An Integrated Vision Touch-Probe System for Dimensional Inspection Tasks

Marilyn Nashman
William Rippey
Tsai-Hong Hong
Martin Herman
Intelligent Systems Division

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Bldg. 220 Rm. B124
Gaithersburg, MD 20899

NIST

An Integrated Vision Touch-Probe System for Dimensional Inspection Tasks

Marilyn Nashman
William Rippey
Tsai-Hong Hong
Martin Herman
Intelligent Systems Division

·U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Bldg. 220 Rm. B124
Gaithersburg, MD 20899

June 1995

An Integrated Vision Touch-Probe System for Dimensional Inspection Tasks

Marilyn Nashman, William Rippey, Tsai Hong Hong, Martin Herman
National Institute of Standards and Technology
Building 220/B124
Gaithersburg, MD 20899

Abstract

This paper discusses the integration of vision and touch sensors in a coordinate measuring machine (CMM) controller used for dimensional inspection tasks. A real-time hierarchical control system is presented in which a vision system extracts positions of features on a part to be inspected and then guides a touch probe to efficiently measure these features. The probe is tracked by the vision system as it scans surfaces so that its motion can be visually servoed. Minimalist sensor-derived representations, involving only task-specific information, are used in this process. Although the camera itself remains uncalibrated, a real-time calibration of very limited scope is performed each processing cycle to transform the task-specific image information into 3-D information used as feedback to guide the probe. The ability to integrate vision and touch sensors for CMM tasks promises expanded capabilities for flexible inspection data acquisition.

1. Introduction

A coordinate measuring machine (CMM) is a highly accurate three degree-of-freedom cartesian robot, often used for dimensional inspection of mechanical manufactured parts. Dimensional inspection involves measuring the relative geometry of surface features and determining whether they are within tolerance. Examples of geometries evaluated by such a system include shapes of smooth surfaces, distances between edges, positions of holes, and diameters and shapes of holes. For many applications, such as in the automobile or aircraft industries, measurement accuracies on the order of 25 μm (0.001 in.) are required. Virtually all CMMs in use today use touch-trigger probes. When such a probe pushes against a surface with enough force to exceed a certain probe deflection, a signal is sent to the CMM controller to read the machine's scales (i.e., the positions of all the machine axes). Collection of data using such probes is very slow, usually about one point per second. Although edge features contain very important information for the measurement system, object edges are difficult to locate using probes. Also, touch trigger probes are not suited for gathering dense surface data which are important for measurement of parts with complex geometries.

This paper describes advanced data acquisition methods being developed at NIST in the Next Generation Inspection System (NGIS) project. The main objectives of this project (involving a consortium of companies organized by the National Center for Manufacturing Sciences) are to increase the speed and flexibility of data acquisition using CMMs while still maintaining accuracy. Our goals are to increase data acquisition rates by two orders of magnitude, to increase the density of data acquisition, and to simplify the process of measuring geometrically complex parts. We are integrating advanced sensors (including a video camera, an analog touch probe, a point laser triangulation probe, and an analog capacitance probe) with an advanced control system. Our ultimate goal is to develop sensor-servoed scanning control algorithms within an overall control system that can be transferred to manufacturing plants.

In this paper, we focus on the video camera, the analog touch probe, and the control system for controlling probe motion, data acquisition, and interactions between the sensors. The analog touch probe differs from the touch-trigger probe in that it extracts a continuous signal which represents

the deflection of the probe tip as the probe scans over a surface. As the probe is scanned by the CMM arm, an accurate synchronization procedure synchronizes the probe data with the machine scales.

The camera is stationary relative to the part being measured and is positioned so that both the part and the probe are within its field of view. In order to make the system easy to use, camera calibrations should be minimal and easy for inexperienced operators to perform quickly. This would allow shop floor personnel to easily place the camera wherever the viewpoint seems most appropriate.

Because of the low magnification of the camera, visual data are not accurate enough to use for precise part measurements. However vision can provide position estimates of features of interest on the part. Using real-time vision, the probe is guided to features of interest, and probe measurements provide the inspection data. Probe motion is controlled by using feedback from the vision system as it tracks the moving probe. This allows parts to be measured even if an accurate *a priori* model is not available, and, in some cases, allows us to bypass the time-consuming step of part set-up, which uses fixturing or other methods to register the part with the model.

As the probe scans across a surface, the motion of the probe is controlled by information from three sources, the camera, the machine scales, and the probe itself. Vision provides information about positions of part features (e.g., edges, holes, grooves, protruberances) as well as position of the probe. The machine scales, when used in conjunction with vision, provide the distance of the probe from these features. The probe data provide the displacement of the probe from the part surface. We intend to use this information to demonstrate the following capabilities:

(a) efficient probe scanning of smooth surfaces. In order to achieve maximal speed during surface scanning, we want to scan quickly over smooth portions of the surface but slowly over portions near edges. Edge proximity can be detected visually and used as feedback to control arm velocity.

(b) controlling the probe to track an edge. Visual feedback will provide the edge contour so that it can be tracked by the probe.

In a non scanning mode, vision will be used to locate features of interest on the part so that the probe can be moved directly to these features to obtain measurement data.

This paper describes in detail a necessary primitive capability for our system: controlling the speed of the scanning probe using visual feedback. Smooth portions of a surface may be scanned very quickly, but, unless an edge is scanned slowly, the probe will fly over it and fail to measure it completely. Our approach achieves real-time intelligent behavior by using minimalist sensor-derived representations. In such representations, a minimal amount of information required to achieve the given task is extracted from the sensors [4][5][13][17]. The representations contain only task-relevant information. In the example of reducing the probe's speed when it approaches an edge, the task-relevant information needed from the sensors is the current distance from the probe to the edge. This is a value in 4-D space-time, and we demonstrate that it can be extracted using an uncalibrated camera.

In Section 2, we discuss the strengths and weaknesses of camera sensors and touch probes. Section 3 discusses various approaches to sensor integration. Section 4 describes the NGIS hierarchical control system architecture. Section 5 describes the integrated vision-probe system algorithms. Section 6 describes the integrated vision-probe experiment, and Section 7 discusses future work.

2. Vision and Touch Sensors

In order to use the combination of camera and touch probe to its best advantage in an inspection task, we compare the strengths and weaknesses of each sensor. In this discussion, we assume low-magnification lenses on the cameras.The most obvious characteristic of a camera is the fact that it is a non-contact sensor. The advantages of visual information are speed and the global nature of the data. Although camera data are generally noisy, an entire image can be read in 16 ms. The bandwidth for visual information is very high: a typical image can contain between 65,000 and 262,000 pixels. Although a low magnification camera image produces less accurate results than a touch probe, it can quickly locate and measure features of interest in the scene such as object edges, corners, and centroids. A small number of CMM manufacturers are providing imaging capabilities with their systems [22]. These vision systems are designed to obtain very accurate measurements rather than to provide global information about the scene in view. The system described in [22] accomplishes its goal by using high magnification lenses that in effect trades a global field of view for accuracy.

The problems associated with using camera data can be divided into two classes: geometric constraint problems and radiometric constraint problems [21]. Geometric constraints include visibility, field of view, depth of field, and pixel resolution. The radiometric constraints include illumination, specularity, dynamic range of the sensor, and contrast. Section 5 discusses our use of polarizing filters and polarized lighting to reduce specularity. We are not currently using an active vision system which could potentially overcome some geometric and radiometric constraint problems.

A touch probe is a contact sensor. The information it extracts is of a local nature; the data apply only to the specific point touched. Since information is read one point at a time, data acquisition is very slow. For touch trigger probes, the rate of data collection is very low which makes them unsuitable for rapid high-density data acquisition.Touch probes are also quite crash-prone. Nevertheless, they are highly accurate measuring sensors, and there is very little noise associated with their data [3]. Touch probes are best suited for measuring simple geometric features.

3. Sensor Integration

When a single sensor is used to sense the environment, its output is often simple to interpret, but the user of such a system must rely completely on the accuracy and integrity of those data. Single sensor systems are limited in their ability to sense and identify meaningful features under varying conditions. A single source of information can only provide partial information about an environment, and that information is often insufficient to constrain possible interpretations and to resolve ambiguities [9]. The use of multiple sensors to perform a task overcomes the problems caused by relying on a single sensory input, but creates other problems concerning the interpretation and possible merging (fusion) of multiple sensory outputs. A great deal of research has been directed at ways of combining the information from a multiple sensory system. Most methods use measures of statistical uncertainty to model sensor readings. Measures of confidence in the individual sensor readings are updated based on the uncertainty measures [2][9][11][15].

Multiple sensory systems offer many advantages over single sensory systems. Their primary benefit stems from the use of diverse sensors that produce logically distinct outputs. There is some level of uncertainty in the output of each sensor that is due to noise in the system, difficulties in obtaining measurements, calibration errors or sensor degradation [6]. A multi-sensory system uses the diversity of information to overcome the limitations of the individual components [9][11].

The outputs from multiple sensors can be classified into three categories based on their

interactions: *redundant, complementary,* and *cooperative.* The features to be perceived are considered dimensions in a space of features [14] and can be either dependent or independent of one another. Redundant information interaction is found among sensors observing the same object in a scene which measure dependent features in the feature space. A high correlation between redundant sensor readings increases confidence in the validity of the extracted information. A low correlation between redundant sensor readings lowers confidence and suggests a possible sensor error. Touch probe data and highly magnified camera information in a CMM are an example of redundant information. Both sensors supply information about the three-dimensional position of a surface point.

Complementary information interaction occurs when two or more sensors observe features that are independent of one another. In such cases, each sensor provides partial information about the feature in the environment. A simple example of complementary sensor interaction is the integration of information returned by a thermometer measuring the temperature of an object and a range finder measuring the distance from the sensor to the object. The information returned from one sensor can neither strengthen nor weaken the information from another sensor in this configuration, but the combination of returned information provides the user with a greater understanding of the sensed object.

Cooperative information interaction occurs when one sensor's observations guide the actions of another sensor. Guidance can be in the form of physical actions or in the form of software processing decisions. The information obtained from one sensor directs the other sensor to obtain new information relative to a feature of interest. Allen extensively discusses the cooperative interaction of vision and touch in [2]. Figure 1 describes the operation of a cooperative information interaction. The first sensor in this system, S_1, processes its output, d_1 in process T_1. The output from this process is used to guide the actions of sensor S_2 which operates on its data, d_2, in process T_2 and produces output C_i. The dotted line from process T_1 to S_1 represents a closed feedback loop in which the processed output from the sensor S_1, is used to guide its own placement. The feedback loop need not be present in a cooperative system, but its presence adds to the sensors' capabilities. The integrated system described in Section 5 takes advantages of cooperative interaction.The goal of this system is to combine the strength of vision systems-- the ability to gather high bandwidth global information--with the strength of a touch probe, the ability to obtain highly accurate 3-D information. Before describing our application, we describe the system architecture.

4. Integrated System Architecture

The NGIS testbed is designed according to the architecture guidelines of the Real-time Control System (RCS) described in [1]. The architecture defines a hierarchy of controller nodes, each with
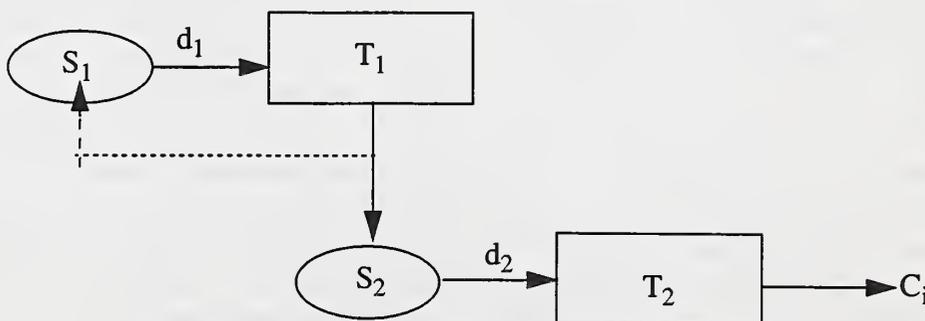
Figure 1 Cooperative Information Interaction
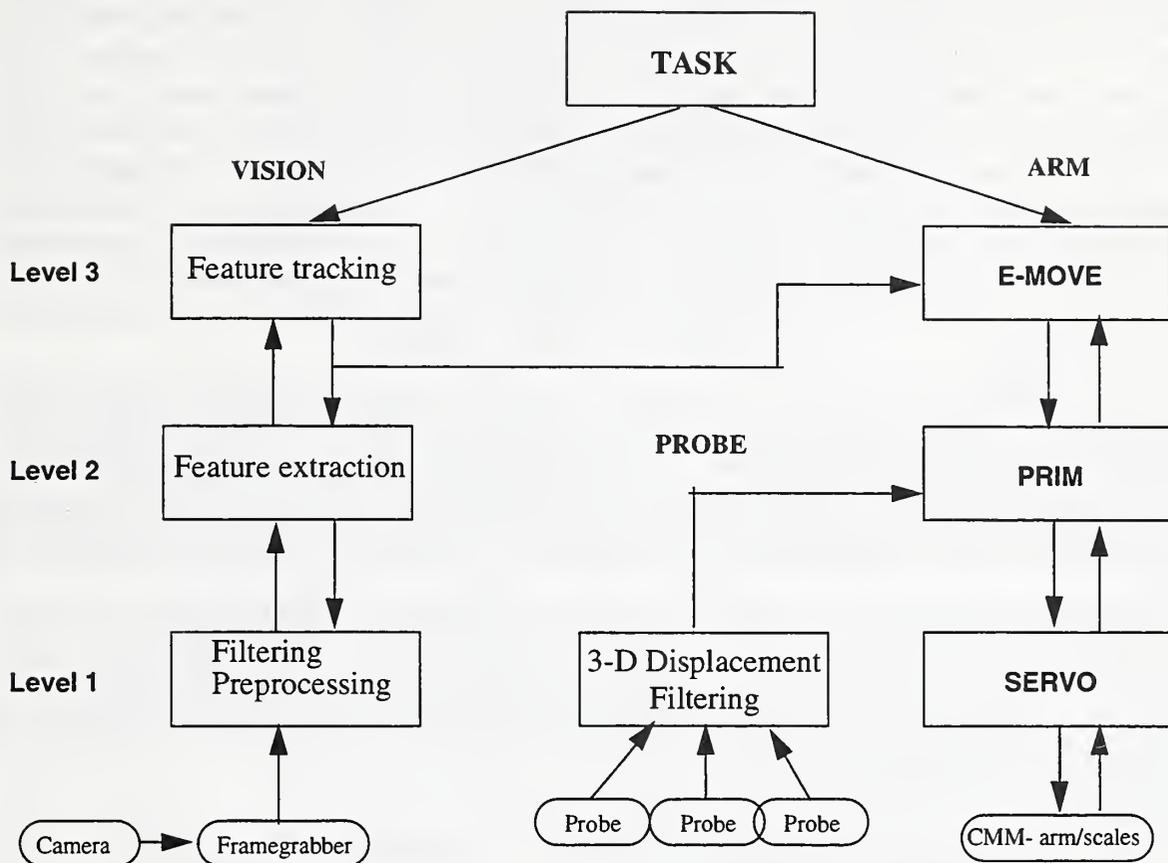
Figure 3 NGIS Control System Architecture

an assigned set of responsibilities that include sensory processing (SP), world modelling (WM) and behavior generation (BG) (Figure 2). Figure 3 describes the functional architecture of the NGIS integrated system. Each rectangle represents a controller node in either the sensory processing hierarchy or the controller hierarchy at a specific level. Although it is not explicitly drawn, each node in Figure 3 consists of SP, WM, and BG modules as shown in Figure 2.

Figure 2 RCS Controller Node

The Servo level is the lowest level of CMM arm control; it translates commanded positions into commands to motors and actuators. The next higher level, Primitive (Prim), computes inertial dynamics and generates smooth trajectories. Its output to the Servo level consists of evenly spaced trajectory points. The third level, elemental (E-move), transforms symbolic commands for elemental movements into strings of intermediate poses which define motion pathways that are free of collisions and kinematic singularities. The fourth level, Task, is the highest level implemented in our system. It transforms goals defined in terms of desired actions to be performed on objects into a series of E-moves designed to achieve these actions.

The role of the sensory processing system is to monitor and analyze information from multiple

sources to recognize objects, detect events, and filter and integrate information. Sensory processing is also hierarchically decomposed into levels which define the scope of the operations at each level (Figure 3). Processing at the lowest level is limited to gathering raw information (readings) from each sensor and filtering the information. In a vision system, one or more cameras act as the sensing agent. Level 1 vision processing reads an image frame digitized from the camera and filters the image to enhance its quality [1]. The analog touch probe reads accurate three dimensional displacement of the probe tip in the probe coordinate system. The information extracted from a touch probe at Level 1 processing consists of continuous values of probe deflection. These readings are filtered before they are made available to the arm control hierarchy. As shown in Figure 3, input to the Probe hierarchy can be read from multiple probes. In our application, a single analog touch probe supplies input.

At the next higher level, Level 2, the image output from Level 1 is analyzed in order to detect two-dimensional image features such as edges, corners, and region attributes such as areas, perimeters, centroids, and cavities (holes). If sufficient information exists, the two dimensional features are transformed into three-dimensional coordinate space. In this way, features extracted from image processing may be expressed in a coordinate system common to that of other positional sensors.

The third level of the vision processing system is responsible for tracking features on moving objects and detecting possible collision. In the arm hierarchy, the filtered displacement values computed by the probe system are integrated with the filtered arm scale readings using calibration constants obtained off-line for the machine and probe. This results in representation of the probe tip in the machine 3-D coordinate system.

5. Integrated Vision Touch-Probe System

The NGIS testbed contains a three degree-of-freedom CMM equipped with a touch probe and a

Figure 4 NGIS testbed

stationary camera mounted on the CMM surface. In Figure 4, the left image displays the probe attached to the CMM arm measuring the surface of a rectangular step object. The right image shows the relationship of the camera mounted on the CMM table to the probe. The camera output is captured by a framegrabber which digitizes it to a 512 x 512 resolution. The CMM table moves in the Z direction (Figure 5), and the arm moves the probe in positive or negative XY directions.

The controller positions the CMM table and the probe in 3-D coordinates relative to a pre-defined
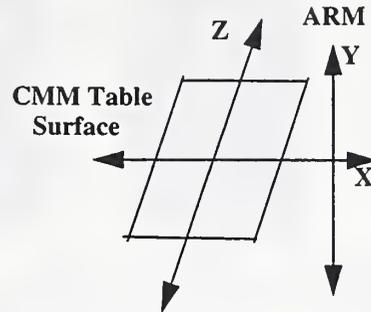
Figure 5 CMM Coordinate System

origin in order to position the probe.

The camera currently used in our experiments is mounted on the CMM table, but it is stationary relative to the part on the table. Lighting conditions and specularity create problems for this application since many of the parts being measured are machine finished to a high gloss. The glare and reflections from overhead lighting introduce shadows and artifacts that interfere with the image processing algorithms. Any efforts to reduce specularity in this environment must be practical as well as effective. For a CMM to be used on the shop floor, it would not be practical to install multiple lighting sources or special purpose lighting to improve the appearance of a computer image. We have introduced a relatively simple and inexpensive technique, polarization of light, to reduce specularity. Reflections are composed of diffuse and specular components; if the specular component could be removed, the remaining information would be easier to interpret [24]. Sheets of polarizing filters attached to the fluorescent light fixtures in the laboratory serve as polarizers at the light source. In addition, we have a rotatable polarizing filter attached to the camera lens that is adjusted to select light polarized at 90 degrees from the overhead filters. Diffuse areas on the part will depolarize the incident light, so half their light will go through the camera polarizer. Specular light will not pass the camera polarizer, and thus is greatly attenuated [8].

The vision-probe system is used in a cooperative interaction integration mode. Global information generated by the vision system is used to guide the movement of the touch probe across the surface of the part. Vision provides information about positions of part features of interest, e.g. edges, and, in conjunction with the machine scales, the distance of the probe from these features. The probe data provide the displacement from the part surface. Using this information, the controller scans the probe quickly over smooth portions of surfaces and slowly over portions near edges.

Visually derived feedback in the control system consists of the distance (in millimeters) between the current position of the probe and the edge of interest. There are five steps in guiding the probe and alerting the controller when the probe is approaching a part edge.

(1) Extracting edge pixel locations and orientations.

(2) Fitting edge pixels to line segments corresponding to the edges of the part.

(3) Extracting probe position and computing goal position.

(4) Filtering the computed probe position and predicting its position in the next image.

(5) Tracking the probe as it moves along the part's surface.

A Sobel edge extraction is performed on the full image [18]. The spatial derivatives $\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$
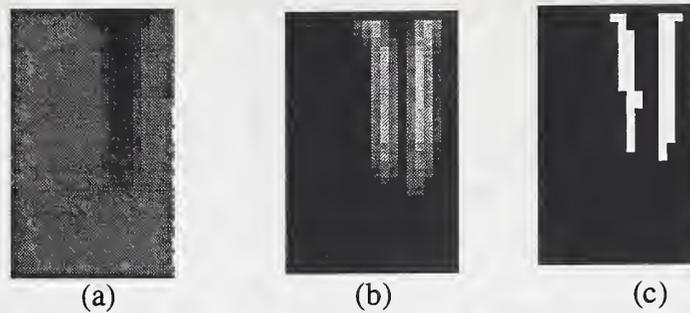
Figure 7 (a) raw image of probe

(b) gradient image of probe (c) thresholded gradient image

are computed using 3 x 3 spatial gradient operators. The magnitude is thresholded at a pre-determined grey level in order to preserve only edgels in high contrast areas. A list is generated of edge positions and orientations in image space.

A Hough transform [12] fits the extracted edges to straight line segments. Figure 6, showing representative results for the step-block, displays a graph of the extracted edge points overlaid with the lines (shown as dotted lines) selected by the Hough transform.
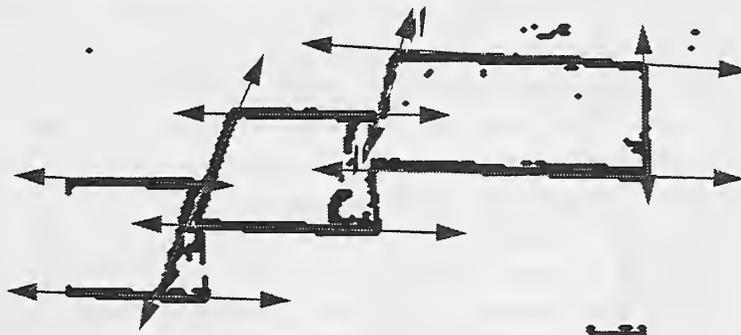
Figure 6 Extracted Edge Pixels and Fitted Lines

Tracking the probe in the video imagery begins with calibration. First the controller places the probe in a manually determined location on the part surface. Then, in an off-line operation, the system operator determines the approximate location of the probe in image coordinates by examining the image. The operator's coarse estimate in image coordinates is used to search a small window in the image from which an accurate 2-D position of the probe is computed. It is assumed that the window contains an image of the probe; if not, an error is generated. To segment the probe in the image, we take advantage of its distinguishing characteristic: its vertical orientation. The horizontal gradient is computed and dynamically thresholded. The use of dynamic thresholds eliminates the problems caused by inconsistent lighting, shadows, poor contrast, and reflection. We choose the bottom-most pixel above threshold as the 2-D location of the probe. Figure 7a shows an enlarged view of the image window containing the touch probe based on the operator's estimate of the initial location. The original image quality is poor because of low contrast in the scene and the shadow cast by the probe on the CMM table (lower left corner). Figure 7b shows the grey-scale results of applying a 3x3 gradient mask to extract the horizontal gradient. Figure 7c shows the results of applying a dynamic threshold to the gradient image based on a cumulative histogram. The probe location is computed to be the coordinates of the bottom-most above-

threshold pixel in the window. A 10 x 10 grey scale template centered at the probe location is generated for the correlation tracking algorithm described below.

Assuming that the probe's motion while scanning along the surface is a straight line path, we next compute a vector in the image plane in the direction of motion using the initial probe position and a second image position computed either as the probe moves or entered by the operator. The intersection in the image plane of the motion vector and the nearest part edge represents an initial estimate of the probe's 2-D goal point. The distance between the initial position and the goal position is measured in pixels.

We track the probe as the arm moves towards its goal position using the predicted probe image velocity and sum of absolute differences (SAD) correlation algorithm[18]. A predictive filter is used to filter and predict the probe position and velocity at the next time interval [7][19][23]. The prediction is based on a weighted sum of the current position and velocity and a history of past positions and velocities. Depending on the weights used, the predictions can be tuned to be more responsive to new readings or to respond smoothly over time. At each processing iteration, the search direction and window used for the SAD correlation are recomputed based on the predicted probe image position and velocity. The size of the search window is determined by the probe velocity magnitude, and the direction of search is biased in the direction of velocity. The probe position is computed to be the position which yields a minimum value for the sum of absolute differences over the search space. The correlation template is updated each cycle to reflect the current grey scale information representing the probe position.

Feedback to the arm control system consists of updates of the 3-D distance between the probe and the nearest edge object. Figure 8 describes the relationships in the image plane that are used to compute this information.
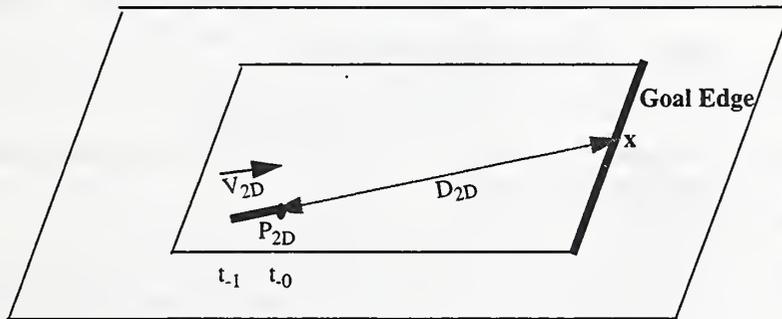
Figure 8 Image Plane Representation

P_{2D} is the probe position in the image plane at time t_0 and V_{2D} is the velocity in the image plane measured between times t_0 and t_{-1}. Both of these parameters can be measured, since the probe's positions in the image have been extracted. At each processing cycle, we recompute point x, the image goal position, by intersecting the motion vector generated by P_{2D} and V_{2D} with the part edge. (This works best for straight-line motion of the probe.) Using the updated 2D goal point and the current probe position, we compute D_{2D} to be the distance between the two points. Assuming constant velocity, the time required to reach the goal point, in both 2-D and 3-D coordinate systems, is simply:

$$T = (D_{nD}/V_{nD}) \quad where \ n = 2 \ or \ 3$$

Using this formula, T is computed from 2-D information. In 3-D space, the probe position and ve-

locity, P_{3D} and V_{3D}, are computed by the arm E-move level and made available to the vision system through a common memory communication process[10]. The 3-D distance to the edge can therefore be computed each cycle as $D_{3D} = V_{3D} * T$. This distance is computed in millimeters every processing cycle and provides feedback to the arm controller. In many situations, the probe motion will be along a curved, rather than straight, path, and certainly the probe velocity is not constant. However, since the image goal point and the time to reach the goal are updated every processing cycle, we expect that as the probe approaches the goal, the estimated distance to the goal will become increasingly accurate. In effect, we are re-calibrating this distance in real-time every processing cycle. This distance is a task-specific value, as it is used to determine the speed to scan the probe. As this distance decreases, the probe must be slowed down. Since no other 3-D information about the environment need be known, the camera itself can remain uncalibrated.

6. Experiments and Results

The equipment used in the NGIS testbed laboratory includes a three degree of freedom coordinate measuring machine as described in Section 5, multiple interchangeable probes, a black and white CCD camera with a 16mm lens, a framegrabber, Sun SPARC4 workstations, and a VME-based multiprocessor system running a vxWorks operating system. Code is written in C++ on the workstations, and downloaded to the vxWorks system.[1] The purpose of the experiment described here is to demonstrate our real-time control system in which computer vision is used to determine the distance from the probe to a goal edge as the probe scans over a planar surface. This distance is used to control the speed of the probe as it nears the edge. Although the testbed is equipped with many probes, this experiment uses only a single touch probe.

The experiment is initiated by the controller task level (Task) shown in Figure 3. Commands are sent to the elemental level (E-move) and to the vision Level 3 process via a common-memory interface[10]. Table 1 describes the commands generated by Task which are sent to either arm

Table 1: Demonstration Scenario

Acting Level	Commanded Action
Arm E-move	Move arm near part
Vision Level 3	Extract and define edges of part in field of view
Arm E-move	Place probe at pre-defined starting position on part surface
Vision Level 3	Extract 2-D position of probe on part surface
Vision Level 3	Determine 2-D goal point of expected probe trajectory
Arm E-move	Scan surface until proximity to goal reported
Vision Level 3	Loop and track probe; report distance to edge
Arm E-move 3	Terminate tracking

---

1. Certain commercial equipment, instruments, or materials are identified in this paper in order to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by NIST, nor does it imply that the materials or equipment identified are necessarily best for the purpose.
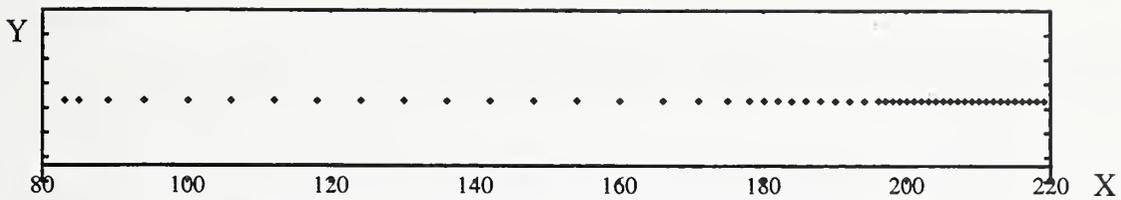
Figure 9 2-D Tracked Probe Positions over Time

E-move or vision Level 3 processes. Note that the command to extract and define edges of the part in the field of view occurs just once for a particular field of view. This means that the edge extraction and Hough transform processes, which are time consuming, are run only once. The probe tracking process, on the other hand, is updated every processing cycle. Both the arm control and vision processes decompose their commands for the levels hierarchically below them. Status information is always sent to the requesting process indicating either completion of the command or an error condition. Visual tracking is terminated when the closest edge is detected as being less than a pre-defined δ millimeters away from the current probe position. Visual feedback in the controller is used to command arm velocity based on distances to features of interest. Currently, the arm is programmed to scan at 15 mm/s when the distance to the goal is greater than 50 mm; 5 mm/s for distances between 50 mm and 10 mm; and 2 mm/s for distances less than 10 mm. Figure 9 is a plot of 2-D probe positions extracted during a typical experiment. Each point represents the probe's x-y image plane coordinates over time. (Time is expressed in updated vision cycle units.) The direction of motion in this case is left to right. At the start of the run, the arm is scanning at 15 mm/s, and since the image acquisition and processing time is relatively constant, fewer points are extracted from the image sequence. Points between X = 178 and X = 196 represent positions extracted during the intermediate velocity, and points extracted between X = 198 and X = 220 represent the extracted probe positions at the slowest velocity.

Figure 10 describes the results of a typical experiment. The y axis represents 3-D distance in mm of the probe from the part edge; the x-axis represents time units. The true distance between the starting position and the edge goal point is 120 mm. The curve represents the real-time estimates of the 3-D distance remaining to the object edge over time. This information is computed during run time. The first velocity change occurs at time t = 30; the second velocity change occurs at time t = 40. The 3-D estimates computed near these changes display discontinuous peaks caused by both misregistration between the 2-D computations and the 3-D readings and instability in the predictive filters. The computed distances resume a smooth path when proper registration is restored and the confidence in predictions increases. Between times 40 and 70, the visually computed distances decrease 0.5 mm each processing cycle. The error between the true probe position and the position reported by the vision system at the conclusion of the experiment was 2.5 mm.

We have demonstrated the integrated vision touch-probe system over many experiments. Our assumptions about the positioning of the part being scanned are that we know the initial probe position and that the start position and goal edge are in the camera's field of view. The visual servo feedback loop has performed reliably in over 100 experiments. Measured differences between computed final distance and true distance to the goal point range between 0 and 4mm. These differences are attributed to digitization error, coarseness of the Hough transform line fit, and communication delays between the vision process and the arm Emove process.

Visual processing is performed on the multiprocessor system; other than the framegrabber, no
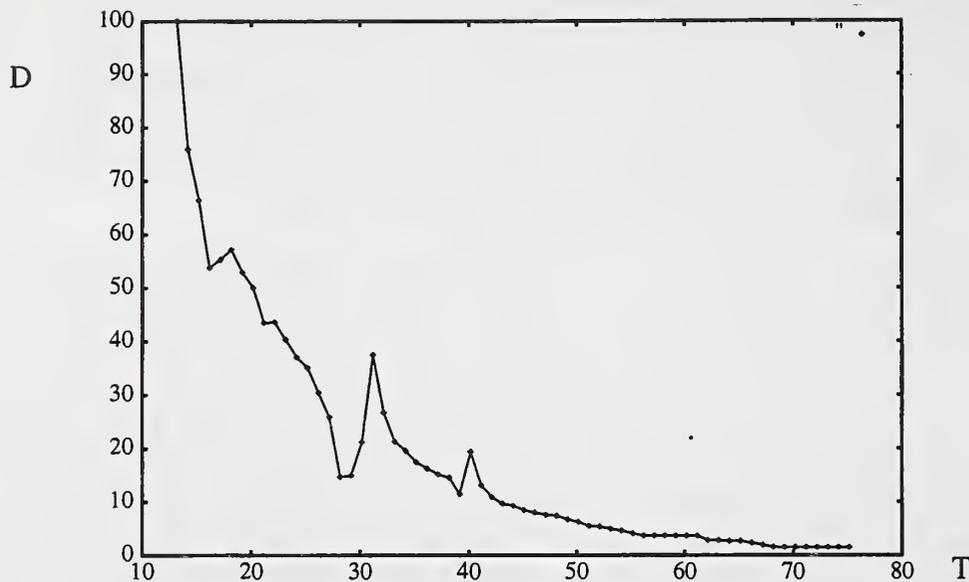
Figure 10 Experimental Results

Table 2: Image Processing Timing

Type of Processing	Time (in 10 millisecond units)
Digitize and store image frame	5
SAD correlation	5
Filter and predict next probe location	1

special image processing hardware is used. This results in long processing times for the more computationally intensive operations such as convolutions, Hough transform, and SAD algorithm. The time critical component for real-time visual processing is data acquisition and the SAD tracking algorithm; edge extraction and line fitting are performed only once for each run. A tracking iteration is performed every 110 ms, or at a rate of a little above 9 Hz. This time is decomposed as shown in Table 2. We expect that the tracking related computations will be reduced when existing faster microprocessor boards are installed.

7. Future Work

Our initial experiments were performed on a simple geometric part (Figure 4). We plan to perform similar experiments on complex parts. We are also implementing additional image processing capabilities in order to detect features such as curved edges, holes, and grooves on complex parts. We are investigating fast, easy to perform, calibration methods in order to register image space with CMM space; this will give us the ability to visually servo the arm to features of interest and to inspect and follow linear and curved contours. Our ultimate goal is to develop sensor-servoed scanning control algorithms that can be transferred to manufacturing plants.

References

[1]   J. Albus, "Outline for a Theory of Intelligence," IEEE Transactions on Systems, Man and Cybernetics, 21(3):473-509, 1991.

[2]   P. K. Allen, Robotic Object Recognition Using Vision and Touch, Kluwer Academic Publishing, Boston, MA., 1987.

[3]   P. K. Allen, R. Bajcsy, "Object Recognition Using Vision and Touch," Proceedings of the 9th International Joint Conference on Artificial Intelligence, Los Angeles, CA, August, 1984, pp 1131-1137.

[4]   J. Aloimonos, I. Weiss, A. Bandyopadhyay, "Active Vision," International Journal of Computer Vision 1:333-356, 1988.

[5]   D. Ballard, C.Brown, "Principles of Animate Vision," CVGIP: Image Understanding, 56(1):3-21, 1992.

[6]   S. S. Blackman, "Theoretical Approaches to Data Association and Fusion," SPIE Volume 931 Sensor Fusion, 1988, pp 50-55.

[7]   Coombs, D., Herman, M., Hong, T., Nashman, M., "Real-Time Obstacle Avoidance Using Central Flow Divergence and Peripheral Flow," Proc. Fifth International Conference on Computer Vision, Cambridge, MA, June 1995.

[8]   R. Daum, K. Harding, "The Machine Vision Lighting Testbed," Applied Machine Vision Conference, June 1992.

[9]   H. F. Durrant-Whyte, "Sensor Models and Multisensor Integration," International Journal of Robotics Research, Vol 7 (6), 1988, pp 97-113.

[10]   J. Fiala, "Note on NASREM Implementation," NIST Internal Report 89-4215, Gaithersburg, MD. 1989.

[11]   G. D. Hager, Task-Directed Sensor Fusion and Planning, Kluwer Academic Publishers, Boston, MA, 1990.

[12]   R. Haralick, L. G. Shapiro, Computer and Robot Vision, Addison-Wesley Publishing Co., 1992.

[13]   M. Herman, D. Raviv, H. Schneiderman, M. Nashman, "Visual Road Following Without 3-D Reconstruction," Proc. SPIE 22nd Applied Imagery Pattern Recognition Workshop - Interdisciplinary Computer Vision: Applications and Changing Needs, Vol. 2103, Washington, D.C., October 1993, 226-236.

[14]   R. C. Luo, M. G. Kay, "Multisensor Integration and Fusion in Intelligent Systems", IEEE Transactions on Systems, Man, and Cybernetics, Vol 19, No 5, September/October 1989, pp 901-931.

[15]   R. C. Luo, M. Lin, R. S. Scherp, "The Issues and Approaches of a Robot Multi-Sensor Integration," Proceedings IEEE International Conference on Robotics and Automation, Raleigh, NC, Volume 3, April 1987, pp 1941-1946.

[16]   Nashman, M., "The Use of Vision and Touch Sensors for Dimensional Inspection

Tasks," Manufacturing Review, Vol. 6, Number 2, June 1993, pp 155-162.

[17]   D. Raviv, M. Herman, "Visual Servoing from 2-D Image Cues." In Active Perception, Y. Aloimonos, ed., Lawrence Erlbaum Associates: Hillsdale, NJ 1993, 191-226.

[18]   A. Rosenfeld and A. Kak, Digital Picture Processing, Vol. 1, Second Edition, Academic Press, 1982.

[19]   H. Schneiderman, M. Nashman, "A Discriminating Feature Tracker for Vision Based Autonomous Driving," IEEE Transactions on Robotics and Automation, Vol. 10, 1994.

[20]   M. J. Swain, M. Stricker, Eds., "Promising Directions in Active Vision," NSF Active Vision Workshop, Chicago, Ill., 1991.

[21]   K. Tarabanis, R. Y. Tsai, "Sensor Planning for Robotic Vision: A Review," Robotics Review II, Ed. O. Khatib, J. Craig, Lozano-Perez, MIT Press, 1992, pp 113-136.

[22]   VIEW Precis 3000 Manual, VIEW Engineering Inc., 1989.

[23]   A. J. Wavering, R. Lumia, "Predictive Visual Tracking,", Proceedings of SPIE Intelligent Robots and Computer Vision 7, Boston, MA,1993.

[24]   L. B. Wolff, "Using Polarization To Separate Reflection Components," CVPR, 1989, pp 363-369.