Human-Scale Haptic Interaction with a Reactive Virtual Human in a Real-Time Physics Simulator

SHOICHI HASEGAWA, ISHIKAWA TOSHIAKI,
NAOKI HASHIMOTO, MARC SALVATI, HIRONORI MITAKE,
YASUHARU KOIKE, AND MAKOTO SATO
Tokyo Institute Technology / PRESTO JST

In this article we propose a framework for haptic interaction with a reactive virtual human in a physically simulated virtual world. The user controls an avatar in the virtual world via human-scale haptic interface and interacts with the virtual human through the avatar. The virtual human recognizes the user's motion and reacts to it. We create a virtual boxing system as an application of the proposed framework. We performed an experiment to evaluate the validity of the reaction of the virtual human. We got confirmation that the proposed framework creates realistic reactions and that users can easily estimate the input motions of the avatar.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces -- *Haptic I/O*; I.3.6 [**Computer Graphics**]: Methodology and Techniques --*Interaction techniques*; I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism -- *Virtual reality*

General Terms: Algorithms, Measurement, Design, Human Factors

Additional Key Words and Phrases: Reactive virtual human, haptic interaction, physics simulator

---

1. INTRODUCTION

The realism of graphics and movement in computer entertainment has recently improved remarkably. Virtual humans are a key feature of computer entertainment, and their appearance was realized via GPU and motion-capture techniques.

While movies and animation require the offline creation of realistic motion, computer games need real-time interactions with virtual humans. Conventional computer games create realistic reactions for virtual humans by choosing and connecting motion data prepared offline that is appropriate for the user's button inputs. This technique is suitable for limited digital input;. but more direct interaction through analog pads or haptic interfaces requires a new paradigm for generating reactions.

In this article we propose to create reactions based on dynamic simulation and on a mental and motion control model based on humans. Moreover, we propose a framework for interacting with the virtual human through an avatar controlled by the user via a haptic interface.

2. RELATED WORK

Many virtual humans are used in computer games like RPGs, sports games, and action games. Fighting games like Virtua Fighter [Suzuki et al.1993] are pioneer users of virtual humans. In these

---

Authors' addresses: S. Hasegawa, Tokyo Institute of Technology, PRESTO JST, 4259 Nagatsuta Midori, Yokohama, Japan; email: hase@hi.pi.titech.ac.jp; T. Ishikawa, Tokyo Institute of Technology, 4259 Nagatsuta Midori, Yokohama, Japan; email: ishikawa@hi.pi.titech.ac.jp; N. Hashimoto, email: naoki@hi.pi.titech.ac.jp; M. Salvati, email: salvati_marc@hi.pi.titech.ac.jp; H. Mitake, email: mitake@hi.pi.titech.ac.jp; Y. Koike, email: koike@pi.titech.ac.jp; M. Sato, email: sato@pi.titech.ac.jp

games, the reactions of virtual humans are created by connecting motions in a motion database such as Motion Graph [Kovar et al. 2002.]. Therefore, variations among the reactions are limited, and there is a large cost to creating a motion database. Using a database approach, Lee and Lee [2004] created realistic motions for boxers; but their method did not create dynamic reaction for contacts.

Advances in human interface and virtual reality technologies make possible a wide variety of user inputs. Hence virtual humans are required to perform a wide variety of reactions. Jeong et al applied a database approach to create a reactive virtual human for haptic interaction [Jeong 2004.]. The interaction in their system is caching a ball, which is not a direct contact interaction due to restrictions on the variety of motions. Space-time constraint methods [Rose et al. 1995; Komura et al. 2000] generate an optimal trajectory and minimize some optimization functions. However, these methods must optimize the trajectory, and it is difficult to run them in real-time.

Dynamic control and simulation methods use controllers to compute joint torques based on current states and desired actions. These methods create dynamically correct motions from specified motions [Zordan and Hodgins 1999], state machines [Hodgins et al. 1995], and environmental physical input [Oshita and Makinouchi 2001]. These methods generate appropriate motions for variations of the input.

State machines are used to create the motions and actions of virtual characters that correspond to environments [Funge et al. 1999] and user instructions [Blumberg et al. 2002]. However, these researches did not include direct haptic interactions.

3. PROPOSED FRAMEWORK

Figure 1 shows an overview of the proposed framework. It consists of a model of a human body, dynamic simulators, a haptic interface, an avatar controller, and a cognitive and motion control model of a virtual human.

The user controls the avatar through a haptic interface. The avatar's model of the human body traces the motion of the user, while forces that act on the avatar's body model are fed back to the user. The cognitive model of the virtual human predicts the virtual world, selects the next action of the virtual human, and gives a motion instruction to the motion-control model. From the instructions, the motion-control model decides the joint torques which act on the virtual human's model of the human body.

3.1 Dynamics Simulator

We employ a dynamics simulator called Springhead [Hasegawa and Sato 2004] that is suitable for haptic interaction. Springhead employs Featherstone's method [Featherstone 1983] for joints, a penalty method for contacts, and requires less computation time for each iteration.

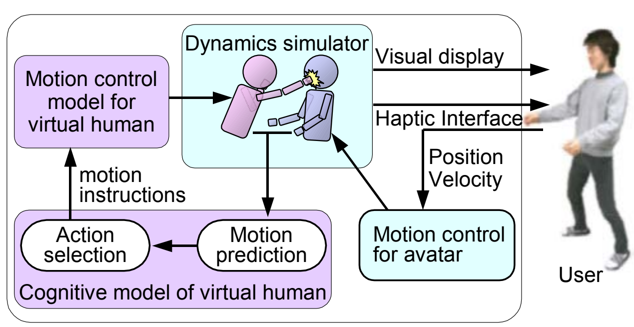

Fig. 1. Overview of the proposed framework.

3.2 Human Body Model

Figure 2 shows the human body model in the proposed framework. In this article we model the upper part of the body for a virtual boxing application, and use this model for both the virtual human and the avatar. We set the dimensions, weights, and inertias of the body parts and the limits of the joint angles by referring to databases on human characteristics [Kouchi et al.1998; National Institute of Technology and Evaluation 2004]. In addition, we give a default angle and a spring damper model for each joint to give a default posture for the model of the human body.

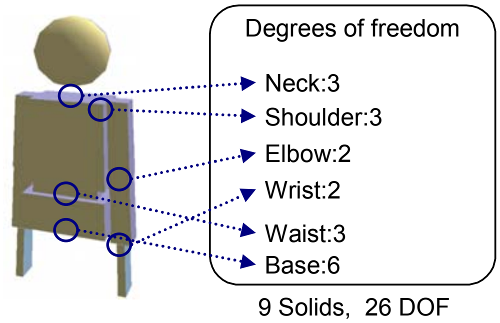

Fig. 2. The model of the human body.

3.3 Haptic Interface and Visual Display

We use a human-scale projection display to show real-time images of the virtual world. In addition, a human-scale both-hand haptic interface called SPIDAR-H [Cai et al.1996] is employed to control the avatar and to feed the forces from the avatar back to the user. Figure 3 shows the hardware setup of the proposed framework. Each hand of the user is pulled by four strings and gets three degrees of freedom (DOF) force feedback.

3.4 Avatar Control and Force Feedback

As shown in Figure 2, the avatar has 26 DOF, while the haptic interface has only 6 DOF. We have to estimate the avatar's entire DOF from the small input DOF. Inverse kinematics methods with a

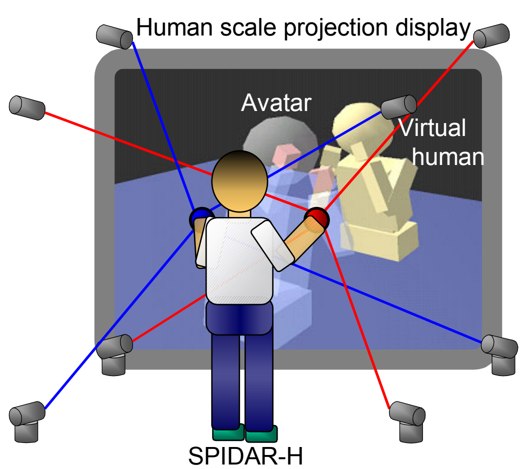

Fig. 3. The proposed framework's hardware setup.

pseudo-inverse matrix are often used for this purpose [Badler et al. 1993; Yamane and Nakamura 2003]. However, these methods do not regard the contact forces that are added by other objects such as an opponent in a boxing match. Hence we use dynamic simulation and PD control. The avatar's hands are pulled by a spring and damper model to the positions of the user's hands.

Because the avatar's human body model has joints and each joint has a spring and damper model to set the posture to default, the avatar comes to a balanced posture between the default posture and the hand positions. In addition, the contact forces, which act on the avatar's body, affect the balanced posture. We provide softer spring and damper models for the default posture, so the positions of the avatar's hands almost reach the positions of the user's hands.

The user should not feel the dynamics of the avatar's hands; but the user should feel the contact force that acts on the avatar's hands. Therefore, the framework does not feed-back the forces from the spring and damper model. The contact forces that act on the avatar's hands are fed back to the user's hands via the strings of SPIDAR-H.

3.5 Cognitive Model of the Virtual Human

Humans recognize environments and predict changes in them. Then they decide on their next actions, thus creating natural and smooth human motion In the proposed framework the virtual human's cognitive model mimics this function of the human mind. The cognitive model has an internal model of the environment to predict the user's movements.

Figure 4 shows the virtual human's cognitive model. In the proposed framework, the internal model is represented by a dynamics simulator, which is a copy of the simulator for the virtual world. With the dynamics simulator, the cognitive model predicts the user's movements and those of the virtual human itself. Then, the cognitive model analyzes those motions and contacts to decide on the next behavior and action. Many movements of the upper body can be represented by combinations of reaching motions. We represent an action as a reaching motion by a body part. The part, duration, and the target are sent to the motion control model.

3.6 The Motion Control Model for the Virtual Human

The motion control model creates a motor control signal for the virtual human from three parameters: the part, the duration, and the target.

Flash [1987] proposes a control model for human "reaching" movements from measurements of human motion and simulations. The model consists of a minimum "jerk" model and PD control of joints. We expand and modify the model and use it to control the movements of the virtual human. We describe the details of our "reaching" motion model in the Appendix.

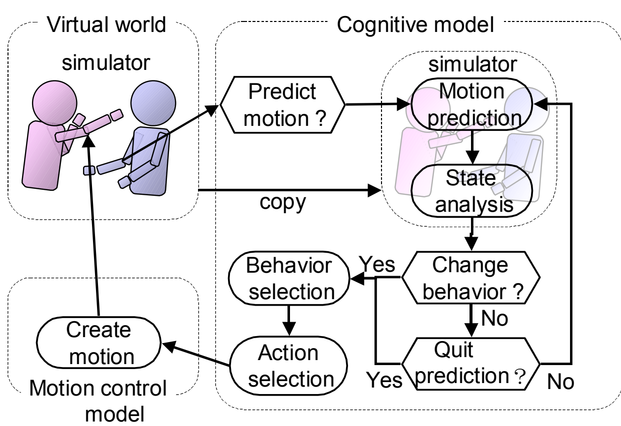

Fig. 4. The cognitive model of the virtual human.

Virtual Boxing Application. We created a virtual boxing application as an example of the proposed framework. The application creates reactions such as attacking, blocking, and dodging in the upper part of the virtual human's body. These movements are created by the "reaching" motion model. The following sections explain the application's implementation.

3.7 Predicting In the Virtual World

The virtual human's cognitive model predicts the environment with a dynamics simulator (see Section 3.5). The virtual human begins to predict when the avatar's hand moves toward the virtual human and the velocity of the avatar's hand achieves sufficient speed. For prediction, we copy the state of the simulator for the virtual world to the simulator for the internal model of the virtual human, which is used for prediction. The Simulator is for the internal model of the virtual human and prediction. We then ran the simulator for prediction faster than the simulator for the virtual world. The simulator for prediction will find contacts between the virtual human and the avatar before they actually occur in the virtual world.

Because the simulator for prediction runs faster than real-time, we have to predict the user's inputs and give them to the simulator. We make two assumptions for the prediction:

- The user's input can be predicted from the past input series by linear extrapolation.
- The virtual human continues the "reaching" motion that it started at the beginning of the prediction.

We give the predicted user's input and avatar's actions to the simulator.

3.8 Analyzing the Predicted Virtual World

The virtual human's cognitive model selects the next behaviors and actions from the analysis of the predicted contacts. We classify the contacts by the contact parts of the human body model for the virtual human and the avatar.

Table I shows the classification and priorities for the contacts. When contact occurs, the cognitive model records contact information such as the contact part of the body, the contact position, the time when the contact will occur, and the priority of the contact. Then, the cognitive model changes the behavior of the virtual human to reflect priorities. The following describes the contact priorities and the workings of the cognitive model.

Table I. The Classification of Contacts

Contact pair (VH, AV)	Recognition	Priority
1. (head, hand)(chest, hand)	AV's attack	high
2. (hand, head)(hand, chest)	VH's attack	high
3. (hand, hand)(arm, hand)	VH blocked	middle
4. (hand, arm)	AV blocked	middle
5. others	other contacts	low

(VH: virtual human, AV: avatar)

High priority: The cognitive model quits prediction and changes behavior.
Middle priority: The cognitive model continues the prediction for a while to test if a high priority contact will occur or not. If a high priority contact does not occur, the cognitive model quits prediction.
Low priority: The cognitive model continues prediction to find a higher priority contact.
In the simulation for prediction, the cognitive model may find more than one contact before ending prediction. The highest and earliest contact decides the behavior of the virtual human.

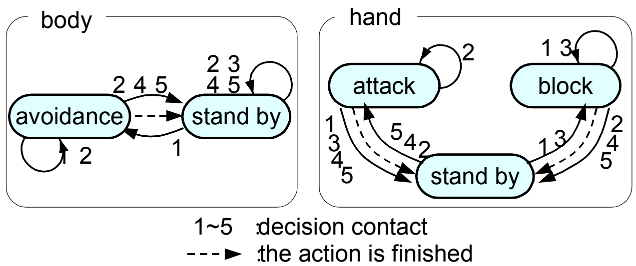

Fig. 5. State transitions for behaviors of the hands and body.

3.9 Selecting Behavior

The cognitive model assigns a behavior for each hand, head, and chest. The body (head and chest) assumes a state of standby and avoidance. The hand assumes a state of standby, attack, and block. The states are changed by a decision on contact (see Section 4.2). Figure 5 shows the status transition. For example, imagine a situation where the simulator for prediction finds a contact between the left hand of the avatar and the head of the virtual human, when the current behaviors of the virtual human are (body: stand by, right hand: attack, left hand: stand by). The behaviors will change to (avoidance, stand by, block). We make one more restriction, i.e., that the virtual human use only one hand for attack at one time.

3.10 Selecting an Action

A virtual human's actions are realized via "reaching" motions, which are defined by a target body part, target position, and duration of the motion. The cognitive model decides the parameters of a "reaching" motion, as shown in Table II.

4.4.1 *Avoidance motion*. To avoid a user's attacks, the virtual human should move the body to a position with enough distance from the trajectory of the user's hand. First, we suppose that the avatar's hand will stay behind the contact position, and then we define the trajectory of the avatar as the segment AP' in Figure 6. The decision contact (see Section 4.2) records the predicted contact point P and duration t. The point A is the center of the avatar's hand, and we define P' by extending the segment AP to a certain length. Next, we find the nearest point Q on the segment AP' from the center of the virtual human's target body part B. Finally, we define the target position R by extending segment QB to a certain length. The duration of the "reaching" motion is set to the duration between the current time and the contact time.

4.4.2 *Block Motion*. To block a user's attacks, the virtual human should move the forearm to a position on the trajectory of the user's hand. The duration of the reaching motion is set for a slightly shorter time than the duration between the current time and the contact time. We define point P' as the center of the avatar's hand at the end of the duration. We find the nearest point on the segment AP' from the center of the virtual human's hand B and set the point R as the target of the reaching motion.

Table II. Actions and their Behaviors

behavior	body part	target position
avoidance	head, chest	A position not reached by avatar's hand.
block	hand, forearm	The position of the avatar's hand.
attack	Hand	The position of avatar's head and chest.

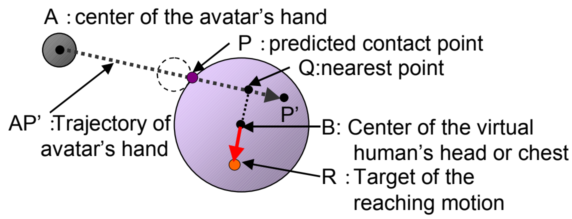

Fig. 6. The avoidance action.

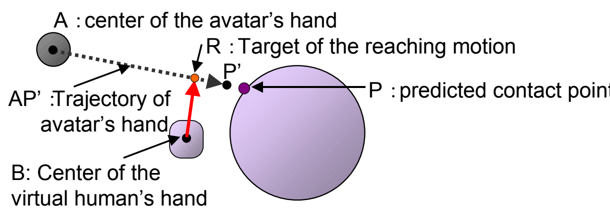

Fig. 7. The block action.

4.4.3 *Attack Motion*. The virtual human selects the target of the attack from the head and chest of the avatar by considering the distance from the virtual human's hand to the target. The target position and the duration are set randomly.

3.11 Controlling Human Body Models

The human body models are fixed to the ground via spring and damper models to keep the position and orientation of their bodies. The human body models have a default posture described in Section 3.2; Figure 8 shows the default posture of the human body models. The avatar's body is controlled by the user as described in Section 3.4. The "reaching" motions of the virtual human are realized via the method described in Section 3.6.

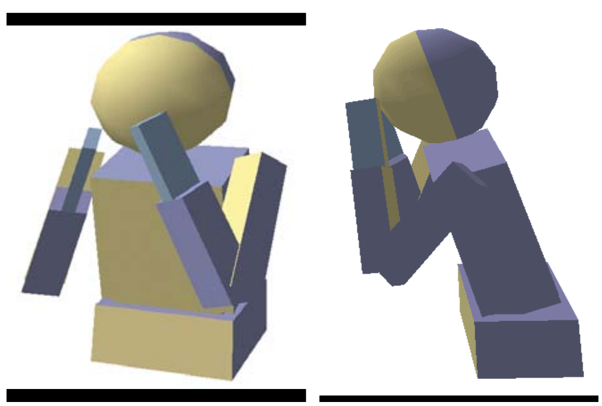

Fig. 8. The default posture of a human body model.

Table III. Processing Time

Simulation for the virtual world	1.5ms
Simulation for prediction	3.1ms
Average of total computation times for a step	1.6ms
Time steps of the simulator for the virtual world	3.0ms

3.12 Force Feed-Backs

The haptic interface feeds back the contact force that acts on the hand of the avatar onto the corresponding hand of the user. In addition, we feed back contact forces that act on other parts of the avatar's body onto the user's hands (with some attenuation) to notify the user of the damage.

4. EVALUATION

We evaluated the proposed framework and the virtual boxing application. We used a PC with a Pentium4 3.2 [GHz] processor for the evaluation.

4.1 Processing Time

Table III shows the processing time for the virtual boxing application. The average total computation time of 1.6ms is smaller than the simulation and update rate time step (3ms). During prediction, the update rate drops to 4.6ms. However, predictions are done in short durations and users are not aware of the delay.

4.2 Test Play

We asked six students to play the virtual boxing application. Figure 9 shows a scene of the interaction; Figure 10 shows the the virtual human's reaction. We cut off the virtual human's cognitive model to create a passive motion, shown on the left side of the figure.

Students reported that the active reactions (within the proposed framework) were more real and attractive than the passive movements. Students also reported that the force feed-backs helped them know about the hits from attacks by both the virtual human and the user.

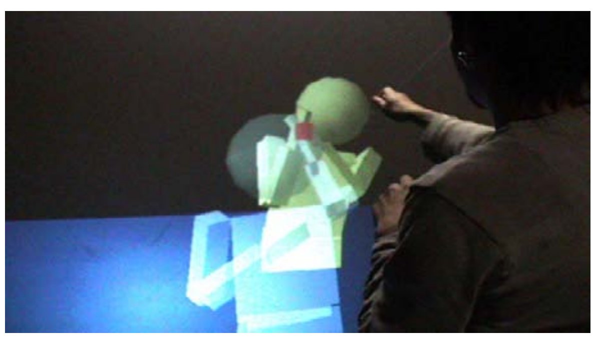

Fig. 9. The interaction.

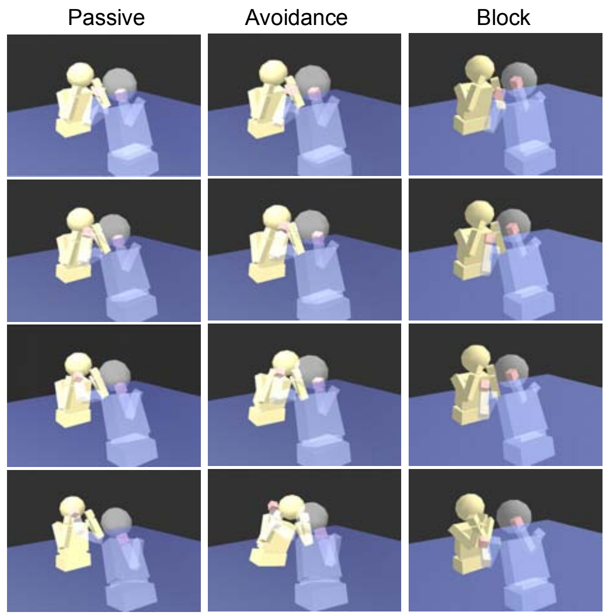

Fig. 10. Reactive motions of the virtual human. (Please access **http://springhead.info/boxing.mpg** for the movie clip.)

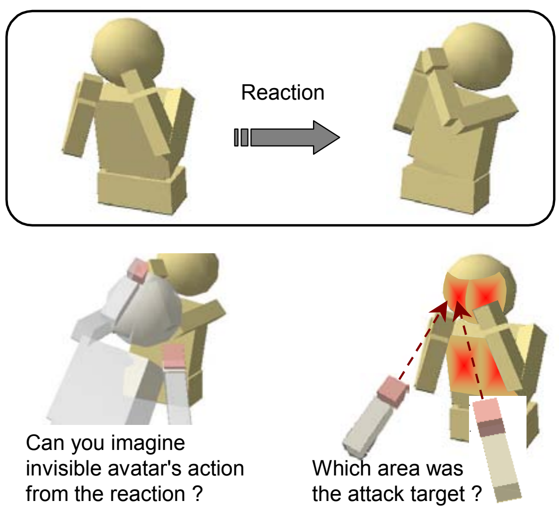

Fig. 11. The experiment's procedure.

4.3 Evaluating the Reactions

In the real world, humans often guess the cause of a reaction from people's reactive movements. Therefore, we expect that the reality of a reactive motion can be evaluated by testing whether the cause of a reaction can be guessed or not. We experimented with this assumption.

Table IV. Results of the Experiment

A. *Answers for the active reactions in the proposed framework.*

correct rate		subject's answer			
		left hand	right hand	left chest	Right chest
Target of the avatar's attack	left hand	0.88	0.08	0.08	0
	right hand	0.20	0.71	0	0.08
	left chest	0.08	0	0.79	0.17
	right chest	0	0	0.13	0.88

B. *Answers for the passive motions without the cognitive model.*

correct rate		subject's answer			
		left hand	right hand	left chest	Right chest
Target of the avatar's attack	left hand	0.63	0.13	0.25	0
	right hand	0.42	0.42	0.08	0.08
	left chest	0.25	0.42	0.17	0.17
	right chest	0.13	0.63	0.13	0.13

C. *Total correct rate.*

	active reaction	passive motion
correct rate	0.80	0.33

Figure 11 shows the procedure for the experiment. Eight subjects were shown the reactions of the virtual human and asked which part of its body was attacked by the avatar. We didn't exhibit the avatar to the subjects. As an answer, the subjects selected one of the four parts of the body. We showed two types of reactions: reactions with the proposed framework and passive movements without the virtual human's cognitive model. We show the reactions in random order. Table IV displays the subjects' answers, which indicate that the subjects predicted the reaction more accurately when it was generated by the proposed framework than via passive motions.

5. CONCLUSION

In this article we proposed a framework for haptic interaction with a reactive virtual human in a physically simulated virtual world. We employed a human-scale projection display with a haptic interface for interaction. The virtual human in the framework has a cognitive model with a dynamics simulator for prediction and state transition machines for selecting behaviors. Realistic "reaching" motions for virtual humans were generated based on the human "reaching" model.

We created a virtual boxing system as an application of the proposed framework. We confirmed that the framework generates realistic reactions and that users can correctly estimate the input motions of the avatar.

Integrating more human sensory and motion models will realize more realistic reactions. Currently, we are integrating models of visual attention and eye-gaze into a virtual human.

APPENDIX

A. GENERATION OF REACHING MOTION

We create a model for a reaching motion that is based on a human motion model proposed by Flash [1987]. His model consists of a minimum jerk model and PD control of joints. The minimum jerk model generates a trajectory that minimizes the total sum of the jerk:

$$\int_0^t \| \dddot{\mathbf{r}} \|^2 dt \rightarrow 0$$

where **r** is the trajectory in the Cartesian coordinates. This optimization has a explicit solution:

$$\int_0^t \| \frac{d\mathbf{r}}{dt^3} \|^2 dt \rightarrow 0$$

$$\mathbf{r}(\mathrm{t}) = \mathbf{r_0} + (\mathbf{r_f} - \mathbf{r_0})(6\tau^5 - 15\tau^4 + 10\tau^3)$$

where $\tau = t/t_d$ for the duration of the motion td. We now have a trajectory of the target for the PD-control of the hand.

Next, the target position r is converted to the joint angles and then the joint torques are calculated by the PD-controller for each joint. The spring and damper coefficient of the PD controllers are set for the stiffness and viscosity of human joints. A dynamics simulator creates the motion of the joints, which follow the target trajectory and pay attention to the dynamics of the hand.

Flash's [1987] model works well for the two-dimensional motions of an arm with two joints. However, this model generates unnatural motions if the elbow starts from a position far from the body and the hand starts from a position near the body. In addition, Fish's model requires converting the position of the hand into joint angles. This is not very easy for redundant arms such as the three-dimensional models of human arms.

Hence, we expand and modify this model for more complex motions. Instead of inverse kinematics, we use a spring and damper model to calculate the joint torques. We put a spring and damper model and two ball joints between the hand and the target of the reaching motion. The spring and damper model is connected to the hand and target via ball joints. We then set the natural length of the spring *l* to

$$l(\mathrm{t}) = l_0(6\tau^5 - 15\tau^4 + 10\tau^3)$$

which minimize the integral of the jerk:

$$\int_0^t \| \frac{dl}{dt^3} \|^2 dt \rightarrow 0$$

$$\int_0^t \| \dddot{l} \|^2 dt \rightarrow 0$$

where $\tau = t/t_d$ is the duration of the motion t_d and l_0 represents the initial length between the hand and the target. The spring and damper models provide the force and move the arm. The dynamics simulator creates the motion of the arm and reflects the dynamics of the hand.

We provide an opposite force to the one at the virtual human's center of gravity to make the forces internal. Converting the force coordinates makes it possible for the sum of the forces to be considered as generated by the joint torques.

REFERENCES

BADLER, N. I., HOLLICK, M. J., AND GRANIERI, J. P. 1993. Real-time control of a virtual human using minimal sensors. *Presence 2*, 1 (1993), 82-86.

BLUMBERG, B., DOWNIE, M., IVANOV, Y., BERLIN, M., JOHNSON, M. P., AND TOMLINSON, B. 2002. Integrated learning for interactive synthetic characters. In *Proceedings of the 2002 SIGGRAPH Conference*. ACM, New York, 417-426.

CAI, Y., ISHII, M., AND SATO, M. 1996. A human interface device for cave: Size virtual workspace. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics* (Oct. 1996), 2084-2089.

FEATHERSTONE, R. 1983. The calculation of robot dynamics using articulated-body inertias. *Int. J. Robotics Research 2*, 1 (1983),13-30.

FLASH, T. 1987.The control of hand equilibrium trajectories in multi-joint arm movements. *Biological Cybernetics 57* (1987), 257-274.

FUNGE, J., TU, X., AND TERZOPOULOS, D. 1999. Motion graphs. In *Proceedings of the 1999 SIGGRAPH Conference*. ACM, New York. HASEGAWA, S. AND SATO, M. 2004. Real-time rigid body simulation for haptic interactions based on contact volume of polygonal objects. In Computer Graphics Forum, Sept.2004.

HODGINS, J. K., WOOTEN, W. L., BROGAN, D. C., AND O 'BRIEN, J. F. 1995. Animating human athletes. In *Proceedings of the1995 SIGGRAPH Conference*. ACM, New York, 71-78.

JEONG, S., HASHIMOTO, N., AND MAKOTO, S. 2004. A novel interaction system with force feedback between real and virtual human. In *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology.*

ACM, New York.

KOMURA, T., SHINAGAWA, Y., AND KUNII, T. L. 2000. Creating and retargeting motion by the musculoskeletal human body model. *The Visual Computer 16* (2000), 254-270.

KOUCHI, M., MOCHIMARU, M., ICHIMURA, R., IWASAWA, H., AND MITANI, S. 1998. Japanese body dimension data. http://www.dh.aist.go.jp/research/bodydimeDB/, 1998.

KOVAR, L., GLEICHER, M., AND PIGHIN, F. 2002. Motion graphs. In *Proceedings of the 2002 SIGGRAPH Conference*. ACM, New York, 473-482.

LEE, J. AND LEE, K. H. 2004. Precomputing avatar behavior from human motion data. In Proceedings of the Symposium on Computer Animation. 2004.

National Institute of Technology and Evaluation. 2004. Human characteristics database. http://www.tech.nite.go.jp/human/, 2004.

OSHITA, M. AND MAKINOUCHI, A. 2001. A dynamic motion control technique for human-like articulated figures. *Computer Graphics Forum 20*, 3 (2001),192-202.

ROSE, C., GUENTER, B., BODENHEIMER, B., AND COHEN, M. F. 1995. Efficient generation of motion transitions using spacetime constraints. In *Proceedings of the 1995 SIGGRAPH Connference*. ACM, New York, 147-154.

SUZUKI, Y., ISHII, S., AND SEGA-AM2. 1993. Virtua fighter. http://en.wikipedia.org/wiki/Virtua_Fighter, http://www.sega.com/gamesite/vf4evo/, 1993.

YAMANE, K. and NAKAMURA, Y. 2003a. Dynamics filter concept and implementation of on-line motion generator for human figures. *IEEE Transactions on Robotics and Automation 19,* 3 (2003), 421-432.

YAMANE, K. and NAKAMURA, Y. 2003b. Natural motion animation through constraining and deconstraining at will. *IEEE Transactions on Visualization and Computer Graphics, 9*, 3 (2003),352-360.

ZORDAN, V. B. AND HODGINS, J. K. 1999. Tracking and modifying upper-body human motion data with dynamic simulation. In *Proceedings of the 1999 Eurographics Workshop on Animation and Simulation*.

Received June 2005; accepted March 2006