Function of Object from Observing Human Action
-Toward improving the operationality of virtual tools-

Kazuhiro Annoura, Kazuaki Tanaka, Jiang Yu Zheng, Norihiro Abe, Shoujie He
Faculty of Computer Science and Systems Engineering
Kyushu Institute of Technology
noura@sein.mse.kyutech.ac.jp
Hirokazu Taki
Faculty of System Engineering
Wakayama University

Commission V, Working Group SIG

KEY WORDS: magnetic sensor, virtual reality, human action, function of object

ABSTRACT

As it is not easy for a user in virtual environmente use virtual tools, improving the feasibility of manipulating them is necessary to help him correctly achieve his intenstion. Tools are used to change the current state of an object into desirable one, so the several constraints are imposed on their movement, a grasping method and an object class to be manipulated. Consequently, it is possible for the system to predict user's intent by watching his behavior. We propose a method to measure the movement of real tools when they are exploited for a particular purpose and to store them so that they are retrieved according to their constraints. Inferring user's intention through the observation of his behavior, the system plays back one of the stored movements. This allows any users to exploit a virtual tool in virtual environment in the similar way in the real one.

1. INTRODUCTION

An artificial reality or a virtual reality have been attracted attention as techniques for having us intuitively understand problems which it is hard for us to experience in daily life. For making it easy for us to intuitively comprehend events occurring in virtual environments, a man machine interface plays an important role to secure fast and correct communication between us and a computer.

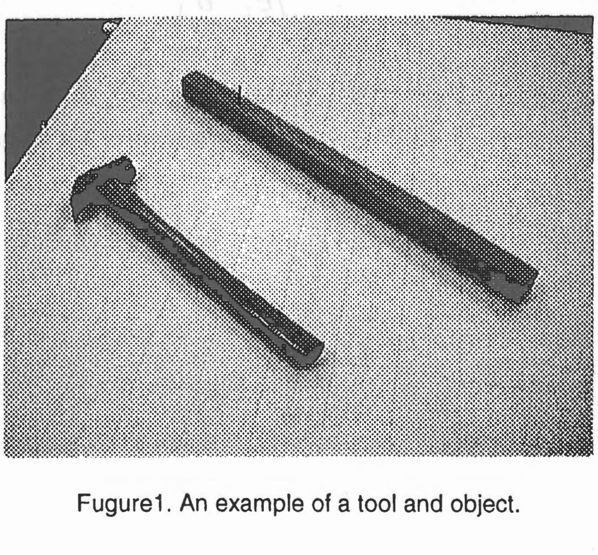
Displays, keyboards or mouses are well known as interface tool for bridging between a person and a computer, but a mouse does not allows a person with no experience in a virtual space to manipulate the 3 dimensional space at his own will. A number of purpose-built input/output devices have been developed in virtual reality studies to facilitate the 3 dimensional manipulation. Input devices such as a handle or an accelerator used in drive simulators are typical examples. These purpose-made devices generally make it difficult to generalize a system for controlling devices including problems of their sizes, costs and popularization into families. An approach of building a knowledge base of tools was proposed to divide input devices from a core system by restricting them to 3 dimensional position sensors and data gloves (Funahashi et al., 1997). In this system, the process of manipulating a tool in a virtual environment is realized by visualizing a gesture of a data-glove, but there is somewhat difference between the real behavior and the virtual one.

Tools are used to manipulate an object for attaining a particular purposive operation, this allows a system to anticipate an operator's behavior. This means that it is possible to store in advance the behavior of hands observed at the time when he manipulated a real tool.

When a system perceives that a person is going to use a tool in order to attain its original goal, by playing back the behavior corresponding to the tool he will get real operational feeling.

Consequently, if a data glove approaches to a tool and he is going to grasp it with a grip which makes the operation possible, the system will get his virtual hand to grasp the virtual tool and generate a sequence of motion from the stored behavior. Of cause, the size of the virtual tool is not always equal to the stored one, then some translation is necessary for generating the motion.

To make clear the effect of the method proposed in this paper, we select a claw hammer as an example which is usable both as hammer and as nail extractor. A nail and a board is given as the objects of the tool (Figure 1).



Fugure1. An example of a tool and object.

2. FUNCTION OF A TOOL

In robotic research, the researches have been studied that try to have robot manipulators to realize a sequence of motion obtained by getting a person to do the works such as trimming, grasp, assembly and elastic object insertion.

The human behavior is measured with the motion capturing method such as a vision system, data glove or 3 dimensional position sensor, and a force torque sensor is used to detect a force needed to the works. The force measurament is required for reproducing actions which need not only the position control but also the force one.

The purpose of this research is to realize a system that helps a person treat virtual tools in a virtual environment which correspond to the tools he uses every day. At present we do not intend to have five fingered robots to use scissors or a nail extractor, but if humanoid robots are introduced into human community, they may be required to adapt themselves to the tools. In the case, the method we propose in this paper will contribute to make clear the relation between tools and their functions.

As you can see from the fact that you will use a tool except a hammer to hit a nail if you find it unavailable, a tool will have not only the function peculiar to itself but also another function which is fulfilled only when there is no tool best fits to a given purpose. The latter one is not difficult to give the definition in advance, it can be inferred from observing which kind of tools the majority of men will use to attain a given goal when the most familiar tool is not available.

A vision system is reported that infers the function of an object by analizing the locus of its movement when a person uses it as tool (Sharma et al., 1997), but it is difficult to measure the grasping method or intricate motion of hands.

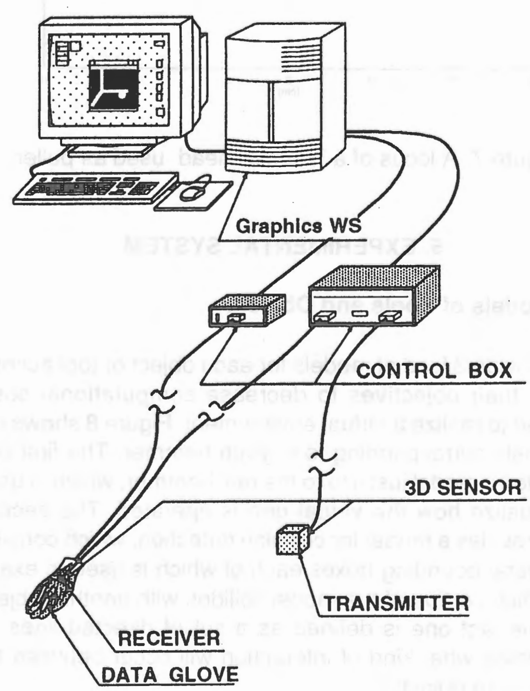


Figure 2. System congifuration.

3. SYSTEM ORGANIZATION

Figure 2 shows the system organization. Both position data acquired from 3 dimensional position sensor attached to a data glove and each joint angle of every fingure are input to the graphic workstation OCTANE, and the scene in a virtual environment is generated from these data. The 6 dimensional data consisting of x, y, z and yaw, pitch, role are acquired from ISOTRAK II of POLHEMUS INC., and the maximum rate of data acquisition is 30 frames/sec per one receiver. 10 bending angles of finger joints are captured with Data Glove Model J of Nissyo Electronics INC.

4. MOTION CAPTURE OF HANDS DURING OPERATING A TOOL

4.1 Environment of Data Acquisition

The motion of a hand during operating a tool in a real environment is captured with the magnetic position sensors.

Magnetic sensors have a weak point that it is affected by environment, and what is worse the compensation of the influence is hard because it is non linear. Fortunately, as the measuring range is from 0 to 50cm, the distortion is very little and there is no need of data calibration. Figure 3 shows an amount of distortion occurring in position data measured with a magnetic sensor; dotted lines show measured values.

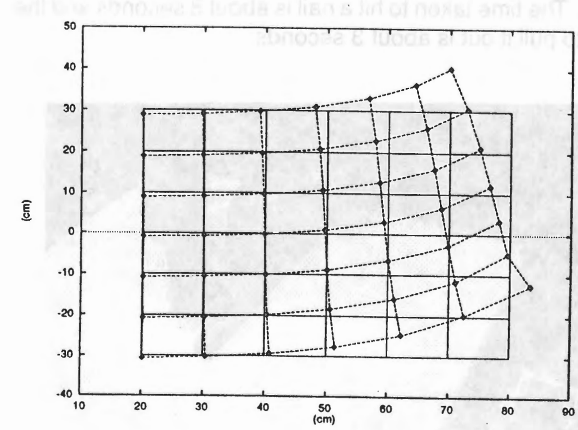


Figure 3. Distortion of magnetic field.

Because the head of a hammer is made of metal, it certainly affects to the value measured with a magnetic sensor. Figure 4 shows the deviation in values of the magnetic sensor located at (0,25), with a transmitter located at (0,0) and a hammer moving along on the horizontal line passing the point (0,20). From the result, it can be concluded that the influence of a hammer head to the magnetic field takes a maximum value when it is on the line passing both the transmitter and receiver, and is negligible when it is 10cm distant from the receiver. The motion measurement of a hammer is conducted without grasping a hammer head in a hand, no calibration is not specially taken into consideration.

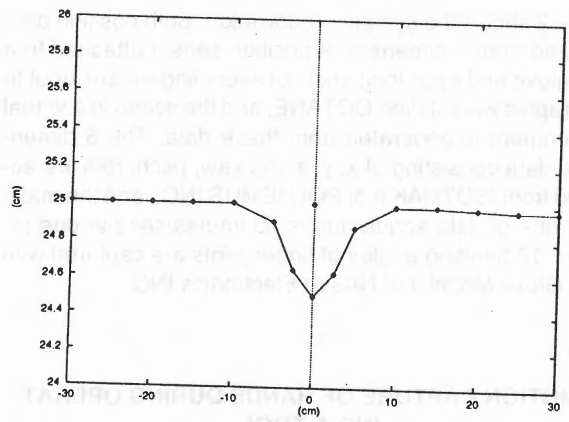


Figure 4. Effect caused with a hammer head.

4.2 Motion Capture of a Hand

The system proposed in this paper allows a person to manipulate a virtual tool according to his intention, and the operationality of the tool to be promoted by giving both his virtual hand and the virtual tool the movement reproduced using the data obtained from the operation of the real tool with his own hand in an ideal real environment. The motion of a hand is captured 30 times in every second at both the cases when he hits a nail with a hammer and when he uses a nail extractor. Figure 5 shows the data acquisition process. The time taken to hit a nail is about 8 seconds and the one to pull it out is about 3 seconds.

Figure 5. Environment for experiment.

4.3 Analysis of Hand Movement

Figure 6 shows the motion of a hand observed during a hammer is hitting a nail. You can see from the figure that the motion of hand draws a small ellipse while the nail is not yet enough driven into the board, but that the size of ellipse becomes big as the nail is firmly fixed to the board. The positional relation between a hammer and a nail is not shown here, it has become clear that the head of a hammer hits a nail not from right above but from a little aslant direction. It seems that this is due to the characteristic of our arm joints. Consequently an simple up and down motion of a hammer is not enough to give an operator feeling that he is really hitting a nail.

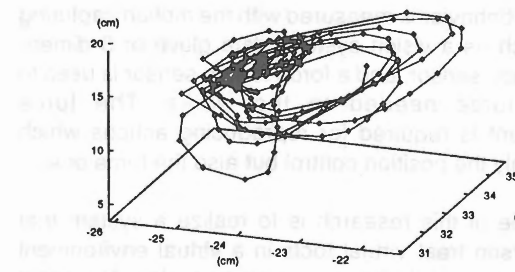


Figure 6. A locus of a hammer head used as hammer.

Figure 7 shows how the hand moves observed while a nail is pulled out with a claw hammer. You can see that it moves along an arc that has the nail as a center and the handle of the hammer as a radius. Comparison with the hitting operation, the motion is a two dimensional one.

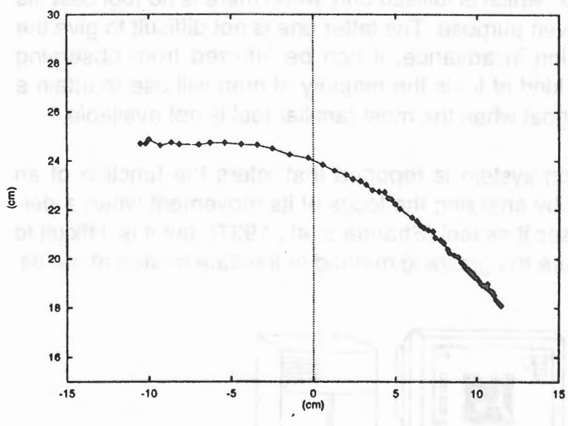


Figure 7. A locus of a hammer head used as puller.

5. EXPERIMENTAL SYSTEM

5.1 Models of Tools and Objects

We provide 3 kind of models for each object or tool according to their objectives to decrease computational costs needed to realize a virtual environment. Figure 8 shows the 3 models corresponding to a given hammer. The first one provides a model just like to the real hammer, which is used to visualize how the virtual one is operated. The second one provides a model for collision detection, which consists of several bounding boxes each of which is used to examine which portion of the model collides with another object, and the last one is defined as a set of directed lines for examining what kind of interaction will occur between the tool and an object.

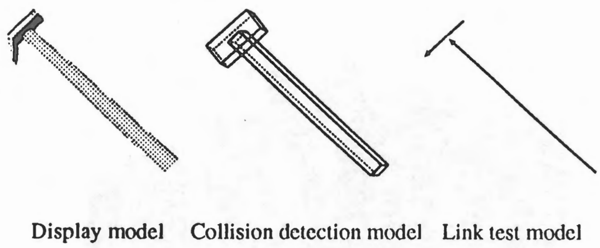


Figure 8. Three models prepared for a hammer.

5.2 Rendering Virtual Hand and Arm

A user is able to see his virtual hand and arm operating in a virtual environment. As shown in Figure 9 the postures of them are measured with two receivers which are attached at the middle point of his lateral brachial region and at the back of his hand. The attitude of his forearm is calculable using the above data and constraint relation on the forearm. The constraint between the brachium and forearm is that round a single axis, and that between the hand and forearm is biaxial one. The upper arm is also rotatable around the line connecting the forearm and hand, but it is not taken into consideration because the models are rigid and are not deformed.

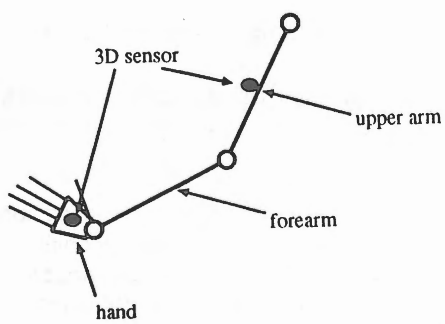


Figure 9. Joint constraint on a hand.

The virtual hand is drawn referring to bending angles of fingers acquired from a data glove. As it returns but 2 value corresponding to the first and second joint of each finger, the value of the third joint for each finger except a thumb must be calculated. The dataglove used is only exploited as switch deciding whether it grasps something, and so 1.3 times the angle of the second joint is set as that of the third one.

5.3 Grasp of a Model by Virtual Hand

Tools or objects are movable when they are grasped. It is determined that a virtual hand holds an object if the following conditions hold.
(1) The virtual hand gets near the object to be grasped.
(2) The shape of hand is similar to that used to grasp the object.
(3) The point at a short distance from the center of the palm, shown in Figure 10, intersects the bounding box of the object.
When the above condition (2) becomes false, the hand is regarded to be separated from the object.

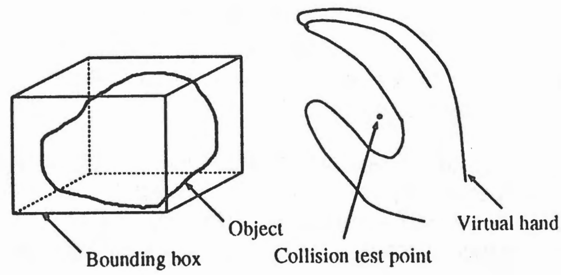


Figure 10. Collision detection method.

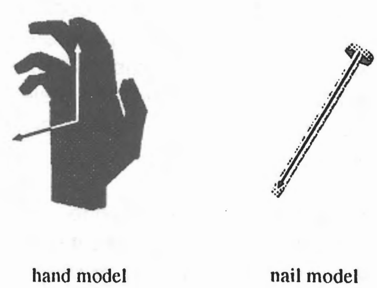


Figure 11. Directed arc models for a hand and hammer.

5.4 Reproduction of Stored Movement

Directed-arc models shown in Figure 8 and 11 are provided for a hammer, virtual hand and nail respectively. The system judges that an operator is going to grasp a hammer to hit a nail when the relation enclosed with the box A in Figure 12 holds between the model of the hammer and that of the nail. With keeping this relation holding, when the relation in the box B also holds, the motion tailored to the operation is generated from the stored one as shown in Figure 14. According to the movement of the hand, the operation is repeated. In the same way, when the relations shown in Figure 13 hold, the behavior for pulling out a nail with a hammer is realized as shown in Figure 15.

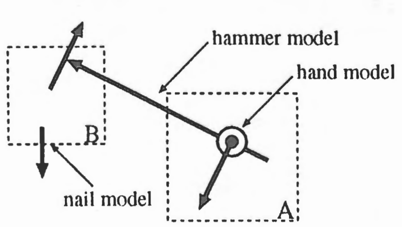


Figure 12. Relation among hammer used as hitter, hand and nail.

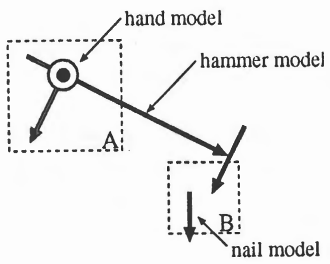


Figure 13. Relation among hammer used as extracter, hand and nail.

The experimental system proposed in this paper successfully improved the operationality by showing an operator how skillfully he manipulates a virtual tool, by taking a hammer for instance which is used for hitting or extracting a nail.

At present, the system simply reproduces the motion of a hammer depending just on the configuration shown in Figure 12 or 13, and does not take account of the situation, where, an operator interrupts hitting or extracting operation. Operator's behavior must be observed to cope with interruption or alteration of his motion. Our experimental system admits that an operator is going to grasp a tool when the difference between his operation and the modeled one remains less than ±10% of that. The validity of the value must be confirmed.

At present, as an operator is not permitted to use any tools not recorded into the system, it must make it possible for a human to use the tool with different size or shape by transforming the corresponding model according to its scale. A discussion about the scale transformation is needed concerning whether a pile driver should be classified into the same class as a hammer.

It is impossible to use a pile driver with the same grip as a hammer in a real world, but in a virtual environment it may be admissible. This problem must be solved depending on what the purpose of the system is.

There are many other tools except for hammers, they must be introduced into this system. The directed arc model provided for a hammer is considered to be adequate for prescribing the relation to a nail. A hammer without a function as nail extractor allows us to hit a nail at either side of its head. This means that bidirectional arc models may be advantageous. Judging which tool a human is going to use requires both segmenting a sequence of hand motion intro articulated parts and matching them against the precondition for applying the tool.

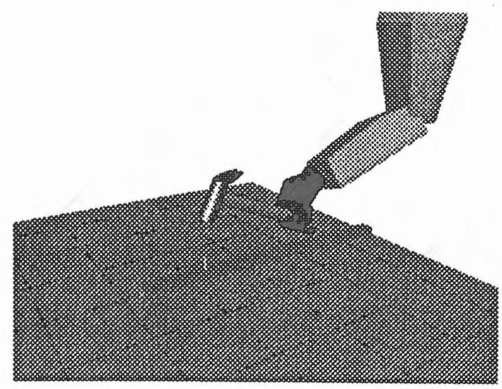


Figure 14. A situation a hammer is used as hammer.

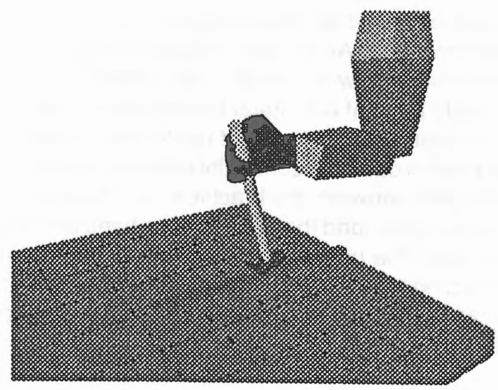


Figure 15. A situation a hammer is used as extracter

References

Duric. Z., Fayman. J. A., Rivlin. E., 1996. Function from motion, PAMI, Vol.18, No.6, pp.579-591

Funahashi er al. 1997. Knowledge base for supporting usage of tools by hand in a virtual sapace. Proc. VRSJ. Vol.2, pp.250-253(in japanese)

Ghazisaedy. M. et al. 1995. Ultrasonic calibration of a magnetic tracker in a virtual reality sapace. Proc, VARIS, pp.179-188

Hirai M. et al. 1997. Human-demonstration based approach to the object motion detection and the recognition of process state transition in insertion of deformable tubes. Jounal of RSJ, Vol.15,No,8, pp.1172-1179(in japanese)

Hirose M., 1993. Virtual reality. (in japanese)Sangyoutosyo Publishers, INC

Ikeuchi K. et al. 1995..Teaching roboto hand using vision. Jounal of RSJ, Vol.13, No.6, pp.599-602(in japanese)

Kuniyoshl, K. et al.1995. Learning by showing. IEEE Trans. R & A. Vol.10, No.6.pp.799-822

Kuroda et al. 1997. Inferring behavior of human's bust using 3 magnetic sencors. Proc. VRSJ. Vol.2, pp.228-229(in japanese)

Nakamae, Nishida. 1986. 3 dimensional computer graphics Syokoudou Publishers, INC (in japanese)

Sharma. R., Molineros.J.,1997. Computer vison-based augmented reality for guiding manual assembly, IEEE Presence. Vol.6, No.3, PP.292-317

Wernecke. J. 1994, The Inventor Mentor. Addison Wesley publishers, INC.