The Use of Vision and Touch Sensors for Dimensional Inspection Tasks

Marilyn Nashman
Robot Systems Division; National Institute of Standards and Technology
Bldg. 220/ B124; Gaithersburg, MD 20899

Abstract

The purpose of dimensional inspection is to verify the geometry of a manufactured part. A machined part is either accepted or rejected based upon the sensed errors between the object and its specified geometry as defined in a CAD (Computer Aided Design) model or other model database. Various sensors can be used for inspection tasks. The use of multiple sensors is relatively new in this application and coordinate measuring machine (CMM) manufacturers have only recently begun supplying machines that provide multiple sensor capabilities. The purpose of this paper is to discuss the current use of vision and touch sensors for inspection tasks and to suggest alternative strategies for the use of these sensors to increase their capabilities.

Introduction

The purpose of dimensional inspection is to verify the geometry of a manufactured part. A machined part is either accepted or rejected based upon the sensed errors between the object and its specified geometry as defined in a CAD (Computer Aided Design) model or other model database. Various sensors can be used for inspection tasks. The use of multiple sensors is relatively new in this application and CMM manufacturers have only recently begun supplying machines that provide multiple sensor capabilities. The purpose of this paper is to discuss the current use of vision and touch sensors for inspection tasks and to suggest alternative strategies for the use of these sensors to increase their capabilities. We briefly describe a hierarchical architectural design for a sensory processing system. In addition, we discuss issues related to sensor integration that have been raised by the robotics research community as they might apply to inspection tasks. We compare the strengths and weaknesses of camera imaging sensors and touch probes. We describe the VIEW CMM [1]system currently being used by the Robot Systems Division and the Precision

Engineering Division at the National Institute of Standards and Technology (NIST) as an example of a multi-sensor CMM and describe its use of vision and touch sensors. Lastly, we offer suggestions for future development of multi-sensor systems for inspection.

Sensory Processing Architectures

Sensory processing systems monitor and analyze information from multiple sources in order to recognize objects, detect events, and filter and integrate information. In a hierarchical system, this processing is divided into levels which define the scope of the operations defined at each level (figure 1). For example, in the NASA/NBS Standard Reference Model for Telerobotic Control System Architecture (NASREM) [1], processing at the lowest level is limited to gathering raw information (readings) from each sensor, filtering the information, and when applicable, enhancing it. When considering machine vision, a camera acts as the sensing agent. Level 1 processing reads an image frame from the camera, digitizes it and performs filtering and/or enhancement in order to improve the image quality[2]. The input data consists of a rectangular array of digitized points (pixels), and the output consists of an array of processed pixels. The touch-trigger probe is a three dimensional sensor capable of computing accurate three dimensional position. The information extracted from a touch probe at Level 1 processing consists of the electrical impulse signaling contact with the part surface.

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 attributes of regions such as area, 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 are expressed in a coordinate system common to that of other positional

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.

sensors. Level 2 processing for the touch-trigger probe involves reading the x, y and z coordinates of the touch probe based on apriori calibration constants inherent to the machine and probe.

Information from the individual sensors is integrated at Level 3. A distinction is made between sensor integration and sensor fusion [3]: sensor integration is defined as the use of multiple sensors to attain more information than any one sensor is capable of delivering. The goal of sensory integration is to create a better understanding of a scene by pooling multiple sources of information in order to emphasize points of agreement and de-emphasize points of disagreement [4]. Sensor fusion, on the other hand, refers to the actual combination of multiple sensory outputs into a single representation. Figure 2 describes a system containing two sensors, S_1 and S_2. The output generated by each of these sensors, d_1 and d_2, is interpreted individually in processes T_1 and T_2 respectively based on the special features of each sensor.The results of the processing, C_a and C_b, can either be fused into form C_f or treated as input components in an integrated system. The information contained in an integrated system more fully describes the sensed environment. In the following section, we discuss the issues related to sensor integration using camera data and touch-trigger probe data.

Sensor Integration

In the current technology, virtually all coordinate measuring machines in the plant environment use touch-trigger probes for manufacturing inspection tasks. The output from a single sensor is relatively simple to interpret, but the user of such a system must rely completely on the accuracy and integrity of that 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 usually insufficient to constrain possible interpretations and to resolve ambiguities [5]. 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. The reader is referred to [5, 6, 7, 8] for in depth discussions of the statistical techniques.

Multiple sensory systems offer many advantages over single sensory systems. Their primary benefit stems from the use of diverse sensors which produce logically distinct outputs. The output from each individual sensor contains some level of uncertainty caused by noise in the system, difficulties in obtaining measurements, calibration errors or sensor degradation [9]. In a multi-sensory system, the diversity of information is used to overcome the limitations of the individual components. Each sensor in the system is modelled to describe its special abilities, the form of its output, and the interpretation of its output relative to the other sensors in the system. A sensor can be classified according to its strengths relative to the task to be performed as well as to its limitations relative to that task. Among the factors influencing this categorization are the complexity of the individual sensor, observation errors caused by sensor placement, understanding of the implicit and explicit calibration constants, sensor noise, and an understanding of techniques for expressing sensor output in a common measurement system. These issues are discussed in detail in [5, 7].

The outputs from multiple sensors can be classified into three categories based on the interactions of the outputs generated by the sensors. These interactions are labelled competitive interactions, complementary interactions, and cooperative interactions. The features to be perceived are considered dimensions in a space of features [3] and can be either dependent or independent of each other. Competitive information interaction is defined as the interaction between sensors perceiving the same object in a scene which measure dependent features in the feature space. The information provided by these sensors is called redundant information since the sensed information is not adding additional knowledge to the system. The integration of competitive sensors results in either an increased confidence in the validity of the extracted information when the readings support each other, or, conversely, a lowered confidence indicating

a possible sensor error when the readings disagree. The use of a touch-trigger probe and camera information as combined in a multi-sensory CMM provides an example of competitive information. Both sensors supply information about the three-dimensional position of an object feature. When dealing with competitive sensory information, the issue of a common coordinate system for interpreting individual sensor readings is very important. Each sensor is first calibrated relative to its own internal representation, and then transformations are computed which convert the sensor readings into a common coordinate system.

Figure 3 describes a competitive information interaction in a system containing two sensors, S_1 and S_2 which produce redundant outputs d_{1a} and d_{1b}. Processes T_{1_s1} and T_{2_s2} contain algorithms which convert the outputs into units relating to the coordinate systems of the individual sensors. These results are passed into a third process which transforms the outputs into a common coordinate system.

Complementary information interaction occurs when two or more sensors supply information about the same object in a scene but each sensor perceives features that are independent of the features perceived by the other sensors. 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. A second example is the use of a touch probe to measure a feature occluded in a camera scene. The returned information 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 a greater understanding of the sensed object. Figure 4 describes an example of a complementary information interaction. Using the same notation as in figure 3, processes T_1 and T_2 produce independent information. There is no need for a process to combine the outputs of T_1 and T_2, but the individual outputs C_a and C_b create a fuller understanding of the sensed object.

Cooperative information interaction occurs when one sensor's observations guide the actions

of another sensor. The information obtained from one modality 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 [6]. Figure 5 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 lines from process T_1 to S_1 represent a closed feedback loop in which the processed output from the sensor S_1, is used to guide its own placement. The feedback loop does not have to be present in a cooperative system, but its presence adds to the sensors' capabilities.

Imaging and Touch Sensors

In order to use the combination of a camera and a touch probe to its best advantage in an inspection task, we compare the strengths and weaknesses of each sensor. We then discuss the cooperative interactions possible with this combination of sensors.

The most obvious characteristic of a camera is the fact that it is a non-contact sensor. (Cameras are sometimes called passive sensing devices because of this characteristic, but that term is confusing in the context of active vision research [10].) The advantages of using visual information are speed and the global nature of the data. An entire scene can be read in 16 milliseconds. Depending on the image processing hardware and the complexity of the algorithms performed, simple features can be extracted very easily. The bandwidth for visual information is very high: a typical image can contain between 65,000 and 262,000 pixels depending on the camera resolution. In addition, camera data is generally noisy. A full resolution non-magnified camera image produces less accurate results than a touch probe, but can quickly locate and measure object edges, corners, and centroids. The issue of reduced accuracy has been addressed by the manufacturers of coordinate measuring machines equipped with vision systems. In effect, global information capabilities are traded for increased accuracy by using high- magnification camera lenses for which the field of view is reduced. This results in each pixel in the field of view covering a smaller area

which in turn improves the accuracy of the measurements.

The greatest problems associated with using camera data are caused by poor lighting conditions, surface reflectance, and occlusion. In cases where static cameras are used, problems also occur due to a limited viewpoints. Active vision systems [10] are designed to overcome the problem of restricted viewing situations.

A touch-trigger probe is a contact sensor. The information it extracts is of a local nature; the data applies only to the specific point touched. Since information is read one point at a time, data acquisition is very slow. The bandwidth of touch probe data is very low and there is much less noise associated with the data [11]. Touch-trigger probes are highly accurate measuring sensors and are best used for measuring simple geometric features. Within the constraints of their range of motion, they are not affected by viewpoint considerations.

Current Experimental Testbed

The VIEW 3000 measurement machine (figure 6) is being used to learn about multi-sensor interactions for inspection tasks at NIST. The system is equipped with a camera and a touch-trigger probe [12]. In this section, we describe the physical characteristics of the VIEW machine and the sensor strategies available to the user.

The camera mounted on the VIEW machine has three camera lenses: a low-magnification lens (1.6X), a medium-magnification lens (3.33X) and a high-magnification lens (6.66X). These magnification factors result in an effective field of view of 103.2 mm^2 (0.16 square inches), 51.6 mm^2 (0.08 square inches), and 25.8 mm^2 (0.04 square inches) respectively. The video signal from the camera is digitized into a rectangular array of picture elements that is 512 pixels wide and 480 pixels high.

System software provides tools for extracting three dimensional position information from the camera data. The system is programmable and offers routines to calibrate each camera lens in order to convert the two dimensional pixel coordinates into three dimensional position. The z dimension

(the height of the camera over the object surface) is measured by using the autofocus feature of the camera. The system contains software routines to extract edge points, compute region centroids, and fit lines or curves to edge points.

The system is also equipped with a 5-way touch-trigger probe (figure 7) which is mounted on the measuring machine. The stylus or tip is mounted directly into the probe and is the part of the measuring system that makes contact with the component. The probe acts as an on/off switch which freezes the readings of the x, y, and z machine scales when the probe tip touches the part surface. It measures in both the positive and negative x and y directions and in the positive z direction. The system provides software for calibrating the touch probe in order to convert the electrical signals into three dimensional world coordinates

Data from the touch probe and the camera system are predominantly competitive in terms of the sensory interaction between the two sources of information. Both sensors measure three dimensional position of the feature. Because of the very small field of view afforded by the magnification lenses, global information is not available.

In the VIEW system, measurements from the camera and from the touch probe are interchangeable for most tasks, and the accuracy of the readings is equivalent [13]. The exceptions to this statement are discussed below. Calibration standards and algorithms are provided for determining the 2-D to 3-D transformations between image coordinates and world coordinates. The computed calibration constants are specific to the lens being used. The system also contains code for computing the radius of the probe stylus in order to interpret probe readings relative to the probe center. A cross calibration algorithm is provided for determining the offsets between vision system measurements and touch probe measurements. In this way, measurements from either sensor are expressed in a common coordinate system.

Sensor Characteristics

As mentioned earlier, there are cases where one sensor is better suited to a measuring task than another. An understanding of the strengths and weaknesses of the sensors enables the user to

choose the sensor whose data has the higher level of confidence in a given situation. One such example occurs when poor viewing conditions exist. Camera accuracy is affected by the quality of the image viewed and is degraded by poor contrast, specularity, shadows, and image noise. Specularity is a common problem when measuring machined parts using computer vision because the machined surfaces are often highly polished. Factors contributing to image noise are object surface texture, part contamination, dirt particles, surface scratches, etc. When these conditions are present, the level of confidence in the processed camera data decreases. The level of confidence of data from the touch probe is not as affected by these conditions, and thus the probe would be used under such circumstances.

Similarly, any situation in which the feature of interest is either partially or fully occluded in the field of view is not suited for camera measurement. Occlusion can occur when an object feature overlaps another feature in the x-y plane but not in the z plane. Touch probe measurements can be used when visual occlusion is present providing that there is sufficient clearance for the probe to contact the occluded feature.

Another example where probe data is likely to be more accurate that camera data occurs when measuring the depth of a part containing a vertical cavity such as a cylinder or a deep threaded nut. The camera is unable to auto-focus inside the hole and cannot accurately measure the height of the object. The touch probe would be the sensor of choice to measure such features.

The probe cannot however provide information about edges, and thus burrs and other irregularities are measurable only by the vision system. The large number of data points available from vision also makes measurement of shape features such as straightness and roundness much easier. In addition, situations can exist where the touch probe cannot be used, and in those cases, camera information is available. For example, consider a part consisting of a surface with a hole drilled in it. If the diameter of the hole is smaller than the diameter of the probe stylus, the touch probe cannot be used. Camera data could be used to determine the dimensions of the drilled cavity although the accuracy of the measurements in this case would be very dependent on the depth of

the hole and lighting conditions.

In the current mode of operation, the user locates the feature to be measured, uses the image processing tools to compute the position of the feature, and then commands the touch probe to that position for a contact reading. Currently, experiments are being done to determine the relationship between three dimensional readings of the same part from each of the two sensors. It is believed that x, y, and z offsets can be determined between camera position data and touch probe data which will result in correlated repeatable and reliable readings. Additional tests are planned to quantify the repeatability and uniformity of the measurements.

Although the combined use of vision and touch sensors for measurement and inspection tasks as described in this section is a positive step towards integrating information from multiple sources, the limitations placed on the user of the system prevent him from taking full advantage of the available resources. Integration of the sensory output is performed by the operator rather than by the system. The following section suggests alternative strategies for using these sensors to better advantage.

Future Strategies

The measuring system described above is designed as a competitive information interaction system. As shown in figure 2, the positional information obtained from the camera and the touch probe can be treated separately or can be fused into a single output. Without a plan for sensory integration, knowledge of the strengths and weaknesses of the sensors is not incorporated into the computation of the three dimensional position. In the latter case, difficult issues associated with sensor fusion must be addressed. In this section, we propose alternative uses of vision and touch sensors which would result in a cooperative information interaction system.

Cooperative interaction offers great promise in sensor integration techniques and lends itself to the task of inspection of manufactured parts. The types of features measured during machine inspection are divided into two classes: edge properties and area properties. Edge properties define the position of points, lines, or circles that represent part features. Area properties define the area of

a region, the diameter of a circular region, and the center-of-mass of the region. By concentrating on the global nature of visual processing and the local nature of touch probe processing, the cooperative interaction between these sensors can be used to extract the required features in an accurate and efficient manner. The effective use of touch requires that the sensor be directed to points of interest. Without some form of guidance, touch sensing is both too slow and too difficult to interpret. Global image processing can extract features of interest from an image and convert the locations of those features to world coordinates. The feature coordinates can then be used to guide the touch probe.

The remainder of this section discusses alternative strategies for using multiple sensors in measuring machines. In all cases, feature extraction capabilities would benefit from lower magnification camera lenses. Although camera accuracy is reduced when lower magnification lenses are used, a camera's main advantages -- its speed and ability to quickly extract global information-- are unaffected. The strength of the touch probe is its ability to accurately measure surface points. Using the strengths of each sensor in a cooperative interaction would result in an intelligent approach to feature extraction and measurement.

One modification to the existing CMM entails the addition of a low-magnification camera lens which would allow the camera to be used to greater advantage as a global sensor. The camera is currently being used as a local sensor because of the very small field of view afforded by the magnification lenses. A non-magnification lens cannot be used because of the geometry of the camera mounting on the CMM. However, a low-magnification lens would result in a greater field of view and thus would increase the global information available from the camera. In this mode, global edge point information can be extracted from the image and features of interest on the surface of the part can be identified. A simple example is the inspection of a hole. There is considerable research [14] that shows that the number and positions of touch points needed to gage a hole depends on its shape. With a measurement of the hole shape at low resolution, the system could select the appropriate touch pattern for the probe.

An alternative to the above scenario involves the addition of a second camera to the system. As part of a competitive information system, the existing camera and touch probe can be used interchangeably, or, under specific conditions, the operator can choose the sensor with the greater level of confidence for a given situation. The second camera would have a non-magnification lens and could be mounted in a fixed location above the base of the measuring machine. Its field of view would encompass the entire part being measured. The camera signals would be digitized and sent to a programmable image processing system. Such a system allows the user to experiment and develop algorithms suitable for a particular application. This form of operation can be considered to be an off-line teach mode in a manufacturing environment. It consists of software tools for learning about specific situations and studying the effects of camera noise, lighting conditions, part reflectance, shadows, etc. The output of this process consists of an algorithm capable of extracting the required feature measurements. This algorithm can be programmed into the system and used in a cooperative sense in a production mode to guide the actions of either (or both) of the other sensors. The sensor processing architecture design discussed earlier would be incorporated into this system. Measurements would be guided by the description of the part provided by the CAD design.

A third and more sophisticated sensor system for measurement would incorporate closed loop feedback between an actively controlled camera and the touch probe. An active vision system has "mechanisms that can actively control camera parameters such as position, orientation, focus, zoom and aperture...in response to the requirements of the task [10]." In this configuration, the results of the image processing algorithm could be used to either direct the motion of the touch probe or to reposition the camera in order to permit a better viewing angle. Figure 8 illustrates such a system. The touch-trigger probe placement is guided by the output of the image processing algorithm and produces accurate three-dimensional position data, $P_{x,y,z}$. When the image processing algorithm fails to compute a motion command for the touch probe because of insufficient image information, it generates a command to move the camera in order to obtain a better view of the scene.

Conclusion

We have discussed some of the issues associated with the use of multiple sensors in a manufacturing inspection task. In particular, we have described the strengths and weaknesses of cameras and touch-trigger probes relative to their use in a measuring machine. The interaction of elements of a multi-sensory system have been characterized. We have described the CMM in use at NIST and have offered alternative uses of sensory information to improve its effectiveness in the inspection task.

References

[1] Albus, J. S., H. G. McCain, R. Lumia.,"NASA/NBS Standard Reference Model for Telerobot Control System Architecture (NASREM)," NIST Technical Note 1235, Gaithersburg, MD, July, 1987.

[2] Chaconas, K., M. Nashman, "Visual Perception Processing in a Hierarchical Control System," NIST Technical Note 1260, Gaithersburg, MD, March, 1989.

[3] 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.

[4] T. D. Garvey, "A Survey of AI Approaches to the Integration of Information," Proceedings SPIE, Vol. 782, Infrared Sensors and Sensor Fusion, R. G. Buser and F. B. Warren, Eds., Orlando, FL, May 1987, pp 68-82.

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

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

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

[8] 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, pp 1941-1946.

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

[10] M. J. Swain, M. Stricker, Eds., “Promising Directions in Active Vision”, NSF Active Vision Workshop, Chicago, Ill., 1991.

[11] 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.

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

[13] Conversation with L. Savage, Technical Representative, VIEW Engineering Inc.

[14] C. Plazek, Ed., “Efforts to Relate CMM Data to Design Specifications Underway”, Quality Magazine, December, 1988.

Biography

Marilyn Nashman received the B.S. degree in mathematics from the University of the City of New York in 1961. Since 1979, she has been with the Robot Systems Division of the National Institute of Standards and Technology specializing in real-time hierarchical image processing.

List of Figures

Figure 1. Hierarchical Sensory Processing System

Figure 2. Sensor Integration Process

Figure 3. Competitive Information Interaction

Figure 4. Complementary Information Interaction

Figure 5. Cooperative Information Interaction

Figure 6. Data Gathering Unit

Figure 7. Touch-Trigger Probe

Figure 8. Active Sensing System

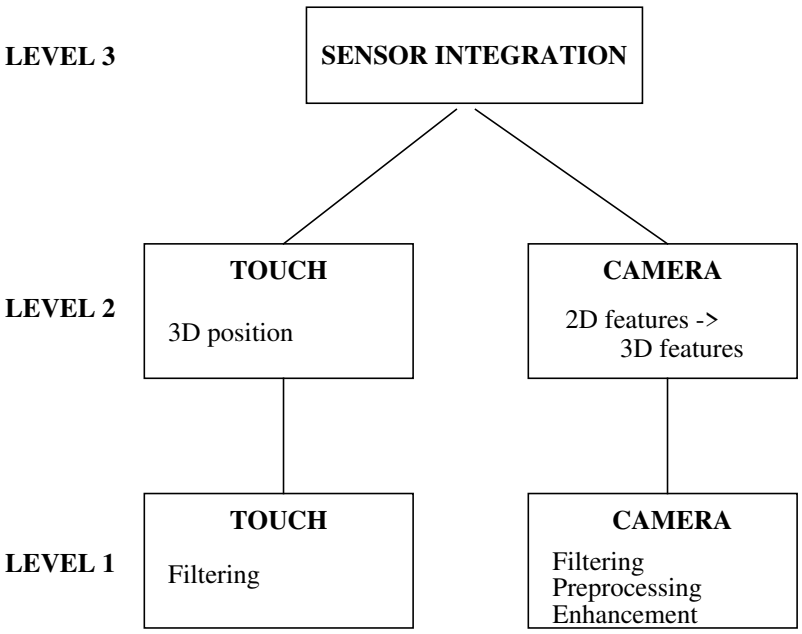

Figure 1. Hierarchical Sensory Processing System

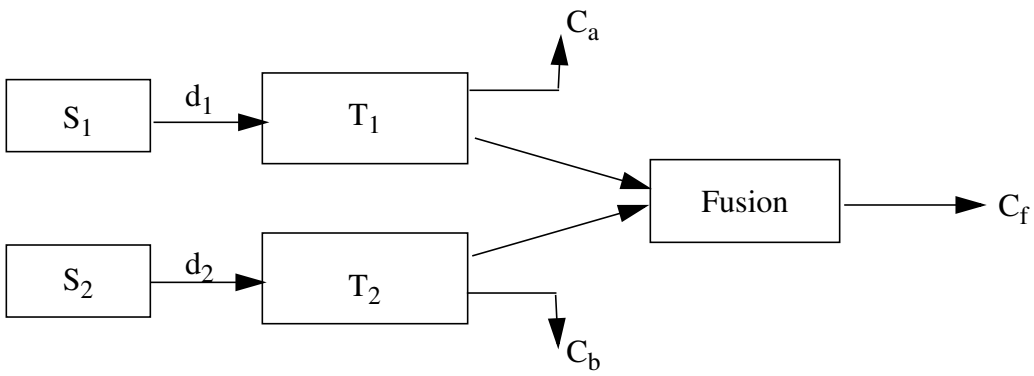

Figure 2. Sensor Integration Process

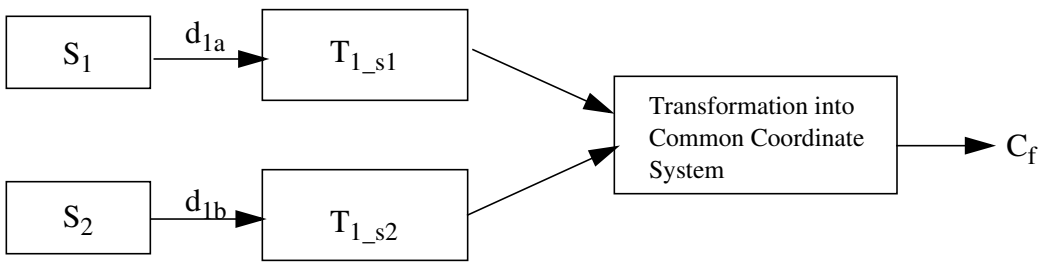

Figure 3. Competitive Information Interaction

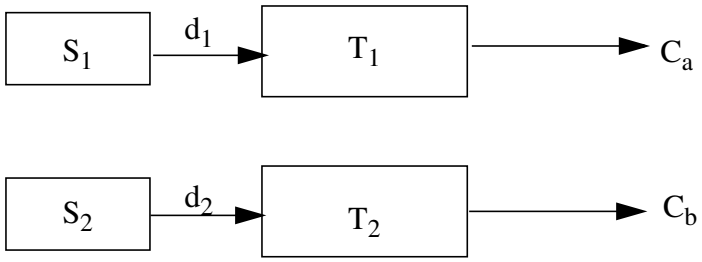

Figure 4. Complementary Information Interaction

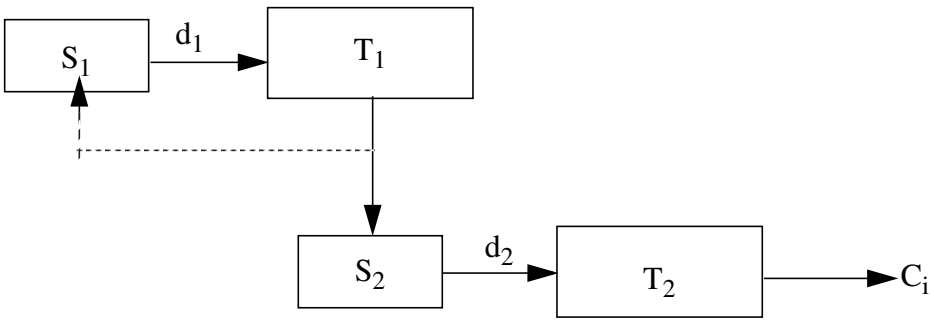

Figure 5. Cooperative Information Interaction

Figure 6. Data Gathering Unit

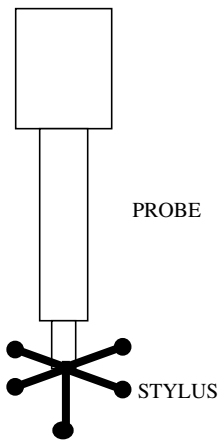

Figure 7. Touch-Trigger Probe

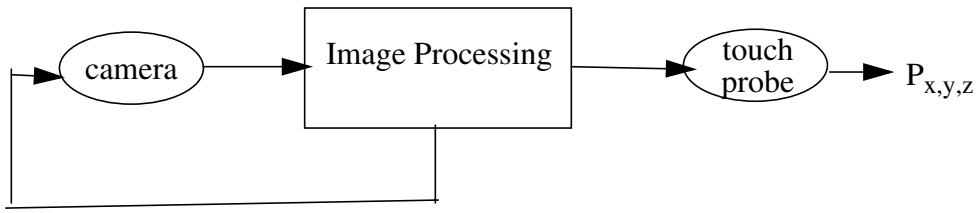

Figure 8. Active Sensing System

Marilyn Nashman is a computer scientist working for the Robot Systems Division specializing in real-time hierarchical image processing. She has worked on projects funded by the Automated Manufacturing Research Facility (AMRF), the Army (TEAM), NASA (Telerobotic Control System Architecture), the Intelligent Machines Initiative (IMI) and the Intelligent Vehicle/Highway Systems program(IVHS). Her research interests include real-time image processing and sensor integration.