COGNITIVE MODELING IN DESIGN BASED ON HUMAN EMOTIONAL REASONING

Computer based Cognitive interaction based on mimesis of human emotional behavior

HAMIDO FUJITA, JUN HAKURA AND MASAKI KUREMATSU
Iwate Prefectural University, Iwate - Japan
issam@soft.iwate-pu.ac.jp

Abstract. This paper presents a progress development results of Virtual intelligent interface based on human facial and voice recognition. We this is new challenge for sensing the user emotional space and interact with it. It is part of the cognitive spatial design needed to have the mentality of the designer been part of the system recognition. This is experimental built prototype. 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 architect can use these design practices to inhale the emotional practices into the design using such experiment.

1. Introduction

The architect 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. Architect is a creative designer that needs to stimulate his/her thinking from spatial design prospective, which cooperatively and intellectually participates to establish an engaging harmony to best design practices based in 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 architectural design, is to bring a close look on the adaptability of such innovation on architecture spatial design, that should have cognitive process and user usability go be integrated as well, to examine 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. Architect should integrate or reflect this aspect into the design principles.

Awareness of their framing within for example, the architecture of the church, 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 do not. It offers the opportunity to lift actions and intentions out of the moment into multiple versions of new concept. Metaphors apply to forms to give meaning. Form is therefore a 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 W. (2005) 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 contributes to such bridge.

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 architecture as collaborative cognitive design process. This participate in generative and evolutionary techniques in architecture (JANSSEN 2006). There are digital spaces that participate in architectural design in such digital world that involve people to interact through its space (PENG, W., et.al.,2007), (BILDA, 2006).

The architect can use these design practices to inhale the emotional practices into the design using such experiment. (SINGH, and MINSKY, 2004, MINSKY, 2006) describe 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. (HOLLAND, 2000) identifies 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, entries in logbooks, 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 (Neisser, 1967), contributes to understand human thought from an individual perspective.

We project the conceptual framework through the parts of Miyazawa Kenji project. We present the main four parts of the project. We show our technology on the interaction between human and virtual system representing the cognitive mental model of other human subject.

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. 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 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. 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.

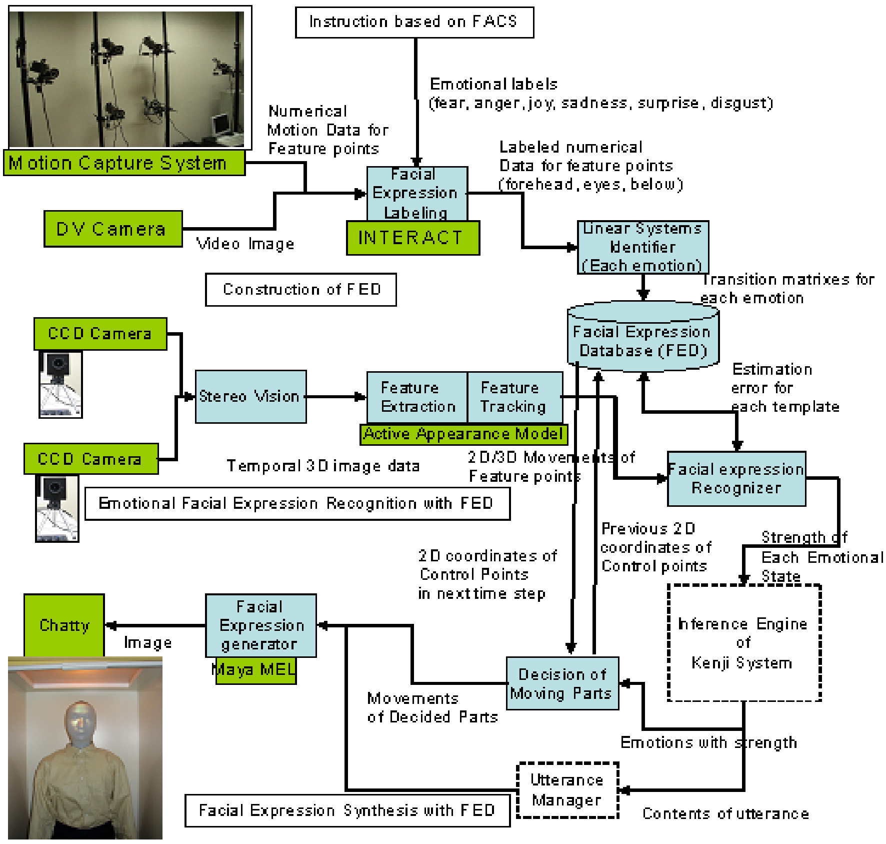


Figure 1. Outline of Virtual Kenji System

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 (MINSKY, 2006, Ch.9) has introduced the “self,” as mechanism 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 by testing actual person act with some knowledge on Kenji scripts. The analysis data have been classified according to six emotional modes of EKMAN. We use such style of reasoning to label and understand on how to use the gesture. We have selected scripts from Kenji artwork.

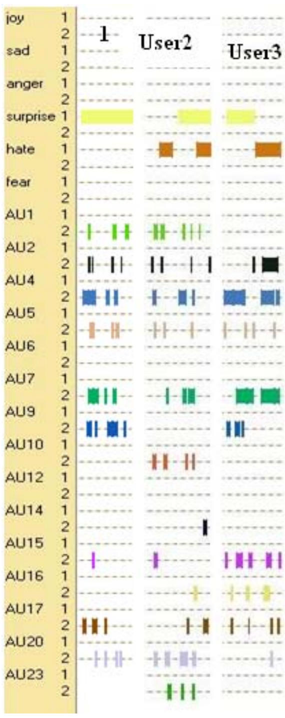


Figure 3. (INTERACT) software usage in video analysis of human behavior.

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). This all is in reference to Kenji own utterance, cognitive thinking extraction, referenced by customs, habits and other self-based extracted personality related to different views analysis. The resulted patterns or templates reflect what we called Kenji Style. Kenji style has been captured using the above-mentioned two case studies approaches. 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 Figure 3, that among 10 observed users we have collected 3 users with 90% match related to the Action units and six Ekmann emotional modes. 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 (Fujita, 2006), (INTERACT, 2007).

3. Emotion Estimation from Facial Expressions of Users

Part 1 of Kenji System presents the hologram, it is as shown in Figure 5. The total image of Kenji Hologram is on Figure 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). The OutputVideo program make automatic connection to realtime generated animation emotional template. KANJO processor (Figure 4-1) is synchronizing the MAYA images that is generating in real-time animated facial images, and synchronized through KServer (Figure 4-2), and the emotional sound file extracted Kenji text (refer to: Figure 8), all this is synchronized through KNAJO processor. All this; is referenced as Part_1 in our system.

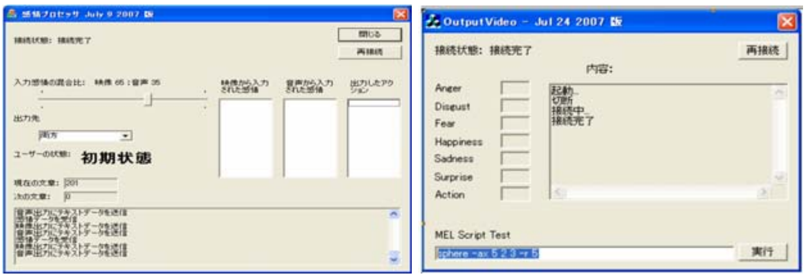


Figure 4-1. KANJO processor *Figure 4-2.* KENJI-Engine Server

Figure 4. Virtual Engine component

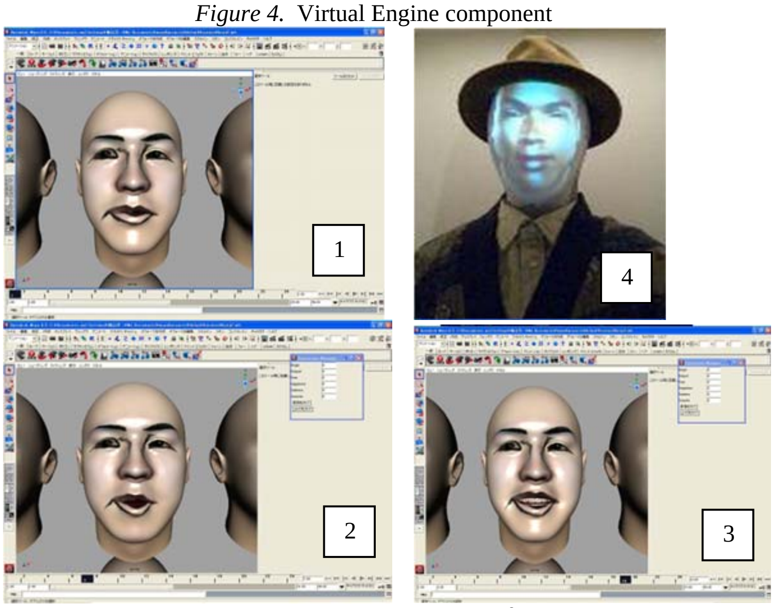


Figure 5. Shows part 1 of the hologram

The details of the software development (Part_1) are omitted for space and technological securities issues related reasons. You can reference to demos on this application by reference to (INTERACT, 2007) or the link (http://www.fujita.soft.iwate-pu.ac.jp/prof_dir/issam/others/KenjiOnly.wmv). Please notice that all output is done in real time. This section will be reference again in Sec.5, and related to creating facial images in harmony with the contents of the spoken text.

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, Ekmann and Friesen, 1975; for a survey of the literature see, Pantic and Rothkrantz, 2000). 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 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 (Hakura et. al., 2007). The approach is able to fully utilize the dynamic aspects of facial expressions.

The approach assumes that the movements of the facial feature points are the consequence of the facial muscles modeled as linear systems with several modes. The modes of the system correspond to the represented emotions, and each identified mode of the system can estimate the movements of the facial feature points in expressing the emotion. As shown in Figure 6, the differences of the estimated movements and actual movements are used to detect the emotional expressions. While we are planning to merge the approach with the FACS approach, yet at the current stage, this section concentrates on the system identification approach. The rest of the section is assigned to the following four topics: Face and feature points' detection's method; Facial expression database realizing system identification approach; Emotion estimation with facial expression database, and an experimental results showing that the proposed method can detect the six basic emotional expressions, i.e., happiness, surprise, sadness, fear, anger, and disgust, with their exact starting points.

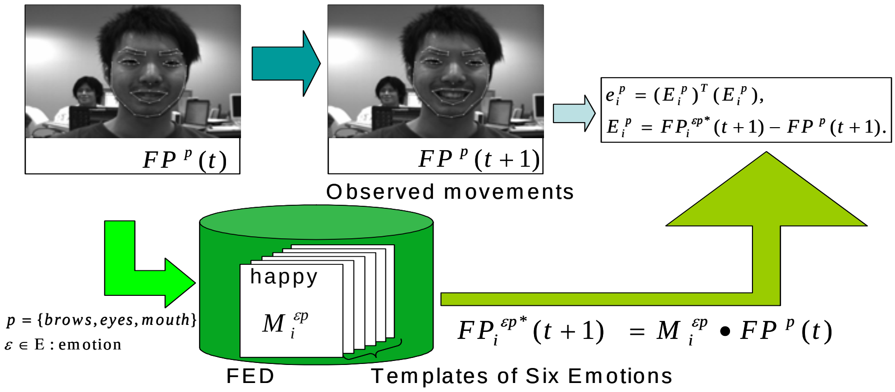


Figure 6. System Identification Approach to Facial Expression Analysis.

4.1. A FACE AND FACIAL FEATURE POINTS DETECTION METHOD

In contrast to our previous results on this project (e.g., Fujita, et.al., 2006 Hakura, et. al., 2007) that require a motion capture system to detect face and facial feature points, we have adopted Active Appearance Model (AAM) (Cootes et. al., 1998) to raise the availability of our system. While there is a defect in the accuracy, there are several advantages in this change such as free from markers (used in motion capture systems), and availabilities of the device and needed real time processing. The AAM is a statistical model/method to detect deformable objects in images. It makes a statistical model of the shapes and appearances from training images with manually added feature points. When applied to the novel image, it finds parameters to adapt the model to the novel image by means of an optimization method. Today's computational advances enable the method to be processed in real time so that it can be applied to object tracking tasks. The points on the facial images in Figure 4 are automatically detected by the implemented AAM system.

4.2. FACIAL EXPRESSION DATABASE (FED)

Facial expressions in this study are considered as the results of movements of facial feature points. Facial feature points are the points that are needed to estimate the target emotional expressions by observing the movements of these points. The movements of the feature points are assumed to be caused by a set of systems as in the Facial Score (Nishiyama, et. al., 2005). The system identification method, i.e., LSM: Least-Square Method is adopted to identify the systems. A face is divided into roughly three parts: eye brows, eyes, and mouth. Each part is assumed to contain its own emotional signals. The results of the identification for each part would be the six transition matrixes; each of those can estimate the movements of the facial feature points for particular emotional facial expressions. The acquired knowledge is collected as a database, called Facial Expression Database (FED).

To extract typical movements of the facial feature points that express the particular emotion, we should collect facial expressions that represent the particular emotion. For this aim, subjects who have trained to express the six emotions act on each emotion for several times, and the movements at that time are labeled as an emotion. As mentioned above, the movements are represented as result of the system's output. The system has some modes; initially, the duration is assumed to be a mode. A merging algorithm of the modes reduces the number of the modes. The similar movements are considered as results of the system in the same mode. Namely, each mode expresses an emotional category. The detail of the algorithm is described in (Fujita, et. al., 2006). As the results of applying the algorithm, each mode corresponds to the most typical movements of these points to express an emotion. For example, when there exists N durations labeled as "fear", the modes that representing fear are described as follows:

$$\mathbf{M}^{\mathbf{fear}} = \{\mathrm{M}^{\mathbf{brows}}, \mathrm{M}^{\mathbf{eyes}}, \mathrm{M}^{\mathbf{mouth}}\}, \quad \mathrm{M}^{p} = \{M_i^p \mid 1 \le i \le m, i \in I, 1 \le m \le N\}. \tag{1}$$

where, $p = \{brows, eyes, mouth\}$, I denotes a set of positive integers. M^{p} is obtained for each of the six emotions.

The modes are represented as the transition matrixes. Namely, every facial feature point can be estimated by using the matrixes. Let eye brows requires six feature points (three points for each brow) at time t denoted as $FP^p(t) = (px_1, py_1, px_2, py_2, \Lambda, px_6, py_6)^T$, where px_i and py_i are the x and y coordinate of i-th feature point respectively. Then, the points at the next time step can be estimated by the following equation:

$$FP_i^{p*}(t+1) = M_i^p \bullet FP^p(t) \tag{2}$$

where, p* means that it is the estimated value. Note that M_i^p is a 12 x 12 matrix in this example. Note also that every estimation value is calculated by the actual value of the facial points. The estimation process compares the error between actual and estimated values of the facial points for every mode, part, and emotion. The following describes our method to do estimation with the modes setting.

4.3. EMOTION ESTIMATION WITH FACIAL EXPRESSION DATABASE

As mentioned in the previous section, FED provides system identifiers of the systems that control the facial feature points. Therefore, emotional estimation from the facial expression with FED uses these identifiers to know to what emotional categories the presented facial expression belongs. The presented facial expressions can be detected as the movements of the facial feature points with a vision system and the AAM, so that $FP^p(t)$ in Equation (2) are available at every time step. Every identifier estimates $FP_i^{p*}(t+1)$ at the next time step through Equation (2). These estimated points are compared with the actually observed points at t+1, and then we can calculate the error e_i^p:

$$\begin{aligned} e_i^p &= (E_i^p)^T (E_i^p), \\ E_i^p &= FP_i^{p*}(t+1) - FP^p(t+1). \end{aligned} \tag{3}$$

Note that e_i^p is a scalar value, and E_i^p is a vector. Then, according to Equation (1), we have $|M^p| (= m)$ errors for each facial part with respect to each emotional category. To determine which emotion is observed, we have to cumulate the error values of each part. We simply employ the minimum value for the aim, because the identifier with the minimum error itself implies that the emotion is detected:

$$e^p = \min_i \{e_i^p\} \tag{4}$$

Namely, we have now error vectors Δ_e consist of three elements, i.e., on eyebrows, eyes, and mouth, for each emotion:

$$\begin{aligned} \Delta_e &= (e^{brows}, e^{eyes}, e^{mouth}), \\ e &\in \{\text{joy, sadness, anger, disgust, fear, surprise}\} \end{aligned} \tag{5}$$

While the error vector can be used as differently, we simply add the errors. Then, the error is used to estimate the expressed emotions: Here we simply set an empirically obtained threshold to each error. Namely, the emotion with the estimation error below the threshold is considered as the expressed emotions.

4.4. An Experiment on detecting emotions

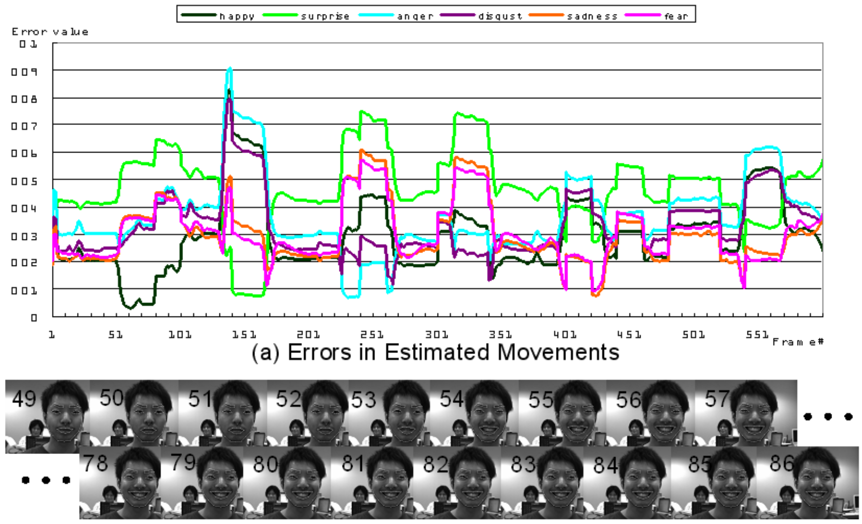
We have implemented the AAM by means of the aam-api (Stegmann et. al., 2003) coupled with OpenCV library (Open Computer Vision Library). The frame rate of the camera is approximately 20 frames per second.

An experiment to check basic detective abilities of the proposed approach is conducted. For this aim, only a subject who acts on six emotions according to the FACS is assumed to be the target person. Thus, the training data for AAM and the FED is prepared only for the subject. After construction of the AAM and the FED, another sequence of the actions is observed. In the observation phase, emotion estimation process works in real time, and the results are sent to the conceptual cognition engine in response to the facial expressions of the subject. To confirm the above-mentioned abilities of the approach, the subject acts on the six emotions in the following order: happy, surprise, anger, disgust, sadness, and fear. The frames corresponding to the facial expressions are listed in Table 1.

TABLE 1. Frames and Acted Emotional Facial Expression.

Acted Emotional Facial Expression	Frames
Happy	51 to 105
Surprise	131 to 170
Anger	222 to 268
Disgust	310 to 346
Sadness	394 to 430
Fear	535 to 569

The result of the experiment is depicted in Figure 7. Figure 7(a) shows the transitions of sums of the difference between estimated movements of the feature points and the actual movements of the points, i.e., estimation error, observed in each frame. Namely, the estimations by means of the modes (matrices) corresponds to the six emotions are compared with the actual movements of the feature points. Therefore, the lower error value means that the corresponding emotional facial expression is observed. Comparing Figure 7 with Table 1 reveals interesting facts. The error values rapidly decrease at the beginning of each act, and in most cases, continues decreasing until the facial expression becomes the typical one that can be recognized with FACS. Figure 7(b) is the actual sequence of the facial expression expressing “Happy” from beginning to the end. It means that the proposed method is able to detect the very beginning of the facial expressions. It is very important for our aim, i.e., construction of conceptual cognition system, because the system should know what makes the user having that emotion. Although some emotions, disgust and happy, and fear and sadness, seem hard to distinguished by the proposed method, the other emotions are detectable with the method. The undistinguished emotional expressions are sometimes very hard to identify even by the human when forced to identify it only from the facial expression. We are planning to merge and/or analyze the information from the other modalities, i.e., voice (as explained in the following Section), situation, and so on, to overcome this issue.



(a) Errors in Estimated Movements

(b) Actual Sequence of "Happy"

Figure 7. Result for Estimating Acted Emotions.

5. Emotional Voice recognition

This Section 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 8, 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 references), which is API (application program interface) with MAYA application. However, the templates generated on in Figure 8 used conventional rules that we revised as in the following.

Researchers use discriminant rules to estimate emotion in human speech in general (Oudeyer, 2003). First, researchers record human utterance that participants speak certain phrases emotionally in role act, like actor or actress. Second, researchers extract sound features, for example power and fundamental frequency using frequency analysis and get mean and maximum value of these features. Finally, they make discriminant rules from these features using machine-learning methods, for example decision tree and artificial neural network. These rules depend on recorded data. The set of such rules are not the same among different systems. Therefore, the correctness rate is not high.. We think it is necessary to improve such method. For the voice emotional recognition system, we propose three methodological steps to improve the referenced method.

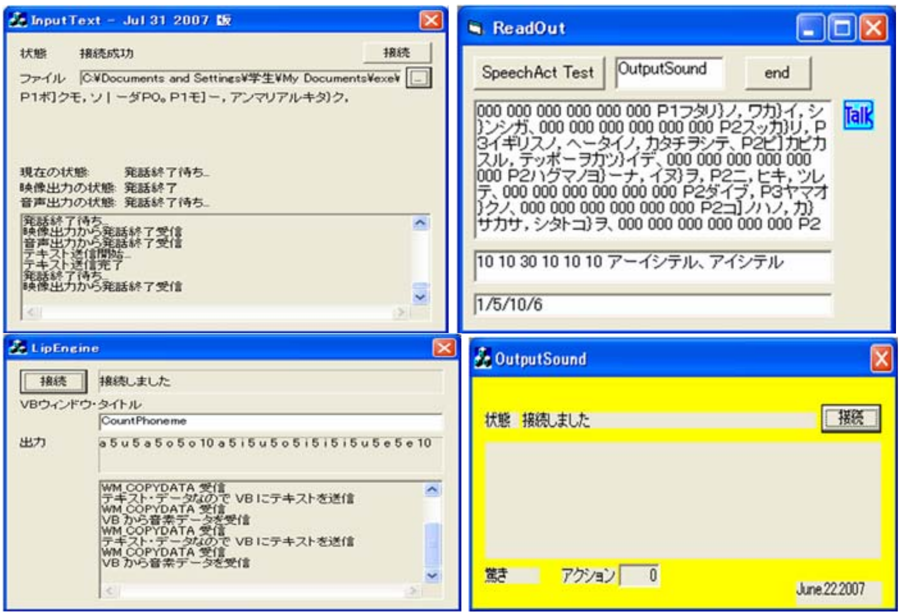


Figure 8. Shows the screens of the Input text window, and readout window

More details in this section have been omitted due to space. However, it should be clear that the voice emotional characteristics have been computed as part of the interface to have consistent emotional integration with intention.

6. Cognitive Scenario generation according to human cognitive state

We explain here part 4 of the system. How to make system act with the user according to the 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 situation 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 the system, which conclude if the user is interesting in the current created scenario or not interesting. Actively engaged or disconnected. There are four states, which the system concludes to reach; according to the decision taken 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 (Tickle, 2003). 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, like architecture design, 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 should include a mechanism to include situations, and user metal background ontological views (vast views: culture views, and mental view and spontaneous views). We human our intellectual communication is not bounded by 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.

7. Conclusions

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. Such anecdotic view based on spatial cognition is essential for architecture design.

This research has other objective in bridging the complexity of communication between human and computer in design issues solutions.

We agree with the reviewers that this research has many different disciplines, but we think this is the issue in design. Design also has different views and integrating it is the creative process itself. This work objective is to have cognitive emotional interface that can be integrated with knowledge base system especially for design in complex systems, like spatial design. The two cognitive spatial spaces (the voice synthesis space, and facial space), contribute to establish the perceptional view of the best harmony of engagement in human user interaction.

We are now integrating with this system colors and music to be part of the presentation to reflect user emotional states in another metaphoric forms. Integrating the emotional space as part in the design is could motivate the design itself. Human mental states can be represented by multidimensional presentation reflected on our five sense, color integration with facial images, sound integration with the generated readout text, all in real-time animated with the transition state of the human user. By the time of the conference in November, we can show a video for our prototype system in the conference.

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, 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.

References

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

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*

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

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

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

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.

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.

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.

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

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

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

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

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

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

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

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.

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.

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

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.

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

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

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.

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

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