2008 IEEE/RSJ International Conference on Intelligent Robots and Systems
Acropolis Convention Center
Nice, France, Sept, 22-26, 2008

A Robotic KANSEI Communication System Based on Emotional Synchronization

Tatsuya Usui, Kazuomi Kume, Misaki Yamano and Minoru Hashimoto, *Member, IEEE*

Abstract - **Human-robot communication is an important subject for housekeeping, elderly care and entertainment robots. To make a natural communication entrainment between human and robot emotions plays a vital role. From this view point we propose to make a KANSEI communication system based on emotional synchronization.**

We developed a robotic KANSEI communication system using entrainment of dynamics, and its effectiveness was examined by experiments of human-robot communication. The robotic emotion was entrained to human emotion by using a vector field of dynamics. The robotic facial expression using a communication robot "Kamin_FA1" was realized dynamically based on the robotic emotion. In the experiment of communication, the human impression was changed by the strength of synchronization of robot. Then we confirmed that this method could utilize human-robot communication to keep a comfortable state.

I. Introduction

In an aging society it is necessary that robots work for housekeeping and elderly care at home and hospital. Such assisting robots have to communicate with humans. Then, robots purposely designed for communicating with humans have attracted our attention. These robots are indispensable for human-robot symbiosis in the near future and need to have not only intelligence but also KANSEI to make natural communication. KANSEI is cerebral activity without logical thinking such as emotion, feeling, impression, sensitivity, intuition, and so on. Although much research has been done in artificial intelligence, there has been no consideration about KANSEI in the scientific field, since KANSEI is a subjective process. It is important to consider KANSEI in the field of communication robots.

It is well known that facial expressions play a very important role in daily communication. In order to familiarize the robot with human society, it is essential to create affinity with facial expressions. From this perspective, we have developed a head robot called Kamin-FA1 (KAnsei MINd robot) [3] to enhance emotional communication with humans. The robot has a facial expression function using a curved surface display. This technique provides a facial expression easily compared with other methods of mechanical facial expression.

We may change the atmosphere of interaction by adjusting our emotion to the other accordingly. For example, the interaction field becomes lively, if a robot makes a pleased reaction when the partner is pleased. On the other hand, if we do not make any reaction, the field does not swell very much. Jonsson et al. [1] found that matching the car voice to the drivers' emotions had enormous consequences. Drivers who interacted with voices that matched their own emotional state had less than half as many accidents on average as drivers who interacted with mismatched voices. Therefore, the emotional synchronization in human-robot interaction is important to swell the interaction state.

In this paper, we propose to make a KANSEI communication system based on the emotional synchronization. The robotic emotion was entrained to human emotion by using a vector field of dynamics. The robotic facial expression using a communication robot "Kamin_FA1" was realized dynamically based on the robotic emotion. The ultimate purpose of this study is to find the interaction technique which makes a comfortable state through adjusting the communication field. If this study is successful, this technique will be widely utilized for human–robot communication to effect human emotional state.

In the next section, we address the outline of KANSEI communication system based on the emotional synchronization. In section III, the emotional synchronization is discussed in detail. We explain the emotional recognition of human voice. The dynamics design for emotional synchronization is also explained. In section IV, we describe the facial expression of Kamin_FA1. In section V, Kamin_FA1 and the communication system are explained. We conduct the experiment in section VI. Finally, in chapter VII we conclude this paper.

II. Outline of Kansei Communication System

The structure of the proposed communication system is shown in Fig.1. The system consists of recognition, emotion generation and expression parts. In the recognition part we use the voice analysis software, "RobEsense" made by Nemesysco Ltd. [11], to recognize the emotions from human voice. In the emotion generation part, the robotic emotion is determined by the emotional entrainment using human emotion. The emotional entrainment is performed by a vector

Manuscript received February 22, 2008.
T.Usui, K.Kume and M.Yamano are graduate students in Shinshu University, Japan.
M.Hashimoto is a professor in Shinshu university, Japan (corresponding author to provide phone: +81-268-21-5524; fax: +81-268-21-5524; e-mail: hashi@shinshu-u.ac.jp).

field of dynamics according to the strength of the synchronization. We use an online design method of dynamics to realize the synchronization between human and robotic emotions [2]. The robotic emotion is mapped to an emotional symbol space, and then the vector filed of a expression space is generated according to the state of the emotional symbol space. The expression space generates the facial expression of Kamin_FA1 based on the state of the characteristics quantity. These information processing systems are performed continuously by using the vector fields of dynamics, since the emotion and the expression of the robot is constantly changing.

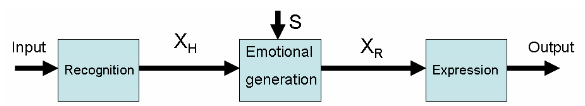


Fig.1 KANSEI communication system based on emotional synchronization

III. Emotional Recognition and Generation Spaces

The voice analysis software "RobEsense" [11] in the recognition part is a tool judging the emotional state from human voices. It is used in a call center and for a diagnosis of depression mainly. It analyses the emotion of human voice using 18 parameters which is obtained by data with high speed sampling. The result of voice analysis does not depend on language and gender.

We use two parameters, "Excited" and "Atmosphere", within 18 parameters of related emotions. Russell's circumplex model of emotion [7] is utilized to express human and robotic emotions in the recognition part. This model can represent the emotions based on a two dimensional space. One of the axes in the space expresses "comfort-discomfort," and the other is "rouse-sleep" as shown in Fig.2. Using the numerical values of the parameters "Excited" and "Atmosphere" in "RobEsense" we express the emotional state based on Russell's model. In addition, we form a vector field of dynamics on the two dimensional space to realize the entrainment between the human and robotic emotions. In the vector field we make a basin of attractor near the position of human emotion to entrain the robotic emotion. We call it the emotional generation space for the robot as shown in Fig.3. The human emotion obtained by "RobEsense" is plotted in the vector field, and an attractor is constructed in the vector field using the online-design method of dynamics. The attractor is updated continuously according to the result of the recognition part. In the on-line design method, the oblivion and weighted parameters explained below, as determined by the strength of the entrainment in the emotional generation space.

In the next subsections, we describe the design of vector field which is used in the emotional generation space.

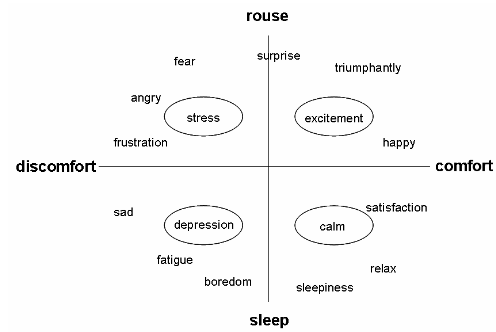


Fig.2 Russell's circumplex model of emotion

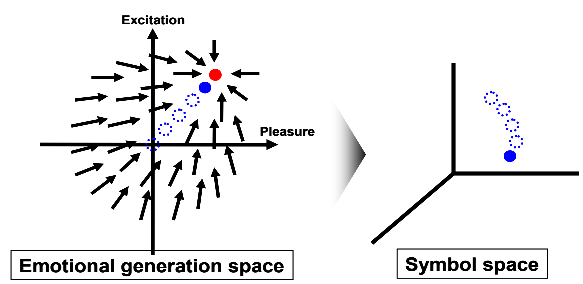


Fig.3 Emotional generation space and symbol space

A. *Polynomial expression approximation*

We need to design dynamics to have arbitrary attractors for information processing. As one of the design methods of dynamics, there is general RNN (Recurrent Neural Network) in a field such as sound recognition. Morita et al. [4] suggested a design method of attractor by a neural network with a non-monotonous function used for pattern recognition. However, they still need considerable time for learning. On the other hand, as a simple design method of dynamics to have arbitrary attractors, Okada et al. [5] proposed a polynomial expression approximation. They use dynamics to design the trajectory of the robotic motion. We apply this method to make synchronization in human-robot interaction. By using this method we can design a vector field around the attractor geometrically and approximate it by a polynomial expression. It is easy to design arbitrary attractors.

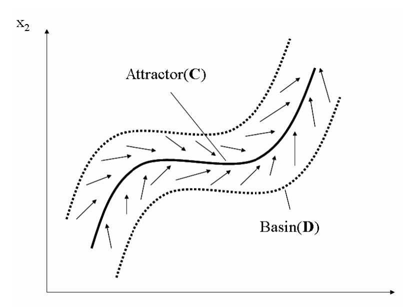


Fig.4: Attractor and basin of a vector field

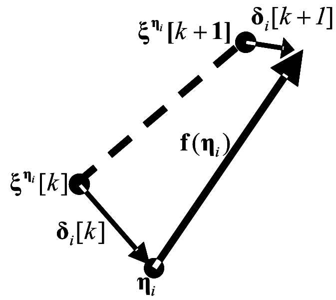


Fig.5: Defining the vector field

An example of a vector field in the two dimensional space is shown in Fig.4. The vector filed is formed around arbitrary curve C. The curve C is assumed as an attractor, and region D is a basin of entrainment around the attractor. The curve C is a function of discrete time *k* and consists of $\xi[k](k=1,2,\ldots)$ as shown in Fig.5. We decided the number of sample point among the basin of entrainment around the attractor which is described as $\boldsymbol{\eta}_i(i=1,2,\ldots,m)$, where m is the number of sample points. $\boldsymbol{\xi}^{\boldsymbol{\eta}_i}[k]$ are located on the attractor is the nearest point from $\boldsymbol{\eta}_i$, and $\boldsymbol{\delta}_i[k]$ is the connection vector between $\boldsymbol{\eta}_i$ and $\boldsymbol{\xi}^{\boldsymbol{\eta}_i}[k]$. Then, $\boldsymbol{\delta}_i[k+1]$ and $\boldsymbol{\delta}_i[k]$ can be defined as follows:

$$\boldsymbol{\delta}_i[k+1]=(\boldsymbol{\eta}_i+\mathbf{f}(\boldsymbol{\eta}_i))-\boldsymbol{\xi}^{\boldsymbol{\eta}_i}[k+1] \quad (1)$$

$$\boldsymbol{\delta}_i[k]=\boldsymbol{\eta}_i-\boldsymbol{\xi}^{\boldsymbol{\eta}_i}[k] \quad (2)$$

And, the sufficient condition for convergence is:

$$\|\boldsymbol{\delta}_i[k+1]\|<\|\boldsymbol{\delta}_i[k]\| \quad (3)$$

We make vectors $\mathbf{f}(\boldsymbol{\eta}_i)$ in the basin D by using equations (1)-(3), and approximate them by a polynomial expression. When $\boldsymbol{\eta}_i\in R^2$, the polynomial expression is as follows,

$$\begin{aligned}\mathbf{f}(\mathbf{x})&=a(50)x^5+a(41)x^4y+a(32)x^3y^2+a(23)x^2y^3+a(14)xy^4+a(05)y^5\\&+a(40)x^4+a(31)x^3y+a(22)x^2y^2+a(13)xy^3+(04)y^4\\&+a(30)x^3+a(21)x^2y+a(12)xy^2+(03)y^3\\&+a(20)x^2+a(11)xy+(02)y^2\\&+a(10)x+(01)y\\&+a(00)\end{aligned} \quad (4)$$

Where, $a(ij)$ are the constant parameters, and by using the vector $\boldsymbol{\eta}_i$, $\mathbf{f}(\boldsymbol{\eta}_i)$ can be described as:

$$\mathbf{f}(\boldsymbol{\eta}_i)=\boldsymbol{\Phi}(a(50)\,a(41)\cdots a(00))\boldsymbol{\theta}(\mathbf{x}) \quad (5)$$

$$\boldsymbol{\theta}(\mathbf{x})=[x^5\ x^4y\ x^3y^2\cdots 1]^T \quad (6)$$

We can determine $\boldsymbol{\Phi}$ by the least-squares method. $\mathbf{F}$ and $\boldsymbol{\Theta}$ can be defined by equations (7) and (8) :

$$\mathbf{F}=[\mathbf{f}(\boldsymbol{\eta}_1)\,\mathbf{f}(\boldsymbol{\eta}_2)\,\mathbf{f}(\boldsymbol{\eta}_3)\ \cdots\ \mathbf{f}(\boldsymbol{\eta}_m)] \quad (7)$$

$$\boldsymbol{\Theta}=[\boldsymbol{\theta}(\boldsymbol{\eta}_1)\,\boldsymbol{\theta}(\boldsymbol{\eta}_2)\,\boldsymbol{\theta}(\boldsymbol{\eta}_3)\ \cdots\ \boldsymbol{\theta}(\boldsymbol{\eta}_m)] \quad (8)$$

And,

$$\boldsymbol{\Phi}=\mathbf{F}\boldsymbol{\Theta}^{\#} \quad (9)$$

Where, $\boldsymbol{\Theta}^{\#}$ means a pseudo inverse matrix of $\boldsymbol{\Theta}$.

B. Online design method

After designing a vector field consists of *m* sample points, by adding one sample point, the sample points become to *m+1*. Appealing to *m+1* sample point, online design can be described as equations (10)-(15) [5]. If we know the values of $\mathbf{P}_m$, $\boldsymbol{\Phi}_m$, $\xi[k+1]$ and $\xi[k]$, and by determining a certain point in vector $\boldsymbol{\eta}_{m+i}(i=1,2,\cdots,l)$ which nears to $\xi[k]$, $\boldsymbol{\eta}_{m+i}$ we can define $\mathbf{f}(\boldsymbol{\eta}_{m+i})$ based on equation (1). By this way, the input signal is given as a new sample point, and the output signal is computed by the dynamics.

$$\hat{\mathbf{F}}=[\mathbf{F}\quad \mathbf{f}(\eta_{m+1})] \quad (10)$$

$$\hat{\boldsymbol{\Theta}}=[\boldsymbol{\Theta}\quad \boldsymbol{\theta}(\eta_{m+1})] \quad (11)$$

$$\boldsymbol{\Phi}=\hat{\mathbf{F}}\hat{\boldsymbol{\Theta}}^{\#} \quad (12)$$

$$\mathbf{P}_{m+1}=\mathbf{P}_m-\frac{\mathbf{P}_m\boldsymbol{\theta}(\eta_{m+1})\boldsymbol{\theta}^T(\eta_{m+1})\mathbf{P}_m}{1+\boldsymbol{\theta}^T(\eta_{m+1})\mathbf{P}_m\boldsymbol{\theta}(\eta_{m+1})} \quad (13)$$

$$\boldsymbol{\Phi}_{m+1}=\boldsymbol{\Phi}_m-\{\boldsymbol{\Phi}_m\boldsymbol{\theta}(\eta_{m+1})-\mathbf{f}(\eta_{m+1})\}\boldsymbol{\theta}^T(\eta_{m+1})\mathbf{P}_{m+1} \quad (14)$$

$$\mathbf{P}_m=\{\sum_{i=1}^{m}\boldsymbol{\theta}(\boldsymbol{\eta}_i)\boldsymbol{\theta}^T(\boldsymbol{\eta}_i)\}^{-1} \quad (15)$$

An oblivion parameter of online method is defined as α, and a weighted parameter for a new sample signal is defined as β. By using these two parameters equations (10)-(12) can be described as follows:

$$\hat{\mathbf{F}}=[\alpha\mathbf{F}\quad \beta\mathbf{f}(\eta_{m+1})] \quad (16)$$

$$\hat{\boldsymbol{\Theta}}=[\alpha\boldsymbol{\Theta}\quad \beta\boldsymbol{\theta}(\eta_{m+1})] \quad (17)$$

$$\boldsymbol{\Phi}=\hat{\mathbf{F}}\hat{\boldsymbol{\Theta}}^{\#} \quad (18)$$

When $0\le\alpha\le 1, 0\le\beta\le 1$, we can get the following equation:

$$\boldsymbol{\Phi}_{m+1}=[\alpha\mathbf{F}\quad \beta\mathbf{f}(\boldsymbol{\eta}_{m+1})][\alpha\boldsymbol{\Theta}\quad \beta\boldsymbol{\theta}(\boldsymbol{\eta}_{m+1})]^{\#} \quad (19)$$

IV. FACIAL EXPRESSION

Yamada [8] found the relationship between the basic emotional category and the structural variables of facial expression based on the displacement structure of the characteristics points in the facial expression. Yamada reported that there are two kinds of the facial structural variables called “Inclination” and “Bend”. “Inclination” means the displacement of the characteristic points concerning the degree of the slant of eyes and eyebrows. In the case of the mouth, it means the strength forming the shape of V and reverse V characters. “Bend” means the curvature of eyebrows, or the strength of the opening of the eyes and the mouth. The facial images of this model are shown in Fig.6. In this study, we referred to this facial expression model in order to make the robotic facial expression of Kamin_FA1.

We make a two dimensional space whose axes indicate “Inclination” and “Bend” to express the facial image. And then we build a vector field of dynamics on the two dimensional space to make the facial expression dynamically as shown in Fig.7. In this study we assume that the facial expression is not static, it is a dynamic process. The facial expression is modeled by an attractor on this vector filed. We designed the attractors of the basic six emotions based on Yamada’s facial expression model. The vector field of the facial expression is formed by the polynomial approximation described in section III.

We symbolized the facial expression space and make a symbol space. In the symbol space a point expresses an emotional state and a vector filed of the facial expression space. Therefore, we can change the vector field of the facial expression by the sate of the symbol space as shown in Fig.7.

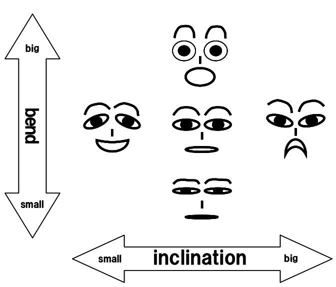


Fig.6 Facial expression depend on inclination and bend

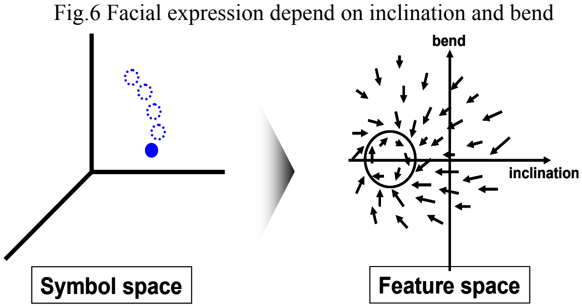


Fig.7 Symbolization of dynamics system

The facial image is made up of the line drawing, and it is determined numerically. The facial image consists of straight lines, bezier curves, and circles. A straight line can be created from specifying two points. Similarly a bezier curve is four points, and a circle is the center point and a radius. That is, the state of facial expression is specified with the parameter vector of the points and the radiuses. An example of face model is shown in Fig. 8.

An emotional state is expressed within the parameters of six basic emotions: happiness, sad, anger, fear, surprise, and disgust. The parameter vector of the facial image is computed from this emotional state. The state of the face model in six emotions is set up beforehand.

The expression model of the emotional state is made by AU (Action Unit) of FACS (Facial Action Coding System) as proposed by Ekman and Friesen [12]. The target state and the change speed of the model are computed from the present value and the change speed of the emotional state. The facial image is transformed based on the parameter and displayed. Moreover, the movement state of a head according to the emotional state is also determined simultaneously. Some kinds of movement are prepared beforehand. One of the movements is chosen from among them and then performed.

Using this method, the subtle expression between emotional states is possible by using the parameter vector. We modeled the parameter vector based on “Inclination” and “Bend”. Then we can determine the parameter vector by the coordinate in the facial expression space.

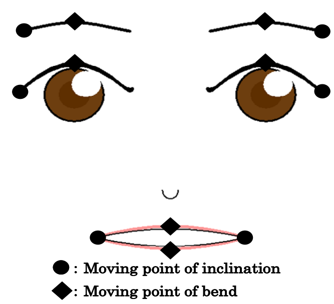


Fig.8 Facial model with control point

V. KAMIN_FA1 AND COMMUNICATION SYSTEM

We used the head robot Kamin_FA1 [3]. The head mechanism is a facial image display, and it consists of a dome screen, a fish-eye lens, and a projector. The face image is projected to the dome screen from the inside. The fish-eye lens is installed on the front of the projector, and it projects the picture on the dome screen. And by using the projector with a DLP (Digital Light Processing) system, the head can be made small and light.

The communication system configuration is shown in Fig.9. We analyze the human emotional state by “RobEsense” in PC1 and output the recognition result into the emotional generation part in PC2. Then, the computation of emotional entrainment is performed by using MATLAB. The result is used to make The facial expression of Kamin_FA1 is calculated in the facial expression space and is exhibited by the robot.

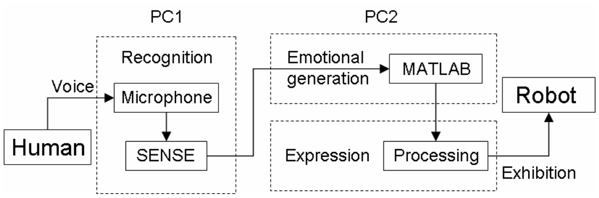


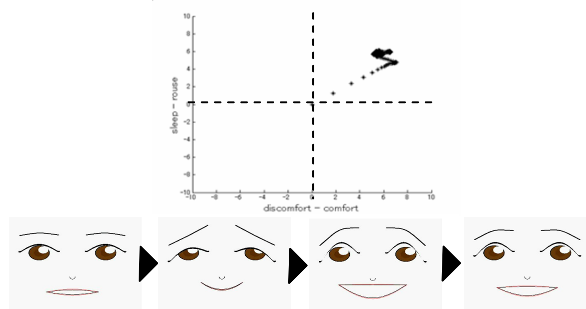
Fig.9 Block diagram of the communication system

VI. Communication Experiment

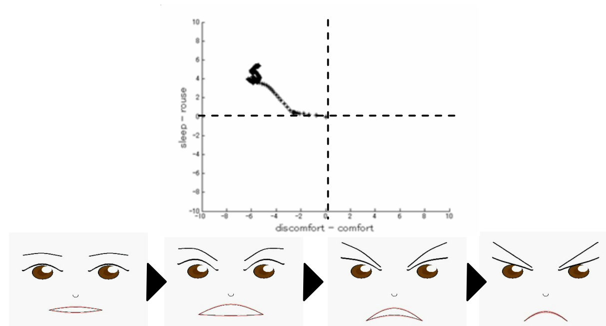
A. An Experiment Method

In order to examine the effectiveness of the proposed communication method, we conducted some experiments using Kamin_FA1. The overview of the communication experiment is shown in Fig.10. The subject sat on the position where was 100[cm] away from Kamin_FA1 and talked to the robot. The number of subjects was ten. The facial expression of Kamin_FA1 changed according to the subject's emotional state. The subject talked while watching a face of Kamin_FA1. In addition, we did not address the contents of communication to the subjects, they talked freely and free time. Each subject took the communication experiment twice. One of the experiments was conducted by using the strong emotional entrainment case, the other was the non entrainment case. In the case of the high entrainment, the facial expression of Kamin_FA1 changed much by reacting the subject's emotional state. Figure 11 shows the examples of happy and angry facial expressions. In the case of the non entrainment, the facial expression did not change so much. In addition, the subjects filled in a questionnaire based on a semantic differential (SD) method [10] after the communication to examine the impression of the subject. Subjects answered in 7-point scoring scale.


Fig.10 Overview of the communication experiment



(a) Happy



(b) Angry

Fig.11 Emotion change of robot and expression change

B. Expression change of the robot by the synchronization

The changes of the robotic emotion and expression during the communication expressions are shown in Figs.12 and 13. Figure 12 shows the results of the high synchronization case ($\alpha = 1, \beta = 1$ in the on-line design method of the vector field). It was shown that the robotic emotion and expression changed much along with the human emotional change during the communication. Figure 13 shows the result in the case of the un-synchronization ($\alpha = 1, \beta = 0$). In this case the changes of the robotic emotion and expression were poor.

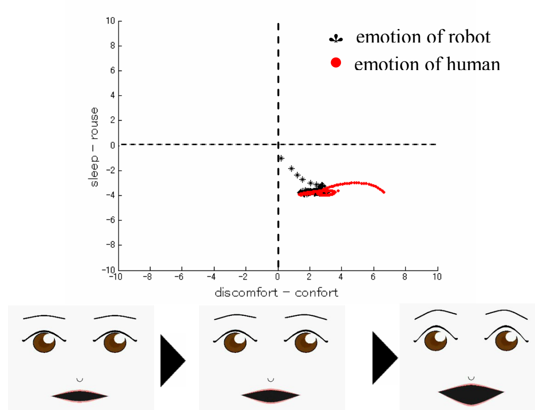


Fig.12 Emotion change of robot and expression change in synchronization

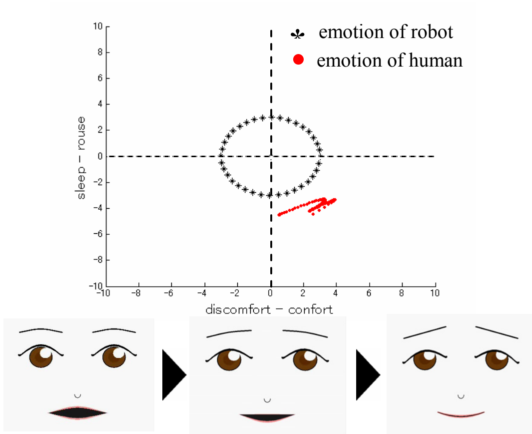


Fig.13 Emotion change of robot and expression change in un-synchronization

C. Relationship between emotion and synchronization

We compared the change of the human's emotion with the strength of synchronization of robot. Table I listed the communication period and the difference between maximum and minimum values of "Atmosphere" and "Excited" of subject's voice during the communication. In the case of synchronization the communication period is longer than that of un-synchronization case. Also, the changes of "Atmosphere" and "Excited" parameters were larger than that of the un-synchronization case. That is, the human emotional state becomes more comfortable in the case of high synchronization.

D. Evaluation of Impression

We examined the difference of the impression using SD method, when Kamin_FA1 expresses the synchronized and un-synchronized responses. The average rated values answered in the questionnaires are shown in Fig.14. It was found that in the case of high synchronization the subjects had positive impressions in terms of active, happy, friendly and so on comparing with the un-synchronization case.

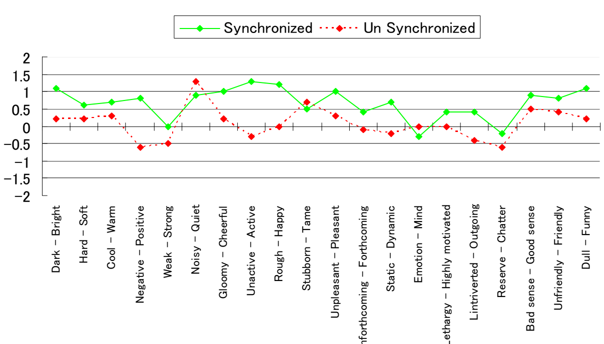


Fig.14 Result of psychological experiment

TABLE I Average period and amplitude

	Synchronization		Unsynchronization	
	Atmosphere	Excited	Atmosphere	Excited
Average period (s)	122.2		84.8	
Difference of max. and min.	6.4	5.8	7.7	6.6

VII. Conclusions

We developed a KANSEI communication system based on emotional expressions, and its effectiveness was verified by experiments in human-robot communication. The robotic emotion was determined by an entrainment to human emotion. The entrainment was accomplished using a vector field of dynamics. The robotic facial expression using a communication robot was realized dynamically based on the emotional space. The human impression was changed by the degree of synchronization of robot . The high emotional synchronization caused positive impressions in terms of active, happy, friendly, and so on.

As the future work, we are planning to investigate the effect of the emotional synchronization in long-term experiments.

Acknowledgment

This research was supported by The Mitsubishi Foundation and Toyota Tsusho Electronics Corporation. A.Prof. Honywood of Shinshu University provided critical assistance on our manuscript. We are grateful to them for their support.

References

[1] I.M.Jonsson, C.Nass, H.Harrris, L.Takayama: "Matching In-Car Voice with Driver State Impact on Attitude and Driving Performance", Proceedings of the Third Informational Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp173-180.

[2] T.Sato, M.Hashimoto and M.Tsukahara: "Synchronization Based Control using Online Design of Dynamics and Its Application to Human-Robot Interaction", Proceedings of the 2007 IEEE International Conference on Robotics and Biomimetics, Sanya, China, pp.652-657, 2007.,

[3] M.Hashimoto, D.Morooka : "Robotic Facial Expression Using a Curved Surface Display" , Journal of Robotics and Mechatronics, Vol.18, No.4 , pp.504-510, 2006.

[4] M. Morita and S. Murakami, "Recognition of Sequential Patterns by Nonmonotone Neural Networks," *The IEICE Transactions*, Vol. J81- D-II, No.7, pp.1679-1688, 1998.

[5] M.Okada, H.Nakamura: "Design of Continuous System Space and Whole Body Motion Generation using Dynamics-based Information Processing System", A society magazine of robot of Japan, Vol.23 No.7, pp.858~863, 2005.

[6] T.Inui, Y.Anzai: "Communication and thought", Iwanami shoten, pp.118~133, 2001.

[7] J.A.Russell: "A circumplex model of affect.", Journal of Personality and Society Psychology, Vol.39, PP.1161~1178,1980.

[8] H.Yamada: "Models of perceptual judgment of emotion from facial expressions", Japanese Psychological Review Vol.43, No.2, pp.245~255, 2000.

[9] M. Young: "*The Techincal Writers Handbook*", Mill Valley, CA: University Science, 1989.

[10] M.Inoue T.Kobayashi, "The research domain and constuction of adjective-pairs in a semantic differential method in Japan", J.of Educ.Phychol, 33, pp.253~260, 1985.

[11] http://www.nemesysco.com/

[12] P.Ekman and W.V.Friesen: Facial Action Coding Consulting, Psychologist Press, 1977.