Touch and Emotion: 
Modeling of developmental differentiation of emotion lead by tactile dominance

Takato Horii, Yukie Nagai, and Minoru Asada 
Graduate School of Engineering Osaka University 
2-1 Yamada-oka, Suita, Osaka, 565-0871 Japan 
Email: {takato.horii, yukie, asada}@ams.eng.osaka-u.ac.jp

Abstract—Emotion is one of the important elements for humans to communicate with others. Humans are known to share basic emotions such as joy and anger although their developmental changes have been studied less. We propose a computational model for the emotional development in infancy. Our model reproduces the differentiation of emotion from pleasant/unpleasant states to six basic emotions as known in psychological studies. The key idea is twofold: the tactile dominance in infant-caregiver interaction and the inherent ability of tactile sense to discriminate pleasant/unpleasant states. Our model consists of probabilistic neural networks called Restricted Boltzmann Machines. The networks are hierarchically organized to first extract important features from tactile, auditory, and visual stimuli and then to integrate them to represent an emotional state. Pleasant/unpleasant information is directly provided to the highest level of the network to facilitate the emotional differentiation. Experimental results show that our model with the tactile dominance leads to better differentiation of emotion than others without such dominance.

I. Introduction

Humans express their internal states, which are shaped by the information from the external sense, the internal sense, and the proprioception. The best known internal state is emotion such as joy, anger, fear, etc. Emotion has influences of triggering decision making, changing perception and action, and enhancing memory and learning [1]. In social contexts, emotion is communicated between humans [2]. People transmit their internal states by changing their facial expressions, vocalizations, and so on.

We, humans, can express a variety of emotional states by using different ways such as facial, vocal, and gestural expressions. Russell [3] proposed a circumplex model representation of emotional relations. The model is represented in two dimensions, which are corresponding to pleasure/unpleasure and arousal/sleep axes. Shaver et al. [4] classified emotions hierarchically. Ekman [5] has defined six basic emotions: joy, surprise, anger, fear, sadness, and disgust. These emotions are expressed in an earlier stage of development than other emotions [6]. However, how humans acquire various states of emotion is still an open issue. It has been argued that the emotional category is nature or nurture. We suppose that emotions differentiate from primitive internal states like pleasure/unpleasure as studied in psychology [6]–[8]. Primitive emotions vary with sensory inputs from the environment. Especially, a tactile sensation directly conveys unpleasant as pain. The tactile interaction amount to over 30% of communication with caregiver in infancy [9]. Hertenstein [10] studied the relationship between emotions and tactile interaction in infant-caregiver communication. He pointed out that tactile interaction has not been studied much in the context of emotional development.

In this study, we propose a computational model for the emotional development in infancy. The model reproduces the differentiation of emotion from a pleasant/unpleasant state to six basic emotions as known in psychological studies. The key idea is twofold: the inherent ability of tactile sense to discriminate pleasant/unpleasant state and the tactile dominance in infant-caregiver interaction. The model consists of probabilistic neural networks called Restricted Boltzmann Machine. The networks are hierarchically organized to first extract important features from tactile, auditory, and visual sensory information and then to integrate them to represent the emotional states. Pleasant/unpleasant information is directly provided to the highest level of the network to facilitate the emotional differentiation.

The rest of this paper is organized as follows: the basic idea of our approach with a general hypothesis of emotional development is given, and then the details of the model with its assumptions are described. Experimental results under different conditions are shown to verify the hypothesis, and future issues are discussed.
II. Behavioral Evidence about Emotional Development

The question of whether the six basic emotions are nature or nurture is still controversial. Tomkins [11] and Izard [12] proposed discrete emotion theory. The theory explains that emotional states are innate sets of facial expressions and bodily reactions. In contrast, Bridges [6], Sroufe [7], and Lewis [8] suggested that emotions are acquired through development from infancy to childhood. According to Bridges’ theory, newborns have only general excitement, which is first differentiated into delight and distress, and then into five emotions such as “Elation”, “Affection”, etc (see Fig. 1).

However, the above studies observed only the behavioral change in emotional expressions. They did not measure brain activations or hormones in infants. It is not clear what triggers the developmental differentiation of emotion. Our approach focuses on tactile communication in infant-caregiver interaction inspired by tactile dominance in infancy.

III. Our Hypothesis about Tactile Dominance

The sense of touch is an important ability for infants to gather information from other individuals and environment. Especially, the tactile organ can perceive emotional information more directly than other sensors [13]. For example, contact with other’s skin can transfer the internal state of the person with less ambiguity than vision and auditory, which have more contextual information. Here we introduce two types of tactile dominance:

- Tactile superiority in interaction:
  It is known that the proportion of tactile interactions between an infant and a caregiver is larger than those using other modalities [10] [9] [14].

- Inherent ability to perceive emotional information:
  There are nerve fibers in human skin, which easily sense primitive emotions like pleasure and displeasure (e.g. noxious stimuli and comfortable touch) [15] [16].

The amount of tactile interaction reaches 30 to 60% of the total amount of infant-caregiver interaction [9]. It has been suggested that experiences of touch affect the development of cognitive functions in infants [10]. Hertenstein [10] investigated the effect of tactile interaction with caregivers on the emotional expressions of infants. His studies showed that frequencies of tactile interaction from a caregiver elicit the expression of pleasant emotion in an infant (e.g. frequencies of infants’ smiles are increased [17] [18]).

Infants more strongly react in their cerebral cortex for tactile stimuli than other sensory inputs. Brain imaging study has revealed that tactile stimulation activates wider range of the cerebral cortex of neonate than visual and auditory stimuli [19]. Björnsdotter et al. [15] examined the anatomical mechanism, which transmit signals of C-fiber. C-fibers are classified into two types: one transmits pain information and the other reacts pleasant touch like 1-10[cm/s] velocity of stroke on the hairy skin. It is assumed that a comfortableness is detected by C-fibers which are activated according strokes [20]. These fibers have connection not only with the somatosensory area in cerebral cortex but also with the insular cortex [15]. Craig [21] measured the activity of the insular cortex to tactile stimuli by using fMRI. Furthermore, the brain of features even at 36 weeks after the conception has been reported to respond to unpleasant stimuli. These evidences support our hypothesis about the tactile superiority in interaction and the inherent ability to perceive emotional information.

IV. Problem Setting and Assumptions

We assume a face-to-face interaction between an infant and a caregiver as shown in Fig. 2. An infant is able to receive three types of modality information: tactile, auditory, and visual stimulation detected through the interaction. The tactile sensor can perceive different types of touch like stroke, push, pinch, etc. Auditory stimuli are vocalizations of the caregiver, and his facial expressions are communicated to the infant as visual stimuli. For the sake of simplicity, no sensory information irrelevant to the emotional interaction is considered in the current experiment (e.g. tactile stimuli produced by the body movement of the infant, sounds from the environment, the facial expression of other people). The sensory input from the three modalities are supposed not to contradict each other in terms of the emotional state. For example, the caregiver gives stroke stimuli to the infant when the caregiver expresses a smile and a cheerful voice. The infant does not realize the detailed state of the caregiver’s emotion from auditory and visual stimuli. However, the infant can discriminate between pleasant and unpleasant directly from tactile information of interaction according to the superiority of emotional perception in sense of touch.

V. A Model for Emotional Differentiation

A. Basic idea

Fig. 3 shows the proposed model for the developmental differentiation of emotion. This model consists of two types of modules: Sensory feature modules and Feature integration module. The lower ones are Sensory feature modules, which receive sensory information through the interaction. The upper one is called Feature integration module, which represents the emotional state by integrating extracted information from each modality. Each module is made of a neural network called Restricted Boltzmann Machine (RBM) [22]. RBM has...
Fig. 3. A computational model of emotional differentiation based on tactile dominance in infancy. Each lower RBM network which is drawn by the blue connection (i.e., the region of enclosed broken lines) constitutes a Sensory feature module for the corresponding modality. The top layer of the RBM which is drawn by the green connection and emotional units which represent a rough emotion according to the superiority of emotional perception in sense of touch constitute a Feature integration module (i.e., the region of enclosed chain lines).

A function to abstract input data in a hidden layer based on non-linear self-organization and thus can reconstruct the data through the hidden layer regardless of a lack of the input information. For example, the proposed model is able to generate absent sensory information only from the existing information.

B. Restricted Boltzmann Machine

RBM [22] is a probabilistic neural network consisting of two layers (i.e., the input (visible) layer and the hidden layer). Units in one layer are connected to all units in another layer, and their connections are symmetrical, that is, the weight $w_{ij}$ from an input unit $i$ to a hidden unit $j$ equals to $w_{ji}$. The activity of $i$ and $j$, $v_i$ and $h_j$, are calculated by:

$$p(h_j = 1) = \frac{1}{1 + \exp(-b_j - \sum_i v_i w_{ij})}$$  \hspace{1cm} (1)$$

$$p(v_i = 1) = \frac{1}{1 + \exp(-c_i - \sum_j h_j w_{ij})}$$  \hspace{1cm} (2)$$

where $b_j$ and $c_i$ are activation biases.

The RBM is trained through the reconstruction calculation as shown in Fig. 4. First, the activation probability of the hidden layer $p(h^j = 1)$ is calculated by Eq. (1) using an actual input $v^0_i$. Then, a reconstruction of $v^1_i$ is calculated by Eq. (2) using $h^0_j$. This process is repeated several times. We define $p(v)$ and $p(v|w)$ as the probabilistic distributions of the input data and data which are reconstructed after $\infty$ reconstruction repeats. The probabilistic distributions are used to minimize the difference between $p(v)$ and $p(v|w)$ by adjusting connection weights. The distance between the two distributions is measured by the cross-entropy error, which is calculated by the following equation:

$$L = -\left( \log(p(v|w)) \right)_{p(v)}$$

$$= - \sum_i p(v_i) \log(p(v_i|w))$$  \hspace{1cm} (3)$$

The connection weights are trained to decrease this error.

C. Sensory feature module

A sensory feature module consists of an RBM. Each module processes sensory information (tactile, auditory, and visual) individually. Units in the hidden layer represent the abstracted information of each modality.

D. Feature integration module

The feature integration module consists of an RBM and units for back propagation. This module receives the combined information from all the Sensory feature modules. We trained the model through two steps. First, the module combines output data which are trained at each sensory feature module. Next, the proposed model executes the back propagation to use the emotional state which is given by the superiority of emotional perception in the sense of touch.

VI. Experiments

A. Experimental setting

We evaluated the proposed model using data acquired through simulated infant-caregiver interaction. A purpose of this experiment is to investigate how the emotional differentiation changes by having the superiority of emotional perception in sense of touch or not. The model uses assumed interaction data which are gathered from a tactile sensor, a microphone, and a camera and an emotional label which was attributed to tactile stimulus (i.e., pleasant, neutral, or unpleasant). Each dataset has one emotional state from among the 6 basic emotions (i.e., joy, surprise, anger, disgust, sadness, and fear) and neutral. Note that the proposed model is not provided with the emotional state but only with the pleasant/unpleasant information included in tactile stimuli. Table I shows samples of dataset. We conducted off-line learning using the dataset.

Fig. 4. The learning process of Restricted Boltzmann Machine.

### TABLE I. SAMPLES OF DATASET FOR THE PROPOSED MODEL FROM ASSUMED INTERACTIONS.

<table>
<thead>
<tr>
<th>Emotional state</th>
<th>Tactile (Emotional label)</th>
<th>Auditory</th>
<th>Visual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Joy</td>
<td>Stroke (Pleasant)</td>
<td>Pitch risen voice</td>
<td>Smiley face</td>
</tr>
<tr>
<td>Neutral</td>
<td>Push (Neutral)</td>
<td>Neutral voice</td>
<td>Neutral face</td>
</tr>
<tr>
<td>Anger</td>
<td>Pinch (Unpleasant)</td>
<td>Loud voice</td>
<td>Anger face</td>
</tr>
<tr>
<td>Joy</td>
<td>Pinch (Unpleasant)</td>
<td>Smiley face</td>
<td>Loud voice</td>
</tr>
</tbody>
</table>

TABLE I. SAMPLES OF DATASET FOR THE PROPOSED MODEL FROM ASSUMED INTERACTIONS.
consideration of all tactile dominance (full dominance: FD).

B. The condition of full dominance

Fig. 6 shows two emotional faces of the caregiver.

We calculate the change in pitch frequency and intensity of voice data which expresses the emotional state of a caregiver. Push stimuli are neutral.

2) Auditory stimuli: Auditory stimuli are gathered from voice data which expresses the emotional state of a caregiver. We calculate the change in pitch frequency and intensity of sounds. Auditory stimulation has an emotional state from among the 6 basic emotion and neutral.

3) Visual stimuli: Visual stimuli are facial expressions of the caregiver. Face images are converted into 30x30 pixels and gray scale. The first 20 principal components calculated by principal component analysis are used as a visual input. Visual stimulation has an emotional state like auditory stimulation. Fig. 6 shows two emotional faces of the caregiver.

B. The condition of full dominance

We trained the proposed model under the condition in consideration of all tactile dominance (full dominance: FD).

1) Tactile stimuli: Tactile stimulation is an important factor to verify our hypothesis that the emotional differentiation are affected by the tactile dominance. We collected tactile stimuli using human-like tactile sensor as shown in Figs. 5(a) and (b). The sensor imitates the structure of human skin which is inspired by Tada et al. [23]. Fig. 5(c) shows the output signal of a PVDF film located at the center of the tactile sensor when the sensor is landed a stroke stimulation. We extract nine features: the duration of the contact, the area of the contact, the maximum contact force, the duration of the contact with the maximum force, the velocity of the contact, the intensity of the frequency in three parts(Low:1-60Hz, Middle:61-100Hz, High:101-200Hz), and the number of skin vibrations from sensory outputs. These are determined based on the knowledge about tactile receptors [24]. There signals are conveyed through four types of tactile interaction: stroke, push, pinch, and pat, each of which correspond to one of three states (i.e., pleasant, unpleasant, and neutral). The interaction of stroking has pleasant emotion, whereas unpleasant emotion is evoked from pinch and pat. Push stimuli are neutral.

Output data constituted clusters which belong to the emotional state of infant-caregiver interaction(Fig. 7(a)). Furthermore, the 1st principal component axis represents the pleasant/unpleasant emotion like the circumplex model proposed by Russell(see Figs. 7(b) and (c)). Fig. 7(b) shows the differentiation of clusters corresponding to unpleasant emotions that include anger, disgust, sadness, and fear. The pleasant emotion which comprise the joy and the surprise is classified into each cluster by 4th principal component axis(see Fig. 7(c)). The experimental result under the FD condition demonstrates an appropriate differentiation of emotions.

C. The condition without the superiority of emotional perception

The second experiment examined the proposed model under the condition without the superiority of emotional perception (WE) which stops the back propagation in the learning sequence. This experimental condition assumed the situation that the superiority of emotional perception is lost one of tactile dominance. Fig. 8 shows the output data of the proposed model under the WE condition. Output data that are shown in Fig. 8(a) constitutes clusters based on each emotional state like Fig. 7(a). However, the feature of pleasant/unpleasant emotions does not represented in 1st principal component axis in Fig. 8(b) and (c). The distribution of pleasant emotion and unpleasant emotion is disordered in the emotional space. Furthermore, output data that belong to the same emotional state constituted plural clusters. Because of this, the RBM does not guarantee to abstract features from input data according to the emotional state of the interaction stimuli. The emotional state of interaction is not given to the model under this condition. The model have no choice but to abstract features based on the distance of input stimuli. It is conceivable that the emotional information ascribable to the tactile dominance takes an important role for the abstraction and the integration of the sensory information based on the emotional state. For example, it is difficult that
The space composed of 1st, 3rd, and 4th principal components

The space composed of 1st and 3rd principal components

The space composed of 1st and 4th principal components

Fig. 7. The emotional space is composed of principal components under the condition of full dominance. The horizontal axis represents the differentiation of pleasant/unpleasant emotions. Unpleasant emotions cluster each category of emotion in (b). Pleasant emotions split into joy and surprise in (c). The vertical axis of (c) represents the clear differentiation of pleasant emotions. We recognize the state of the emotion to use only the intensity as a feature in auditory stimuli. However, it is possible that we separate the sensory information by slight differentiation of feature to use an emotional information from the tactile stimulation. Moreover, the emotional information from tactile stimuli enables to reconfigure the feature space of sensory information which are sampled from same communications. As a result, it is suggested that the superiority of emotional perception affects the developmental differentiation of emotion in infancy.

VII. DISCUSSION AND CONCLUSION

The superiority of emotional perception comes from the nature of C-fibers in human skin. There are two types C-fibers, which are classified by different stimuli to activate them. The first ones are activated by noxious stimuli such as pain stimuli, thermal stimuli, and chemical substances. The second ones are activated by light strokes whose velocity is lower than 10 [cm/s]. Several studies have suggested the importance of C-fibers in emotional recognition. Patients with Congenital Insensitivity to Pain with Anhidrosis (CIPA) are known to inherently lose the first types of C-fibers. They have insensitivity to pain due to the lack of Aß-fibers as well as C-fibers and therefore suffer from emotional impairments. Their weaker reaction to noxious stimuli produces difficulty in estimating others’ emotional state [25]. There is a similar symptom called Congenital Insensitivity to Pain (CIP). Patients with CIP do not have Aß-fibers but have C-fibers. Unlike patients with CIPA, they can recognize others’ emotional states thanks to the sensitivities of C-fibers to noxious stimuli. The influence of CIPA on emotional functions has also been suggested by Indo et al [26]. Our study empirically demonstrates the importance of C-fibers in emotional development. The experimental condition in which the pleasant/unpleasant information was not given simulates CIPA. Our results explain what CIPA causes in emotional differentiation, which provides new insight into the role of tactile interaction.

Our proposed model receives facial images as information of the visual input in the experiment. The facial expression is a very persuasive way to express the emotion. The amygdala and the superior temporal sulcus (STS) are known as brain regions related to recognition of emotional facial expressions [27] [28]. Especially, the amygdala has been studied by many researchers to elucidate the relation to unpleasant emotions. For example, a monkey usually expresses the fear emotion when the monkey meets snakes. It is known that the monkey which is researcted the amygdala does not express the emotion of fear for snakes [29]. The amygdala has connections with the STS [28]. Sugase et al. [30] suggested the hierarchically information of facial images which are represented by the change of nervous activities in the STS. The hierarchically information of facial images means the difference of individuals, the facial expression of emotion, etc. Furthermore, it is suggested that the STS processes not only the recognition of facial expressions, but also the recognition of biological motion and the processing of the sensory information relative
to emotions [31].

The knowledge of these results supports our model which reappears brain functions in terms of integration of emotional information from tactile sensor and abstract features of visual and auditory information. It is considered that neurons which are suggested by Sugase et al. [30] are developmentally acquired. Fortunately, the proposed model is able to acquire features of sensory input to increase the hierarchies of sensory feature modules. However, the model should consider the relationship with other regions of the human brain (e.g. the insular cortex, the pulvinar, the frontal lobe, etc.) and the endocrine system.

We hypothesize that the tactile dominance leads the developmental differentiation of emotion. Our proposed model is experimented under two conditions: the condition of the full tactile dominance and the condition without the superiority of emotional perception. In the first experiment under the condition of the full superiority shows that our model is able to represent the development of differentiation of emotion to learn sensory information like the infant-caregiver interaction. Under the condition which lose the superiority of emotional perception, the model does not clearly differentiate emotions. As a result, it is verified that the proposed model with the tactile dominance leads to better the emotional differentiation than the model without such dominance.

For future issues, our model will be extended to incorporate mutual effects of infant actions. For example, movement of a infant who is added stimuli will be changed infant emotions. Infants are able to cancel unpleasant emotions by own actions. How motor development which influences emotional development and vise versa is most interesting question. We assumed that emotions motivate actions and perceptions. In addition, we implement the extended model for robots to learn in the real-interaction. Relations with caregivers in interaction would change the estimated emotion of others to use learned model, which is another issue to be approached.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Numbers 24000012, 24119003, and 24650083.

References


