Towards a Reactive Virtual Trainer

Zsófia Ruttkay, Job Zwiers, Herwin van Welbergen, and Dennis Reidsma

HMI, Dept. of CS, University of Twente,
POBox 217, 7500AE Enschede, The Netherlands
{zsofi, zwiers, welberge, dennisr}@ewi.utwente.nl
http://hmi.ewi.utwente.nl

Abstract. A Reactive Virtual Trainer (RVT) is an Intelligent Virtual Agent (IVA) capable of presenting physical exercises that are to be performed by a human, monitoring the user and providing feedback at different levels. Depending on the motivation and the application context, the exercises may be general ones of fitness to improve the user's physical condition, special exercises to be performed from time to time during work to prevent for example RSI, or physiotherapy exercises with medical indications. In the paper we discuss the functional and technical requirements of a framework which can be used to author specific RVT applications. The focus is on the reactivity of the RVT, manifested in natural language comments on readjusting the tempo, pointing out mistakes or rescheduling the exercises. We outline the components we have implemented so far: our animation engine, the composition of exercises from basic motions and the module for analysis of tempo in acoustic input.

1 Introduction

You have been spending hours working in front of your computer. All of a sudden your friendly Office Trainer – one whom you like the look of – greets you on your screen, suggests that you perform some 5 minutes of exercises right now (or somewhat later, if you do not want to be interrupted instantly) and 'dictates' the exercises for you. To make it less of a routine, you may choose from different pieces of music to give the tempo. Before the session you may indicate if you have specific complaints such as a stiff neck, or ache in the back or in your fingers. The Office Trainer offers appropriately tailored exercises and after a few sessions she asks about your progress.

Another scenario from real life: after a severe illness and several weeks in hospital a patient needs to do regularly special exercises to regain the functioning of certain muscles, to be able to use his hands. However, regular visits to a physiotherapist are not possible due to some forbidding constraint (shortage of experts, distance, lack of money etc.). But there is the Virtual Physiotherapist (VP), programmed by an expert with the sequence of exercises to be done at home. The VP explains each exercise and then coaches the patient, adjusting for example the tempo and the number of repetitions of the exercises, if needed. And she keeps the patient motivated by giving encouragement, feedback, even small talk. The real physiotherapist comes along only occasionally to supervise the progress and to 'instruct' the VP for the next sessions.

J. Gratch et al. (Eds.): IVA 2006, LNAI 4133, pp. 292 – 303, 2006.

In yet another scenario, a patient needs to do recuperation with the aid of a dedicated device which measures biomedical signals and motion characteristics. The Virtual Trainer gives feedback on whether he is doing well, what he should do in order to improve, or explains another, easier exercise to switch to.

These examples all come from real life. There are more and more people in need of such services due to the growth of the elderly population, the alarming increasing number of people who are overweight [28] and the more and more 'white-collar' jobs done via the computer in Western societies. The application context is a new domain where the usefulness of IVAs for society can be demonstrated.

In general, we will use the terminology Reactive Virtual Trainer (RVT) for IVAs in the previous and similar scenario. The RVT is assumed to be a she, while the real human a he, for sake of simplicity. The practical potential of a RVT is that she is universal: with minor investment she can be applied in any number to suit different users and objectives. Moreover, the RVT shares the most essential characteristics of real trainers such as physiotherapists: she is *reactive*, both on a strictly professional and on a psychological/sociological basis. The latter has proven to be essential to keep the client motivated (not to give up an exercise or drop out) and to make the repetitive exercises more fun [3].

The scientific challenges are related to the three major capabilities of the RVT:

(1) perceiving the user's performance in non-lab environments by using robust and non-intrusive devices;
(2) evaluating the performance of the user and deciding how to react;
(3) generating the reaction of the RVT accordingly and presenting it in a natural, subtle way, where motion, speech and (eventually) music are fine-tuned and coordinated.

For our research group, who have been working on multimodal interaction and human body animation, the application poses the following new challenges in these fields:

- to develop a framework where multiple paradigms (forward/inverse kinematics, motion capture) can be used to define 'building blocks' of pieces of motions, and these different pieces can be parameterised and combined into more complex and/or longer exercises in a uniform way;
- to develop strategies, models and real-time algorithms for smooth adaptation of scheduled motion sequences to observed (re)actions of the user.

This paper reports on our ongoing work towards a RVT. In section 2, we give an overview of related work. In section 3, we articulate the requirements from the point of view of the application and the scientific tasks associated, and describe the modular architecture of the RVT. Section 4 is devoted to already implemented components: the low-level real-time animation engine, the GESTYLE language tailored to author single elements or a complete sequence of exercises, with parameters to be adjusted at run-time, and the acoustic perception module. Finally we outline the issues to be tackled next.

2 Related Work

In *health care applications* where motion is central virtual agents have been used in two roles:

(1) as a *medium* to present certain motions to be mimicked by the patients, and
(2) as an *empathic consultant* to aid and coach the patient with the same empathy and psychological insight as a real expert would do.

For the first type of application, a Tai Chi exercise performer was developed recently [6]. An exercise is generated by finding motion samples in a huge indexed database appropriate for a natural language utterance. The motion samples have been gained by motion capture, and are not altered, in any respect, for example tempo.

There are more examples of medical and psychological consultancy applications, where empathic feedback has turned out to be as a basis for success [4, 16]. Finally, we mention a web-site [7] and two recent workshops devoted to the topic [11, 17]. The possibility of using computer vision for physiotherapy was raised [22]. At one of these workshops [17] the interest was aroused by novel devices to sense physical activity. Such devices, as well as traditional ones providing biomedical signals, may be useful in specific application contexts for a RVT. At the same event there have been industrial proposals to use portable handheld devices during exercises, to keep personalized training instructions downloaded from a web location or a PC and to collect logs of the performance. A physical robot has also been considered as a training companion, but again only as a social coach, without the function of 'making the user move'.

Two industrial projects are the closest to ours. The virtual Fitness Trainer of Philips, projected onto an immersive screen in front of the trainee exercising on a home-trainer, evaluates his performance based on heart rate feedback [9, 10]. Our goals, however, differ substantially, by giving an active, motion-presentation role to the RVT, endowing it with synthetic vision and hearing, empathic multimodal response capabilities and also intelligent motion plan revision. We expect that in our case, where the RVT has more intense and more natural contact with the trainee, the positive effect will be much more significant than what the Philips researchers have reported so far [27].

Sony's new EyeToy: Kinetic 'game' [24] offers personalised as well as ready-made sequences of fitness exercises, commented on by one of two virtual trainers. Though we have not been able to try out the game, on the basis of the descriptions on the web it seems that our project is more challenging just in the two aspects which are the basis for the success of the game: personalization and active feedback from the trainers. Moreover, the EyeToy game is closed; it offers no real authoring tool to customise it.

Finally, we rely on our own earlier work on a dancing ECA which shares many characteristics with a RVT [20]. The virtual dancer times her dance movements to the (external) audio of the music through real-time tempo and beat analysis. She also reacts to the motion of the user - her dancing partner - in front of a camera. The selection of dance moves from a repertoire and the adjustment of the style of these movements reflect global properties of the movement of the user. The dancing movements of the ECA are constructed dynamically, on the fly by selecting from a limited repertoire of basic movements labelled from different aspects.

3 The Framework for an RVT

In this section we discuss the functional requirements and outline a corresponding modular architecture for the RVT. In Figure 1 the components of the RVT are indicated. The major components are *user monitoring*, *action planning* for the RVT and *presentation* of the actions (exercises movements and/or multimodal feedback) by the RVT. Optional modules such as sensors capable of monitoring biosignals are also indicated. A very important non-technological issue is the interaction with the expert who should author the exercises and monitor the global progress of the user, but who may be not very experienced technologically.

3.1 Authoring Exercises

How, when and by whom should the RVT be 'programmed'? Depending on the specific application, the RVT can be pre-programmed for a general, and from a medical point of view not critical, task such as providing morning gymnastic for healthy women around 30. The other extreme is when the RVT acts as a kind of 'assistant' to an expert, who instructs the RVT regularly what exercises to do in the next two weeks with a given patient. The definition of exercises is the task of an authorised person such as an expert physiotherapist. The authorized person should be able to prescribe exercises built up from pieces of basic motions and earlier defined sequences. The authoring should be easy, using high-level scripting language, templates for parametrizing exercises, or restricted natural language. Relative directions and references to body parts should be used, as done e.g. in our earlier work [21]. The tempo of the exercise may be set by acoustic cues (e.g. given by counting, clapping), or in a qualitative or quantitative way. The library of basic motions may be defined in terms of poses and paths, also using some reference points on or around the body. Once some key poses are defined, the motion path and parameters may be chosen from a predefined set. However, it should be possible to define very specific, new motion pieces which cannot be given in the previous frameworks. There should be motion capture means that allow the authoring person to directly 'act out' a sophisticated motion of the repertoire.

3.2 Calibration

The *calibration* of the RVT may have different motivations, affecting different aspects. The *calibration of motion characteristics* to those of the user should assure that the RVT's motion performance remains within the range of the actual, or realistically achievable capabilities of the user. The calibration may take place on a general level, by specifying motion parameters to be used throughout the exercises, as appropriate for the age and physical state of the user. But it may also involve the specification of certain joint parameters (such as stiffness limiting acceleration, or extreme position).

The *calibration of body geometry* similar to the user's *may* have a positive effect on presenting the exercises: a short-legged, bulky person will not be able to mimic the movements of a tall, thin trainer. On the other hand, it is not true per se that a user prefers a RVT with body geometry, weight and age characteristics of his own. Besides the mimicking of exercises, other factors such as aesthetic appeal and a

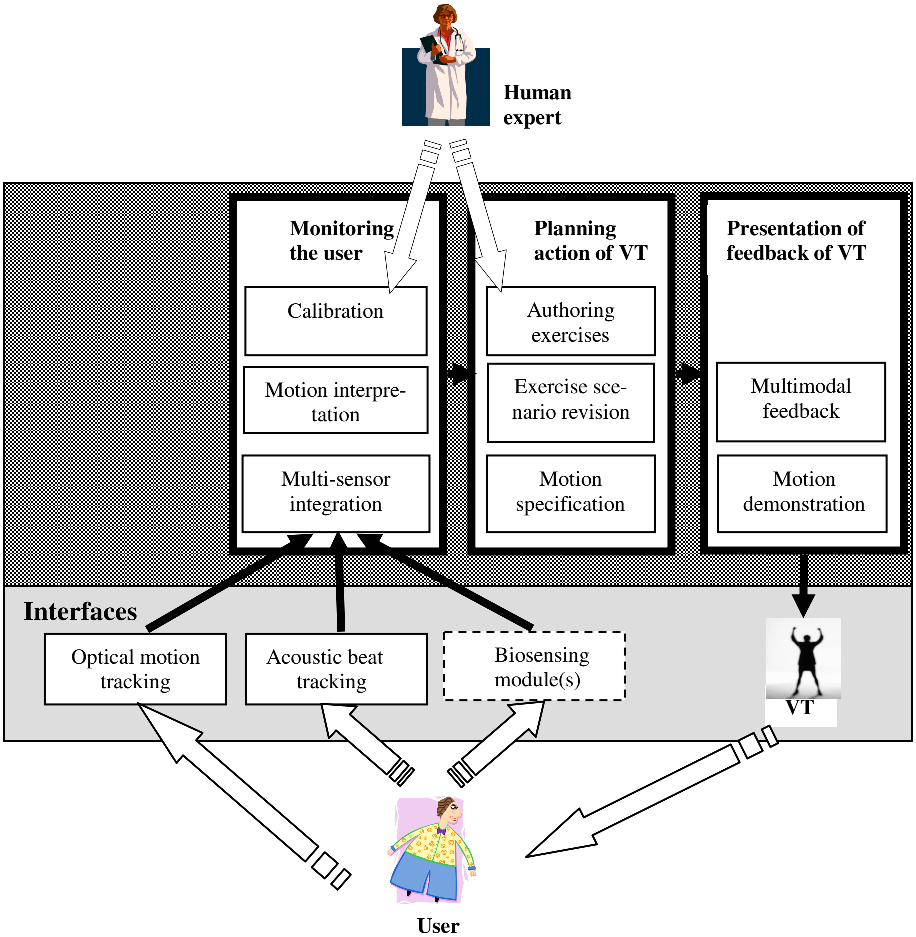

Fig. 1. Architecture of the RVT

preference for matching or different gender or young and good-looking RVTs may override the benefits of letting the RTV mirror the user.

These calibration tasks may involve an expert and in some cases even motion capture, but in less critical cases simple body parameters of the user suffice. An interesting possibility is when the RVT himself is prepared to do an initial self-calibration session by asking the user to perform a few poses and adjusting his own geometry and motion characteristics accordingly. It is also an option to re-calibrate the user from time to time, as his shape and movements improve.

3.3 Motion Demonstration

The most essential capability of the RVT is to act out the exercises prescribed by a script according to the initial body calibration parameters and (maybe dynamically changing) motion parameters. Concatenation of unit motions, automatic transition to start or rest poses and changes in timing parameters as well as 'graceful suspension' of an exercise should be taken care of by the 'motion intelligence', a component of the animation engine.

3.4 Monitoring User Performance

The RVT should monitor the performance of the user in near real time, and in a non-obtrusive way. In the basic scenario this should be achieved by robust synthetic vision getting input from a single, every-day camera connected to the PC hosting the RVT. In the case of some applications biosignals (e.g. heart-rate) are also appropriate. Through these signals, the presence, tempo of motion, the (basic) morphology of the performed movements as well as the physical state of the user should be perceived. By tracking the face too, not only the facial expression, but also coloration and reflection (sweat) may be used as source of information about the physical state.

3.5 Perceiving Acoustic Signals

Besides monitoring the user by synthetic vision, acoustic signals may also be helpful to detect if he is jogging in the right tempo, or if he is out of breath. Another function of the acoustic perception is to define the tempo in the authoring stage, for example by acoustic feedback such as counting or tapping with the feet. Finally, musical beat detection could be useful in assigning pieces of background music to exercises, either stored on the computer in digital form or played for the microphone on some device of the user.

In the first case, detection of tempo in the acoustic cue should be done real-time. The detection of the tempo and the beat of music may be done off-line, if the pieces of music to be used are pre-selected, or on-line, if the user provides new pieces of music from external sources at the time of doing the exercises.

3.6 Reactive Adjustment of Exercises

The RVT observes the user continuously and reacts to the situation by adjusting the current scenario of exercises and accompanying verbal comments. The reaction may involve:

- the adjustment of certain motion parameters of the current exercise;
- the re-scheduling of the exercises to be performed;
- some speech and/or nonverbal feedback to acknowledge performance or to inform about the modifications above (discussed in the next section).

Any single reaction or a subset of the above type of reactions may be triggered. For example, if the RVT notices that the user's rhythm is slower, the reactions may be:

- slowing down the RVT's tempo, to help the user to remain in sync;
- deciding to finish the current exercise and to shorten the number of repetitions still to come, as the user is possibly too tired to be able to keep up with the original schedule;
- warning the user about his delay, giving encouragement, and helping him by counting in a raised voice to get back to the tempo.

Note that in the first case the user drives the RVT indirectly, by his motion. The mechanism to decide when and how to react to the user's performance should be based on expert knowledge, pedagogical goals and a model of the user, reflecting also his physical state. Variations in the style of the RVT may be covered too [12].

3.7 Empathic Multimodal Feedback Generation

The RVT should address the user from time to time, during or in between performing an exercise. As the feedback is a crucial component for success, both from the point of view of the *effect* and *engagement*, it should be believable and subtle, in the following aspects:

- for each type of feedback a set of different natural language utterances should be used to achieve variety, accompanied (or even replaced) by facial expressions and possibly hand-gestures;
- some utterances should be synchronized to the tempo of the exercises (counting, indicating when to finish a sequence);
- addressing the user should be made clear by gaze behavior and head movement, whenever the exercise being performed allows.

4 Reactive Motion Generation

In this section we explain the computational means we have developed so far for the following major tasks: the representation of motion building blocks, the definition of parameterized exercises and driving the RVT by acoustic cues.

4.1 Defining Basic and Compound Motions

Three Kinds of Basic Motions

A *basic motion* is the smallest unit of motion that can be used (possibly in an altered form) to build up compound motions and exercises. A basic motion involves joints of different parts of the body, according to a taxonomy corresponding to the hierarchy of joints (left/right limbs, arms, hands; for a taxonomy see [21]). We use three kinds of basic motions to be explained below. In all cases, a basic motion has at least 2 key postures (the start and end pose), as well as possible in-between poses defining the path through animation space.

In the case of *forward kinematics motion* (FKM), the key poses are defined explicitly by the rotation of the joints involved. Movements are defined as interpolations through key poses. The interpolation determines for a large part the style of execution of movements, and can range from simple linear interpolation between the poses through nonlinear functions giving certain expressivity to a movement to possibly user-specific interpolation functions. A function may assigning the rotation to each time moment of the entire duration, see in Figure 2 the rotation of the wrists.

In the case of *inverse kinematics motion* (IKM), a position is specified on or around the body, and one of the so-called end-effectors - such as the end of a finger or one of the wrists - has to touch or reach for the given position. Similarly, the expert, when authoring an exercise, may specify a motion path by identifying points around the body, to be followed by the right hand, and the IK automatically calculates the rotations of the shoulder and elbow needed to place the hand on this path. The path may be given in terms of Hermite splines. In Figure 2, the path the wrist has to follow is given as a 3D curve parameterized by time.

Basic motions may also be specified in terms of *knowledge-based animation models*, specific for some body parts, as we have used in our earlier work on a virtual presenter [26]. Such models are based on human movement theory, such as Donder's law for head movement [25], detailed analysis of captured motion [1] or more ad-hoc solutions, such as the use of stochastic noise to add expressiveness to a character [19].

Motion Parameters

The motion of the RVT can be modified in real time by adjusting – for the time being – three motion parameters: *tempo*, indicating the timing of the repetitive, rhythmic motions; *amplitude*, indicating the amount of motion performed; and *effort*, indicating the acceleration profile of the motion. Our effort and amplitude parameters are similar to the effort and shape parameters of the Laban Notation, also taken as a starting point in [2] and [8]. In these related works for expressive gesture accompanying speech motion path, hand shape and acceleration profile are taken as dependent variables, and more than one of these are influenced by the higher-level gesture motion parameters. For the motion exercises domain we find it more useful to keep parameters influencing the motion dynamics, the motion path and the morphology separate, as these features are often addressed individually in fitness, and often need to be carefully controlled in physiotherapy exercises. By parameterizing movement, one can put the variants of a movement to different uses: for example tailoring general exercises to specific users (tempo, amplitude, target positions for IKM), accentuating certain movement features (effort) and performing at a slow tempo for a demonstration, and aligning movement in the exercise to the tempo according to music or clapping.

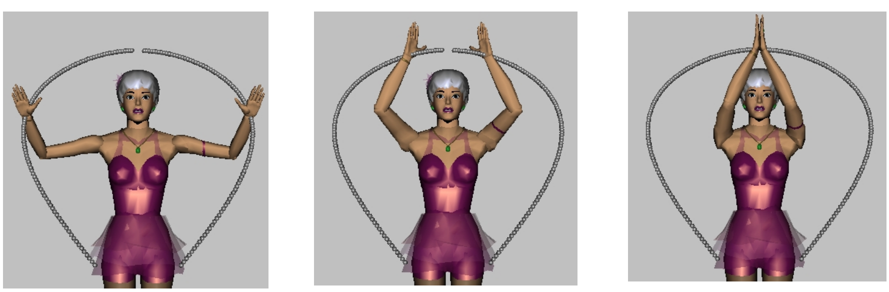

Fig. 2. A clapping exercise in our animation tool, where the movement path shown for the hands and the rotation of the wrists are defined by functions of time t, assigning a 3D point and a triple of Euler angles to each time moment of the entire duration

The position of the left hand:	Rotation of the left wrists:
X(t) = 0.18 + sin(t*3.14*1.185)*0.3	Around-x-axis(t) = 0.5*t
Y(t) = 0.79 + sin(t* 3.14 *0.5)*0.9	Around-y-axis(t) = 0
Z(t) = 0.2	Around-z-axis(t) = 0.16* 3.14*t

Combining Motions

Basic motions can be used in their parameterized form to compose more complex ones. This is done by using the GESTYLE scripting language [18]. Gestures may be defined by *motion expressions*, built up from basic motions, by using the *parallel* and *sequential* compositions and *repeat* operators in an embedded way. For example, a clap above the head is defined as the parallel execution of two single-armed movements above the head in the plane of the body, each composed of two basic motions, one for the arm movement, and one for the palm orientation. The definition of this clap in GESTYLE is:

```
<DefGest Name="LeftHand_ClapAboveHead">
  <PAR>     <UseGest Name="LeftHand_Lift" />
            <UseGest Name="LeftWrist_Rot"/></PAR> </DefGest>
<DefGest Name="RightHand_ClapAboveHead">
  <PAR>     <UseGest Name="RightHand_Lift" />
            <UseGest Name="RightWrist_Rot"/></PAR> </DefGest>

<DefGest Name="ClapAboveHead">
  <PAR>     <UseGest Name=" LeftHand_ClapAboveHead" />
            <UseGest Name=" RightHand_ClapAboveHead"/> </PAR> </DefGest>
```

A more sophisticated clapping movement where the wrists are rotated in the last quarter of the entire duration of the motion may be defined by exploiting that GESTYLE which allows delayed start of parallel threads. For the left hand, this looks like:

```
<DefGest Name="LeftHand_ClapAboveHead_1">
  <PAR>     <UseGest Name=" LeftHand_Lift" sub_start_time = "0" />
            <UseGest Name=" LeftWrist_Rot" sub_start_time="0.75"/> </PAR> </DefGest>
```

Compound motions and longer exercises may be given names and have higher-level parameters of their own, which may be used to control the parameters of the building blocks. Timing can be subtly defined by adding ‘wait’ durations, and thus need not be perfectly parallel for each component. Times are given in terms of abstract units of beats; the actual timing will be computed according to the (maybe changing) specification of the duration of a beat. On this level, GESTYLE is used to author the repertoire which may be used to compose exercises.

Besides the tempo, the amplitude and effort parameters of the motion to be performed as well as the number of repetitions may be left unspecified for an exercise sequence. Moreover, wherever applicable, left, right or both limbs may be given as parameters. Hence GESTYLE also functions as a higher-level scripting language whith parameters referring to specific fine-tuned motions or (one of the half of) symmetric motions or, regarding motion direction, opposite motions. We are aware of the fact that even GESTYLE is too technical for a physiotherapist or a trainer, so ideally there should come a scripting language that is close to natural, as the authoring tool for exercise sequences.

4.2 Performing Motions

All the above types of animations are represented on the lowest level in terms of a function $f(t,\boldsymbol{a})$, mapping time t and the parameter vector $\boldsymbol{a}$ to a 3D point or a vector of a triple of rotational angles, in order to define animation path for IKM or rotation values for a certain joint for FKM, respectively. The time t ($0 \leq t \leq 1$) is relative, and the 1 time duration can be time-warped to real durations. $\boldsymbol{a}$ is a vector of low-level motion parameters that can be modified to allow animation changes in real time. By manipulating t, we can adjust the tempo and the velocity profile of an animation. Key positions given in the definition of basic motion may be used to align their time to external sources, for example to the beat of the music.

4.3 Driving Motion by Acoustic Cues

We use different types of auditory cues for RVT applications with different purposes (see 3.5). For all cases, the audio can be processed for tempo and beat information, which is then used to adapt the movements of the RVT to properly align them to the audio input or to determine whether the user is still performing the exercise in the expected tempo. In all of these situations the beat and tempo tracking can be performed by our improved implementation of Klapuri's beat tracking algorithm, capable of real-time performance. Klapuri's algorithm uses different frequency bands to detect accentuation in the audio signal, then a bank of comb filter resonators to detect the beat [14].

5 Further Work

Our current work follows three lines. First, yet missing modules are being developed. As for synthetic vision, we are going to check how useful the global information is on amplitude and tempo of motion that we can get from our real-time single camera image analysis module used to categorize dance motions [23, 20]. As it is known what motion the user should be performing, the recognition task may be easier. Moreover, the time evolvement of basic motions and exercises corresponds to certain patterns in the global characteristics, the timing of which may thus be captured. On the other hand, for subtle visual perception, such as detection of a 'hanging elbow' multiple cameras and special visual markers (e.g. wearing a dress with a color code for different parts of the body) will be necessary. Whether these extensions will be sufficient to gain information detailed enough for interpreting the correctness of physiotherapeutic motions is as yet an open question.

In cooperation with an expert physiotherapist we will develop a knowledge base of standard exercises with feasible default parameters, as well as coaching strategies for a specific application context. Then we will test how likeable and effective the RVT is.

In the second stage, we shall refine the system by extending it with calibration facilities and a choice of different RVTs and with interpretation of the user's state with respect to logged recent performance. Another important extension we are aiming at is the parameterization and re-use of motion-captured samples [13, 15].

Finally, we will continue experimenting with coordination of acoustic cues and motion. In particular, the synchronization of accentuated speech to the predefined motion tempo is an interesting option.

Our ultimate goal is to have different settings for RVTs, as suggested in the introduction, and to collect feedback from real users.

Acknowledgement

The acoustic beat detector has been implemented by P. Bos. We are thankful to the anonymous reviewers for their comments, and to D. Kiss and L. Packwood for proof-reading the paper.

References

1. K. Abdel-Malek, J. Yang, Z. Mi, V.C. Patel, K. Nebel: Human Upper Body Motion Prediction, Conference on Applied Simulation and Modeling (ASM) 2004, June 28-30, 2004, Rhodes Greece
2. N. Badler, R. Bindiganavale, J. Allbeck, W. Schuler, L. Zhao, M. Palmer: *Parameterized Action Representation for Virtual Human Agents*, In: [5] pp. 256-284.
3. T. Bickmore, J. Cassell: *Social Dialogue with Embodied Conversational Agents*, In: J. van Kuppevelt, L. Dybkjaer, and N. Bernsen (eds.), Natural, Intelligent and Effective Interaction with Multimodal Dialogue Systems. In press, New York: Kluwer Academic.
4. T. Bickmore, R. Picard: *Towards caring machines*, Proc. of CHI, 2004.
5. J. Cassell, J. Sullivan, S. Prevost and E. Churchill (eds.), *Embodied Conversational Agents*, MIT Press, 2000.
6. S-P.Chao, C-Y Chiu, S-N, Yang, T-G. Lin: *Tai Chi synthesizer: a motion synthesis framework based on key-postures and motion instructions*. Computer Animation and Virtual Worlds, Vol. 15. pp. 259-268, 2004.
7. Cybertherapy URL: http://www.cybertherapy.info/pages/main.htm
8. B. Hartmann, M. Mancini, C. Pelachaud: Implementing Expressive Gesture Synthesis for Embodied Conversational Agents, Gesture Workshop, LNAI, Springer, May 2005.
9. HomeLab Philips, http://www.research.philips.com/technologies/misc/homelab/
10. W. IJsselsteijn, Y. de Kort, J. Westerink, M. De Jager, R. Bonants: *Fun and Sports: Enhancing the Home Fitness Experience*, Proc. of ICEC 2004.
11. International Workshop on Virtual Rehabilitation, http://www.iwvr.org/2005
12. W. Lewis Johnson, P. Rizzo, W.E. Bosma, S. Kole, M. Ghijsen, H. van Welbergen: Generating socially appropriate tutorial dialog, in ISCA Workshop on Affective Dialogue Systems, Kloster Irsee, Germany, ISBN 3-540-22143-3, pp. 254-264, 2004.
13. T. H. Kim, S. I. Park, S. Y. Shin: *Rhythmic-motion synthesis based on motion-beat analysis*, ACM Trans. on Graphics, ACM Press, 22(3), pp. 392-401, 2003.
14. A. Klapuri, A. Eronen, J. Astola: *Analysis of the meter of acoustic musical signals*, in IEEE transactions on Speech and Audio Processing, (2006)
15. M. Mandel, V. Zordan: Beyond Ragdolls: *Generating Versatile Human Behaviors by Combining Motion Capture and Controlled Physical Simulation*, Proceedings of Game Developer's Conference (GDC), 2005.
16. S. Marsella, L. Johnson, and C. LaBore: *Interactive Pedagogical Drama for Health Interventions,* AIED 2003, 11th International Conference on Artificial Intelligence in Education, Australia, 2003.
17. Monitoring, measuring and motivating exercise ubiquitous computing to support fitness, Workshop at Ubicomp 2005, Tokyo, September 11, http://seattleweb.intel-research.net/projects/ubifit/ubicomp05workshop.html
18. H. Noot, Zs. Ruttkay: Style in Gesture, In: A. Camurri, G. Volpe (Eds.), Gesture-Based Communication in Human-Computer Interaction, LNCS 2915, Springer-Verlag, 2004.
19. K. Perlin, Ken: *An Image Synthesizer.* Proceedings of SIGGRAPH '85: 287-296.

20. D. Reidsma, A. Nijholt, R. Poppe, R. Rienks, H. Hondorp: Virtual rap dancer: *Invitation to dance*. In: Proceedings ACM CHI 2006 (Extended Abstracts), Montreal, April 2006, to appear.
21. Zs. Ruttkay, Z. Huang, A. Eliëns: *Reusable Gestures for Interactive Web Agents*, In: R. Aylett, D. Ballin, T. Rist (Eds.), Intelligent Virtual Agents, IVA-2003 Proceedings, LNAI 2792, Springer-Verlag, pp. 80-87.
22. S. Shafaei, M. Rahmati: *Physiotherapy Virtual Training by Computer Vision Approach*, Third International Workshop on Virtual Rehabilitation, 2004. http://www.iwvr.org/2004/
23. S.Takaaki, N. Atsushi and K. Ikeuchi: Rhythmic Motion Analysis using Motion Capture and Musical Information, Proc. of 2003 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems 2003, pp. 89-94.
24. Sony: EyeToy: Kinetic, http://www.us.playstation.com/Content/OGS/SCUS-97478/Site/ and http://www.eyetoykinetic.com
25. D. Tweed: A three-dimensional model of the human eye-head saccadic system. J. Neurophysiol. 77, 1997. pp. 654-666.
26. H. van Welbergen, A. Nijholt, D. Reidsma, J. Zwiers: *Presenting in Virtual Worlds: Towards an Architecture for a 3D Presenter explaining 2D-Presented Information*, IEEE transcactions on Intelligent Systems, special issue "Intelligent Technologies for Interactive Entertainment" to appear Sept./Oct. 2006
27. J. Westerink, M. de Jager, M., Y. de Kort, W., IJsselsteijn, R., Bonants, J. Vermeulen, J. van Herk, M. Roersma: Raising Motivation in Home Fitnessing: Effects of a Virtual Landscape and a Virtual Coach with Various Coaching Styles. *ISSP 11th World Congress of Sport Psychology,* 15 - 19 August 2005, Sydney, Australia.
28. The World Health Organization warns of the rising threat of heart disease and stroke as overweight and obesity rapidly increase http://www.who.int/mediacentre/news/releases/2005/pr44/en/