A Human Robot Interaction by a Model of the Emotional Learning in the Brain

Satoshi Sonoh, Shuji Aou, Keiichi Horio, Hakaru Tamukoh, Takanori Koga, Takeshi Yamakawa

Abstract:

In this paper, we proposed an emotional expression system as a brain-inspired system. The emotional expression was achieved by an Emotional expression Model of the Amygdala (EMA), which was an engineering model inspired by an emotional learning in the brain. EMA can realize both recognition of sensory inputs and a classical conditioning of emotional inputs. Furthermore, a specific hardware of EMA was developed with a massively parallel architecture by using an FPGA, and achieved a calculation speed that is over 20 times faster than an embedded general-purpose computer. Finally, we confirmed an effectiveness of a human-robot interaction with the emotions, which were generated by the proposed emotional expression system.

Keywords: *emotional learning, amygdala, classical conditioning, human-robot interaction.*

1. Introduction

Recently, emotions are introduced into intelligence robots [1]. An expression of the emotion has advantageous effects on a communication between a human and a robot, as well as behaviors of the robot. The communication through the expression of the emotions is a non-verbal and intuitive for the human. For example, internal states of the robot, which are not usually visualized, can be communicated to the human by using facial expressions based on the emotions that the robot generates. As conventional methods for the expression of the emotion, artificial intelligence models have emulated some emotions of the robots [2], [3]. However, the expression of the emotion must be programmed with defined inputs and certain situations in advance, and it is not an emergence as a result of interactions between unknown environments. Therefore, a novel model, which can generate wide variety of emotions by learning in the unknown environments, has growing requirements to improve the human robot interaction.

Emotional-expression Model of the Amygdala (EMA) has been proposed as an artificial neural network of the amygdala from a viewpoint of an engineering approach [4]. EMA has been established based on neuroscience findings of the amygdala that is an emotional learning system in the brain [5]. The learning of the emotions by EMA is interactively achieved by both recognition and a classical conditioning of inputs from environments. Furthermore, EMA is suitable for applications to the robot, because it has superior recognition abilities compared to other models of the emotion [6], [7].

In this paper, we apply EMA to the expression of the emotion for an autonomous mobile robot as a brain-inspired system. First, we demonstrate an effectiveness of EMA using a robot simulator. Furthermore, we develop an accelerator of EMA, which allows real-time interactions, using FPGA. Finally, we implement EMA in the robot as an emotional expression system including sensors, expression-devices and the accelerator of EMA. An effectiveness of the developed system is confirmed by an interactive training of the expression of the emotion between the human and the robot.

2. Emotional-expression model of the amygdala

2.1. Amygdala

A limbic system of the brain is an information processing system involved in an emotion and a memory. The amygdala, which is a part of the limbic system, involves in the emotional learning. The amygdala receives various sensory stimuli from an inside and an outside of the body *via* a sensory thalamus [5]. The sensory stimuli are integrated in a lateral nucleus of the amygdala (LA) and are localized and recognized based on the characteristics. Furthermore, a value of the stimulus is evaluated for corresponding emotions in a central nucleus of the amygdala (CE). As a consequence, emotional reactions, such as freezing and stress-hormone release, arise in whole of the body as emotional responses.

A relationship between the sensory stimulus and the emotional responses is acquired by a classical conditioning by using the sensory stimulus and the emotional stimulus [8]. The sensory stimulus, such as an auditory stimulus, acts as a conditioned stimulus (CS), and is naturally irrelevant to the emotional responses. On the other hand, the emotional stimulus, such as an electrical shock, is called an unconditioned stimulus (US) since the stimulus potentially generates the emotional responses. The classical conditioning is achieved by simultaneously presenting CS and US. After the conditioning, the emotional responses are induced by observing CS only. The amygdala is related to the conditioning with a fear emotion in particular.

2.2. Architecture of EMA

EMA emulates two essential functions in complex and diverse functions of the amygdala, recognition of the sensory stimulus and conditioning to the emotional response. Two functions are absolutely imperative in the emotional human robot interaction. It is preferable that the recognition system is self-organized thorough the interaction. Furthermore, the recognition should be

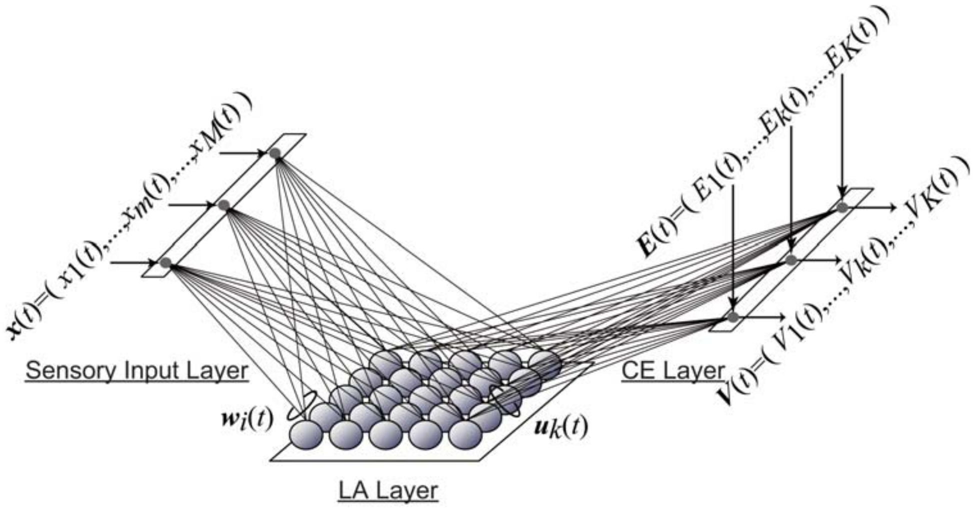

Fig. 1. Architecture of EMA. EMA is inspired by anatomical findings of the amygdala and consists of three layers, the sensory input layer, the LA layer and the CE layer.

adaptive to environmental changes. In order to realize interactive and adaptive recognition system, we adopted Self-Organizing Map (SOM) [9] and its adaptive learning rule in EMA.

The EMA architecture to satisfy two functions is inspired by the anatomical findings of the amygdala. The architecture is three-layers to integrate the sensory stimuli shown in Fig. 1. The sensory input layer has several input units, and corresponds to an entrance area of the amygdala including the sensory thalamus. The LA layer has a number of competitive units, and receives the sensory stimulus. The competitive units are arranged in a two-dimensional array, and can extract characteristics of the sensory stimulus as a feature map in a self-organizing manner. Finally, emotional values, which represent strength of the emotional responses, are evaluated in the CE layer. Several types of emotions, such as fear, pleasure and surprise, are available in EMA although the amygdala is specifically related to the fear emotion.

2.3. Algorithm of EMA

Let $\mathbf{x}(t)=(\mathrm{x}_1(t),\ldots,\mathrm{x}_m(t),\ldots,\mathrm{x}_M(t))$ be an input vector that represents the sensory stimulus at time step t, and $\mathbf{w}_i(t)=(\mathrm{w}_{1,i}(t),\ldots,\mathrm{w}_{m,i}(t),\ldots,\mathrm{w}_{M,i}(t))$ be a reference vector of the *i-th* unit on the LA layer. The best matching unit (BMU) for the sensory stimulus is selected by the following equation

$$BMU(t) = \arg\min_i (error_i(t)) = \arg\min_i (\|\mathbf{x}(t) - \mathbf{w}_i(t)\|^2). \quad (1)$$

EMA regards the BMU as a classified CS of the sensory stimulus. The reference vectors are updated by the following equations.

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + \alpha(t) h_{i,BMU(t)} (\mathbf{x}(t) - \mathbf{w}_i(t)) \quad (2)$$

$$h_{i,BMU(t)} = \exp\left(-\frac{D(i, BMU(t))^2}{2\sigma(t)^2} \right) \quad (3)$$

Here, $h_{i,BMU}(t)$ means a neighborhood function and $D(\mathrm{i},\mathrm{BMU}(t))$ means a distance function between the *i-th* unit and BMU on the LA layer. $\alpha(t)$ and $\beta(t)$ are the learning ratio and the neighboring width at time step t, respectively. These parameters in EMA are determined based on an adaptation degree to the sensory stimulus, because an adaptive learning is significant for an interactive learning between the human and the robot. The adaptive learning is achieved by the following equations [11].

$$\alpha(t) = \frac{error_i(t)}{error_{\max}} \quad (4)$$

$$\sigma(t) = \max(\alpha(t)\sigma_{\max}, \sigma_{\min}) \quad (5)$$

Here, $error_{max}$ is the maximum error until time step t from the initial state, $\alpha(t)$ means a normalized error as the adaptive degree. σ_{max} and σ_{min} are the maximum and the minimum neighboring widths, respectively. These equations mean that the learning ratio and the neighboring width increase when the normalized error is large, but these decrease when the normalized error is small.

Furthermore, a relationship between the sensory stimulus and the emotional stimulus is acquired by using a learning model of the classical conditioning [10].

Let $\boldsymbol{E}(t)=(E_1(t),\ldots,E_k(t),\ldots,E_K(t))$ be an input vector that represents strength of the emotional stimulus, $\boldsymbol{V}(t)=(V_1(t),\ldots,V_k(t),\ldots,V_K(t))$ be an output vector of the emotional value that represents strength of the generated emotional responses at the time step t, where the suffix k means an index corresponding to a kind of the emotion. The emotional value is calculated by the following equations, when the sensory stimulus is presented.

$$\boldsymbol{V}(t) = \sum_t act_i(t)\boldsymbol{u}(t), \quad (6)$$

$$act_i(t) = \frac{h_{i,BMU(t)}}{\sum_j h_{i,BMU(t)}}. \quad (7)$$

Here, $act_i(t)$ is a normalized activity of the *i-th* unit to the sensory stimulus and $\boldsymbol{u}_i(t)=(u_{1,i}(t),\ldots,u_{k,i}(t),\ldots,u_{K,i}(t))$ is an emotional weight of the *i-th* unit. The emotional weight is updated by the following equation.

$$\boldsymbol{u}_i(t+1)=\boldsymbol{u}_i(t)+\delta act_i(t)(\boldsymbol{E}(t)-\boldsymbol{V}(t)). \tag{8}$$

Here, δ is a conditioning ratio. The algorithm of EMA is summarized two computational processes; (1) the stimulus recognition process and (2) the conditioning process. Recognition of the sensory stimulus can be achieved in the stimulus recognition process. In parallel, a prediction and an update of the emotional value can be achieved in the conditioning process.

In application such as human robot interaction, an advantage of EMA over other classical conditioning models, for example TD model [11], is the self-organizing and adaptive recognition of the sensory stimulus (See [4]). We confirm an effectiveness of EMA by software simulations and experiments with developed hardware of EMA in the following section.

3. Computational simulations

3.1. Simulation environment

An experiment of the expression of the emotion for a dog-like robot is performed by computational simulations. A simulation environment is created by a robotics simulator "Webots" [12], and is shown in Fig. 2. The simulated robot (SONY AIBO ERS210) in the environment has a vision sensor and a touch sensor. The robot can detect a color intensity of the front ball as the sensory stimulus by using the vision sensor. Furthermore the robot sometimes receives tactual stimulus such as hitting and gentle stroking from the environment, when the sensory stimulus is presented. Here, we assume that the tactual stimulus induces a corresponding emotional response to the robot as the emotional stimulus. EMA implemented in the robot performs the recognition and the conditioning for the emotional learning like the amygdala.

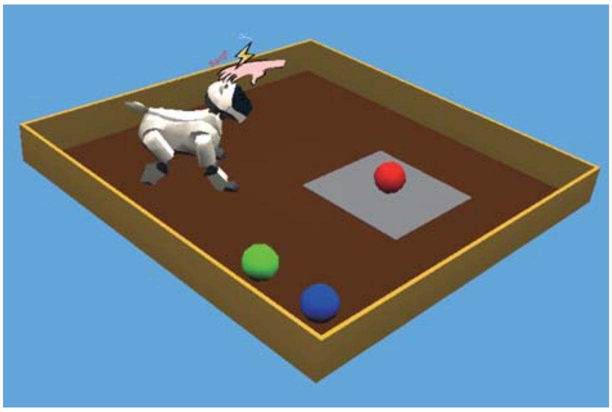

Fig. 2. A simulation environment. The robot has a vision sensor to detect the sensory stimulus and has a touch sensor to detect the emotional stimulus. The sensory stimulus is a color of ball objects and the emotional stimulus is a tactual sense of the robot.

The sensory stimulus is represented as a three-dimensional vector $\mathbf{x}(t)=(I_R, I_G, I_B)$, where I_R, I_G and I_B are color intensities of red, green and blue, respectively. Strength of the emotional stimuli is represented as $\boldsymbol{E}(t)=(E_p, E_f)$. In this simulation, the parameters of EMA are as follows; the number of the LA unit is 64 (8 x 8), σ_{max}=1.0, σ_{min}=0.1, σ=0.3.

3.2. Recognition of the sensory stimulus

At the beginning, a performance of the stimulus recognition was confirmed. We randomly presented the color ball to the robot. Fig. 3(a) shows the feature map of EMA after presenting 1000 times. In the feature map, each color patch shows features of the sensory stimulus that are represented by the reference vector of the LA unit. The neighboring LA units have similar feature, and the distant LA units have different features, for For BMUR, BMUG and BMUB, R, G and B are subfix. Although uniformed feature map was obtained by the randomly sensory stimuli, we presented red-biased stimuli in order to confirm effectiveness of the adaptive learning rule, additionally 500 times. The feature map was changed and specialized in red feature by additional stimulus, as shown in Fig. 3(b). The feature map is updated depending on the adaptive degree to the sensory stimulus. Thus, the recognition of EMA works well even if the environment is dynamically changed, although the conventional model [11] cannot accommodate.

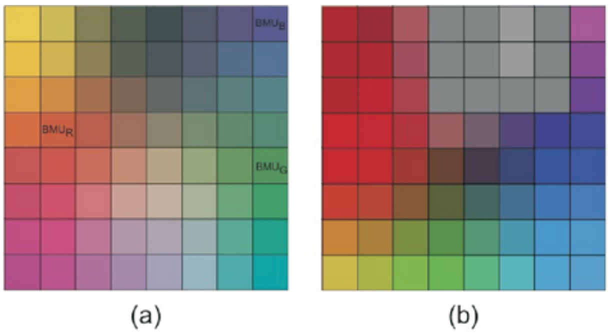

Fig. 3. A feature map in EMA obtained by the stimulus recognition process. Each color patch represents the reference vector of unit on the LA layer. (a) The obtained map by presenting 1000 random stimuli. (b) The map adapted additional stimuli (red color).

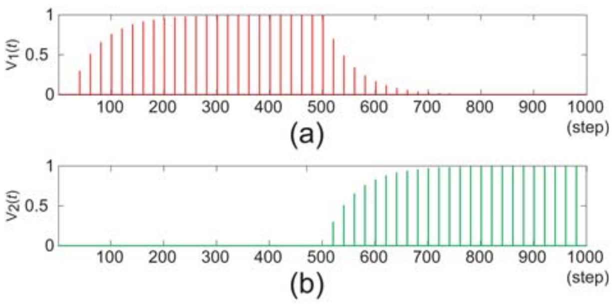

Fig. 4. Emotional values in the basic classical conditioning experiment. (a) The emotional value to fear emotion. (b) The emotional value to pleasure emotion. Emotional values represent strength of emotional responses to the sensory stimulus. EMA can achieve the acquisition and the extinction for more than one emotion.

3.3. Emotional conditioning

Next, the classical conditioning experiment was performed by simultaneously presenting both the sensory stimulus and the emotional stimulus. The emotional stimulus, $\boldsymbol{E}=(1,0)$, was associated with the sensory stimulus, $\mathbf{x}=(1,0,0)$, every 20 steps in the first 500 steps. In the next 500 steps, the emotional stimulus, $\boldsymbol{E}=(0,1)$, was associated with the same sensory stimulus every 20 steps. Fig. 4 shows the emotional values in the learning steps, where V_1 and V_2 correspond to fear and pleasure emotions, respectively. Emotional values were acquired by the classical conditioning. As a result, the emotional responses to the sensory stimulus were generated without the emotional stimulus. In the last of the conditioning, the emotional value (V_1) was eventually lost because the corresponding emotional stimulus was not presented. Thus, the "acquisition" and the "extinction", which are a basic principle of the classical conditioning, can be achieved in EMA.

In this simulation, it was confirmed that the robot with EMA was able to generate the emotions by a combination of the recognition and the conditioning. This contribution by EMA makes behaviors of robots more intelligent, and the human robot interaction becomes natural and interactive.

4. FPGA implementation of EMA

4.1. Architecture of EMA hardware

To realize real-time processing of EMA in robotics applications, we proposed specific hardware architecture of EMA in which a hardware-oriented algorithm is employed. The proposed EMA hardware (EMAHW) was developed based on a massively parallel architecture as well as conventional SOM hardware [14], [15] because the algorithm of EMA was based on SOM. Fig. 5 shows the massively parallel architecture of EMAHW including 81 units. The architecture is achieved by several local circuits and one global circuit. Each local circuit corresponds to one LA unit, and has a memory of the reference vector and the emotional vector. The global circuit calculates following processes; finding BMU, adapting the learning parame-

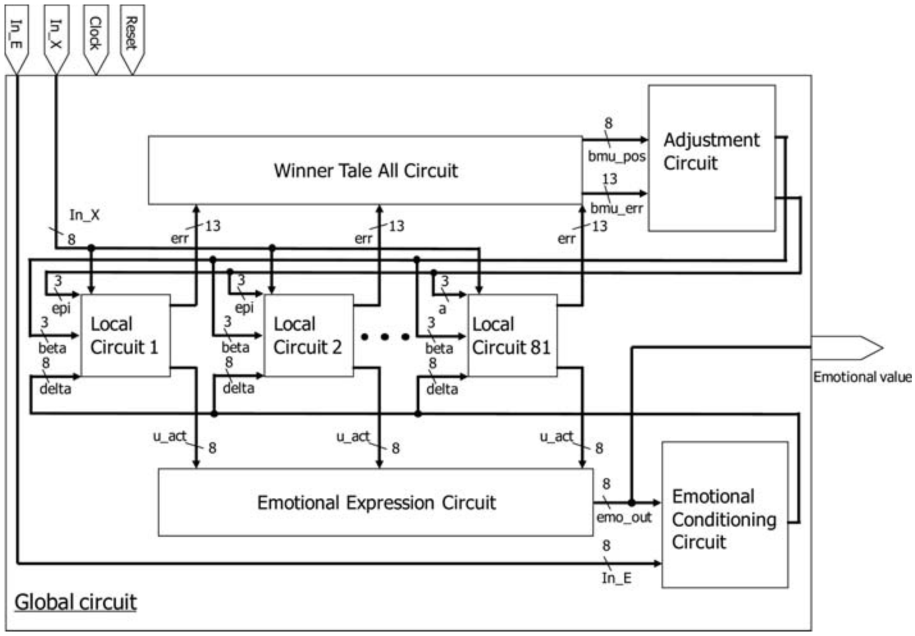

Fig. 5. A massively parallel architecture of the EMA hard-ware. The EMA hardware consists of several local circuits and one global circuit.

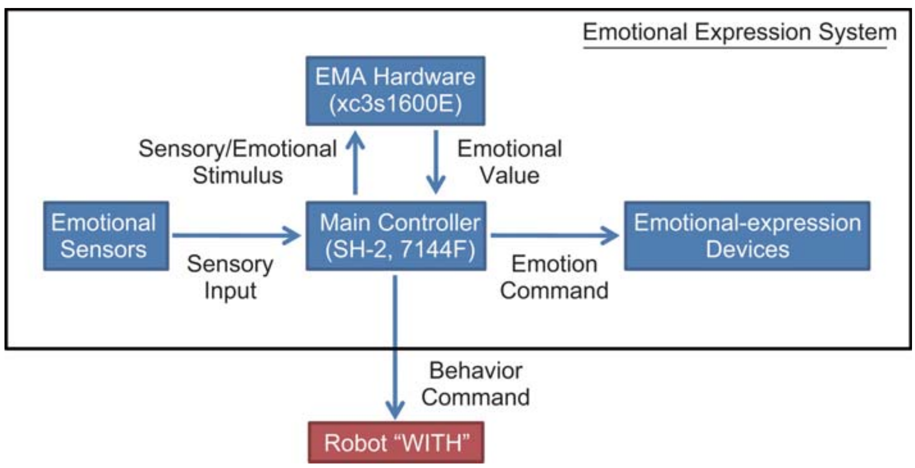

Fig. 6. A block diagram of the emotional expression system including the digital hardware of EMA, the emotional sensors and the emotional expression devices.

ters, calculating the emotional value and conditioning. Furthermore, the learning parameters (Eq. (3) and Eq. (4)) were modified in EMAHW as follows;

$$h_{i,BMU(t)} \in 2^{-n} \quad (n = 0,1,\ldots,N)\,, \tag{9}$$

$$\alpha(t) \in 2^{-n} \quad (n = 0,1,\ldots,N)\,. \tag{10}$$

These learning parameters allow to EMAHW use shift register as a substitute for multiplier. Thus, it drastically reduces a circuit area and calculation cost.

4.2. Performance of EMA hardware

EMAHW was implemented on FPGA (Xilinx Spartan-3E, sc3s1600E) with 81 LA units and 8-bit accuracy. EMAHW calculated at clock frequency up to 50 MHz. A performance of EMAHW was estimated at 632.8 MCUPS (Million Connection Update Per Second), when the reference vectors were three dimensions and the emotional weights were two dimensions. For comparison, a performance of a general-purpose PC (Intel Core 2 Duo, 2.66GHz) is 43.2 MCUPS in the original algorithm.

Furthermore, a performance of EMAHW was confirmed by comparison with embedded processors that were available an autonomous mobile robot. A benchmark test using three-dimensional 1000 sensory stimuli was performed by each processor. Table 1 shows a comparative result. The calculation speed of the EMAHW was twenty-times or more as fast as the portable PC in spite of the lowest clock frequency.

5. Human robot interaction with EMA

5.1. Robot with emotional expression system

To verify an effectiveness of EMA in the human robot interaction, an emotional expression system was implemented in an autonomous mobile robot "WITH" [16]. A block diagram of the emotional expression system is shown in Fig. 6, and the robot including the sensors and the devices is shown in Fig. 7.

The emotional sensors include a CMOS camera to capture front images of the robot, and a capacitance sensor array to detect tactual senses from the human to the robot. The main controller receives the captured image from the CMOS camera, and sends an averaged the color value as the sensory stimulus to EMAHW. Furthermore, the main controller estimates tactile information, hitting or gentle stroking, by using the number of responded sensors, and sends the information as the emotional stimulus to EMAHW. EMAHW calculates the emotional value by using the sensory and emotional stimulus provided. The emotional-expression devices were developed based on an ear and a tail of a dog in order to communicate the robot's emotions to human. The dog is the most famous pet, and their emotional expressions have been investigated in detail. The robot's emotions, which are generated by EMAHW, are expressed by simple movements of the emotional-expression devices. For example, the ear is laid back and the tail is wagged in small motions, when the robot is feeling a fear emotion. In addition, pleasure, confusion and attention emotions have been elaborated.

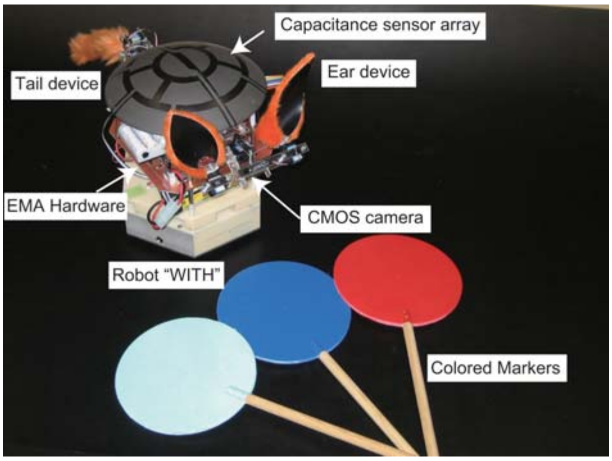

Fig. 7. The robot "WITH" equipped the emotional expression system.

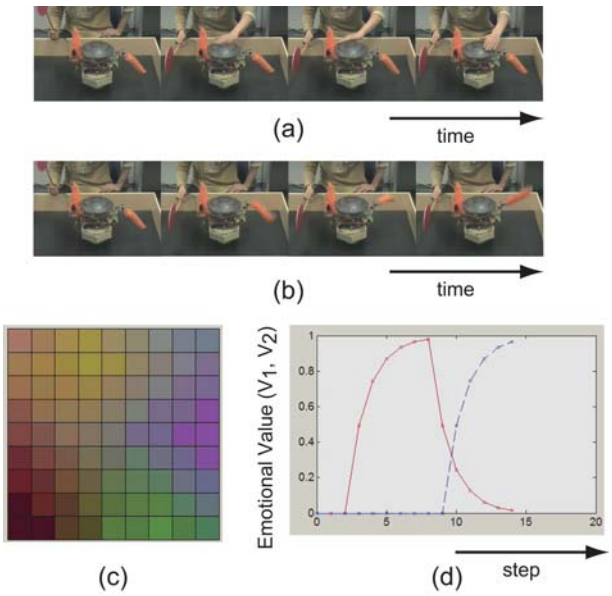

Fig. 8. The experimental results of the human robot interaction. Here, the colored marker is the sensory stimulus as CS and the touch to the robot is the emotional stimulus as US. (a) A scene that the human is training the robot with the red maker and gentle stroking. (b) A scene that the robot is expressing the pleasure emotion to the only red marker. (c) The feature map obtained by the interaction. (d) Emotional values acquired by the interaction.

Table 1. A comparison of calculation time to benchmark test between EMAHW and other embedded processor.

	EMA Hardware	Microprocessor	Portable PC
Specification	Xilinx Spartan-3E FPGA xc3s1600E, 35MHz	Renesas Microprocessor SH-2 7144F, 50MHz	Intel CPU Pentium M, 1.1GHz
Calculation Time (ms)	0.68	7500	15.31

5.2. Experiment in human robot interaction

The emotional expression system including EMAHW was implemented in the robot as a robot's amygdala system. As a result, the robot got to recognize the sensory stimulus from the environment, and to generate the emotional value and the expression of the emotion. The human robot interaction was perfumed in the real environment. In the interaction, the human as a trainer presented a colored-marker as the sensory stimulus, and touched with one's hands as the emotional stimulus to the robot. Here, the sensory stimulus as CS is a presentation of colored-marker, and the emotional stimulus as US is a touch of the robot with gentle stroking or hitting. At the begging, the robot unconditionally responded using the specific emotion to the touch, but the colored-marker did not induce any emotions.

Fig. 8 (a) shows a scene that the human was training the robot with the red marker and the gentle stroking. As the interaction was repeated, the robot became to express the emotion to the red marker only. Fig. 8 (b) shows a scene that the robot was expressing pleasure emotion using the tail wagging to the red marker without the touches after the training. Fig. 8 (c) shows the obtained feature map in EMAHW. The feature map preserves the topology of the sensory stimulus as well as the result of the computational simulation. Fig. 8 (d) shows the emotional value corresponding to the interaction step. The red and blue lines represent the emotional value of fear and pleasure emotion, respectively. In the real environment, the acquisition and the extinction can be successfully achieved.

In the human robot interaction experiment, the robot recognized the sensory stimulus, and predicted the emotional value as well as animals. The emotional expression of the robot makes the interaction to the human more intelligent and human-friendly.

6. Conclusion

In this paper, we implemented EMA in the simulated and real robot in order to realize the expression of the emotions from the interaction. The expression of the emotions can be achieved by the recognition of the sensory stimulus and the classical conditioning in EMA. Furthermore, we proposed the emotional expression system including the accelerator of EMA. The EMA hardware has the computational ability, which is 20 times or more fast than other embedded processors. The human robot interaction with the expression of the emotion can be achieved in simulated and real environment. In the future work, a system that estimates a stimulus involving the expression of the emotion by consideration of internal and contextual status of the robot is needed as an extension of EMA. Then, we believe that the brain inspired system will achieve a breakthrough in the human robot interaction.

ACKNOWLEDGMENTS

This work was supported by a 21st Century Center of Excellence Program, "World of Brain Computing Interwoven out of Animals and Robots (PI: T. Yamakawa)," granted in 2003 to the Department of Brain Science and Engineering (Graduate School of Life Science and Systems Engineering), Kyushu Institute of Technology by Japan Ministry of Education, Culture, Sports, Science and Technology. This work is also supported by Grant-in-Aid for Scientific Research (A), 2006, 18200015.

AUTHORS

Shuji Aou, Keiichi Horio, Takeshi Yamakawa - Department of Brain Science and Engineering, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu-shi, 808-0196 Japan. E-mails:
aou@brain.kyutech.ac.jp,
horio@brain.kyutech.ac.jp,
yamakawa@brain.kyutech.ac.jp.
Satoshi Sonoh*- Corporate Research & Development Center, Toshiba Corporation, 1 Komukai-Toshiba-cho, Saiwai-ku, Kawasaki-shi, JAPAN. E-mail:
satoshi.sonoo@toshiba.co.jp.
Hakaru Tamukoh - Institute of Symbiotic Science and Technology, Tokyo University of Agliculture and Technology, 2-24-16 Nakamachi, Koganei-shi, 183-8509 Japan. Email: tamukoh@cc.tuat.ac.jp.
Takanori Koga - Fuzzy Logic Systems Institute, 680-41 Kawazu, Iizuka-shi, 820-0067, Japan. E-mail:
koga@flsi.or.jp.
* Corresponding author

References

[1] Breazeal C.L., *Designing Sociable Robots*, Cambridge, MA: The MIT Press, 2002.
[2] Fujita M., Kuroki Y., Ishida T., Doi T., "Autonomous Behavior Control Architecture of Entertainment Humanoid Robot SDR-4X". In: *Proc. of 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2003, pp. 960-967.
[2] Sawada T., Takagi T., Fujita M., "Behavior Selection and Motion Modulation in Emotionally Grounded Architecture for QRIO SDR-4X II". In: *Proc. of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2004, pp. 2514-2519.
[4] Sonoh S., Horio K., Aou S., Yamakawa T., "An Emotional Expression Model Inspired by the Amygdala", *International Journal of Innovative Computing, Information and Control*, vol. 5, 2009, no. 5, pp. 1147-1160.
[5] Aggleton J.P., *The Amygdala: A Functional Analysis*, New York: Oxford University Press, 2000.
[6] Armony J.L., Servan-Schreiber D., Cohen J.D., LeDoux J.E., "An anatomically constrained neural network model of fear conditioning", *Behavioral Neuroscience*, vol. 109, 1995, no. 2, pp. 246-257.
[7] Mor'en J., Balkenius C., "A computational model of emotional learning in the amygdala". In: *Proc. of the 6th International Conference on the Simulation of Adaptive Behavior*, 2000, pp. 383-391.
[8] Phelps E.A., LeDoux J.E., "Contributions of the amygdala to emotion processing: from animal models to human behavior", *Neuron*, vol. 48, 2005, pp. 175-187.
[9] Kohonen T., *Self-Organizating Maps*, Berlin: Springer-Verlag, 1997.
[10] Rescorla R.A., Wagner A.R., "A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement", *Classical conditioning II:*

current research and theory, New York: Appleton-Century-Crofts, 1972, pp. 64-99.
[11] Sutton R.S., "Learning to predict by the methods of temporal difference learning", *Machine Learning*, vol. 3, 1988, no 1, pp. 9-44.
[12] Cyberbotics Webots "*http://www.cyberbotics.com/*"
[13] Berglund E., Sitte J., "The parameterless self-organizing map algorithm", *IEEE Transaction on Neural Networks*, vol. 17, 2006, pp. 305-316.
[14] Tamukoh H., Horio K., Yamakawa T., "Fast Learning Algorithm for Self-Organizing Map Employing Rough Comparison WTA and its Digital Hardware Implementation", *IEICE Trans. on Electronics*, vol. E87- C, 2004, no. 11, pp. 1787-1794.
[15] Hikawa H., "FPGA implementation of self organizing map with digital phase locked loops", *Neural Networks*, vol. 18, 2005, pp. 514-522.
[16] Takemura Y., Sato M., Ishii K., "Toward Realiation of Swarm Interlligene Mobile Robots", *Brain-Inspired IT II*, Amsterdam: Elsevier B.V., 2006, pp. 273-276.