The Seventeenth International Symposium on Artificial Life and Robotics 2012 (AROB 17th '12),
B-Con Plaza, Beppu, Oita, Japan, January 19-21, 2012

Toward Artificial Kansei Based on Mental Image Directed Semantic Theory

Tzu-Hsuan Huang Kaoru Sugita Masao Yokota
Fukuoka Institute of Technology
yokota@fit.ac.jp

Abstract

The authors have proposed a human mind model of human mind consisting of Stimulus, Knowledge, Emotion and Response Processing Agents and simulated human-robot communication based on it. This paper proposes 'Artificial *Kansei* (AK)', namely, *Kansei* for a robot as tight collaboration of Knowledge and Emotion Processing Agents of our mind model, and considers its application to a *Kansei* information system for Buddhism statues made to order.

1. Introduction

In recent years, there have been developed various types of robots in Japan. However, they are to play their roles according to programmed actions to stimuli and have not yet come to understand such a mental function of their human partners as is called *Kanse*i. The authors have proposed a human mind model of human mind consisting of Stimulus, Knowledge, Emotion and Response Processing Agents (see Fig.1) and simulated human-robot communication based on it [1]. Besides this, we have tried to describe the meanings of *Kansei* expressions such as 'heart-calming', 'fantastic', 'soft', 'thick', 'grotesque', etc related to visual images of crafts and to retrieve them by these expressions as queries [4-6]. In this paper, we describe 'Artificial *Kansei* (AK)', namely, *Kansei* for a robot as tight collaboration of Knowledge and Emotion Processing Agents of our mind model, and verbalization of *Kansei* information so called '*Kansei* expression' by Response Processing Agent in the view of artificial or robotic individuality.

2. Multi-agent mind model

Figure 1 shows the multi-agent mind model proposed by the authors [1]. This is a functional model of human central nervous system consisting of the brain and the spine. The basic performances of its agents are as follows.

(1) **Stimulus Processing Agent** (**St**) receives stimuli from the world (**W**) and encodes them into mental images (i.e. encoded sensations) such as *"I sensed something oily."* (if verbalized in English.)

(2) **Knowledge Processing Agent** (**Kn**) evaluates mental images received from the other agents based on its memory (e.g. knowledge), producing other mental images such as *"It is false that the earth is flat."*

(3) **Emotion Processing Agent** (**Em**) evaluates mental images received from the other agents based on its memory (e.g. instincts), producing other mental images such as *"I like the food."*

(4) **Response Processing Agent** (**Re**) converts mental images (i.e. encoded physical actions such as *"I'll walk slowly."*) received from the other agents into real physical actions against **W**.

A *performance* ***P*** for a *stimulus X* with a *result Y* at each agent can be formalized as a function by the expression (1).

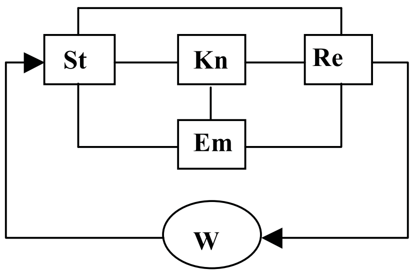
$$Y=\boldsymbol{P}(X), \quad (1)$$

where

P: a combination of ***atomic performances*** described later,

X: a spatio-temporal distribution of stimuli from **W** to **St** or a mental image for another agent, and

Y : a series of signals to drive an actuator for **Re** or a mental image for another agent.



St: Stimulus Processing Agent.
Kn: Knowledge Processing Agent.
Em: Emotion Processing Agent.
Re: Response Processing Agent.
W: World surrounding human mind, including his/her body.

Fig.1. Multi-agent model of human mind.

A performance ***P*** is assumed as a function formed either consciously or unconsciously. In a conscious case, a set of atomic performances are to be chosen and combined according to *X* by a meta-function, so called, '***Performance Selector*** (PS)' assumed as '***Conscience***'. On the contrary, in an unconscious case, such a performance as associated most strongly with X is to be applied automatically [8]

3. Mental image description

Mental Image Directed Semantic Theory (MIDST) has modeled mental images as "Loci in Attribute spaces" [3], [7]. An attribute space corresponds with a certain measuring instrument just like a thermometer, map measurer or so and the loci represent the movements of its indicator. The performance of 'Attribute space' is the model of '***Atomic performance***' introduced in Section 2.

A general locus is to be articulated by "Atomic locus" formalized as the expression (2) in first-order logic, where "L" is a predicate constant.

$$L(x,y,p,q,a,g,k) \quad (2)$$

The expression (2) is called "Atomic locus formula" whose arguments are referred to as 'Event Causer', 'Attribute Carrier', 'Initial Attribute Value', 'Final Attribute Value', 'Attribute Kind', 'Event Kind' and 'Standard Attribute Value', respectively.

The interpretation of (2) is as follows, where "matter" refers to "object" or "event".

"Matter 'x' causes Attribute 'a' of Matter 'y' to keep (p=q) or change ($p \neq q$) its values temporally ($g=G_t$) or spatially ($g=G_s$) over a time-interval, where the values 'p' and 'q' are relative to the standard 'k'."

When $g=G_t$ and $g=G_s$, the locus indicates monotonic change or constancy of the attribute in time domain and in space domain, respectively. The former is called 'temporal event' and the latter, 'spatial event'.

For example, the motion of the 'bus' represented by S1 is a temporal event and the ranging or extension of the 'road' by S2 is a spatial event whose meanings or concepts are formalized as expressions (3) and (4), respectively, where the attribute is "physical location" denoted as A_{12}. We think that the verb 'run' used in S2 must reflect the motion of the observer's attention [4].

(S1) The bus runs from Tokyo to Osaka.

$$(\exists x,y,k)L(x,y,Tokyo,Osaka,A_{12},G_t,k) \wedge bus(y) \quad (3)$$

(S2) The road runs from Tokyo to Osaka.

$$(\exists x,y,k)L(x,y,Tokyo,Osaka,A_{12},G_s,k) \wedge road(y) \quad (4)$$

The expression (5) is the conceptual description of the English word "fetch", implying such a temporal event that 'x_1' goes for 'x_2' and then comes back with it, where 'Π' and '•' are instances of the tempo-logical connectives, 'SAND' and 'CAND', standing for "Simultaneous AND" and "Consecutive AND", respectively.

In general, a series of atomic locus formulas with such connectives is called simply 'Locus formula'.

$$(\exists x_1,x_2,p_1,p_2,k)\ L(x_1,x_1,p_1,p_2,A_{12},G_t,k)$$
$$\bullet (L(x_1,x_1,p_2,p_1,A_{12},G_t,k) \Pi L(x_1,x_2,p_2,p_1,A_{12},G_t,k))$$
$$\wedge x_1 \neq x_2 \wedge p_1 \neq p_2 \quad (5)$$

4. Artificial *Kansei*

It is well known that emotion in a human can be affected by his/her world, namely, ***W*** in Fig.1. For example, a person's evaluation of live image of an object (i.e. image output from **St**) expressed by such words as 'favorite', 'beautiful', 'tasty', etc can vary depending on his/her emotional bias such as 'hungry', 'depressed', etc.

Kansei is one of mental functions with emotion involved but has a more complicated phase than pure emotion originated from instincts or imprinting. For example, sweet jam may be nice on toast but not on pizza for certain people knowledgeable about these foods. For another example, people can be affected on their evaluation of an art by its creator's name, for example, 'Picasso'. These are good examples of *Kansei* processing as emotional performance affected by knowledge in humans.

Therefore, *Kansei* can be defined as human emotion toward an object affected by its information, so called, 'concept', including his/her intellectual pursuits, traditions, cultures, etc concerning it. In this sense, *Kansei* is assumed to be reasonable among the people sharing such concepts unlike pure emotion. These hypothetic considerations are formalized as (7) and (8).

$$I_P(\mathrm{x}) = \boldsymbol{P_E}(\mathrm{S}(\mathrm{x})) \quad (7)$$

$$I_K(\mathrm{x}) = \boldsymbol{P_E}(\mathrm{S}(\mathrm{x}) \wedge \mathrm{O}(\mathrm{x})) = \boldsymbol{P_E}(\mathrm{S'}(\mathrm{x})) \quad (8)$$

where

$\boldsymbol{P_E}(\mathrm{X})$: Performance of **Em** for mental image 'X',

$I_P(\mathrm{x})$: Mental image as pure emotion for object 'x',

$I_K(\mathrm{x})$: Mental image as *Kansei* for object 'x',

S(x): Live image of object 'x' from **St**,

O(x): Concept of object 'x' from **Kn**,

S'(x): Unified image of live image and concept.

Figure 3 shows an example of *Kansei* processing in our mind model, where perceived, induced and inspired images correspond to S(x), S'(x) and $I_K(\mathrm{x})$, respectively, while Fig.2 is for pure emotion with $I_P(\mathrm{x})$ as the inspired image.

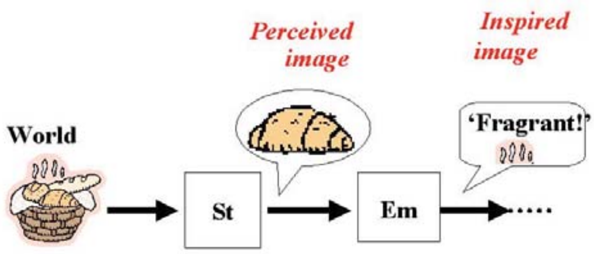


Fig.2. Example of pure emotion

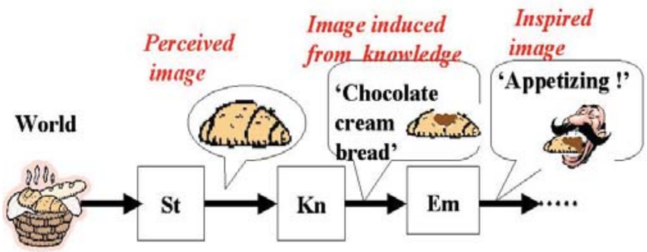


Fig.3. Example of *Kansei* processing

These two inspired images are to be verbalized in **Re** as 'Fragrant!' and 'Appetizing!', labeled in Fig.2, respectively. The essential difference between them is assumed to reside in whether or not they are affected by O(x), namely, the concept of 'chocolate cream bread', inferred by **Kn** from the shape and the smell. Whereas, pure emotion for an object can be a special case of *Kansei* processing without knowing or recognizing what it is.

In MIDST, the concept of an object '*x*' is given as an integrated omnisensory mental image of its properties and its relations with other objects involved. For example, the concept of 'chocolate cream bread' can be given by (9), reading that *x* is bread, sweet due to chocolate cream, fragrant of itself, etc, where A_{29} and A_{30} refer to 'Taste' and 'Odour', respectively.

$$(\lambda x)chocolate\_cream\_bread(x) \Leftrightarrow (\lambda x\exists y,k_1,k_2)L(y,x,Sweet,Sweet,A_{29},G_t,k_1)\Pi L(x,x,Fragrant,Fragrant,A_{30},G_t,k_2) \wedge bread(x) \wedge chocolate\_cream(y) \wedge \ldots \quad (9)$$

5. Human language understanding in robots

For comprehensible communication with humans, robots must understand natural language *semantically* and *pragmatically*. Here, semantic understanding means connecting symbols to conceptual images of objects and pragmatic understanding means connecting symbols to real objects by unifying conceptual images with perceptual images. However, humans and robots can be equipped with sensors, actuators and brains of different performances and their vocabularies may well be grounded on quite different sensations, physical actions or mental actions. And in turn such a situation may bring inevitably different kinds of semantics to them, so called, "Natural Semantics (NS)" for humans and "Artificial Semantics (AS)" for robots.

For example, consider such a scenario as follows.

…A human 'Kate' and a humanoid robot 'Robbie' encounter at the terrace in front of the room where a Christmas party is going on merrymaking. Kate says "Robbie, please fetch me some nice food from the gaudy room." Robbie replies "OK, Kate."….

For a happy end of this dialog, Robbie must have a good knowledge of Kate's NS for *Kansei* and translate it into its AS appropriately enough to find out the real objects referred to by her words. In this case, Robbie needs at least to interpret Kate's statement as the expression (10) reading "If <u>*Robbie* fetches *Kate* some food nice ***for her*** from the room noisy ***for her*** (***E1***),</u> then consecutively <u>it makes *Kate* happier (***E2***)</u>". It is notable that (10) is the canonical form of the meaning of an imperative sentence.

$$\boldsymbol{E_1} \rightarrow_c \boldsymbol{E_2} \quad (10)$$

where

$$\boldsymbol{E_1} \Leftrightarrow (\exists x_1,x_2,k_1,k_2,k_3,k_4)\ (L(R,R,K,x_2,A_{12},G_t,k_1)\bullet (L(R,R,x_2,K,A_{12},G_t,k_1)\Pi L(R,x_1,x_2,K,A_{12},G_t,k_1)))$$
$$\Pi(L(K,x_1,Nice,Nice,B_{08},G_t,k_2)$$
$$\Pi\ L(K,x_2,Gaudy,Gaudy,B_{08},G_t,k_4)$$
$$\wedge food(x1) \wedge room(x2)$$
$$\boldsymbol{E_2} \Leftrightarrow (\exists e_1,e_2,k_7)\ L(\boldsymbol{E_1},K,e_1,e_2,B_{04},G_t,k_7) \wedge e_2>e_1.$$

The special symbols and their meanings in the expressions above are:

'$X \rightarrow_c Y$' ='If *X* then consecutively *Y*', *R*='Robbie', *K*='Kate', B_{08}='*Kansei*' and B_{04}= 'Happiness (=degree of happiness)'.

As easily imagined, these values of the attribute *Kansei* (B_{08}) greatly depend on their standards (i.e. k_2 and k_4) that are most closely related to 'Individual' or 'Purposive' standard shown in Table 1 (see APPENDIX).

By the way, Robbie's task is only to make ***E1*** come true where each atomic locus formula is associated with his actuators/sensors. Of course, Robbie believes that he will become happier to help Kate, given by expression (11) where 'B_{03}' is 'trueness (=degree of truth)'and 'K_B' is a certain standard of 'believability'. That is *emotionally* to say, Robbie likes Kate. Therefore, this example is also very significant for intentional sensing and action of a robot driven by logical description of its belief.

$$(\exists p)L(R,\boldsymbol{E},p,p,B_{03},G_t,K_B) \wedge p>K_B \wedge \boldsymbol{E} = \boldsymbol{E_1} \rightarrow_c \boldsymbol{E_2} \quad (11)$$

6. Affective analysis of Buddhism statues

Many psychologists have claimed that certain emotions are more basic than others (Ortony & Turner, 1990). We have assumed that human emotion consists of 5 primitives representing the degrees of 1) Anger, 2) Disgust, 3) Anxiety, 4) Happiness, and 5) Superiority. For example, the degree of Happiness is measured by using such a word set as {anguish, distress, sorrow, gloom, content, joy, ecstasy}, whose each element is possibly arranged on a coordinate axis and fuzzified with a certain characteristic function. Therefore, we have assumed Kansei as a certain function to evaluate totally the loci in the attribute spaces of these primitives. Based on the 5 primitive emotional parameters, we have been analysing Buddhism statues as shown in Fig.4 in order to plot them in the attribute space of Kansei (B_{08}) with some purposive standard of Buddhism. Table 1 shows an example of such analysis, where H, M and L denote high, middle, and low in degree, respectively. These results are to be associated with Kansei words such as divine, gentle, reverential, noble, valiant, etc. analyzed in the same way and to be applied to the customer servicing interface of a Buddhism statue ordering system.

a) Dainichi-nyorai (DN) b) Fudo-myo-o (FM)

Fig.4. Samples of Buddhism statues

Table 1 Affective analysis of Buddhism statues

Sample	anger	disgust	anxiety	happiness	superiority
DN	L	L	L	H	H
FM	H	M	H	L	H

7. Discussion and conclusion

Our mind model is much simpler than Minsky's [2] but the locus formula representation can work for representing and computing mental phenomena fairly well [1, 3]. For realizing a plausible *Kansei*, it is most essential to find out functional features of **Em** and to deduce from them such laws that rule P_E. The most important problems to be solved are how to realize the attribute space of *Kansei* and how to build its corresponding atomic performance. In order to solve these problems, focusing on Buddhism statues, we will consider the application of soft computing theories such as neural network, genetic algorithm, fuzzy logic, etc. in near future.

References

[1] Yokota,M., Shiraishi,M. and Capi,G.: "Human-robot communication through a mind model based on the Mental Image Directed Semantic Theory," Proc. of the 10th International Symposium on Artificial Life and Robotics (AROB '05), pp.695-698, 2005.
[2] Minsky, M: The society of mind, Simon and Schuster, New York, 1986.
[3] Yokota, M. & Capi,G.: "Cross-media Operations between Text and Picture Based on Mental Image Directed Semantic theory," WSEAS Trans. on INFORMATION SCIENCE and APPLICATIONS, Issue 10, 2, pp.1541-1550, Oct. 2005.
[4] M.Fukuda, K.Sugita and Y.Shibata, "Perceptional Retrieving Method for Distributed Design Image Database System," IPSJ Journal, Vol.39, No.2, pp.158-169, 1998.
[5] K.Sugita, A.Miyakawa and Y.Shibata, "Relation Between Kansei Words and the Room Space in Digital Traditional Japanese Crafting System," The International Conference on Advanced Information Networking and Applications (AINA), 159-162, 2003.
[6] K.Sugita, T.Ishida, A.Miyakawa, Y.Shibata, "Kansei Retrieval Method using the Quantitative Feature of Traditional Japanese Crafting Object," Proc. of The 6th International Workshop on Multimedia Network Systems and Applications (MNSA'2004), 112-117, 2004.
[7] Prinz, W.: "A common coding approach to perception and action", In *Relationships between perception and action* (Neumann, O. and Prinz, W. eds.), pp.167-201, Springer-Verlag, 1990.
[8] Brooks, R. A.: "A robust layered control system for a mobile robot", IEEE Journal of Robotics and Automation, RA-2, pp.14-23, 1986.
[9] Ortony,A. and Turner,T.J.: "What's basic about basic emotions?", Psychological Review, Vol.97, pp.315-331,1990.

APPENDIX

Table 2 Examples of standards.

Categories of standards	Remarks
Rigid Standard	Objective standards such as denoted by measuring *units* (meter, gram, etc.).
Species Standard	The *attribute value ordinary* for a species. A *short train* is ordinarily longer than a *long pencil*.
Proportional Standard	'*Oblong*' means that the width is greater than the height at a physical object.
Individual Standard	*Delicious* food for one person can be too *poor* for another.
Purposive Standard	One room *comfortable* enough for a person's *sleeping* may be *uncomfortable* for his *jogging*.
Declarative Standard	The origin of an order such as 'next' must be declared explicitly just as 'next *to him*'.