US 20050033580A1

(19) United States

(12) Patent Application Publication (10) Pub. No.: US 2005/0033580 A1

Wang et al. (43) Pub. Date: Feb. 10, 2005

(54) SPEECH INTERFACE FOR AN AUTOMATED ENDOSCOPE SYSTEM

(75) Inventors: **Yulun Wang**, Goleta, CA (US); **Darrin Uecker**, Santa Barbara, CA (US)

Correspondence Address:
TOWNSEND AND TOWNSEND AND CREW, LLP
TWO EMBARCADERO CENTER
EIGHTH FLOOR
SAN FRANCISCO, CA 94111-3834 (US)

(73) Assignee: **Computer Motion, Inc.**, Sunnyvale, CA

(21) Appl. No.: **10/942,374**

(22) Filed: **Sep. 16, 2004**

Related U.S. Application Data

(63) Continuation of application No. 10/095,488, filed on Mar. 11, 2002, which is a continuation of application No. 08/310,665, filed on Sep. 22, 1994, now Pat. No. 6,463,361.

Publication Classification

(51) Int. Cl.[7] G10L 11/00
(52) U.S. Cl. 704/275

(57) ABSTRACT

A robotic system which controls the movement of a surgical instrument in response to voice commands from the user. The robotic system has a computer controlled arm that holds the surgical instrument. The user provides voice commands to the computer through a microphone. The computer contains a phrase recognizer that matches the user' speech with words stored in the computer. Matched words are then processed to determine whether the user has spoken a robot command. If the user has spoken a recognized robot command the computer will move the robotic arm in accordance with the command.

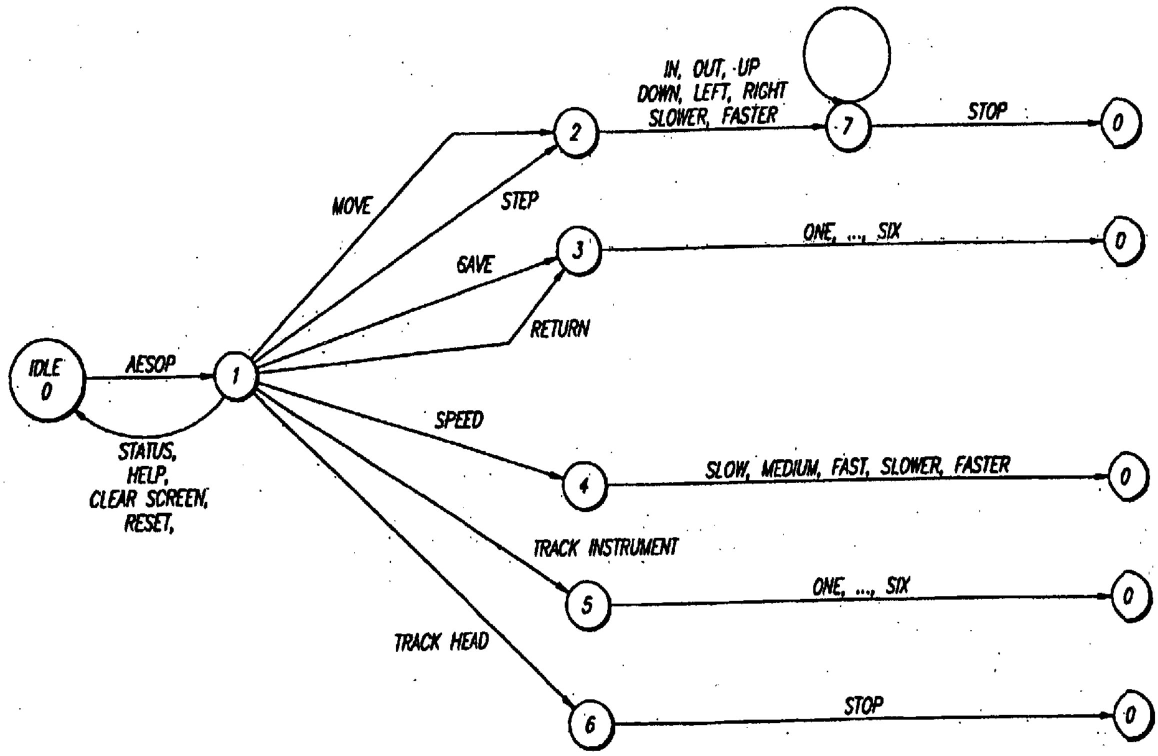
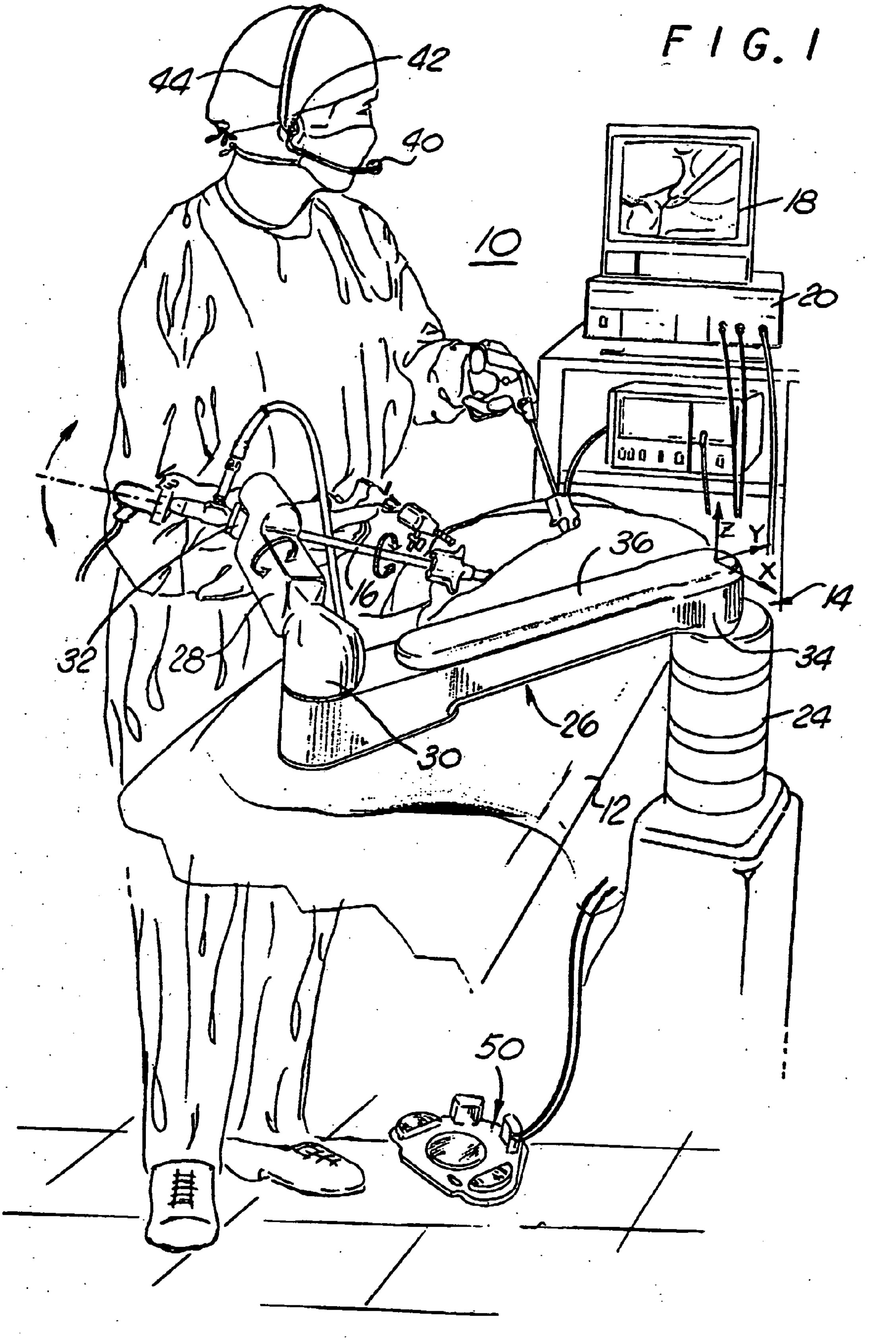


FIG. 1



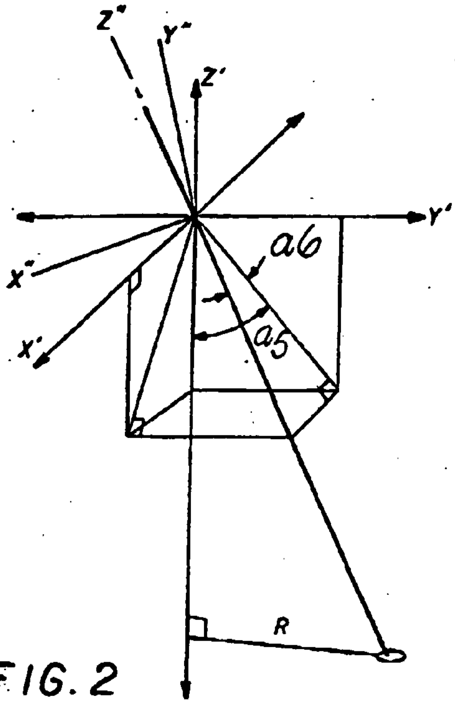
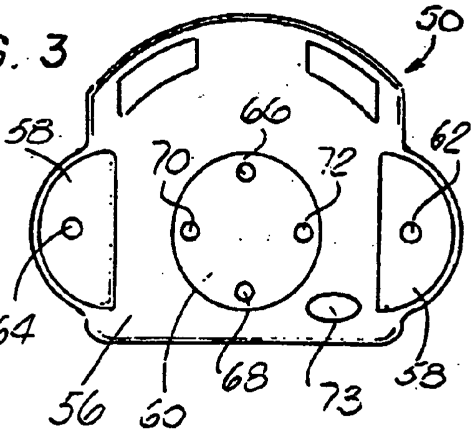


FIG. 2

FIG. 3



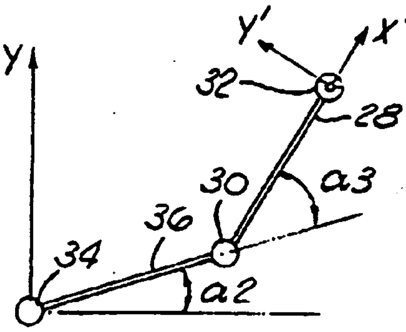


FIG. 6

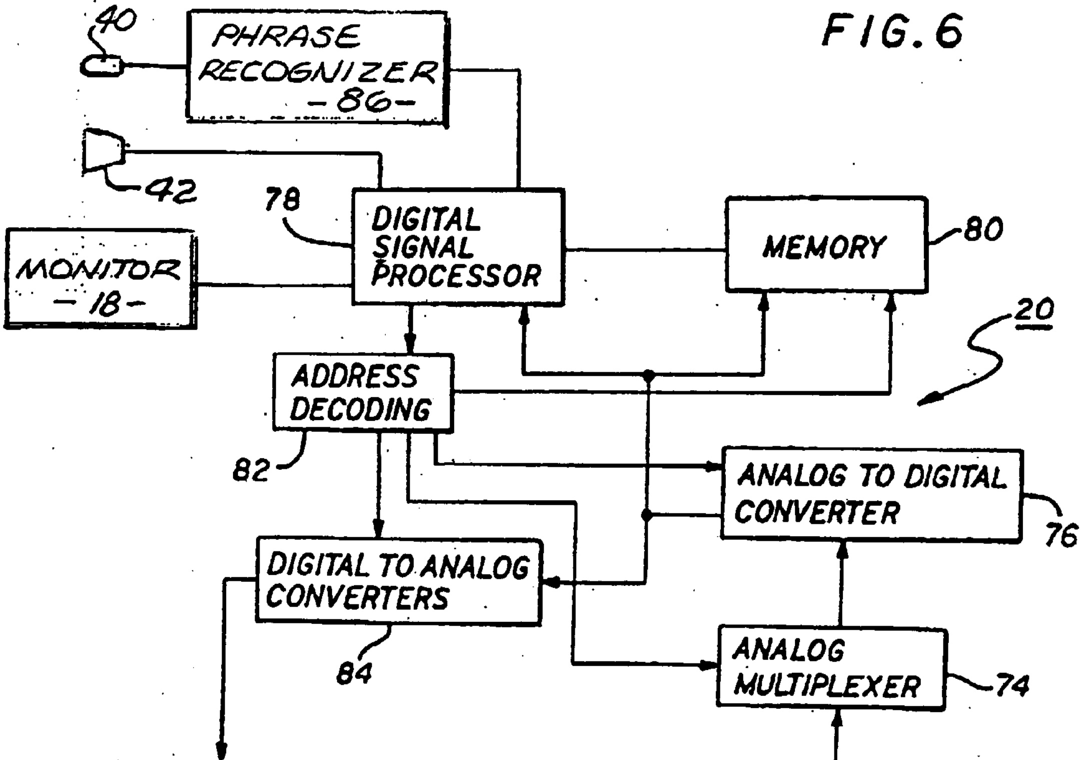


FIG. 4

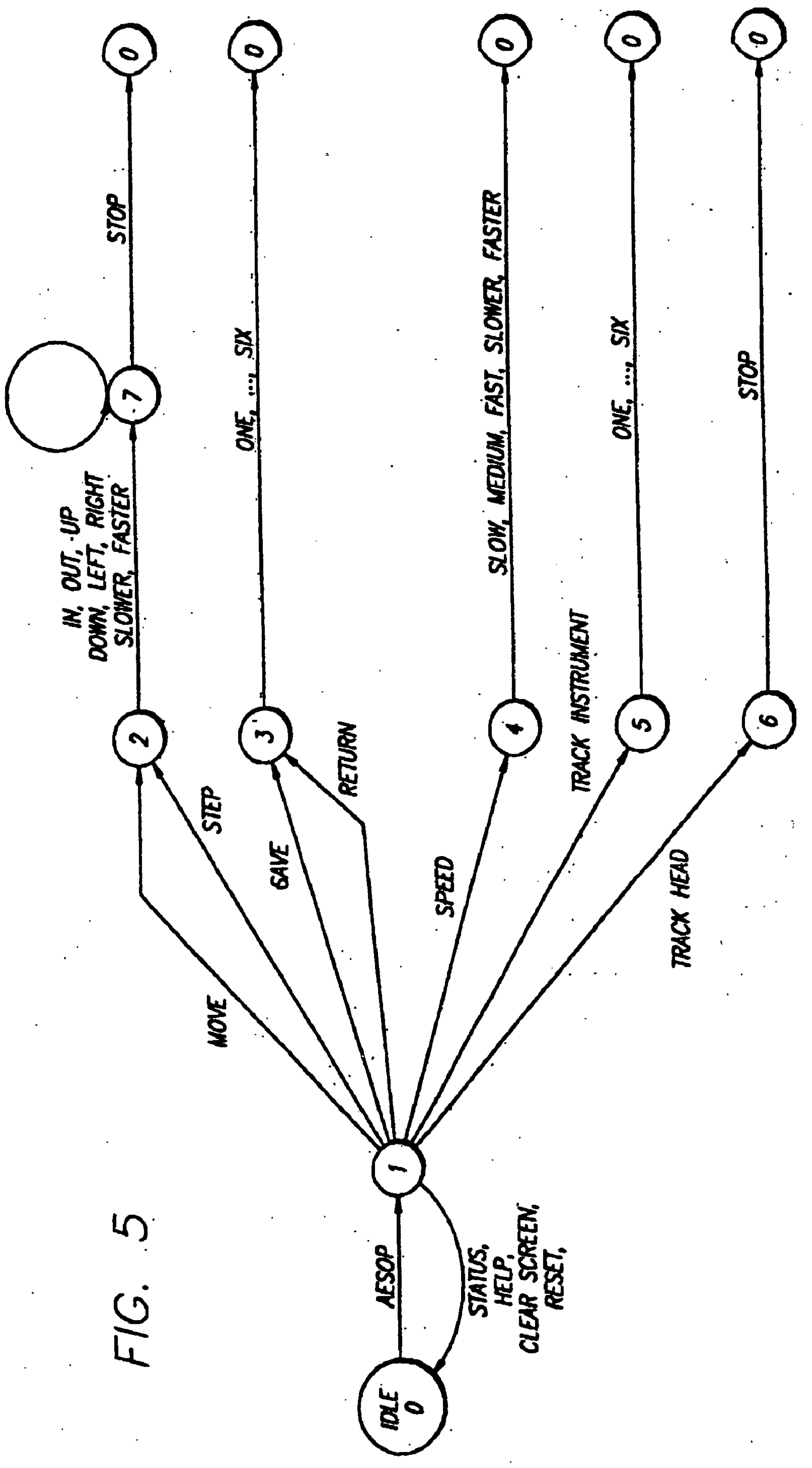

IDLE 0
AESOP
STATUS, HELP, CLEAR SCREEN, RESET,
1
MOVE
STEP
SAVE
RETURN
SPEED
TRACK INSTRUMENT
TRACK HEAD
2
3
4
5
6
7
IN, OUT, UP DOWN, LEFT, RIGHT SLOWER, FASTER
STOP
ONE, ..., SIX
SLOW, MEDIUM, FAST, SLOWER, FASTER
ONE, ..., SIX
STOP
0
0
0
0
0

FIG. 5

SPEECH INTERFACE FOR AN AUTOMATED ENDOSCOPE SYSTEM

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This is a continuation patent application of U.S. patent application Ser. No. 10/095,488 filed Mar. 11, 2002, the full disclosure of which is incorporated herein by reference.

BRIEF SUMMARY OF THE INVENTION

[0002] The present invention is a robotic system which controls the movement of a surgical instrument in response to voice commands from the user. The robotic system has a computer controlled arm that holds the surgical instrument. The user provides voice commands to the computer through a microphone. The computer contains a phrase recognizer that matches the User's speech with words stored in the computer. Matched words are then processed to determine whether the user has spoken a robot command. If the user has spoken a recognized robot command the computer will move the robotic arm in accordance with the command.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] The objects and advantages of the present invention will become more readily apparent to those ordinarily skilled in the art after reviewing the following detailed description and accompanying drawings, wherein:

[0004] **FIG. 1** is a perspective view of a robotic endoscope system of the present invention;

[0005] **FIG. 2** is a schematic of an endoscope within two separate coordinate systems;

[0006] **FIG. 3** is a top view of a foot pedal;

[0007] **FIG. 4** is a schematic of a computer system;

[0008] **FIG. 5** is a schematic of a grammar process;

[0009] **FIG. 6** is a schematic of a robotic arm.

DETAILED DESCRIPTION OF THE INVENTION

[0010] Referring to the drawings more particularly by reference numbers, **FIG. 1** shows a robotic system **10** of the present invention. The system **10** is typically used in a sterile operating room where a surgeon performs a surgical procedure on a patient. The patient is placed on a operating table **12**. Attached to the table **12** is a robotic arm assembly **14** which can move a surgical instrument **16** relative to the table **12** and the patient. The surgical instrument **16** is typically an endoscope which is inserted into the abdomen of the patient **12**. The endoscope **16** enters the patient through a cannula, wherein the scope **16** rotate about a cannula pivot point. The endoscope is typically connected to a monitor **18** which allows the surgeon to view the organs, etc. of the patient. Although an endoscope is described and shown, it is to be understood that the present invention can be used with other surgical instruments.

[0011] The robotic arm assembly **14** controlled by a computer **20**. In the preferred embodiment, the robotic arm assembly **16** includes a linear actuator **24** fixed to the table **14**. The linear actuator **24** is connected to a linkage arm assembly **26** and adapted to move the linkage assembly **26** along the z axis of a first coordinate system. The first coordinate system also has an x axis and a y axis.

[0012] The linkage arm assembly **26** includes a first linkage arm **28** attached to a first rotary actuator **30** and an end effector **32**. The first rotary actuator **30** is adapted to rotate the first linkage arm **28** and end effector **32** in a plane perpendicular to the z axis (x-y plane) The first rotary actuator **30** is connected to a second rotary actuator **34** by a second linkage arm **36**. The second actuator **34** is adapted to rotate the first actuator **30** in the x-y plane. The second rotary actuator **34** is connected to the output shaft of the linear actuator **24**. The actuators **24**, **30** and **34** rotate in response to output signals provided by the computer **20**. As shown in **FIG. 2**, the junction of the endoscope **16** and the end effector **32** define a second coordinate-system which has an x' axis, a y' axis and a z' axis. The junction of the end effector **32** and endoscope **18** also define the origin of a third coordinate system which has a x' axis, a p axis and a z" axis. The z" axis parallel with the longitudinal axis of the endoscope **16**.

[0013] The arm assembly may have a pair of passive joints that allow the end effector to be rotated in the direction indicated by the arrows. The actuators **24**, **30** and **34**, and joints of the arm may each have position sensors (not shown) that are connected to the computer **20**. The sensors provide positional feedback signals of each corresponding arm component.

[0014] The system has a microphone **40** that is connected to the computer **20**. The system may also have a speaker **42** that is connected to the computer **20**. The microphone **40** and speaker **42** may be mounted to a headset **44** that is worn by the user. Placing the microphone **40** in close proximity to the user reduces the amount of background noise provided to the computer and decreases the probability of an inadvertent input command.

[0015] As shown in **FIG. 3**, the system may also have a foot pedal **50**. The foot pedal **22** has a housing **56** that supports a pair of outer first foot switches **58** and a second foot switch **60**. One outer foot switch **58** has a first pressure transducer **62** and the other switch has a second pressure transducer **64**. The second foot switch **60** has third **66**, fourth **68**, fifth **70** and sixth **72** pressure transducers. The transducers are each connected to a corresponding operational amplifier that provides a voltage input to the computer **20**. The pressure transducers **62**-**72** are preferably constructed so that the resistance of each transducer decreases as the surgeon increases the pressure on the foot switches. Such a transducer is sold by Interlink Electronics. The decreasing transducer resistance increases the input voltage provided to the computer **20** from the operational amplifier. Each transducer corresponds to a predetermined direction within the image displayed by the monitor. In the preferred embodiment, the first pressure transducer **62** corresponds to moving the endoscope toward the image viewed by the surgeon. The second transducer **64** moves the scope away from the image. The third **66** and fourth **68** transducers move the image "up" and "down", respectively, and the fifth **70** and sixth **72** transducers move the image "left" and "right", respectively. The pedal may have a button **73** that enables the foot pedal **50** and disable the voice command feature, or vice versa.

[0016] **FIG. 4** shows a schematic of the computer **20**. The computer **20** has a multiplexer **74** which is connected to the pressure transducers of the foot pedal **50** and the position sensors of the arm. The multiplexer **74** is connected to a single analog to digital (A/D) converter **76**. The computer **20** also has a processor **78** and memory **80**.

[0017] The processor **78** is connected to an address decoder **82** and separate digital to analog (D/A) converters **84**. Each D/A converter is connected to an actuator **24**, **30** and **34**. The D/A converters **84** provide analog output signals to the actuators in response to output signals received from the processor **78**. The analog output signals have a sufficient voltage level to energize the electric motors and move the robotic arm assembly. The decoder **82** correlates the addresses provided by the processor with a corresponding D/A converter, so that the correct motor(s) is driven. The address decoder **82** also provides an address for the input data from the A/D converter **76** so that the data is associated with the correct input channel.

[0018] The computer **20** has a phrase recognizer **86** connected to the microphone **40** and the processor **78**. The phrase recognizer **86** digitizes voice commands provided by the user through the microphone **40**. The voice commands are then processed to convert the spoken words into electronic form. The electronic words are typically generated by matching the user's speech with words stored within the computer **20**. In the preferred embodiment, the recognizer **86** is an electronic board with accompanying software that is marketed by SCOTT INSTRUMENTS of Denton, Tex. under the trademark "Coretechs Technology".

[0019] The electronic words are provided to the processor **78**. The processor **78** compares a word, or a combination of words to predefined robot commands that are stored within a library in the memory **80** of the computer **20**. If a word, or combination of words match a word or combination of words in the library, the processor **78** provides output commands to the D/A converter **84** to move the robotic arm in accordance with the command.

[0020] **FIG. 5** shows exemplary words and combinations of words that provide robot commands. A grammar process is performed to determine whether the voice commands satisfy certain conditions. The process contains a number of states advanced by the satisfaction of a condition. If the voice command provided by the user satisfies a first condition, then the process proceeds to the first state. If a condition of a next state is satisfied then the process proceeds to the next corresponding state, and so forth and so on. For example, to prevent a robot command from being inadvertently spoken, it is desirable to predicate all voice commands with a qualifier. For example, the qualifier may be a name given to the robot such as "AESOP". Therefore when the user provides a voice command, the process initially determines whether the spoken word is AESOP. If the spoken word is not AESOP then the process ends. The term "stop" may be an exception to this rule, wherein the computer will stop arm movement when the user provides a simple "stop" voice command.

[0021] If the spoken word is AESOP the process continues to state 1. The process next determines whether the user has spoken a word that satisfies a condition to advance to states **2-6**. These words include "move", "step", "save", "return", "speed", "track instrument" and "track head". The track instrument command is for a system which has the ability to move an endoscope to automatically track the movement of a second instrument that is inserted into the patient. The track head command may enable the system so that the endoscope movement tracks the user's eyes. For example, if the user looks to the right of the image displayed by the monitor, the robot will move the endoscope to move the image in a rightward direction. The move and step commands induce movement of the scope in a desired direction. The save command saves the position of the endoscope within the memory of the computer. The return command will return the scope to a saved position.

[0022] From states **2-6** the process will determine whether the user has spoken words that meet the next condition and so forth and so on. When a certain number of conditions have been met, the processor **78** will provide an output command to the D/A converter **84** in accordance with the voice commands. For example, if the user says "AESOP move left", the processor **78** will provide output commands to move the endoscope **12**, so that the image displayed by the monitor moves in a leftward direction. The microphone **40** phrase recognizer **86** and grammar process essentially provide the same input function as the foot pedal **50**, multiplexer **74** and A/D converter **76**.

[0023] The processor **78** can also provide the user with feedback regarding the recognized command through the speaker **42** or the monitor **18**. For example, when the user states "AESOP move right", after processing the speech, the processor **78** can provide an audio message through the speaker **42**, or a visual message on the monitor **18**, "AESOP move right". Additionally, the processor **78** can provide messages regarding system errors, or the present state of the system such as "speed is set for slow".

[0024] Referring to **FIG. 6**, the processor **78** typically computes the movement of the robotic arm assembly **16** in accordance with the following equations.

$$a3 = \pi - \cos^{-1}\left(\frac{x^2 + y^2 - L1^2 + L2^2}{-2L1L2}\right) \qquad 1)$$

$$\Delta = \cos^{-1}\left(\frac{x^2 + y^2 + L1^2 - L2^2}{2 \cdot L1\sqrt{x^2 + y^2}}\right)$$

$$a0 = \tan^{-1} 2\left(\frac{y}{x}\right)$$

$$a2 = a0 + / - \Delta$$

[0025] where;

[0026] a2=angle between the second linkage arm **36** and the x axis.

[0027] a3=angle between the first linkage arm **28** and the longitudinal axis of the second linkage arm **36**.

[0028] L1=length of the second linkage arm.

[0029] L2=length of the first linkage arm.

[0030] x=x coordinate of the end effector in the first coordinate system.

[0031] y=y coordinate of the end effector in the first coordinate system.

[0032] To move the end effector to a new location of the x-y plane the processor **78** computes the change in angles a2 and a3 and then provides output signals to move the actuators accordingly. The original angular position of the end effector is provided to the processor **78** by the position sensors. The processor moves the linkage arms an angle that corresponds to the difference between the new location and the original location of the end effector. A differential angle Δa2 corresponds to the amount of angular displacement provided by the second actuator **34**, a differential angle Δ3 corresponds to the amount of angular displacement provided by the first actuator **30**.

[0033] To improve the effectiveness of the system **10**, the system is constructed so that the desired movement of the surgical instrument correlates to a direction relative to the image displayed by the monitor. Thus when the surgeon commands the scope to move up, the scope always appears to move in the up direction. To accomplish this result, the processor **78** converts the desired movement of the end of the endoscope in the third coordinate system to coordinates in the second coordinate system, and then converts the coordinates of the second coordinate system into the coordinates of the first coordinate system.

[0034] Referring to **FIG. 2**, the desired movement of the endoscope is converted from the third coordinate system to the second coordinate system by using the following transformation matrix:

$$\begin{pmatrix} \Delta x' \\ \Delta y' \\ \Delta z' \end{pmatrix} = \begin{pmatrix} \cos(a6) & 0 & -\sin(a6) \\ -\sin(a5)\sin(a6) & \cos(a5) & -\sin(a5)\cos(a6) \\ \cos(a5)\sin(a6) & \sin(a5) & \cos(a5)\cos(a6) \end{pmatrix} \begin{pmatrix} \Delta x'' \\ \Delta y'' \\ \Delta z'' \end{pmatrix} \quad 2)$$

[0035] where;

[0036] Δx"=the desired incremental movement of the scope along the x" axis of the third coordinate system.

[0037] Δy"=the desired incremental movement of the scope along the y" axis of the third coordinate system.

[0038] Δz"=the desired incremental movement of the scope along the z" axis of the third coordinate system.

[0039] a5=the angle between the z' axis and the scope in the y-z' plane.

[0040] a6=the angle between the z' axis and the scope in the x'-z' plane.

[0041] Δx"=the computed incremental movement of the scope along the x' axis of the second coordinate system.

[0042] Δy"=the computed incremental movement of the scope along the y' axis of the second coordinate system.

[0043] Δz"=the computed incremental movement of the scope along the z' axis of the second coordinate system.

[0044] The angles a5 and a6 are provided by position sensors located on the end effector **32**. The angles a5 and a6 are shown in **FIG. 2**.

[0045] The desired movement of the endoscope is converted from the second coordinate system to the first coordinate system by using the following transformation matrix:

$$\begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \begin{pmatrix} \cos(\pi) & -sn(\pi) & 0 \\ \sin(\pi) & \cos(\pi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x'' \\ \Delta y'' \\ \Delta z'' \end{pmatrix} \quad 3)$$

[0046] where;

[0047] Δx'=the computed incremental movement of the scope along the x' axis of the second coordinate system.

[0048] Δy'=the computed incremental movement of the scope along the y' axis of the second coordinate system.

[0049] Δz'=the computed incremental movement of the scope along the z' axis of the second coordinate system.

[0050] π=is the angle between the first linkage arm and the x axis of the first coordinate system.

[0051] Δx=the computed incremental movement of the scope along the x axis of the first coordinate system.

[0052] Δy=the computed incremental movement of the scope along the y axis of the first coordinate system.

[0053] Δz=the computed incremental movement of the scope along the z axis of the first coordinate system.

[0054] The incremental movements Δx and Δy are inserted into the algorithms described above for computing the angular movements (Δa2 and Δa3) of the robotic arm assembly to determine the amount of rotation that is to be provided by each electric motor. The value Δz is used to determine the amount of linear movement provided by the linear actuator **24**.

[0055] The surgical instrument is typically coupled to a camera and a viewing screen so that any spinning of the instrument about its own longitudinal axis will result in a corresponding rotation of the image on the viewing screen. Rotation of the instrument and viewing image may disorient the viewer. It is therefore desirable to maintain the orientation of the viewing image. In the preferred embodiment, the end effector has a worm gear (not shown) which rotates the surgical instrument about the longitudinal axis of the instrument. To insure proper orientation of the endoscope **16**, the worm gear rotates the instrument **16** about its longitudinal axis an amount Δθ6 to insure that the y" axis is oriented in the most vertical direction within the fixed coordinate system. Δθ6 is computed from the following cross-products.

[0056] Δθ6=zi" (yo".yi")

[0057] where;

[0058] Δθ6=the angle that the instrument is to be rotated about the z" axis.

[0059] yo"=is the vector orientation of the y" axis when the 20 instrument is in the first position.

[0060] yiΔ=is the vector orientation of the y" axis when the instrument is in the second position.

[0061] zi"=is the vector orientation of the z" axis when the instrument is in the second position.

[0062] The vectors of the yi" and zi" axis are computed with the following algorithms.

$$[zi''] = \begin{pmatrix} \cos a6 & 0 & -\sin a6 \\ -\sin a5 \sin a6 & \cos a5 & -\sin a5 \cos a6 \\ \cos a5 \sin a6 & \sin a5 & \cos a5 \cos a6 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$xi''=z \times zi''$

$yi''=zi'' \times xi''$

[0063] where;

[0064] a5=is the angle between the instrument and the z axis in the y-z plane.

[0065] a6=is the angle between the instrument and the z axis in the x-z plane.

[0066] z=is the unit vector of the z axis in the first coordinate system.

[0067] The angles a5 and a6 are provided by position sensors. The vector yo" is computed using the angles a5 and a6 of the instrument in the original or first position. For the computation of yi" the angles a5 and a6 of the second position are used in the transformation matrix. After each arm movement yo" is set to yi" and a new yi" vector and corresponding Δθ6 angle are computed and used to re-orient the endoscope. Using the above described algorithms, the worm gear continuously rotates the instrument about its longitudinal axis to insure that the pivotal movement of the endoscope does not cause a corresponding rotation of the viewing image.

[0068] The system may have a memory feature to store desired instrument positions within the patient. The memory feature may be enabled either by voice commands or through a button on an input device such as the foot pedal. When a save command is spoken, the coordinates of the end effector in the first coordinate system are saved in a dedicated address(es) of the computer memory. When a return command is spoken, the processor retrieves the data stored in memory and moves the end effector to the coordinates of the effector when the save command was enabled.

[0069] The memory feature allows the operator to store the coordinates of the end effector in a first position, move the end effector to a second position and then return to the first position with a simple command. By way of example, the surgeon may take a wide eye view of the patient from a predetermined location and store the coordinates of that location in memory. Subsequently, the surgeon may manipulate the endoscope to enter cavities, etc. which provide a more narrow view. The surgeon can rapidly move back to the wide eye view by merely stating "AESOP return to one".

[0070] In operation, the user provides spoken words to the microphone. The phrase recognizer **86** matches the user's speech with stored words and provides matched electronic words to the processor **78**. The processor performs a grammar process to determine whether the spoken words are robot commands. If the words are commands, the computer energizes the actuators and moves the endoscope, accordingly. The system also allows the user to control the movement of the endoscope with a foot pedal if voice commands are not desired.

[0071] While certain exemplary embodiments have been described and shown in the accompanying drawings, it is to be understood that such embodiments are merely illustrative of and not restrictive on the broad invention, and that this invention not be limited to the specific constructions and arrangements shown and described, since various other modifications may occur to those ordinarily skilled in the art:

What is claimed is:

1. A voice recognition system for use with a surgical instrument in surgery on a patient, the system comprising:

a microphone for inputting a plurality of spoken surgical instructions and a spoken stop command, each of the spoken instructions including a spoken qualifier and a spoken command, the plurality of instructions including a first instruction having a first spoken qualifier and a first spoken command;

a memory for storing a first plurality of allowable alternative commands;

a processor coupled to the microphone and the memory, the processor having a first state and a second state, the processor in the first state configured to change to the second state in response to the first spoken qualifier, the processor in the second state configured to compare the first spoken command to the first plurality of commands, the processor having an output for transmitting a first command signal to the surgical instrument in response to the first spoken instruction having the first spoken qualifier and the first spoken command matching one of the first plurality of commands;

the processor configured to transmit a stop command signal to the instrument in response to the spoken stop command without an associated spoken qualifier.

2. The voice recognition system of claim 1, the surgical instrument including an endoscope coupled to a monitor, the output of the processor being coupleable to the endoscope so as to alter a displayed image from the endoscope as shown on the monitor.

3. The voice recognition system of claim 2, the surgical instrument including a robotic arm supporting the endoscope, the distal end of the endoscope comprises a tip defining a viewing coordinate frame, wherein the processor is configured to calculate transformations between the viewing coordinate frame and a coordinate frame of the robot, so that the command signal comprises motor signals derived from the transformations so that the camera tip moves in an internal surgical site to effect an instructed change in the image shown in the display.

4. The voice recognition system of claim 1, wherein the output is configured to be coupled to the surgical instrument when the surgical instrument comprises a member selected from the group consisting of:

a surgical tool;

a forceps;

a device positioned within the patient's body for transmitting an image outside of the patient's body;

a laparoscope; and

a medical telescope.

5. The voice recognition system of claim 4, wherein the stop command signal from the processor is configured to inhibit potential injury to the patient that might otherwise be inflicted by the surgical instrument.

6. The voice recognition system of claim 1, wherein the plurality of instructions include a second spoken instruction having a second spoken qualifier and a second spoken command, the memory storing a second plurality of allowable alternative commands, the processor changing to a third state in response to the second spoken qualifier, the processor in the third state configured to compare the second spoken command to the second plurality of commands, the processor generating a second command signal in response to the second spoken instruction having the second spoken qualifier and the second spoken command matching one of the second plurality of commands

7. The voice recognition system of claim 6, wherein the processor has a fourth state, the processor changing to the fourth state in response to the second command signal, the voice recognition system in the fourth state accepting a third plurality of allowable commands stored in the memory and associated with the fourth state.

8. The voice recognition system of claim 1, further comprising a speaker coupled to the processor for generating audible messages to a surgeon regarding operation of the system.

9. The voice recognition system of claim 8, wherein said audible messages comprise audible feedback indicating successful receipt of said spoken instructions.

10. The voice recognition system of claim 9, wherein said audible messages comprise synthesized voice messages.

* * * * *

专利名称(译)	用于自动内窥镜系统的语音接口		
公开(公告)号	US20050033580A1	公开(公告)日	2005-02-10
申请号	US10/942374	申请日	2004-09-16
[标]申请(专利权)人(译)	电脑动作公司		
申请(专利权)人(译)	COMPUTER MOTION，INC.		
当前申请(专利权)人(译)	Intuitive Surgical公司运营，INC.		
[标]发明人	WANG YULUN UECKER DARRIN		
发明人	WANG, YULUN UECKER, DARRIN		
IPC分类号	A61B17/00 A61B19/00 G06F19/00 G10L15/26 G10L11/00		
CPC分类号	A61B19/22 A61B19/50 G10L15/26 G06F19/3406 A61B2017/00203 A61B1/00041 A61B1/00149 A61B1/0016 A61B34/10 A61B34/70 G16H40/63		
其他公开文献	US7395249		
外部链接	Espacenet USPTO		

摘要(译)

一种机器人系统，其响应于来自用户的语音命令控制手术器械的运动。机器人系统具有计算机控制的臂，其保持手术器械。用户通过麦克风向计算机提供语音命令。计算机包含一个短语识别器，它将用户的语音与存储在计算机中的单词相匹配。然后处理匹配的单词以确定用户是否已说出机器人命令。如果用户说出了识别出的机器人命令，则计算机将根据该命令移动机器人手臂。