Berthouze, L., Prince, C. G., Littman, M., Kozima, H., and Balkenius, C. (2007).
Proceedings of the Seventh International Conference on Epigenetic Robotics: Modeling
Cognitive Development in Robotic Systems. *Lund University Cognitive Studies, 135*.

Parental Signal Indicating Significant State Change in Action Demonstration

Yukie Nagai **Katharina J. Rohlfing**
Faculty of Technology, Bielefeld University,
33594 Bielefeld, Germany
{yukie, rohlfing}@techfak.uni-bielefeld.de

Abstract

This paper presents a new insight from a computational analysis of parental actions. Developmental behavioral studies have suggested that parental modifications in their actions directed to infants versus to adults may aid the infants' processing of the actions. We have been analyzing parental actions using a bottom-up attention model so as to take the advantage in robot action learning. Our latest result indicates that parental social signals can be used for a robot to detect significant state changes in the demonstrated action.

1. Introduction

A difficulty in robot action learning is that a robot does not know what aspects of the demonstrated action it should attend to. While it is exposed to a huge amount of sensory signals, only a part of them might be important for performing the action. We propose that parental social signals in forms of emotional expressions and speech observed in infant-directed interactions (IDI) enable a robot to detect significant state changes in their actions.

Parents tend to address an infant partner in the middle of an action demonstration. They suppress their task-relevant movements and start to talk to the infant, whereas they rarely stop their movements when addressing an adult partner (Rohlfing et al., 2006). It seems that parents try to draw the infant's closer attention to the demonstration by giving the social signals. Our former analysis focusing on the spatial saliency of the demonstrated action showed that such a parental modification in IDI can highlight the parent's face even during the task demonstration (Nagai and Rohlfing, 2007).

Here we focus on the temporal characteristic of the highlighted information. Although we revealed that the parent's face can be salient enough to attract the robot's attention, we do not know yet when it becomes salient and how it could help a robot to learn the action. This paper presents a new insight on the role of parental social signals in the action demonstrations.

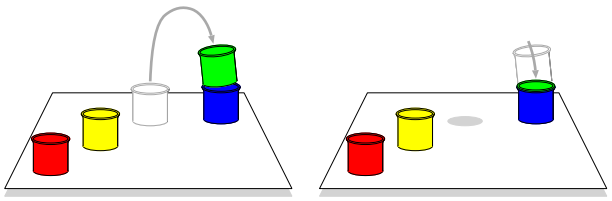
(a) Input image with attended locations by saliency model (b) Saliency map corresponding to (a)

Figure 1: A scene of parental action demonstrations to infants

2. Analysis of Parental Actions

We introduced a saliency-based attention model (Itti et al., 1998) to examine the visual information highlighted by parental action modifications. The model, simulating the bottom-up attention mechanism of primates, detects salient points in a scene in terms of primitive features: color, intensity, orientation, flicker, and motion. For example, a human face can be detected as a salient location because of the color outstanding against natural backgrounds, rich edge features, and movement. Objects used in an action demonstration can also be salient in terms of their intrinsic features without any knowledge about the task or the context. Especially, we consider that parental action modifications in IDI have the effect of emphasizing important locations in their actions, which causes higher saliency for them enough to attract the model's attention.

Figure 1 (a) and (b) show a sample scene from the experiment and the corresponding saliency map. A father was demonstrating a stacking-cups task to his infant; specifically he was closely showing the green cup held by his right hand. As denoted by circles in Figure 1 (a), the cup attracted the model's attention due to the outstanding color and the motion. Refer to (Nagai and Rohlfing, 2007) for a more detail explanation.



(a) Period of moving a cup (denoted by grayed window in Figure 3)

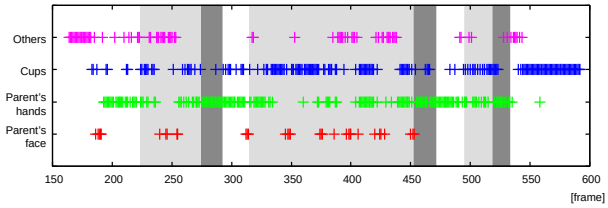
(b) Period of putting a cup into another (denoted by a darker window in Figure 3)

Figure 2: Two periods in a stacking-cups task

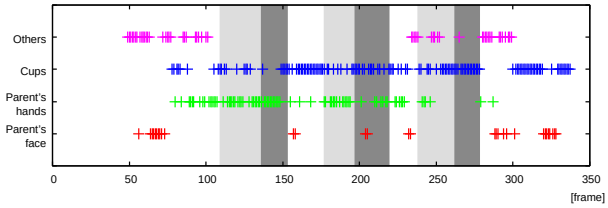
3. Experimental Results

To investigate when the parent's face attracts the model's attention and how it can help to learn the demonstrated action, we analyzed the history of the attended locations. Figure 3 plots the locations attended to by the saliency model over the duration of the task demonstration; (a) is the result for a mother interacting with her 10-month-old infant and (b) is another mother interacting with her 9-month-old infant. The attended locations, denoted by "+" every frame, are classified into four regions: the parent's face, his/her hands, the cups, and others. The grayed time windows and darker ones show the periods during which the parents were moving a cup from the tray to another cup (see Figure 2 (a)) and those during which they were putting the cup into the other (see Figure 2 (b)), respectively. It is therefore considered that the darker windows in Figure 3, corresponding to the illustration of Figure 2 (b), represent the periods of the significant state changes in the stacking-cups task.

From the result shown in Figure 3 (a), we can see that the mother-A's face attracted more attention shortly before the significant state changes. Figure 4 (a) gives the scene captured at 401st frame, which is just before the second cup was being stacked into another. She was talking to her infant while suppressing her cup-handling movement, which made her face salient enough to draw the model's attention. She seemed to send social signals to alert her infant to the following movement, i.e., putting the cup held by her right hand into another one, before demonstrating it. As another typical phenomenon, Figure 3 (b) indicates that the mother-B emphasized the significant state changes after fulfilling them. She took a long pause and addressed her infant immediately after stacking each cup. Figure 4 (b) shows the scene captured at 159th frame, where her face attracted the model's attention. Note that the cup held by her hand had already been put in another one, meaning that the significant event had just been demonstrated. Although a further statistical analysis is required, we can suggest from these results that the highlighted parent's face indicates the significant

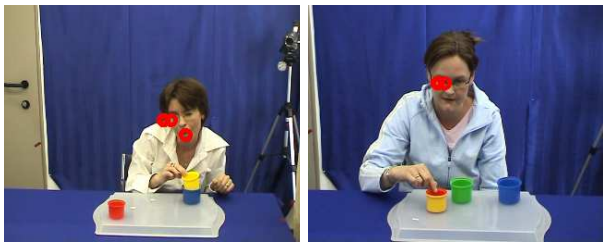


(a) More attention on the parent's face before the significant state changes (mother-A)



(b) More attention on the parent's face immediately after the state changes (mother-B)

Figure 3: History of the attended locations by the saliency model



(a) Mother-A talking to her infant before putting down the held cup into another

(b) Mother-B addressing her infant just after stacking the cups

Figure 4: Sample scenes for parental social signals toward infants

events in the demonstrated action and could enable a robot to detect those events beforehand and/or afterward in learning for the action.

References

Itti, L., Koch, C., and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259.

Nagai, Y. and Rohlfing, K. J. (2007). Can motionese tell infants and robots "what to imitate"? In *Proceedings of the 4th International Symposium on Imitation in Animals and Artifacts*, pages 299–306.

Rohlfing, K. J., Fritsch, J., Wrede, B., and Jungmann, T. (2006). How can multimodal cues from child-directed interaction reduce learning complexity in robot? *Advanced Robotics*, 20(10):1183–1199.