Human Scale Haptic Interaction with a Reactive Virtual Human in a Realtime Physics Simulator

Shoichi Hasegawa
Tokyo Institute of Technology / PRESTO JST
4259 Nagatsuta Midori
Yokohama, Japan
hase@hi.pi.titech.ac.jp

Ishikawa Toshiaki
Tokyo Institute of Technology
4259 Nagatsuta Midori
Yokohama, Japan
ishikawa@hi.pi.titech.ac.jp

Naoki Hashimoto
Tokyo Institute of Technology
4259 Nagatsuta Midori
Yokohama, Japan
naoki@hi.pi.titech.ac.jp

ABSTRACT
In this paper, 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 get confirmation that proposed framework creates realistic reactions and users can easily estimate the input motions of the avatar.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interface—*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*

1. INTRODUCTION
Virtual humans are one of the key feature in computer entertainments. Recently, the reality of graphics and motions is remarkably improved in computer entertainments. The appearances of virtual humans come to be real with GPUs and motion capture techniques.

While movies and animations require offline creation of realistic motions, computer games need realtime interactions with virtual humans. Conventional computer games create realistic reactions of virtual humans by choosing and connecting offline-prepared motion data which are appropriate for the user's button inputs. This technique is suitable for limited digital input. However, more direct interaction through analog pads or haptic interfaces requires new paradigm of reaction generation.

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

2. RELATED WORKS
Many virtual humans are employed in computer games such as RPGs, sports games and action games. Fighting games such as Virtua Fighter [17] are pioneer users of virtual humans. In these games, reactions of virtual humans are created by connecting motions in a motion database such as Motion Graph [12]. Therefore, variations of the reactions are limited and it takes large costs to create a motion database. Lee and Lee [13] created realistic motions of boxers with a database approach. However, their method does not create dynamic reaction for contacts.

The advances of the human interface and virtual reality technologies for entertainment computing bring wide variations of user's inputs. Therefore, virtual humans are required to perform many variations of reactions. Seungzoo et al. applied a database approach to create a reactive virtual human for haptic interaction [9]. The interaction in their system was a catch ball. It was not a direct contact interaction because of the restriction of the variation of motions.

Space time constraint methods [16, 10] generate optimal trajectory which minimize some optimization function. However, these methods must solve optimization of the trajectory and it is difficult to run them in realtime.

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 [20, 18], state machines [8] and environmental physical input [15]. These methods generate appropriate motions for variations of the input.

State machines are used to create motions and actions of virtual characters corresponding to the environments [6] and user's instructions [2]. However, these researches did not include direct haptic interactions.

3. PROPOSED FRAMEWORK

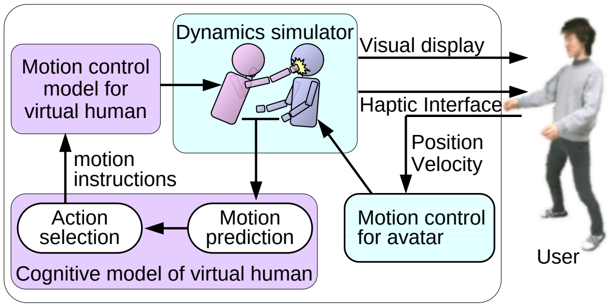

Figure 1: An overview of the proposed framework.

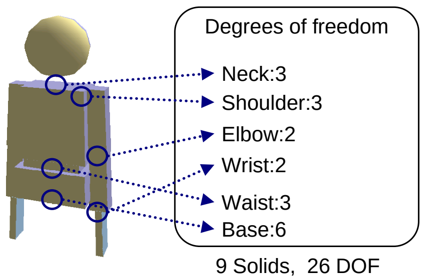

Figure 2: The human body model.

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

The user controls the avatar through a haptic interface. The human body model for the avatar traces the motion of the user, while forces which act on body model for the avatar are feedbacked 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. The motion control model decides the joint torques, which act on the human body model of the virtual human, from the instruction.

3.1 Dynamics Simulator

We employ a dynamics simulator named Springhead [7], which is suitable for haptic interaction. The dynamics simulator, Springhead, employs Featherstone's method [4] for joints and penalty method for contacts and requires less computation time for each iteration.

3.2 Human Body Model

Figure 2 shows the human body model in the proposed framework. In this paper, we model the upper part of the body for virtual boxing application. We 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 referring databases of human characteristics [11, 14]. In addition, we give a default angle and a spring-damper model for each joint to give a default posture of the human body model.

3.3 Haptic Interface and Visual Display

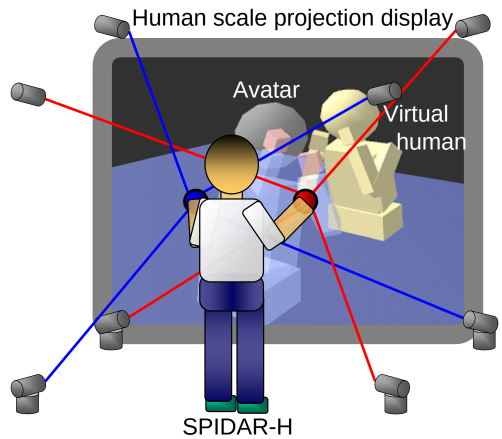

Figure 3: The hardware setup of the proposed framework.

We use a human scale projection display to show realtime images of the virtual world. In addition, a human scale both-hand haptic interface named SPIDAR-H [3] is employed to control the avatar and to feedback the forces on the avatar 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 whole DOF of the avatar from the small input DOF.

Inverse kinematics methods with pseudo-inverse matrix are often used for this purpose [1, 19]. However, these methods do not regard contact forces added from other objects such as opponents of the boxing. Therefore, we use dynamics simulation and PD control. The avatar's hands are pulled by spring and damper model to the measured positions of the user's hands.

Because the human body model of the avatar has joints and each joint has spring and damper model to set the posture to the 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 give softer spring and damper models for the default posture. Therefore, the positions of the avatar's hands almost reach to the positions of the user's hands.

The user should not feel the dynamics of the avatar's hands. Instead, the user should feel the contact forces, which act on the avatar's hands. Therefore, the framework does not feedback the forces from the spring and damper model. The contact forces, which act on the avatar's hands, are feedbacked to the user's hands through the strings of the SPIDAR-H.

3.5 Cognitive Model of the Virtual Human

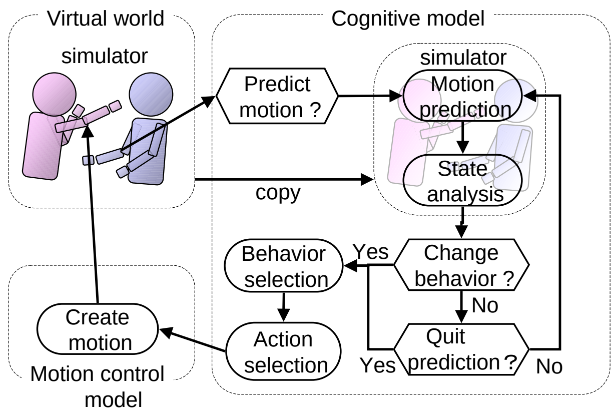

Figure 4: The cognitive model of the virtual human.

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

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

3.6 The Motion Control Model for the Virtual Human

The motion control model creates motor control signal of the virtual human from the three parameters (the part, the duration and the target).

Flash [5] proposes a control model of human reaching motion from measurements of human motion and simulations. The model consists of a minimum jerk model and a PD control of joints. We expand and modify the model and use for the motion control of the virtual human. We describe the detail of our reaching motion model in section A.

4. VIRTUAL BOXING APPLICATION

We create a virtual boxing application as an example of the proposed framework. The application creates reactions of the upper part of the virtual human's body such as attacking, blocking and dodging motion. These motions are created by the reaching motion model.

Following sections explains the implementation of the application.

4.1 Prediction of the Virtual World

The cognitive model of the virtual human 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 is fast enough. We copy the state of the simulator for the virtual world to the simulator for the prediction. Then, we run the simulator for the prediction faster than the simulator for the virtual world. The simulator for the prediction will find contacts between the virtual human and the avatar before they really occur in the virtual world.

Because the simulator for the prediction runs faster than realtime, we have to predict the user's inputs and give it to the simulator. We suppose 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 it does at the beginning of the prediction.

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

4.2 Analysis of the Predicted Virtual World

The cognitive model of the virtual human 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 1 shows the classification and priorities of the contacts. When the 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 reflecting the priority. Followings describe priorities of the contacts and the workings of the cognitive model.

High priority: The cognitive model quits the prediction and changes the behavior.

Middle priority: The cognitive model continues the prediction for a while to test if high priority contact will occur or not. If high priority contact does not occur, the cognitive model quits the prediction.

Low priority: The cognitive model continues the prediction to find higher priority contact.

In the simulation for the prediction, the cognitive model may find more than one contact before the end of the prediction. The highest and earliest contact decides the behavior of the virtual human. We call such a contact *decision contact* in followings.

4.3 Behavior Selection

The cognitive model assigns a behavior for each hand and body (head and chest). The body takes state of standby and avoidance. The hand takes state of standby, attack and block. The states are changed by a *decision contact* in section 4.2. Figure 5 shows the status transition. For example,

Table 1: 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)

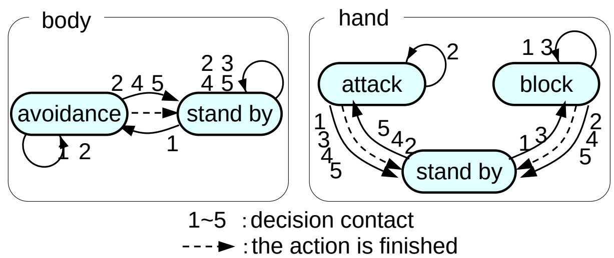

Figure 5: The state transitions of the behaviors of the hands and body.

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 give one more restriction that the virtual human uses only one hand for attack at one time.

4.4 Action Selection

Actions of the virtual human are realized by reaching motions, which are defined by a target body part, target position, and duration of the motion. The cognitive model decides the parameters of reaching motion shown in Table 2.

4.4.1 *Avoidance Motion*

To avoid user's attacks, the virtual human should move the body to a position taking enough distance from the trajectory of the user's hand. First, we suppose that the avatar's hand will proceed behind the contact position. Then, we define the trajectory of the avatar as the segment AP' in Figure 6. The *decision contact* (see section 4.2) records 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 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 certain length. The

Table 2: Actions for the 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 the avatar's head and chest.

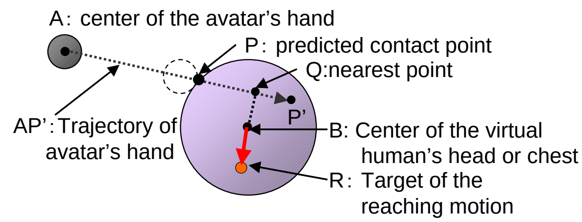

Figure 6: The avoidance action.

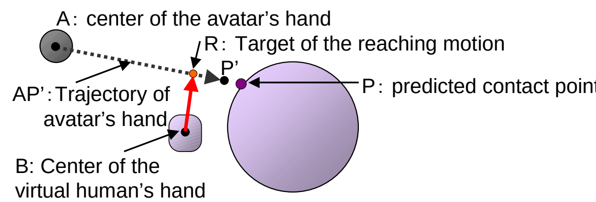

Figure 7: The block action.

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 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 to a little shorter duration than the duration between the current time and the contact time. We define the 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.

4.4.3 *Attack Motion*

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

4.5 Control of the Human Body Models

The human body models are fixed to the ground via spring and damper models to keep the position and orientation of the 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 by the method described in section 3.6.

4.6 Force Feedbacks

The haptic interface feedbacks contact forces which act on the hand of the avatar to the corresponding hand of the user. In addition, we feedback contact forces, which act on other parts of the avatar's body to both hands with some attenuation, to notify damages to the user.

5. EVALUATION

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

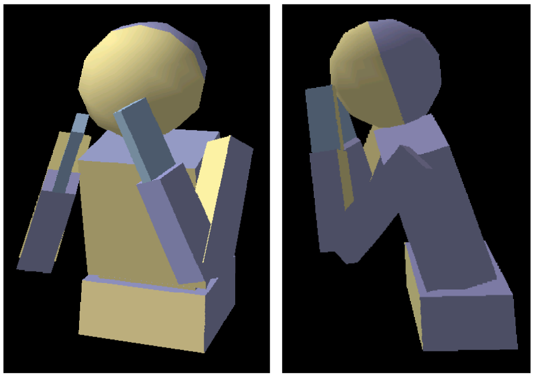

Figure 8: Default posture of a human body model.

Table 3: Processing time

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

5.1 Processing Time

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

5.2 Test Play

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

Students reported that the active reactions with the proposed framework were more real and attractive than the passive motions. Students also reported that the force feedbacks helped to know the hits of the attacks of both the virtual human and the user.

5.3 Evaluation of the Reaction

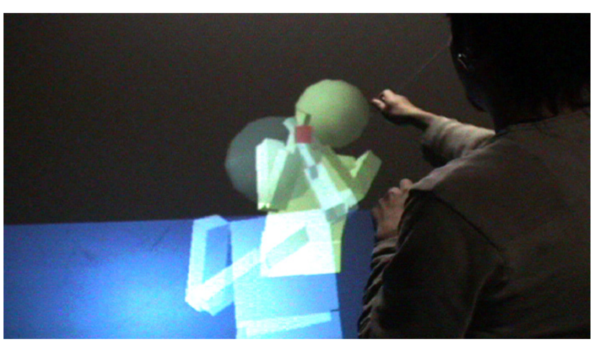

Figure 9: The interaction.

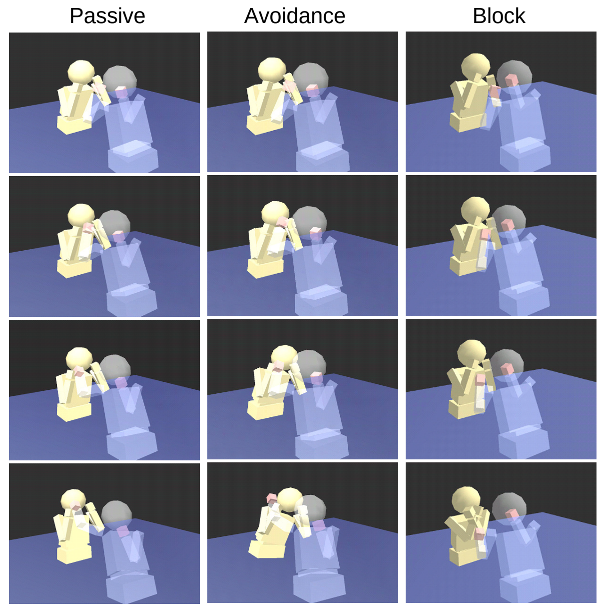

Figure 10: Reactive motions of the virtual human. Please access http://springhead.info/boxing.mpg for the movie clip.

In the real world, humans often guess the cause of a reaction from the reactive motions of people. Therefore, we expect that the reality of the reactive motion can be evaluated by testing if the cause of a reaction can be guessed or not. We made an experiment with this assumption.

Figure 11 shows the procedure of the experiment. Eight subjects were shown the reaction of the virtual human and asked which part of the body was attacked by the avatar. In this experiment we didn't show the avatar to the subjects. The subjects select one of the four parts of the body for the answer. We showed two types of the reaction: reactions with the proposed framework and passive motions without the cognitive model of the virtual human. We show the reactions with a random order. Table 4 shows the answers of the subjects. The answers show that the subjects predicted the reaction more accurately when it was generated by the proposed framework than the passive motions.

6. CONCLUSION

In this paper, we proposed a framework for haptic interaction with a reactive virtual human in a physically simulated virtual world. We employed human scale projection display, haptic interface for the interaction. The virtual human in the framework has a cognitive model, which has a dynamics simulator for prediction and state transition machines for the behavior selections. Realistic reaching motions of virtual humans are generated based on the reaching motion model of humans.

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

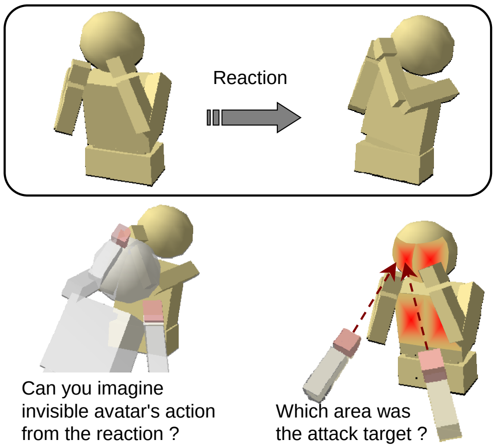

Figure 11: The procedure of the experiment.

Table 4: The result of the experiment.

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

correct rate		subject's answer			
		left head	right head	left chest	right chest
target of the avatar's attack	left head	0.88	0.08	0.08	0
	right head	0.20	0.71	0	0.08
	left chest	0.08	0	0.79	0.13
	right chest	0	0	0.13	0.88

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

correct rate		subject's answer			
		left head	right head	left chest	right chest
target of the avatar's attack	left head	0.63	0.13	0.25	0
	right head	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

7. ADDITIONAL AUTHORS

Additional authors:
Marc Salvati (Tokyo Institue of Technology, email: `salvati_marc@hi.pi.titech.ac.jp`),
Yasuharu Koike (Tokyo Institue of Technology, email: `koike@pi.titech.ac.jp`) and
Makoto Sato (Tokyo Institue of Technology, email: `msato@pi.titech.ac.jp`)

8. REFERENCES

[1] N. I. Badler, M. J. Hollick, and J. P. Granieri. Real-time control of a virtual human using minimal sensors. *Presence*, 2(1):82–86, 1993.

[2] B. Blumberg, M. Downie, Y. Ivanov, M. Berlin, M. P. Johnson, and B. Tomlinson. Integrated learning for interactive synthetic characters. In *Proc. of SIGGRAPH 2002*, pages 417–426, 2002.

[3] Y. Cai, M. Ishii, and M. Sato. A human interface device for cave: size virtual workspace. In *Proc. of IEEE International Conference on Systems, Man, and Cybernetics*, pages 2084–2089, October 1996.

[4] R. Featherstone. The calculation of robot dynamics using articulated-body inertias. *Int. J. Robotics Research*, 2(1):13–30, 1983.

[5] T. Flash. The control of hand equilibrium trajectories in multi-joint arm movements. *Biological Cybernetics*, 57:257–2740, 1987.

[6] J. Funge, X. Tu, and D. Terzopoulos. Motion graphs. In *Proc. of SIGGRAPH 99*, 1999.

[7] S. Hasegawa and M. Sato. Real-time rigid body simulation for haptic interactions based on contact volume of polygonal objects. *Computer Graphics Forum*, September 2004.

[8] J. K. Hodgins, W. L. Wooten, D. C. Brogan, and J. F. O ’Brien. Animating human athletes. In *Proc. of SIGGRAPH 95*, pages 71–78, 1995.

[9] S. Jeong, N. Hashimoto, and S. Makoto. A novel interaction system with force feedback between real - and virtual human. In *Proc. of ACM SIGCHI International Conference on Advances in Computer Entertainment Technology 2004*, 2004.

[10] T. Komura, Y. Shinagawa, and T. L. Kunii. Creating and retargeting motion by the musculoskeletal human body model. *The Visual Computer*, 16:254–270, 2000.

[11] M. Kouchi, M. Mochimaru, R. Ichimura, H. Iwasawa, and S. Mitani. Japanese body dimension data. *http://www.dh.aist.go.jp/research/bodydimeDB/*, 1998.

[12] L. Kovar, M. Gleicher, and F. Pighin. Motion graphs. In *Proc. of SIGGRAPH 2002*, pages 473–482, 2002.

[13] J. Lee and K. H. Lee. Precomputing avatar behavior from human motion data. In *Symposium on Computer Animation*, 2004.

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

[15] M. Oshita and A. Makinouchi. A dynamic motion control technique for human-like articulated figures. *Computer Graphics Forum*, 20(3):192–202, 2001.

[16] C. Rose, B. Guenter, B. Bodenheimer, and M. F. Cohen. Efficient generation of motion transitions using spacetime constraints. In *Proc. of SIGGRAPH 95*, pages 147–154, 1995.

[17] Y. Suzuki, S. Ishii, and Sega-AM2. Virtua fighter. *http://en.wikipedia.org/wiki/Virtua_Fighter, http://www.sega.com/gamesite/vf4evo/*, 1993.

[18] K. Yamane and Y. Nakamura. Dynamics filter - concept and implementation of on-line motion generator for human figures. *IEEE Transactions on Robotics and Automation*, 19(3):421–432, 2003.

[19] K. Yamane and Y. Nakamura. Natural motion animation through constraining and deconstraining at will. *IEEE Transactions on Visualization and Computer Graphics*, 9(3):352–360, 2003.

[20] V. B. Zordan and J. K. Hodgins. Tracking and modifying upper-body human motion data with dynamic simulation. In *Proc. of Eurographics Workshop on Animation and Simulation '99*, 1999.

APPENDIX

A. GENERATION OF REACHING MOTION

We create a model for reaching motion based on a human motion model proposed by Flash [5]. His model consists of minimum jerk model and PD control of joints. The minimum jerk model generates a trajectory which minimize the total sum of the jerk:

$$\int_0^t ||\dddot{\mathbf{r}}||^2 dt \to 0 \tag{1}$$

where $\mathbf{r}$ is the trajectory in Cartesian coordinates. This optimization have a explicit solution:

$$\mathbf{r}(t) = \mathbf{r_0} + (\mathbf{r_f} - \mathbf{r_0})(6\tau^5 - 15\tau^4 + 10\tau^3) \tag{2}$$

where $\tau = t/t_d$ for the duration of the motion t_d. Now, we got a trajectory of the target for the PD-control of the hand.

Next, the target position $\mathbf{r}$ is converted to the joint angles. Then, the joint torques are calculated by the PD-controller for each joint. The spring and damper coefficient of the PD-controllers are set regarding the stiffness and the viscosity of human's joints. A dynamics simulator creates the motion of the joints, which follow the target trajectory and regards the dynamics of the hand.

His model works well for two dimensional motions of an arm with two joints. However, his model generates unnatural motion if the elbow starts from a position far from body and the hand starts from a position near the body. In addition, his model requires to convert the position of the hand into joint angles. It is not very easy for redundant arms such as three dimensional models of human arms.

Therefore, we expand and modify his 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. Then, we set the natural length of the spring l to:

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

which minimize the integral of the jerk:

$$\int_0^t ||\dddot{l}||^2 dt \to 0 \tag{4}$$

where $\tau = t/t_d$ for 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 give force and move the arm. The dynamics simulator creates the motion of the arm and it reflects the dynamics of the hand.

We give an opposite force of the force to the center of the gravity of the virtual human to make the forces an internal force. Converting the coordinates of the forces, the sum of the forces can be considered to be generated by the joint torques.