Modeling and Evaluating of Human 3d+t Activities in Virtual Environment

Djadja Jean Delest Djadja, Ludovic Hamon, Sébastien George

▶ **To cite this version:**

HAL Id: hal-02298496

https://hal-univ-lemans.archives-ouvertes.fr/hal-02298496

Submitted on 26 Sep 2019

Modeling and Evaluating of Human 3d+t Activities in Virtual Environment

Djadja Jean Delest Djadja$^{(\boxtimes)}$, Ludovic Hamon$^{(\boxtimes)}$, and Sébastien George$^{(\boxtimes)}$

LIUM - EA 4023, Le Mans Université, 72085 Le Mans, Cedex 9, France
{djddja,ludovic.hamon,sebastien.george}@univ-lemans.fr

Abstract. This paper studies the problem of evaluation of human 3d+t activities in Virtual Environments (VE) for Learning (VEL). Current evaluation methods focus mostly on: (i) the automatic identification of an ordered sequence of actions and/or (ii), an empirical analysis made by experts through the VE. In many cases, the learner's activity can be represented by some specific time series made of geometrical data of 3D artefacts. For the extraction and analysis of such Motions Of Interest (MOI), one can manually segment them among the traces, and/or use automatic approaches requiring a database of annotated examples. Both cases usually require too many resources to design such environments. Consequently, this work presents a method allowing teachers to quickly build, compare and evaluate a 3d+t learning activity in VE. This method is based on a semi-automatic approach combining the Dynamic Time Warping algorithm, with 3D reference shapes and few expert's demonstrations of the task to learn.

Keywords: Virtual reality · Activity evaluation · Motion · Human learning

1 Human 3d+t Activity in Virtual Environment

Virtual Environments (VE) for learning (VEL) offer advanced interaction possibilities such as human movements and gestures for 3D objects selection, manipulation and navigation [2,6]. Regarding the human activity, most of the time, it is evaluated after the activity execution in VEL [4], according to an empirical evaluation and/or the use of task-dependant metrics characterizing the activity at the action level (done or not), and/or considering the search of an ordered sequence of actions. However, in many cases, it is crucial to characterize and evaluate each human action in terms of interactions and 3d+t features to provide an appropriate and real-time feedback in VEL [5]. Although there are a lot of 3d+t metrics [3] to characterize a motion (velocity, jerk, curvature, etc.). They can have a pedagogical value only if they are related to a domain and a task-dependent analysis of the learning situation with some experts [7]. VEL based on this approach, require a heavy re-engineering process if the task or the application domain change. Usually, the re-engineering process cannot be made

by the expert or the teacher [1]. Furthermore, the identification of the Motions Of Interest (MOI) that would be analyzed, is a difficult task if we consider a VE where the user can start, pause and resume the learning scenario at anytime. Therefore, given a predefined task to learn, our method consists in automatically capturing, segmenting and evaluating targeted MOIs during the learning activity in VE. These MOIs represents the evolution in time and space (3d+t) of some monitored virtual artefacts resulting from the learner's activity in VE. Our approach relies on the comparison between the expert performing the task and the learner thanks to some reference shapes considered as checkpoints, the Dynamic Time Warping (DTW) algorithm and some kinematic metrics [5]. The details of our method and the implemented system are presented in the next section.

2 Activity Modeling and Evaluation: System Functioning and Use Cases

Suppose a toy problem like a navigation task, where the learner has to walk according to two specific paths for reeducation purposes (Fig. 1, left). To set up the system, the expert (*i.e.* the expert of the task to learn) has to firstly select the virtual element to monitor (*e.g.* the learner's body). Next, the expert places three kinds of 3D reference shapes, acting as CheckPoints (CP) with which the artefact must collide with: a unique Starting and a unique Ending CP (SCP and ECP) for the task beginning and ending, and some optional Intermediate CPs (ICP). The ICPs represent sub-sequences of the task to learn. In case of several intermediate ICPs, the artefact must collide with them according to a sequential order. For the navigation example, boxes are used as CPs and the expert makes several demonstrations of a walking made of a curving path followed by a turning path. The SCP acts as an oriented local landmark from which, the positions and orientations of the artefact are computed and recorded until the ECP is reached. Finally, the teacher chooses one demonstration as a task to imitate.

Fig. 1. (Left) Navigation task (Bird's eye); (Right) State machine for the expert's demonstration and learner evaluation with one ICP

The learner tries to reproduce the task. The system will then record the learner's MOI from the SCP to the ECP. The shape of the learner's path will not necessarily match with the expert's demonstration. Therefore, the system compares the learner's MOI to the expert one by using DTW that gives a value

indicating the similarity of two signals in terms of shape regardless of the duration (the lower the value is, the closer the two signals are, [5]). If the value is acceptable (under a threshold), the learner can compare the performance to the expert's one thanks to a visual feedback, on a wall within the VE, that can display: the value of the DTW, the values of two kinematic metrics (for this example) the velocity and the jerk *i.e.* the rate of change of the movement acceleration. The lower the value is, the smoother the motion is [3] (Fig. 2, right). A replay of the learner's performance and/or the expert's demonstration is also available. Finally, the system functioning, with one ICP, is represented by the finite-state machine with the following transitions (Fig. 1, right):

– a: SCP collision, go to state S *i.e.* the artefact evolution is monitored and recorded, the previous recording is deleted if existing
– b: ICP collision, go to state I *i.e.* the ICP collides
– c: ICP collision, go to state F, *i.e.* the expert's demonstration or the learner task fails
– d (for the expert's demonstration): ECP collision, go to the ending state M, *i.e.* the recording stops, the metrics are computed and displayed within the VE.
– e (for the learner evaluation): ECP collision, go to state D, *i.e.* the recording stops, the learner's performance is compared to the expert's one thanks to DTW
– f: the DTW value is under or equal to the threshold, go to the ending state M
– g: the DTW value is above the threshold go to the ending state F

Two other simple toy problems related to (a) a throwing task of a ball and a manipulation task of a glass has been implemented (Fig. 2, demonstrations video can be found by following the link below[1]). The task (a) was built to address, in the future, an example of a throwing task. The monitored artefact is the ball that must be thrown in a bin according to a basketball launch (hand close to the head on start and elbow down, Fig. 2, left).
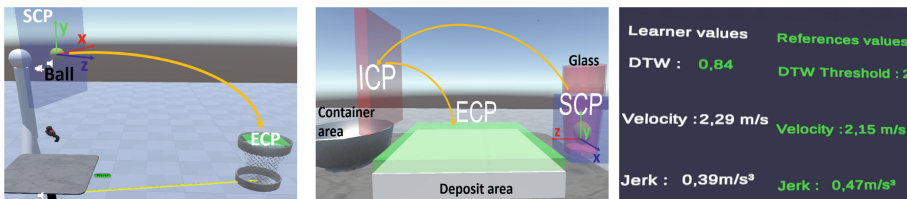

Fig. 2. Left: throwing task (Side view); Middle: manipulation task of glass (front view); Right: Metrics display within the VE

The SCP is placed close to the user's head on her/his right (if right-handed) and attached to her/him to allow the launch of the ball around the bin. The task (b) illustrates the manipulation of object such a beaker in chemistry. This use case consists in monitoring a glass with a ball inside. The user must take the glass and put the ball into another container. The user must finish the task by turning over the glass and put it on the deposit area (Fig. 2, middle). In both tasks, the motions of the object are compared to the demonstration, thanks to DTW.

3 Conclusion and Future Work

With the proposed method, we hope that experts and teachers will be able to easily build a learning situation in VE and evaluate the learner's 3d+t activities. Once the checkpoints are placed in the scene and the demonstration is performed, the system can automatically record and analyse every learner's MOI, colliding an ordered sequence of virtual checkpoints and fitting the expert's demonstration. The future works will focus on the evolution of this prototype with the implementation of an appropriate interaction paradigm to allow teachers to build their own learning situations in VE. The limits and advantages of our method will also be studied by conducting an experiment to evaluate the authoring aspects of the proposed solution. The dart game and some simple chemistry exercises will be simulated to allow experts to put in place learning situations. The goal is to study the operationalization capacity of our system regarding the proposed learning scenarios. The learner evaluation process based on DTW and some kinematic metrics must be improved on three main points: (a) an inclusion of temporal aspects as the DTW algorithm works regardless of the signal duration, (b) a contextual and application-domain-dependent model to choose a minimal set of appropriate metrics to display within the VE and (c) visual feedbacks with affordance properties.

References

1. Buche, C., Querrec, R., De Loor, P., Chevaillier, P.: MASCARET: a pedagogical multi-agent system for virtual environment for training. Int. J. Distance Educ. Technol. **2**, 41–61 (2004)
2. Emma-Ogbangwo, C., Cope, N., Behringer, R., Fabri, M.: Enhancing user immersion and virtual presence in interactive multiuser virtual environments through the development and integration of a gesture-centric natural user interface developed from existing virtual reality technologies. In: Stephanidis, C. (ed.) HCI 2014. CCIS, vol. 434, pp. 410–414. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07857-1_72
3. Larboulette, C., Gibet, S.: A review of computable expressive descriptors of human motion. In: 2nd International Workshop on Movement and Computing, pp. 21–28 (2015)

4. Lee, G.I., Lee, M.R.: Can a virtual reality surgical simulation training provide a self-driven and mentor-free skills learning? Investigation of the practical influence of the performance metrics from the virtual reality robotic surgery simulator on the skill learning and associated cognitive workloads. Surg. Endosc. **32**(1), 62–72 (2018)
5. Morel, M., Kulpa, R., Sorel, A., Achard, C., Dubuisson, S.: Automatic and generic evaluation of spatial and temporal errors in sport motions. In: 11th International Conference on Computer Vision Theory and Applications, pp. 542–551 (2016)
6. Penichet, V.M.R., Peñalver, A., Gallud, J.A.: New Trends in Interaction, Virtual Title Reality and Modeling. Human-Computer Interaction Series. Springer, London (2013). https://doi.org/10.1007/978-1-4471-5445-7
7. Toussaint, B.M., Luengo, V., Tonetti, J.: Towards using similarity measure for automatic detection of significant behaviors from continuous data. In: Proceedings of the 7th International Conference on Educational Data Mining, pp. 427–428 (2014)