A FLEXIBLE MACHINE VISION GUIDANCE SYSTEM
FOR 3-DIMENSIONAL CONTROL TASKS

H.F.L. Pinkney and C.I. Perratt
National Aeronautical Establishment
National Research Council of Canada
Ottawa, Ontario. CANADA K1A 0R6

ABSTRACT

This paper describes a machine vision guidance system capable of providing real-time control information at the 30 Hz framing rate of video cameras. From the observation that objects in machine control tasks are human made, the question of interest is not "what is it?" but rather "where is it?" In response to the latter, a generic machine vision system was developed based on known discriminable features (targets) and photogrammetric solution algorithms.

This paper describes the integration and functional attributes of the adaptive measurement/analysis/control processes of the system as they relate to real-time machine vision control applications. Its industrial development as a flexible guidance system for industrial robotics and as an experimental Space Vision System for testing during the next Canadian Astronaut shuttle flight is highlighted.

INTRODUCTION

Based on the functional concept that the basic information necessary to meet guidance control needs is the relative position, orientation and corresponding rate of change of two coordinate systems, a flexible machine vision system architecture has been developed, capable of real-time operation using standard North American Closed Circuit Television (CCTV) systems. This development was conceptualized and initiated during the Canadian development of the Shuttle Remote Manipulator System (SRMS), Figure 1, to provide, if necessary, enhanced CCTV displayed guidance information for SRMS manipulation tasks.

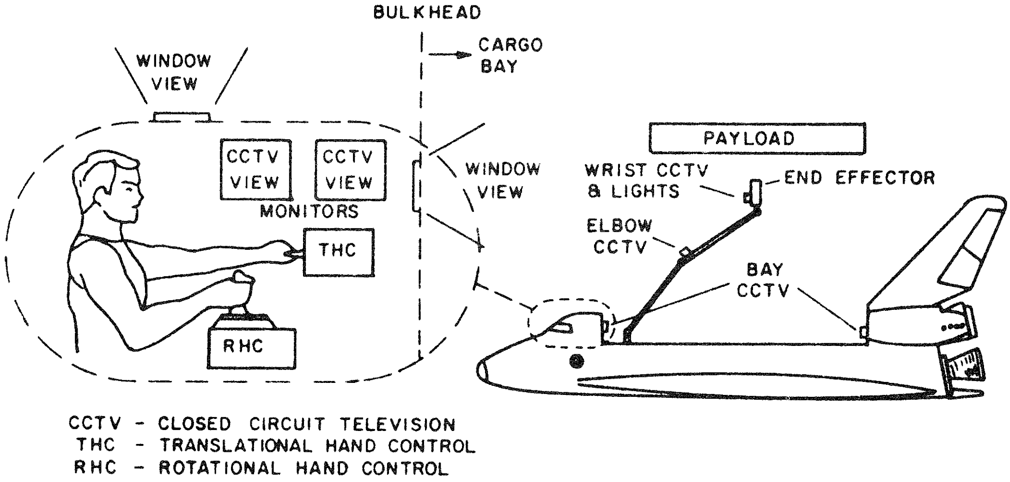


FIG. 1: POTENTIAL SHUTTLE APPLICATION

Given the Shuttle CCTV system, the concept used the method developed during a study of motor vehicle redirection by highway barriers as described by Van Wijk and Pinkney (1972). In this method the use of "specific" targets having accurate known locations with respect to the vehicle sprung mass coordinate system permitted the use of a single camera photogrammetric solution, for each frame of the high speed movie film, to accurately derive the position and orientation of the sprung mass coordinate system with respect to the camera coordinate system.

Given this concept of apriori defined discriminable target features it was then possible to avoid the complexities, uncertainties and computational power required for the general problem of image analysis of objects under ambient lighting conditions and to proceed with the formulation and the technical evaluation of a real-time video photogrammetry system for 3-dimensional control as described by Pinkney, Perratt et al (1976). Using the building block technique of systems engineering, it was then possible to further delineate the functional processes as modules within a "structured form" and develop a laboratory prototype as described by Pinkney (1978).

Essential to the development of this prototype was a hardware implementation of a Video Sampling Processor (VSP) with four independent computer controlled window aperture processors. These processors provided the basic capability for discriminating, tracking and deriving target centroid related data on-line with the video raster scan. By means of a modified ("tailored" for control) form of a single camera photogrammetric solution studied by Kratky (1979), it was then possible to derive, with minimum delay, the position and orientation of the object target coordinate system with respect to the camera coordinate system at 30 Hz for real-time control applications.

Given the laboratory prototype with process modules (1) to (4) in Figure 2 integrated by a flexible real-time system software architecture, the system concept was further developed for potential space and industrial applications via a NRCC Program for Industry/Laboratory Projects (PILP) contract with Leigh Instruments (Ottawa) for a microprocessor based application/demonstration system. In addition to a bit slice processor based VSP, the system included a video distribution module (2a) to accept up to two camera inputs and to provide VSP enhanced scene and synthetic display video to separate monitors, as well as an operator control panel (7) with essential VSP and system control functions. Since in 1980 a hardware floating point capability was not available for microprocessors, the microprocessor based system was interfaced to a laboratory minicomputer via a DMA link. With the flexibility and processing power of the minicomputer it was then possible to continue the development and characterization of the module processes themselves for integration as a Real-time Photogrammetry System (RPS) for guidance/control tasks.

As now developed, this NRCC RPS with its patented VSP provides the basis for a flexible machine vision system. An experimental flight (prototype) version, called a Space Vision System (SVS), is now being built for testing on the next Canadian Astronaut shuttle flight and an industrial version is also being developed for industrial robotics.

With these application systems now under development, the objective of this paper is to describe the formulation of the integrated system concept with its basic functional module processes for real-time control as implemented using a flexible software architecture.

FORMULATION OF THE INTEGRATED SYSTEM PROCESS

Given the functional formulation of the Video Sampling Processor itself, together with its frame synchronized VSP software processes, which derive the target centroid data from the area and moment data for the current frame and use these data to derive the window position and size for target discrimination during the next frame, it was recognized in principle that:

(1) The integrated system process involves the execution of multiple module processes in a scan synchronized sequence ordered form as shown by the upper ("process") portion of modules (3) to (6) in Figure 2. Necessary to the execution,

data and related control/status information is passed between modules as indicated by the lower ("state") portion of the modules.

(2) For each module process, there is basically a measurement/analysis process which, by means of algorithmic procedures, processes measurement related data and generates related control/status information.

(3) For each module process, there is also a control/analysis process which may interrogate control/status information from any or all modules and conditionally control internal process flow and/or update parametric data used by the module for the next sequence (e.g. a feedback loop).

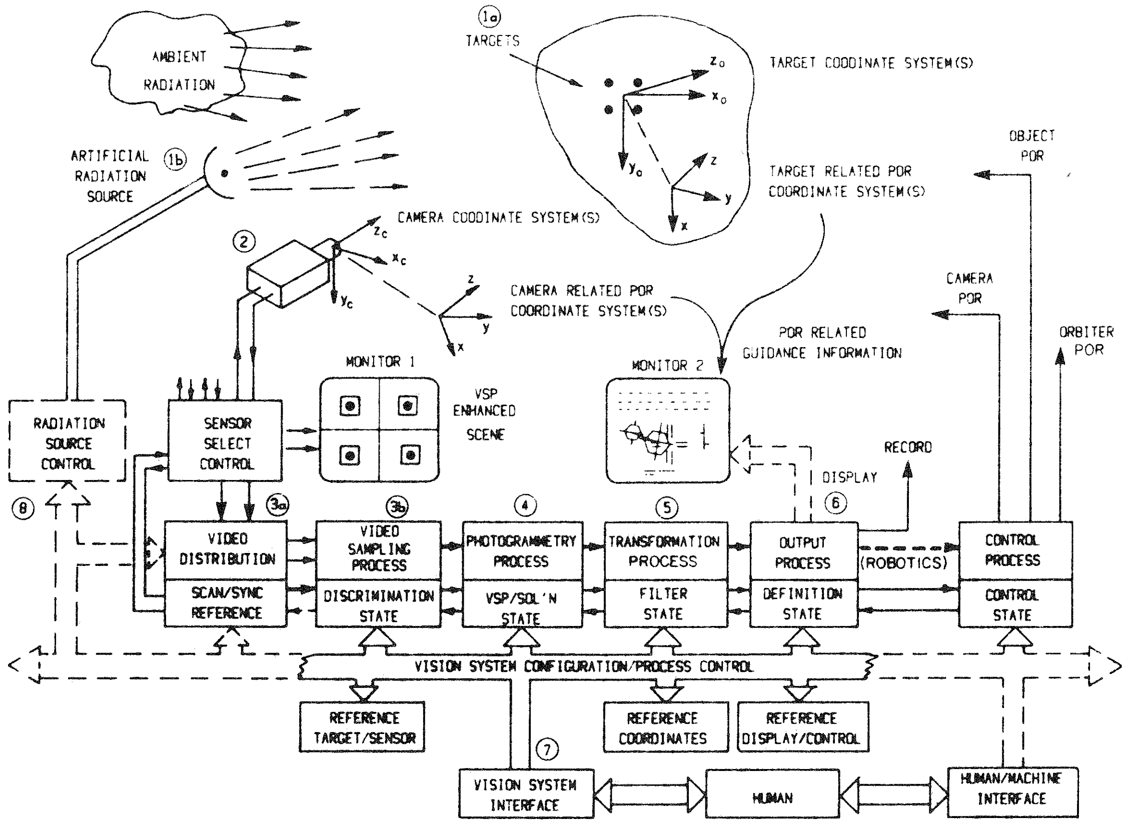


FIG. 2: THE BASIC INTEGRATED SYSTEM PROCESSES

During the execution of a guidance/control task, the module processes use a "current" parametric data set as selected from a "reference" parametric data base. This data base will include the target array(s) geometric data, the camera(s) focal length(s) and calibration data, as well as data defining the position and orientation of the camera and the target array and their control related point of resolution (POR) coordinate systems with respect to their respective reference frames.

To link this reference data to the frame synchronized module processes, a Vision System Configuration/Process Control function is required. This function is shown in Figure 2 by a pseudo data flow bus. It is this high level system function which allows the operator to configure and initialize the system for an application task and during a task it provides the mechanism by which adaptive control, based on control/status information, is imposed on the overall system process as required. As indicated symbolically by the dashed lines in Figure 2, this adaptive control function may in the future accommodate integration with other sensor data and machine control processes.

THE MODULE PROCESSES OF THE INTEGRATED SYSTEM

The modular decomposition of the functional processes of the integrated system is shown in Figure 2. The functional attributes of these processes will be highlighted as they relate to four high level system processes; a Target/Sensor Process (TSP), a Video Sampling Process (VSP), a Real-time Photogrammetry Process (RPP) and an Application Process (AP).

The TSP is comprised of (1a) and (1b) as an application dependent target/source radiation module process together with the camera (2) and its selection and distribution functions (3a). The VSP and the RPP are depicted by (3b) and (4) respectively while the AP is comprised of the application dependent transformation and output/display processes (5) and (6) respectively.

Target/Sensor Process (TSP)

The basic functional objective of the target/source radiation process of the TSP is to provide a discriminable target signal within the output of a scan ordered sensing process (e.g. the raster scan of a video camera) for threshold discrimination and processing by the VSP.

As investigated during the prototype development, Pinkney (1978), this objective can be achieved, for a broad range of ambient sunlight conditions, by using retroreflective targets and a narrow band spectral source with a beam splitter and filter mounted in front of the camera. Where the ambient lighting conditions are favorable or can be controlled (such as indoors), white targets on a black background or black targets on a white background will suffice. Of course, the target radiation could simply be lights or, in other applications, the object surface reflected radiance of a laser projected line or set of dots. Still further, the output signal from the TSP could come from a scanning radar or laser process, or a 3-dimensional (topological) camera or even an ultrasonic C-scan process.

Video Sampling Process (VSP)

The VSP, (3b) in Figure 2, accepts as input the TSP video output signal(s) containing discriminable target signals representing images of target elements having greater or less radiance than the local background. Typical target element types as shown in Figure 3 include; dot, corner solid and line, end of line, vertical and horizontal edge, and vertical and horizontal line.

The VSP determines the centroid of up to four target images by processing the video signal within independent sub-frame apertures (see "windows" in Figure 3) to determine the area and first moments of the threshold discriminated target signal with respect to the window origin. The target centroid is then derived from these data and related to the video frame coordinate system allowing the position and size of each window to be adjusted for the next frame such that the targets are "tracked" as they move within the camera FOV. In addition, when several "dot" type target element sets (arrays) are defined as concentric neighbours, the VSP will effect "autoranging" between the sets to ensure that the track function is maintained on the largest defined array that is within the FOV. The window positioning and sizing and the autoranging functions are implemented as functional components of a "tracking" task whose execution is conditional on the "current" state of enable/disable controls and on control/status states as set during task initialization and are updated each frame by internal procedures which detect error conditions such as no target area, FOV encroachment etc. Attendantly, the discrimination "state" is available for interrogation by subsequent processes as required.

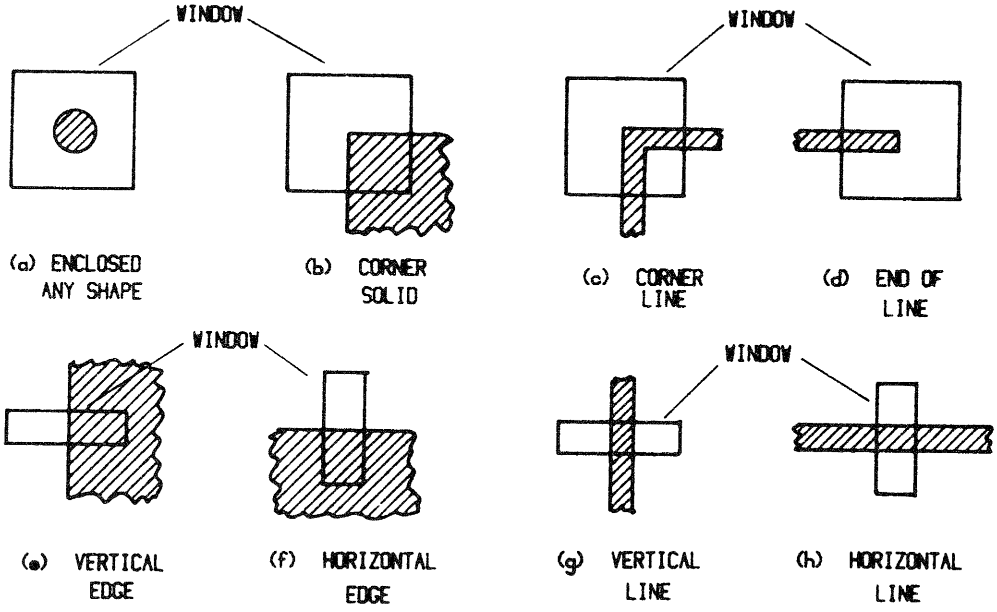


FIG. 3: WINDOW APERTURE DISCRIMINATED TARGET TYPES

To derive target photocoordinates, their centroids in frame coordinates are adjusted according to target type and referenced to camera coordinates by the application of optical axes offsets. These "raw" photocoordinates are then corrected for camera/lens distortions by the application of correction terms derived from a correction matrix generated during calibration. The corrected

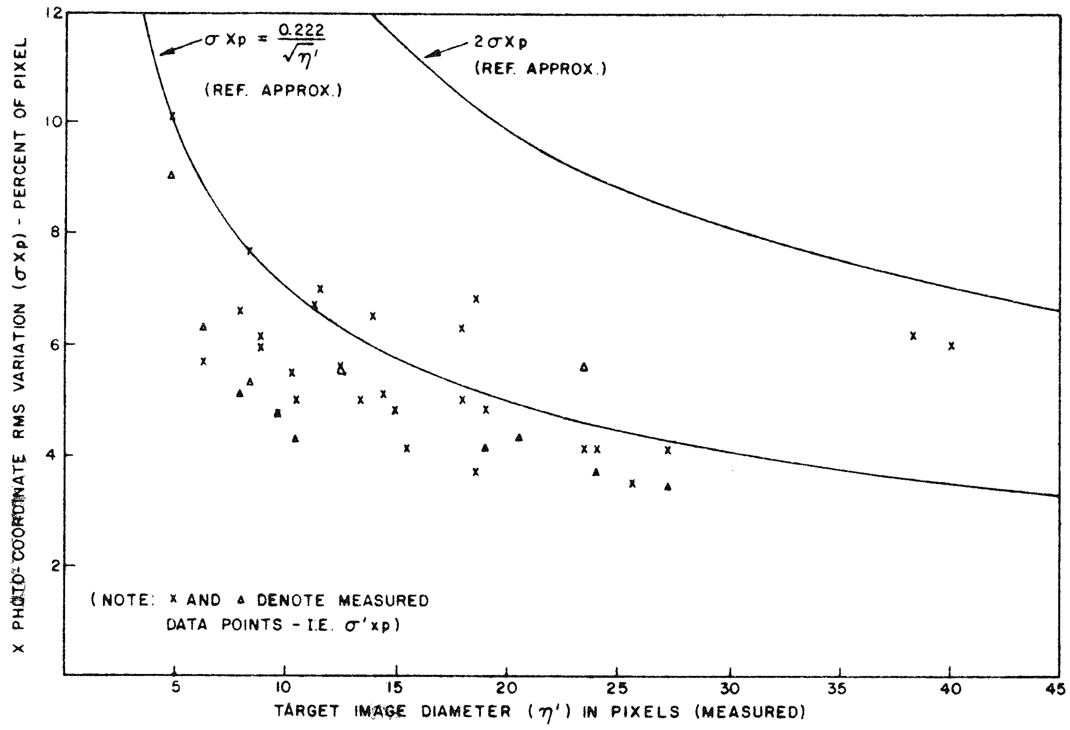


FIG. 4: PHOTO-COORDINATE UNCERTAINTY FOR "DOT" TARGETS

photocoordinates (in the dimensional units of the VSP, i.e. lines/pixels) are then converted to millimeters by the application of calibration determined scalars for output to the photogrammetry process.

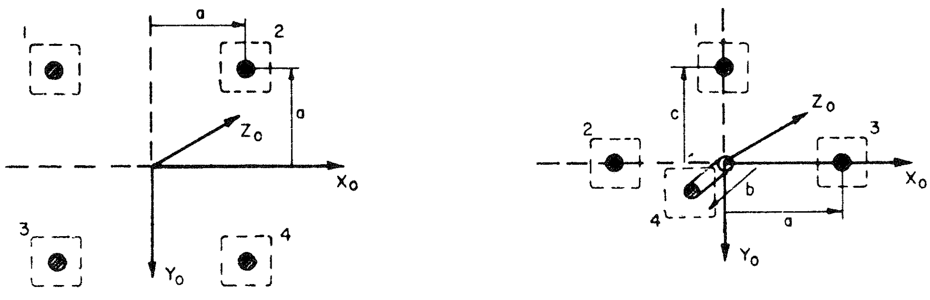
In a study of marking techniques to aid in robotic tasks Bales and Barker (1981) showed that the accuracy of predicting the center of a circular target image improves as the image size increases and that an approximation for uncertainty, σ , (in a least squares sense) is given by $\sigma = 0.222/\sqrt{n}$ where n is the image diameter in pixels. A comparison of this approximation with VSP derived photocoordinates is shown in Figure 4.

Although the tracking and photocoordinate derivation for various target types is a prerequisite to the execution of the photogrammetry process, this inherent VSP capability enabled the NRCC RPS to be used in a study to determine, by independent means, the dynamic response characteristics of the Shuttle Remote Manipulator Arm as described by Basso and Kulchyski (1985). This capability also led to the purchase of a similar system from Leigh by the Industrial Materials Research Institute of the NRCC for application to adaptive robotic welding tasks as described by Dufour and Bégin (1983). In this application the target elements were formed by the projection of laser light.

Real-time Photogrammetry Process (RPP)

Referring to Figure 2 and Figure 5, the RPP accepts as inputs the photocoordinates derived by the VSP and, conditional on VSP status, executes a "tailored" photogrammetric solution process to provide filtered values of the position (Xco, Yco, Zco) and the rotation matrix [Aco] , (for the object/target coordinate system o with respect to the camera coordinate system c), for output to the AP. In addition, by means of a least square straight line fit algorithm the position and orientation rates are derived for output to the AP.

The tailoring of a photogrammetric solution for real-time control is based on control theory considerations. Thus, one solves for the position and orientation of the target with respect to the camera (Xco, Yco, Zco, [Aco]) rather than the inverse relationship (Xoc, Yoc, Zoc, [Aoc]); since in a guidance control task at longer range one primarily uses range and heading (with rates) for control. This approach largely uncouples rotation and translation in the solution form and the equations for deriving the photocoordinates for each of the target elements in terms of the Xco, Yco, Zco, the components of [Aco] and its target coordinates have a direct relationship with the solution. As a result one can then determine which combination of the photocoordinate equations provide the best three equations for the position Xco, Yco, Zco and which components of the rotation matrix [Aco] provide the best three equations for the orientation.



(a) SQUARE TARGET IN-PLANE (b) PEG TARGET

FIG. 5: BASIC TARGET ARRAY GEOMETRIES

As described by Pinkney (1978) this method was first applied to the square array of Figure 5a for small orientation angle use, in which case the selected rotation matrix components for the three orientation equations were approximated by the Euler angles themselves (yaw, pitch and roll) and approximations were similarly used for the rotation components in the Xco, Yco, Zco equations, (feedback with coupling). Filtering was therefore used to control noise and convergence. Since the yaw and pitch noise was proportional to $(Zco/a)^2$ (where a is as shown in Figure 5a) and their contribution to the noise in Zco was proportional to their magnitude times $(Zco/a)^2$, a concentric set of square arrays was employed using "autoranging" to ensure small values of Zco/a within the required operating range. Of course, a peg target, Figure 5b, had a much stronger "tailored" solution for yaw and pitch and better large angle performance.

Based on this experience, for potential SVS use, small angle "tailored" solutions for a planar quadrilateral array and the peg target were formulated using true rotation matrix components, with filtered values from the previous frame. To support the SVS related experiments on flight STS 41-G, October 1984, a rectangle solution for small angle use and a "pitched" rectangle solution for large pitch angle use were formulated. Based on the good results obtained using this "pitched" solution [Tryggvason, Pinkney et al (1985)] these "tailored" solution forms were then reformulated into the NASA coordinates and further refined and evaluated as described in a companion paper by Hughes (1986).

To check the validity and convergence of the current frame derived filtered values of (Xco, Yco, Zco) and [Aco] , the corresponding photocoordinates for each target element are calculated. The square root of the sum of the squared differences between these calculated values and the current VSP photocoordinates is then calculated and output as a system error to the AP for display. The magnitude of this error value is monitored by process control functions to detect system inconsistencies and appropriate action taken when it exceeds a programmable threshold value. It should be noted that, as the error derivation is not a time critical function, it is executed at the reduced cycle time of the application process.

Application Process (AP)

The AP receives data from the RPP, at a 30 Hz rate, which defines the current position, orientation and rate of change of the target coordinate system with respect to the camera coordinate system. These data are processed at a rate not less than 10 Hz by the transformation process of the AP, (5) in Figure 2, to determine the position, orientation and rate of change of a predefined target related Point of Resolution (POR) control coordinate system with respect to a predefined camera related POR. The target and camera related POR coordinate systems are defined in vector form during task configuration and, during task initialization, the angular components of the POR vectors are expressed in rotation matrix form consistent with the form of the orientation data received from the RPP.

Following transformation, an Euler angle solver is used to derive the angular components from the transformed rotation matrix. These values, together with the transformed position and rate values are then converted to alphanumeric representation by the output process (6) for display. The transformed position, orientation and rate data (which may be pre-filtered as required) are also used by the output process to drive selected "patterns" of a synthetic graphics display. The display patterns are designed from straight line segments to provide an operator with a graphical representation of position and orientation and/or rates during "man in the loop" guidance tasks. Sensitivities and offsets appropriate for the application are assigned to each pattern during task configuration. In addition any one of the variables which drive a pattern may be assigned to produce a "trajectory" plot providing a time history of motion during a task.

As shown in Figure 6, lateral motion of objects in the camera scene view during

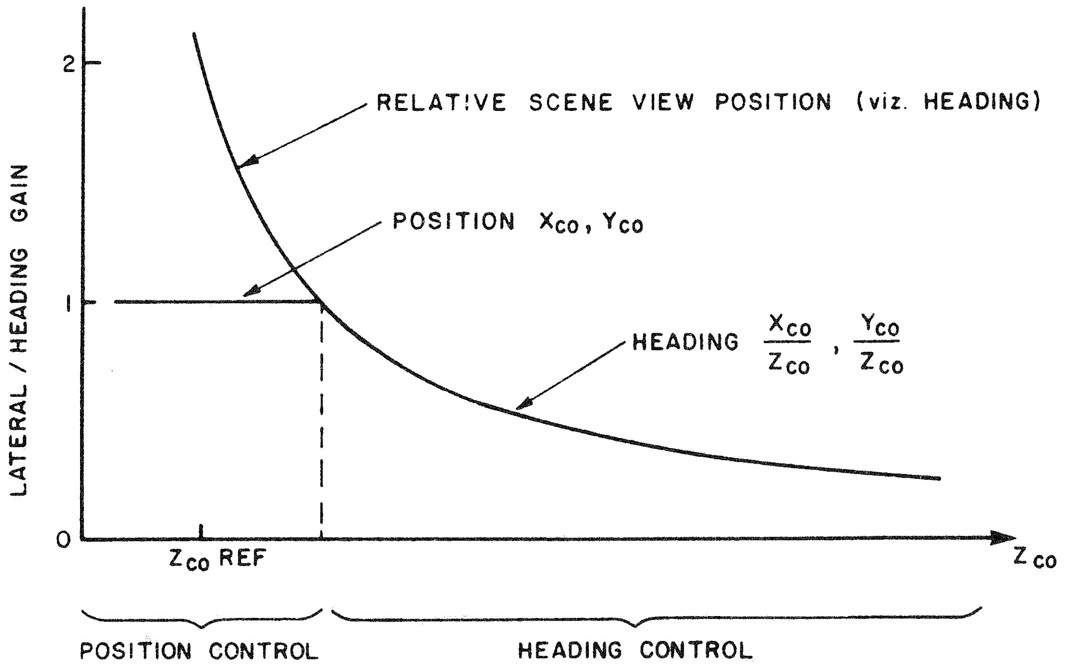


FIG. 6: TAILORING FOR GUIDANCE CONTROL NEEDS

control tasks exhibits an inherent decrease in sensitivity with increasing range. Attendantly, the gain for the variables which drive translational representations in the synthetic display may be dynamically adjusted to provide high sensitivity at close range for "position" control with the sensitivity decreasing as a function of range for "heading" control at longer ranges.

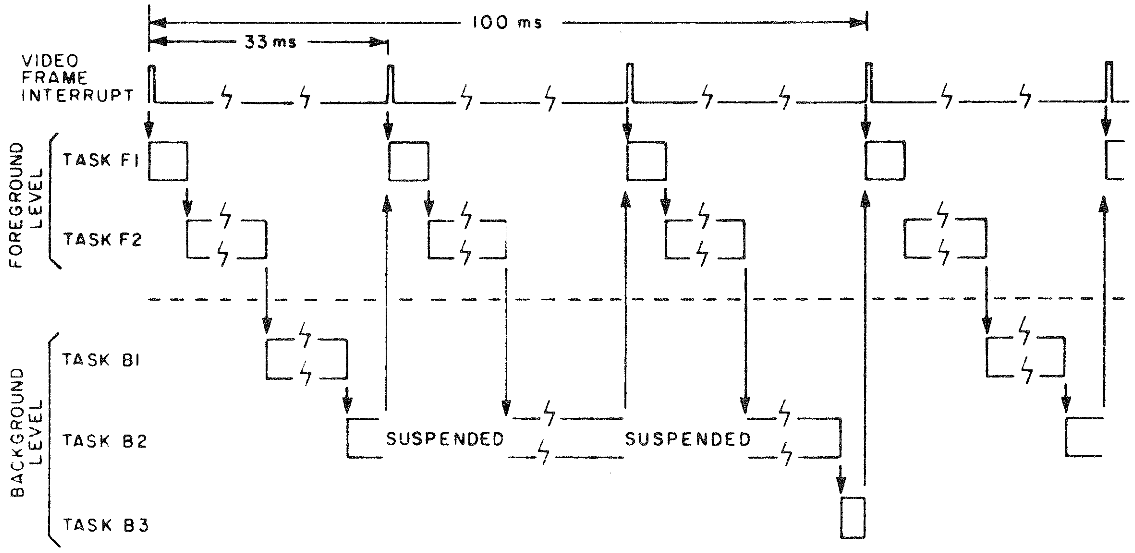
THE INTEGRATED SYSTEM SOFTWARE ARCHITECTURE

Integration of the processes described above, for application to control tasks, requires a flexible software architecture enabling multiple processes to execute within specific time constraints. The time constraints, in general, are imposed to ensure satisfactory system interaction with the dynamics of the task.

In the current system, the video sampling and photogrammetric processes must be executed at the 30 Hz video framing rate to ensure maximum update rates for target tracking and for the derivation of position and orientation rates. The transformation and display output processes must be executed at a rate not less than 10 Hz to ensure satisfactory update rates of guidance information for "man in the loop" control.

To achieve this real-time requirement with a single processor, the processes (delineated according to cycle times) are assigned to execute on separate levels of a multi-level round-robin scheduled software environment as depicted in Figure 7. The highest priority (foreground) level is interrupt scheduled by the video frame signal which also suspends execution of the background level. Once initiated, the foreground level executes to completion at which time the background level is scheduled (or reinstated at the point of interruption).

Within each level, the processes are separated into functional "tasks" which are executed in sequence, i.e. upon completion of the current task, control is passed to the next task in sequence. Thus, the overhead required for task scheduling and context switching is minimized. Also, since the same structure is used for all tasks on both levels, tasks may easily be added, deleted or reassigned within a level (or between levels) according to the application requirements.



TASK F1: TRACKING (WITH AUTOSIZE AND AUTORANGE)
TASK F2: PHOTOGRAMMETRIC SOLUTION
TASK B1: SOLUTION ERROR DERIVATION/TRANSFORMATIONS
TASK B2: DISPLAY GENERATION
TASK B3: IDLE STATE

FIG. 7: MULTI-LEVEL ROUND-ROBIN SCHEDULED SOFTWARE ARCHITECTURE

Each task is structured to allow the transfer of information (status, command and data) between tasks. This exchange of information allows tasks to be designed which optionally execute unconditionally, skip execution or execute only certain procedures on command, or execute any or all of its internal procedures based on status information from other tasks executed in the current or preceding execution cycle. This mechanism allows adaptive tasks to be designed which control process functions such as data and parameter initialization, conditional execution (based on composite error codes) and switching between reference parameter sets (as required by autoranging, for example).

The structured order of task execution within a known context switching sequence greatly simplifies system integration and time based execution analysis. The timing of task modules may be facilitated by embedding DAC commands into the entry and exit segments of each task module and monitoring the DAC outputs. In this manner, compliance with real-time requirements may be easily verified.

DEVELOPMENTS FOR SPACE AND INDUSTRIAL APPLICATIONS

The experimental (prototype) Space Vision System, now being built by SPAR Aerospace and Leigh Instruments under contract from NRCC is fully microprocessor based. When installed on the Shuttle aft flight deck it will accept two video camera signals from the orbiter CCTV system and will return VSP enhanced and synthetic graphics video signals for distribution to the orbiter monitors.

The SVS has improved VSP performance including an autothreshold capability and improved window tracking and sizing resolution. Its "reference" data base provides for 6 cameras, 40 target arrays, 7 synthetic displays, etc. Its software includes 6 "tailored" single camera solutions and a 2 camera, single target "stereo" solution as well as the ability to pre-configure the system for up to 10 tasks (experiments) such as berthing, grappling, assembly, etc. Also included is the ability to autorange between non-concentric arrays of different geometric types (i.e. requiring different solutions).

An industrial robotics application system is being developed by Diffracto Ltd. (Windsor) under a NRCC PILP contribution arrangement for unloading parts from monorail conveyors in automobile factories. In this application retroreflective targets are used together with a ruggedized integrated beam splitter, source, lens, solid state array sensor unit. Using a similar system to the RPS with a DMA link to a minicomputer they were able to use the VSP measurement techniques and a single camera solution to derive vision control data. By means of developing a Vision System Configuration/Process Control function to integrate with a carrier sensor, a parts location sensor and the robot controller, the capability of adaptive switching from manipulator control in its global reference frame using its joint sensors to vision guided control has been achieved. By means of this integrated control capability they are now successfully unloading parts from the parts carrier within only a few seconds after the vision system has acquired the targets.

CONCLUSIONS

An integrated functional module process structure for a flexible machine vision system has been described. Its VSP architecture can be extended to more apertures per frame and can accept inputs from a broad range of target/sensor processes as developed for different applications. The integrated system software architecture provides the adaptive capability necessary to support further development of "tailored" solution forms executing as adaptive algorithmic procedures for use with more general target geometries with good convergence while maintaining a fast real-time processing capability for guidance control tasks.

REFERENCES

Bales, J.W. and Barker, L.K. (1981), "Marking Parts to Aid Robot Vision," NASA Technical Paper 1819, April 1981.

Basso, G.L. and Kulchyski, R.B. (1985), "Use of a Video-Photogrammetry System for the Measurement of the Dynamic Response of the Shuttle Remote Manipulator Arm," Report LTR-ST-1545, National Aeronautical Establishment, National Research Council of Canada, Ottawa, June 6, 1985.

Dufour, M. and Bégin, G. (1983), "Adaptive Robotic Welding using a Rapid Image Preprocessor," 3rd Int. Conf. on Robot Vision and Sensory Controls.

Hughes, R.C. (1986), "Enhanced Single Camera Photogrammetry Algorithms for Real-Time Control Applications," Proc. of ISPRS Commission V Symposium, Ottawa, June 1986.

Kratky, V. (1979), "Real-Time Photogrammetry Support of Dynamic Three-Dimensional Control," Photogrammetric Eng. and Remote Sensing, Vol. 45, No. 9.

Pinkney, H.F.L. (1978), "Theory and Development of an On-Line 30 Hz Video Photogrammetry System for Real-Time 3-Dimensional Control," ISP Symposium of Commission V, PHOTOGRAMMETRY FOR INDUSTRY, Stockholm, August, 1978.

Pinkney, H.F.L., Perratt, C.I., Kratky, V., Ayad, A.A. (1976), "On the Application of Automatic, Real-Time, Single Camera Photogrammetry to Derive the Relative Spatial Orientation and Position of Objects in Machine Tasks. A Conceptual Outline and Preliminary Evaluation," NAE Internal Report, 1976. National Research Council of Canada, NAE Report LTR-ST-1007, August 1981.

Tryggvason, B.V., Pinkney, H.F.L., et al (1985), "Mission 41-G Tests Supporting the Development of a Space Vision System," Canadian Aeronautics and Space Journal, Vol. 31, No. 3, September 1985.

Van Wijk, M.C. and Pinkney, H.F.L. (1972), "A Single Camera Method for the 6-Degree of Freedom Sprung Mass Response of Vehicles Redirected by Highway Barriers," Proceedings of the SPIE Seminar on Optical Instrumentation - A Problem Solving Tool in Automotive Safety, November 1972.