AD-A056 440   WHITE SANDS MISSILE RANGE N MEX   F/G 17/8

NOVEL CONCEPTS IN REAL-TIME OPTICAL TRACKING,(U)

JUN 78 A L GILBERT, M K GILES

UNCLASSIFIED   NL

| OF |

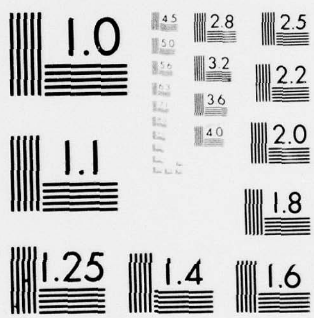
AD A056440

END

DATE FILMED

8-78

DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

NOVEL CONCEPTS IN REAL-TIME OPTICAL TRACKING

ALTON L. GILBERT, DR.
MICHAEL K. GILES, DR.
US Army White Sands Missile Range
White Sands Missile Range, New Mexico 88002

JUN 1978
15p.

INTRODUCTION: Optical tracking has been a mainstay of accurate metric range instrumentation since the first testing of modern rocketry. The accuracies that were possible from optical instruments exceeded those from other available instruments. Improvements in encoders, optical testing, modelling of the atmosphere, and optical design continuously improved the accuracies of optical instruments. The major drawback is the required film processing which delayed the delivery of boresight corrected optical data.

Recent changes in technology have created the potential for relieving part of the delay in data delivery. Automatic tracking methods using high-speed microprocessors, artificial intelligence, and pattern recognition techniques, together with special modifications to the existing optical systems, are now available to perform most of the film reading function in an on-line, real-time mode. These methods far exceed the conventional contrast, edge, and correlation trackers in sophistication and capability, since they are based upon an understanding of some definable properties of the image involving many parameters as compared to only a few.

THE INTELLIGENCE OF OBJECT IDENTIFICATION: Pattern recognition is a mathematical science based upon the separation of a parameter space into two or more regions, so that when the parameter is measured it may be classified as belonging to one of the appropriate regions. It follows that a vector parameter will give rise to a parameter space of dimensionality equal to the number of independent elements in the vector. Thus, for an N-vector, the required separation is a hyperplane in N space. If the parameter is a single element vector, an assignment can be made on the basis of a single threshold on the real numbers, and a tracker can be built that uses this decision rule. An example we call a contrast tracker uses a threshold on brightness for the assignment. A preprocessing algorithm may be placed before the decision. If we preprocess for the magnitude of change in intensity, the same thresholding rule will yield an edge tracker. These are amongst the simplest applications of pattern recognition to the object identification problem. Since these algorithms are easily confused, many spurious objects in the field of view (FOV) often meet the classification criteria.

A somewhat different approach that uses an array of points and measures the closeness of fit to a subsequently measured similar array, while choosing the best match as the correct location, is generally known as a correlation tracker. The decision is again based upon a single element parameter vector (the closeness of fit), but the preprocessing of data is much more elaborate. While this approach offers an improvement in confidence that the correct object has been located if the object description is known, it suffers from two principal problems. The first is that generally the object being tracked changes appearance continually while objects in the background may not. This requires an adaptive object description which may slowly converge to the acceptance of an undesired object as the desired one. The second problem is that this approach requires a very large amount of processing to do a good job, since the optimal linear process would be a convolution of the NxM object description array over the PxQ array of data points, and generally P>>N, Q>>M. A commonly used simplification is an additive (subtractive) algorithm that seeks the best fit of the desired array to the data, instead of the convolution. This approach necessarily results in loss of tracker performance. For these reasons, the correlation tracking method is generally limited to very restricted window tracking and fairly slow update rates.

Approaches to real-time optical tracking have generally been limited to these approaches for the following principal reasons. The first and most important has been the magnitude of the real-time processing requirement for the more elaborate approaches, which have exceeded computational resources generally available. The second has been a lack of image understanding that would allow the formulation of more reliable, yet simple, approaches. Substantial progress has recently been made in the former, and there are many encouraging new developments in the latter.

A variety of methods of image data processing have become known over the past decade. Applications-oriented research at the US Army White Sands Missile Range (WSMR) has lead recently to a system of reasonably high sophistication using concepts developed in-house and through sponsored research to solve complex identification and tracking problems. WSMR has concentrated on objects in the visible spectrum and in real-time. Many other systems, not necessarily real-time, have been developed for applications in medicine, meteorology, and space research.

Many of the newer methods involve the use of many elements in the parameter vector to glean more information from the data. In applying pattern recognition methods to the object identification problem, the engineer is trying to minimize the amount of data he must handle and maximize his confidence that he made the correct decision. Any linear process will preserve the quantity of data (260,000 points for a 512x512 image, possibly 8 bits per point) which is obviously not desirable if much processing is required to make a decision. The engineer is forced to require a high degree of parallel processing on linear processes, and to perform nonlinear operations to reduce the data quantity prior to determining the values of the parameter elements used in the decision rule. Ideally, the dimensionality of the decision space should be kept reasonably small to allow decisions to be made in real-time or in near real-time.

Some of the preprocessing methods currently in use are:

Filtering: Filtering operations generally involve the convolution of a point spread function array with the image to achieve some desired objective with the image. Examples include removing spatially invariant degradations due to the optics of the atmosphere, boosting the high frequency content of the image to enhance edges, removing noise in the image, making the image more pleasing to the eye, and other such operations. Generally those operations which remove degradations are called estimation and those which emphasize certain spatial frequencies or certain aspects of the image are called enhancement. It must be noted that enhancement is an intentionally introduced distortion to produce some desired effect.

Transforms: Operations that map the image into a new domain are called transforms. The elements in the new domain are a measure of some property of the original image. The most common example is the discrete Fourier transform (DFT), especially in the fast algorithms (FFT). The DFT identifies the spatial frequency content of the image, which allows further processing based upon these components. A class of binary Fourier (BIFORE) transforms has been developed over the past decade which are similar to the DFT but are more suited to computer applications. These might be called lesser transforms since they do not represent the information in the image as completely. Because they are much more efficiently run on a computer than the DFT, they have important applications in image transformations. Among these lesser transforms are the now popular Hadamard transform based upon Walsh functions, and the less known but simple Haar transform. These transforms can be useful for identifying features of interest in the image. It is necessary, of course, to apply all of these transforms in a two-dimensional algorithm to process the two-dimensional images.

Point Processing: In point processing, individual points in the image are assigned new values based upon some assignment rule. This may take a variety of forms with a large variation in apparent results. One point processing algorithm averages the corresponding point of several frames or sequential images to produce a weighted composite and remove transient degradations. Another assigns all values above a given threshold to 1 and all values below to 0. This is known as thresholding. A variation on thresholding is to assign predetermined gray levels to 1 even though these may not be in a continuous range. Still another algorithm, known as contrast stretching, assigns all values below some intensity I_0 to 0; all values above another intensity I_1 to the maximum gray level, say 256; and stretches the intermediate values to occupy the full range. Generally, point processing methods are nonlinear, yielding fewer bits in the output than in the data array.

The next step in the process is to identify the values of the elements in the parameter vector. These elements may include such things as size, orientation, number of corners, brightness, etc. When joined in a single parameter vector, they describe all we think we need to know to adequately describe the object for purposes of identification.

A REAL-TIME TRACKING SYSTEM: By using the above concepts together with high-speed microprocessors and special optics, a real-time tracking system may be devised that demonstrates a substantial advantage over the contrast, edge, and correlation trackers currently on the market. The greatest challenge is that of doing "intelligent" processing of video data at the extremely high data rates of standard TV.

The development of an intelligent real-time video (RTV) tracking system has been accomplished through the cooperative efforts of research and development personnel at WSMR, New Mexico State University (NMSU), and the Optical Sciences Center of the University of Arizona. The prototype RTV processor is being assembled at NMSU, the automatic zoom lens and image rotator at the University of Arizona, and the system interfaces at WSMR. The system components will be integrated and the system deployed early in fiscal year 1979 as an add-on modification to the Contraves Model F cinetheodolite at WSMR.

Figure 1 is a block diagram of the RTV tracking system which shows the RTV processor as the central element. The RTV processor

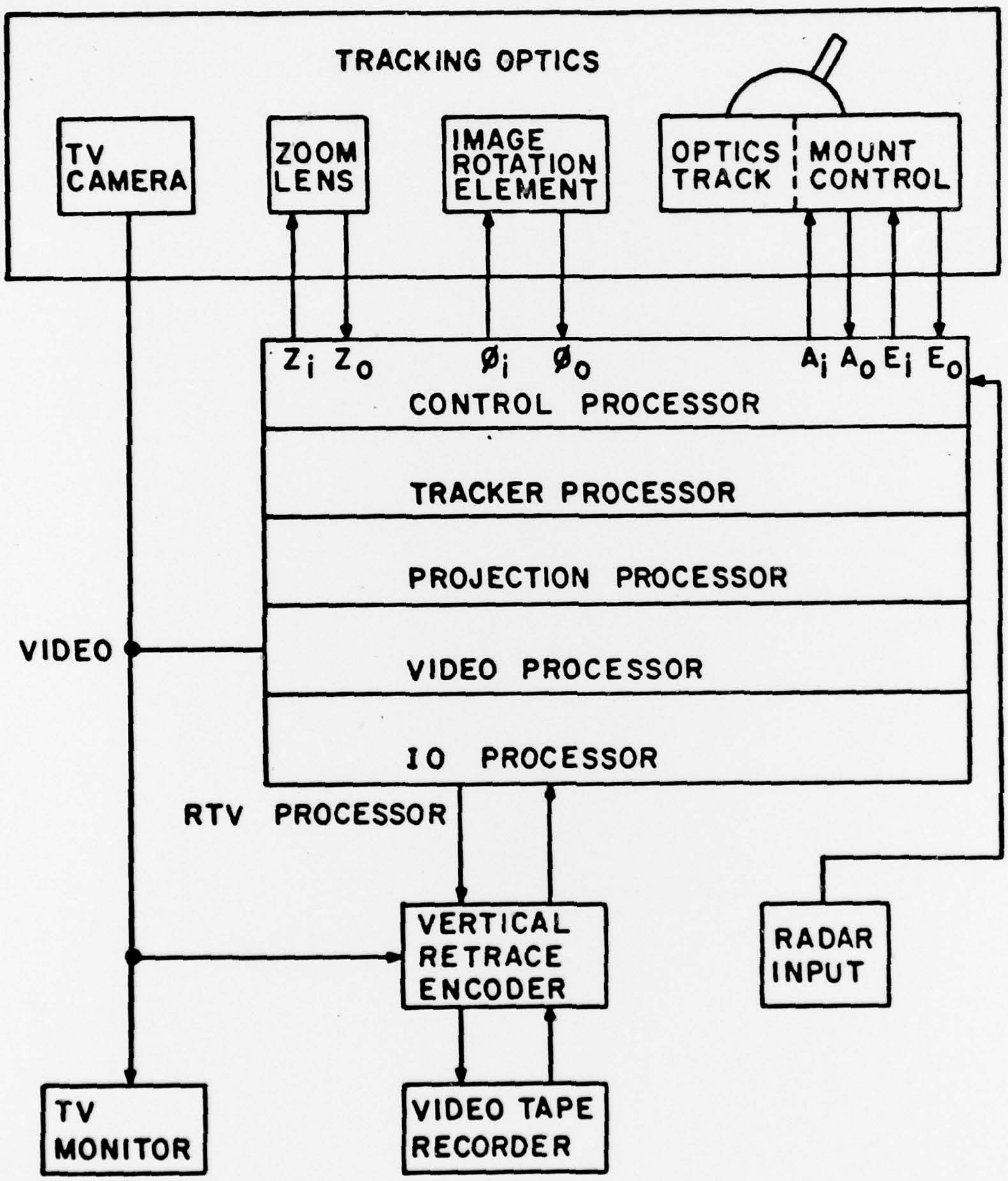


FIGURE 1. RTV TRACKING SYSTEM

receives standard composite video from a television camera, locates the target image, and provides control signals which drive the zoom and image rotation elements and point the Contraves tracking optics at the target. It also provides boresight correction signals and target attitude angles which are recorded into the vertical retrace period of the video tape used to record the tracking sequence.

The RTV processor consists of a distributive array of five processors, shown in Figure 1. The video processor synchronizes and digitizes the video signal from the TV camera, performs a statistical analysis of the digitized image, and separates the target images from the background. The projection processor accumulates binary projections of the target and plume images and establishes the structural parameters which locate and describe the shape of the target and plume images. The tracker processor establishes a structural confidence in the data and implements an intelligent tracking strategy. The control processor utilizes the structural confidence to combine current target coordinates with previous target coordinates to orient the optics toward the next expected target position, forming a fully automatic system. The input/output (I/O) processor provides a user interface to the tracking processors and is responsible for recording the tracking data with a video tape recorder.

A Research Oriented Processor Configuration: Four of the five distributive processors (excluding the I/O processor) which comprise the RTV processor are high-speed microprogrammable processors, each of which requires a stored microprogram to control its designated tracking function. To provide a powerful tool for future research in video tracking algorithms and to facilitate operational testing of the RTV system, the control store of each processor is realized with a read/write random access memory.

These four distributive processors are being built with a standard microprogrammable processor architecture to simplify the development and maintenance of the RTV tracking system. This standard architecture has been designed, built, and tested at NMSU. Based on the new Texas Instruments (TI) 74S481 Schottky processor chip, it provides a microinstruction cycle time of under 200 nanoseconds with sufficient computational power to implement the required RTV tracking algorithms. The standard architecture requires several LSI chips which may be partitioned into control and processing sections. Overlapping the execution of one microinstruction with the fetch of the next one allows the processor to achieve a minimum microinstruction cycle time equal to the larger of either the fetch time or the execution time, significantly increasing the speed of the processor.

The four high-speed processors included in the RTV tracking loop are described in some detail in the following paragraphs. In each case, the processor is built around the standard architecture outlined above. Some specialized hardware is added to the standard configuration in each case to accommodate the specific functions of the individual processors.

The Video Processor: The video processor decomposes each video field into target, plume, and background pixels at the standard video rate of 60 fields per second. As the TV camera scans the scene, the video intensity is digitized at m equally spaced points across each horizontal scan line. A resolution of m = 512 pixels per line results in a pixel rate of 96 nanoseconds per pixel. Within 96 nanoseconds, a pixel intensity is digitized and quantized into 8 bits (256 gray levels), counted into one of six 256-level histogram memories, and then converted by a decision memory to a 2-bit code indicating its classification (target, plume, or background). The 2-bit classification code is passed to the projection processor via the target data (TD) and projection data (PD) lines. TD is high for target points; PD is high for plume points.

The basic assumption of the image decomposition method is that the target image has some video intensities not contained in the immediate background. A tracking window is placed about the target image, as shown in Figure 2, to sample the background intensities immediately adjacent to the target image. The window frame is partitioned into two regions, B and P. Region B is used to provide a sample of the background intensities, and region P is used to sample the plume intensities when a plume is present. Using the sampled intensities, a very simple decision rule is used to classify the pixels in region T as follows:

- Background points--All pixels in region T with intensities found in region B are classified as background points.

- Plume points--All pixels in region T with intensities found in region P, but not found in region B, are classified as plume points.

- Target points--All pixels in region T with intensities not found in either region B or P are classified as target points.

A tracking window placed about the target image provides a method for sampling the pixel features associated with the target and background images. The background sample should be taken relatively close to the target image, and it must be of sufficient size to accurately characterize the background intensity distribution in the vicinity of the target. The tracking window also serves as a bandpass filter

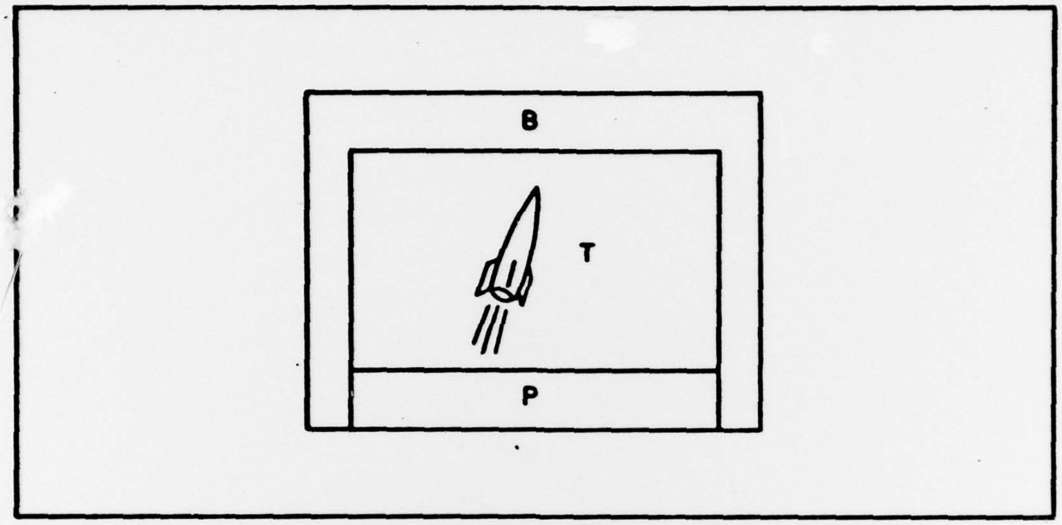


FIGURE 2. TRACKING WINDOW

by restricting the target search region to the immediate vicinity of the target. Although one tracking window is satisfactory for tracking missile targets with plumes, two windows provide additional reliability and flexibility for independently tracking a target and plume, or two targets. Having two independent windows allows each to be optimally configured and provides reliable tracking when either window can track.

The Projection Processor: The projection processor consists of a projection accumulation memory (PAM) and a standard processor which are designed to form projections of simultaneous target and plume windows and to compute structural parameters from the projections. The pixel data from each tracking window enters the PAM in real-time as a synchronized serial stream on lines TD and PD. As the classified pixel data is received, the PAM accumulates the projection data while the processor monitors the y-projections, accumulates the total number of target and plume points, and determines the midpoints used to split the x-projections. Each x-projection is split to allow the computation of target and plume attitude angles based on the locations of the median centers of the x- and y-projections of the top half and bottom half of the target and plume images.

During the vertical retrace interval, the projection processor divides each projection into eight segments of equal mass using a simple algorithm to sequentially address each line of the projection and multiply the number of pixels in the line by eight. If the result exceeds the total number of pixels in the projection, a flag is sent to the PAM forcing the next line to be placed at the beginning of the next 1/8 segment of the projection. If the result is less than the total number of pixels in the projection, additional lines of pixels are accumulated until the line containing the 1/8 percentile point is located.

The 1/8 percentile points for each of the six projections are computed within 410 μsec of the vertical retrace period and then passed to the communication memory along with the total number of target and plume points. These parameters constitute the structural parameters used by the tracker processor to define an intelligent tracking strategy. Figure 3 illustrates the accumulation of the projections and the computation of the percentile points and, for simplicity, omits the splitting of the x-projection.

Tracker Processor: The tracker processor receives the structural parameters from the projection processor, locates and characterizes the structure of the target and plume images, and decides on a tracking strategy to maintain track. It then outputs control signals to place the window frames in the video processor and outputs target location and orientation data to the control processor along with a confidence in the measured data. Since it operates on the projection data from field n while the projections for the next field (n+1) are being accumulated, the tracker processor is always one field behind the video and projection processors. The tracker and control processors must both finish their calculations before the vertical retrace interval begins for field n+1. This constraint requires the tracker processor to output its data to the control processor within 7 milliseconds after it receives the projection data.

Since the tracker processor is the only processor that communicates with all of the other three processors, each of which has its own coordinate system, the tracker processor must interpret the input data intelligently and then output the appropriate data to the video and control processors in their respective coordinate systems. The inputs are positive 16-bit integers defined for a coordinate system whose origin is the first pixel scanned inside the appropriate tracking window. The outputs to the video processor are 9-bit positive integers defined for a coordinate system whose origin is the first pixel scanned within the FOV. The 16-bit outputs to the control processor are defined for a coordinate system whose origin is the boresight.

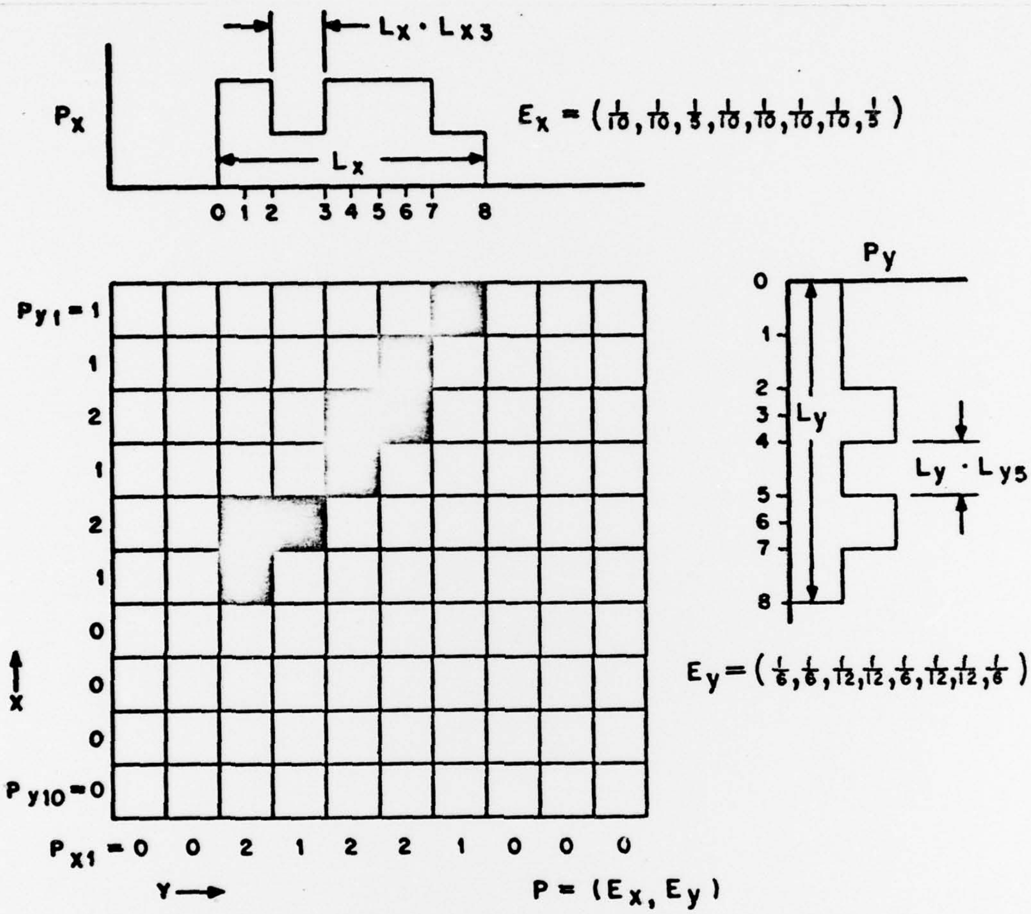


FIGURE 3

PROJECTIONS AND PERCENTILE POINTS

An overall view of the functions of the tracker processor is given in Figure 4. It has two modes of operation, the initial acquisition mode and the autotrack mode. The initial acquisition mode is used when the RTV system is trying to lock onto the target of interest. During this mode, the video processor does little or no learning on the target and plume intensities. The tracker processor will not instruct the control processor to begin predicting the target location until it is sure of the existence of at least the plume within the plume window.

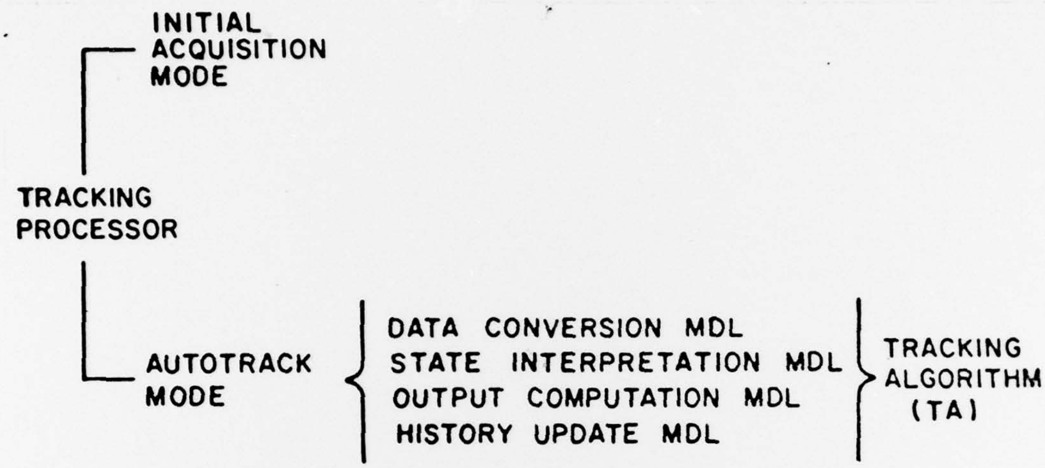


FIGURE 4. TRACKER PROCESSOR FUNCTIONS

When the plume image moves into an appropriate region of the FOV, the tracker processor will notify both the video processor and the control processor with a flag indicating that it is now ready to shift into the autotrack mode.

The autotrack algorithm is divided into the four main modules shown in Figure 4. The data conversion module transforms the projection input data into physical variables; such as, target and plume size, position, and shape. These variables are then combined with previous target activity data from the history update module to obtain additional variables; such as, the changes in target and plume position and size. All of these variables are compared with preassigned reference constants to obtain a set of binary inputs which are used directly by the state interpretation module to define the current tracking situation and produce an optimum tracking strategy. The strategy is implemented by the output computation module in the form of control signals to the video and control processors.

The Control Processor: The function of the control processor is to generate the four control signals that drive the real-time video tracker; i.e., the tracker azimuth A_i and elevation E_i which are sent to the RTV-Contraves system interface and the optics rotation ϕ_i and zoom Z_i which are sent to the RTV-zoom/rotation interface (Figure 1). In addition, the control processor outputs the following tracking data to the I/O processor after each field so they can be recorded in the vertical retrace period of the video tape: field count, tracker status, time, x-displacement from boresight, y-displacement from boresight, tangent of the target orientation angle from vertical boresight, target azimuth, target elevation, tracker azimuth, tracker elevation, image rotation angle, and zoom ratio.

The tracking optics feeds the target image to the video processor portion of the RTV processor (Figure 1) which establishes the target coordinates with respect to the optics boresight. The control processor combines current target coordinates with previous target coordinates to point the optics toward the next expected target position. The predicted control equations are based on the combination of linear and quadratic optical estimates taken from a five-deep history stack. Since the input data is derived from field (K-1), and the estimates are being computed during field K, the control estimates must predict ahead two time increments to provide control signals which will place the boresight at the correct position during frame K+1.

COMPUTER SIMULATION OF THE REAL-TIME VIDEO TRACKER. A computer simulation of the RTV tracking system, incorporating the algorithms used in the control stores of the four distributive processors, has been developed and implemented on the PDP 11/35 system at WSMR. The purpose of this simulation is to provide a method for testing new design concepts and evaluating the RTV tracking system under realistic tracking conditions. The simulation model includes dynamic models for the target trajectory and the Contraves Model F cinetheodolite tracking system, in addition to the RTV processor algorithms, for simulating the complete tracking system. The recent development of an image processing laboratory at WSMR has enabled research personnel to digitize sequential video fields of typical tracking imagery. These fields of digitized video are now being used in the RTV simulation and in the development of improved image segmentation and structural analysis algorithms.

The RTV simulation is being used as a research tool at WSMR. It is especially effective in evaluating the RTV system performance and in identifying and seeking solutions to real-time tracking problems before the RTV tracking system is deployed. With the added capability of using digitized video from a variety of tracking sequences as inputs to the video processor, the simulation can now test the system performance under a variety of tracking conditions, thus allowing thorough evaluation and possible refinement of the tracking and processing algorithms and the state transitions of the tracker processor.

CONCLUSION: RTV tracking is not new, but recent developments have added new capabilities that enhance the advantages of these systems. Video tracking offers some distinct advantages over electronic tracking (such as ECM immunity), but suffers from some disadvantages as well (such as restrictions in visibility). Several other aspects of system development for RTV tracking are discussed in the papers and reports listed in the bibliography.

A continuing research need exists for better understanding of imagery. A human incorporates many elements into the parameter vector that he uses to identify an object. The difficulty of understanding the human visual process has caused rather slow progress in teaching computers to "see." We know that the human uses such things as texture, orientation, color, size, shading, shape, context, etc. to identify objects. Parameter vectors which incorporate these elements are difficult to quantify. It is not necessary, however, to require the computer to see the same things a human does. It is difficult to visualize elements of parameter vectors that do not have a physical meaning to a human, but which may be useful for computer recognition processes. Much work remains to be done to produce a highly sophisticated sight process in a computer.

The concepts described in this paper have, however, been tested and will result in a prototype system deployed in 1979. Through a process of simulation and breadboard verification, WSMR has determined that such a system is well within the current capabilities of technology. A great deal of national (and some international) attention has been focused on this project because of the unique applications of pattern recognition in a tracking situation.

CHRONOLOGICAL BIBLIOGRAPHY: The papers and reports listed below have been produced during the course of this project.

1. Flachs, G. M., "A New Target Tracking Algorithm for the Automatic Programmable Film Reader," WSMR Technical Report, STEWS-ID-74-3, October 1974.

2. Flachs, G. M. and A. L. Gilbert, "Automatic Boresight Correction in Optical Tracking," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, June 1975.

3. Flachs, G. M., "Real-Time Tracking Algorithm," Final Report for Grant DAHCOA-15-G-0038, Army Research Office, October 1975.

4. U, Yee Hsun and G. M. Flachs, "Structural Feature Extraction by Projections," Proceedings of the IEEE Region V Conference, 1976.

5. Black, R. J., J. T. Whitney, G. M. Flachs, and A. L. Gilbert, "A Pre-Prototype Real-Time Video Tracking System," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, May 1976.

6. Flachs, G. M., W. E. Thompson, Yee Hsun U, and A. L. Gilbert, "A Real-Time Structural Tracking Algorithm," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, May 1976.

7. Thompson, W. E. and G. M. Flachs, "A Structure and Dynamic Mathematical Model of a Real-Time Video Tracking System," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, May 1976.

8. Flachs, G. M. and J. A. Vilela, "A Structural Model for Polygon Patterns," Proceedings of the IEEE 1976 National Aerospace and Electronics Conference, May 1976.

9. Gilbert, A. L. and G. M. Flachs, "A New Concept in Optical Tracking using Pattern Recognition," International Symposium on Information Theory, June 1976.

10. Vilela, J. A., "Application of Automata Theory to Picture Modeling," PhD Dissertation, New Mexico State University, 1976.

11. Whitney, J., "A Pre-Prototype Real-Time Tracking Filter," MS Thesis, New Mexico State University, 1976.

12. Flachs, G. M., W. E. Thompson, R. J. Black, J. M. Taylor, W. Cannon, R. B. Rogers, and Yee Hsun U, "A Pre-Prototype Real-Time Video Tracking System," Final Report for Contract DAAD07-76-C-0024, WSMR, January 1977.

13. Cannon, W., "A Microprocessor System Development Interface," MS Thesis, New Mexico State University, 1977.

14. Rogers, R. B., "Specification for a Two-Pass Symbolic Assembler Generating Absolute Microcode for Arbitrary Microprogram Control Store Architectures," MS Thesis, New Mexico State University, 1977.

15. Perez, P. I. and G. M. Flachs, "A Microprogrammable Processor Architecture for System Design," Technical Report, New Mexico State University, 1977.

16. Ritchell, J., "A General Purpose Microprogrammable Computer Architecture," MS Thesis, New Mexico State University, 1977.

17. Flachs, G. M., W. E. Thompson, J. M. Taylor, W. Cannon, R. B. Rogers, and Yee Hsun U, "An Automatic Video Tracking System," Proceedings of the 1977 National Aerospace and Electronics Conference, May 1977.

18. Rogers, R. B. and G. M. Flachs, "Mathematical Modeling and Simulation in a Programmed Design Methodology," Proceedings of the First International Conference on Mathematical Modeling, August 1977.

19. Fukunaga, K., A. L. Gilbert, M. K. Giles, O. R. Mitchell, R. D. Short, and J. M. Taylor, "Segmentation and Structure Analysis for Real-Time Video Target Tracking," WSMR Technical Report, STEWS-ID-77-1, October 1977.

20. Flachs, G. M., P. I. Perez, R. B. Rogers, S. J. Szymanski, J. M. Taylor, and Yee Hsun U, "A Real-Time Video Tracking System," Annual Report for Contract DAAD07-77-C-0046, WSMR, January 1978.

21. Gilbert, A. L., "Pattern Recognition and Real-Time Boresight Correction--A Tutorial," SPIE Proceedings, Vol. 134, Photo and Electro-Optics in Range Instrumentation, March 1978.

22. Giles, M. K. and A. L. Gilbert, "Concepts in Real-Time Video Tracking," WSMR Technical Report, STEWS-ID-78-1, May 1978.

23. U, Yee Hsun, "Applications of Projection Concepts to Image Processing," PhD Dissertation, New Mexico State University, 1978.

24. Gilbert, A. L. and M. K. Giles, "A Real-Time Video Tracking System," Proceedings of the 24th ISA International Instrumentation Symposium, May 1978.