

## Toward Visualization of Skill in VR: Adaptive Real-Time Guidance for Learning Force Exertion through the "Shaping" Strategy

Mikko J. Rissanen  
*Grad. School of Informatics,  
Kyoto University*  
mikko@kuhp.kyoto-u.ac.jp

Yoshihiro Kuroda  
*Grad. School of Engineering  
Science, Osaka University*  
ykuroda@bpe.es.osaka-u.ac.jp

Megumi Nakao  
*Grad. School of Information  
Science, NAIST*  
meg@is.naist.jp

Naoto Kume  
*Grad. School of Medicine,  
Kyoto University*  
kume@kuhp.kyoto-u.ac.jp

Tomohiro Kuroda  
*Dept. of Medical Informatics,  
Kyoto University Hospital*  
tkuroda@kuhp.kyoto-u.ac.jp

Hiroyuki Yoshihara  
*Dept. of Medical Informatics,  
Kyoto University Hospital*  
lob@kuhp.kyoto-u.ac.jp

### Abstract

*The authors aim to make principles of expert's haptic interaction explicit on a Virtual Reality (VR) based simulator. Our approach is based on visualization of significant components of the interaction with consideration on their importance which acts as the foundation for manual skill.*

*Delicate force exertion, which is the basis for various fine-motor skills, was chosen as an example. Expert's haptic interaction was recorded and presented to novices. Two visualization techniques were compared as a training aid by overlaying guidance on a simulator's screen: 1) tracking force curve (traditional technique), and 2) tracking the components of pre-defined "skill" (proposed technique). The visual presentation adapts to the components' importance: maximum power, duration, and force curve. The results support the possibility of using the proposed visualization technique for mediating principles of haptic interaction from experts to novices through a VR system.*

© 2007 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

### 1. Introduction

Our work aims to make manipulation-level principles of expert's manipulation and components of skill explicit in VR-based training. Today, training simulators with haptic feedback are a major research field, but relatively little is known about how to enhance simulators' efficiency for training. The future's training simulator should be capable of visualizing what should be learned, i.e. what the expert's "skill" consists of. This approach would add value to simulator-based self-learning by making invisible aspects of interaction visible to the learner.

Training simulators are envisioned to become complete learning environments. Shaffer et al. [1] introduced the idea of a surgical simulation system that enhances the simulation beyond instruction that could be given in the real world. Their vision covers artificial intelligence based instruction for decision making, such as guidance by "highlighting performance problems and by pointing out the relevant cues that the trainee should attend to," developed for flight simulators [2].

"Shaping" is a learning strategy for decomposing complex tasks, and it is one of the approaches desired in simulator-based training, e.g. in surgery [3]. The task is constructed little by little so that the difficulty can be mastered gradually. In haptic VR systems, the amount of information that can be measured from the expert's demonstration can be overwhelming. The expert's

insight is required to clarify the importance of each component of the interaction.

This study presents a concrete design that aims to the above-mentioned visions. A visualization technique which represents only the essential components of skill of the expert's demonstrations is proposed. The technique is demonstrated in a force exertion experiment with a medical expert's pre-recorded example motions. Force exertion is considered as a general root level skill. The skill was defined to consist of correct maximum force, duration, and complete force curve as time-series. Following the "shaping" learning strategy, the visualization was "built" little by little to meet all the criteria of the skill.

## 2. Related work

There are two research fields significant to this study. VR training systems have been studied extensively for the last decade. Psychophysics studies provide the basis for our visualization technique.

### 2.1. Skill training systems

**2.1.1. Traditional approach.** The traditional approach to learn movements and motor skills is based on observation and mimicking expert's examples. This approach has been incorporated into VR-based training environments. For example Just Follow Me (JFM) [4] supports learning from expert's recorded example movements by displaying the examples as a "ghost" that guides the user in a virtual environment. However, this approach does not draw attention to any components of skill, and it is capable of presenting the example as a single state only. In practice, the first trials with this approach require playback of the example as it is and practice using slow-motion playback. The learning process is based strongly on "shaping", but the insight of what the skill is composed of has to come outside the system.

**2.1.2. Haptic training systems.** As one of the early works, Yokokohji et al. [5] studied what data should be recorded and how it should be presented to the learner in order to transfer skills from human to human via a computer system. The initial results did not show any significant advantages in haptic guidance. Since the early studies, much research has been carried out. Feygin et al. [6] reviewed latest developments of haptic training systems. Most systems try to enhance learning by some kind of restrictions to the novice's actions on

the haptic modality. Haptic video [7] presented a pro-active motor-skill training method but the results on the learning effect were not fully conclusive.

Haptic guidance has not yet met a grand theory and more sophisticated methods are continuously studied.

Our goals require discussion on definitions of what should be learned from the simulator and a design for presenting that information to the user. Haptic training systems have proposed novel approaches for learning manual skills, but have not really addressed what the skill actually is. All observable information about the example motion is not always relevant to the success of a task, which leaves room for visualization of skill from the expert's point of view.

The intention of this paper is to demonstrate the visualization of skill as a method that draws attention and abstracts the actual example to consist only of the relevant components of skill. Since only visual cues are used, the proposed technique does not have any restriction to pro-activity during interaction.

### 2.2. Psychophysics

Studies by Pang et al. [8] and Tan et al. [9] have shown that Just Noticeable Difference (JND) of human is about 7% of force (tested at 2.5-10N), i.e. the perception of force through tactile feedback cannot be expected to be more accurate than 7%.

Srinivasan and Chen [10] studied human's force control ability with the aid of visual feedback. Subjects followed example time profiles: constant, ramp and sinusoidal shape. Their target maximum forces ranged 0.5-1.5N. Mean average error during each moment of time was about  $0.046N \pm 0.007SD$  (standard deviation) in a case of sinusoidal force profile with 0.5N maximum power and duration about 7 sec. In comparison to the results of Pang et al. [8], it was found that the error was about 11-15%. Later, Chen and Srinivasan [11] examined the similar interaction further with soft objects with different levels of force: 2, 4 and 8N. When tracking visually displayed level of constant target force, the subjects usually managed to maintain the force level at about 2% error. When trying to maintain the same force level without visual feedback, the error became 4-11%.

These results suggest that a graphical presentation can be used as a training aid for force exertion. In our study, expert's recorded force exertion is presented to subjects, who practice the task by mimicking the

visualized components of skill one by one in a way that follows the "shaping" learning strategy.

### 3. Visualization of skill

Visualization of skill is yet an unexplored area of VR. In the real world it is not possible for a novice to perceive the skill as it is, since the skill itself cannot be presented, only the appearance of one of its instances at a time. VR offers total freedom for any kind of presentation, which is often undermined.

Our design for visualization of skill is based on simplicity of the 2D visual cues and hiding the exact appearance of the example interaction so that only the essential components of the skill are presented to the user. The possibility for this type of presentation exists only in VR. Visualization serves four purposes:

- Principle of interaction is directly perceivable, i.e. playback of the interaction itself may not be needed.
- Accurate real-time feedback is achieved during training.
- Fully proactive training, i.e. the haptic modality is not restricted in any way.
- Also applicable to other phenomena than direct haptic interaction, for example indirect effects of actions on the target.

#### 3.1. Problems with the traditional visualization

Normally time-series data is shown as a curve (as in Fig. 1) to display maximal amount of details of the phenomena presented. This technique is capable of presenting the future states of an example, thus giving the user information that aid planning ahead how skilful motion should be produced.

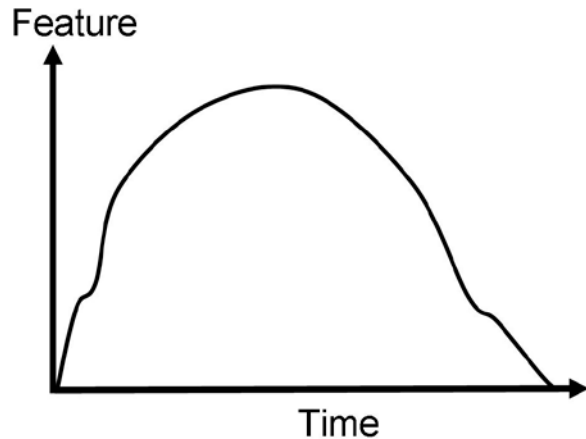


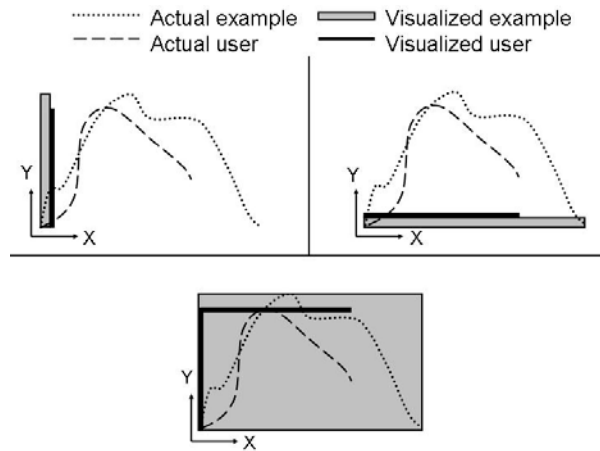
Fig. 1. TRADITIONAL visualization of time-series data as a curve.

By presenting all the details, the user's attention is drawn to every aspect of the example. Presentation of all the details may not be relevant to what the skill actually is composed of. The traditional curve presentation does draw the attention to different levels of significance of its components.

#### 3.2. General design for visualization of skill

The abstract presentation for skill can be constructed by using the simplest two-dimensional graph presentation for the chosen components. This is illustrated in Fig. 2. Assignment of the X and Y axes determines what the skill is composed of. This presentation is flexible to be adjusted to various components of skill.

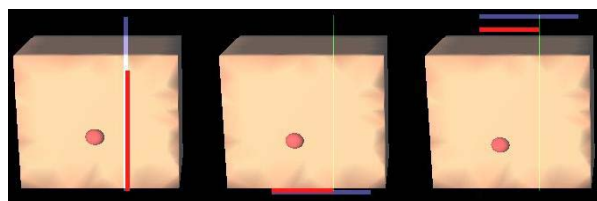
By presenting combinations of the components in the order of importance during training, the "shaping" learning strategy is supported. Practice of a complex skill can be started with a simple one-dimensional and later two-dimensional representations of one feature (e.g. use of force). When the first feature has been learned, another can be introduced (e.g. inflicted stress on the target).



**Fig. 2. GENERAL design for visualization of skill with a few example and user's performance presented in 1 and 2 dimensions. X and Y can be assigned to any feature of interaction.**

### 3.3. Design for force exertion skill

Force exertion was chosen to be a concrete example to demonstrate skill visualization of a fundamental interaction. If the proposed technique is beneficial in a relatively simple task, more complex tasks can be introduced to the user by adding more features to the presented example. Fig. 3 presents the visualization for each of the components and their combination (compare to the general design in Fig. 2.).



**Fig. 3. VISUALIZATION of force exertion that consists only of maximum power (vertical) and duration (horizontal). Blue: example, red: user.**

By displaying only the maximum preferred power of force exertion on the Y axis, the user's attention is drawn to exactly that component of skill. Duration is also presented only as one dimension, the X axis.

The combined display (Fig. 3, right) is capable of presenting two aspects accurately: maximum power

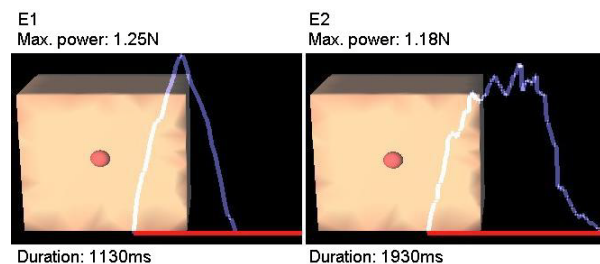
and duration as a box. Further information about the exact features of the example is hidden from the user.

Force exertion is presented here as a demonstration, yet, the same visual presentation could be applied to any other phenomena in the interaction with objects. For example, skill could be defined as an ability to move an object from a place to another without exceeding certain pressure threshold, to touch an object at frequent distance interval, or to tear tissue without inflicting too high stress on any part of the object.

## 4. Experiment

### 4.1. Expert's example data

A cardiovascular surgeon from Kyoto University Hospital performed palpation of the aorta on the MVL simulator [12] (Xeon 3.2GHz dual CPU with 4GB RAM) capable of real-time Finite Element Method based computation of deformation and reaction force for haptic feedback. His example performance using a Sensable PHANTOM™ was recorded at 100Hz sampling. Two excerpts were selected from the recorded data. The chosen curves and their appearance on the simulator are presented in Fig. 4.



**Fig. 4. EXAMPLES E1 and E2 selected from the expert's recorded performance. Blue: example force curve, red: user's curve (no force).**

### 4.2. Conditions

Two main conditions (training modes) were tested: curve visualization (CC, as in Fig. 4) and the proposed visualization technique, skill visualization (CS, Fig. 3).

Evaluation criteria for the force exertion skill were correct maximum power ( $max.p$ ), duration ( $dur$ ) and curve ( $cur$ ) compared to the expert's example.  $max.p$  was evaluated as error percentage to the example's maximum power,  $dur$  as error percentage to the

example's duration and *cur* as error percentage to the example's power at each moment sampled at 100Hz.

Following the "shaping" strategy, training phases were defined as *max.p*, *max.p+dur* and *cur* for E1, and *dur*, *dur+max* and *cur* for E2. The training mode differed for the two first phases so that for E1, CS consisted of vertical bars indicating the maximum powers in Phase 1 (Fig. 3, left), 2-dimensional bars in Phase 2 (Fig. 3, right) and the curves in Phase 3 (Fig. 4, left). The example was shown as a template into which the user's performance was drawn in real-time (visual RT feedback). For E2, CS contained horizontal bars displaying the durations in Phase 1 (Fig. 3, middle), 2-dimensional bars in Phase 2 (Fig. 3, right) and the curves in the last phase (Fig. 4, right). CC showed the curves as they were. CS was expected to draw attention to the selected components of the skill before displaying the full curve and, thus, to provide smaller errors at least in the first two phases.

On a screen of 1280x1024 pixels resolution 1 pixel on the X axis equaled to 10ms. The Y axis minimal perceivable difference for both E1 and E2 was 0.00625N/pixel. Thus, height of the curve was about 200 pixels on the y-axis for all the presentation modes. The visual aid was aligned near the contact location but not exactly on top of the manipulator.

Virtual mesh models (782 triangles) representing elastic cubes were prepared with two stiffness parameters: 1.0MPa and 0.1MPa Young modulus. Poisson's ratio was set to 0.4. The 1.0MPa model had the same parameters as the aorta model that was used during the expert's recording. The stiffness difference was to a) make the user to perform larger motions and b) give clearer visual cues about the cube's deformation, which could produce different results.

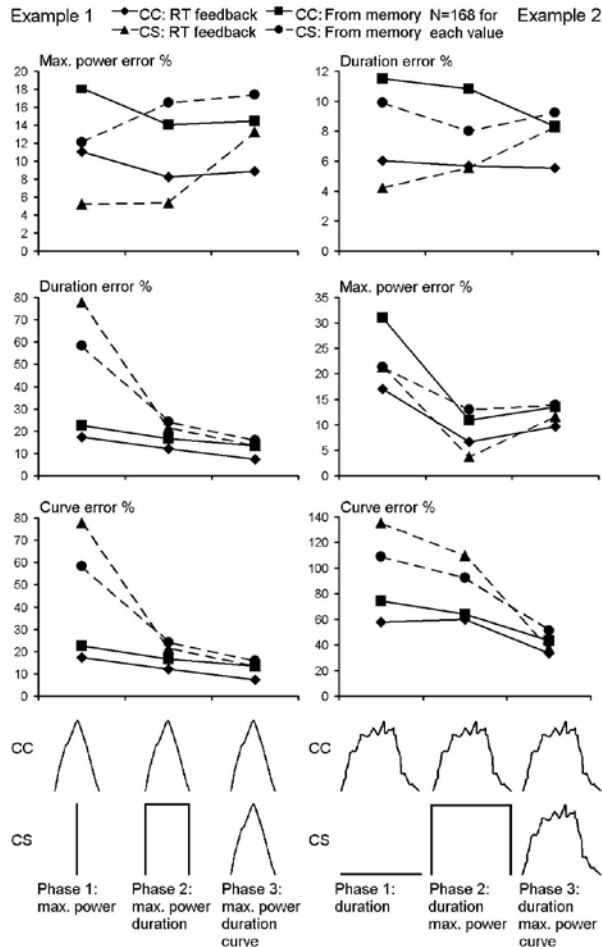
Table 1 summarizes the conditions and the order of tasks. 6 subjects were divided into Group A and B. Group A started with Session 1 and continued to Session 2 the next day. Group B started with Session 2. Each task was performed in trial pairs: tracking the example with RT feedback and repetition from memory with knowledge of result presented for 2 seconds after the trial as two overlapping bars or curves. There were 7 trial pairs in each training phase, resulting in 42 trials per task, 336 per subject and 2016 in total. In each phase, the subjects were advised to focus on the evaluated component(s) of skill. With CC the subjects were told to focus on each component at the time, even though the curve was fully visible. CS visualized only the evaluated components.

**Table 1. CONDITIONS of the experiment.**

Training mode	Example task: name, <i>max.p</i> (N), <i>dur</i> (ms)	Stiffness (MPa)	Order of training phases 1, 2 and 3
Session 1			
CC	E1, 1.25, 1130	1.0	<i>max.p</i> , <i>max.p+dur</i> , <i>cur</i>
CC	E2, 1.18, 1930	1.0	<i>dur</i> , <i>max.p+dur</i> , <i>cur</i>
CC	E1, 1.25, 1130	0.1	<i>max.p</i> , <i>max.p+dur</i> , <i>cur</i>
CC	E2, 1.18, 1930	0.1	<i>dur</i> , <i>max.p+dur</i> , <i>cur</i>
Session 2			
CS	E1, 1.25, 1130	1.0	<i>max.p</i> , <i>max.p+dur</i> , <i>cur</i>
CS	E2, 1.18, 1930	1.0	<i>dur</i> , <i>max.p+dur</i> , <i>cur</i>
CS	E1, 1.25, 1130	0.1	<i>max.p</i> , <i>max.p+dur</i> , <i>cur</i>
CS	E2, 1.18, 1930	0.1	<i>dur</i> , <i>max.p+dur</i> , <i>cur</i>

## 5. Results and discussion

Mean averages from 168 recorded trials per training mode in each phase were compared in order to see the training effect differences between CS and CC. ANOVA was performed with  $p < 0.05$  to each finding discussed below. Fig. 5 summarizes the results.



**Fig. 5. RESULTS:** mean averages of error % for each component of skill in the force exertion tasks grouped by the training phases 1-3. At the bottom: visual representations of the examples. Trials were performed in 7 trial pairs: with RT feedback, then from memory.

### 5.1. First example task: E1

E1 in Phase 1, where CS presented only *max.p*, CS provided for clearly more accurate tracking of the example (mean average of error 5.22%, standard error 0.99) than CC (11.05%, 0.99). This resulted also in better performance from memory: 12.14%, 1.77 (CS) against 18.11%, 1.77 (CC). With CC, when the subjects saw the curves all the time but were told to pay attention only to *max.p*, they tended to follow the example curve anyway (duration error 17.28%, 4.69 with RT feedback and 22.61%, 3.76 from memory).

*Dur* and *cur* were not presented in CS, which helped the subjects to focus on *max.p*, as expected. Errors were smaller on the softer cube, which was suspected to result from longer motions and better hand control.

In Phase 2, where *dur* was added to the evaluation criteria, the training modes provided similar results in *max.p*: 8.25%, 0.79 (CC) and 5.34%, 0.78 (CS) with RT feedback, and 14.04%, 1.32 (CC) and 16.55%, 1.29 (CS) from memory. In terms of *dur*, CS provided worse results (21.58%, 1.98 with RT feedback and 24.26%, 1.97 from memory) than CC (12.05%, 2.0 and 16.6%, 2.02). This was due to the fact that it was the first time that duration was displayed in CS. Results with CS in Phase 2 correspond to the results with CC in Phase 1.

In Phase 3, error percentages provided by CC were smaller than with CS when RT feedback was present, but statistically the same when performed from memory. *Max.p* errors: 8.88%, 1.13 (CC) and 13.11%, 1.11 (CS) with RT feedback, and 14.49%, 1.73 (CC) and 17.43%, 1.69 (CS) from memory. *Dur* errors: 7.32%, 1.19 (CC), 13.52%, 1.18 (CS) with RT feedback, and 13.53%, 1.62 (CC) and 15.86%, 1.58 (CS) from memory. *Cur* errors: 28.12%, 1.67 (CC) and 34.51%, 1.65 (CS) with RT feedback, and 42.06%, 2.47 (CC) and 48.09%, 2.42 (CS) from memory.

### 5.2. Second example task: E2

E2 demonstrated similar results as E1. In Phase 1, *dur* errors were statistically significant (6.02%, 0.55 with CC and 4.22%, 0.55 with CS) with RT feedback but about the same when performed from memory (11.5%, 1.3 with CC and 9.93%, 1.3 with CS).

In Phase 2, *dur* errors were the same with RT feedback (5.7%, 0.57 with CC and 5.56%, 0.57 with CS), but statistically different when performed from memory (10.83%, 0.79 with CC and 8.0%, 0.79 with CS). In *max.p*, instead, the difference was found from the performances with RT feedback (6.69%, 0.55 with CC and 3.75%, 0.55 with CS), but not from memory (10.9%, 1.19 with CC and 12.97%, 1.19 with CS).

In Phase 3, only one significant difference was found: CC provided for *dur* error 5.55%, 0.67 whereas CS only 8.31%, 0.67 with RT feedback. However, performance from memory was about the same (8.33%, 0.89 with CC and 9.24%, 0.89 with CS). Errors in *max.p* (RT: 9.69%, 0.82 with CC and 11.59%, 0.82 with CS; From memory: 13.5%, 1.15 with CC and 13.94%, 1.15 with CS) and *cur* (RT feedback: 33.62%, 1.53 with CC and 37.38%, 1.53 with CS; From

Mikko J. Rissanen, Yoshihiro Kuroda, Megumi Nakao, Naoto Kume, Tomohiro Kuroda and Hiroyuki Yoshihara. Toward Visualization of Skill in VR: Adaptive Real-Time Guidance for Learning Force Exertion through the "Shaping" Strategy, In: Proceedings of Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WorldHaptics), Tsukuba, Japan, March 22-24, 2007, pp. 324-329. (WHC'07; 324-329) IEEE Computer Society. DOI: [10.1109/WHC.2007.117](https://doi.org/10.1109/WHC.2007.117)

memory: 43.39%, 2.85 with CC and 51.19%, 2.85 with CS) were not significant. *Cur* errors in E2 were clearly higher than in E1 due to more complex shape of the curve, but still smaller in Phase 3 than in 1 or 2.

### 5.3. Conclusions

When comparing the performances from memory after the trial, training with CS demonstrated advantages in comparison to CC: since CC made the subjects follow the curve during training with RT feedback despite the verbal instructions, the subjects could not perform as well as with CS when evaluating the performance in the first two phases. Restricted training with CS did not hinder the performance in the last phase which was evaluated by all the criteria. CS reached CC's results despite the fact that with CC the curves were shown to and followed by the subjects in all phases whereas with CS the full curves were shown only in the last phase. With CS, the most essential components of skill were mastered in Phases 1 and 2.

### 6. Summary

We presented the design for skill visualization for VR training simulators and demonstrated its benefits in a simple haptic sensation based task. The technique was proven to have benefits in "shaping" style training of delicate force exertion so that the user's attention can be drawn to specific aspects of skill without hindering the overall outcome of the training later on.

Here, the example data consisted only of individual excerpts. Our future goal is to include variation of several experts' examples recorded into a database, which would give information about true limits of "good" performance to trainees. More complex skills have to be examined using various learning strategies.

### 7. Acknowledgements

This research is funded by Grant-in-Aid for Scientific Research (S) (16100001), Young Scientists (A) (18680043) and Exploratory Research (18659148) from The Ministry of Education, Culture, Sports, Science and Technology, Japan, and Nakajima Fund.

### 8. References

[1] D. Shaffer, D. Meglan, M. Ferrell, and S. Dawson, "Virtual Rounds: Simulation-based Education in Procedural

Medicine," in SPIE Vol. 3712: Battlefield Biomedical Technologies, 1999.

[2] E.J. Bass, "Architecture for an Intelligent Instructor Pilot Decision Support System," in IEEE Int. Conference on Systems, Man, and Cybernetics, San Diego, CA, 1998, pp. 891-896.

[3] A.G. Gallagher, E.M. Ritter, H. Champion, G. Higgins, M.P. Fried, G. Moses, C.D. Smith, and R.M. Satava, "Virtual Reality Simulation for the Operating Room: Proficiency-Based Training as a Paradigm Shift in Surgical Skills Training," *Annals of Surgery*, Lippincott Williams & Wilkins, Vol. 241, No. 2, Philadelphia, PA, 2005, pp. 364-372.

[4] U. Yang, and G.J. Kim, "Implementation and Evaluation of "Just Follow Me": An immersive, VR-based, Motion-Training System," *Presence: Teleoperators and Virtual Environments*, MIT Press, Vol. 11, No. 3, Cambridge, MA, 2002, pp. 304-323.

[5] Y. Yokokohji, R.L. Hollis, T. Kanade, K. Henmi, and T. Yoshikawa, "Toward Machine Mediated Training of Motor Skills -Skill Transfer from Human to Human via Virtual Environment," in IEEE Int. Workshop on Robot and Human Communication, Tsukuba, Japan, 1996, pp. 32-37.

[6] D. Feygin, M. Keehner, and F. Tendick, "Haptic Guidance: Experimental Evaluation of a Haptic Training Method for a Perceptual Motor Skill", in 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Orlando, FL, 2002, pp. 40-47.

[7] S. Saga, K. Vlack, H. Kajimoto, S. Tachi, "Haptic Video," in SIGGRAPH, 2005, DVD-ROM.

[8] X.D. Pang, H.Z. Tan, and N.I. Durlach, "Manual Discrimination of Force Using Active Finger Motion," *Perception and Psychophysics*, Psychonomic Society Publications, Vol. 49, Austin, TX, 1991, pp. 531-540.

[9] H.Z. Tan, X.D. Pang, and N.I. Durlach, "Manual Resolution of Length, Force, and Compliance," in Int. Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, ASME Dynamic Systems and Control Division, Vol. 42, Anaheim, CA, 1992, pp. 13-18.

[10] M.A. Srinivasan, and J.S. Chen, "Human Performance in Controlling Normal Forces of Contact with Rigid Objects," *Advances in Robotics, Mechatronics, and Haptic Interfaces*, ASME, Vol. 49, 1993, pp. 119-125.

[11] J.S. Chen, and M.A. Srinivasan, "Human Haptic Interaction with Soft Objects: Discriminability, Force Control, and Contact Visualization," *Touch Lab Report 7*,

Mikko J. Rissanen, Yoshihiro Kuroda, Megumi Nakao, Naoto Kume, Tomohiro Kuroda and Hiroyuki Yoshihara. Toward Visualization of Skill in VR: Adaptive Real-Time Guidance for Learning Force Exertion through the "Shaping" Strategy, In: Proceedings of Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WorldHaptics), Tsukuba, Japan, March 22-24, 2007, pp. 324-329. (WHC'07; 324-329) IEEE Computer Society. [DOI: 10.1109/WHC.2007.117](https://doi.org/10.1109/WHC.2007.117)

MIT Press, RLE Technical report No. 619, Cambridge, MA, 1998.

[12] Y. Kuroda, M. Nakao, T. Kuroda, H. Oyama, and H. Yoshihara, "MVL: Medical VR Simulation Library," in Medicine Meets Virtual Reality (MMVR13), IOS Press, Long Beach, CA, 2005, pp. 273-276.