Cognitive Modeling in Software and Relation to Human Emotional Reasoning [1]

HAMIDO FUJITA[2], JUN HAKURA, MASAKI KUREMATU
Intelligent Software Systems Laboratory
Iwate Prefectural University
Iwate, 020-0193
JAPAN

Abstract: - The paper reports on our experience to in adapting emotional experiences of the software engineers in evolutionary design of software systems. The works here reported present development progress report in relation to the state-of art that need to create the multidisciplinary technologies, needed to establish best harmony engagement between human user the software application, based on human cognitive analysis. We approach the user best engagement from facial and voice analysis. And through it, we can measure (collectivized and quantified), and observe the user behavior, and accordingly enhance the engagement by generative interactive scenario. The approach has been experimented using famous literature person (Keni Miyazawa).

Keywords. Intelligent software, cognitive modeling, Human computer Interaction, software development, facial analysis, sound analysis.

1 Introduction

The software designer needs to extend certain cognitive view on the user emotional behaviour in the design process. Having the user mental behaviour be reflected into the system can enhance creativity process on such understanding. Software design is a creative process that needs to stimulate thinking from spatial design prospective, which cooperatively and intellectually participates to establish an engaging harmony to best design practices based on distributed cognition. Computer can stimulate such engagement and create multilayer situated enact for best practices in creative design process. A possible elaboration on the work presented in this paper for software design, is to bring a close look on the adaptability of such innovation on software spatial deployment, that should integrate cognitive process and user usability be integrated, by examining the nonverbal communication skills between the designer and the systems. For example in visual art, of all cultures mental work is done to bring separation together into a whole. Designer should integrate or reflect this aspect into the design principles.

Awareness of their framing within for example, software interface, as well as the knowledge of the separate narratives captured within them, links our visual experience to a known genre, event, or tradition creating a congruence of understanding. This mental work of making fragments whole or of shaping clues and cues into a pattern is often thought if as an individual responses. Connections between perceptual and conceptual or linguistics representations emerge in socially interactive situations that punctuate underline and enlarge individual understanding. Verbal language has minimal units of meaning whereas visual and (generally) gestural units would not. It offers the opportunity to lift actions and intentions out of the moment into multiple versions of something else. Metaphors apply to forms to give meaning. Form is therefore a

[1] This research is supported by Iwate Prefectural University Research Grants on “Kenji project”..

vehicle for inference, and the content of the inference depends on the metaphor.

Mental model of the designer engaged with the tool depends on the role and level of engagement. The minute usage of each component in the tool and also on the system capability to enhance and empower the designer enacts to prompt his/her emotional collective intuition for best practices. Digital virtual worlds as mentioned by Mitchell [13] in not only more pervasive and efficient than ever, it is also generating new cultural complexities. Words not as literature but as signs within the context of space, to confirm and enliven our urban setting in our information age. Urban spaces and places provide setting for communication and at how they conduct complex flows of information through new architectural design. It is the essential interaction between digital media and the built environment. We think our project reported here contributes to such bridge.

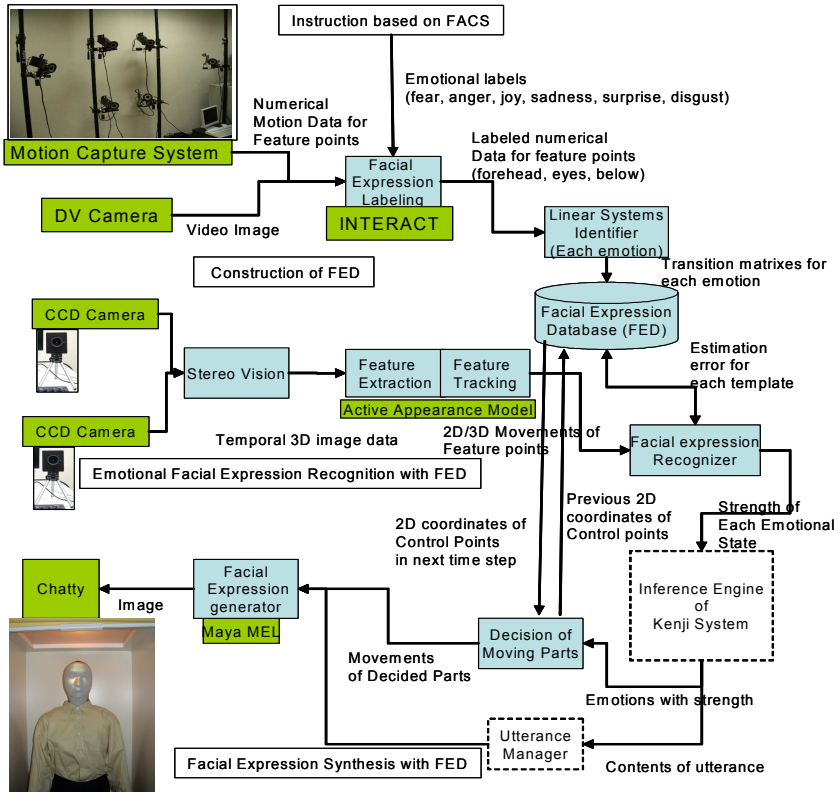


Figure 1. Outline of Virtual Kenji System

People interact with digital technologies through touch manipulation and gesture interaction is increasingly being embodied. People move through environments embedded with digital artefacts, and interact with and through technologies in new ways. This act as collaboration in design, and specify the design concept as collaborative cognitive process. This participate in generative and evolutionary techniques in architecture[10]. There are digital spaces that participate in architectural design in such digital world that involve people to interact through its space [20], [2].

We think that the practices reported in this work contribute to integrate (corporate) the cognitive intention of the designer with the knowledge of the system. The system designer can use these design practices to inhale the emotional practices into the design using such experiment. [24], [26] described a possible architecture for organizing agents into a flexible, human-like Society of Mind. Rather than seeking a best way to organize agents, their architecture supports multiple 'ways to think', each a different architectural configuration of collaborative agents. [8] identified three different kinds of distribution of cognitive process' across people, across representation, and across cultures. Socially distributed cognition focuses on the role that a group of people have in thinking and knowing and on the phenomena that emerge as a result of these social interactions. Cognitive process makes use of external as well internal representations. These external representations are things such as notes, scripts, and other information artifacts. It is a metaphoric representation, collected from different dimensional representation, (i.e., disciplines), collectively to enact for example Miyazawa Kenji and revive him through such conceptual cognitive representation. Psychology, linguistics,

computer science, and philosophy, collectively can lead to cognitive science disciplines. Cognitive Psychology [14], contributes to understand human thought from an individual perspective.

representing the cognitive mental model (Self) of other human.

Computer program reflecting some concept does have a self, this reflected by the conceptual design, of certain personality, so the self is defined by the

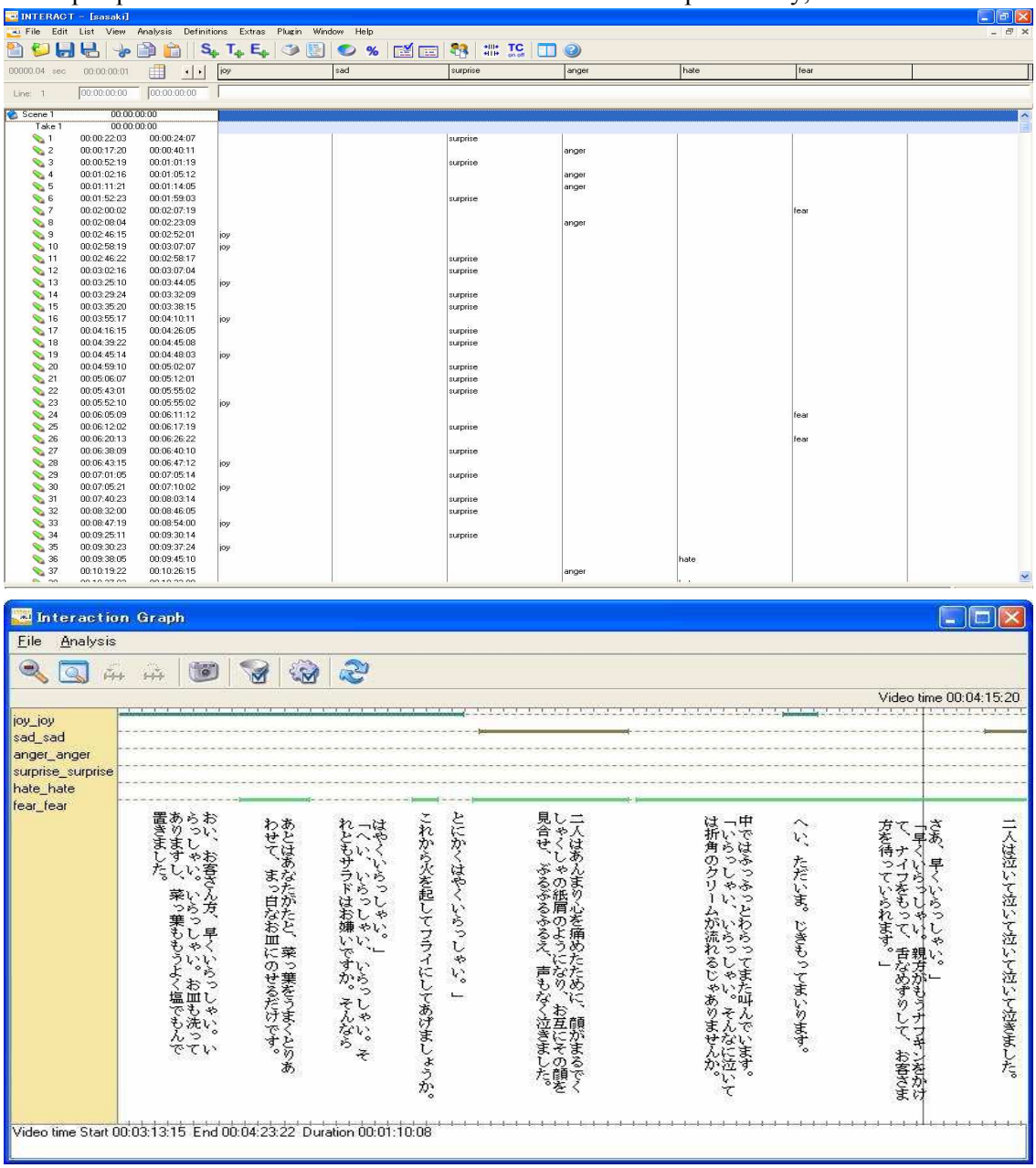


Figure 2. Cognitive analysis of Kenji Scripts using INTERACT software

We project the general conceptual framework above through the parts of Miyazawa Kenji project (Fig.1). We present the main four parts of the project. We show our technology on the interaction between human and virtual system

personality, of that person reflected in the program. The face and voice emotional representation belong to human mind, or mental representation. They are part of human emotional states representation, reflecting mind transition in different representation reflected as modes. It is

part of the language context that be integrated into our vocabulary and language representations.

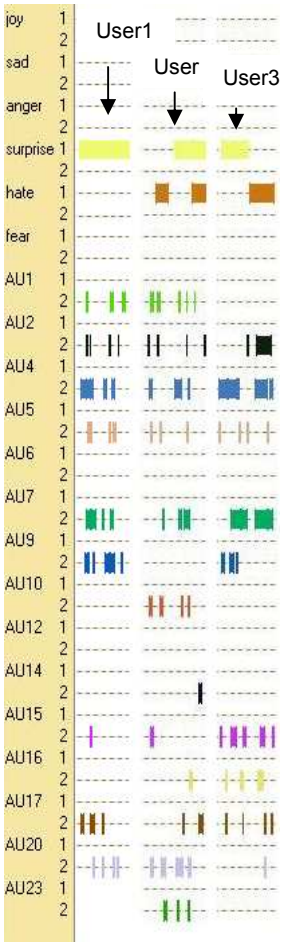
In this paper, we are using this case study to bridge these issues and move to the direction of intelligent human centric computing that can mimic a specific human cognitive behavior and based on this cognitive modal we can reason on real human interactive behavior for spatial design. The rest of this paper is organized, to show the major part in this case study. In Section 2, we will discuss the example of presenting certain human cognitive model. We have used Miyazawa Kenji cognitive mode. In Sec. 3 we present part 1 of the system, that will create the emotional feature of Kenji system as virtual world. In Sec. 4 we present part 2 of the system, that collect human user cognitive interaction and mental behavior based on Kenji Style reasoning and other knowledge related to common sense reasoning. In Sec. 5 we preset part 3 of the system, that related to voice emotional recognition (but in short outline). In Sec. 6 we deal with Part 4, that initiate the scenarios and responses to the user in role act style to the cognitive behavior of the user state. In Sec. 7 we present conclusions.

2. Cognitive style of human cognition: Kenji Style case study

As mentioned in the introduction that we have selected Kenji Miyazawa(MK) to be the virtual model of our experiment on intelligent human interaction cognitive based conceptual model. http://www.kenji-world.net/english/who/who.html; this link give an overview about who is Kenji. Such cognitive behavior reasoning system interacts with human user based on cognizing-based reasoning, and factorized through, based on MK cognitive studies.

Our system thinks on which action it may take to appropriately interact with the user. The outline of the system is shown in Fig.1. This decision making process is based on MK thinking style. The way in which we use our mind becomes the way in which we use our body and the attitudes of mind so that to create its own manifestation in the function of the muscles that implements deliberately the concessions behavior behind it.

Previous or old thoughts (from Aristotle to Darwin) saw facial expressions as the result of internal emotional states. Facial expressions were seen as pre-warning of emotional responses on others. However, why do humans need such non-verbal communication and complex facial muscles when we have language? Darwin tried to extend his theories on evolutions of structures to behavior. He felt that behavior also evolves, and concluded from the universality of many facial expressions (sadness, happiness, etc.) that such behaviors also evolved from lower life forms. Facial expressions are "serviceable habits" that helped the organism react to sensations and internal states.



In [26] has introduced the “self,” as mechanism

Figure 3. *[INTERACT] software usage in video analysis of human behavior*

to logic related to how to define or represent and put in structures the self to reason cognitively on it. Kenji style is the emotional voice and facial animation that virtual MK is able to speak through in role act to the user.

These are the extracted cognitive feature reasoned templates. As stated in the introduction, we need to construct creatively and physiologically Kenji style featured by his personality implicitly hidden within his scripts and from scholars who have acquaintance on his personality reflected through his published artwork. This style is constructed from collected data from testing actual person act and has some knowledge on Kenji scripts, and from general person who can read and cognitively understand Japanese scripts.
The analysis data have been classified according to six emotional modes of EKMAN[4]. We use such style of reasoning to label and understand on how to use the gesture. We have selected scripts from Kenji artwork.
The analysis is based on cognitive feature extraction referenced on reading of above-mentioned Kenji scripts (1) by specialist in Kenji literature and his art pieces. Also, the same has been done by: (2) reading observation analysis on non-expert people, (i.e., general Japanese people).
These extracted patterns based on experimental analysis and reasoning of Kenji scripts projected through (life style, physiological view, philosophical, linguistics reasoning referencing analysis (onomatopoeia and mimesis) and other analytical observations as shown in Fig.2, that shows the emotional analysis using INTERACT tool [9]. As shown in Figure 3, that among 10 observed users we have collected 3 users with 90% match related to the Action units and six Ekman emotional modes [4]. These collected templates have been used to reflect Kenji Style. We used these templates to establish part 1 of the system. For more details on these templates, please refer to [6], [9].

3. Emotion Estimation from Facial Expressions of Users

Part 1 of Kenji System presents the hologram, it is as shown in Fig. 5. The total image of Kenji Hologram is on Fig. 5, image_4. The other photos(1-3) snapshots are taken while Kenji is talking through the emotional templates that are created in real time by the Emotion processor (We called it *KANJO* processor, *KANJO* means emotion in Japanese language). KANJO processor (Figure 4-1) is synchronizing the MAYA images generating in real-time animated facial images, and synchronized through KServer (Figure 4-2), and the emotional sound file extracted Kenji text (refer to: Fig.8), all this is synchronized through KNAJO processor. All this is referenced as Part_1 in our system. The details of the software development (Part_1) here are omitted for space and technological securities issues related reasons.

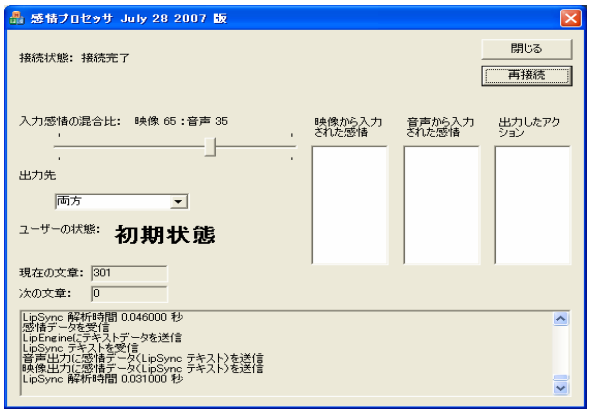


Figure 4-1. KANJO processor

Figure 4-2. KENJI-Engine Server

Figure 4. Virtual Engine component

You can reference to demos on this application by reference to [9] or the link (http://www.fujita.soft.iwate-pu.ac.jp/prof_dir /issam/others/ KenjiOnly.wmv). Please notice that all is done in real time. This section will be reference again in Sec.5, related to creating

facial images in harmony with the contents of the spoken text.

Figure 5. Shows part 1 of the hologram

4. Emotion Estimation from Facial Expressions of Users

In this section, we preset part 2 of Virtual Kenji system, that to make the interaction between Kenji and human user, to achieve the conceptual cognition engagement with a user, the system is required to react to emotional states of the user. Emotional states of the user can be perceived through emotional signs exhibited in several modalities, such as words, vocal features, and gestures, and recognized collectively through situated reasoning. Gestures are known as one of the essential modalities to perceive the emotional states of the user. Among the gestures, facial expressions afford a great deal of emotional information in human natural communications. In proportion to the importance, there have been a lot of studies concerning facial expressions are conducted not only in psychology and philosophy, but also in computer science. One of the most popular approaches to automatic facial expression analysis is relying on the Facial Action Coding System (FACS), (for FACS, Ekman and Friesen, 1975[4]; for a survey of the literature see, [19]. The FACS uses the combinations of movements of facial parts, named Action Units (AUs).

Namely, detecting the AUs is the main subject of the approaches relying on the FACS. The AUs' are defined as typical results of movements of facial parts in facial

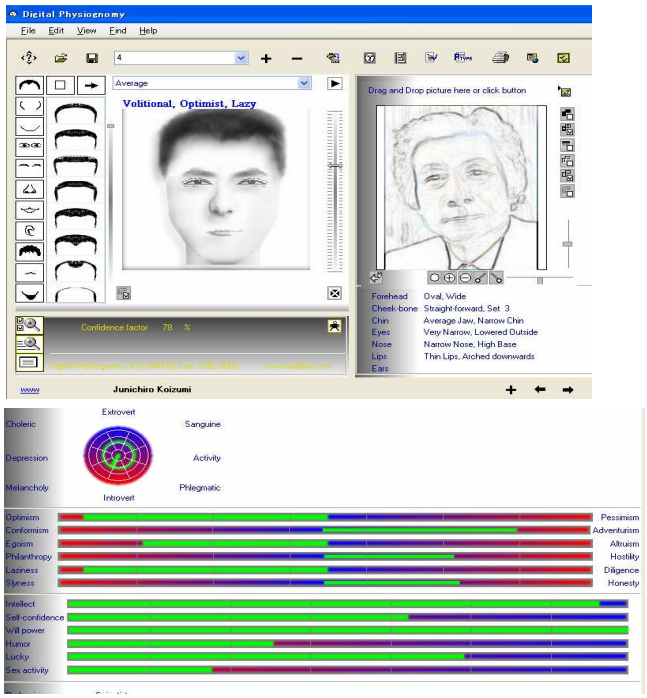
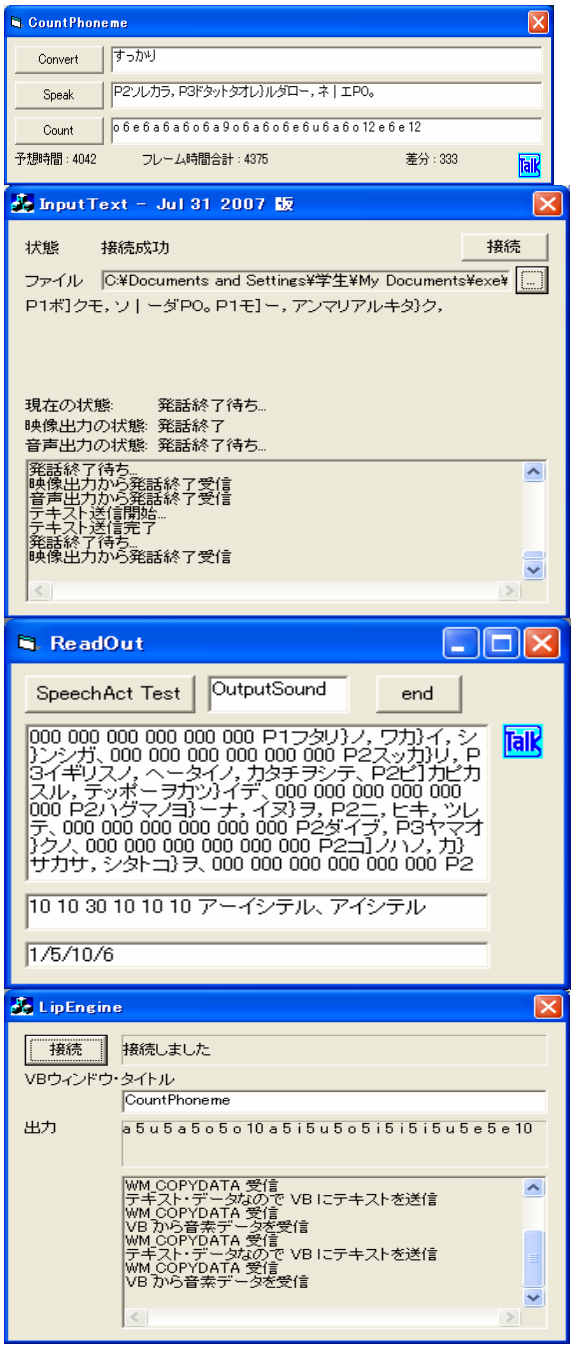


Figure 6. Physiognomy analysis of J. Koizumi

expressions, such as “left eyebrow up” and so on. Thus, they are apt to focus on the static images of the facial expressions, and require the completion of the expressions.
Namely, they do not fully utilize the dynamic aspects of the facial expressions.
This would result in the misleading at on the reasoning about the situation: “What triggers the facial expression?”
To know the emotional states of the user together with the exact timing of their appearance is one of the important requirements to the conceptual cognition. For this aim, we have introduced a linear system identification approach to the facial expression analysis [7]. The approach is able to fully utilize the dynamic aspects of facial expressions.

4.1 First impression recognition
In order to enhance the recognition of the subject user, we need to have a means to cognitively recognize the user. This inference is the first sep to collect the impressive behavior of the user physiognomy that means reading one’s character by face. However, the correctness of such science cannot be predicated, but we have used such analysis view to establish the initial state of the user impression that we would start with to navigate in discovering the recursive property of user emotional states. Most personality features are manifested at different times in response to different situation. We computed these situations with trace using techniques mentioned in Sec.3. The emotional estimation mentioned in Sec.3, is based on the active appearance model, through the interaction sequence with VK. However, the user estimated emotional state is calculated though the physiognomy (study of facial characteristics) techniques. Like Fig. 6 that shows the extraction of Koizumi (previous Japanese PM), emotional features, based on his face shape. Ekman[4] discovered that the face is such an efficient instrument of communication that govern the way we interpret facial expressions. We would like to learn to identify and analyze the feature of the face and gestures that characterize personality traits. The Choleric face is full of energy, but lacks of self-control. We have used digital physiognomy tool http://www.uniphz.com in this analysis for 1st impression extraction. This is based on defining classes library for face parts, an instantiates the attribute collected from the a taken user photo to



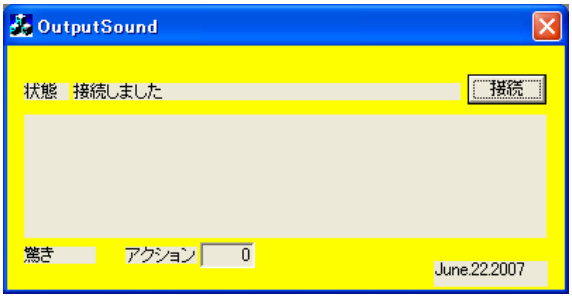


Figure 7. Shows the screens of the Input text window,
and readout window

specify the feature representing that attribute. These classes libraries, have been organized as class FACESHAPE, Class Eyes, Class FOREHEAD and so on. The initial function for the subjective user is recursive as the initial input is 1st set by the physiognomy system, 1st impression emotional attribute model for the subjective user. Then after this would be the basic model that VK would use to interact with the subjective user (part 4 of the system). The change of the emotional behavior would be extracted using the model presented in Sec.3, based of Active appearance that attributed to Kenji system.
We extract these details due to space.

5. Emotional Voice recognition

This Section in very short resembles what we call as part_3 of Virtual Kenji System. In our system reference to Figure 5 (emotionally generated facial images), the corresponding text with emotional features (represented as templates) generated by the system and spoken by Kenji system as shown in Figure 7, are been synchronized to create the total cognitive real image interface for talking person with cognitive personality specialized as Kenji. The facial movement of the lips has been (real time) synchronized by interface (LipSync mentioned on the reference [11]), which is API (application program interface) with MAYA application.
However, the templates generated on are shown in Figure 7. These software components resemble the modules on outputting VK voice been synchronized with real-time generated facial images. For more details on this, please reference to another paper presented in this conference.

6. Cognitive Scenario generation according to human cognitive state

We explain here part 4 of the system. How to make a system can act with the user according to a situation, what type of possible scenario or knowledge that the system can provide to the user? These issues should be reflected on memory structure and situated computing.
The User cognitive states been examined and analyzed using Part 2 of the system. The user engagement with Kenji system is computed using Sec.4 that analyze the facial expression of the user to examine the degree of user engagement with Virtual system mode. The system would conclude if the user is interesting in the current created scenario or not interesting. Actively engaged or disconnected. There are four states that the system concludes to reach; according to the conclusion, it takes through the facial analysis of users. Also, this is the same for voice emotional recognition as well. So, we have a view on facial analysis state (using Sec. 4), and view on voice analysis state (using Sec. 5). These two views are integrated to create the cognitive state of user engagement with Virtual Kenji system. We have created several scenarios that prepared according to the expertise in Kenji cognitive style views. Also other views have been prepared according to several situations, that been classified according what is called as 1st imprecision [21]. We use a tool named as digital physiognomy from www.uniphiz.com to test user physiological emotional states before they engaged with KENJI so that to create emotional 1st impression model. According to the type of the user face, we created a scenario that Kenji virtual model can interact. These systematic guidelines are to simplify the best engagement between the human user the virtual system.
Though the experiment is to have the system, be Kenji and the user is general user who has certain knowledge and interest on such

famous writer. We think the system can be useful for HCI design to complex creative artwork or eliciting a complex requirement, where the user nonverbal communication work in hand with to stimulate the designer thinking for best harmony with system and user cognitive thinking mutually, with emotional integration of the design.

The templates mentioned in Sec.4 and Sec.5 include mechanism to imbed situations, and user mental background ontological views (vast views: culture views, and mental view and spontaneous views). We human; our intellectual communication is not bounded by fixed templates. Though we use them in learning and adapting our self through them, but we modify them for best performance. For example, we learn templates on driving skills by theory and practices. But on road, we modify these templates to match it to our behavior and cognitive mental performance. Such adaptability is related to the best adjustment that our body system and condition can fit into to create the best harmony that we think such driving performance is best. For the same human, driving style (templates) in downtown Cairo is not as driving style in downtown Tokyo. Looking into the contents of cognitive actions, we notice different patterns, between the proposed virtual system and human user in terms of perceiving a certain space. The spatial space relationship cognitive integration between human and virtual system is essential to best harmony in communication. These issues can be reflected into the Architecture design, when the integration of spatial space in design is essential to evaluate the whole layout of the architecture, where designers compete to new ideas based on their perception of the current state of the integrated spatial design in total be interfered by the cognitive user (mind).

We think the development of new interactive environment (like virtual Kenji system) can employ user interface with spatial cognition integration. This can contribute to reduce the load of mental visual reasoning.

We have in addition this, we added other sub-space, we called tonal sub-space that created a short of musical tone on the generated synthesized voice. Reference to rules of compositions by Marc-Antoine Charpentier (1692), we could notice the perpetual aspect between music and text. Pitch (Melody and harmony), rhythm (tempo and meter), and sonic qualities, (timbre, articulation dynamics, and texture) are the three parts defining the music structure. Such patterns construction, as combination of natural stimuli in well-specified generative forms reflects the source of our voice synthesis to produce sounds with such constructed nature. From this table (below) we have cognitive six Ekman modes be classified into tonal sub-space. The style of wording pronunciation is based on the compositional synthesis summarized in the below table. This sub-space would be integrated with the voice emotional recognition (Sec.5), sub-space. These two sub-spaces constitute the voice emotional synthesis and recognition part.

We still are adding issue in regard to the colors and emotional states of the systems. From that prospective, we will have facial recognition space (Sec.4) and color sub-space to be integrated as part of the emotional space (facial, and voice), representation, all together to construct the distributed spatial recognition for best engagement between man and, machine.

The color sub-space has not explained yet in this paper. This will be referenced in another to submit paper.

We still are adding issue in regard to the colors and emotional states of the systems. From that prospective, we will have facial recognition space (Sec.4) and color sub-space to be integrated as part of the emotional space (facial, and voice), representation, all together to construct the distributed spatial recognition for best engagement between man and, machine. The color sub-space has not

explained yet in this paper. This will be referenced in another to submit paper.

7. Conclusion:

We think that the two sub spatial cognitive spaces model would participate collectively to establish a cognitive interface between human and machine. The voice synthesis explained in Sec.5, integrated with tonal presentation outline in Sec. 6, these two subspaces contributes to create the voice synthesis space of machine recognition. Facial sub space explained in Sec. 4 and the color subspace, contribute in establishing the perceptional view of the cognitive interaction (the perceptional synthesis and recognition of human emotional reasoning). This research project contributes to establish the best harmony and engagement between human user and the system based on defining and constructing the "self" in computing style, based on voice and facial construction definition of the self. The self defined in the system interacts with the human user based on the reasoning and recognition of the collected user voice and facial informational features and reasoned according to the self. This experiment would contribute to create new generation of cognitive user interface between man and machine.

Acknowledgement

We would like to thank Iwate Prefectural University for the financial support provided fro this project. Also, gratitude goes to (www.bitage.co.jp), Bitage limited for the hard work development on KANJO processor. Thanks also to Mangold International (www.mangold-international.com/), Mr. Pascal T. Mangold(CEO), for the support on INTERACT software tool, also thanks go to Dr. Kevin Erler of Automatic Sync Technologies (www.automaticsync.com), for technical advises on LipSync. Also, many thanks go to MIYAZAWA KENJI Museum (Mr. Ushizaki), Prof.T. Sasaki (Kenji's Scripts analysis), and Prof. Mochituki.

References

[1] ADELMANN, P. K. and ZAJONC, R. B. 1989. "*Facial efference and the experience of emotion*." Annual Review of Psychology, 40: pp249-280.

[2] BILDA Z., 2006. The role of mental imagery in conceptual designing, Doctor Thesis, *Key Centre of Design Computing and Cognition, Faculty of Architecture, University of Sydney, Australia*

Sad	C_{Minor}, B_{Minor}, F_{Minor}
Happy	D_{Major}, G_{Major}, A_{Major}
Fear	$B_{FlatMinor}$, E_{Minor}
Surprise	$B_{FlatMajor}$, G_{Minor}, G_{Ma}
Anger	F_{Major}, E_{Major}
Disgust	F_{Minor}, M_{Major}

[3] COOTES, T., EDWARDS, G. J., Taylor, C. J., Burkhardt, H., Neuman, B., 1998. Active Appearance Models, *Proc. Eur. Conf. Computer Vision*, 2, 484-498.

[4] EKMAN, P. and FRIESEN, W. V., 1975. Unmasking the Face, *Prentice Hall*, NY.

[5] EVANS, W.A., 1994. Approaches to intelligent information retrieval. *Information processing and management,* 7 (2), 147-168.

[6] FUJITA, H., HAKURA J., KUREMATSU, M., 2006. Virtual Cognitive Model for Miyazawa Kenji Based on Speech and Facial Images Recognition, *WSEAS Transactions on Circuits and Systems,* 10(5), 1536-1543.

[7] HAKURA, J., KUREMATSU, M., FUJITA, H., 2007. Facial Expression Recognition and Synthesis for Virtual Miyazawa Kenji System, *WSEAS Transactions on Circuits and Systems,*3(6), 288-295.

[8] HOLLAN, J., et al. 2000. *Distrubuted Cogntion:Toward a New Foundation for Human-Computer Interaction Research*. ACM Trans. On Computer-Human Interaction 7, No.2: 174-196.

[9] INTERACT, 2007. http://www.mangold-international.com/ News: *Japanese Scientist Revive Famous Writer, April 2007. Mangold International*

[10] JANSSEN, P., 2006. A Generative Evolutionary Design Method, *Digital Creativity*, Vol. 17, Issue 1, 2006, pp. 49–63.

[11] LIPSYNC, Automatic Sync Technologies, http://www.automaticsync.com ProductionSync and InstantSync.

[12] MERCER, P.A. AND SMITH, G., 1993. *Private view data in the UK.* 2nd ed. London: Longman.

[13] MITCHELL, W. J., 2005: *Placing Words: Symbols, Space, and the City*, MIT press, ISBN: 10-0-262-63322-1, Sept., 2005

[14] NEISSER, U. 1967. *Cogntive Psychology*, Englewood Cliffs, NJ' Prentice-Hall.

[15] NISHIYAMA, M., KAWASHIMA, H., HIRAYAMA, T., MATSUYAMA, T. 2005, Facial Expression Representation Based on Timing

Structures in Faces, *IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, 140-154.

[16] R-Project *http://www.r-project.org/*

[17] OUDEYER Pierre-Yves, 2003. The production and recognition of emotions in speech: features and algorithms *International Journal of Human Computer Interaction* Vol.59(1-2) pp.157-183.

[18] OPEN COMPUTER VISION LIBRARY, http://sourceforge.net/projects/opencvlibrary/

[19] PANTIC, M. and ROTHKRANTZ, L J. M., 2000. Automatic Analysis of Facial Expressions: The State of the Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 1424-1445.

[20] PENG, W and GERO, J.S., 2007: "*Computer-aided design tools that adapt"*, *CAADFutures2007* (to appear)

[21] TICKLE N.R., 2003. *You can read a face like a book*, Daniels Publishing, 2003.

[22] SILVER, K., 1991. Electronic mail: the new way to communicate. *In:* D.I. RAITT, ed. *9th international online information meeting, 3-5 December 1990 London.* Oxford: Learned Information, 323-330.

[24] SINGH, P. , and MINSKY M., 2004. *An Architecture for Cognitive Diversity*, In*:* D. DAVIS , ed. Visions of Mind.

[25] STEHMANN, M. B., ERSBOLL, B. K., LARSEN, R., 2003. *FAME – a flexible appearance modeling environment*, IEEE Transactions on Medical Imaging, 22(10), 1319-1331.

[26] MINSKY M., 2006. *The Emotion Machine,* Simon & Schuster publisher,