Reactive Virtual Creatures for Dexterous Physical Interactions

Hironori Mitake[1], Shoichi Hasegawa[1 2], and Makoto Sato[1]

[1] P&I Lab., Tokyo Institute of Technology, Kanagawa, Japan
[2] PRESTO JST, Tokyo, Japan
{mitake, hase, msato}@pi.titech.ac.jp

Abstract. Dexterous physical interactions with virtual creatures are important to bring the fun of playing with animals into arts and entertainment. For reality of interaction, virtual creatures need to react to highly varied user inputs in a variety of ways according to physical and psychological laws. We propose constructing virtual creatures using a physical simulator, sensor/attention models, and physical motion controllers. The physical simulator and motion controllers generate highly varied physically real reactions, while sensor/attention models provide psychologically feasible target selection for motion controllers. Having constructed a virtual creature prototype, we realize communicative physical interactions such as guessing and attracting attention by touching it via a haptic device. We have confirmed the prototype's effectiveness experimentally.

Keywords: Virtual Creature, Attention, Sensorimotor System, Motion Generation, Physical Interaction

1 Introduction

Characters in games play important roles expressing personalities and stories. They interact with players and other characters, act expressively, and evoke players' empathy. Such roles are required for attractive game experiences and are similar to those of creatures (humans/animals) in the real world. Therefore, we call such characters **virtual creatures**.

Improving enjoyment of games is possible by introducing touch interaction with the virtual creatures as if they were living creatures. Because recent game devices using motion trackers or touch panels allow physical interaction between virtual worlds, there are increased expectations of real-time reaction generation methods with physical effects. Several games [9, 13] already implement physical interaction between animals, albeit with reactions occurring in limited prepared ways.

Touching real creatures gives us a sense of life. Dexterous interaction by exchanging force between the parties and reacting with various expressive motions

http://haselab.net/~mitake/en/research/VirtualCreature.html showing situational thinking (e.g., paying attention by moving their gaze) is possible. Reactions showing what is on their mind create empathy and enable us to deal with a creature's mentality.

However, dexterous physical interactions and empathetic expressions are difficult to combine. Expressive motions should be created to satisfy the nature of the human mind. Currently, this requires intervention by animators. However, dexterous physical interactions require that virtual creatures respond with highly varied reactions according to continuous player input.

This research aims to implement a method to construct virtual creatures for dexterous physical interactions with lifelike expressive reactions that depend on the interaction.

2 Related Works

Motion generation methods for characters have been widely researched to realize physical interaction. Physically simulated articulated body models are often employed [11, 2, 5], while prepared motion databases have been included in recent studies [15, 14].

To ensure that characters' expressive behaviors happen automatically, eye-movement generation to display attention has also been studied. A psychological model for human attention [10, 7] or a statistical model [6] is typically used for offline animation generation or dialogue interaction.

Psychological models are more commonly used in research on cognitive robots [3, 8]. To make human-shaped robots that behave naturally and physically in the real world, sensor devices, information processing, and motor controls are constructed to mimic the psychological function of humans or animals. However, cognitive processes such as finding objects in real-world sensor data, are complicated and real robots are limited by their driving mechanism.

Our approach is a combination of real-time articulated body simulation and psychological models. Full physical simulation based motion generation ensures a variety of reactions in dexterous physical simulation. In addition, cognitive models to emulate higher functions of human mind are much simpler in virtual world, because they can obtain cognitive information directly from the simulator as the virtual world itself.

3 Approach and Proposal

We focus on several aspects of the nature of real creatures.

Sense and action cycle is a common characteristic of creatures. Real creatures use sensory organs to understand the environment and then decide what action to take. When we observe the creatures' surroundings and their reactions, we can guess what they have sensed and thought in choosing the action. Reacting without sensing appears unnatural. Limitations in the range of sense (e.g., field of view) necessitate paying attention to important targets. Motion of sensory organs in the direction of the most important target highlights the focus.

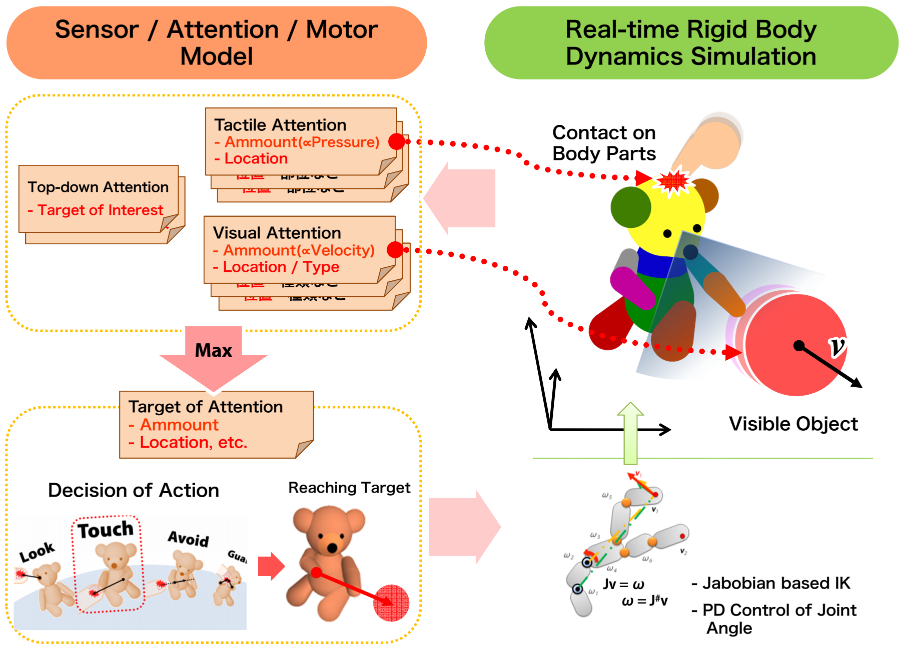

Fig. 1. Overview of the implementation of the proposed virtual creature system

Therefore, we propose constructing virtual creatures as follows. First, the system should be based on sensor models similar to real creatures to make decisions regarding actions. Second, they should also express the focus of their attention and interest in other virtual creatures using motion. Finally, these movements should be generated using a physical simulator to ensure variety of reactions.

4 Realization

In this section, we give an example of the realization of a virtual creature from the proposal. The system consists of 4 parts: a real-time physical simulator, physical motion controllers, sensor/attention models, and character AI. We describe how the proposal should be applied in the virtual creature system.

Fig. 1 shows an overview of the example system, while Fig. 2 gives a definition of the symbols used in following sections.

4.1 Physical Simulator

The virtual creature's motion is generated using a physical simulator and a simulation model of the virtual creature's body. The physical simulator generates various physically realistic motions capable of physical interaction.

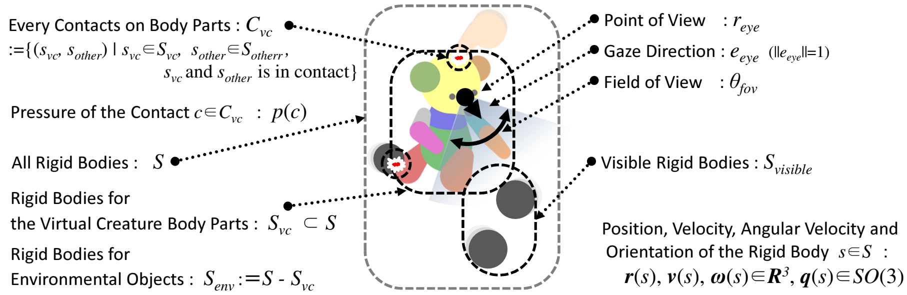

Fig. 2. Definition of symbols

We employed a real-time rigid-body dynamics simulator. Such a simulator calculates translational/rotational movements of a rigid-body object and collisions/friction between multiple objects. Virtual creature is modeled as an articulated body model, constructed with several rigid-bodies and joints. Each joint of the model is actuated by joint torque control. Not only virtual creature, but also interacting environmental objects including players' hands, are modeled as rigid-body/joint models.

The simulator calculates effect on motions or poses by forces applied to virtual creature's body parts from player's hands and fingers. This is the basis of dexterous physical interaction. The simulator also calculates forces applied to player's virtual hands from the virtual creature according to the interaction. The player can feel accurate force feedback if a haptic device is available. This enables certain interactions such as feeling movement of the virtual creature from the force feedback, or disrupting the character's movement by pushing moving body part of the character. In addition, the simulator automatically generates physically realistic action by the virtual creature depending on the situation. For example, speed of motion or position of the center of gravity will change in a carrying motion if the weight of the object changes.

4.2 Physical Motion Controller

Several motion controllers actuate physically simulated articulated body model by controlling torque of each joint. Each controller is prepared to implement particular type of motion (e.g., gaze motion or hand reaching motion).

Nevertheless, each controller must generate variety of motions. For example, changing the creature's gaze implies different motion according to the direction of the target. Generally, each controller must actuate the body to realize particular motion for any selected target. We realized such motion controller with 3 steps: calculation of target position, inverse kinematics, PD-control of the joint angle.

First, rigid body object comprising the virtual creature $s_{vc} \in S_{vc}$ (Fig. 2) is selected as an 'end effector' depending on the type of motion (e.g., a hand should

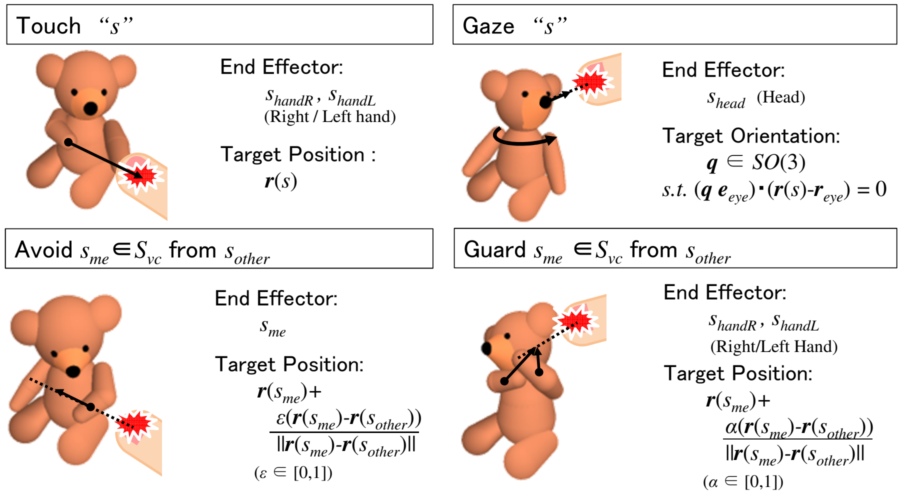

Fig. 3. Physical motion controllers

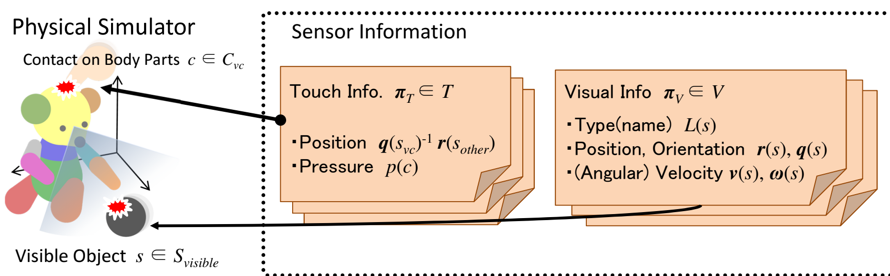

Fig. 4. Sensor information

be end effector of reaching motion). Then, target location for the end effector is calculated according to the location, orientation, and motion of the target.

Several types of motions can be done at the same time, such as hand reaching while looking at the target. We employed inverse kinematics (IK) based on a pseudo inverse of the Jacobian. The IK calculates target angular velocity of each joint to make each end effector approximate target position. Then, torque of each joint is determined with PD-control to realize calculated angular velocity.

In this implementation, we constructed 4 types of motion controllers. Fig. 3 shows details of each controller.

4.3 Sensor Models

Sensor model mimics a real creature's senses, thereby contributing to realizing lifelike actions such as reacting to objects entering the field of view. We focused particularly on visual and touch senses, because vision gives the greatest amount of information on the environment for most real creatures, while physical interaction including direct touch to the virtual creatures causes them to react to tactile sensation. This sense information is recreated from the information on the physical simulator.

In our implementation, visual sense is defined as information on type, location and movement of the object. Since minimum unit of an object in the rigid-body simulator is a single rigid body, the visual information is calculated for each rigid body object in the virtual creature's field of view.

As for touch sense, location and pressure of contact on virtual creature's body surface should be sensed, because players can touch each part of the virtual creature's body in various ways through physical interaction. Contact and contact forces of rigid bodies are calculated by the simulator for every pair of rigid bodies. Therefore, tactile information is created for contact on each rigid body comprising the virtual creature's body.

Definition of visual/tactile information is shown in Fig. 4 and given by Eqs. (1) and (2). Each symbol is defined in Fig. 2.

$$V := \{(s, \boldsymbol{r}(s), \boldsymbol{q}(s), \boldsymbol{v}(s), \boldsymbol{\omega}(s), L(s)) | s \in S_{visible}\} \tag{1}$$

$$T := \{(c, p(c), \boldsymbol{q}(s_{vc})^{-1}\boldsymbol{r}(s_{other})) | c := (s_{vc}, s_{other}) \in C_{vc}\} \tag{2}$$

Here, $S_{visible}$ is a rigid body in the field of view, defined as Eq. (3).

$$S_{visible} := \{s \in S_{env} | \left(\frac{\boldsymbol{r}(s) - \boldsymbol{r}_{eye}}{\|\boldsymbol{r}(s) - \boldsymbol{r}_{eye}\|} \cdot \boldsymbol{e}_{eye} \right) \geq \cos(\theta_{fov}/2)\} \tag{3}$$

$L(s)$ is the type of the object (e.g., "right hand", "head", "apple"). The third element of tactile information $\boldsymbol{q}(s_{vc})^{-1}\boldsymbol{r}(s_{other})$ is the relative position of the touching object in the body centered coordinates.

4.4 Attention Model

Selective attention is a common psychological mechanism for higher animals. The mechanism causes motion of sensor organs towards the object with highest priority. This motion, typified by eye movements, reflects which target has the greater interest. Furthermore, this behavior is a strong clue in guessing creature's intention for us as the observer.

In our implementation, we mimicked the mechanism of selective attention. Real animals have 2 types of attention mechanisms: bottom-up attention caused by salient sensation stimuli and top-down attention, which is intentional. To reproduce these, we implemented attention model as follows. First, amount of bottom-up attention $A_{bot}(\pi)$ is calculated for each piece of sense information π. Next, top-down attention $A_{top}(\pi)$ is calculated according to the decision of character AI, which is a mechanism for making decisions. Finally, sensor information with maximum amount of attention for bottom-up/top-down attention is selected as focus of attention.

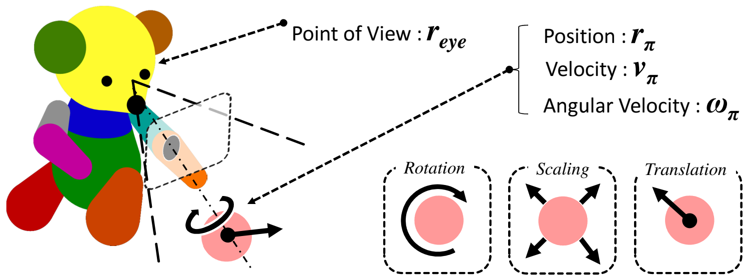

Fig. 5. Bottom-up visual attention for motion

The model is described by Eq. (4), which selects one aspect of the sensor information $\pi_{attention}$ as the focus of attention.

$$\pi_{attention} := \arg\max_{\pi}(A_{bot}(\pi) + A_{top}(\pi)) \tag{4}$$

Here,

$$\pi(\in V \cup T) = \begin{cases} (s_\pi, \boldsymbol{r}_\pi, \boldsymbol{q}_\pi, \boldsymbol{v}_\pi, \boldsymbol{\omega}_\pi, L_\pi) \\ \qquad\qquad\qquad if\ \pi \in V \\ (c_\pi, p_\pi, \boldsymbol{r}_\pi) \\ \qquad if\ \pi \in T(c_\pi = (s_{vc_\pi}, s_{other_\pi})) \end{cases}$$

Calculating amount of bottom-up attention is described in detail below. As for visual attention, bottom-up attention for motion of visual stimuli is important, because motion of the player's hands and objects is important in physical interaction. In our model, amount of visual attention corresponds to motion in a retinal image of the virtual creature. In the retinal image, a moving rigid body causes 3 types of optical flow: translational, rotational, and scaling flow. Therefore, the amount of visual attention is defined by Eq. (5). According to this definition, objects moving faster attract greater attention.

$$A_{visual}(\pi) := k_T\|\boldsymbol{v}_\pi - \boldsymbol{d}_\pi(\boldsymbol{d}_\pi \cdot \boldsymbol{v}_\pi)\| + k_S\|\boldsymbol{d}_\pi \cdot \boldsymbol{v}_\pi\| + k_R\|\boldsymbol{d}_\pi \cdot \boldsymbol{\omega}_\pi\| \tag{5}$$

Here,

$$\boldsymbol{d}_\pi = \frac{\boldsymbol{r}_\pi - \boldsymbol{r}_{eye}}{\|\boldsymbol{r}_\pi - \boldsymbol{r}_{eye}\|}$$

Also, k_T, k_S, k_R are parameters to adjust each type of optical flow to affect the amount of attention.

Tactile attention is defined similarly to visual attention. The amount of tactile attention corresponds to contact pressure for each contact on the rigid bodies comprising the virtual creature (Eq. (6)). Since the definition is based on contact pressure, greater attention follows when touched with sharp edge, and not just when touched with a strong force.

$$A_{touch}(\pi) := k_P p_\pi \tag{6}$$

Here, k_P is a parameter to adjust the correspondence between the pressure and the amount of attention.

Then, sum of visual/tactile attention gives the total amount of bottom-up attention (Eq. (7)).

$$A_{bot}(\pi) := A_{visual}(\pi) + A_{touch}(\pi) \tag{7}$$

With the attention model, virtual creatures can behave in such a way as to express paying attention. As the definition shows, when different targets attract bottom-up/top-down attention, attention conflicts may occur. In such situations, the virtual creature acts as if it is trying to concentrate on target of interest, but being distracted by any other moving objects in sight.

The definition includes various parameters, k_T, k_S, k_R, k_P. These parameters or amounts of top-down attention for each object can be manually adjusted to change behavior of the virtual creature. The parameters should be tuned to realize expected behavior that is derived from settings of the virtual creature. It is easy to guess effects of the parameters, and adjustment can be made smoothly by observing resulting behavior.

4.5 Character AI

Character AI is a conventional mechanism to make current characters behave according to the situation, such as if-then rules or state machines. Our implementation employs character AI to make action decisions.

The character AI for the proposed method needs to make decisions according to the nature of real creatures, and is required to be used in combination with sensor/attention models and physical motion controller. To satisfy this, conditions of rules or state machines are based on the sensor information. This prevents virtual creatures from making unnatural decisions using a super sense. Then, character AI determines amount of top-down attention for each piece of sense information and selects the relevant motion controller to be used. In this way, virtual creature acts and pays attention as determined by the AI.

Our implementation uses a simple character AI with the following 2 rules, which can be activated at the same time, in which case, 2 motion controllers would run simultaneously. t_{look}, t_{touch} are thresholds for the amount of attention to activate the rules and should be adjusted manually.

- Rule 1 :
  if $A_{bot}(\pi_{attention}) + A_{top}(\pi_{attention}) > t_{look}$
  use controller Eye Movement
- Rule 2 :
  if $A_{top}(\pi_{attention}) > t_{touch}$
  use controller Reaching Movement (Hand)

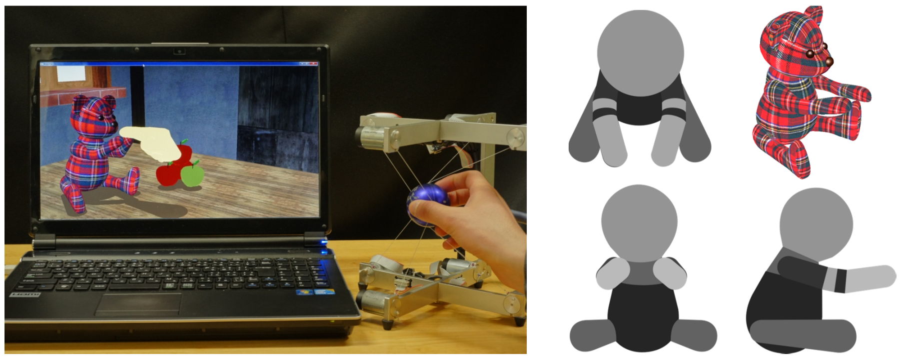

Fig. 6. Experimental environment and the rigid body model of a virtual creature

Table 1. Parameter settings for the experiment

Paramter	Value	Parameter	Value
$L(s_{pointer})$	“pointer”	$A_{top}(s_{pointer})$	0.5
$L(s_{greenapple})$	“greenapple”	$A_{top}(s_{greenapple})$	0.0
$L(s_{redapple})$	“redapple”	$A_{top}(s_{redapple})$	1.0
k_T	0.03	t_{look}	0.05
k_S	0.03	t_{touch}	0.2
k_R	0.03	k_p	2.0

5 Evaluation

We evaluated effectiveness of the method in realizing various physical interactions between the virtual creatures with lifelike expressions of attention. Experiments were carried out with the implemented virtual creature system described in Section 4.

5.1 Environment and Configuration of Virtual Creatures

We set up an environment to enable physical interaction with a virtual creature using a 6-DOF haptic device SPIDAR [12] (Fig. 6). Users can touch and drag both the virtual creature and environmental objects around the virtual creature.

Springhead2 [4], a real-time rigid-body dynamics simulator that runs on commonly used PCs, is employed as the physical simulator. The virtual world in the physical simulator contains the virtual creature and several different objects (apples) as targets for the attention of the virtual creature.

The implemented articulated body model of the virtual creature is shown in Fig. 6. The model has 17 rigid bodies, 16 joints, and 44 DOFs. Parameter

Major Division	Action of Player	Reaction
Trying to attract attention while the virtual creature gazing at an apple	Shake hands in front of the virtual creature	Look at the hand when it moves
		Still looking the apple
	Push the virtual creature	Look at the hand when it touches
	Take the apple which is a gaze target	Chase apples with eyes and try to catch
	Take the apple which is not a gaze target	Change gaze target to another apple
		Look both apple alternately
		No change in gaze target
	Throw down apple from out of sight	Gaze and reach out a hand
	Push with an apple	Look at the apple when it touches
Trying to make various reaction by touching the virtual creature	Push the virtual creature	Arm bends
		Upper body bends
		Neck bends
	Try to put the apple on the head	Catch apples with hands and a head
	Approach the virtual creature by hand	Reach hands to player's hand

Fig. 7. Patterns and examples of resulting interaction

settings for the virtual creature are given in Table 1. These values were adjusted manually to achieve appropriate behavior in the constructed environment.

5.2 Experiments and Results

In the experiments, subjects were required to interact freely with a virtual creature for 10 minutes. All interactions were recorded on video. The virtual creature's top-down attention for the apple was increased every minute, in an attempt to change the subject's experience with the virtual creature in terms of a different top-down/bottom-up ratio. On conclusion of the experiments, we extracted patterns of interaction from the recorded video images. The subjects were 7 male adults.

The results are shown in Fig. 7. Sixteen patterns were extracted, roughly divided into 2 types: trying to invoke various reactions from the virtual creature, and trying to attract the attention of the virtual creature. Actually, reactions all differ in terms of the detail even for reactions classified as same pattern.

5.3 Exhibition

We also demonstrated the virtual creature as an interactive exhibit [1]. Over 1000 people participated, and the attendees' interactions and voices were recorded. After the exhibition, we analyzed the recorded video.

Three suggestions came out of the analysis. First, participants seemed to recognize the virtual creature's attention and interests. Several attendees tried to find out what the virtual creature was interested in, by placing apples one by one in front of the virtual creature. Second, the virtual creature seemed to be regarded as a living creature. Most of the attendees tried to feed the virtual creature an apple. Third, certain words spoken by the attendees included observation of aspects of the virtual creature's emotion. When the virtual creature was pushed strongly by a participant and the virtual creature pushed back, some comments were that the virtual creature was scared of the attendee. This suggests that attendees empathized with the virtual creature.

5.4 Discussion

In the experiments, the virtual creature reacted continuously according to the position and motion of players' hands, showing that it was paying attention to the hands or apples. The virtual creature also reacted to force applied from the haptic interface. Players could disrupt the movement of each of the body parts of the virtual creature through the haptic device, and could feel movements through accurate force feedback. In this way, dexterous physical interaction with the virtual creature reacting in various lifelike motions was realized.

The virtual creature gazed at any object suddenly entering into sight, and tried to take an apple, which was a target of interest. The virtual creature's gaze was easily disturbed by shaking a hand in front of it or pushing it strongly when it was looking at an apple with only a small amount of interest. On the other hand, it was hard to deflect the creature's gaze when directed at an apple with high interest. This behavior is similar to that of a real creature.

6 Conclusion

We proposed the construction of a virtual creature using a physical simulator and sensor/attention models. The prototype implementation shows the effectiveness of the approach experimentally.

The proposed method can be used in combination with current action determination systems for game characters. The method also has high affinity to recent motion input devices for games, or open-world games with physical simulators. Using our methods, virtual creatures can behave with various empathizing actions depending on the situation or interaction.

Sensor/attention models and motion controllers form the basis of the virtual creature system to enable creatures to behave naturally and autonomously. The next step will be the reproduction of higher functions of real creatures. In particular, memory and prediction are basic functions that determine actions. These are also important for attention models, in that memorized or expected objects attract less attention than unknown and unexpected ones.

In the future, using such virtual creature systems, designers will be able to describe higher level instruction for character AI, and virtual creatures will act automatically in indicated ways.

Acknowledgement

This work was partially supported by the International Communications Foundation Research Promotion Program.

References

1. Interactive tokyo 2010 (2010), `http://interactivetokyo.jp/2010/`
2. Abe, Y., Liu, C.K., Popović, Z.: Momentum-based parameterization of dynamic character motion. In: SCA '04: Proc. of the 2004 ACM SIGGRAPH/Eurographics symp. on Computer animation. pp. 173–182. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland (2004)
3. Brooks, R.A., Stein, L.A.: Building brains for bodies. Auton. Robots 1(1), 7–25 (1994)
4. Hasegawa, S., Sato, M.: Real-time Rigid Body Simulation for Haptic Interactions Based on Contact Volume of Polygonal Objects. Computer Graphics Forum 23(3), 529–538 (2004), `http://springhead.info/`
5. Hasegawa, S., Toshiaki, I., Hashimoto, N.: Human scale haptic interaction with a reactive virtual human in a realtime physics simulator. In: ACE '05: Proc. of the 2005 ACM SIGCHI Intl. Conf. on Advances in computer entertainment technology. pp. 149–155. ACM, New York, NY, USA (2005)
6. Lee S. P., Badler J. B., Badler N. I.: Eyes alive. ACM Trans. Graph. 21(3), 637–644 (July 2002)
7. L.Itti, Dhavale N., Pighin F.: Realistic avatar eye and head animation using a neurobiological model of visual attention. In: Proc. SPIE Intl. Symp. onOptical Science and Technology. vol. 5200, pp. 64–78 (Aug 2003)
8. Mitsunaga, N., Miyashita, T., Yoshikawa, Y., Ishiguro, H., Kogure, K., Hagita, N.: Robovie-iv : An every day communication robot. Technical report of IEICE. PRMU 105(534), 47–52 (2006), `http://ci.nii.ac.jp/naid/10017302543/en/`
9. NINTENDO: nintendogs (2005), `http://www.nintendo.co.jp/ds/adgj/`
10. Peters C., Sullivan C. O.: Bottom-up visual attention for virtual human animation. In: Proc. of Computer Animation for Social Agents (2003)
11. Rose, C., Guenter, B., Bodenheimer, B., Cohen, M.F.: Efficient generation of motion transitions using spacetime constraints. In: SIGGRAPH '96: Proc. of the 23rd annual conf. on Computer graphics and interactive techniques. pp. 147–154. ACM Press, New York, NY, USA (1996)
12. Sato M., Hirata Y., Kawarada H.: Space interface device for artificial reality. Journal of Robotics and Mechatronics 9(3), 177–184 (1997)
13. SCEE: Eyepet (2009), `http://www.eyepet.com/`
14. da Silva, M., Abe, Y., Popović, J.: Interactive simulation of stylized human locomotion. ACM Trans. Graph. 27(3), 1–10 (2008)
15. Zordan, V.B., Majkowska, A., Chiu, B., Fast, M.: Dynamic response for motion capture animation. ACM Trans. Graph. 24(3), 697–701 (2005)