Dolan, J.M., Friedman, M.B., Nagurka, M.L., and Gotow, J.K., "Gestural Control of Industrial Robots: An Application to Surgical Instrument Positioning," Advanced Topics in Manufacturing Technology: Product Design, Bioengineering, and Space Commercialization, *ed.* Francis, P.H., ASME, December 1987, pp. 29-34.

GESTURAL CONTROL OF INDUSTRIAL ROBOTS:

An Application to Surgical Instrument Positioning

J.M. Dolan, M.B. Friedman*, M.L. Nagurka, and J.K. Gotow

Department of Mechanical Engineering
and *The Robotics Institute
Carnegie-Mellon University
Pittsburgh, PA 15213

INTRODUCTION

A major reason that robotic systems are not viable substitutes for human workers in many tasks is not due to limitations in the strength or dexterity of robotic manipulators, but rather due to the inability of machine guidance systems to perform in unstructured environments. Machine vision systems, which use the most sophisticated of presently available sensor and signal processing systems, are primitive compared to human perceptual capabilities.

It is often impractical (because of economic and physical limitations) to structure a task so that industrial vision systems can segment out, let alone recognize, relevant objects. Given the state of the art of image segmentation research -- both the statistical and the heuristic approaches -- it is unlikely that machine vision systems will be capable of working in unconstrained environments in the near future. In such environments, it would be ideal if humans could rapidly, easily, and reliably designate locations and pathways to industrial robots in machine-interpretable coordinates. Then, integrated human-robot interactive systems could be realized that are both more cost-effective than fully automated systems and more productive than current systems based on job segmentation between manual labor and fixed automation.

Constrained to use in highly structured environments, most of today's industrial robots do highly repetitive tasks that require no human intervention for extended periods of time. Since the destination of the end effector of such a robot must be modified only occasionally, it is efficient to have a human operator reprogram the robot controller. In contrast, where daily tasks are highly varied and dependent on operator perception and judgement, robots have been excluded. We are investigating the use of gesture to specify "where" and in "what orientation" a robotic action is to be performed while voice or a keypad is used to determine "which" pre-programmed subroutine is to be executed by the robotic tool at the specified site. We are evaluating the relative advantages of voice and gesture for modifying end-effector position.

We believe that the combination of gesture and voice for robot control will allow skilled personnel to efficiently and productively supervise multiple robotic tools working on non-repetitive tasks that have previously been resistant to automation. Skilled humans working interactively with industrial robots may be more productive than either agent working independently. Before considering a specialized application niche - surgical robotics - we will review the history of the development of modern robotic control systems to see why robots are usually viewed as substitutes for, rather than enhancers of, skilled human performance.

HISTORY

In the early 1950's, servo motors, electronic controls, and digital memory were added to standard milling machine designs to make the first numerically controlled (NC) milling machines. The early NC machines were not developed so much for their labor saving properties as for their improved accuracy and uniformity in machined parts production.

Preceding the development of NC machine tools, however, servo driven remote actuators were used to handle radioactive materials. The earliest of these telemanipulators used mechanical linkages to allow an operator to remotely control a six degree-of-freedom (DOF) manipulators from behind the safety of a radiation shield. Later, electric motor powered "slave" manipulators were developed to move and position containers of radiochemicals based on information encoded from the joints of a human-controlled "master" linkage located at a safely remote site. The operator relied solely on visual feedback to guide the first servo powered teleoperators since force feedback was eliminated when the mechanical linkage between master and slave was removed.

By using the same digital information storage technology that was adapted a few years later to control the servo powered NC milling machines, teleoperators could have been developed to operate remotely in time as well as in space -- true robotic manipulators. Instead, because of the critical nature of the radiochemistry being performed using teleoperators, development efforts were concentrated on giving real-time force feedback to the human operator to increase the operator's sensitivity and precision of control. To maintain human decision making in the "loop," there was no attempt to develop autonomous (*i.e.*, pre-programmed) operation.

It was not until the late 1950's and early 1960's that programmable articulated servo powered linkages, *i.e.*, robots, were developed and patented. Early industrial robots were driven by hand through a sequence of positions which were recorded in digital memory. As in the NC milling machines, the digital program could be replayed to drive the servo positioning motors in a precisely repeatable pattern. In routine operation, there was no interaction with humans and only limited interaction between the robot and the workpieces it manipulated. The commercial success of robots depended on their application to jobs where repeatability and reliability of workpiece placement were of paramount importance. In tasks where workpiece position could be absolutely defined, the advantages of repeatability surpassed those of adaptability.

GESTURAL CONTROL

There are several ways of combining the desirable features inherent in NC technology and teleoperators. In the development of industrial robotics, the playback capability of digitally stored programs was coupled with the configurational flexibility of telemanipulators. An alternative approach is to use the computational power of digital processing to exploit the unmatched sensory capabilities of the human. We have adopted this strategy of merging the advantages of robotic controllers with the real-time human control of telemanipulators, and call this interactive approach **gestural control**. This name stems from its association with using body motions (gestures) to guide the operation of mechanical systems.

Gestural control has many potential applications, from the industrial manufacturing environment to the surgical theatre. In the former case, gestural control may be used to input new control points, by using a laser pointer, for example, to designate specific locations where a task is to be performed. In the latter application, gestural control may be used to provide continuous positioning of surgical instruments, as described below. Thus, gestural control can supply occasional inputs to a preprogrammed, semi-autonomous operation, or provide real-time, dedicated position feedback.

Surgical Robotics

Surgical microscopes are currently used in a significant fraction (about 20 percent) of surgical procedures. A typical microscope weighs 10 kg, is stereoscopic (binocular), and has magnification of 10x - 20x. With the necessary auxiliary equipment (secondary viewing tube, laser, camera, *etc.*) the total weight can be as great as 20 kg.

In order to reduce stress on the surgeon, speed surgery, and convey timely visual information about the operating site, it is desirable to have a flexible and unobtrusive way of repositioning and refocusing the microscope. Current methods are cumbersome, requiring either manual readjustment, which slows surgery, or the use of voice control or foot pedals to command individual axis motions, which makes it difficult to attain complex orientations. Neither method allows the surgeon to easily perform such actions as scanning along a blood vessel or nerve while doing continuous repair, or rapidly pivoting about a fixed focus, say, to determine the full extent of a tumor.

The alternative which suggests itself in the context of this paper is to continuously sense the surgeon's head position and orientation (hereafter referred to as position) and to reposition the microscope accordingly, thus exercising a form of gestural control. We have implemented such a system using an ISOTRAK low-frequency magnetic sensor by Polhemus Navigation Sciences, Inc. to obtain head position and a Merlin six DOF robot by American CIMFLEX to hold the microscope and track the sensed position. (See Figure 1.) The microscope and sensor system were also implemented on a General Electric P-50, a five DOF electric robot. The Merlin offers the advantage of being counter-balanced and of having an additional DOF.

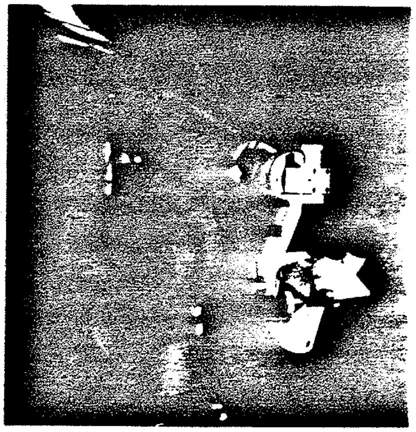

Figure 1. Photograph of Surgical Microscope Mounted on Industrial Robot. (Auxiliary Microscope Equipment Absent.)

There are three basic modes of operation: 1) tracking, in which the microscope duplicates head orientation at a desired fixed (presumably small) distance from the eyes as the surgeon scans an area; 2) pivoting, in which the microscope pivots about a fixed focus as the surgeon seeks different views of the operating site; and 3) Cartesian, which permits fine adjustments. The various modes are invoked under voice control to afford the surgeon maximum freedom of motion.

Determination of the appropriate microscope position is made by assigning Cartesian frames to points on the microscope and the user's head and performing the transformations necessary to move from one frame to the next. (See Figure 2.) Each frame is represented by a 4 x 4 matrix containing 9 direction cosines and 3 position coordinates. The origin of the frame of ultimate interest (frame M) is located midway between the microscope eyepieces in the plane formed by the upper surfaces of the eyepieces. Frames of secondary interest are associated with a point midway between the user's eyes (frame E), and with the magnetic sensor (frame S), which may be attached to the head in an arbitrary position, but must remain fixed with respect to frame E. Once frame S has been sensed, a series of relative transformations results in frame M.

The proposed scheme requires three preliminary steps. First, the robot must be taught a tool position for the microscope. The tool tip should clearly be located at the origin of frame M. Although the tool orientation is arbitrary in principle, it is computationally and conceptually convenient to align one axis (say, the x-axis) with the binocular tubes to which the eyepieces attach, and one axis (say, the y-axis) along a line connecting the eyepieces. The z-axis is then given by the cross-product of the x- and y-axes. Once the tool is taught, the robot can be directly commanded to position frame M rather than the default frame associated with its faceplate. A further necessary piece of information which is also dependent on the microscope geometry is the vector position f of the focal point in M-coordinates. We will refer to f as the focal vector.

Second, a global reference frame G must be chosen and taught both to the robot and to the magnetic source to which the sensor is referenced. The location of this frame within the robot's work envelope is arbitrary, but placing it centrally in the area of operation of the microscope will minimize errors due to inexact correspondence between the taught source and robot frames. Note that frame G remains stationary, whereas frames S,E, and M move about, and are expressed in G-coordinates.

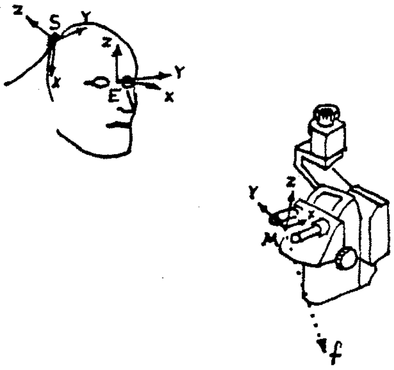

Figure 2. Schematic of S, E, and M Coordinate Frames.

Third, the constant transformation from frame S to frame E for a given wearing of the sensor must be found. This calibration permits the sensor to be attached to the head in any position, and allows calculation of frame E for a given sensor position S by matrix postmultiplication. First, we choose the x- and y-axes of frame E to be directed along the outward normal to the plane of the user's face (the "direction of gaze"[1]) and along a line connecting the user's eyes, respectively. If the microscope is then positioned so that frame M coincides with frame G, and the user assumes normal viewing posture with his eyes resting on the eyepieces, then frame E also coincides with frame G, and the transformation ${}^{S}E$ (*i.e.*, frame E expressed in S-coordinates) can be obtained by sensing frame S, in this case in E-coordinates, and inverting that transformation.

Once these steps have been carried out, it remains to implement the tracking, pivoting, and Cartesian modes mentioned above. In the case of tracking, the microscope frame M duplicates the eye frame E's orientation at a distance D along the direction of gaze. The user specifies D, and the fixed transformation ${}^{S}M = {}^{S}E * DX$ is calculated, where DX is a matrix representing a translation D along the x-axis, and $*$ represents matrix multiplication. Frame M is then continuously updated using $M = S * {}^{S}M$. The direction cosines in M must be converted to Euler angles in order to specify a move to the robot. If D is set to zero, then frames E and M are coincident, and the eyes ideally rest directly on the eyepieces during tracking.

Pivoting requires that the microscope maintain a fixed focus while frame M follows the motion of frame E as closely as possible. The problem is overconstrained in that it is no longer possible for M to exactly follow the orientation of an arbitrary E and still maintain focus. Nevertheless, we can use the most important information from E to construct M. If we note that M's origin is constrained to lie on the surface of a sphere of radius $|f|$ centered on the current focal point, and make the natural assumption that the user's eyes will be close to the surface of this sphere, then M's origin should be placed at the intersection of the focal sphere with the line segment joining E's origin and the current focal point. In order to align the eyepieces as nearly as possible with the user's eyes, frame E's y-axis is resolved within the plane it forms with the focal vector to give an M-frame y-axis which agrees with the relationship between frame M and the focal vector determined in the first preliminary step above. Once the y-axis is known, the same relationship uniquely determines the x- and z-axes.

Cartesian adjustments of position can be accomplished in any desired frame C by defining C with respect to the microscope, teaching and selecting a corresponding tool, and instructing the robot to move relative to C. A useful frame is one having an axis along which the zoom function can be performed. This mode is envisioned as being best suited for fine motion once a desired position has been nearly attained. The current implementation allows for motion in six Cartesian directions and specification of increment size under voice control.

The repeatability of the sensor and robot are below the optical resolution of the microscope, so that focus is well-maintained as long as the sensor-robot frame correspondence is good. Position errors due to inexact correspondence are typically on the order of 2 percent of the current distance from the origin of frame G. The current position update rate is 2.5 Hz, which is sufficiently responsive for changing views in pivoting mode, but is probably too jerky for user comfort in tracking mode. We are hoping to significantly increase this update rate with a newer version of the Merlin controller.

CONCLUSIONS

Modern industrial robot systems evolved from the combination of the controls developed for NC machine tool systems, emphasizing automatic repeatability, and general purpose manipulators similar to teleoperated systems, where mechanical configuration was optimized to take advantage of direct human control. Current industrial practice has fostered robotic systems based on the flexible machine tool model. In the next step of robot evolution, human sensory capabilities need to be used effectively in concert with computer control feedback loops.

This paper shows that industrial robot control systems can be complemented by direct human guidance to enhance the performance of skilled human workers. Gestural control exploits the processing power of the robot controller to provide the human operator with an intuitive method of interaction. The application of controlling a surgical microscope via gestural control has demonstrated that the implementation does not tax available instrumentation and promises clinical significance.

---

[1] Although the eyes can clearly gaze in directions other than straight ahead, this is the only way to conveniently look through a microscope.

GESTURAL CONTROL OF INDUSTRIAL ROBOTS: AN APPLICATION TO SURGICAL INSTRUMENT POSITIONING

J. M. Dolan, M. B. Friedman* M. L. Nagurka, and J. K. Gotow
Department of Mechanical Engineering and
*The Robotics Institute
Carnegie-Mellon University
Pittsburgh, Pennsylvania

INTRODUCTION

A major reason that robotic systems are not viable substitutes for human workers in many tasks is not due to limitations in the strength or dexterity of robotic manipulators, but rather due to the inability of machine guidance systems to perform in unstructured environments. Machine vision systems, which use the most sophisticated of presently available sensor and signal processing systems, are primitive compared to human perceptual capabilities.

It is often impractical (because of economic and physical limitations) to structure a task so that industrial vision systems can segment out, let alone recognize, relevant objects. Given the state of the art of image segmentation research -- both the statistical and the heuristic approaches -- it is unlikely that machine vision systems will be capable of working in unconstrained environments in the near future. In such environments, it would be ideal if humans could rapidly, easily, and reliably designate locations and pathways to industrial robots in machine-interpretable coordinates. Then, integrated human-robot interactive systems could be realized that are both more cost-effective than fully automated systems and more productive than current systems based on job segmentation between manual labor and fixed automation.

Constrained to use in highly structured environments, most of today's industrial robots do highly repetitive tasks that require no human intervention for extended periods of time. Since the destination of the end effector of such a robot must be modified only occasionally, it is efficient to have a human operator reprogram the robot controller. In contrast, where daily tasks are highly varied and dependent on operator perception and judgement, robots have been excluded. We are investigating the use of gesture to specify "where" and in "what orientation" a robotic action is to be performed while voice or a keypad is used to determine "which" pre-programmed subroutine is to be executed by the robotic tool at the specified site. We are evaluating the relative advantages of voice and gesture for modifying end-effector position.

We believe that the combination of gesture and voice for robot control will allow skilled personnel to efficiently and productively supervise multiple robotic tools working on non-repetitive tasks that have previously been resistant to automation. Skilled humans working interactively with industrial robots may be more productive than either agent working independently. Before considering a specialized application niche - surgical robotics - we will review the history of the development of modern robotic control systems to see why robots are usually viewed as substitutes for, rather than enhancers of, skilled human performance.

HISTORY

In the early 1950's, servo motors, electronic controls, and digital memory were added to standard milling machine designs to make the first numerically controlled (NC) milling machines. The early NC machines were not developed so much for their labor saving properties as for their improved accuracy and uniformity in machined parts production.

Preceding the development of NC machine tools, however, servo driven remote actuators were used to handle radioactive materials. The earliest of these telemanipulators used mechanical linkages to allow an operator to remotely control a six degree-of-freedom (DOF) manipulators from behind the safety of a radiation shield. Later, electric motor powered "slave" manipulators were developed to move and position containers of radiochemicals based on information encoded from the joints of a human-controlled "master" linkage located at a safely remote site. The operator relied solely on visual feedback to guide the first servo powered teleoperators since force feedback was eliminated when the mechanical linkage between master and slave was removed.

By using the same digital information storage technology that was adapted a few years later to control the servo powered NC milling machines, teleoperators could have been developed to operate remotely in time as well as in space -- true robotic manipulators. Instead, because of the critical nature of the radiochemistry being performed using teleoperators, development efforts were concentrated on giving real-time force feedback to the human operator to increase the operator's sensitivity and precision of control. To maintain human decision making in the "loop," there was no attempt to develop autonomous (*i.e.*, pre-programmed) operation.

It was not until the late 1950's and early 1960's that programmable articulated servo powered linkages, *i.e.*, robots, were developed and patented. Early industrial robots were driven by hand through a sequence of positions which were recorded in digital memory. As in the NC milling machines, the digital program could be replayed to drive the servo positioning motors in a precisely repeatable pattern. In routine operation, there was no interaction with humans and only limited interaction between the robot and the workpieces it manipulated. The commercial success of robots depended on their application to jobs where repeatability and reliability of workpiece placement were of paramount importance. In tasks where workpiece position could be absolutely defined, the advantages of repeatability surpassed those of adaptability.

GESTURAL CONTROL

There are several ways of combining the desirable features inherent in NC technology and teleoperators. In the development of industrial robotics, the playback capability of digitally stored programs was coupled with the configurational flexibility of telemanipulators. An alternative approach is to use the computational power of digital processing to exploit the unmatched sensory capabilities of the human. We have adopted this strategy of merging the advantages of robotic controllers with the real-time human control of telemanipulators, and call this interactive approach **gestural control.** This name stems from its association with using body motions (gestures) to guide the operation of mechanical systems.

Gestural control has many potential applications, from the industrial manufacturing environment to the surgical theatre. In the former case, gestural control may be used to input new control points, by using a laser pointer, for example, to designate specific locations where a task is to be performed. In the latter application, gestural control may be used to provide continuous positioning of surgical instruments, as described below. Thus, gestural control can supply occasional inputs to a preprogrammed, semi-autonomous operation, or provide real-time, dedicated position feedback.

Surgical Robotics

Surgical microscopes are currently used in a significant fraction (about 20 percent) of surgical procedures. A typical microscope weighs 10 kg, is stereoscopic (binocular), and has magnification of 10x - 20x. With the necessary auxiliary equipment (secondary viewing tube, laser, camera, *etc.*) the total weight can be as great as 20 kg.

In order to reduce stress on the surgeon, speed surgery, and convey timely visual information about the operating site, it is desirable to have a flexible and unobtrusive way of repositioning and refocusing the microscope. Current methods are cumbersome, requiring either manual readjustment, which slows surgery, or the use of voice control or foot pedals to command individual axis motions, which makes it difficult to attain complex orientations. Neither method allows the surgeon to easily perform such actions as scanning along a blood vessel or nerve while doing continuous repair, or rapidly pivoting about a fixed focus, say, to determine the full extent of a tumor.

The alternative which suggests itself in the context of this paper is to continuously sense the surgeon's head position and orientation (hereafter referred to as position) and to reposition the microscope accordingly, thus exercising a form of gestural control. We have implemented such a system using an ISOTRAK low-frequency magnetic sensor by Polhemus Navigation Sciences, Inc. to obtain head position and a Merlin six DOF robot by American CIMFLEX to hold the microscope and track the sensed position. (See Figure 1.) The microscope and sensor system were also implemented on a General Electric P-50, a five DOF electric robot. The Merlin offers the advantage of being counter-balanced and of having an additional DOF.

There are three basic modes of operation: 1) tracking, in which the microscope duplicates head orientation at a desired fixed (presumably small) distance from the eyes as the surgeon scans an area; 2) pivoting, in which the microscope pivots about a fixed focus as the surgeon seeks different views of the operating site; and 3) Cartesian, which permits fine adjusments. The

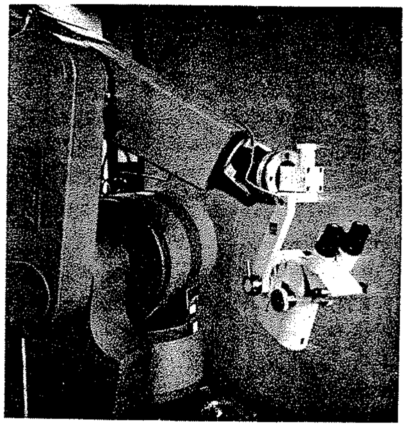

Figure 1. Photograph of Surgical Microscope Mounted on Industrial Robot. (Auxiliary Microscope Equipment Absent.)

various modes are invoked under voice control to afford the surgeon maximum freedom of motion.

Determination of the appropriate microscope position is made by assigning Cartesian frames to points on the microscope and the user's head and performing the transformations necessary to move from one frame to the next. (See Figure 2.) Each frame is represented by a 4 x 4 matrix containing 9 direction cosines and 3 position coordinates. The origin of the frame of ultimate interest (frame M) is located midway between the microscope eyepieces in the plane formed by the upper surfaces of the eyepieces. Frames of secondary interest are associated with a point midway between the user's eyes (frame E), and with the magnetic sensor (frame S), which may be attached to the head in an arbitrary position, but must remain fixed with respect to frame E. Once frame S has been sensed, a series of relative transformations results in frame M.

The proposed scheme requires three preliminary steps. First, the robot must be taught a tool position for the microscope. The tool tip should clearly be located at the origin of frame M. Although the tool orientation is arbitrary in principle, it is computationally and conceptually convenient to align one axis (say, the x-axis) with the binocular tubes to which the eyepieces attach, and one axis (say, the y-axis) along a line connecting the eyepieces. The z-axis is then given by the cross-product of the x- and y-axes. Once the tool is taught, the robot can be directly commanded to position frame M rather than the default frame associated with its faceplate. A further necessary piece of information which is also dependent on the microscope geometry is the vector position f of the focal point in M-coordinates. We will refer to f as the focal vector.

Second, a global reference frame G must be chosen and taught both to the robot and to the magnetic source to which the sensor is referenced. The location of this frame within the robot's work envelope is arbitrary, but placing it centrally in the area of operation of the microscope will minimize errors due to inexact correspondence between the taught source and robot frames. Note that frame G remains stationary, whereas frames S,E, and M move about, and are expressed in G-coordinates.

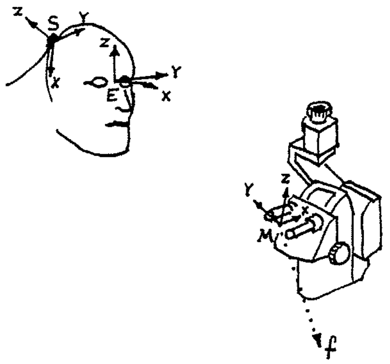

Figure 2. Schematic of S, E, and M Coordinate Frames.

Third, the constant transformation from frame S to frame E for a given wearing of the sensor must be found. This calibration permits the sensor to be attached to the head in any position, and allows calculation of frame E for a given sensor position S by matrix postmultiplication. First, we choose the x- and y-axes of frame E to be directed along the outward normal to the plane of the user's face (the "direction of gaze"[1]) and along a line connecting the user's eyes, respectively. If the microscope is then positioned so that frame M coincides with frame G, and the user assumes normal viewing posture with his eyes resting on the eyepieces, then frame E also coincides with frame G, and the transformation ^{S}E (*i.e.*, frame E expressed in S-coordinates) can be obtained by sensing frame S, in this case in E-coordinates, and inverting that transformation.

Once these steps have been carried out, it remains to implement the tracking, pivoting, and Cartesian modes mentioned above. In the case of tracking, the microscope frame M duplicates the eye frame E's orientation at a distance D along the direction of gaze. The user specifies D, and the fixed transformation $^{S}M = {}^{S}E * DX$ is calculated, where DX is a matrix representing a translation D along the x-axis, and * represents matrix multiplication. Frame M is then continuously updated using $M = S * {}^{S}M$. The direction cosines in M must be converted to Euler angles in order to specify a move to the robot. If D is set to zero, then frames E and M are coincident, and the eyes ideally rest directly on the eyepieces during tracking.

---

[1] Although the eyes can clearly gaze in directions other than straight ahead, this is the only way to conveniently look through a microscope.

Pivoting requires that the microscope maintain a fixed focus while frame M follows the motion of frame E as closely as possible. The problem is overconstrained in that it is no longer possible for M to exactly follow the orientation of an arbitrary E and still maintain focus. Nevertheless, we can use the most important information from E to construct M. If we note that M's origin is constrained to lie on the surface of a sphere of radius $|f|$ centered on the current focal point, and make the natural assumption that the user's eyes will be close to the surface of this sphere, then M's origin should be placed at the intersection of the focal sphere with the line segment joining E's origin and the current focal point. In order to align the eyepieces as nearly as possible with the user's eyes, frame E's y-axis is resolved within the plane it forms with the focal vector to give an M-frame y-axis which agrees with the relationship between frame M and the focal vector determined in the first preliminary step above. Once the y-axis is known, the same relationship uniquely determines the x- and z-axes.

Cartesian adjustments of position can be accomplished in any desired frame C by defining C with respect to the microscope, teaching and selecting a corresponding tool, and instructing the robot to move relative to C. A useful frame is one having an axis along which the zoom function can be performed. This mode is envisioned as being best suited for fine motion once a desired position has been nearly attained. The current implementation allows for motion in six Cartesian directions and specification of increment size under voice control.

The repeatability of the sensor and robot are smaller than the depth of field of the microscope, so that focus is well-maintained as long as the sensor-robot frame correspondence is good. Position errors due to inexact correspondence are typically on the order of 2 percent of the current distance from the origin of frame G. The current position update rate is 2.5 Hz, which is sufficiently responsive for changing views in pivoting mode, but may be too jerky for comfortable use in tracking mode. We are hoping to significantly increase this update rate with a newer version of the Merlin controller.

CONCLUSIONS

Modern industrial robot systems evolved from the combination of the controls developed for NC machine tool systems, emphasizing automatic repeatability, and general purpose manipulators similar to teleoperated systems, where mechanical configuration was optimized to take advantage of direct human control. Current industrial practice has fostered robotic systems based on the flexible machine tool model. In the next step of robot evolution, human sensory capabilities need to be used effectively in concert with computer control feedback loops.

This paper shows that industrial robot control systems can be complemented by direct human guidance to enhance the performance of skilled human workers. Gestural control exploits the processing power of the robot controller to provide the human operator with an intuitive method of interaction. The application of controlling a surgical microscope via gestural control has demonstrated that the implementation does not tax available instrumentation and promises clinical significance.