Toward Machine Mediated Training of Motor Skills
-Skill Transfer from Human to Human via Virtual Environment-

Yasuyoshi Yokokohji †*, Ralph L. Hollis ††, Takeo Kanade ††
Kazuyuki Henmi †, and Tsuneo Yoshikawa †

†Department of Mechanical Engineering, Kyoto University, Kyoto 606, JAPAN
††The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A

{yokokoji | henmi | yoshi}@mech.kyoto-u.ac.jp   {rhollis | tk}@cs.cmu.edu
http://www.cs.cmu.edu/~msl

Abstract

In this paper, we investigate a possibility of skill mapping from human to human via a visual/haptic display system. Our goal in the future is to develop a training system for motor skills such as surgical operations. We have proposed a new concept of visual/haptic display called a WYSIWYF Display *(What You See Is What You Feel). The proposed concept ensures correct visual/haptic registration which is important for effective hand-eye coordination training. Using the prototype WYSIWYF display, we did a preliminary experiment of skill training. Our idea of skill transfer is very simple; basically it is a "record-and-replay" strategy. Questions are "What is the essential data to be recorded for transferring the skill?" and "What is the best way to provide the data to the trainee?". Several methods were tried but no remarkable result was obtained, presumably because the chosen task was too simple.*

1 Introduction

Today effective motor skill training is becoming more important, especially in the area of medicine. Many new technologies are being introduced, such as laparoscopic, endoscopic, arthroscopic surgeries, and it gets difficult for medical students to complete their training for these new technologies within a limited period. It also gets difficult for medical doctors and surgeons in service to adapt new technology and update the knowledge and skills.

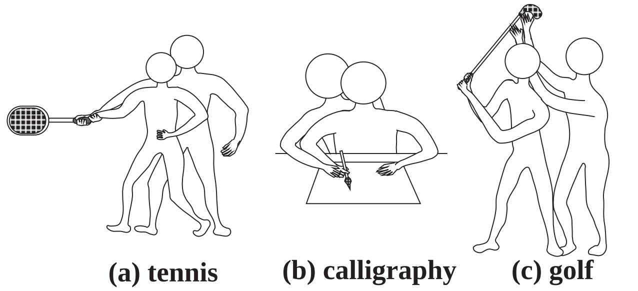


Figure 1: Motor skill training by human instructor

Training is a kind of skill mapping from a person (trainer or expert) to another (trainee or learner), sometimes directly and sometimes through medium such as textbook and video tape. In general, transferring skills from person to person is not easy, and training takes a lot of time to complete. Especially transferring motor skills through media is very difficult because most skills are nonlinguistical, difficult to describe and even invisible, such as haptic senses. Physical guidance by the instructors, as shown in Fig. 1, is effective for motor skill training. Recently, virtual reality (VR) technology has been getting considerable attention as a new kind of medium for training, which could substitute the human instructor in Fig. 1.

In this paper, we investigate a possibility of skill mapping from human to human via a visual/haptic display system. Our goal in the future is to develop a training system for motor skills such as surgical operations. We have proposed a new concept of visual/haptic display called a *WYSIWYF Display* (What You See Is What You Feel) [18]. The proposed concept ensures correct visual/haptic registration which is important for effective hand-eye coordination training.

The first prototype of WYSIWYF display was built. A simple tube manipulation was chosen as a target task to be transferred. Our idea of skill transfer is very simple; basically it is a "record-and-replay" strategy [21]. Questions are "What is the essential data to be recorded for demonstrating the skill?" "What is the best way to provide the data to the trainee?". Several methods were tried in the preliminary experiment, but no remarkable results were obtained, presumably because the chosen task was too simple.

*Experimental study in this work was done when Y.Yokokohji was with The Robotics Institute, Carnegie Mellon University.

2 VR Training

2.1 Machine mediated training concept

One of the authors proposed a new direction of robotics: mechanical medium, i.e., using robotic mechanisms as media of motion intelligence or motor skill from human to human[17]. In Fig. 2, the top figure represents conventional robotic application where a human intelligence is somehow mapped to the target task via robot by means of program or teleoperation. The bottom figure, on the other hand, shows another possible application of robots (or more precisely saying, robotic mechanism), where a motion intelligence of someone is transferred to other via robotic mechanism. This paper aims at this new direction.

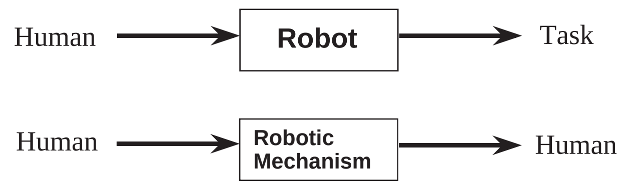


Figure 2: Mechanical medium concept

2.2 Previous research on VR training

The theory of identical elements, well known among behaviorist, says that the degree of transfer is a function of the identity of stimulus-response pairs between the original (training) task and the transfer (target) task[12]. There is no doubt that VR training is effective if the system can provide high enough fidelity. Flight simulator of airplanes is a typical example that has been successful.

In 60's, Hammerton examined visual factors affecting transfer from simulated to a real situation[5][6]. Recently, Kozak et al.[9] compared the value of real-world training, virtual reality training, and no training in the transfer of learning to the same task performed in real-world conditions. Although VR training made an improvement of the performance in VR environment, there was no transfer from a VR training to a real-world task. In Kozak's experiment apparatus, only HMD and data-glove were used without haptic interface. Reference[12] pointed out that the chosen task (picking and placing cans) was too easy and people are already proficient in that task in their daily life. To evaluate the value of transfer in training, the task should be more difficult, really requiring training.

There are several researches of VR training involving haptics. Colgate's group has been developing a haptic device to train astronauts for the docking task in space[11]. Kawamura et al.[8] developed a sports simulator with a wire-driven mechanism. Yoshikawa and Henmi[21] proposed a "virtual lesson" concept and constructed a virtual calligraphy training system. For medical applications, there are many challenges of VR simulator for training purpose[3][7][13]. Brett's lumber puncture simulator is one example using haptic interface[4].

3 Psychological Issues

3.1 Theories of motor learning

So far several psychological theories of motor learning had been proposed. Adams proposed a closed-loop theory of motor learning[1]. Contrary to earlier closed-loop theorists, Adams realized that in order to have the capability for the system to detect its own errors, two memory states must be exist: one to produce the action and one to evaluate the outcome. He called them the perceptual trace and the memory trace, respectively. Translating them into robotics terminologies, the perceptual trace would correspond to a series of joint sensor data for a specific skill motion (i.e., a desired trajectory in joint space), while the memory trace would correspond to a series of motor command data to generate that motion.

Inheriting some ideas from Adams, Schmidt proposed the Schema Theory[15]. Schmidt's theory also holds that there are two states of memory, a recall memory responsible for the production of movement, and a recognition memory responsible for response evaluation. In his theory, Schema is thought of as a general rule that can be used for generating, or selecting, a motor program. This theory also makes a distinction between slow positioning responses and rapid movements for which the recognition memory (or perceptual trace) presumably cannot be used during the response to guide the limb.

The Schema Theory explains the learning mechanism as follows. Given initial condition and desired outcomes, the recognition schema generates expected sensory consequences (expected proprioceptive feedback and expected exteroceptive feedback). The recall schema then generates a motor program, which is hopefully appropriate to the given situation. Getting the knowledge of results (KR), both schemas are refined. Once the correct recognition schema was established, the learner can continue to improve his motor program, even after KR is withdrawn, from the error between the expected sensory consequences and the actual sensory feedback. Schmidt and White[14] showed an evidence of this error detection mechanism based on the perceptual trace.

3.2 Some aspects of skill training

3.2.1 Adaptive training

One advantage of VR training over real-world training is that VR training system can provide supplemental cues to help the learner to improve the performance. A typical example is graphical cue for airplane landing in the flight simulator[10]. Adding augmenting cues is, however, concerned that subjects might learn to depend on the augmenting cues during early training, and are unable to use the important intrinsic cues effectively when the augmenting cues are withdrawn, especially when the task has no clear intrinsic feedback.

To prevent such a negative effect, an off-course schedule, in which the augmenting cues appeared only when subjects exceeded a prespecified error, might aid learning in a poor-intrinsic-feedback task while not permitting subjects to rely on the augmenting cues when they were performing at a reasonable standard. Such a strategy of showing augmenting cues can be regarded as adaptive training; the task is automatically and continuously transformed from easier to a more difficult version at a rate determined by individual learning. Lintern[10] showed in a simulator-to-simulator transfer-of-training design that adaptively trained subjects performed best in a transfer task than the control groups with constant-augmented-feedback and nonaugmented-feedback training.

3.2.2 Guidance

Another technique frequently used in teaching/training involves guidance, whereby the learner is in some way guided through the task that is to be learned[16]. Guidance could be a variety of procedures, such as physically pushing and pulling the learner through a sequence, visual cues to be followed by the learner, and verbal instructions. Guidance could be regarded as a kind of supplementary cues, but it is more direct assistance than other cues, especially in case of physical guidance. These guidance procedures tend to prevent the learner from making errors in the task. There have been two opposing views of whether or not guidance should be effective[16]. First, it can be argued that it is important for the learner to avoid making errors and guidance can effectively prevent many kinds of errors some of which could be even dangerous. Alternatively, we could argue that learning is most effective by trial-and-error. Guidance, under this view, prevents the person from receiving experience about errors, and thus learning might not be as effective as practicing the task under unguided procedures[16].

In machine mediated learning case, physical guidance by a robotic mechanism would be possible. As discussed in the motor learning theory section (3.1), there is an evidence of the error detection mechanism based on the perceptual trace. Physical guidance might be effective for strengthening this perceptual trace (in Adams' theory) or recognition memory (in Schmidt's theory). Although it might depends on the given target task, guidance may be most effective in early practice stage when the task is unfamiliar to the learner.

4 WYSIWYF Display

The authors has been emphasizing the importance of correct visual/haptic registration for visual-motor skill training, and proposed a new concept of WYSIWYF (What You See Is What You Feel)[18][19].

Figure 3 shows an overview of the prototype system in use. The user can see the virtual environment in the LCD (Liquid Crystal Display) as though looking through a window. Figure 4 shows an example of the displayed image to the user. Note that the virtual

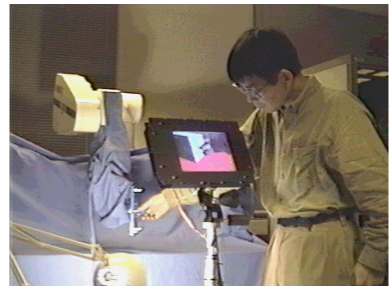


Figure 3: Overview of prototype WYSIWYF display

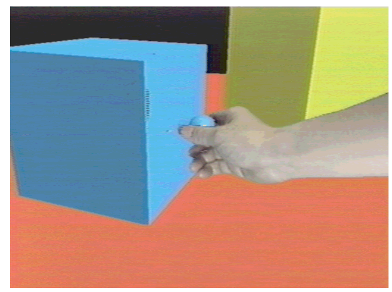


Figure 4: Displayed image example

object is a synthetic image but the user's hand image is a real image. WYSIWYF display provides visual and haptic sensations in a consistent manner. Therefore it would be suitable for the platform of visual-motor skill training. For more details of WYSIWYF display, see [18][19].

5 Motor Skill Training via WYSIWYF Display

5.1 Basic idea

In considering movements, especially skills, it is often difficult to isolate a movement from its environment. In other words, skill cannot exist alone but has its mean through an interaction with the environment. To construct a training system, therefore, it would be reasonable to build a virtual environment of the target task first and add a training mechanism on it.

Let us suppose that the target task is to manipulate an object. Our idea of skill transfer is very simple; basically it is a "record-and-replay" strategy[21]. First, an expert performs his/her skill in the virtual environment via a WYSIWYF display, and the system records all available data such as position and velocity of the target object and the applied force, locations of contact points, and constraint forces (See Fig. 5 (a)). Our WYSIWYF display uses the physically-based simulation algorithm by Baraff[2], which is a kind of *measuring force and displaying motion* approach[20]. As shown in Fig. 5 (a), the user's applied force, f_{input}, is first measured, and the algorithm solves constraint forces or impulses, if necessary. The algorithm then

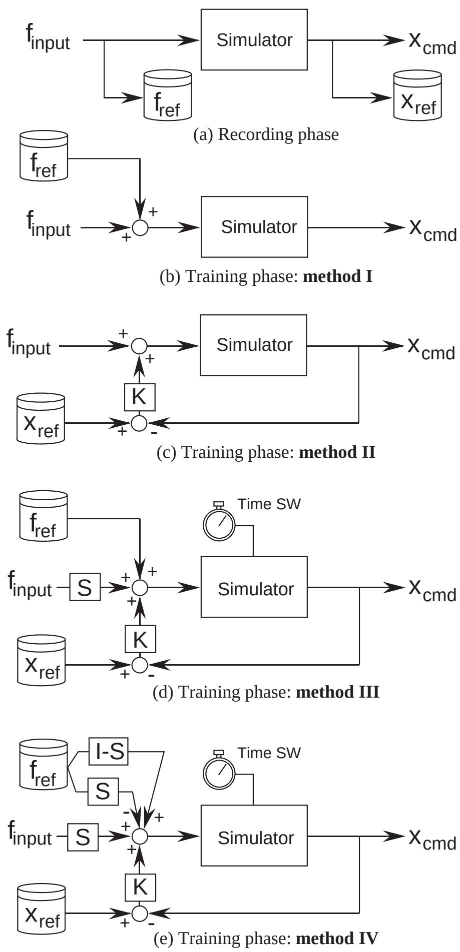


Figure 5: Training methods

outputs the position command $\boldsymbol{x}_{cmd}$ to the haptic device controller, solving ordinary differential equations. In the next subsection, we discuss some possible ways for motor skill training with VR system.

5.2 Some possible ways

Yoshikawa and Henmi[21] discussed the "virtual lesson" concept based on this strategy[21]. They proposed two methods for motor skill training: (i) motion playback with visual cue to display the desired force, and (ii) force playback with visual cue to display the desired motion, in which the desired but inverted force is replayed while eliminating the original constraint from the virtual environment. They also discussed switching the two methods, depending on the situation. Here five possible methods, which are extensions and/or modifications of their two methods, will be given as follows:

Method 0 (Visual cue)

Since we have the reference motion, we can visually display this motion to the learner so that he can follow it. If the target task is to manipulate an object, the reference motion can be displayed by a transparent object as shown in Fig.6. One problem of this method would be that the response time from visual stimulus is usually slower than from haptic or acoustic stimuli. If the target motion is fast, it would be difficult for the learner to follow it without delay. There is also a concern that the learner might rely on this supplemental cue and might not be able to perform the task well without it. As discussed in 3.2.1, it might be necessary to display this cue only when the motion error exceeds a certain threshold.

Method I (Visual cue + Force playback)

The next method to be consider is force playback method where the recorded force is "fed" to the simulator, upon the actual force applied by the learner. (See Fig.5 (b).) Since this method has no feedback mechanism, once the motion largely deviates from the reference motion, it would become no meaning to continue to feed the reference force.

Method II (Visual cue + Motion playback)

Instead of feeding the reference force in an open-loop manner, one can introduce a position feedback mechanism, regarding $\boldsymbol{x}_{ref}$ as the desired trajectory (See Fig.5 (c)). If we set the feedback gain K large, the system pulls the learner along with the reference motion by force. If the gain K is small, the simulator accepts the learner's force and the resultant motion can deviate from the reference motion. This method is a kind of physical guidance discussed in 3.2.2. If the learner can follow the trajectory with a small error, contribution from the feedback loop is small. Therefore this method has an adaptive mechanism in nature. We can also change the feedback gain adaptively, making it smaller as the learner improves his performance.

Method III (Visual cue + Hybrid playback)

One major problem of motor skill training is the difficulty of displaying the desired force, especially constraint forces. Suppose that the target task is to move a cube on a frictionless table while applying certain force in the normal direction of the table surface. In the training phase, we may want to replay the desired forces, $\boldsymbol{f}_{ref}$, as shown in Fig.7 (a). It is impossible, however, for the learner to know how much force is applied in the normal direction from the tangential motion. Visual cue such as bar graph would be an alternative way to show the desired force to the learner, but again response time to visual stimulus would be a problem. In addition, the desired force could be three dimensional force and moment in general, which would be difficult to graphically display.

One possible method to display the desired force is to introduce the time switch, by which the user can suspend and resume the simulation time, as shown in Fig.5 (d). We also introduce the selection matrix of the constraint, $\boldsymbol{S}$, so that the learner's force is accepted only along with the direction in which the object is constrained, expecting that the learner can understand the constraint state and the constraint

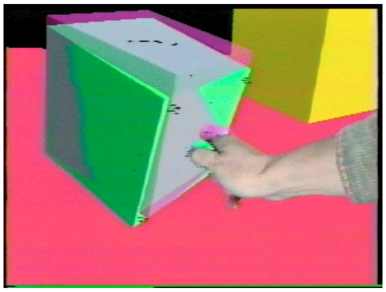
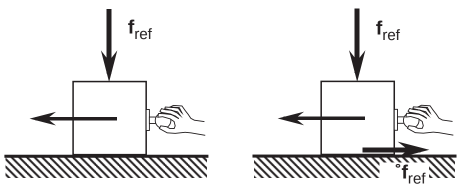


Figure 6: An example of skill training



(a) frictionless case (b) friction case

Figure 7: Sliding a cube on a rigid table

force more clearly When the learner pushes the time switch, the simulation is suspended. If the object is constrained at that time, he/she can examine how much the desired force is, by applying force in an opposite direction until the constraint contact is broken. By pushing the time switch again, he/she can resume the simulation.

Method IV (Visual cue + Hybrid playback with inverted force)

One problem of the previous method is that the learner has to suspend the simulation to examine the magnitude of constraint forces. One possible way to avoid this problem is to replay the constraint force in the opposite direction so that the object pushes back the learner (Fig. 5 (e)). The constraint may be broken, unless the learner applies correct constraint force in the original direction. The expectation is that the learner can simply apply forces in such a way that the constraint may not be broken, feeling the desired force without suspending the simulation.

5.3 Preliminary experiment and discussion

To evaluate the training methods discussed above, a preliminary experiment was carried out. Manipulating a cube on a frictionless flat table was chosen as the target task. One of the authors performed a cube manipulation, making face-to-face contact, edge-to-face contact, and apex-to-face contact, applying various forces against the table, and so on, in a virtual environment through the WYSIWYF display. Position of the cube, applied force, contact points, constraint

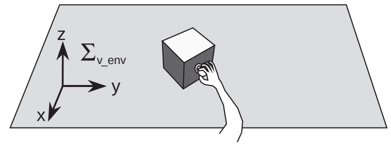


Figure 8: Base coordinate frame attached to the virtual table

forces and impulses at each simulation cycle (20 msec) are all recorded. The task lasts about 30 seconds.

The learner should be someone who has never performed the target task before, but in our preliminary experiment the same person, who demonstrated the target task, also tried the target task as a learner several days later. All of the above methods were tried with our prototype WYSIWYF display. Figure 6 shows the training phase in the experiment. To be honest, however, no remarkable result was obtained. One of the reason would be that the given task was too easy, simply manipulating a cube on a table, and there was no essential element that can be called “skill”. However we got some insights from this preliminary experiment.

First, Method I (Visual cue + Force playback) was not good as we expected. Once the motion runs off the target motion, force playback becomes no meaning and may even disturb the learner. With Method III (Visual cue + Hybrid playback), one can certainly know the constraint force when he suspends the simulation. But it was not an effective way. Method IV (Visual cue + Hybrid playback with inverted force) gives somewhat unnatural feeling to the learner (the object actively pushes against the learner). Even though the learner could apply the desired force with this method, he only worried about keeping the object being contacted to the table, probably ending up with a completely different task from the original target.

Finally Method II (Visual cue + Motion playback) with a small feedback gain seems most promising. Figure 9 shows plots of the desired/applied forces of one component (moment around x-axis of the base coordinates shown in Fig. 8). The dashed line represents the desired force and the solid one is the learner's response. Note that the learner's response is delayed (about 40 msec) from the desired trajectory because there is no force playback and the position feedback gain was set small; the learner tried to follow the motion mainly from the visual cue.

We cannot derive any conclusion like which method is the best from this preliminary experiment. At least the target task must be more complex (e.g., peg-in-hole with very tight clearance). For further investigation, upgrading the computation power of the system would be necessary so that more complicated constrained dynamics can be simulated.

We do not find any good way to display the de-

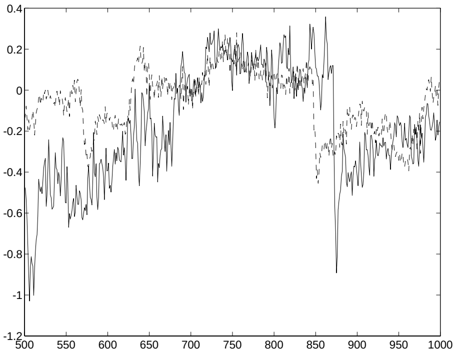


Figure 9: Experimental result (horizontal axis is simulation cycle (20 msec) count, vertical axis is moment [Nm])

sired force yet. Let us consider the situation of Fig. 7 again. If the table is frictionless and the object and table are rigid, it does not matter whether 10 N or 20 N is applied. When friction exists between the object and table, on the other hand, the normal force can be estimated through the tangential force that the learner can feel, assuming that he knows the friction coefficient. This discussion might lead us to the following hypothesis: "We need not worry about displaying forces which cannot be displayed anyhow".

6 Conclusion

In this paper, we investigated a possibility of skill mapping from human to human via a visual/haptic display system. Using the prototype WYSIWYF display, we did a preliminary experiment of skill training. Based on the "record-and-replay" strategy, several training methods were tried. Since the chosen task was too easy, simply manipulating a cube on a flat table, no remarkable result was obtained.

It was found that displaying constraint forces in a natural manner is difficult. If the situations shown in Fig. 1 were ideal, two haptic devices might be necessary: one for simulating the virtual environment and another for simulating the instructor. For further investigation to evaluate the effectiveness of training strategies, the target task must be difficult, really requiring training.

References

[1] J. A. Adams, "A Closed-Loop Theory of Motor Learning", Journal of Motor Behavior, vol.3, no.2, pp.111-149 (1971)

[2] D. Baraff, "Fast Contact Force Computation for Non-penetrating Rigid Bodies", In *Proc., SIGGRAPH'94*, pp.23-34 (1994)

[3] A. Bauer et al., "Virtual Reality in the Surgical Arthroscopical Training", In *Proceedings, 2nd Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, pp.350-354 (1995)

[4] P. Brett et al., "A Computer Aided Tactile Simulator for Invasive Tools", In *Proceedings 2nd Annual International Symposium on Medical Robotics and Computer Assisted Surgery (MRCAS'95)*, pp.324-328 (1995)

[5] M. Hammerton, "Transfer of training from a simulated to a real control situation", Journal of Experimental Psychology, vol.66, no.5, pp.450-453 (1963)

[6] M. Hammerton and A. H. Tickner, "Visual factors affecting transfer of training from a simulated to a real control situation", Journal of Applied Psychology, vol.51, no.1, pp.46-49 (1967)

[7] I. W. Hunter et al., "A Teleoperated Microsurgical Robot and Associated Virtual Environment for Eye Surgery", Presence, vol.2, no.4, pp.265-280 (1993)

[8] S. Kawamura et al., "Development of A Virtual Sports Machine using A Wire Drive System —A Trial of Virtual Tennis—", In *Proc. IROS'95*, pp.111-116 (1995)

[9] J. J. Kozak et al., "Transfer of Training from Virtual Reality", Ergonomics, vol.36, no.7, pp.777-784 (1993)

[10] G. Lintern, "Transfer of Landing Skill after Training with Supplementary Visual Cues", Human Factors, vol.22, no.1, pp.81-88 (1980)

[11] P. Millman et al., "A System for The Implementation and Kinesthetic Display of Virtual Environments", SPIE Vol.1833 Telemanipulator Technology, pp.49-56 (1992)

[12] National Research Council, "Virtual Reality - Scientific and Technological Challenges", National Academy Press (1995)

[13] R. M. Satava, "Virtual Reality and Telepresence for Military Medicine", Compt. Bid. Med., vol.25, no.2, pp.229-236 (1995)

[14] R. A. Schmidt, "Evidence for An Error Detection Mechanism in Motor Skills: A Test of Adams' Closed-Loop Theory", Journal of Motor Behavior, vol.4, no.3, pp.143-153 (1972)

[15] R. A. Schmidt, "A Schema Theory of Discrete Motor Skill Learning", Psychological Review, vol.82, no.24 (1975)

[16] R. A. Schmidt, "Motor Control and Learning (Second edition)", Human Kinetics Publishers, Inc. (1988)

[17] Y. Yokokohji, "Does Telepathic Robot Need a Heart?", Journal of the Robotics Society of Japan, vol.12, no.1, pp.65-66 (1994) *(in Japanese)*

[18] Y. Yokokohji, R. L. Hollis, and T. Kanade, "What You can See Is What You can Feel —Development of a Visual/Haptic Interface to Virtual Environment—", In *Proc. IEEE Virtual Reality Annual International Symposium (VRAIS'96)*, pp.46-53 (1996)

[19] Y. Yokokohji, R. L. Hollis, and T. Kanade, "Vision-based Visual/Haptic Registration for WYSIWYF Display", In *Proc. IROS'96*, (1996) *(to appear)*

[20] T. Yoshikawa et al., "Display of Feel for the Manipulation of Dynamic Virtual Objects", ASME J. of Dynamic Systems, Measurement, and Control, vol.117, no.4, pp.554-558 (1995)

[21] T. Yoshikawa and K. Henmi, "Availability of Virtual Lesson and Construction of Virtual Shuji System", In *Proceedings 1996 Annual Conference of The Virtual Reality Society of Japan* (1996) *(to appear in Japanese)*