Comparison of Naturalness: Virtual Assembly with a Sophisticated Aid vs. Real Assembly in Building Block Task

Yoshifumi Kitamura, Amy Yee* and Fumio Kishino

ATR Communication Systems Research Laboratories
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-02, Japan
<kitamura, kishino>@atr-sw.atr.co.jp

Abstract

We propose a sophisticated manipulation aid method in a virtual environment based on the visual feedback using the dynamic constraints among object faces. It enables the user to manipulate and place objects in a virtual environment efficiently, as it likely would be done in the real world. The method is general and independent of manipulation device, therefore it can also reduce the work strain on the user. We design two experiments to compare the task of constructing a simple toy in a real versus virtual environment with two manipulation devices. In the first experiment, we use a 6-DOF mechanical tracker (ADL-1™) as the manipulation device in the virtual environment. And in the second experiment, we use a 6-DOF magnetic tracker (Fastrak™) attached to a lightweight block. For the real environment, we use our hand to manipulate a set of real blocks from which the virtual objects were modeled. The results from both experiments show that the virtual task with manipulation aid is close to the real task in distance accuracy and completion time. Also, results from both tracking devices suggest that the manipulation aid method is independent of manipulation device.

Keywords: Virtual Environment, Manipulation Aid, Natural and General Method, Comparison between Virtual and Real Task

1 Introduction

A virtual environment created by computer graphics and having appropriate user interfaces can be used in many applications [1, 2]. A sophisticated interface for virtual environment is possible with virtual reality techniques, which can be used to provide an intuitive user interface using human spatial perception, unfortunately, perfect virtual environments are difficult to achieve due to limitations in computational power, so a simple task in a real environment often becames an operation requiring skill in a virtual environment. In order to accomplish even an easy task in a virtual environment efficiently, as it likely would be done in the real world, it is considered to be necessary to calculate and simulate such factors as the avoidance of intersection from the test of interference among virtual objects, the fall of virtual objects caused by gravity, and friction between objects.

A useful way to provide the user with such a natural user interface in a virtual environment is to restrict the degrees of freedom (DOF) concerning the motion of objects or the user's hand. Two main ideas exist to restrict the DOF. The first is to restrict the DOF of the user's hand motion with devices such as force feed-back tools which generate a reaction between two faces touching each other. The scond is to restrict the DOF of the object motions without restricting the motion of the user's hand. In the first case, the user must be equipped with special hardware (force feed-back devices) capable of generating an accurate reaction to restrict his hand motion or between two faces touching each other [3–5]. On the other hand, a simple configuration is sufficient for the second one; however, careful verification is necessary, because the user may have a sense of incompatibility caused by a difference between visual feedback and motor control. Moreover, if the provided interface is not natural nor intuitive, a user's moter system will have to adapt to the sensory conflict with taking a lot of time.

There has been considerable study in assisting the manipulation of objects in virtual reality [6–12]. However, most examples are limited to a single level of constraint complexity or special functions are attached to the objects in advance. Moreover they consider only single face-to-face interactions and the number of attracting faces is limited to only one for each object. Therefore it is not flexible and cannot apply to a variety of tasks employing multiple objects with complicated shapes. Non of them discussed on the "naturalness" of the method in campare with the object manipulation task which we usually do in the real world.

In this paper, we propose a manipulation aid method based on the visual feedback using the dynamic constraints among object faces. It enables the user to manipulate and place objects in a virtual environment efficiently, as it likely would be done in the real world. The method is general and independen-

*Currently with the Simon Fraser University, Canada, amyy@sfu.ca

t of manipulation device, therefore it can also reduce the work strain on the user. The experimental results which show the effectiveness of the proposed method particularly when the user is requested to precisely place a virtual object in a certain location is reported in [13]. In this paper, we design two experiments to compare the task of constructing a simple toy in a real versus virtual environment with two manipulation devices. In the first experiment, we use a 6-DOF mechanical tracker (ADL-1™) as the manipulation device in the virtual environment. And in the second experiment, we use a 6-DOF magnetic tracker (Fastrak™) attached to a lightweight block. For the real environment, we use our hand to manipulate a set of real blocks from which the virtual objects were modeled.

2 Manipulation Aid Based on the Dynamic Constraints among Faces

2.1 Constraints among Faces

Suppose a simple task is to place blocks on a table. The block has six planar sides which are connected perpendicularly. We consider the following three conditions.

One Face Constraint When a block is placed on the table as shown in figure 1 (a), the motion of the manipulated block is constrained by the upper surface of the table. In this case, by constraining one pair of faces, the DOF of block motion is restricted to 3 (2 translations and 1 rotation) from 6 which it originally had.

Two Faces Constraint Subsequently, if the second block is aligned adjacent to the first one on the table as shown in figure 1 (b), the motion of the manipulated block is constrained by two faces (i.e. the upper surface of the table and a touching face of the first cube). In this case, the DOF of the block motion has to be restricted to 1 (1 translation).

Three Faces Constraint When a block on the table is aligned against two other blocks as shown in figure 1 (c), the aligned block has no DOF. It is constrained by three faces (i.e. the upper surface of the table and two touching faces of the blocks).

2.2 Method for Manipulation Aid

The method using the constraints among object faces that are dynamically selected while the user manipulates the object is described (for details see [13]).

2.2.1 Outline of Method

Our manipulation method is a visual technique which restrict virtual object motion but not user's hand motion, therefore the distinction between actual and displayed object position is made. The actual object position is controlled by the user's hand while the displayed object position is what the user sees. The key

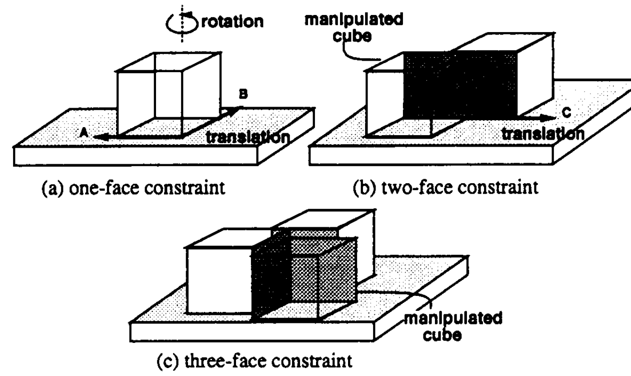

Figure 1: Constraints among faces for manipulation aid.

to the method is how the manipulation aid modifies the object position to provide intuitive and compatible assistance for the user during object interaction. In order to determine how an object should be constrained, accurate collision data is required. The efficient collision detection algorithm [14] is used. This algorithm detects colliding pairs of faces in real-time for three-dimensional graphical environments where objects are undergoing arbitrary motion. The algorithm can be used directly for both convex and concave objects.

The manipulation aid algorithm consists of three parts: constraint selection, constrain object motion, and constraint release. The next sections describe how these work.

2.2.2 Constraint Selection

An object with a complicated shape may have a number of colliding face pairs detected by the collision detection algorithm. By examining the geometry between the collision pairs and speed of interaction, we can guess the intention of the user and dynamically select the best faces to constrain. Since we are only considering face to face interaction, we can apply several conditions to each colliding face pair to reduce the number of possible pair candidates. These conditions are (1) at least one face in the pair is moving, (2) angle between the two face normals is more than 120 degrees, and (3) ratio of overlap area between the two faces to the smaller face is more than a certain threshold

For face pairs which satisfy all of the above conditions, an attraction value is calculated using the following equation.

$$attraction = rC_r + vC_v \qquad (1)$$

where, r is the rotation angle between the two face normals, v is the angle between the moving object's velocity vector and the colliding face of the target object, d is the distance of the moving face projected onto the target face, C_r, C_v, and C_d are parameter coefficients. The collision face pair which have the highest attraction value is selected for constraint.

2.2.3 Constraining Object Motions

If a valid face pair is found as described in the previous section, the displayed object position must be modified to reflect the new assisted position. The grasped object is translated to the surface of the selected face and a rotation matrix is applied to make the chosen face pair parallel. The distance to translate is calculated by projecting the center of the moving face onto the target face. The rotation matrix is found by:

$$M_{rot} = T(vtx)A(\theta, \vec{v})T(-vtx) \qquad (2)$$

where, M_{rot} is the rotation to apply to the current object position, vtx is the center of gravity of the moving face, and $A(\theta, \vec{v})$ is the matrix for making the moving face parallel to the target face, $\vec{v}$ is the normal vector orthogonal to the two face normals, θ is the angle of rotation found by the dot product of the two face normals.

After the object position is constrained onto the target face, further motion of the object is restricted by the currently constraining faces until the user deliberately exits from constraint. The method for release from constraint is described in the next section.

2.2.4 Release From Constraint

Since the manipulation method uses an intuitive "magnetic" attraction to constrain an object, the method for release from a constraint should be intuitive also. We call the release action "unsnap" for obvious reasons. Two conditions allow a constrained object to unsnap from a particular face: overlap ratio and distance from face. A release detected by either one of these conditions is sufficient for unsnap.

Overlap Ratio Checking the overlap area ensures that the object is constrained only when it is still touching another object. To take into account different object sizes and scaling, the overlap ratio is used. The Overlap Ratio is the intersection area of the assisted faces over the area of the smaller face in the pair. Hence the overlap ratio is' a percentage of the smaller face area. The following condition is tested:
when $overlap\_ratio < overlap\_threshold$
unsnap from face
$overlap\_threshold$ value should be a value near zero with hysteresis to prevent object snap and unsnap in borderline cases.

Distance from Face We have discussed how the displayed object position is modified from the actual position using various translation and rotation constraints. And we saw that the delta movement of the actual position (from hand input) is used to determine the constrained movement. Here, we look at how the distance of the actual position from the constrained face can be used to unsnap from a surface.

The Distance from Face is the distance of the current hand position to the constrained face of the target object. Checking this distance allows the user to deliberately unsnap from a face by pulling far enough away.

when $distance > dist\_threshold$ unsnap from face where $distance$ is the distance from current hand position to constrained face of target object, $dist\_threshold$ is a dynamic threshold calculated below.

To better simulate the magnetic property of our constraint method, we modeled a simple "magnet" which has equal magnetic fields perpendicular to the surface. Hence, we assume the greater the contact area, the more force is required to pull the object away from the surface, thus a greater distance is required to unsnap from the surface. We call this a dynamic threshold because it changes with overlap area.

$$dist\_threshold = k\sqrt{A} \qquad (3)$$

where, A is the overlap area, and k is a user-adjustable parameter The parameter k is adjusted accordingly to provide the best feeling of magnetic behavior in a simulation.

3 Experimental Setup

Figure 2 shows the hardware configuration of our experimental system. All input and output devices and sensors are controlled by an SGI ONYX workstation. A 70-inch CRT projector displays position tracked stereoscopic images. User eye position is derived from a 6 DOF magnetic sensor attached to LCD shutter glasses used for stereo viewing. Accordingly, the system can present non-distorted images with depth sensations and motion parallax. The user can grasp and manipulate objects using the ADL-1™, a 6 DOF mechanical tracker or the Fastrak ™, a 6 DOF magnetic tracker, both connected to a serial port of the workstation.

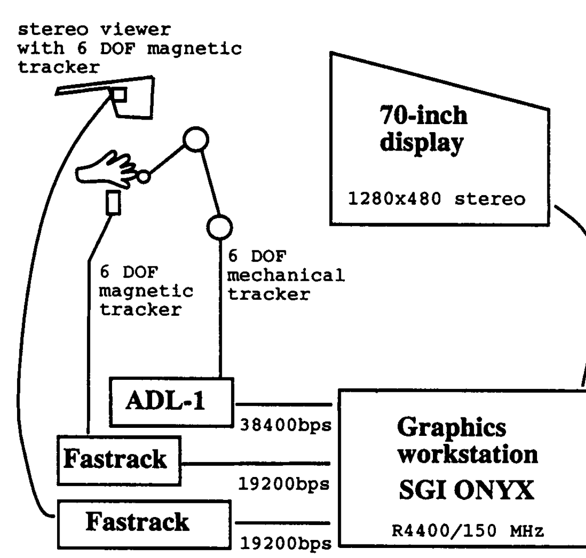

Figure 2: Hardware configuration for virtual and real task experiments.

4 Experimental Method

The task is to construct a snail using five predefined blocks. Figure 3 shows the initial positions of the blocks and figure 4 shows the completed snail. The virtual blocks are modeled to resemble the real blocks in geometry, size and color.

The assembly of the toy snail is carried out in three different modes:

Virtual without aid: assemble virtual blocks without manipulation aid and no collision color cue

Virtual with aid: assemble virtual blocks with manipulation aid and color cue

Real: assemble real blocks

In order to compare the virtual and real manipulations as best as possible, several rules were established for the manipulation of real blocks. These rules are:

- only the thumb and one other finger of one hand are allowed to grasp an object
- turning of the object with fingers is not allowed; instead turn the wrist or arm
- only the grasped object may be touched
- during the time between object grasp and release the subject's elbow or palm cannot rest on any structure (i.e. table, other blocks)
- objects already placed should not be moved while placing other blocks

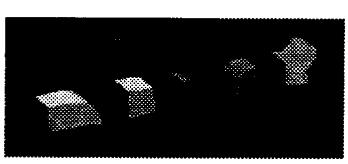

Figure 3: Initial positions of blocks for toy snail.

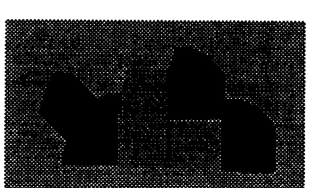

Figure 4: Finished construction of toy snail.

4.1 Method for Accuracy Evaluation (Stage 1)

The purpose of stage 1 is to compare the distance accuracies and completion times for the three modes. Subjects were asked to build the toy snail as quickly and accuracy as possible. Two measurements were taken: *completion time* and *distance accuracy.*

Completion time is the real clock time from grasp of the first block to release of the last block in the assembly sequence. *Distance accuracy* is the sum of the distance error between all adjacent vertices.

In the real environment, the distance error is difficult to measure precisely because of the imprecise shape of the real blocks. However, rough estimates of several trials yielded errors of less than 2mm per vertex. Hence, we assume a maximum distance error of 2mm per vertex for the real task. In this stage, eight trials were completed in each of the three modes.

4.2 Method for Efficiency Evaluation (Stage 2)

The purpose of stage 2 is to compare the time required to construct the snail within a certain distance accuracy. The accuracy requirement was selected to be 3mm per vertex. Ideally, 2mm per vertex would best correspond to the real task case, but the task becomes considerably difficult for the virtual task without manipulation aid, particularly with limitations of the mechanical tracker (which are discussed in the results section). The task is to place each object until the distance error falls below 3mm per vertex, which is indicated by a change of color in the object. In this stage, eight trials were also carried out in each of the three modes and completion times were measured.

5 Results with Mechanical Tracker

This section presents the results from the first experiment using the mechanical tracker. The results from stages 1 and 2 are given, followed by user feedbacks.

Five young subjects participated in the experiments, their ages ranging from mid-twenties to early thirties. There were three males and two females; all had technical and/or VR experience. The experiments consisted of two stages, accuracy evaluation and efficiency evaluation. Prior to stage 1, the subjects practiced using the system in the various modes to become familiar with the virtual environment behaviors and system hardware in the experiment. During the trials for both stages, subjects took short rest breaks as were required.

5.1 Accuracy Evaluation (Stage 1) Results

A plot of the completion time versus distance accuracy for one subject is shown in figure 5. There are three distinct groups of data, corresponding to the three modes. For the real task, the distance accuracy of all the points has been set to a maximum estimated value, with the actual accuracy somewhere between 0 and the maximum. The maximum error was estimated by assuming a 2mm error per vertex, and multiplying by 14 vertex pairs yields 28mm for maximum distance accuracy for the real task.

From the plot we see the group with manipulation aid is closer to the real task than without; it has distance accuracies in a close range to the real task, but taking slightly more time. The time delay may be due to the nature of a virtual environment and limitations in using a mechanical device instead of hand and fingers. Feedbacks on the mechanical tracker will be discusssed further in section 5.3. As expected, the tasks without manipulation aid had higher errors and completion times.

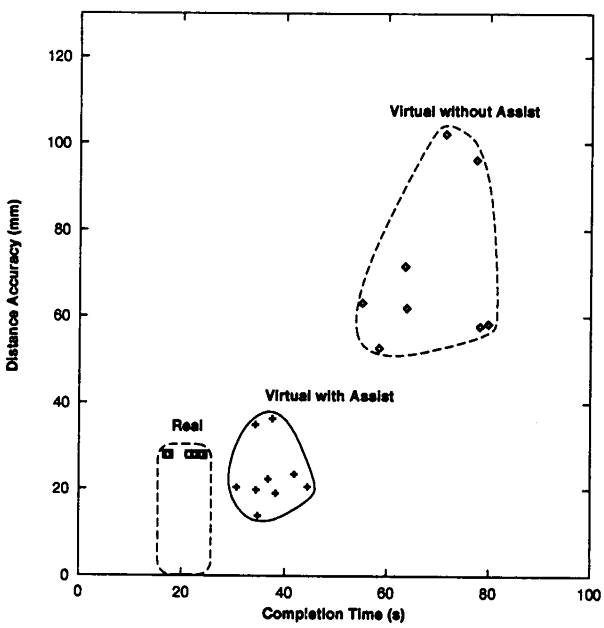

Figure 5: Stage one results with mechanical tracker for one subject.

5.2 Efficiency Evaluation (Stage 2) Results

In this stage, objects were placed within a distance accuracy of 3mm per vertex. The completion times were measured and the average times for each subject are plotted in Figure 6. The average times with manipulation aid are generally very close to the real task. The difference may be due to the limitations of the virtual environment, such as graphics resolution and device awkwardness.

To compare the time difference across all subjects, time percentages were calculated for each subject as follows: $percentage = \frac{t_M - t_{realavg}}{t_{realavg}} \times 100\%$, where t_M is the average time for either with aid or without, and $t_{realavg}$ is the average time for real task. Figure 7 shows the percentage results. The dashed line separates the two modes. With manipulation aid, all subjects took less than twice the real task time, whereas without the aid took up to five times longer. The most efficient virtual task took only 8% longer than the real task. These percentages indicate that while the virtual task with manipulation aid requires more time than the real task, the additional time is typically less than twice.

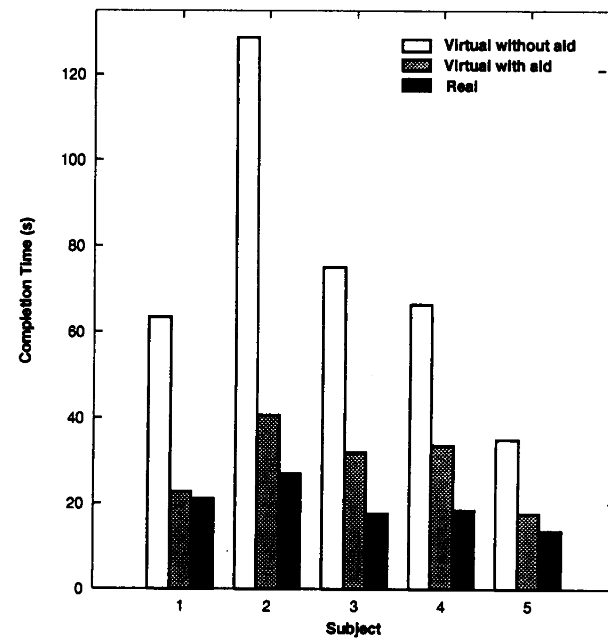

Figure 6: Stage two results with mechanical tracker for all subjects.

5.3 Observations and Feedbacks

The purpose of this section is to discuss some of the user feedbacks and observations from the previous experiment to see what factors may affect the performance of the virtual task. The feedbacks also help us to understand more about the experimental results and lead us to further experiments to eliminate potential biases.

1. Object grasping with mechanical tracker
   The mechanical tracker (ADL-1) has a limited roll angle which placed an undesirable constraint on the rotation of objects especially without manipulation aid. Three of the five objects in the snail assembly required rotations. The user had to learn to grasp the objects at certain angles to facilitate the rotations required for placement of objects. This restriction in object grasping may be the main cause for longer completion times and lower accuracy for the virtual task without manipulation aid.

   Another cause for longer time is the grasp misses that happen occasionally when the user attempts to grasp objects too quickly. This occurs due to the difference between virtual object grasping and real object grasping.

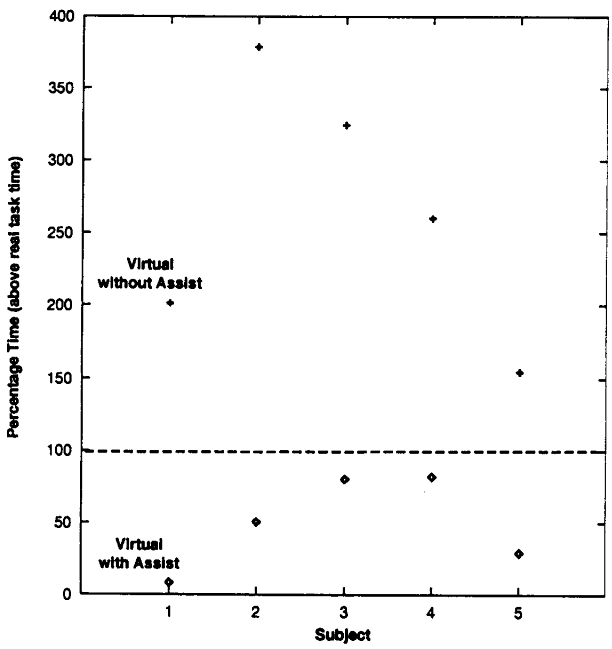

Figure 7: Time percentages from stage two with mechanical tracker.

2. Object handling with mechanical tracker
Most of the users found the mechanical arm to be slightly heavy, causing some arm strain and requiring more work compared to the real task. In addition, the grasping of the mechanical tracker end effector (EE) is unnatural compared to the grasping of a real block. A real block is grasped with thumb and one other finger, while the mechanical tracker EE is held in the palm. The EE can be held with two fingers similar to a real block, but the weight of the mechanical tracker would cause too much strain. The unnatural way of manipulating objects using the mechanical tracker may account for the longer completion time and lower accuracy.

3. Absence of force feedback
In the real environment, there is force feedback when a collision with another block occurs. When one object is placed onto another object, the bottom one supports the top one, therefore less work is required in horizontal movement. In the case of the mechanical tracker, the user still has to hold the mechanical tracker weight without any support or force feedback for movement in any direction.

4. Virtual versus real environment
In the virtual environment, users tend to spend some time observing or mentally evaluating the placement of the previous block before continuing. This does not occur in the real environment. Since the completion time is measured from beginning of the assembly to the end, the slower behavior of the user in the virtual environment results in a longer completion time.

   The user's hand position in the virtual world is sensitive to the movement of the user's head during object alignment. The viewing position of the user is tracked by a magnetic sensor attached to the side of stereographics glasses. When the user adjusts his view, the alignment is often upset, hence resulting in longer task completion time. The quality of 3-D effect, contrast and screen resolution of the virtual system may also contribute to more error and time compared to the real task.

The above observations and feedbacks suggest that the mechanical tracker may be a major factor in determining the performance of the virtual tasks. To test this hypothesis and compare the performance of the manipulation aid method with another device, we replaced the mechanical tracker with Fastrak, a compact and lightweight 6-DOF magnetic sensor which can be held naturally with the hand. The next section contains the results from experiments using this new device.

6 Results with Magnetic Tracker

This section presents the results from the second experiment in the construction of a toy snail using magnetic tracker instead of the mechanical tracker. In this experiment, only stage 2, the task requiring a 3mm distance accuracy, was done. The results for this stage and user feedbacks are given.

6.1 Efficiency Evaluation (Stage 2) Results

All of the users found the virtual task to be easier with the magnetic sensor than mechanical tracker. The average completion time for each subject are shown in figure 8. For comparison with mechanical tracker, the time scale is the same as the graph in figure 6 and the times for the real task are also the same as before. Looking at the two graphs, we can see a significant decrease in task time for the magnetic sensor in the virtual task without manipulation aid. This result agrees with the hypothesis that mechanical tracker limitations affect the task time. However, this effect seems to dominate only when manipulation aid is not used, because the times for virtual with aid (middle bar) do not differ by much; that is there was not much gain from using magnetic tracker over the mechanical tracker with manipulation aid. The reasons for this is discussed in the device comparison later in this section.

The time percentages are plotted in figure 9 with the same method as in section 5.2. Again, the dashed line separates the two modes, except for one point. With manipulation aid, all subjects took less than about 70% additional time in the virtual task, whereas without aid took 65 to 200% more time. Comparing these results to the mechanical tracker graph, we see a gain of a factor of two for the virtual task with manipulation aid, whereas the percentages with the aid are similar. A direct comparison of the mechanical tracker and magnetic tracker is discussed next.

The completion times for both devices are graphed together in figure 10. Three subjects performed better with the mechanical tracker while two subjects were better with magnetic tracker. The time differences are small, with the largest difference being about 4 seconds. There are two possible explanations for the same or slightly worse performance in using magnetic tracker. The first is that the manipulation aid method may be independent of device so that the time required are almost the same for a particular task. The second explanation is the difference of physical configuration of the mechanical tracker and magnetic tracker. The mechanical tracker end effector is located directly in front of the user's body, therefore arm motions are relatively small and close to the user. On the other hand, the operating space of the magnetic sensor is further away from the user and closer to the projection screen. This configuration is more natural since the virtual object location corresponds closer to the hand location (but not exactly because of occlusion by hand). Hence, to complete the task with the magnetic sensor, the user has to reach further away from the body, which can account for an increased completion time or counterbalance any time savings over the mechanical tracker, in the case with manipulation aid.

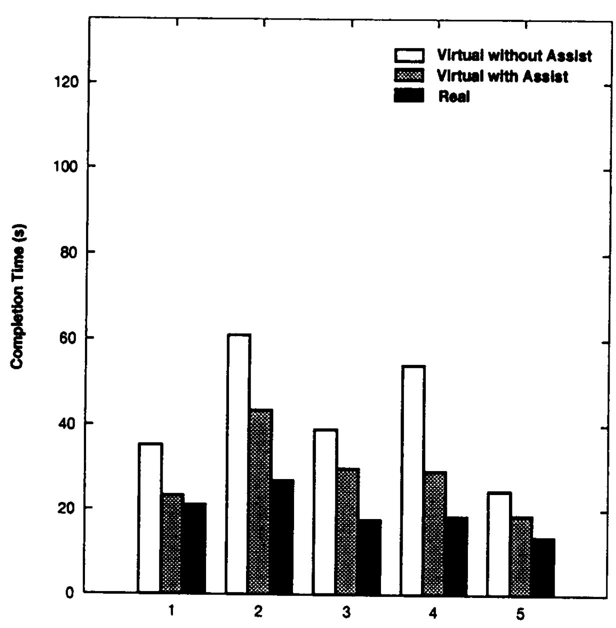

Figure 8: Stage two results with magnetic tracker device for all subjects.

6.2 Observations and Feedbacks

1. Easier to use than mechanical tracker
All the users felt that the magnetic sensor was easier to use than the mechanical tracker because it was lighter and rotations were no longer restricted by the shape of the device, but only in the limitations of the user's wrist and arm. This accounts for time saving when the placement was difficult because manipulation aid was not used.

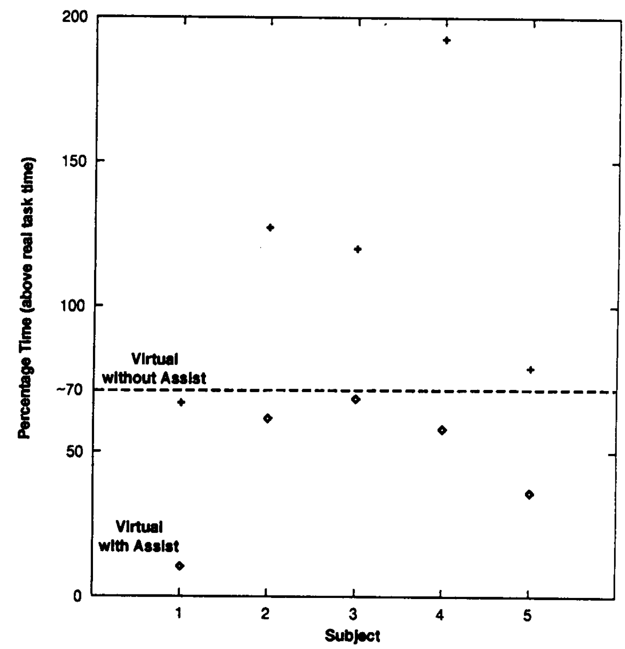

Figure 9: Time percentages from stage two with magnetic tracker.

2. Virtual versus real environment
The method of grasping is the same as for the mechanical tracker, hence the problem of grasp misses still exists. Even though the user's hand correspond closer with the graphics, it does not meet exactly, therefore the user requires more time to grasp and place objects than in a real task. When asked about the difference between the real and virtual task with manipulation aid, one subject commented on the absence of sound in the interaction of the blocks.

7 Conclusion

In this paper, we presented two experiments for the construction of a toy snail with simple blocks in both real and virtual environments. The results from both experiments show that the virtual task with manipulation aid is close to the real task in distance accuracy and completion time.

Two manipulation devices were used in the experiments, a mechanical tracker (ADL) and a magnetic tracker (Fastrak). The magnetic sensor allowed for more natural manipulations similar to handling real blocks. When the performance of both devices are compared for tasks with aid, the completion times are close. This suggests that the manipulation aid method is independent of manipulation device, yielding better

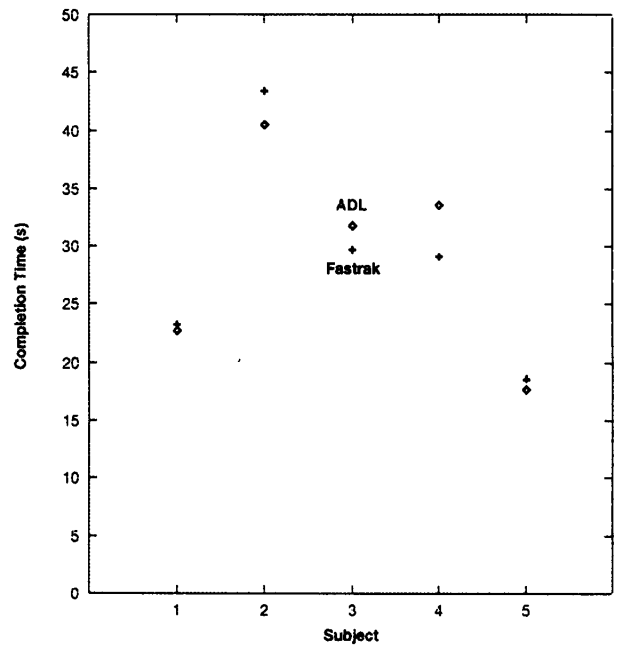

Figure 10: Comparing time percentages of two devices.

accuracy and times over a task without such an interface aid.

The manipulation aid method has proven to be a valuable tool for the virtual task of constructing a simple toy. With this user interface aid, a virtual task is more natural and simpler, but is still a step away from realism since only visual feedback is exploited. Perhaps when the three senses of sight, touch and sound are all influenced together with the integration of manipulation aid, tactile feedback and sound feedback, then we may approach towards a true virtual reality.

Acknowledgements

The authors would like to thank Dr. Haruo Takemura, Associate Professor of the Nara Institute of Science and Technology, Japan for his useful discussions.

References

[1] Barfield, Woodrow and Furnes, Thomas, editors. *Virtual environments and advanced interface design.* Oxford University Press, 1995.

[2] Earnshow, R. A., Vince, J. A., and Jones, H., editors. *Virtual Reality applications.* Academic Press, 1995.

[3] Iwata, Hiroo. Artificial reality with force-feedback: development of desktop virtual space with compact master manipulator. *Computer Graphics*, Vol. 24, No. 4, pp. 165–170, 1990.

[4] Ishii, M. and Sato, M. A 3D interface device with force feedback: a virtual workspace for pick-and-place tasks. In *Virtual Reality Annual International Symposium*, pp. 331–335. IEEE, 1993.

[5] Hirota, K., Hashimoto, T., and Hirose, M. CAD/CAE application of force display. In *7th Symposium on Human Interface*, pp. 95–98, 1991. (in Japanese).

[6] Baraff, David. Analytical methods for dynamical simulation of non-penetrating rigid bodies. *Computer Graphics*, Vol. 23, No. 3, pp. 223–232, 1989.

[7] Bier, Eric A. Snap-dragging in three dimensions. In *Computer Graphics, 1990 Symposium on Interactive 3D Graphics*, pp. 193–204. ACM, 1990.

[8] Venolia, Dan. Facile 3D direct manipulation. In *INTERCHI*, pp. 31–36. ACM, 1993.

[9] Chanezon, A., Takemura, H., Kitamura, Y., and Kishino, F. A study of an operator assistant for virtual space. In *Virtual Reality Annual International Symposium*, pp. 492–498. IEEE, 1993.

[10] Fa, M., Fernando, T., and Dew, P. M. Interactive constraint-based solid modeling using allowable motion. In *Symposium on Solid Modelling and Applications*, pp. 243–252. ACM/SIGGRAPH, 1993.

[11] Snyder, John M. An interactive tool for placing curved surfaces without interpenetration. In *Computer Graphics, Annual Conference Series*, pp. 209–218. ACM, 1995.

[12] Kijima, R. and Hirose, M. The impetus method for the object manipulation in virtual environment without force feedback. In *Symbiosis of Human and Artifact*, pp. 479–484, 1995.

[13] Kitamura, Y., Yee, A., and Kishino, F. A sophisticated manipulation aid in a virtual environment based on the dynamic constraints among object faces. In *International Conference on Systems, Man and Cybernetics*. IEEE, 1995.

[14] Smith, A., Kitamura, Y., Takemura, H., and Kishino, F. A simple and efficient method for accurate collision detection among deformable polyhedral objects in arbitrary motion. In *Virtual Reality Annual International Symposium*, pp. 136–145. IEEE, 1995.