Understanding Information Transfer in Caregiver-Infant Interaction

*Shibo Qin(Osaka Univ.), Yukie Nagai(Osaka Univ.),
Akiko Nakatani(Osaka Univ.), Hiroshi Fukuyama(Kyoto Univ.),
Masako Myowa-Yamakoshi(Kyoto Univ., JST ERATO), and Minoru Asada(Osaka Univ.)

1. Introduction

Cognitive developmental robotics (CDR) aims to provide new understandings of how human acquires higher cognitive functions by means of a synthetic approach [1]. Among various aspects of human development, we put emphasis on social interaction, that is, how interaction with others, especially with caregivers, facilitates infants' development.

In caregiver-infant interaction, infants try to receive various kinds of information from caregivers while caregivers exaggerate their actions and speech to help infants' learning [2, 3]. Unlike conventional studies based on qualitative analysis, Yu, Smith, and colleagues applied information transfer to quantify how participants orchestrated speech, visual attention, and body movements. Their experiment showed that a better coordination between infants and caregivers leads to a higher accuracy for leaning the names of objects [4, 5].

However, they have not dealt with an issue of developmental aspect that both caregivers and infants change their behavior through. We propose a new method to quantitatively evaluate the dynamic structure of information exchange between caregivers and infants in different age groups. It is supposed that their body movements for task learning convey various signals to each other. Our study intends to reveal developmental changes in their information exchange using transfer entropy [6].

2. Information Transfer in Caregiver-Infant Interaction

2·1 Introduction of Transfer Entropy

Transfer entropy (TE) is an information theoretic measure that quantifies the statistical coherence between systems evolving in time [6]. For time series data I and J, TE from J to I is defined as follows:

$$T_{J\to I} = \sum p(i_{n+1}, i_n^{(k)}, j_n^{(l)}) \log \frac{p(i_{n+1}|i_n^{(k)}, j_n^{(l)})}{p(i_{n+1}|i_n^{(k)})}, \quad (1)$$

which represents the influence of the last l steps of J on the I's next step after deducting its own history for k steps.

A. Nakatani works at Panasonic Corporation since April 2012.

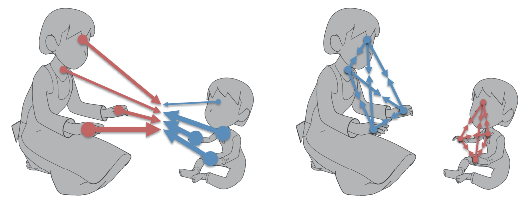


Fig. 1 Information flows in caregiver-infant interaction. The left side shows information transfer between two participants (called inter-TE); the right side shows information transfer within each participant (called intra-TE).

2·2 Our Hypotheses about Development

We apply TE to measuring information flows *within* and *between* caregivers and infants. In Fig. 1, the left side illustrates information flows *between* participants (called inter-TE), which represents social contingency. The right, on the other hand, shows information flows *within* each participant (called intra-TE), which corresponds to body coordination.

Studies on developmental psychology have suggested that infants' social contingency develops as they grow [7]. For example, infants start tracking caregivers' gaze after 6 months of age [8]. Based on such studies on developmental psychology and the theories of information flow, our hypotheses for the information transfer are made as follows:

(a) The development of infants' social contingency causes an increase of inter-TEs from caregivers to infants.
(b) In the opposite direction, the inter-TEs from infants to caregivers also increase as the infants' social contingency develops. Specifically, the inter-TE from infants' dominant (most cases, right) hand to caregivers, as its unequal development, may increase more significantly than that from infants' non-dominant (most cases, left) hand.
(c) The development of body coordination causes an increment in infants' intra-TEs, while caregivers' intra-TEs may also rise to adapt infants' development.

3. Analyzing Method

Two KINECT sensors with OpenNI [9] were used to record the 3D skeleton and the depth information

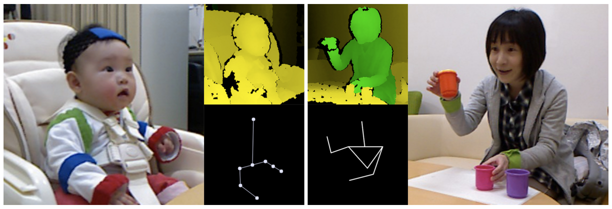


Fig. 2 RGB, depth, and 3D skeleton information recorded by KINECT sensors.

of participants' bodies. Fig. 2 shows an example of the skeleton generation. For caregivers, OpenNI provides user recognition, which can automatically generate the coordinate of the skeleton. For infants, however, OpenNI does not provide the function of infant recognition. Therefore we applied markers to detect the position of the skeleton in the 2D image so that it can be combined with the depth information. We used this information to generate the coordinate of infants' skeleton.

Four kinds of time series data: right hand position (rhand), left hand position (lhand), body orientation (torso), and gaze information (gaze) are used to describe the state of each participant. Therefore we have 32 kinds of inter-TEs (16 from caregivers to infants and 16 for the other way around) and 24 kinds of intra-TEs (12 for each participant). Note that intra-TEs between the same data sequence cannot be calculated (e.g., from infants' right hand to the same right hand).

Let x, y, and z be the horizontal, vertical, and depth positions in the image frames recorded by KINECT. We detected the right and left hands position directly from the skeleton coordinate and generated the torso information at time t $(z_{\mathrm{torso},t}, \theta_{\mathrm{torso},t}, z_{\mathrm{head},t})$ using the shoulders and head positions:

$$\begin{cases} z_{\mathrm{torso},t} = (z_{\mathrm{lshld},t} + z_{\mathrm{rshld},t})/2 \\ \theta_{\mathrm{torso},t} = \arctan\left(\dfrac{z_{\mathrm{lshld},t} - z_{\mathrm{rshld},t}}{x_{\mathrm{rshld},t} - x_{\mathrm{lshld},t}}\right) \\ z_{\mathrm{head},t} = z_{\mathrm{head},t} \end{cases} \quad (2)$$

Only the parameter of gaze was determined subjectively in the following rules. When a participant was looking at the objects used in the interaction, gaze information gaze_t at time t was assigned a label "1", assigned "2" if looking at the opponent, and "3" for looking at other things. For each participant the four data sequences were:

$$\begin{cases} \mathrm{rhand}_t(x_{\mathrm{rhand},t}, y_{\mathrm{rhand},t}, z_{\mathrm{rhand},t}) \\ \mathrm{lhand}_t(x_{\mathrm{lhand},t}, y_{\mathrm{lhand},t}, z_{\mathrm{lhand},t}) \\ \mathrm{torso}_t(z_{\mathrm{torso},t}, \theta_{\mathrm{torso},t}, z_{\mathrm{head},t}) \\ \mathrm{gaze}_t \end{cases} \quad (3)$$

Next we applied robust singular spectrum transform [10] to segment each data sequence, and then assigned a label to each frame. The label l_t between two segment points $t = s_n$ and s_{n+1} was determined

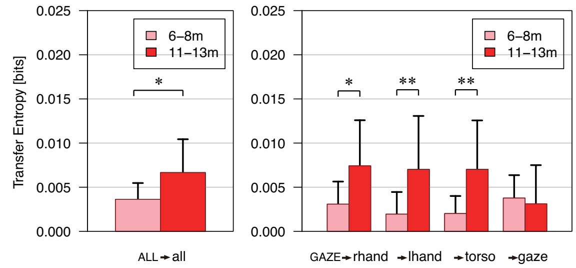


Fig. 3 Inter-TEs from caregivers to infants. The graph shows the average and standard deviation. A significant increment appears in $T_{\mathrm{ALL}\to\mathrm{all}}$, indicating the development of social contingency.

according to the motion direction. For example, l_t for $x_{\mathrm{rhand},t}$ was:

$$l_{x_{\mathrm{rhand},t}} = \begin{cases} 1 & \text{if } x_{\mathrm{rhand},s_n} < x_{\mathrm{rhand},s_{n+1}} \\ 2 & \text{if } x_{\mathrm{rhand},s_n} = x_{\mathrm{rhand},s_{n+1}} \\ 3 & \text{if } x_{\mathrm{rhand},s_n} > x_{\mathrm{rhand},s_{n+1}} \end{cases} \quad (4)$$

In this way, if rhand_t increases in every coordinate the label is "111", whereas "222" if no significant changes appear. For gaze_t there is no need for labeling it because the segmentation of gaze_t can be treated as labels. We applied these labels to calculating inter-TEs and intra-TEs in the following analyses.

4. Experiment and Result

4·1 Task Information

In order to elucidate information transfer from the developmental viewpoint, we divided 26 caregiver-infant pairs into two groups: a younger group for 6- to 8-month-old infants ($M = 186.4$ days, $SD = 22.6$ days, $N = 16$) and an older group for 11- to 13-month-old infants ($M = 359.5$ days, $SD = 16.7$ days, $N = 10$).

Each experiment took about 3 minutes, in which caregivers were asked to show a cup-nesting task to infants (Fig. 2). We detected the skeleton of both caregivers and infants using two KINECT sensors and then analyzed the data by the method explained in Section 3.

4·2 Transfer Entropy Analysis

In the first analysis, we calculated information transfer using the shortest history length of TE, i.e., $k = l = 1$ in Eq. (1). The following sections describe our findings related to the three hypotheses. For the sake of simplicity, we use upper case letters to represent data sequence of caregivers while lower case letters for infants.

4·2·1 Development of Infants' Social Contingency

Fig. 3 shows the inter-TEs from caregivers to infants. Generally, the average TE $T_{\mathrm{ALL}\to\mathrm{all}}$ increases significantly as infants grow, which indi-

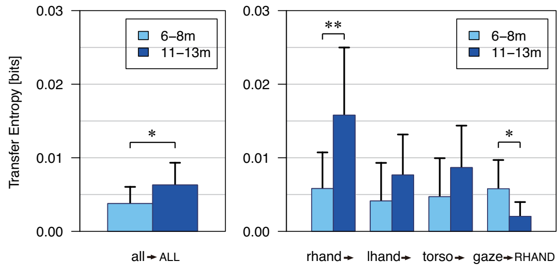


Fig. 4 Inter-TEs from infants to caregivers. A significant increase in $T_{\mathrm{all}\to\mathrm{ALL}}$ indicates caregivers' adaption to infants' social contingency.

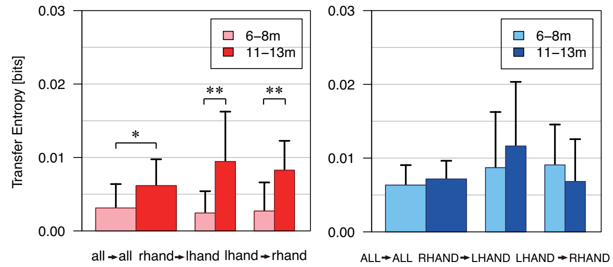


Fig. 5 Intra-TE for infants (left) and for caregivers (right). Infants improve body coordination while caregivers do not change it.

cates the infants' development of social contingency. For instance, $T_{\mathrm{GAZE}\to\mathrm{rhand}}$, $T_{\mathrm{GAZE}\to\mathrm{lhand}}$, and $T_{\mathrm{GAZE}\to\mathrm{torso}}$ show significant increments, which demonstrate stronger influence of caregivers' gaze on older infants. Gaze is an important social signal from caregivers. These results thus support our first hypothesis, i.e., the development of infants' social contingency. The reason for no significant difference in $T_{\mathrm{GAZE}\to\mathrm{gaze}}$ is that this inter-TE may contain two factors of development: eye contact and gaze following. The former phenomenon is known to appear shortly after birth whereas the latter takes 8 months. The different time of appearance for these two factors may eliminate the change in $T_{\mathrm{GAZE}\to\mathrm{gaze}}$. We will thus separately analyze them by modifying the gaze segmentation way in the future study.

4·2·2 Adaption of Caregivers' Social Contingency

The average inter-TE from infants to caregivers $T_{\mathrm{all}\to\mathrm{ALL}}$ (Fig. 4) shows a significant increment between two groups. This result supports our second hypothesis, that is, infant development causes increase in social contingency of caregivers.

Moreover, closer investigation of the result yielded an interesting finding related to infants' gaze. The right side of Fig. 4 shows a typical example: $T_{\mathrm{gaze}\to\mathrm{RHAND}}$ decreased as infants grow while the others increased to support the average of inter-TE. We conjecture a reason for it as follows: younger infants responded to caregivers mainly by shifting their gaze, which resulted in relatively higher influence from their gaze. On the other hand, older infants moved their body in a relatively larger extent to show their interest. Their body movements thus influenced on caregivers more strongly than their gaze. Furthermore, in the comparison between $T_{\mathrm{rhand}\to\mathrm{RHAND}}$ and $T_{\mathrm{lhand}\to\mathrm{RHAND}}$, the former has a significant increase while the latter does not. This difference can be regarded as the development of infants' dominant hand.

4·2·3 Development of Body Coordination

The left graph in Fig. 5 shows higher intra-TEs for older infants than that for younger infants. It suggests that the body movement of older infants are highly coordinated than that of younger infants. For instance, $T_{\mathrm{rhand}\to\mathrm{lhand}}$ and $T_{\mathrm{lhand}\to\mathrm{rhand}}$ are significantly higher for older infants than for younger infants. In contrast, caregivers' intra-TE has no significant changes regardless of infants' age, which may suggest no adaptation in caregivers. Our third hypothesis was partially supported by these results.

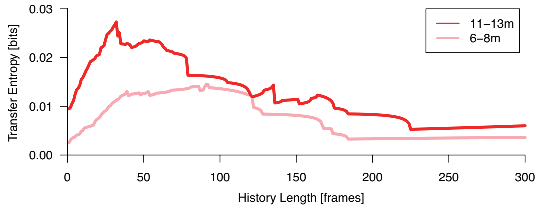
4·3 Effects of History Length on TE

The first analysis in Section 4·2 is based on TE with history length 1, which calculates influence of the current step on the next step. However, responses of caregivers and infants may have some temporal delays and/or uncertainties, and such temporal factors may change with the development of infants. We thus analyze the information exchange with a longer historical length and compare the result between two age groups.

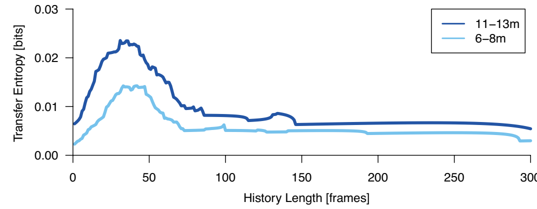
The second analysis calculated TE from the past l steps to the next step. The parameters k and l in Eq. (1) changed from 1 to 300 under a condition of $k = l$. As a preliminary experiment, we selected one example from each group: a 172-day-old infant from the younger age group and a 355-day-old infant from the older age group.

Fig. 6(a) shows the inter-TE in different history length from the caregiver to the infant. Over nearly the whole history length, we can easily find that the inter-TE from the caregiver to the older infant is higher than that to the younger infant. Of particular interest is that the inter-TE for the older infant comes to a small peak at around 37 frames whereas the younger infant does not show such a significant peak. The peak at around 37 frames means that the infant responded to the caregiver with a regular temporal delay at about 1.2 sec. The emergence of this time delay further supports the development of social contingency; more specifically, the initiation of turn-taking.

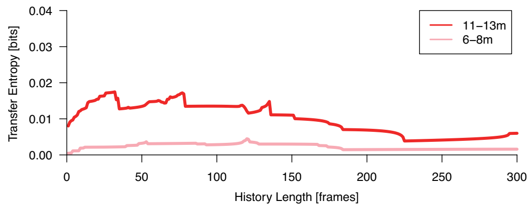
In Fig. 6(b), we applied the same analysis to inter-TE from the infant to the caregiver. The older infant gave stronger influence on the caregiver than the younger infant over the whole history length. The inter-TEs for both caregivers shows a significant peak at around 38 frames (i.e., 1.3 sec), which means that



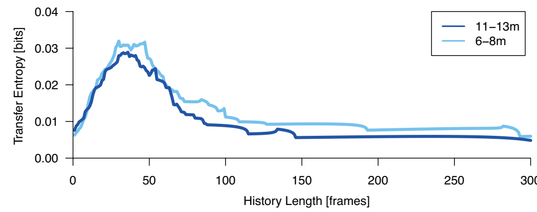
(a) Inter-TE from the caregiver to the infant



(b) Inter-TE from the infant to the caregiver



(c) Intra-TE within the infant



(d) Intra-TE within the caregiver

Fig. 6 TE analysis with varied history length. (a) and (b) are inter-TEs between participants. (c) and (d) are intra-TEs within participants.

social contingency of the caregiver was highly developed and thus did not change the response time regardless of the infant's age.

Comparing the result in Fig. 6(c), we can see a developmental change in infants' body coordination. The intra-TE of the older infant is higher than that of the younger infant. However, the TE for both infants does not show a significant peak unlike in Fig. 6(a), which needs to be further investigated.

Regarding the body coordination of the caregivers, Fig. 6(d) shows no significant difference between two groups. Caregivers may not change how to demonstrate the task regardless of infants' age.

5. Conclusion and Discussion

We proposed a new method to quantitatively evaluate the dynamic structure of information exchange between infants and caregivers. By calculating TE between and within participants, the development of infants such as social contingency, dominant hand, and body coordination were found. From the analysis with varied TE history length, we verified the development in temporal aspects of social contingency and body coordination for infants. Their response time converged at a certain delay, which is equivalent to the caregivers' one.

The development of response time is especially important in forming turn-taking in caregiver-infant interaction. The analysis with TE history length is a preliminary experiment since only one pair of data from each age group was examined. We will analyze more data to find out common properties in infants' turn-taking.

Our study focusing on social interaction provides a new design for CDR as well as elucidates the development of human beings. The probability and timing of infants' response can be embedded into robots so that the robots can establish infant-like interaction with human caregivers.

Acknowledgment

This study is supported by JSPS Grants-in-Aid for Scientific Research (Research Project Number: 24000012, 24119003) and by JST ERATO Okanoya Emotional Information Project.

Reference

[1] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida: "Cognitive developmental robotics: a survey", IEEE Transactions on Autonomous Mental Development, vol.1, no.1, pp.12-34, 2009.

[2] R. J. Brand, D. A. Baldwin, and L. A. Ashburn: "Evidence for 'motionese': modifications in mothers' infant-directed action", Developmental Science, vol.5, no.1, pp.72-83, 2002.

[3] A. Fernald and T. Simon: "Expanded intonation contours in mothers' speech to newborns", Developmental Psychology, vol.20, no.1, pp.104-113, 1984.

[4] A. F. Pereira, L. B. Smith, and C. Yu: "Social coordination in toddler's word learning interacting systems of perception and action", Connection Science, vol.20, no.2, pp.73-89, 2008.

[5] H. Choi, C. Yu, L. B. Smith, and O. Sporns: "From Data Streams to Information Flow: Information Exchange in Child-Parent Interaction Experiment and Data Preprocessing ", Proc. of the 33nd Annual Conference of the Cognitive Science Society, pp.868-873, 2011.

[6] T. Schreiber: "Measuring information transfer", Physical Review Letters, vol.85, no.2, pp.461-464, 2000.

[7] J. Colombo: "Infants' detection of contingency: a cognitive-neuroscience perspective", Bulletin of the Menninger Clinic, vol.65, no.3, pp.321-334, 2001.

[8] G. Butterworth and N. Jarrett: "What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy", Developmental Psychology, vol.9, no.1, pp.55-72, 1991.

[9] "OpenNI documentation homepage", http://openni.org/Documentation/

[10] Y. Mohammad and T. Nishida: "Robust Singular Spectrum Transform ", Lecture Notes in Computer Science, vol.5579, pp.123-132, 2009.