A Novel Interface for Simulator Training: Describing and Presenting Manipulation Skill Through VR Annotations

Mikko Rissanen[1,*], Yoshihiro Kuroda[2,**], Tomohiro Kuroda[3,*], and Hiroyuki Yoshihara[3,*]

[1] Graduate School of Informatics, Kyoto University
[2] Graduate School of Engineering Science, Osaka University
[3] Division of Medical Informatics, Kyoto University Hospital
[*] Division of Medical Informatics, Kyoto University Hospital Shogoin Kawahara-cho 54, Sakyo-ku, 606-8507 Kyoto, Japan
{Mikko, TKuroda, Lob}@kuhp.kyoto-u.ac.jp
[**] Dept. of Medical Science and Bioengineering, Osaka University, Machikaneyama 1-3, Toyonaka-Shi, 560-8531 Osaka, Japan
YKuroda@bpe.es.osaka-u.ac.jp

Abstract. Virtual reality (VR) based skill training simulators are highly interactive computer systems that allow for training of e.g. manipulation skills. In the most complex domains, such as in surgery, teaching correct manipulation requires expert's judgement and instructions. Video annotation has enabled presentation of explicit principles of manipulation that are highlighted on a recorded demonstration of an expert. In VR, similar annotation methods have been applied even though interactivity of VR allows for novel kinds of representations of annotated data. This paper presents an annotation model and an interface that uses interactivity of VR as an advantage on reading the annotations, which has potential to convey the principles of manipulation through multiple modalities to novices. The design of the model is based on behavioural parameters of the simulation – features which cannot be accessed in the real world. This approach has potential to enhance efficiency of VR simulator based skill training.

Keywords: Virtual Reality, Open Skill, Annotation, Multimodal Feedback.

1 Introduction

To learn manipulation skills, such as sculpting or closing incisions by suturing using surgical instruments, requires much practice. Even though significant learning theories have been proposed, it has been acknowledged that the current theories are incapable of explaining certain aspects of skill. It has always been difficult to distinguish mere motor skill and the cognitive part of human actions [1]. For example attention and proper focusing are difficult to be measured during a performance.

R. Shumaker (Ed.): Virtual Reality, HCII 2007, LNCS 4563, pp. 523–532, 2007.

The most complex manipulation skills are open, i.e. the environment varies during each case in which the skill has to be applied. Domains in which open skills are required are for example team sports, martial arts and surgery.

The traditional approach to learn movements and motor skills is based on observation and mimicry of expert's examples with the aid of the expert's verbal instruction. According to learning theories, in order to gain a skill, one must first learn rules about what to do in which situation and then manage to absorb the rules into almost automatic behaviour through repetition. The use of verbal instruction is mainly dictated by the nature of the real world. In practise, the rules are learned case by case and they become greater in amount and more accurate in details during experience. Further in this text, these rules are called *relationships within manipulation*. True expertise is described as problem-solving without any deliberate attention to the rules learned in the beginning. The relationships are observed almost automatically.

Simulator training offers the possibility to gain experience without consequences in case of failures that are an essential part of learning. Simulators have been extended from aviation to fields that require a safe environment for practise of manipulation skills. Surgical simulators have become a popular subject of research and even industrial development. The focus of research has been strongest on technology. Development of skill assessment methods has begun only recently. Relatively little research exists on teaching methods of manipulation skills. Except for a few attempts on haptic guidance [2,3] and application of Virtual Fixtures [4], the teaching methods in VR usually replicate the ones in the real world even though in a fully digital world there are very few physical limitations. Record and Replay approach (e.g. [2,3,5]), in which expert's actions are recorded and perceptualized to the trainee by using some experimental method, has been found possible. However, as development of Record and Replay training systems usually aim to fully automatic human-machine-human transfer of skill, it may not be the most suitable approach for complex open skills that are taught in the real world under experts' instruction. Research so far has not managed to address the need of expert's insight in the Record and Replay approach.

Using the traditional teaching media, for example text-books and videos, *annotation* has been found useful as the means to highlight the essential from the other content. Through video annotation, general principles of manipulation can be represented: arrows and other illustrations overlaid on a video clip demonstrating an example performance are so called external annotations that intend to convey a message to the viewer. For the learner, presentation of the principles supports the cognitive part of learning. Similar annotations have been implemented into VR training simulators, but the main point of VR – interactivity – has been forgotten in the design. Since interaction modelling in training simulation allows realistic interaction, interactivity should be included into annotations as a basic characteristic.

This paper presents an annotation model that allows expert's insight to be encapsulated into Record and Replay VR and to be represented in intuitive manner to the novices. Two main advantages are introduced: 1) Relationships within manipulation can be explicitly perceptualized to the user in terms of the essential information determined by an expert. 2) The relationships can be presented to several sensory channels making the feedback to the user multimodal. The design of the model is based on behavioural parameters of the simulation that are recorded as time-series – features which cannot be accessed in the real world. This approach can enhance efficiency of VR simulator based skill training by reducing the expert's one-on-one instruction and by providing only the relevant feedback to the user.

2 Related Work

Two aspects of previous work on VR have to be examined: training simulators using advanced teaching methods and VR annotation systems.

2.1 Training Simulators

Conventional training simulators only replicate reality so that realistic interaction modelled in the system allows for practise in a safe environment. Record and Replay approach [2,3,5] has been demonstrated to be possible and even beneficial for some parts in learning of simple skills [2,3]. However, teaching of the most complex open skills would require insight of the experts that is not modelled in the system. In other words, even if Record and Replay was implemented in simulators allowing for training of open skills, the expert would still be needed.

One of the restrictions of the real world is that the intentions, priorities of sub-goals and the significant features of the manipulation can only be described verbally. VR allows access to the whole digital environment, which gives the system's developers freedom to use teaching methods different to the real world. So far, only haptic guidance gives alternative means for representation of the example manipulation than in the real world. Approaches based on visualization, for example Just Follow Me [6], mainly replicate the real world methods with the only difference being that the replayed expert's performance is overlaid onto the user.

This review shows that VR is too often viewed as a mere training environment. Teaching features of VR have not been studied rigorously to the date at the level of complex, open manipulation skills. Yokokohji *et al.* [5] presented their research questions "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?" in 1996. These are questions that have not been answered during this decade. Most skill-transfer systems attempt to transfer skills through a single holistic method. Yet, that approach has not been found to have clear training advantage. Since automatic skill-transfer seems a goal of the far future, current VR training simulation should be designed as a teaching tool. The model presented in this paper is oriented more on the expertise itself than on engineering solutions by harnessing the power of annotation of manipulation level behaviour into the training simulation using multiple modalities.

2.2 VR Annotation Systems

As envisioned by Shaffer *et al.* [7], annotations could be a significant part of a future's training environment that is built around a simulation engine. In the following, the most recent attempts of VR annotation are compared.

Welch *et al.* [8] introduced a VR recording system for surgical training as knowledge mediator and included annotation as description of the surgical procedure's stages and key events. Aubry *et al.* [9] presented VR annotation as an improvement for product design in collaborative virtual environments. Their model is capable of coupling the knowledge level ontologies with the annotations. As the model is intended to support design processes, the manipulation level that would be needed in training simulators has not been addressed.

These text-based systems mainly implement the traditional video annotation methods into 3D VR. As the aim has been on transfer of knowledge, transfer of manipulation skills has not been addressed in the previous VR annotation systems. Annotation systems developed so far have not attempted to include interactivity of VR as a basic feature of annotations, which provides a new research opportunity.

3 VR Annotation for Training Open Manipulation Skills

Annotation required to represent expert's insight on very complex principles and case-dependant relationships of manipulation has to include interactivity of virtual training environments into the fundamental design. Fig. 1 shows the general idea of how interactive training simulation is annotated. The basic features of annotation are discussed in detail below: target and media of annotation, representation of annotated data and anchoring.

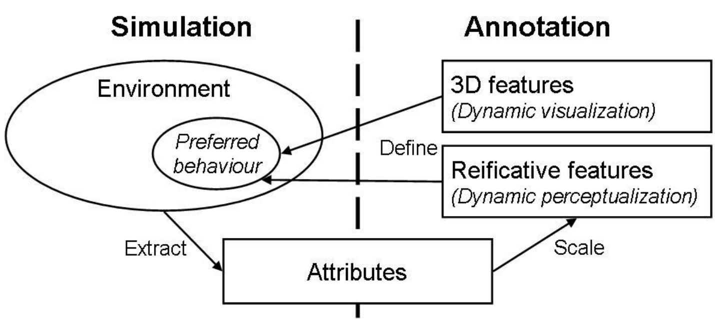

Fig. 1. Annotation of interactive simulation. Preferred behaviour of the simulated environment is annotated as *3D* and *reificative features*. Reificative features facilitate understanding of complex concepts by highlighting the essential from the whole data. They can be perceptualized dynamically using any modality, whereas 3D features can only be visualized. Simulation's attributes act as scaling factors on the reificative features, for example that length of an incision made by a surgeon depends on the size of the limb. 3D features are naturally in right scale without any modifications.

Using the Record and Replay approach, highlighting is done on the expert's recorded performance. *3D features*, such as lines indicating correct directions or highlighted surfaces of a target, are visualized dynamically on the expert's pre-recorded performance. *Reificative features* are used to explain information that is otherwise hard to understand. The concept of reification is explained below by describing the target, representation and anchoring of annotations as well as the process of making them.

3.1 Target of Annotation

Fig. 2 illustrates the data that can be annotated. A recording API allows a simulator's developer to include any behavioural parameter (BP) to be recorded as time-series. Thus, dynamics of the interactive simulation can be accessed.

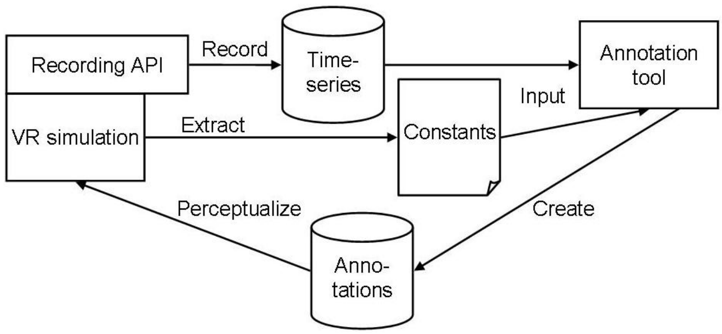

Fig. 2. Annotation of interactive VR simulation. Time-series can be recorded from any variable in the simulation. Constants are extracted attributes of the simulation that can be included into the annotations. Annotations are represented to the user through multi-modal perceptualization.

Time-series are the most significant data form that can be annotated. In practice, time-series are recorded behavioural parameters (RBPs). Time-series data can consist of any number of dimensions. Typically, the data covers a timestamped 3D coordinates (x, y and z). 2-dimensional time-series would consist of a timestamp and a single value changing over time. Thus, dynamics of the simulation are accessible.

Constants are pieces of data that do not change. An example of a constant is for example length of a virtual tool or general stiffness of an elastic model. The annotator binds the constants to BPs. Constants act as scaling factors to time-series data so that, for example, force that the user should exert depends on size or stiffness of the target. The constants are derived from the attributes of the simulation which are included into 3D and interaction models in most of the advanced simulation software.

Using these two simple data formats, their combination allows complex relationships to be made between several variables of the virtual environment: time-series scaled by constants or relationships between several time-series.

3.2 Annotation Process

Visualization of the recorded behavioural parameters is the basis for support of the annotation process. Conventional video annotation software allows sound waves to be represented graphically so that the annotator can select segments on the timeline accurately. For the proposed model, similar approach was chosen with the distinction that there can be multiple layers of data that can be annotated separately. Multiple

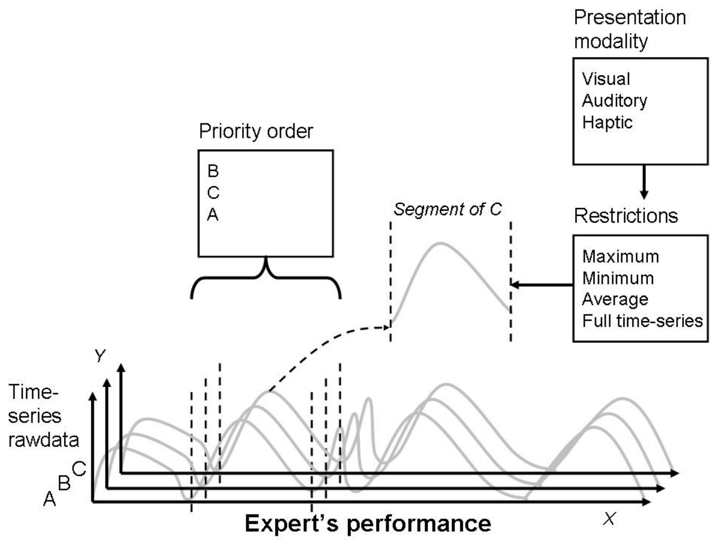

Fig. 3. Annotation process. Expert's performance is first visualized, then segmented, and finally content of annotations determined: priorities between time-series; segment-specific definitions of restrictions and presentation modality.

layers allow different segmentation of the time-series of RBP. Characteristics that can be determined for each segment of time-series are presented in Fig. 3.

Priorities between the factors in manipulation are described as a percentage of what the skill is composed of. For example, the annotator could determine that the most important factor for success is correct amount of power, whereas accuracy of contact location is only secondary factor.

Restrictions are set to every BP: maximum, minimum, average or full time-series can be selected. Restrictions are applied to single segments of example performances which are then aggregated to a so called general description of skill when statistical variance of data in a number of segments is calculated. This allows implementation of rules that, for example, would describe a skilful manipulation as amount of power that is not to be exceeded or acceleration profile during a motion. Restrictions are illustrated later in Fig. 3.

Modality can be chosen between visual, auditory and haptic sensory channels.

3.3 Representation of Annotations

The purpose of this annotation model is to perceptualize principles of manipulation to the user in an intuitive manner. Annotations can be represented on any modality by following the Virtual Fixtures concept [4]. Two time-series of RBPs, the pre-recorded examples and the user's live recording, provide an ideal feedback on performance for simple as well as for complex manipulations. Fig. 4 illustrates how different representations can be generated from the rawdata.

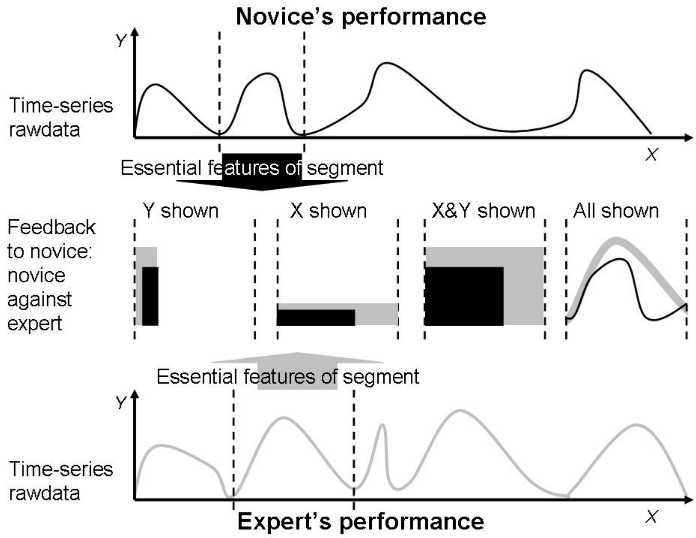

Fig. 4. Simulation's preferred behaviour, caused by expert's performance, is represented on the essential features only as Virtual Fixtures. The novice's performance is shown against the one or many expert's performances according to restrictions annotated. Using several modalities, many layers of time-series can be perceptualized at once. Thus, the annotator is given a choice of priorities of relationships that should be learned from the simulation.

Knowledge of Result (KR) and Knowledge of Performance (or Process, KP) are terms used in motor learning research to categorize feedback to the trainee (see e.g. [1]). *Guidance hypothesis* suggests that simple tasks require KR only after several trials, whereas the most complex skills demand constant KP [10]. KR can be implemented using a simple triggering system coupled with the recording API, so that at the end of recording a segment, KR is displayed.

Mapping of modalities is based on guidelines on multimodal representation of information [11]. Visualization is the most suitable method for representing complex information. In addition to complexity, visualization allows the user to plan the actions when the preferred future behaviour is shown. Auditory feedback is mapped to temporal data and haptic feedback (mostly vibro-tactile) is the best modality to represent simple one-dimensional restrictions as warnings.

3.4 Anchoring

Time-series of RBPs regulate also the anchoring. An example of dynamic anchoring is a contact point (vertex of a mesh) that the user touches virtually through a haptic device. ID of the vertex changes over time during manipulation, and thus the time-series of the vertex' position provides a dynamic anchor for the annotation that appears at the contact point when there is any contact. Appearance of any annotation can be controlled in a similar way by setting the anchor to some time-series.

4 Discussion

In brief, the proposed annotation extends previous VR annotation to truly interactive simulation, displays "the essential information" through a kind of Virtual Fixtures and has the potential to bridge the motor performance and the cognitive part of training by presenting the essential relationships within manipulation explicitly.

4.1 Novelty

Basically, representation of annotations in the proposed annotation model is based on Virtual Fixtures [4]. These sensory overlays were originally intended as a support for telemanipulation tasks, but also applied in training. The most significant advantage of the proposed model is that it conveys only the essential information from the expert's point of view to novices using recorded manipulations as the annotated media. In comparison to previous research, Virtual Fixtures in the proposed model only restrict manipulation on the parts that the expert sees important. For simple closed skills, holistic view of Virtual Fixtures may be enough as a training aid. Complex open skills require more advanced feedback, which is achieved by applying the proposed model to a working interface implemented onto a training simulator.

VR annotation models presented so far has mostly covered somewhat static data. The proposed model is capable of representing highly dynamic data extracted from an interactive simulation. Interactivity, which is the most essential feature of skill training simulators, is now a fundamental part of the design: the simulator shows how the environment should behave and overlays the user's performance against the desired one.

Hutchins *et al.* [12] studied representations of objects in virtual training environments. In their framework, multiplicity and augmentation of representations were some of the main aspects: an object does not need to have only one representation, such as mimetic behaviour of the real world often preferred in training simulation, since virtuality allows several representations. Our work falls into this category: unlike in the real world, the environment's behaviour is perceptualized on multiple modalities and the recorded expert's performance is displayed to the user only for the essential parts. The essential part is annotated into the system by the expert.

4.2 Advantages

The proposed model intends to support especially complex, open skills. Explicit representation of relationships, that in the real world are learned case by case, have the potential of enhancing the link between mere motor performance and the cognitive part of training.

As a note to possible technical issues concerning implementation of the model on a real training simulator, the authors wish to highlight that specifications of the simulator itself do not restrict the annotation model at all. Only the data that can be recorded and the attributes of the simulation act as the annotation media. In practice, there will be only minor technical problems in the implementation since data processing and storage technologies are already very advanced.

4.3 Future Directions

As the model supports teaching (expert's involvement) and learning (novice's involvement), a working interface based on the model is the main topic of the future research. The extent of complexity is to be addressed in future studies on transferring a skill through annotated VR in comparison to non-annotated VR.

Automation for selection of optimal feedback modalities is another topic of the future. The priorities between BP's and their restrictions as well as user's preferences could play a significant role in the design, since there are individual differences on the optimal feedback channels.

5 Conclusion

Previous studies have not managed to address the need of expert's insight in the Record and Replay approach. The proposed annotation model is capable of representing the essential information from recorded experts' examples when replaying the annotations to the user. Replay of annotations is interactive so that the essential parts of both the example and the user's performance are displayed at once. As an enhancement to previous VR annotation designs, the proposed model includes interactivity as the key feature of reading annotations.

The proposed model shows potential to bridge mere motor performance and the cognitive part of skill training, which is expected to promote learning of complex, open manipulation skills. Attention is explicitly drawn to relationships within the manipulation that are difficult to learn. In the real world, the relationships are learned case by case. Using the proposed model, the relationships are represented though several modalities whereas in the real world verbal description is often the only feasible feedback option. Thus, a novel interface for simulator based training of manipulation skills was achieved on a theoretical level. The interface would not depend on specifications of the simulator itself, but only on the data that can be recorded. This approach has potential to enhance efficiency of VR simulator based skill training by reducing the expert's one-on-one instruction and by providing optimal feedback to the simulator's user.

Future work will cover automated representation of annotations on the most suitable modalities. Then, the learning effect can be compared between the annotated and non-annotated training approaches.

Acknowledgements. The authors would like to thank Dr. Naoto Kume, Dr. Megumi Nakao, Mr. Masayuki Kawasaki, Dr. Keisuke Nagase and Mr. Stéphane Aubry for their support. This study was 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 by Nakajima Fund.

References

1. Sherwood, D.E., Lee, T.D.: Schema Theory: Critical Review and Implications for the Role of Cognition in a New Theory of Motor Learning. Research Quarterly for Exercise and Sport, American Alliance for Health, Physical Education, Recreation, and Dance, vol. 74(4), pp. 376–382 (2003)
2. Teo, C.L., Burdet, E., Lim, H.P.: A Robotic Teacher of Chinese Handwriting. In: Proc. of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems 2002, pp. 335–341 (2002)
3. Saga, S., Vlack, K., Kajimoto, H., Tachi, S.: Haptic Video. In: Proc. of SIGGRAPH 2005, DVD-ROM (2005)
4. Rosenberg, L.B.: Virtual fixtures: Perceptual Tools for Telerobotic Manipulation. In: Proc. of the IEEE Annual Int. Symposium on Virtual Reality 1993, pp. 76–82 (1993)
5. Yokokohji, Y., Hollis, R.L., Kanade, T., Henmi, K., Yoshikawa, T.: Toward Machine Mediated Training of Motor Skills -Skill Transfer from Human to Human via Virtual Environment-. In: Proc. of the 5th IEEE International Workshop on Robot and Human Communication 1996, pp. 32–37 (1996)
6. Yang, U., Kim, G.J.: Implementation and Evaluation of “Just Follow Me” : An immersive, VR-based, Motion-Training System. In: Presence: Teleoperators and Virtual Environments, vol. 11(3), pp. 304–323. MIT Press, Cambridge (2002)
7. Shaffer, D., Meglan, D., Ferrell, M., Dawson, S.: Virtual Rounds: Simulation-based Education in Procedural Medicine. In: SPIE vol. 3712: Battlefield Biomedical Technologies (1999)
8. Welch, G., State, A., Ilie, A., Low, K.-L., Lastra, A., Cairns, B., Towles, H., Fuchs, H., Yang, R., Becker, S., Russo, D., Funaro, J., van Dam, A.: Immersive Electronic Books for Surgical Training. In: IEEE Multimedia, vol. 12(3), pp. 22–35. IEEE Computer Society Press, Los Alamitos (2005)
9. Aubry, S., Lenne, D., Thouvenin, I., Guénand, A.: VR Annotations for Collaborative Design. In: Proc. of Human-Computer Interaction International, CD-ROM (2005)
10. Schmidt, R.A., Young, D.E., Swinnen, S., Shapiro, D.C.: Summary Knowledge of Results for Skill Acquisition: Support for the Guidance Hypothesis. Journal of Experimental Psychology: Learning, Memory, and Cognition 15(2), 352–359 (1989)
11. Nesbitt, K.: Designing Multi-sensory Displays for Abstract Data. PhD Thesis, School of Information Technologies, University of Sydney (2003)
12. Hutchins, M., Adcock, M., Stevenson, D., Gunn, C., Krumpholz, A.: The Design of Perceptual Representations for Practical Networked Multimodal Virtual Training Environments. In: Proc. of Human-Computer Interaction International, CD-ROM (2005)