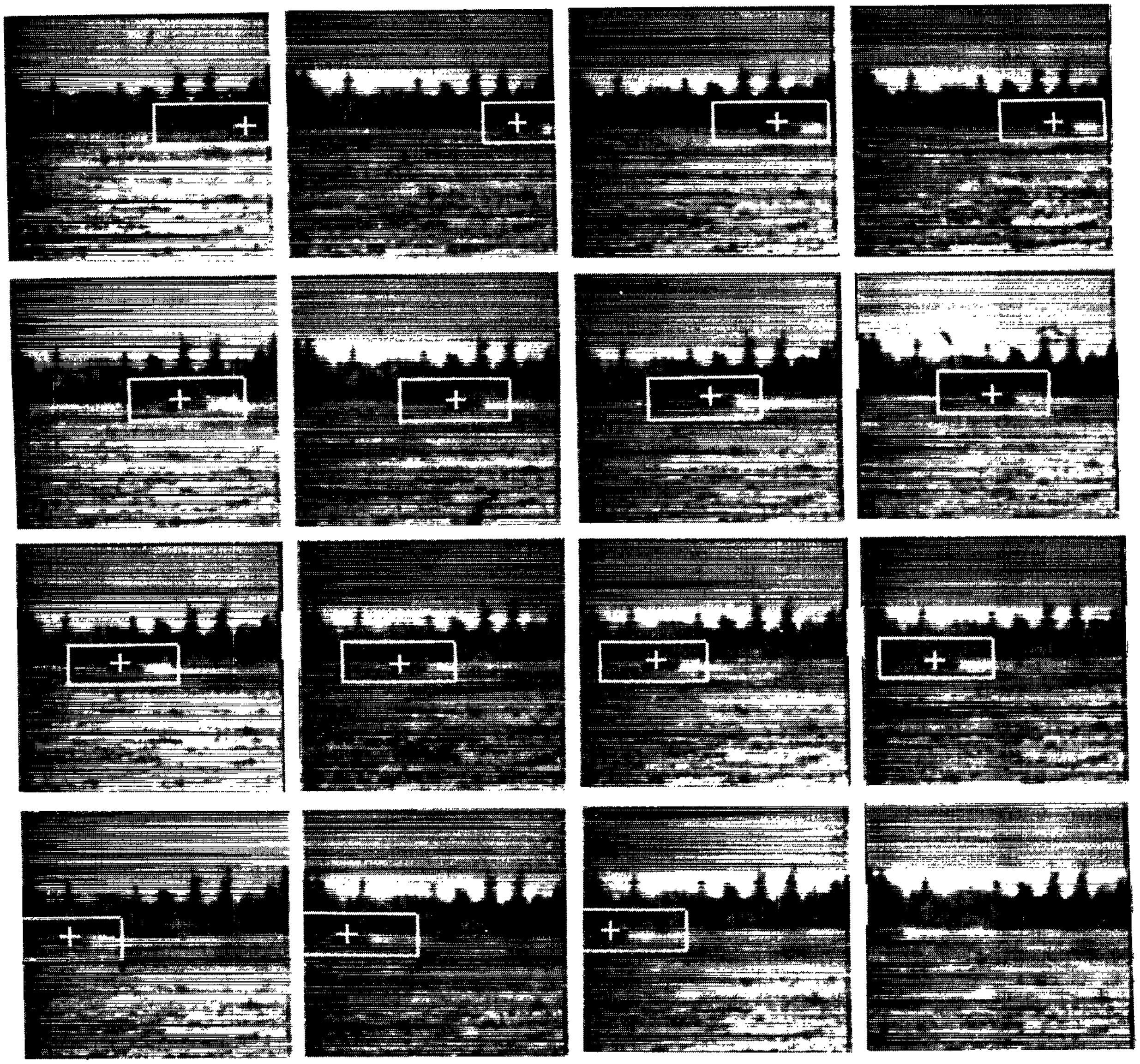


FIGURE 5 - The original cinefilm sequence with the tracking window outlined and the target positions, as determined from the difference images, designated by crosses. The target is easily tracked despite its low contrast and the high image noise level.

## 3.2 IR Sequence

We obtained from the Fort Polk data base (Ref. 12) a series of four digitized IR images taken in the spectral range 8-14 μm. These images have a spatial resolution of 512 by 256 elements. The original images, given in the left photographs in Fig. 6, show a number of moving and stationary M60 tanks and M113 APCs against a background of trees and ground terrain. The targets are hotter than the background and appear as white objects, between 10 and 25 elements wide, that contrast with the darker background. The left part of the third image contains sampling errors visible as dark vertical lines.

The right photographs in Fig. 6 show the three difference images after the addition of a display bias as described in Sect. 3.1. A tank moving just to the left of the center is clearly evident in the difference images, as are several other smaller APC targets moving at the left of the scene. Except for the vertical lines in the second and third difference images caused by the sampling errors in Image 3, all other background areas were cancelled by the subtraction. Note the good cancellation of the large hot stationary target located in the left of the scene identified as a "burning hulk" in the description of the data base.

We segmented and tracked the targets by using the procedures described in Sect. 3.1. Target segmentation was performed by calculating the absolute value of each difference image, and then smoothing and thresholding it. Two initial tracking windows were used; one selected the large tank target near the center of the scene, and the other selected a smaller APC target at the left edge of the first

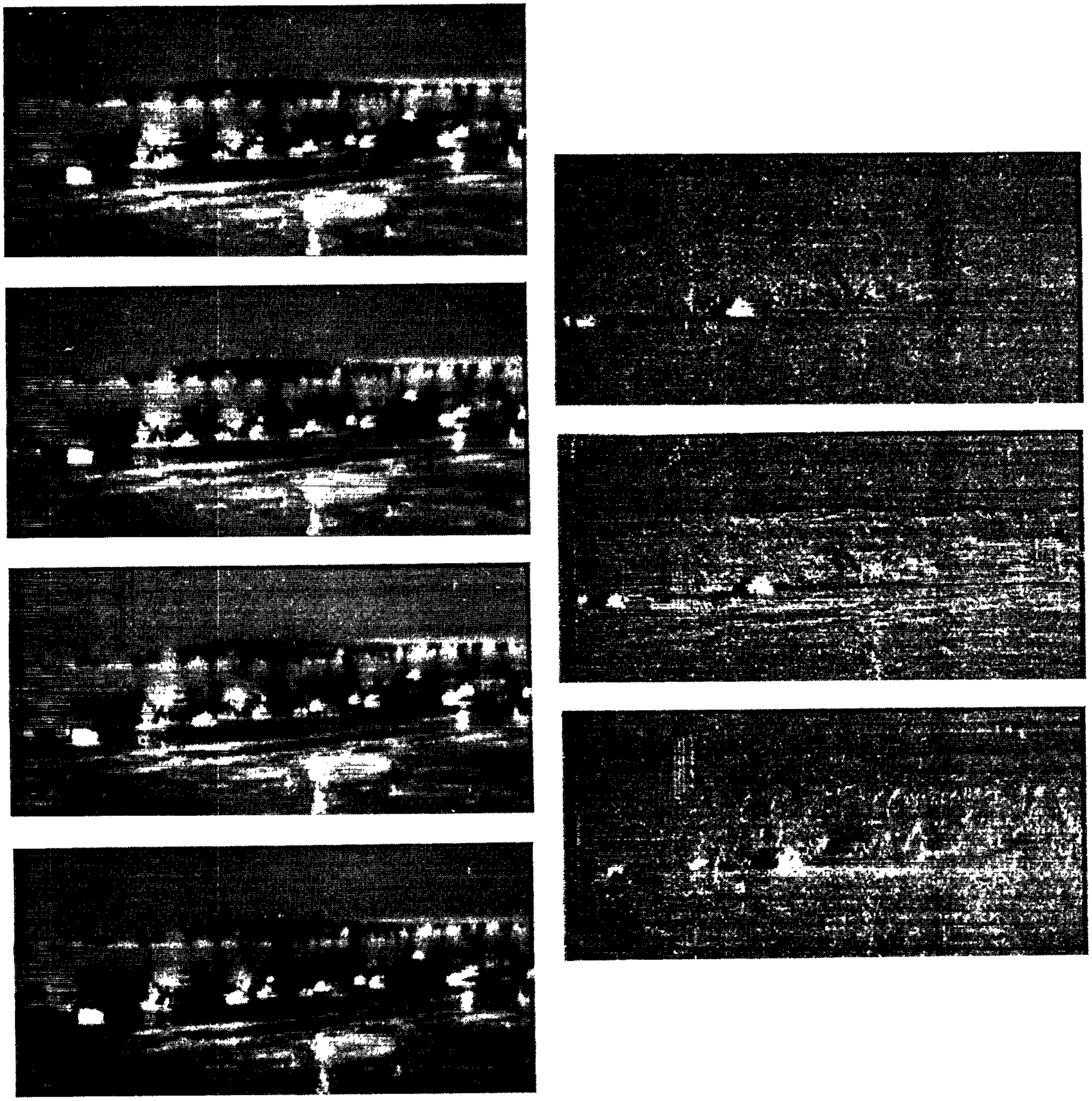


FIGURE 6 - Subtraction of FLIR imagery. The left four photographs show a series of FLIR images, and the three on the right give the differences between successive images. The uncancelled regions in the difference images corespond to moving targets and to errors present in the left part of the third FLIR image.

original image. The three photographs at the left of Fig. 7 show the tracking windows and the target centroid positions superimposed on the segmented target images, and the three photographs at the right show them superimposed on the first three original images.

The tank target in the center window is correctly tracked, but the APC target in the left one is not. The first original image given in Fig. 6 shows the APC entering the scene from the left. The second and the third original images show that a second APC has also entered the tracking window. In the fourth original image, the first APC has moved out of the tracking window to a position to the right of the burning-hulk target, and is partially obscured by small trees.

The first APC is not correctly tracked because of the temporary presence in the tracking window of the second APC. This produces an incorrect estimate for the centroid of the first APC and, therefore, an incorrect estimate of where the next target window should be located. Difficulties, such as these, that arise when the trajectories of multiple targets cross, or run close to one another, may be reduced by recording the past motion of all the known targets, and combining this information with the calculated centroid positions before obtaining the new estimates for the target positions. In addition, a priori information, such as the acceleration limits of a missile or an aircraft, can be used to reject short-term deviations in the calculated centroid position that represent unreasonable target trajectories. To achieve adequate performance in complex tracking situations involving effects such as multiple close or crossing targets, temporary target obscuration, the use of decoys, etc., we must employ intelligent tracking and image-analysis algorithms that make efficient use of all the available information.

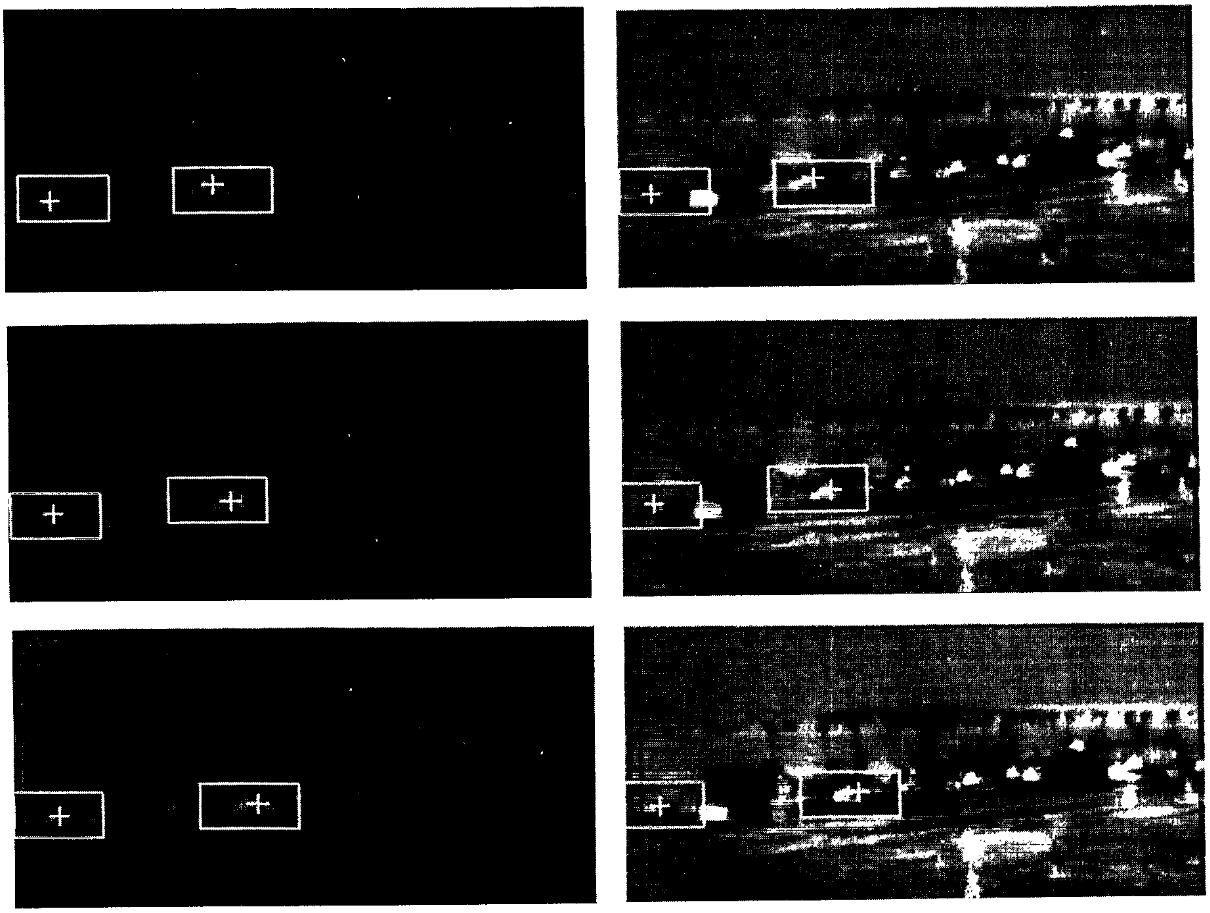


FIGURE 7 - Target tracking in FLIR imagery. The three photographs on the left show the tracking windows and the designated target positions superimposed on the thresholded difference images. The three on the right show them superimposed on the first three original FLIR images. The tank to the left of the center is correctly tracked, but the APCs entering from the left edge of the scene are not because more than one target is present in the tracking window.

## 3.3 Spatial and Intensity Misregistrations

Figure 8 shows the effect on the difference image of spatially shifting Image 3 of the APC sequence described in Sect. 3.1 relative to Image 4. The photograph in the upper left is the difference image we obtain for correct registration. The difference images in the second, third and fourth columns correspond to horizontal misregistrations of 1, 2 and 4 elements, whereas those in the second, third and fourth rows correspond to vertical misregistrations of 1, 2 and 4 elements. For example, we obtained the difference image in the third row and the third column by shifting Image 4 two elements to the right and two elements down relative to Image 3 before performing the subtraction.

Increasing the spatial misregistration reduces the cancellation of the background. Regions containing rapid changes in intensity, such as the borders between the trees and the sky, are more strongly affected than regions with a more uniform intensity distribution, such as the ground terrain in the lower part of the image. For this scene, the target becomes difficult to distinguish from uncancelled background areas when the misregistration exceeds approximately three elements. Relative to the target-image parameters given in Sect. 3.1, this misregistration corresponds to 10% of the target length, and to 23% of the average distance the target moves between images.

Figure 9 shows how intensity misregistration, produced by changing the gain of Image 4 relative to that of Image 3, affects the difference image. The first photograph gives the difference image we obtain when Image 4 is correctly registered in intensity with Image 3. The following photographs show the results of multiplying Image 4 by a gain

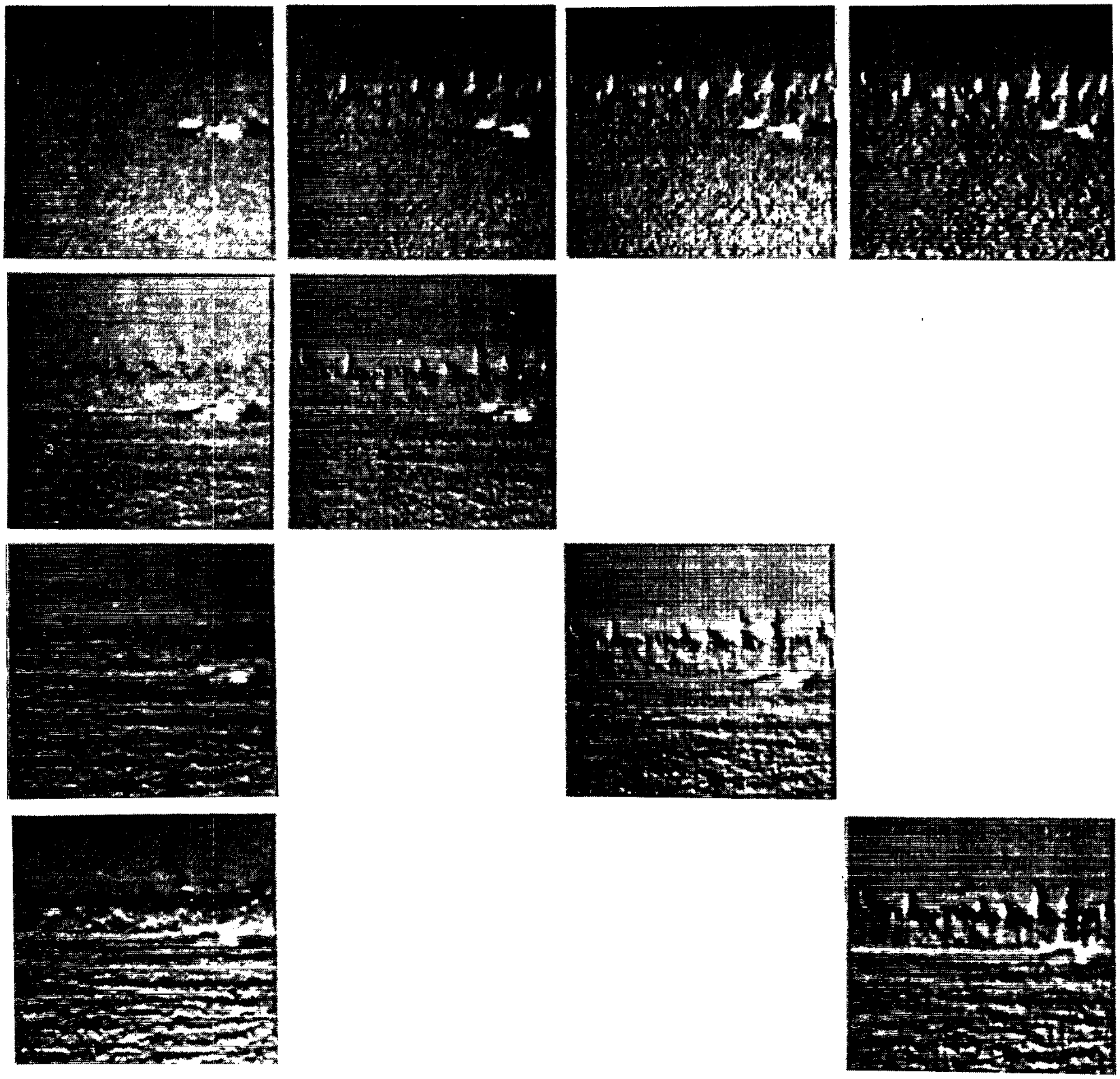


FIGURE 8 - Effect of spatial misregistration. The photograph in the upper left shows the difference obtained by subtracting Image 3 of the APC sequence from Image 4 of it. The two images are correctly registered in intensity and the background cancels. The photographs in Columns 2, 3 and 4 correspond to horizontal misregistrations of 1, 2 and 4 elements, whereas those in Rows 2, 3 and 4 represent vertical ones of 1, 2 and 4 elements. The target is difficult to distinguish from uncancelled background regions when the shift exceeds approximately two elements.

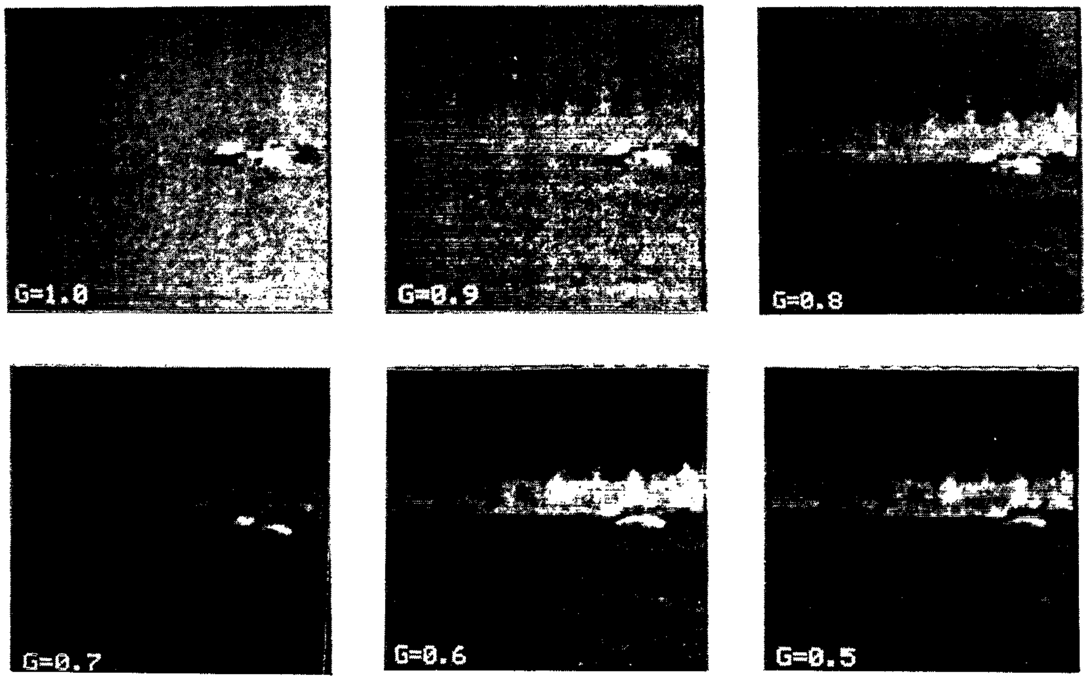
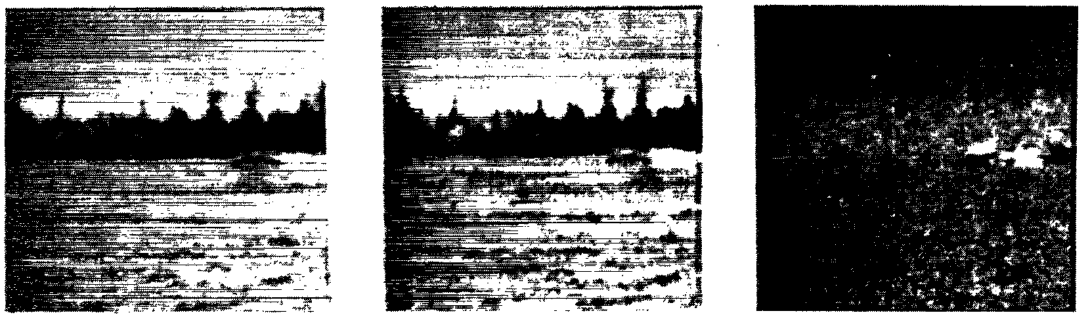


FIGURE 9 - The effect of intensity misregistration. The difference images were obtained by multiplying Image 4 of the APC sequence by the gain factor shown in the lower left of each photograph before subtracting Image 3 from it. A gain error of up to 30% can be tolerated before the target becomes difficult to distinguish from uncancelled regions of the background.

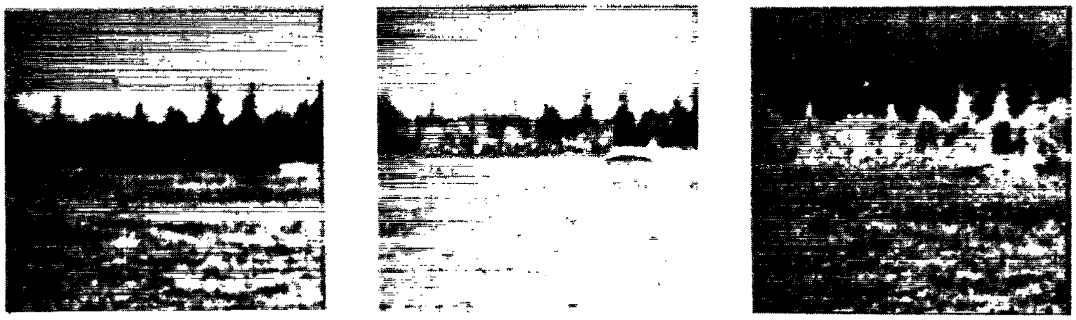
factor between 0.9 and 0.5 before subtracting Image 3. In the present example, a gain error of up to 30% can be tolerated before the target becomes difficult to distinguish from uncancelled background regions.

As noted in Sect. 2.4, one way to restore the intensity registration of two images which have been modified by different gray-scale transfer functions is to histogram equalize the regions common to both. Figure 10(a) shows the original APC Images 3 and 4 along with their difference. In Fig. 10(b), we have modified Image 3 with a square transfer function and Image 4 with a square-root one. (These transfer functions correspond to using photographic films with gammas of 2.0 and 0.5.) The background regions do not cancel in the difference image since they are no longer registered in intensity. Figure 10(c) shows the result of histogram equalizing the two modified images and then subtracting them. The two images are now visually similar and the background is correctly cancelled.

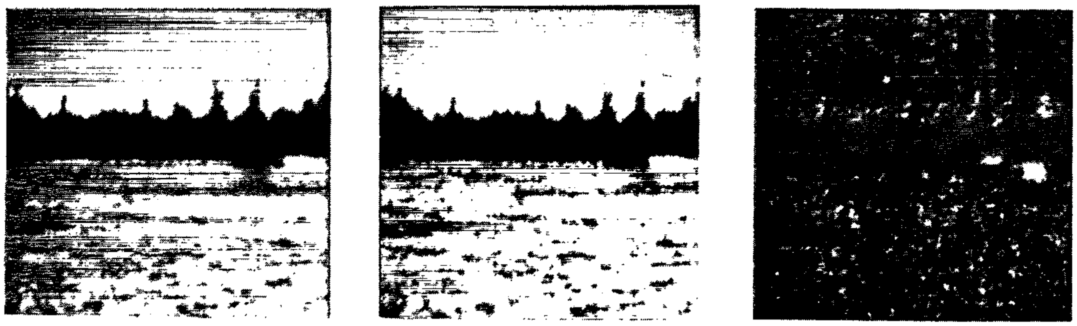
The amount of spatial and intensity misregistration that is acceptable depends on a variety of factors, such as the spatial-frequency and the edge content of the target and the background areas, how well the target contrasts in intensity with its immediate background, the distance the target moves between successive images, etc. The error tolerances considered acceptable for the present example will not necessarily apply for other image sequences, and only represent rough guidelines. However, it is evident that an adequate spatial registration will usually be more difficult to achieve in practice than an adequate intensity registration, except for extreme situations, such as when a high-intensity object moves into or out of the field of view of a television camera with a fast reacting automatic gain control.



(a) Correct intensity registration



(b) Intensity misregistration



(c) Intensity misregistration corrected by histogram equalization

FIGURE 10 - Correction of intensity misregistration by histogram equalization. Images 3 and 4 of the APC sequence are shown in the two left photographs in (a) along with their difference on the right. In (b), Image 3 has been modified with a square transfer function and Image 4 with a square-root one. The background does not cancel in the difference image because of the intensity misregistration. As shown in (c), the intensity registration can be restored without specific knowledge of the gray-scale nonlinearities by histogram equalizing the two images before subtracting them.

## 4.0 HARDWARE IMPLEMENTATION

Figure 11 shows a block diagram of a possible hardware implementation of a target-tracking system based on image subtraction. The images originate from a multispectral television-line-scan camera which can be positioned in azimuth and elevation to track a selected target. The camera outputs a and b represent the images of the scene in different spectral ranges. Two imaging modes are employed: intensity and spectral ratio. We obtain the intensity image by adding a and b, and the ratio image by subtracting their logarithms and exponentiating.

As described in Sect. 2.3, the optimum number of frame intervals between the subtracted images will not necessarily equal one; it will depend on the characteristics of the imaging system and the target. Figure 11 shows a variable frame delay which permits this interval to be selected to suit particular imaging conditions. After the subtraction, the absolute value of the difference image is calculated and the result is smoothed by using a spatial filter. We show a filter that produces a three- by three-element spatial average, i.e., each element is replaced by the average of a square region three elements wide and three rows high. However, the amount of smoothing could be adjustable to permit the target contrast to be optimized for particular target and background conditions.

The smoothed image is then passed to a series of centroid trackers that operate in parallel. Each has its own adjustable tracking window to permit a number of targets to be simultaneously tracked. A microprocessor reads the estimated target locations and positions the camera to track a selected target. One could also present a visual

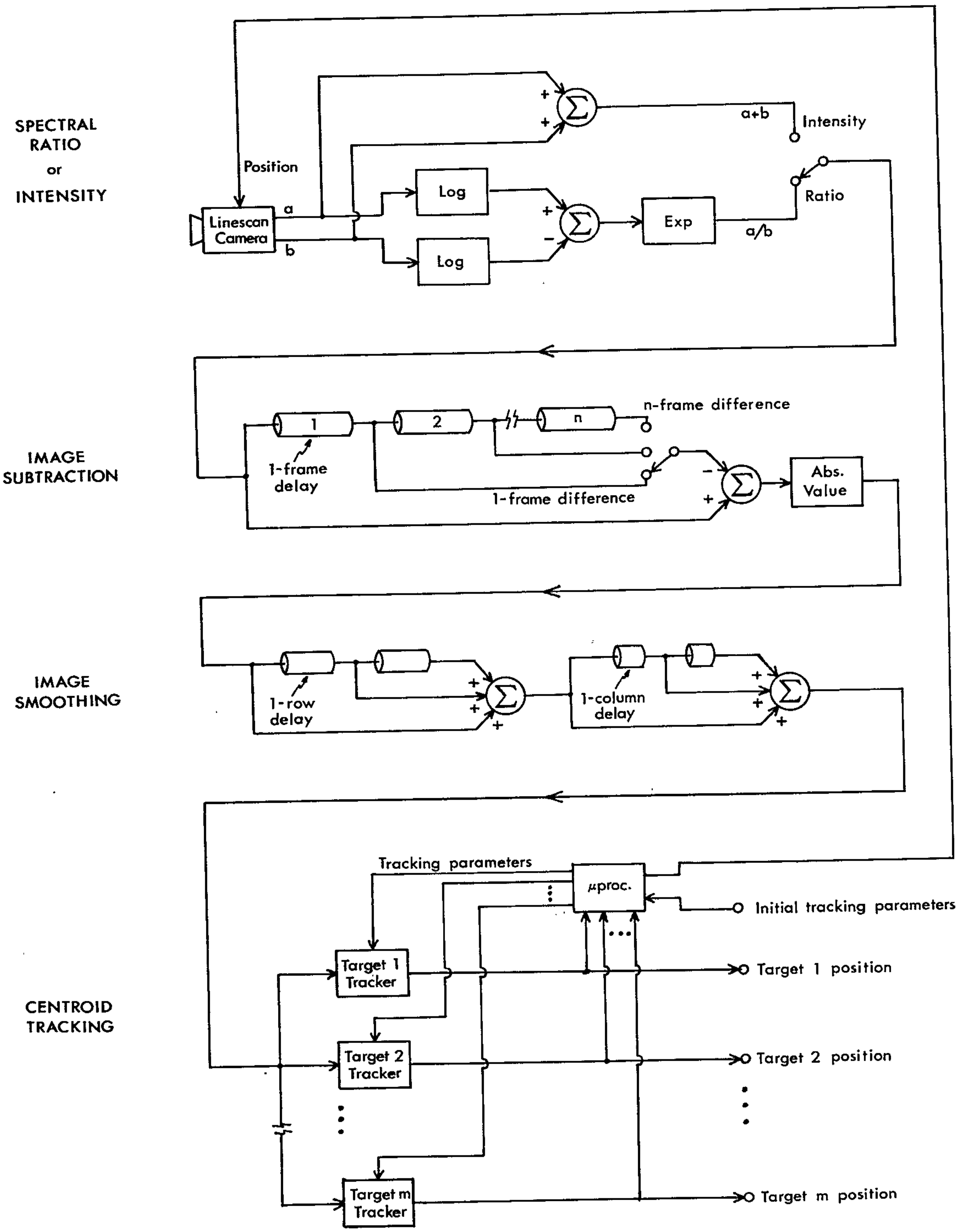


FIGURE 11 - Block diagram of a system based on image subtraction for tracking multiple targets in television imagery. The number of frames between the subtracted images is adjustable to permit optimizing a particular tracking situation. The microprocessor could be programmed with intelligent tracking algorithms, such as those that record and use the past history of the target trajectory, intensity or spectral ratio.

display of the original intensity, spectral-ratio or difference image, and designate each target centroid with a different graphic symbol. The microprocessor could be programmed with intelligent tracking algorithms that record and employ the time history of the trajectory, the amplitude, or the ratio value of each target as well as any available a priori information.

Figure 11 indicates the frame delay, the amount of smoothing and the initial tracking parameters as being manually selected. However, we could achieve a more autonomous system by programming the microprocessor with automatic target-acquisition algorithms (Refs. 13-16) and with routines that adaptively adjust the smoothing and the frame-delay parameters to optimize the tracking of a specified target.

All image-processing and target-tracking operations required here can be implemented in the analog or the digital domains. Analog hardware that operates in real time at television-video rates is available for calculating the logarithm, exponentiating, calculating the absolute value, summing, and centroid tracking, as are the analog delay lines required to perform the image subtraction and the spatial-filtering operations.

A digital implementation would require that the two camera outputs be digitized with a real-time video digitizer. The nonlinear logarithm, exponentiation and absolute-value operations could be achieved by using digital lookup tables, the smoothing by using a two-dimensional convolution filter, and the image addition and subtraction by using an image combiner. A relatively large amount of high-speed digital memory would be required to store and subtract 525-line television imagery digitized with a spatial resolution of 512 by 480 elements. Purely digital target centroid trackers that operate at real-time television rates are not commercially available, but they could be constructed with existing high-speed circuit components. Hardware capable of digitizing television video and processing it in real time has recently come on the market (Ref. 17), and an experimental image-processing system (a Comtal Vision One/20) with real-time two-dimensional convolution, image-combine and intensity-transform processors is in operation at DREV.

A fully digital implementation has advantages over an analog one, such as flexibility and more dynamic range, but it is more costly based on current technology. However, the rapid advances that are being made in digital signal-processing circuits and memories may change this situation within five years. In the shorter term, a hybrid implementation that uses both analog and digital devices may be more appropriate. For example, we may implement all processing up to and including the image subtraction with analog circuits, and then digitize the difference image and perform all subsequent operations digitally.

## 5.0 CONCLUSION

We described, and examined by computer simulation, how targets moving against a fixed background can be enhanced and tracked by subtracting the images of the scene obtained at different times. Adequate tracking performance requires that the two subtracted images be well registered spatially, but intensity registration is less critical. For example, a relative gain shift of 30% may be acceptable in some cases, but a spatial misregistration of only a few percent could seriously reduce the cancellation of the background.

In many cases, such as that of a television line-scan system, we obtain the best tracking performance by subtracting images separated by a number of frame intervals, rather than by just one. The optimum frame delay depends on the characteristics of the imaging system, such as its effective exposure time, frame rate, resolution, etc., as well as on those of the target, such as its dimensions, velocity, contrast, etc.

We proposed a television-line-scan target-tracking configuration that incorporates spectral ratioing to enhance the target contrast, image smoothing to reduce noise, and which has a variable frame delay to permit the subtraction to be optimized. It can track multiple targets and can be programmed with intelligent tracking algorithms. These algorithms may make use of a priori information and the past history of the target trajectory, intensity or spectral ratio. All the required processing operations could be implemented with existing hardware in the analog or the digital domains, or by using a combination of both.

## 6.0 ACKNOWLEDGEMENTS

The author wishes to thank Mr. Léandre Sévigny for providing the images from the Alabama Data Base.

## 7.0 REFERENCES

1. Munteanu, C., "Digital Area Correlation Tracking by Sequential Similarity Detection", DREV R-4097/77, November 1977, UNCLASSIFIED

2. Munteanu, C., "A Proposed Scene-Matching Tracking System", DREV M-2473/79, May 1979, UNCLASSIFIED

3. Psaltis, D. and Casasent, D., "Deformation Invariant Optical Processors Using Coordinate Transformations", Appl. Opt., Vol. 16, pp. 2288-2292, August 1977.

4. Casasent, D. and Psaltis, D., "New Optical Transforms for Pattern Recognition", Proc. IEEE, Vol. 65, pp. 77-84, January 1977.

5. Munteanu, C., "Comparison of Updating and Nonupdating Adaptation in Correlation Guidance", DREV M-2482/79, June 1979, UNCLASSIFIED

6. Munteanu, C., "Area Correlation Tracking with Automatic Compensation for Magnification and Rotation", DREV R-4140/79, July 1979, UNCLASSIFIED

7. Hummel, R. A., "Histogram Modification Techniques", Computer Graphics and Image Processing, Vol. 4, pp. 209-224, 1975.

8. Boulter, J.F., "Image Enhancement by Spectral Ratioing", DREV R-4170/80, June 1980, UNCLASSIFIED

9. Boulter, J.F., "Recognition of Reflection and Illumination Edges by the Human Visual System", Applied Optics, Vol. 19, pp. 2077-2079, 15 June 1980. Also published as DREV M-2488/79, November 1979, UNCLASSIFIED

10. Boulter, J. F., "Removal of Emissivity Contrast from Thermal Emission Images by Spectral Ratioing", DREV M-2525/80, December 1980, UNCLASSIFIED

11. Boulter, J.F., "Interactive Digital Image Restoration and Enhancement", Comp. Graphics and Image Proc., Vol. 9, pp. 301-312, December 1979. Also published as DREV R-4143/79, July 1979, UNCLASSIFIED

12. Goehrig, G. A., "The Night Vision and Electro-Optics Laboratory (NV & EOL) Tactical Target Data Base", Proc. SPIE, Vol. 219, pp. 2-6, February 1980.

13. Sévigny, L., "La reconnaissance de forme et l'acquisition d'objectif en infrarouge: Nouvel algorithme de détection", DREV R-4099/78, March 1978, UNCLASSIFIED

14. Sévigny, L., Extracteurs séquentiels pour l'acquisition de cibles sur images", DREV R-4153/79, August 1979, UNCLASSIFIED

15. Sevigny, L., "Evaluation of a Class of Segmenters for IR Imagery", DREV R-4172/80, May 1980, UNCLASSIFIED

16. Sevigny, L., "Segmentation Algorithms for Detection of Targets in IR Imagery", DREV R-4180/81, January 1981, UNCLASSIFIED

17. Andrews, H. C., "Semiconductor Advances Boost Digital Image Processing System Performance", Computer Design, September 1979.

CRDV R-4226/81 (NON CLASSIFIE)

Bureau - Recherche et Développement, MDN, canada.
CRDV, C.P. 8800, Courcelette, Qué. GOA 1R0

"La poursuite des cibles par la soustraction des images"
par J.F. Boulter

On peut rehausser et poursuivre des cibles qui se déplacent devant un arrière-plan stationnaire en soustrayant les images obtenues de la scène à des temps différents. On étudie, par la simulation numérique, les conditions nécessaires pour qu'un système de cette sorte donne une performance adéquate, et on montre deux exemples de la poursuite automatique de cibles. Le premier démontre la poursuite d'un véhicule blindé de transport de personnel (VBTP) enregistrée sur un film pris dans la lumière visible, et le deuxième, la poursuite de chars d'assaut et de VBTP dans une séquence d'images infrarouges. On décrit comment on peut construire un système fondé sur la soustraction des images et capable de poursuivre, en temps réel, des cibles multiples dans des images prises avec une caméra de télévision. (NC)

---

CRDV R-4226/81 (NON CLASSIFIE)

Bureau - Recherche et Développement, MDN, canada.
CRDV, C.P. 8800, Courcelette, Qué. GOA 1R0

"La poursuite des cibles par la soustraction des images"
par J.F. Boulter

On peut rehausser et poursuivre des cibles qui se déplacent devant un arrière-plan stationnaire en soustrayant les images obtenues de la scène à des temps différents. On étudie, par la simulation numérique, les conditions nécessaires pour qu'un système de cette sorte donne une performance adéquate, et on montre deux exemples de la poursuite automatique de cibles. Le premier démontre la poursuite d'un véhicule blindé de transport de personnel (VBTP) enregistrée sur un film pris dans la lumière visible, et le deuxième, la poursuite de chars d'assaut et de VBTP dans une séquence d'images infrarouges. On décrit comment on peut construire un système fondé sur la soustraction des images et capable de poursuivre, en temps réel, des cibles multiples dans des images prises avec une caméra de télévision. (NC)

---

CRDV R-4226/81 (NON CLASSIFIE)

Bureau - Recherche et Développement, MDN, canada.
CRDV, C.P. 8800, Courcelette, Qué. GOA 1R0

"La poursuite des cibles par la soustraction des images"
par J.F. Boulter

On peut rehausser et poursuivre des cibles qui se déplacent devant un arrière-plan stationnaire en soustrayant les images obtenues de la scène à des temps différents. On étudie, par la simulation numérique, les conditions nécessaires pour qu'un système de cette sorte donne une performance adéquate, et on montre deux exemples de la poursuite automatique de cibles. Le premier démontre la poursuite d'un véhicule blindé de transport de personnel (VBTP) enregistrée sur un film pris dans la lumière visible, et le deuxième, la poursuite de chars d'assaut et de VBTP dans une séquence d'images infrarouges. On décrit comment on peut construire un système fondé sur la soustraction des images et capable de poursuivre, en temps réel, des cibles multiples dans des images prises avec une caméra de télévision. (NC)

---

CRDV R-4226/81 (NON CLASSIFIE)

Bureau - Recherche et Développement, MDN, canada.
CRDV, C.P. 8800, Courcelette, Qué. GOA 1R0

"La poursuite des cibles par la soustraction des images"
par J.F. Boulter

On peut rehausser et poursuivre des cibles qui se déplacent devant un arrière-plan stationnaire en soustrayant les images obtenues de la scène à des temps différents. On étudie, par la simulation numérique, les conditions nécessaires pour qu'un système de cette sorte donne une performance adéquate, et on montre deux exemples de la poursuite automatique de cibles. Le premier démontre la poursuite d'un véhicule blindé de transport de personnel (VBTP) enregistrée sur un film pris dans la lumière visible, et le deuxième, la poursuite de chars d'assaut et de VBTP dans une séquence d'images infrarouges. On décrit comment on peut construire un système fondé sur la soustraction des images et capable de poursuivre, en temps réel, des cibles multiples dans des images prises avec une caméra de télévision. (NC)

DREV R-4226/81 (UNCLASSIFIED)

Research and Development Branch, DND, Canada.
DREV, P.O. Box 8800, Courcelette, Que. G0A 1R0

"Target tracking by image subtraction"
by J.F. Boulter

Targets moving against a stationary background can be enhanced and tracked by subtracting images of the scene obtained at different times. We examine by digital simulation the conditions required for such a system to give adequate tracking performance, and we give examples of the automatic tracking of an APC target in a visible-light cinefilm and of tank and APC targets in a sequence of FLIR images. We propose a possible hardware implementation of a system based on image subtraction for tracking multiple targets in real time in television-line-scan imagery. (U)

DREV R-4226/81 (UNCLASSIFIED)

Research and Development Branch, DND, Canada.
DREV, P.O. Box 8800, Courcelette, Que. G0A 1R0

"Target tracking by image subtraction"
by J.F. Boulter

Targets moving against a stationary background can be enhanced and tracked by subtracting images of the scene obtained at different times. We examine by digital simulation the conditions required for such a system to give adequate tracking performance, and we give examples of the automatic tracking of an APC target in a visible-light cinefilm and of tank and APC targets in a sequence of FLIR images. We propose a possible hardware implementation of a system based on image subtraction for tracking multiple targets in real time in television-line-scan imagery. (U)

DREV R-4226/81 (UNCLASSIFIED)

Research and Development Branch, DND, Canada.
DREV, P.O. Box 8800, Courcelette, Que. G0A 1R0

"Target tracking by image subtraction"
by J.F. Boulter

Targets moving against a stationary background can be enhanced and tracked by subtracting images of the scene obtained at different times. We examine by digital simulation the conditions required for such a system to give adequate tracking performance, and we give examples of the automatic tracking of an APC target in a visible-light cinefilm and of tank and APC targets in a sequence of FLIR images. We propose a possible hardware implementation of a system based on image subtraction for tracking multiple targets in real time in television-line-scan imagery. (U)

DREV R-4226/81 (UNCLASSIFIED)

Research and Development Branch, DND, Canada.
DREV, P.O. Box 8800, Courcelette, Que. G0A 1R0

"Target tracking by image subtraction"
by J.F. Boulter

Targets moving against a stationary background can be enhanced and tracked by subtracting images of the scene obtained at different times. We examine by digital simulation the conditions required for such a system to give adequate tracking performance, and we give examples of the automatic tracking of an APC target in a visible-light cinefilm and of tank and APC targets in a sequence of FLIR images. We propose a possible hardware implementation of a system based on image subtraction for tracking multiple targets in real time in television-line-scan imagery. (U)

35539