SELF-PACED LEARNING WITH 3D HUMANOID VIRTUAL TUTORS

Zysk, Wolfgang, Dipl.-Psych., Dr.phil.
University of Duisburg, Dept. of mechatronics, Germany
Filkov, Rumen, Dipl-Inform.
FH-Vorarlberg, UCT, Austria

Abstract

For acquiring specific vocational skills, it is essential that trainees are able to observe their trainers closely while these demonstrate particular working movement processes in an exemplary fashion. It would further be desirable if each trainee was able to absorb these observations in accordance to his or her own, individual learning habits and capabilities. Yet circumstances in real-life classroom situations often make individual supervision of each trainee to a large extent impossible. Therefore computer-animated 3D humanoid characters, precisely emulating the correct, realistic behavior of trainers, with a computer program which can be operated by the trainee in a self-paced mode, could present a viable training aid. Our work has consequently been focused on the conceptualization and creation of tools to realize visible aspects of the training process for the purpose of self-paced learning through 3D humanoid characters. After a short introduction to the didactical concept, this paper presents the creation and the preceding development of such a self-paced learning tool.

Keywords: Selfpaced-learning, vicarious learning, 3D humanoid virtual tutors, ScanimFBX-AnimationPlayer

Introduction

Inspired by the trend-setting preliminary work of renowned researchers (Benjamin, 1988), the famous psychologist B.F. Skinner constructed and presented his mechanical Teaching Machines in the 1950s (Skinner, 1958). Notwithstanding broad criticism of his views on learning theory (e.g. Chomsky, 1959; Staddon, 2001), which is not central to the deliberations here, Skinner was the first to establish a thematic connection between the positive effect of the didactical concept of selfpaced learning and useful technical aids for application in this type of learning (http://www.youtube.com/watch?v=jTH3ob1IRFo). In accordance to Skinner, the concepts of self-paced or un-paced refer to a process of preparing and presenting information which allows the user to completely independently decide when, at which speed, in which order and intensity he or she wishes to absorb the given information, taking their individual levels of competency into consideration. The opposite to this self-determined learning is extraneously determined learning (paced or other-directed). In the latter, prepared and presented information can only be absorbed in the way dictated by others. In this way, one is forced to, for example, listen to an entire speech as dictated by the speaker to avoid missing some important information. An example of self-determined learning would be reading a book. Here the reader is able to decide for himself, when and where he or she starts, stops, continues reading, takes a break, etc.

Though it had been customary to primarily employ learning material in *written form*, a change accured when the the psychologist, Albert Bandura, showed that *vicarious learning*- meaning the observation and emulation of visible behavior of other people- offers a highly significant source for acquiring new skills (Bandura et.al., 1961). Henceforth the application of *visual material* for self-paced learning became an obvious choice. Enthusiastically proclaimed learning film productions of the early twentieth century proved to be less effective for the learning process as anticipated (Edison, 1913). On the one hand, these films were frequently only consumed passively and could thus reach limited learning effects. On the other hand, trainees were less motivated or unable to independently analyze the film material, or search for relevant information. Although films clearly offer the user the possibility of deciding where to start, stop, go forward and back, pause, use slow motion or fast forward- similar to reading a book- creating and presenting films for learning purposes remain extraneously determined to a large extent, Figure 1. This is because the temporal sequence, perspectives, dimensions and focus, meaning the focus on specific and relevant optical details, are predetermined by the film- to name a few examples. Should one succeed in eliminating these limitations, allowing trainees to make their own decisions on what they wish to see, how long, how often, from which perspectives, from which distance, in which sequence and speed at any given time, the visual potrayal of movement sequences in specific fields of work could provide them with extraordinarily useful training sources.

Fig.1 Film = other-directed acquisition of knowledge

To incorporate the Selfpaced concept and the deliberations of Banduras regarding *vicarious learning* into the creation of a user-friendly computer program for self-paced learning with the aid of visible movement sequences in specific fields of study, intensive preliminary work had to be done. This is outlined step-by-step below.

Videotaped Critical incidents

During our extensive field studies in a wide range of training areas (metal processing and nursing) both trainers and trainees were filmed from the frontal as well as the lateral perspectives while conducting activities typical for their lines of work, e.g. sawing, filing, repositioning patients. Then trainers and persons with longstanding experience in their specific professions, carefully selected clips from the filmed material. A large number of individual movement sequences in specific vocational fields, which are vital for determining whether the actions were executed optimally or less optimally, could be observed in numerous short clips selected, so-called *Critical Incidents* (Flanagan, 1954). Since an adequate tool for such complex video analyses, which is at the same time user-friendly and suitable for use in a self-paced mode by computer laypersons, was not as yet available on the market, we designed and produced a new tool for video analysis, the *AYP-MoviePlayer* (Zysk & Filkov, 2013).

Lifelike reproduction of activities with 3D humanoid characters

The critical incidents in movement sequences in the clips as described above, are to be reproduced with the goal of reanimation through 3D humanoid characters. The movements of the humanoid characters must be highly precise and natural, so as to reproduce the original vocational activities exactly. The demand for this high degree of realism posed the threat of falling into the *uncanny valley* (Mori, 1970). Therefore, a decision had to be made from the outset as to which animation procedure was best suited for the precise and lifelike capturing of human movement. Performance animation as well as motion capturing were at that time not suitable for capturing these finely differentiated movements. Why is this? In addition to restrictions put on the movements of the wearer of sensor fitted suits, particularly complex and concurrent harmonious shifts between slow non-linear accelerated movements and faster non-linear accelerated movements of various human limbs can evidently not be sufficiently portrayed. Motion Capturing by way of optical systems (e.g. VICON), also present major problems. Markers are attached to the skin in order to record locations within a space, which are recorded by special cameras to steer a computer model. These markers are often covered during movements (Wu et.al., 2011) or slip back and forth between the skin and bones of the wearer (soft tissue artifacts) (Kuo et.al., 2011). Both problems produce numerous motion artifacts. Systems without markers (e.g. MS KINECT) frequently suffer from interferences (Dutta, 2012) or covering of body parts resulting in a loss of trackings (Stoyanov et.al., 2011), notwithstanding a very low number of body parts captured (Obdrzalek et.al., 2012). Thus the possibilities offered by these techniques are at present very limited for the purpose of vividly capturing natural human movement. As a basis for the development work presented below, as lesser known procedure was selected: The *Scriptanimation* Method (Kempter, 1999): The startingpoint of this technique is a 2D video. The observed behavior of the persons in the video is transferred by an operator onto a 3D stick figure consisting of 16 parts, remotely resembling a human. Each part of this 3D figure is steerable. The operator then positions each part of the 3D figure exactly like the observed movements of a person on the video (25 fps), frame by frame. The positions generated with the 3D figure are written into a program-internal executable script, arranged according to time (columns) and spatial position of individual stick figure parts (rows). Through this position-time notation (script), the same figure can be reanimated by the values entered into the script. Advantages: All individual parts of the 3D computergenerated model (stick figure) can be precisely positioned. Since these positions are captured frame by frame and exactly replicated in the temporal resolution of movements eventually reanimated, both negative and positive accelerations in movement behavior can be recorded and reconstructed accurately. However, this procedure has three major disadvantages. 1. The character used has very little in common with the realistic possibilities of human movements. This consequently also applies for its operation. 2. It was not possible to manipulate objects with the character and it was not possible to insert the character into the setting of the video. No interface (e.g. FBX-Data Format) for the purpose of commercial animation programs exsisted (e.g. 3DsMax Autodesk). 3. The graphic user interfaces (GUI) and the control and feature elements were not particularly user-friendly. Operating the software demanded a long training phase, proved to be quite bulky and often led to frustration on the part of the operator. These disadvantages needed to be eliminated.

Since no anatomically correctly constructed 3D humanoid character had been available on the open market, we first had to create a bone structure which matched the functional anatomy and the average proportions of an adult human. This constructed character consists of 160 elements, connected to 29 functional bodyparts in a parent-child relation and can be moved in 80 dimensions. The skin of the character is adapted to this bone structure and the body parts are *limited* in regards to their angles of movement to realistically resemble that of natural human movement and saved in FBX format. As a result of using this format, the settings reconstructed in 3DsMax (Autodesk) can be connected with the 3D humanoid characters. The 3D humanoid characters, which are movable entirely in accordance with the natural human anatomy, could then be incorporated into our newly constructed tool- *ScanimFBX* (Zysk, Filkov, Feldmann, 2013).

The *ScanimFBX* is a software package with two separate tools, with separate functions:

- The tool 1 *ScanimFBX* serves for the user-friendly production of natural 3D movement data for 3D humanoid characters [described in detail in (Zysk, Filkov, Feldman, 2013)] and
- The tool 2 *ScanimFBX-AnimationPlayer*, which serves as a user-friendly, in selfpaced mode operable animation player.

Briefly to Tool 1, *ScanimFBX*: The user interface and control and feature elements are designed in such a user-friendly way, as to allow computer laypersons to produce between 10 and 30 images per day in a frame-to-frame procedure, after a short introductory training (2 to 3 times, 2 hours per person, individual or group sessions with 2 to 3 participants). In the process, the operator adjusts the 29 bodyparts of the character in correspondence with the positions of the bodyparts of the person observed in the video clips (see above: videotaped critical incidents). The deviation of the humanoid character from the movements of the person observed in the video is minimal- less than 5%- in all angles. The entire production process is illustrated in Figure 2.

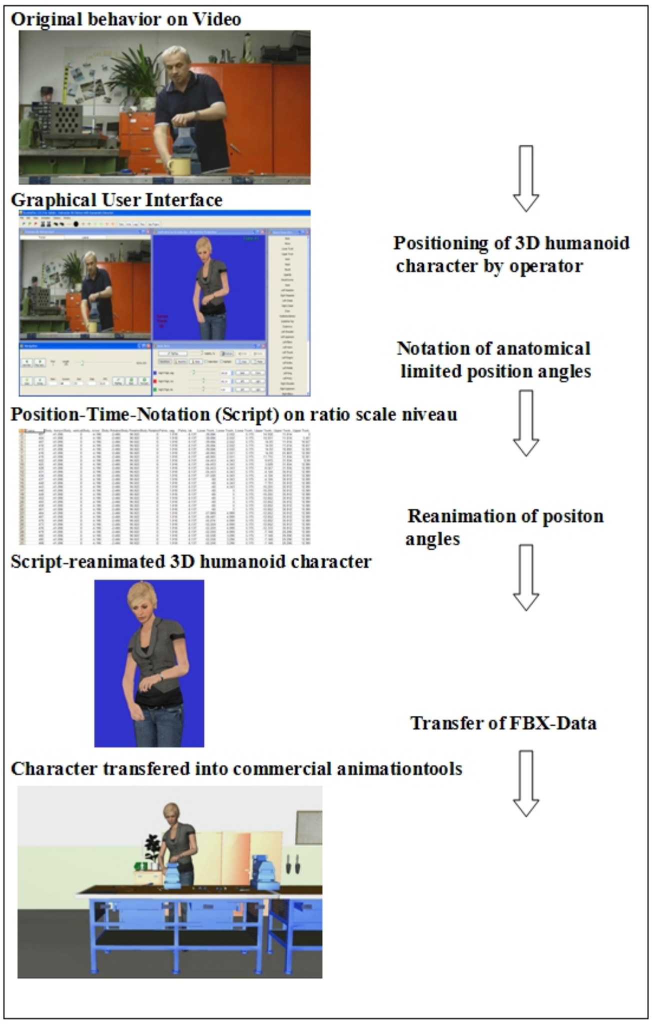

Fig. 2 Reproduction and reanimation of visible human behavior with *ScanimFBX*

Tool 2: *ScanimFBX-AnimationPlayer* - A new Tool for selfpaced learning

The ScanimFBX-AnimationPlayer is a programmable animation tool by way of FBX-SDK 2010.2 (Autodesk). It allows for stored movements of the 3D humanoid character from ScanimFBX in FBX-Data Format to be studied, with or without an animated setting, from all perspectives, at any desirable speed (1-100 fps) or image step width, at random close-up (zoom function), frame precise and in loop and swing (see Figure 3).

Fig. 3a Original 2D videotaped critical incident scene. (Here an instructor demonstrates the optimal movement sequences during filing).

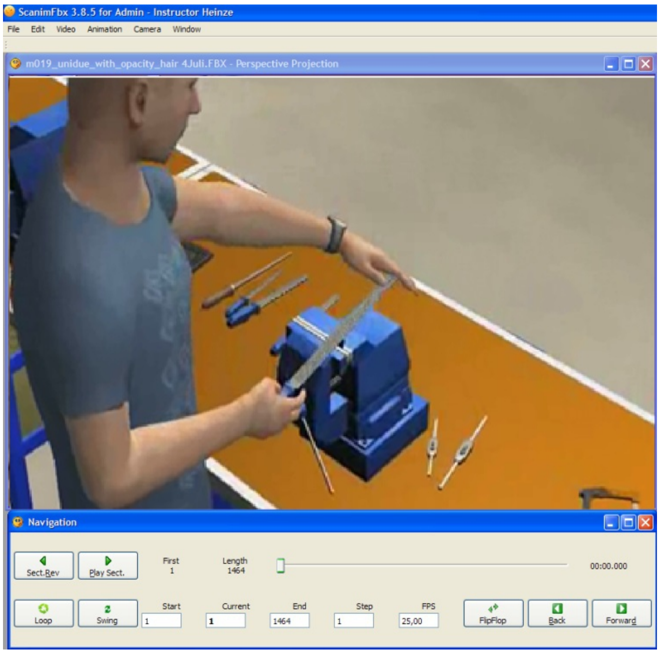

Fig. 3b The exact same movement sequences from Figure 3a can be studied with the *ScanimFBX-AnimationPlayer* from all perspectives, etc., by using the 3D humanoid characters.

Thus the user is able to freely decide what he or she wants to see, for how long, how often, from which perspective, which proximity, sequence and speed. Furthermore, the user can view both optimal and less-optimal movement sequences movements to be studied in isolated movement phases, as selected by experts on the grounds of which of these movements are decisive for optimal movement. Filing serves as an example: 1. What is the optimal leg position, 2. what is the optimal corresponding leg movement. 3. what is the optimal handle movement during filing? and so forth.

Virtually all functions of the *ScanimFBX-AnimationsPlayer* are controllable by a mouse as well as buttons in a navigation window. In the process, the humanoid characters can be manipulated by pointing the mouse directly over the animation window of the AnimationPlayer, in all movement directions. The function and operation features in the navigation window have proven superior in terms off user-friendliness for learning purposes in the first usability screening (20 test persons) in comparison to the QUICK-TIME-PLAYER (Apple) and FBX-REVIEW (Autodesk), the former off which has an installed FBX-Plugin (see Figure 4).

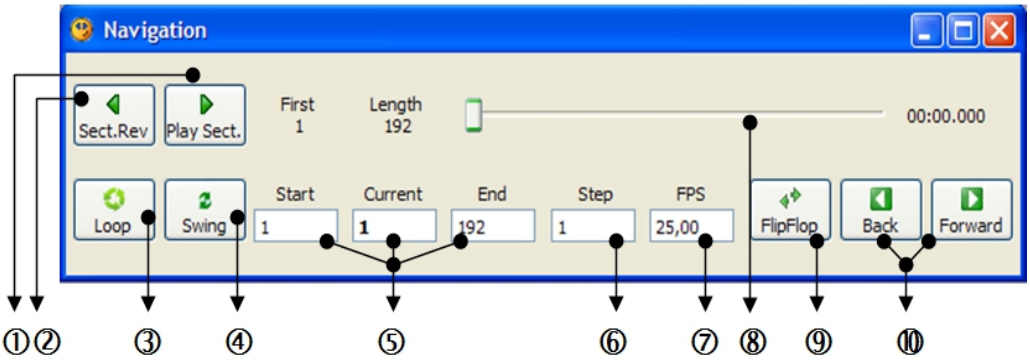

Fig.4 Screenshot section ScanimFBX AnimationPlayer Navigation-Window (Windows-Platform): Function buttons: 1. “Sect.Rev.” A defined section of the animation in reverse start/stop resume; 2.“Play Sect.”: A defined section of the animation in forward start/stop resume;3. “Loop”; 4. “Swing”; Fields: All contents or figures can be randomly defined. It is possible to script directly onto the fields. 5. Definition of Start-Current-End. 6. Step width 7. Speed („FPS“ 1-100 fps). 8. Timeline with date information and flexible slider. 9. “FlipFlop“: Flipping back and forth between two defined pictures. 10. “Back” and “Forward”.

The following results of usability screenings should be emphasized:

- The absence of a “start/stop button” was greeted by test persons as an advancement
- When the animation is played to the end, the last frame remains visible. If the “Play Sect.” button (Fig. 2,4), is pressed again, the first frame is displayed, without the sequence immediately starting again. The button must be pressed again for the sequence to start playing. Also this function contributed to giving test persons the feeling of safe acting control over the Player.
- The possibility of viewing the animation in reverse play, frame accurate and at any desired speed by using the “Sect.Rev” button, was described by all test persons as “very interesting.”
- Both the “loop” and the “swing” functions were described as “very helpful.” This can only be attributed to the fact that the speed, frame accuracy and the section of the sequence were randomly definable, which made it possible for each test person to study different relevant positions in detail, at any speed and as often as necessary.
- The “FlipFlop” function was rated by all test persons as “very informative.” This was especially the case when a frame-accurate study of transitions or changes in movement sequences became possible.
- Addition: Fourteen of the test persons reported that they found this type of observation and working processes “enjoyable” causing them them to be more interested in the theme itself. One was now also able to notice the clear difference between doing it “correctly” and “completely wrong,” and thereby learn to avoid mistakes.

Challenges

The *ScanimFBX-AnimationPlayer* should in the very near future become a standalone device with the use of the newest FBX-SDK programming. The independence of this self-contained tool from the *ScanimFBX* should thereby be underlined and the should be usable detached from it. Numerous additional didactical features should be demonstrated, e.g. the simultaneous comparison of optimal versus sub-obtimal and vice versa, through step-by-step changes in movement. Notwithstanding, ongoing experiments are focused on examining possibilities to improve Motion-Capture techniques in recording certain movements (coarse, highly repetitive), so as to incorporate these into the animation of humanoid characters. Further field of application, such as capturing movement for use in operating instructions or first aid manuals are being considered**.**

Conclusion

This paper presents an overview of preliminary research, results and ongoing work on a new user-friendly FBX AnimationPlayer. This tool, as an additional autodidactical learning aid, should serve trainees as a means to study 3D humanoid characters, which move in a precise and natural fashion, in selfpaced mode, at the hand of expertly controlled and optimized subject- and profession-specific movement sequences performed by the 3D humanoid characters. The first usability screenings suggest that through the implementation of new steering and functional components, a higher usability and thus, acceptance, exists in comparison to animation players commercially available at this time. The additional motivational aspect could suggest that persons who are less motivated when faced with learning from training manuals, find more fitting access to vocationally relevant study content or feel more motivated through such a tool to carry through with their study objectives.

References:

Benjamin, L.T. A History of Teaching Maschines. American Psychologist, Vol. 43, No.9 (1988), pp. 703-712.

Bandura, A. & Huston, A.C. Identification as Process of Incidental Learning. Journal of Abnormal and Social Psychology, Vol.63, No.2 (1961), pp. 311-318.

Chomsky, N. (1959). Verbal behavior by B.F. Skinner. Language, Vol.35, No.1, Jan.-Mar., 1959, pp. 26-58.

Dutta, T. Evaluation of the Kinect sensor for 3-D kinematic measurement in the workplace. Applied Ergonomics, 43 (2012), pp.645-649.

Edison, T.A. The Evolution of the Motion Picture: VI - Looking into the Future with Thomas A. Edison by Frederick James Smith. Interview in: The New York Dramatic Mirror, Page 24, Column 3, July 1913.

Flanagan, J.C. The Critical Incidents Technique. Psychological Bulletin, Vol.51, 4 (1954), pp.327-358.

Kempter, G. Das Bild vom Anderen. Skriptanimation als Methode zur Untersuchung spontaner Attributions-prozesse. PABST (1999), Germany.

Kuo, M.Y., Tsai, T.Y., Lin, C.C., Lu, T.W., Hsu, H.C., Shen, W.C. Influence of soft tissue artifacts on the calculated kinematics and kinetics of total knee replacements during sit-to-stand. Gait & Posture, 33, 3 (2011), pp. 379-384.

Mori, M. The uncanny valley. Energy 7, 4 (1970), pp. 33-35.

Obdrzalek, S., Kurillo, G., Ofli, F., Bajcsy, R. Accuracy and rubustness of Kinect pose estimation in the context of coaching of elderly population. Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE (2012), pp.1188-1193.

Skinner, B.F. Teaching Machines. Science, 128 (1958), pp.969-977.

Staddon, J. (2001). The new behaviorism. Taylor & Francis Group.

Wu, Q. & Boulanger, P. Real-time Estimation of Missing Markers for Reconstruction of Human Motion. Publish in: Virtual Reality (SVR), XIII Symposium (2011), pp. 161-168.

Zysk, W., Filkov, R., Feldmann, S. Bridging the uncanny valley – From 3D humanoid Characters to Virtual Tutors. The Second International Conference on E-Learning and E-Technologies in Education, ICEEE (2013). ISBN: 978-1-4673-5093-8 ©2013 IEEE

Zysk, W. & Filkov, R. Searching for Critical Incidents with AYP-MoviePlayer – A userfriendly new tool for analyzing videographed visible human behavior in social research. The Second International Conference on E-Learning and E-Technologies in Education, ICEEE (2013). ISBN: 978-1-4673-5093-8 ©2013 IEEE