

A Robot-based Learning Companion for Storytelling

Yi-Zeng HSIEH^a, Mu-Chun SU^{a*}, Sherry Y. CHEN^b, Gow-Dong CHEN^a and Shih-Chieh LIN^a

^a *Department of Computer Science & Information Engineering, National Central University, Taiwan, R.O.C.*

^b *Graduate Institute of Network Learning Technology, National Central University, Taiwan, R.O.C.*

*muchun@csie.ncu.edu.tw

Abstract: Storytelling is a joyful and educational activity for most of children. Storytelling has powerful effects on improving children's skills in language expression (e.g., listening, reading, and comprehension). In this paper, we present a robot-based learning companion to accompany children for storytelling. We hope this system can promote children's interest at storytelling and improve their engagement.

Keyword: robot, learning companion, storytelling, motion capture.

Keywords: Storytelling, Learning companion, Image matching.

Introduction

Over the last two decades, human motion capture systems have been gaining more and more interest from many different research disciplines because they have many potential applications. The automatic analysis of human motion is a very challenging research topic due to its inherent complexity (e.g., body self-occluding problem, diverse body poses). Basically, there are two approaches to implement a human motion capture system: 1) the marker-based approach and 2) the computer vision-based approach. While the marker-based approach requires that a certain amount of rigid markers are attached to the user's body, the computer vision-based approach provides a non-invasive and friendly solution which makes the approach more appealing than the marker-based approach. The potential applications of computer vision-based capture systems can be roughly grouped into three main application areas: surveillance, control, and analysis. Surveillance applications try to provide solutions to the monitoring and tracking of one or multiple subjects at a special location (e.g., people counting, people tracking). Control applications are used to provide human computer interfaces to control something (e.g., games, animation in virtual world). As for analysis applications, they involve in the detailed analysis of the motion data and diagnostics of subjects' performance (e.g., athletes' performance). Some recent review literatures on these related topics can be found in [1]-[7].

In this paper, we present a playful human-computer interaction system which is based on a computer motion capture system. We try to implement a robot-based learning companion to accompany children for storytelling. Storytelling is a joyful and educational activity for most of children. Storytelling has powerful effects on improving children's skills in language expression (e.g., listening, reading, and comprehension). In addition, storytelling can foster children's creativity and encourage social interactions [8]. Via this robot-based learning companion, children can tell a story in front of a robot-based learning companion and then they act as listeners to watch the robot acts out the story just told by them. We hope

this robot-based learning companion can promote children's interest at storytelling and improve their engagement.

The remaining of this paper is organized as follows. Section II will briefly review some related work. The proposed robot-based learning companion is introduced in Section III. The conclusion of this paper is given in Section IV.

1. Related Work

Storytelling is essential for children to develop their language expression, logic thinking, imagination, and creativity [9]. There are several existing storytelling systems which adopt emerging technologies for supporting children to enjoy the activities of storytelling [10]-[22]. MIT's SAGE (Storytelling Agent Generation Environment) which uses a stuffed rabbit provided a storytelling tool for supporting children develop interactive stories [10]. StoryMat developed by Ryokai and Cassell stores children's storytelling play by recording their play stories and movements of stuffed animals they play with [11]. These stories are then played as animations on the mat when another child tell the same story at the same place. PETS (Personal Electronic Teller of Stories) is a robotic pet built with LEGO bricks that could support children in the storytelling process [12]. After children have snapped together special animal parts (e.g., excited, sad) to build a robot, they can give their robot emotions (e.g., excited, sad) to act out throughout their story. Plaisant et al. developed a prototype storytelling robot for use with children in rehabilitation [13]. Children can remotely control a large furry robot by using a variety of body sensors attached to children's bodies. The story told by children can be "played" by the remote robot. TellTale is a toy made of body pieces of a plastic caterpillar [14]. Each body piece can record a piece of a story. The children can put the pieces together and hear the story organized in the order the child has chosen. Sam (an embodied conversational storyteller) was an attempt to have technology play a social role in supporting young children's literacy learning [15]. Dolltalk composed of a platform with a variety of hardware (e.g., tag sensor, accelerometers) allows children to record their stories and hears their voices with the same content but with a different voice [16]. KidPad is a collaborative storytelling tool that supports children to author physical storytelling experiences to share with other children [17]. StoryRoom is a room-sized immersive storytelling environment where it supports children as storytellers from the very start of their experiences [18]. FearNot! is an application of virtual drama to anti-bullying education [19]. Mutlu et al. conducted an experiment to assess the efficacy of the gaze effect as a storytelling robot (Honda's ASIMO) recited a Japanese fairy tale [20]. The aim of the storytelling Alice which is a programming environment is to make the process of learning to program more motivating for middle school girls [21]. A 3D story authoring system for children was introduced to allow children to create roles, design roles' actions, makes and share their own stories [22]. Each of the aforementioned systems has its own considerations and application scenarios.

While some of those systems are purely programming environments, some of them truly adopt embodied companions (e.g., a stuffed rabbit in SAGE, two dolls in Dolltalk, a plastic caterpillar in TellTable, a large furry robot in [13]). Embodied companions to some extent can enrich children's storytelling engagement and experiences.

In our system, a tiny furry robot bear is adopted to mimic the actions of a child when he or she is telling a story. To teach the robot bear to tell stories, the child is not required to attach a certain amount of sensors on his or her body. All the child needs to do is to tell a story in front of a low-cost Web camera. Then the robot bear will mimic how the child tells the story. Compared to the storytelling robot in [13], our system is more user-friendly. The comparisons between those systems and our system are presented in Table I.

Table I. The use of embodied companions.

	Sensors	Physical Objects	Interactive Modes
SAGE[10]	Circuitry Resistor Sensors	Stuffed Rabbit	Converse, Listen & Tell
StoryMat [11]	RF Sensors	Stuffed Bunny	Listen, Tell & Projected Movies
PETS [12] [13]	Light & Touch Sensors	Robotic Pet	Listen, Tell, Sapping, Emotions & Behaviors
TellTale [14]	No	Caterpillar	Listen, Tell & Recombination
Sam [15]	Tags, Motion Detector & Audio Threshold Sensor	Toy Castle & Virtual Avatar	Listen & Tell
Dolltalk [16]	Tags & Accelerometers	Dolls	Listen, Tell in altered voices & Gestures
KidPad [17]	No	Computer	Drawing Tool & Collaboration
StoryRoom [18]	Tags, Touch, Heat, & Light sensors	Augmented Objects	Authoring, Theatrical Experience & Visualization
FearNot! [19]	No	Computer	Graphics
Storytelling robot[20]	No	ASIMO Robot	Gaze behaviors
Storytelling Alice[21]	No	Computer	Programming & Animations
3D story authoring system[22]	No	Computer	Pen, Speech & 3D Graphics
Our System	Web camera	Robotic Bear	Listen, Tell

			in altered voices & Imitating Teller's behaviors
--	--	--	--

2. The robot-based learning companion system

Our robot-based learning companion called the RobotTell system is composed of a furry robotic bear, a low-cost Web camera, and a notebook as shown in Fig. 1.



Fig.1. The RobotTell System which is composed of a Web camera, robotic bear, and a notebook.

In order to make the robotic bear be able to mimic the storyteller, the RobotTell system must have a simple but effective human pose analysis algorithm to extract meaningful pose information from the storyteller. Based on the extracted pose information, the robotic bear can mimic the storyteller to tell the story which was just told by the storyteller. Based on the computational efficiency consideration, we decided to adopt a template-matching scheme as the core unit of the automatic human pose analysis algorithm. The basic idea of the template-matching scheme is as follows. We assume that the poses of the storyteller can be composed of a set of basic poses. Each basic human pose corresponds to a robot structure poses (i.e., a set of specific joint angles). Each captured image is matched with the set of basic poses to find the most matched basic pose. Then the whole sequence of images can be transformed to be a numerical sequence. Based on the numerical sequence, the robotic bear can sequentially stretch its joint angles to corresponding robot structure poses.

The stages of the human pose analysis algorithm on individual image frames over the image sequence recorded from the storyteller are described below. The flowchart of the human pose analysis algorithm (shown in Fig. 2) involves 6 steps as follows:

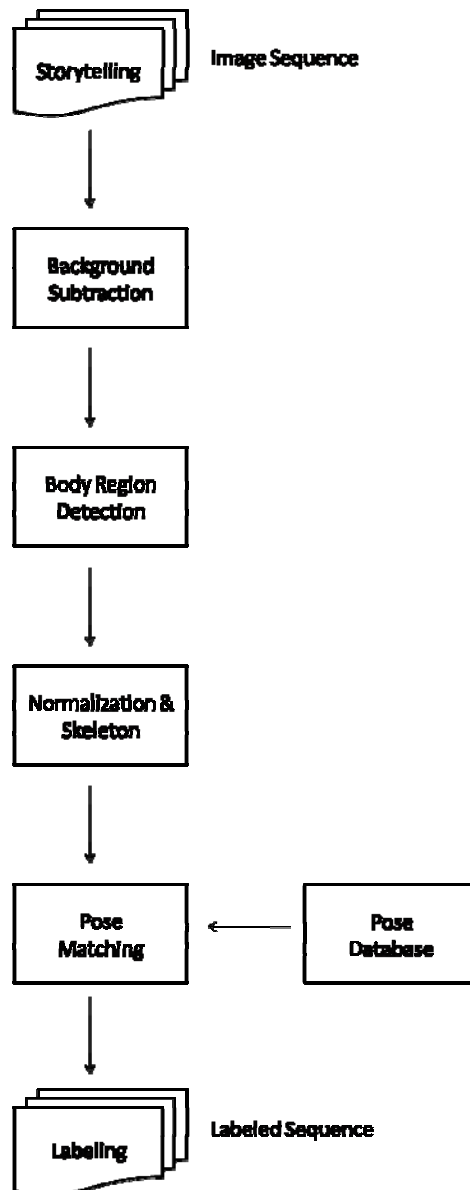


Fig. 2. The flowchart of the pose analysis algorithm in the RobotTell system.

Step 1: Storytelling

Before a child starts to tell a story, a background image without the storyteller has to be acquired for the following background subtraction step. When the child is telling the story, the whole image sequence is recorded and then is analyzed based on the template-matching scheme.

Step 2: Background subtraction

The background of the storytelling environment is assumed to be static. Then, the background subtraction method is utilized as the preprocessing step to segment the foreground from the background so that the storyteller's body can be clipped from the image.

Step 3: Body Region Detection

The silhouette of the storyteller can then be identified from the foreground since the body region will be the largest connected region in the foreground.

Step 4: Normalization and skeleton

The detected body region is firstly normalized as an image with the size of $M_b \times N_b$ (e.g., 60×80 in our system). Then, a thinning algorithm is adopted to find the skeleton of the silhouette of the storyteller.

Step 5: Pose Matching

The skeleton of the silhouette is then matched to the basic poses in the pose database.

Note that the basic poses are all normalized and skeletonized in advance. Since the skeleton images are binary, we can then use the correlation coefficient computation method to select the most matched basic pose from the database.

We adopted 17 basic poses (shown in Fig. 3) in our prototype of the RobotTell system. For each basic pose, we tune the joint angles of the robotic bear to stretch its body structure to display a pose which is similar to the corresponding basic pose. An example of the matching result from an image is shown in Fig. 4.

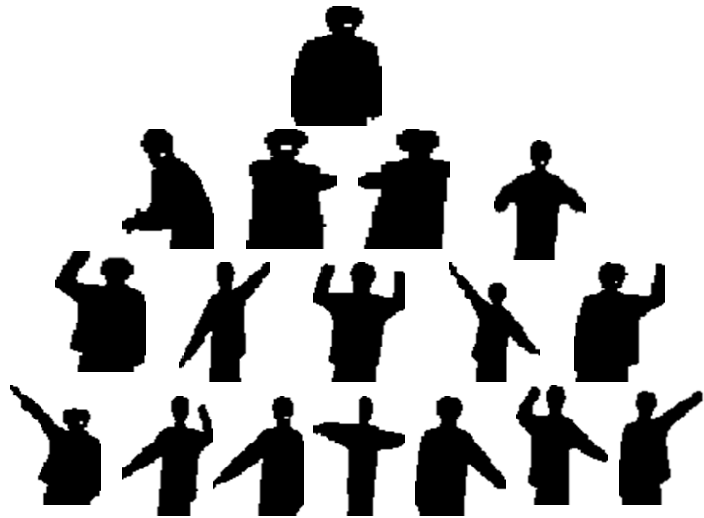


Fig. 3. The 17 pre-defined basic poses stored in the RobotTell system.

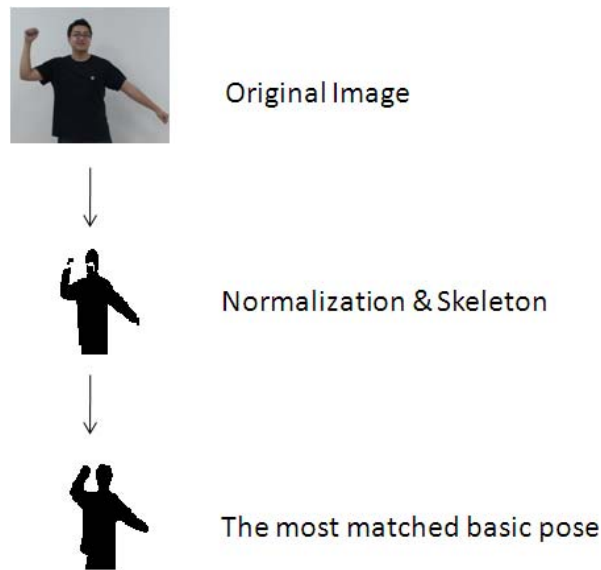


Fig. 4. A sample result of the detected basic pose stored in the RobotTell system.

Step 6: Labeling

After we have found the most matched basic pose for the present frame, we then use a token to label the image frame. The image sequence recorded from the storytelling event is then transformed to a labeled numerical sequence. Based on the labeled numerical sequence, the robotic bear can tell the story as what the child did before. Explain colloquial language and puns. Phrases like “red herring” require a cultural knowledge of English. Humor and irony are difficult to translate.

3. Conclusion

While most of the existing storytelling systems did not incorporate the education method “learning by teaching” into their functionalities, our RobotTell tried to explore the possibility of learning how to tell stories by teaching a robotic companion to tell stories. In our system, a tiny furry robotic bear is adopted to mimic the actions of the children when they are telling stories. To teach the robotic bear to tell stories, the children are not required to attach a certain amount of sensors on their bodies. What the children need to do is to tell stories in front of a low-cost Web camera. Then the robotic bear will mimic how the children tell the stories. One of the aims of the RobotTell is to make storytelling process a joyful and interactive activity. The other aim is to make the system setup as simple and low-cost as possible.

Acknowledgements

This work was partly supported by the National Science Council, Taiwan, R.O.C, under the NSC 98-2221-E-008-094-MY3, the NSC 98-2221-E-008-085-MY2, the NSC 99-2631-S-008-004 -, and the 99-2631-S-008-002-.

References

- [1] Moeslund, T. B., & Granum, E. (2001). A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, 8(3), 231-268.
- [2] Buxton, H. (2003). Learning and understanding dynamic scene activity: a review. *Image and Vision Computing*, 21(8), 125-136.
- [3] Wang, L., Hu, W., & Tan, T. (2003). Recent development in human motion analysis. *Pattern Recognition*, 36(8), 585-601.
- [4] Hu, W., Tan, T., Wang, L., & Maybank, S. (2004). A survey on visual surveillance of object motion and behavior. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 34(3), 334-352.
- [5] Aggarwal, J. K., & Park, S. (2004). Human motion: modeling and recognition of actions and interactions. in: *Second International Symposium on 3D Data Processing, Visualization and Transmission*.
- [6] Moeslund, T. B., Hilton, A. & Kruger, V. (2006). A survey of advance in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2), 90-126.
- [7] Poppe, R. (2007). Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding*. 108(1), 4-18.
- [8] Tingöy, Ö., Günefler, A., Öngün, E., Demirag, A., & Köroğlu, O. (2006). Using storytelling in education. in *4th International Symposium of Interactive Media Design*.
- [9] Wright, A. (1995). *Creating Stories with Children*, Oxford University Press, England.
- [10] Bers, M., & Cassell, J. (1998). Interactive Storytelling Systems for Children: Using Technology to Explore Language and Identity. *Journal of Interactive Learning Research*, 9(2), 183-215.
- [11] Ryokai, K., & Cassell, J. (1999). Computer support for children's collaborative fantasy play and storytelling. In *Proceedings of Computer Support for Collaborative Learning*, 510-517.
- [12] Druin, A., Montemayor, J., Hendler, J., McAlister, B., Boltman, A., Fiterman, E., Plaisant, A., Kruskal, A., Olsen, H., Revett, I., Plaisant, C., Schwenn, T., Sumida, L., & Wagner, R. (1999). Designing PETS: A Personal Electronic Teller of Stories. In *Proceedings of CHI'99*, ACM Press.
- [13] Plaisant, C., Druin, A., Lathan, C., Dakhane, K., Edwards, K., Vice, J. M., & Montemayor, J. (2000). A Storytelling Robot for Pediatric Rehabilitation. In *Proceedings of the fourth international ACM conference on Assistive technologies*, 50-55, Arlington, VA.
- [14] Ananny, M., & Cassell, J. (2001). Telling tales: A new toy for encouraging written literacy through oral storytelling, In: *Proceedings of Society for Research in Child Development Biennial Meeting, Minneapolis, MN*.
- [15] Ryokai, K., Vaucelle, C., & Cassell, J. (2002). Literacy Learning by Storytelling with a Virtual Peer. *Proceedings of Computer Support for Collaborative Learning*.
- [16] Vaucelle, C., & Jehan, T. (2002). Dolltalk: a computational toy to enhance children's creativity. *ACM CHI 2002 Conference Proceedings*, Minneapolis, USA.
- [17] Hourcade, J.P., Bederson, B.B., Druin, A., & Taxen, G. (2002). KidPad: Collaborative Storytelling for Children. *Extended Abstracts of Human Factors in Computing Systems*.
- [18] Montemayor, A., Druin, G., Chipman Farber, A., & Guha, M.L. (2004). Tools for children to create physical interactive StoryRooms. *Computers in Entertainment: Educating children through entertainment*.
- [19] Hall, L., Vala, M., Hall, M., Webster, M., Woods, S., Gordon, A., & Aylett, R. (2006). FearNot's appearance: Reflecting Children's Expectations and Perspectives. *Lecture Notes in Computer Science*, 4133, 407-419, Springer, Heidelberg.
- [20] Mutlu, B. J., Hodgins, K., & Forlizzi, J. (2006). A Storytelling Robot: Modeling and Evaluation of Human-like Gaze Behavior. *IEEE-RAS International Conference on Humanoid Robots*.
- [21] Kelleher, C., Pausch, R., & Kiesler, S. (2007). Storytelling Alice Motivates Middle School Girls to Learn Computer Programming. *Conference on Human Factors in Computing Systems*.
- [22] Wang, D., Ying, T., Peng, F., Xiong, J., Wang, H., & Dai, G. (2009). A Story Authoring System for Children. *Edutainment 2009*, 228-238.