論文 / 著書情報
Article / Book Information

題目(和文)	エージェント依存知識を用いた深層的感情理解とその遠隔インタラクションにおける可視化法
Title(English)	Deep Level Emotion Understanding using Customized Knowledge and its Visualization Method in Long Distance Interaction
著者(和文)	GARCIA SANCHEZ JESUS ADRIAN
Author(English)	GARCIA SANCHEZ JESUS ADRIAN
出典(和文)	学位:博士（学術）, 学位授与機関:東京工業大学, 報告番号:甲第9903号, 授与年月日:2015年3月26日, 学位の種別:課程博士, 審査員:廣田　薫,柴田　崇徳,長谷川　修,小野　功,董　芳艶
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第9903号, Conferred date:2015/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Deep Level Emotion Understanding using Customized Knowledge and its Visualization Method in Long Distance Interaction

（エージェント依存知識を用いた深層的感情理解と
その遠隔インタラクションにおける可視化法）

Jesus Adrian Garcia Sanchez

Supervisor: Prof. Kaoru HIROTA

Doctoral Thesis

Tokyo Institute of Technology
Interdisciplinary Graduate School of Science and Engineering
Department of Computational Intelligence and Systems Science

February 2015

Deep Level Emotion Understanding using Customized Knowledge and its Visualization Method in Long Distance Interaction

（エージェント依存知識を用いた深層的感情理解と
その遠隔インタラクションにおける可視化法）

Garcia Sanchez Jesus Adrian
ガルシア　サンチェス　ヘスス　アドリアン

Supervisor: Prof. Kaoru HIROTA

Doctoral Thesis

Tokyo Institute of Technology
東京工業大学
Interdisciplinary Graduate School of Science and Engineering
大学院総合理工学研究科
Department of Computational Intelligence and Systems Science
知能システム科学専攻

February 2015

Abstract

A deep level emotion understanding method is proposed to facilitate better human-robot communication, where customized learning knowledge of an observed agent (human or robot) is used with the observed input information from Kinect. It aims to realize agent dependent emotion understanding by utilizing special customized knowledge of the agent, rather than ordinary surface level emotion understanding that uses visual/acoustic/distance information without any customized knowledge. In the experiment employing special demonstration scenarios where a company employee's emotion is understood by a secretary eye robot equipped with a Kinect sensor device, it is confirmed that the proposed method provides deep level emotion understanding different from ordinary surface level emotion understanding. The proposal is being planned to be applied to a part of the emotion understanding module in the demonstration experiments of an ongoing robotics research project entitled "Multi-Agent Fuzzy Atmosfield".

A visualization method of users' emotion information is proposed for long distance interaction such as telecommuting and distance learning, where 3D emotion vectors in Affinity Pleasure-Arousal space are illustrated by using shape-brightness-size (SBS) figure. It gives users easily understandable emotional profile information, and provides administrator strategic suggestion to improve the interaction between the users and the system. In the matching experiment between 7 basic emotions and 7 SBS figures for 8

subjects, 83.93% matching is achieved, and the administrator finds contents improvement hint in the questionnaire of emotions for 5 reading-text-tasks by 8 subjects. It is planning to be implemented in a language learning application to provide more comfortable learning experience by contents selection based on user's emotion.

The application of deep level emotion understanding is proposed for distance learning, where 3D deep level emotional emotion concept and adaptive e-learning are combined. The deep level emotion is processed and analyzed to control the adaptive e-learning system providing the best matching content, studying form and practice way to the learning, as well as giving to the administrator reports of the content ranking by the user perception. In the matching experiment, the comparison between the normal system and the proposal is done via web application in an English-German learning language, where users used the system while it is adapting to them, also a report page for the administrator localize the contents to improve. It is planning to adapt it and implement it in different language learning applications.

Acknowledgements

In the first place I will like to express my sincere gratitude to my professor Hirota Kaoru, for accepting me in his laboratory, and all his guidance during my master and doctoral course studies in Japan. His encouragement and inspirational words helped me to move and continue with my studies and research even in the most difficult times. Also I will like to express my sincere gratitude to Professor Shibata Takanori, Professor Hasegawa Osamu, Professor Ono Isao, and Professor Dong Fangyan for their priceless time, valuable comments and advices that have guided me through my doctoral thesis, and I could not forget all the support in my school life of Ms. Hoshino Harumi, all my gratitudek for you.

I would also say thank you so much to my seniors, Phuc Quang Le, Abdullah M, Iliyasu, Chastine Fatichah, Martin Leonard Tangel, Fei Yan, Bo Sun, Yongkang Tang, Adam Witold Pedrycz, Ricardo Antonio Sotolongo Trujillo, Clement Nthambazele Nyirenda, Vu An Hai, Kazushi Okamoto, Hrvoje Markovic and Janet Pomares Betancourt, they shared with me their knowledge and experience me. Also I want to say thank you to my first friends in Japan, Momoko Matsushita, Daisuke Masano, Masashi Yamaguchi, Shibata Atsushi, and Toshihiro Akamatsu, they made my arrival to Japan very warm and make me feel like in home.

I am very grateful to not only laboratory mates but also friends in Hirota Laboratory for all the big lessons and tasks during my studies, Shibata Atsushi, Akamatsu Toshihiro, Ohnishi Kazihiro, Luegeng Chen, Hidenori Sakaniwa, Chehung Lin, Jiajun Lu, Tianyu Li, Maslina Brinti Zolkepli, Takaaki Iwamasa, Shohei Kosaka, Hiroaki Doi, Kyoda Masuyama.

I would like to thank my family and friends in Mexico for their support and encouragement. Thank to my mother (Maria Guadalupe Sanchez Ortega) to bringing to this world, to my grandmother (Francisca Ortega Rodriguez) that passed way while my first year in Japan, my sister (Nora Patricia Garcia Sanchez) for always supporting me in every challenge in my life, my brother (Jose de Jesus Garcia Sanchez) who always challenge me to be more, and to my uncles, cousins and nephews for their patience and prayer to support me to in spirit to finish my studies in Japan. Thank you so much for your love and support to me. I want say especial thanks to my friends, Zahara Castaneda, Victor Perez, Alberto Solis, and Eiji Yokoo, they always cheered me up.

Finally, I will like to give my deepest gratitude to the Japanese Government and Japanese people for the support via the Monbukagakusho Scholarship Program. The experience of living and learning in Japan changed my life forever, I will be eternally grateful.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Emotions are very important in the normal life of human being, helping in the communication and interaction between them. Humans and animals show emotions as part of their normal behavior, creating strong links or alerting other of something wrong is happening. The actions and reactions are important in the acceptation of any individual in a group, how well this individual understands others is how the perception of kindness, adaptation, and empathy of any person, that is why understanding emotion is an important task in the communication between any human, but understanding emotions is a complicated task for computers, so the affective phenomena represents a challenge for the modern computer science [1].

The definition of emotion [2] is

"A strong feeling deriving from one's circumstances, mood, or relationships with others"

As is read in the definition, emotions are important for the relationship between individuals, and they depend of the circumstances and the mood of the person. The issues to understand and to model an emotion is often complex. Modelling and recognizing emotions is important in the interaction with any human. Interdisciplinary approach from philosophy, psychology, neuroscience, and biology is necessary to understand them [3].

Emotions are not general in the world, each culture have a unique way to express and understand them [4]. Emotions are so different in the each culture that even the emotion related with a color is different too [5].

How emotion should measure should be dynamic, culturally mediated, socially constructed and experience, understanding, interpreting, and having the emotion in its full ambiguity and complexity [6], system should learn from each person when they interact, gain experience and then use that experience in the futures interactions.

Different approaches about how to obtain the emotions are being investigated with inputs [7] [8] such as: cardiovascular system [9], electro dermal activity [10], respiratory system [11], brain activity [12], muscular system [13], and activity capture in video camera media.

1.1 Background

There are different studies about the recognition of the human emotions, and divided in two different categories:

A) Invasive methods: those that need to plug devices, cables or any gadget to the human to take some signal to obtain the emotion, examples of the devices are electro dermal sensors, brain activity sensors, etc.
B) Non-invasive methods: those that do not need a direct contact or attachment with the person, examples of this are microphones, cameras, thermo sensors cameras, etc.

A problem with the invasive methods is that emotions are very susceptible to be influenced and changed by the circumstances, implying that the least invasive method, it should be the best method.

Different emotion studies are available, where the use of the voice [14] [15] [16], facial expressions [17] [18] [19], body gestures [20] and also combination of them [21][22], had make great progress in the recognition of extreme human emotion like anger, disgust, fear, happiness, sadness, and surprise, but those are not the normal emotion in humans, the representation need to be more flexible due the complexity of the emotions. That is why, Deep Emotion Level Understanding are proposed and tested. Deep Emotion Level Understanding is a way to simulate the form of how humans understand each other by not only using the physical expression of the emotion, but also using customized knowledge about the person to achieve a closer emotion to the real emotion of the person.

1.2 Emotion visualization

The visualization of the data is very important part to confirm achievements, not only for experts, but also for users without any special training [23]. An effective emotion visualization method is necessary from the point of view of affective computing to developing and realizing smooth interaction. Some approaches to visualize emotions have been studied [1][2] [24], where the visualization of the emotion plays an important role in the design the system, and trying to create a communication to the users, but those methods are not easy to understand emotions intuitively and/or completely compressible. Also human like approaches [3], trying to imitated a humanlike facial expressions, that result in the rejections because they are not sufficient human.

1.3 Deep Emotion Level Understanding Applications

The proposed method to achieve deep emotion level understanding can be used in any of the current application emotions like costumer services, video games, etc. The usage of the emotion in long distance interaction is studied and propose some implementation to improve the interaction between user – system, system – administrator, and user – administrator.

The main approach is to used and adaptive distance learning with deep emotion level understanding engine to create a profile of the learner with the preferences, strong points, weak points information to provide the appropriated information to the learner to improve his studies. Distance Learning is taking a very important part in the modern life, big companies and opportunity to business opened a big interest in this area, being able to reach all areas, without having the limitation of space of time, it is just wonderful. But distance learning have a problem is the adaptation to the different users around the work, and how to support their experience actively, while getting static of the content.

In the chapter 2 the concept of deep level emotion understanding, the details of the features and the test environment is presented, also some scenarios are proposed to corroborate the theory. A visualization method of the emotion information for long distance interaction, the proof of the efficiency of its understanding by the general and expert users, and some possible features of how to analysis the outcome is made in chapter 3. The combination of the deep level emotion understanding and adaptive e-learning system is introduced, and how a possible implementation and application could be designed in chapter 4.

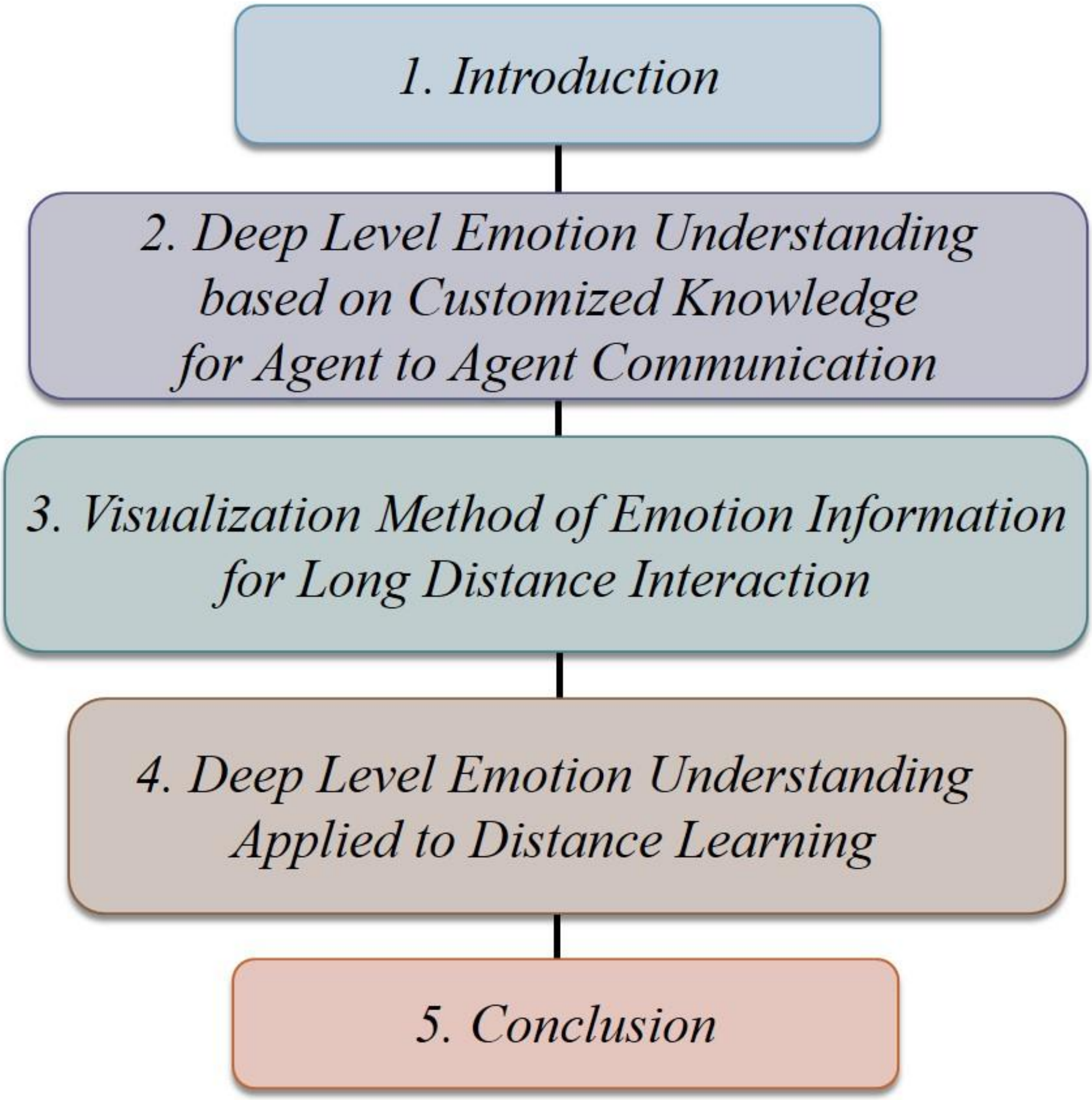


Fig. 1.1. Research roadmap

Chapter 2

Deep Level Emotion Understanding using Customized Knowledge for Human-Robot Communication

2.1. Studies on Emotion Understanding and Importance of Customized Knowledge

Emotion recognition has been studied in human-robot communication using different types of devices [25], the ones that are more close to a human way of understanding the emotions are those which are based on voice, face, and body gesture information [26] [27]. Different researches have been done to make this realizable; most of them based on the facial expressions from the eyes and mouth [28]. Others include the speech [29], gestures [30] [31], or a combination thereof [22] [32], but the missing part from these approaches is the lack of experience [33]; learning from interactions and creating knowledge is what gives humans the power to deeply understand the emotions of other people. Human emotions are complex and in many situations the emotion displayed in the face, voice or body gesture sometimes may not indicate the real or absolute emotion of the individuals [27], making the need to create an algorithm to model this human ability in order to improve human-robot communication [25].

To address and model this problem, a deep level emotion understanding method is proposed for human-robot communication, where customized learning knowledge from communication history and a basic knowledge base about the observed agent are utilized with the observed visual/acoustic/depth information input. In this proposal, the voice, facial image, and body gestures are captured by using a Kinect device. Each input is fed to a corresponding neural network to obtain a 6 dimensional [0, 1] vector representing six basic emotions: anger, disgust, happiness, fear, sadness, and surprise. The three emotion vectors obtained are transformed to fuzzy memberships that are to be combined with the customized knowledge about the observed agent to create the 3D deep emotion vector in the affinity pleasure-arousal space [34]. After the final emotion information is obtained, the knowledge about the observed agent is checked to determine whether a small modification is necessary or not.

By using visual/acoustic/depth information about the observed agent, emotion understanding may be possible to some extent but it is a surface level understanding because surface level face/voice expression may sometimes be different from the real emotion. If, however, the observer agent has enough customized knowledge about the observed agent then the real emotion may be understood by taking the situation and the agent customized knowledge into consideration which is called deep level emotion understanding.

To validate the proposal, two demonstration scenarios are created. The scenarios involve communication in a Japanese company setting where an interaction between a human employee (observed agent) and a robot secretary (observer agent), who is supposed to have enough customized knowledge about the employee. The employee's face, voice, and body gestures are captured using a Kinect sensor device attached to the robot secretary. The communication topic is a 'meeting room reservation' request made by the employee to the secretary and is later modified because of an earlier mistake made by the employee.

The surface level emotion understanding and the possibility of deep level emotion understanding are investigated by utilizing knowledge of the observed agent in 2.2. A deep level emotion understanding method is proposed in 2.3. Demonstration scenarios to confirm the validity of the proposal are tested in 2.4.

2.2. Surface Level Emotion Understanding vs. Deep Level Emotion Understanding

In a multi-agent society consisting of many humans and many robots, studying human-robot communication is essential [26], understanding of emotions plays an especially important role. Recognizing and understanding the emotions of other human beings are easy tasks for human brains, but not for the robots [27]. To realize emotion understanding functions for robots, consider human to human communication from a view point of emotion understanding. Normally, when two people are talking, they understand the emotional state of each other by using two senses: sight and hearing to recognize the voice, facial expressions, and body gestures as shown in Fig. 2.1. When humans meet for the first time they use general rules to understand the emotions of each other, thereby, creating a first impression of one another. After knowing each other for a long time and becoming more familiar, the observing person starts to deeply understand the observed person's emotions by using the experience and acquired knowledge that in this case is called customized knowledge.

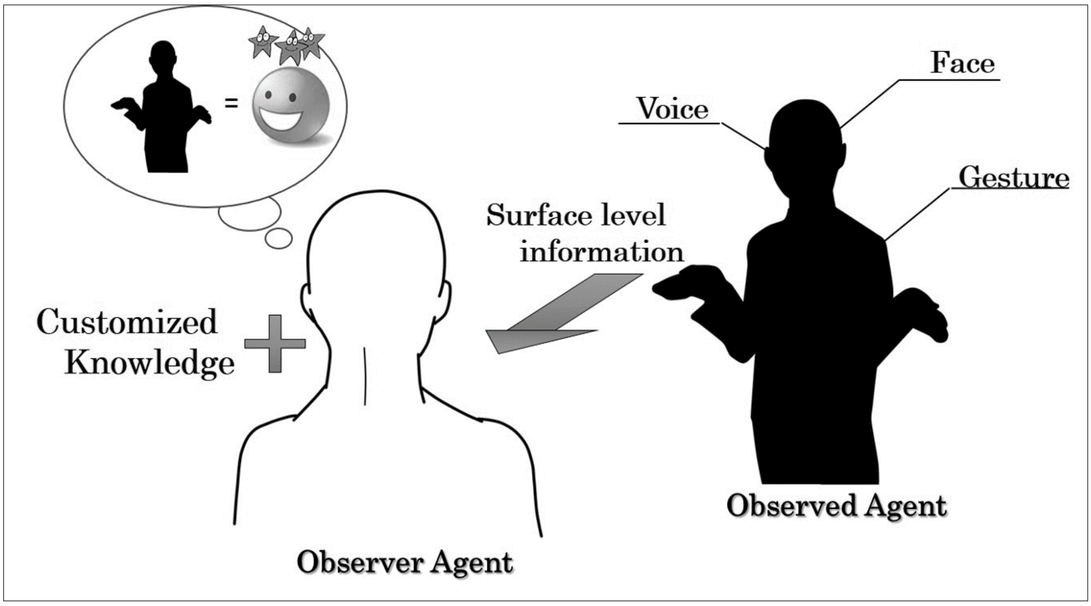


Fig. 2. 1. Two humans meeting for the first time, they use general rules to know the other's emotion, but they start creating custom knowledge for that specific person

Because of these human behaviors, two types of emotion understanding are observed: the first is surface level emotion understanding and the second is deep level emotion understanding. Fig. 2.2 makes a graphical visualization of this concept by comparing human behavior to an iceberg, where on the surface level, only the physical expressions and oral communication are obtained, but under the water the larger hidden part of the iceberg contains the way of thinking and feeling like perceptions, intension, beliefs, knowledge, atmosphere, and emotions of each individual [25].

Available research on emotion recognition has been confined to how to realize and understand human emotions for long time. They have utilized different types of approaches of how to analyze the facial features [28], the voice [29], and the body gestures [30] [31] described in Fig. 2.1. This information is only for the surface level understanding, it just gives an output as the emotion. Another problem with using the surface level emotions is that the same method and parameters are used for every person who interacts with the system.

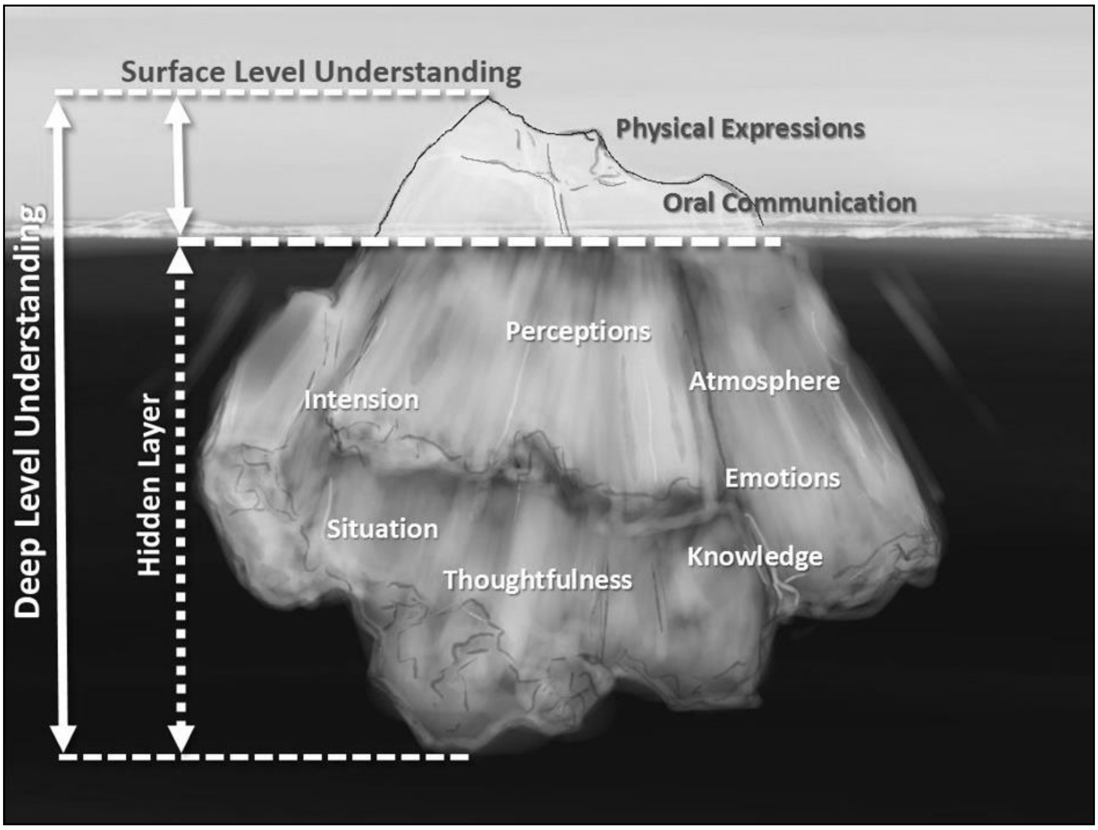


Fig. 2. 2. The iceberg illustrating the complexity and levels of emotion in the human communication

This is similar to what people do when meeting for the first time, but this approach is always general and will never improve. On the other hand, people are different and each one has a different way to express themselves. In addition, after some interaction humans start to understand and know the specific way of how other person expresses themselves, even predicting what the reaction will be to a circumstance. This is deep level understanding. That is what the proposed method is trying to simulate: a human way of understanding the emotions, learning, and knowing what characterizes a person and their unique way of expressing themselves.

2.3. Deep Level Emotion Understanding Utilizing Customized Knowledge of the Observed Agent

As described in 2, the idea of creating a deep level emotion understanding is based on the utilization of customized knowledge in the interaction between two agents, the observer and the observed. This information is updated and tuned with each interaction, until it reaches the level of a "well-known" observed agent, where technically no changes will be needed to the customized knowledge anymore for that observed agent. The diagram of the proposed method, showing its 3 steps is represented in Fig. 2.3. The first step is the surface information acquisition consisting of three engines: the voice emotion extraction, the face emotion extraction, and the body gesture emotion extraction. The second step is the combination method where the surface emotions and the customized knowledge are used to calculate the deep emotion. The third step is the visualization and graphics where the final results are recovered.

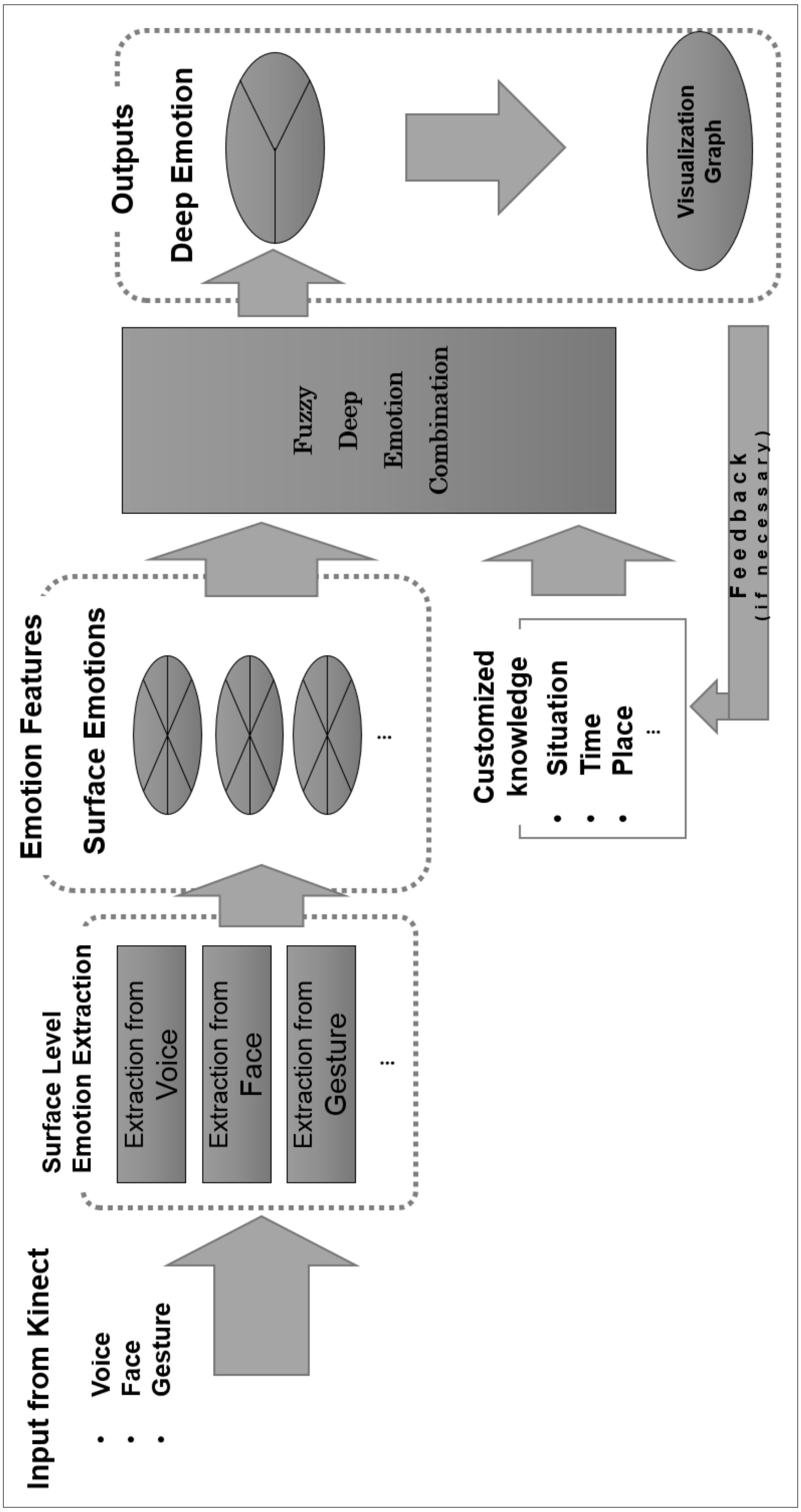


Fig. 2. 3. Deep level emotion understanding method diagram

The steps to obtain the deep level understanding are to get the Kinect inputs from the acoustic and visual sources. To extract the voice emotion energy entropy, short time energy, spectral roll off, spectral centroid, and spectral flux features are used. To extract the face emotion lower eyebrow, raiser eyebrow, upper lip, lip corners depressor, lip stretcher, and lower jaw features are used. To extract the body gestures emotion the head, hands, elbows and shoulders features are used. Each input feature is fed to the corresponding neural network to obtain a 6-dimensional [0, 1] vector representing six basic emotions: anger, disgust, happiness, fear, sadness, and surprise [22] [32] [33], which represent the surface emotion labels for the surface level emotions. Those values are fuzzyfied and combined with the customized knowledge using MAX-MIN, weight and shift values. The output is an Affinity Pleasure-Arousal 3D vector, and represents the deep level emotion. The deep level emotion is displayed in the Affinity Pleasure-Arousal space [34]. After, the method starts the feedback process to create/update the customized knowledge if it is necessary. It is a type of profile for a specific observed agent and it is created at the first interaction.

The customized knowledge is constructed in an XML file with the identification information that consists of the collection of specific data for each different observed agent. The customized knowledge of the observed agent is stored in the observer agent based on his/her/its experience and acquired knowledge. It may consist of, type of person, neutral state, key words, key faces, key gestures, situation, time, place, and so on. Using the customized knowledge the weight and shift values are calculated and implemented in the fuzzy deep combination algorithm. The shift values are determined by the neutral emotion state and are expressed by a triplet (sv_1, sv_2, sv_3) in an emotion space $[-1, 1]^3$. It is given by the average of the emotion output of the system of the observed agent, where the neutral emotion state is achieved after multiple measurements of the emotion of the observed agent in neutral state (normal behavior). The average values are saved in the customized knowledge. For example, if a person had a high tone of voice, the voice emotion will show high arousal and less affinity. The weight values (w_1, w_2, w_3), where $w_1 + w_2 + w_3 = 1$, represent the adaptation of the algorithm to the way of the emotion communication of the specific agent. A human usually uses one main channel (voice, face, or body gestures) to show their emotions. It depends generally on the ethnicity or regions

[35] [36], for example: Asian/ Japanese people tend to use more their voice to communicate their emotions, while European/Italian people use more body gestures. The weight values are given by the designer step by step based on the profile of the observed agent, e.g., in the experiment of Japanese observed agent more priority is assumed and assigned to voice. Fig. 2.4 shows the fuzzy deep emotion combination function where the deep knowledge is being used.

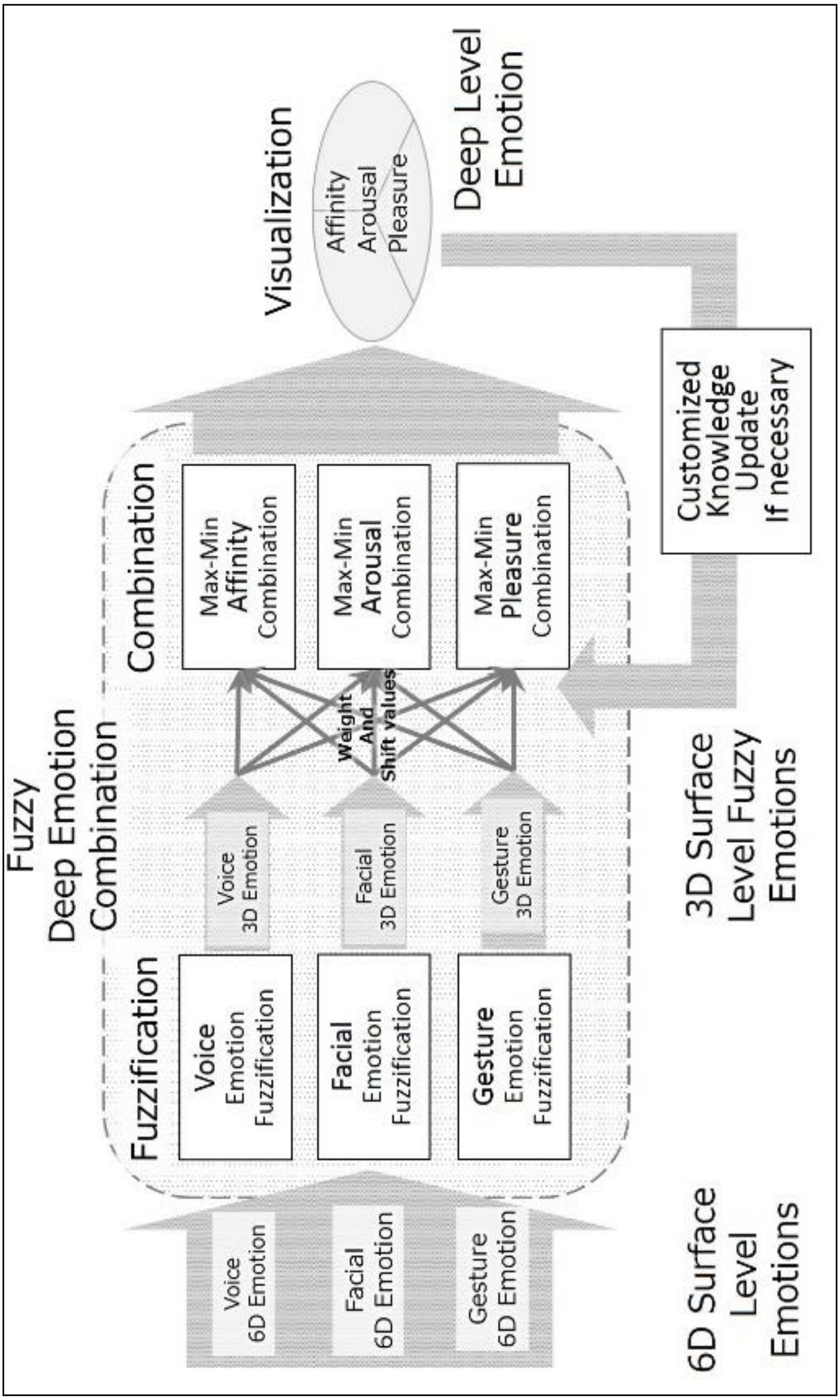


Fig. 2. 4. Fuzzy deep emotion combination algorithm diagram

All emotions are represented in the affinity pleasure-arousal space as,

$$E = \left(e_{affinity},\ e_{pleasure},\ e_{arousal}\right)$$
$$\forall e_{affinity}, e_{pleasure}, e_{arousal} \in [-1,1]^3, \qquad (1)$$

where E is the emotion state, $e_{affinity}$, $e_{pleasure}$, and $e_{arousal}$ are the attributes for "Affinity-No-affinity", "Pleasure-Displeasure", and "Arousal-Sleep" axes, respectively. Each surface level emotion is transformed from one of the 6 basic emotions to a point in the affinity pleasure-arousal space expressed by Eq.(1). In the case of the surface emotions, the *SEL* (surface emotion labels) from each 3 layered feed forward neural network, i.e., 6 inputs 6 outputs for the face, 6 inputs 6 outputs for the voice, and 12 inputs 6 outputs for the body gesture, are expressed by 6 binary vectors as,

$$SEL = \{angry, disgust, happyness, fear, sadness, surprise\}$$
$$\in \{0,1\}^6, \qquad (2)$$

where elements of *SEL* are binary emotion labels (anger, disgust, happiness, fear, sadness, and surprise). In the proposed method, there are three SEL labels from the face, voice, and gesture. In the fuzzification process, each *SEL* label is converted to *SE* (surface emotion) in 3D space of affinity pleasure-arousal using fuzzy sets to represent each emotion based on across-cultural circumplex [4] [37], by membership functions μ(*) as,

$$SE = \mu(SEL)$$
$$= \left(se_{af}, se_{pl}, se_{ar}\right) \in [-1,1]^3, \qquad (3)$$

where se_{af}, se_{pl}, and se_{ar} are the surface emotion affinity, surface emotion pleasure, and surface emotion arousal, respectively. To calculate the surface emotion values membership as shown in Fig. 2.5, Fig. 2.6, and Fig. 2.7 are used.

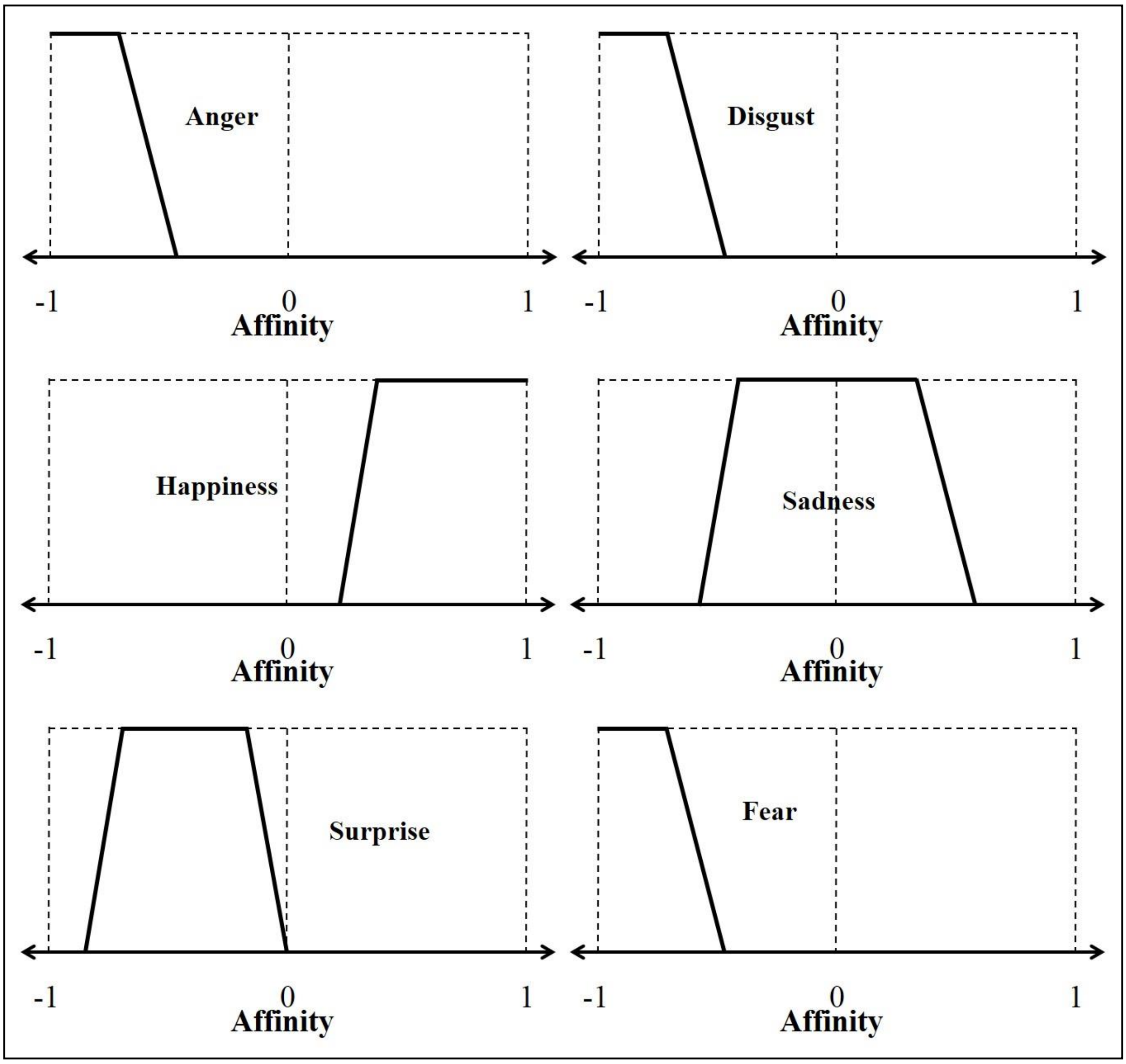


Fig. 2. 5. Affinity axis membership functions for the surface emotion labels

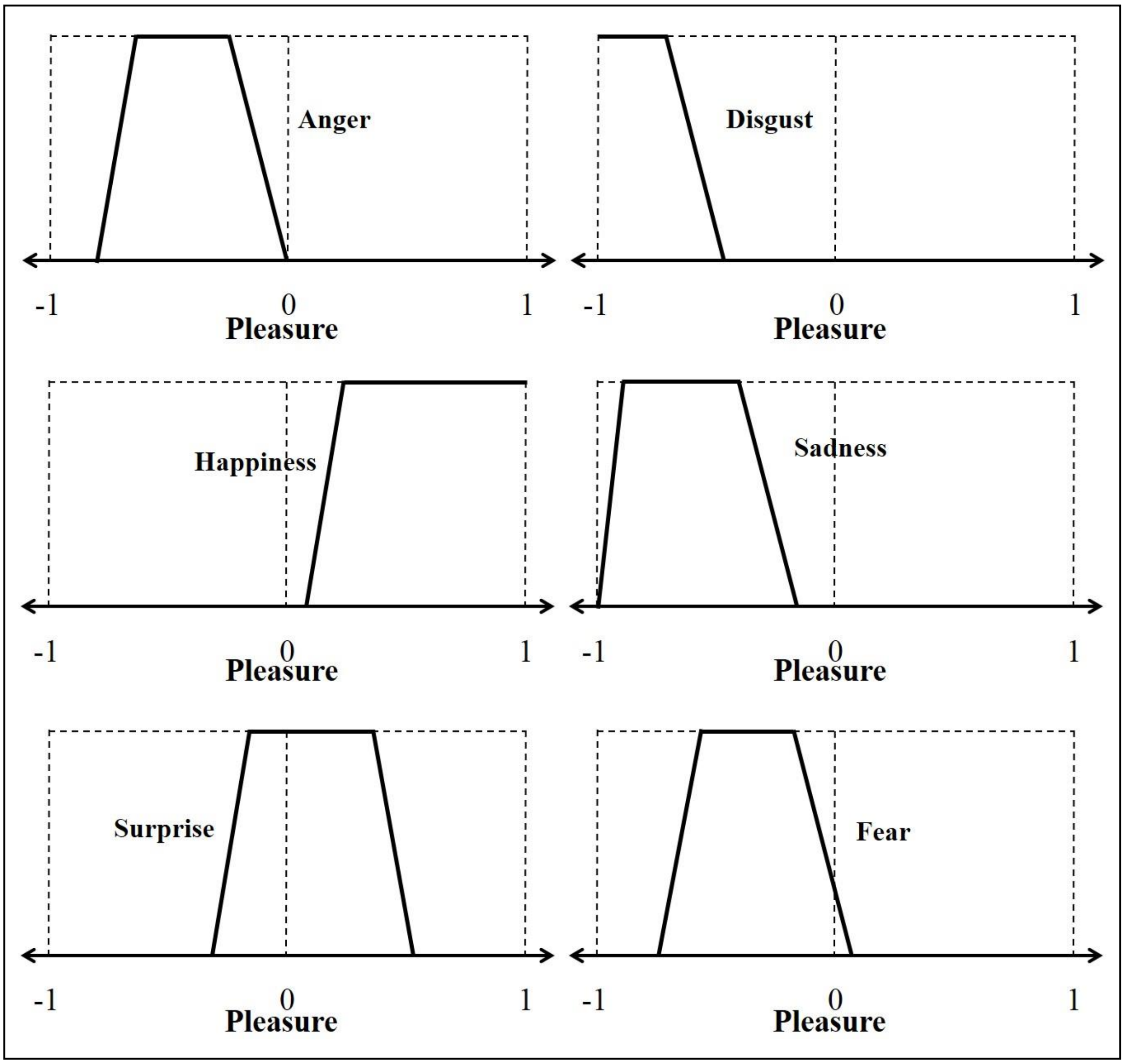


Fig. 2. 6. Pleasure axis membership functions for the surface emotion labels

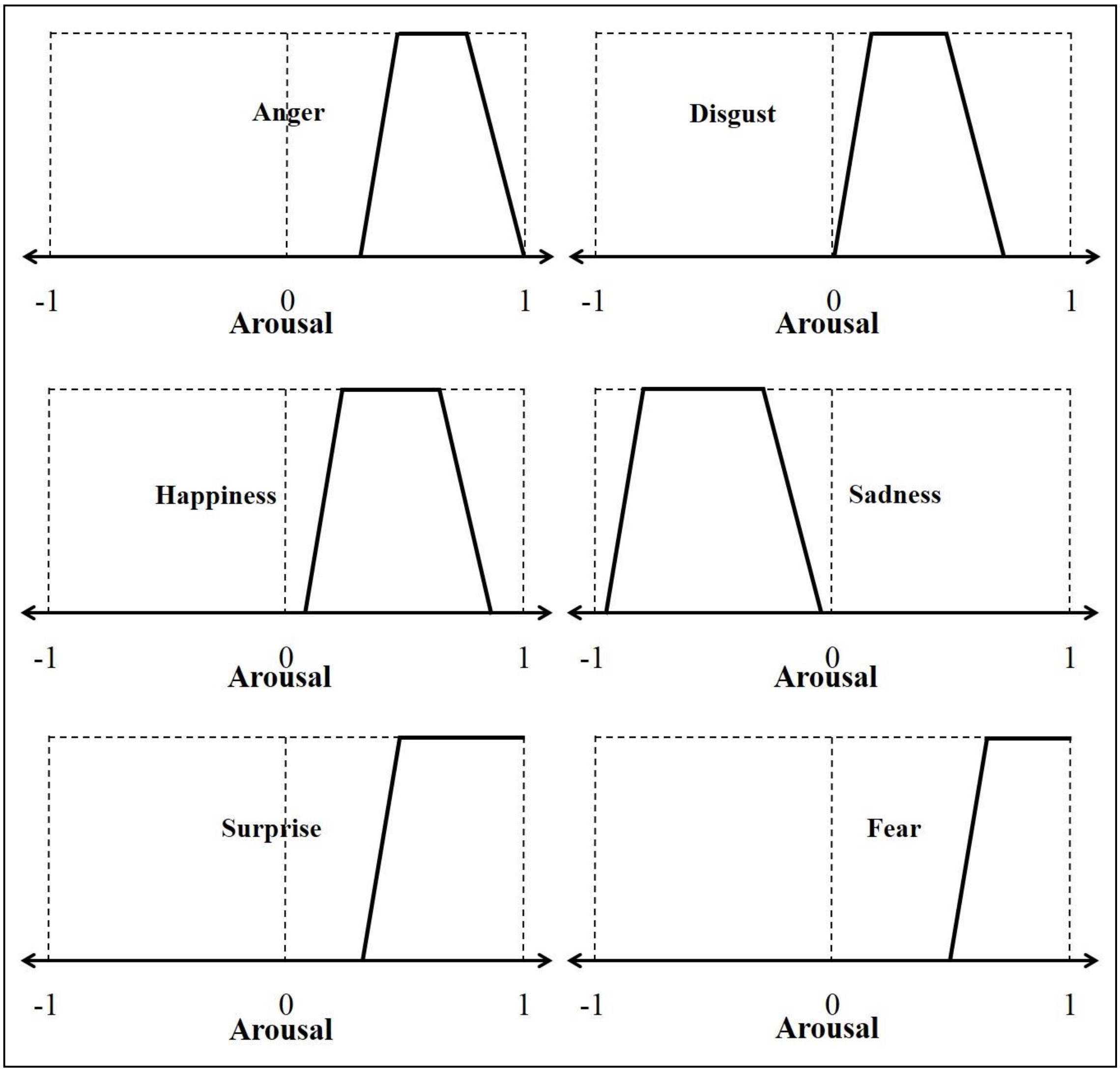


Fig. 2. 7. Arousal axis membership functions for the surface emotion labels

In the fuzzy deep emotion combination process, each value of DE (deep emotion) vector is calculated with SE from each device n as,

$$DE_{af} = \frac{\int af \max_{n \in N} \left(w_n se_{af\,n} \right) daf}{\int \max_{n \in N} \left(w_n se_{af\,n} \right) daf} + sv_{af}, \quad (4)$$

where af is a variable of affinity axis, n is one device out of all devices N, w_n is the weight value for the specific input based on the customized knowledge, $se_{af\,n}$ is the membership in affinity axis, sv_{af} is the shift value for the affinity axis,

$$DE_{pl} = \frac{\int pl \max_{n \in N} \left(w_n se_{pl\,n} \right) dpl}{\int \max_{n \in N} \left(w_n se_{pl\,n} \right) dpl} + sv_{pl}, \quad (5)$$

where pl is variable of pleasure axis, n is one device out of all devices N, w_n is the weight value for the specific input, $se_{pl\,n}$ is the membership in pleasure axis, sv_{pl} is the shift value for the pleasure axis, and

$$DE_{ar} = \frac{\int ar \max_{n \in N} \left(w_n se_{ar\,n} \right) dar}{\int \max_{n \in N} \left(w_n se_{ar\,n} \right) dar} + sv_{ar}, \quad (6)$$

where ar is variable of arousal axis, n is one device out of all devices N, w_n is the weight value for the specific input, $se_{ar\,n}$ is the membership in arousal axis, sv_{ar} is the shift value for the arousal axis.

In the equations (4), (5), and (6) the components from the surface emotion from each neural network are multiplied by the weight value (w_n) depending on the customized knowledge, the *max* value of components is detected, its centroid is calculated, and, finally the corresponding shift value is applied based on customized knowledge.

The resulting values in equations (4), (5), and (6) are the components of the deep emotion affinity pleasure-arousal space axes. Also the standard deviation between the surface emotions is calculated to create a distorted cone to show where the deep emotion is located. The emotion centroid (deep emotion) and standard deviation are plotted in the affinity pleasure-arousal space as in Fig. 2.8.

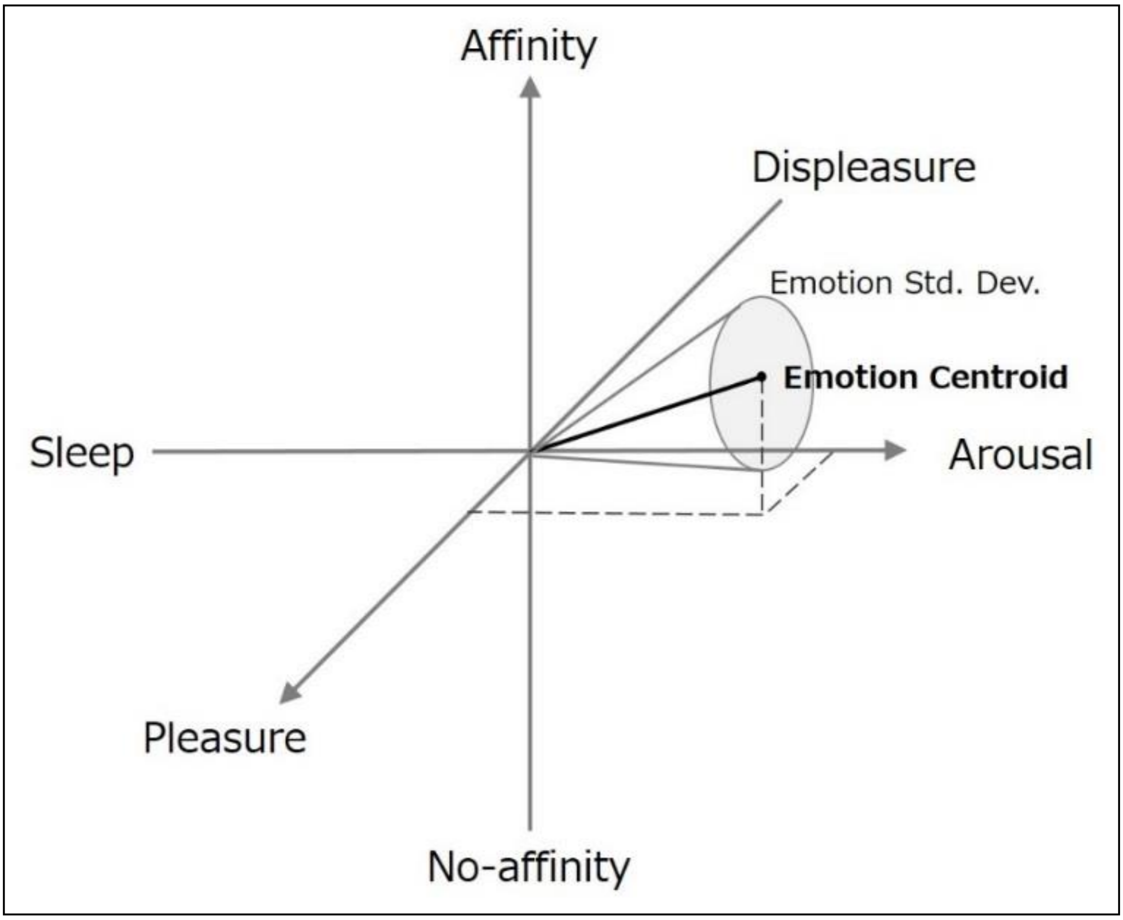


Fig. 2. 8. Representation of deep emotion output where the distorted cone and the centroid of the emotion are shown

2.4. Demonstration Scenarios for Human-Robot Communication

Two demonstration scenes are created to validate the proposed method. Therein, communication between a human employee (observed agent) and a robot secretary (emotion observer) in a Japanese company setting is mimicked as shown in Fig. 2.9.

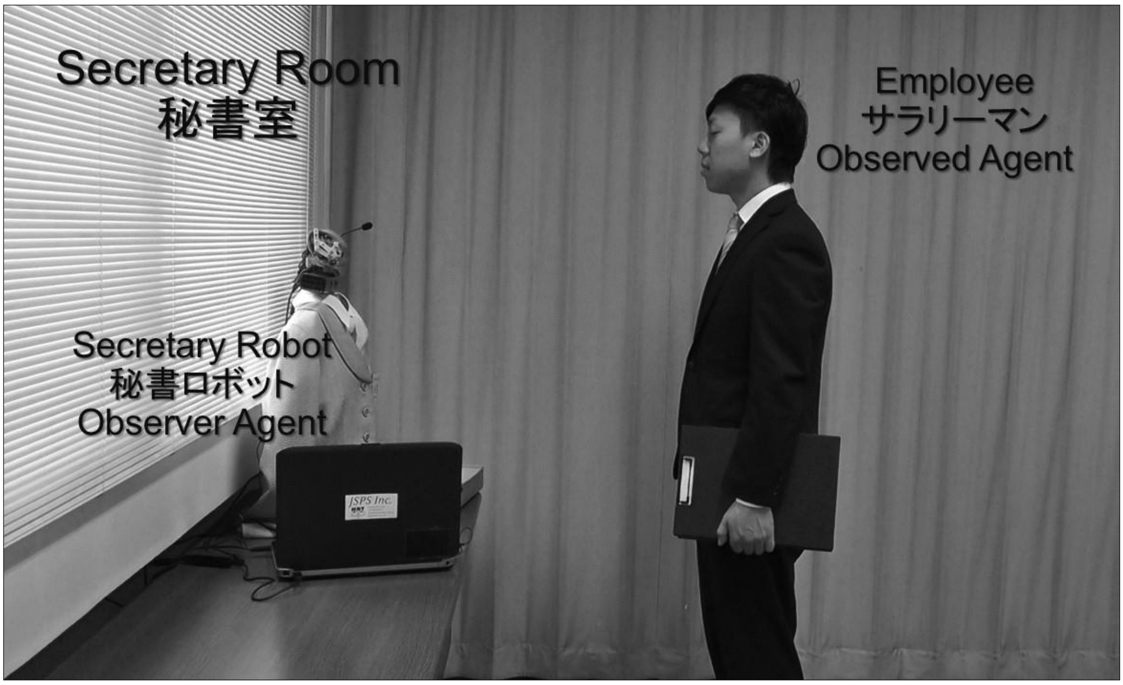


Fig. 2. 9. Scenario consisting of a secretary robot (observer agent) and human employee (observed agent)

The employee's face/voice/body-gesture are captured by Kinect sensor device attached to the robot secretary. The topic is a 'meeting room reservation', where a reservation made by the employee is subsequently modified because of a mistake made in the earlier reservation. This script is created and recorded in Japanese language, and the translated version in English of the script is presented in Table 2.1.

Table 2.1. Script of the interaction entitled "a routine of beloved employee", where a meeting room reservation conversation between an employee and the secretary (robot) is done.

Scene 1:
Employee: *Are you busy?* ***Secretary:*** *No, would you like to reserve a room?* ***Employee:*** *Is the meeting room for 10 people vacant at 3 o'clock this Thursday?* ***Secretary:*** *They are available from 15:30.* ***Employee:*** *Great. A quiet room is preferable.* ***Secretary:*** *How about the regular conference room on the 17th floor?* ***Employee:*** *Sounds good! It's for a remote conference with the branch office, please reserve it until 17 o'clock.* ***Secretary:*** *In addition, I will reserve the video conference system, too* ***Employee:*** *Thanks!* ***Secretary:*** *You are welcome. Good luck.*
Change of Scene:
The employee later realizes his mistake after a conversation with his manager. The employee goes back to the secretary room to change the date of the earlier reservation.
Scene 2:
Employee: *Are you busy now?* ***Secretary:*** *Go ahead!* ***Employee:*** *Excuse me, I want to change the meeting with the branch office to next Thursday.* ***Secretary:*** *Sure! I will check it now.* ***Employee:*** *Yes, please.* ***Secretary:*** *For next Thursday, all the conference rooms on 17th floor have been scheduled already. How about the 11th floor conference room?* ***Employee:*** *Great. I feel relieved.* ***Secretary:*** *I will also update the reservation of the remote conference system.* ***Employee:*** *Thank you very much.* ***Secretary:*** *You're always welcome.*

In this scenario it is assumed that the secretary robot and the employee have already known each other for a long time, this means that the customized knowledge about the employee has already been checked, adjusted, and updated, about the employee. In scene 1, the observer agent senses the normal behavior of the employee and understands the importance of the meeting, and assigns the necessary priority. In scene 2, the secretary robot is able to understand deeply the emotions of the employee, and then uses more adequate words to try to calm the observed agent.

To capture the inputs for the proposed method, a Kinect motion sensor device is used. Kinect sensor has different libraries that make the detection and extraction with enough precision/speed [38]. To capture and record the audio, the Audio library is used, for the facial features the Face Tracking library is used, and the Skeletal Tracking library is used to get the information of the head and the upper extremities [39]. Fig. 2.10 shows the output image of Kinect sensor of the Skeleton and Face Tracking in the center.

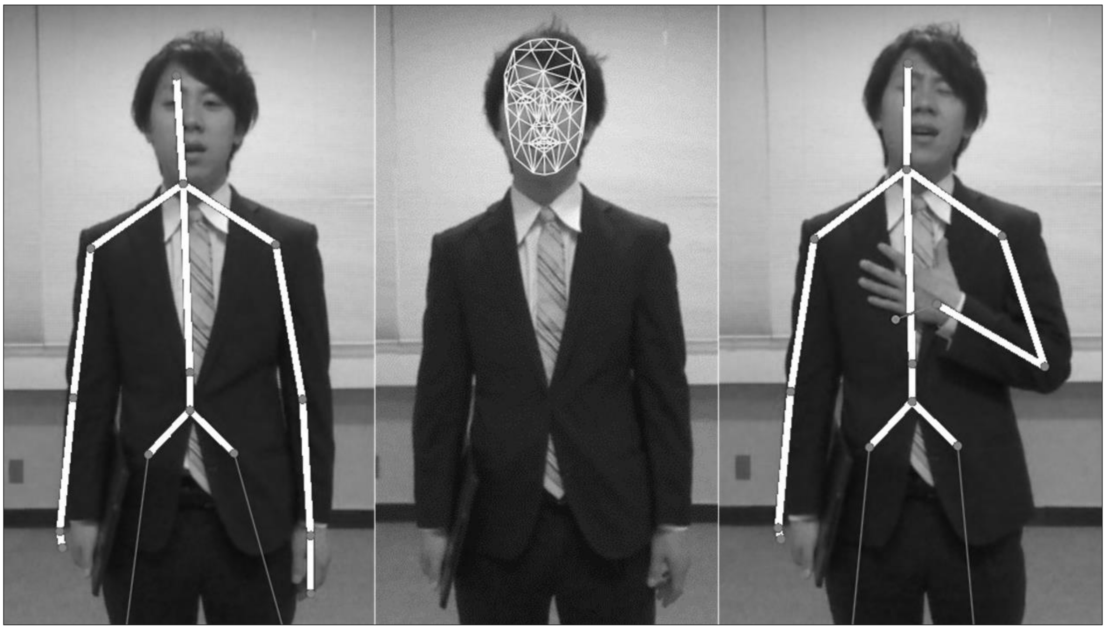


Fig. 2. 10. Output from Kinect to be analyzed

Each of the three emotion extraction engines is created and trained separately; also the programming experiments are divided into two parts. The coding to process the video and depth information is generated in Visual Studio 2012 using C# language and the audio input is processed using the Audio Analysis Tool of Matlab.

To train the 3 neural networks two different datasets are used. One dataset containing the depth and video information (created originally for this project) consists of 10 people showing the six different basic emotions in front of the Kinect sensor two times each. Samples of the records created using Kinect Studio v1.7.0 to make a total of 120 inputs are shown in Fig. 2.11. The second dataset used for the voice train is the Berlin Database of Emotion Speech [40] which consists of 10 different people showing 6 basic emotion plus neutral voice multiple times, making 535 audio voice files.

Three different feed forward neural networks are used to obtain the emotion for face, body gesture and voice. The configuration for face neural network consists of 6 nodes as inputs, 18 nodes in the hidden layer, and 6 nodes as outputs, which achieves 75.5 % accuracy. Using the Kinect sensor makes it possible to recognize and process the skeleton of the people using support vector machine, detecting the position of the human body extremities and head. The body skeleton values outputs are directly input for the gesture neural network. The gesture neural network consists of 12 inputs nodes, 22 nodes in the hidden layer, and 6 nodes in the outputs layer with an accuracy of 85.0%. The emotion from the voice is processed in the voice neural network, a configuration consisting of 6 nodes as inputs, 20 nodes in the hidden layer, and 6 nodes as the outputs, achieving 76.9% of accuracy.

Fig. 2. 11. Some of the participants of the Kinect emotion dataset created for the project

After the surface emotion is calculated, the three 6D emotion vectors are transformed to membership functions for each of the axis values (Affinity, pleasure, and arousal), and combined based on the properties of the observed agent and the situation information. This combination is realized firstly using the type of person who gives more weight to the stronger channel of emotion communication, and then takes into consideration the observed situation information. These values are combined in a fuzzy based algorithm with the surface emotion vectors to obtain the deep level emotion 3D vector. The deep emotion vector is displayed in the Affinity Pleasure – Arousal space [34].

The surface emotion outputs from the voice, face, and body gesture are different, and Fig. 2.12 shows an example of them after transformation along the Affinity Pleasure-

Arousal space axes.

The resultant emotions in the experiment still do not closely match the results described in a survey administrated to Japanese people based on similar scenarios about observed agent emotion, but it gives a good approximation of the emotion described by the observed agent. The results obtained by the system are better than those given by people that are not familiar with the observed agent. An example of the resulting deep emotion is shown in Fig. 2.13, where all the surface emotions and customized knowledge are combined. Fig. 2.14, 2.15, and 2.16 show the comparison of the centroid of the deep level emotion (dot line), centroid of the surface level emotion (double line), and the emotion reported from the observed agent (solid line). The emotion reported from the observed agent is closer to the deep emotion centroid than to the surface emotions.

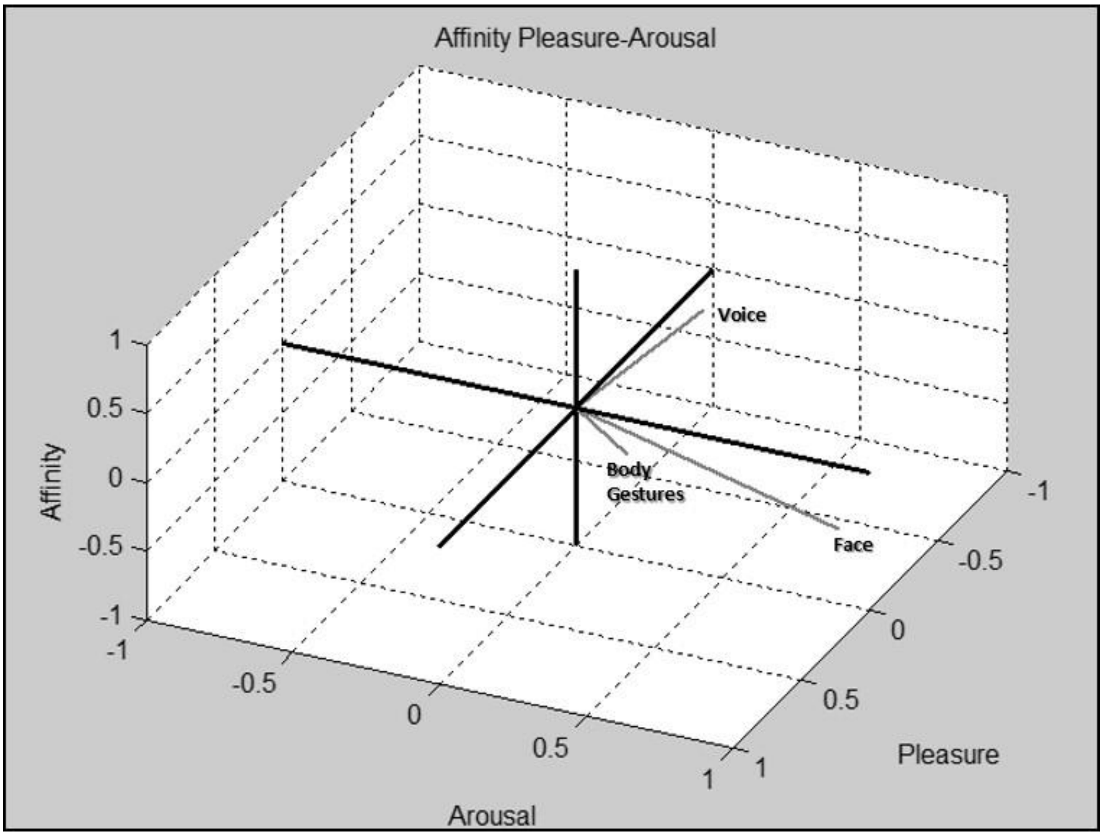


Fig. 2. 12. Affinity Pleasure - Arousal space showing the surface outputs

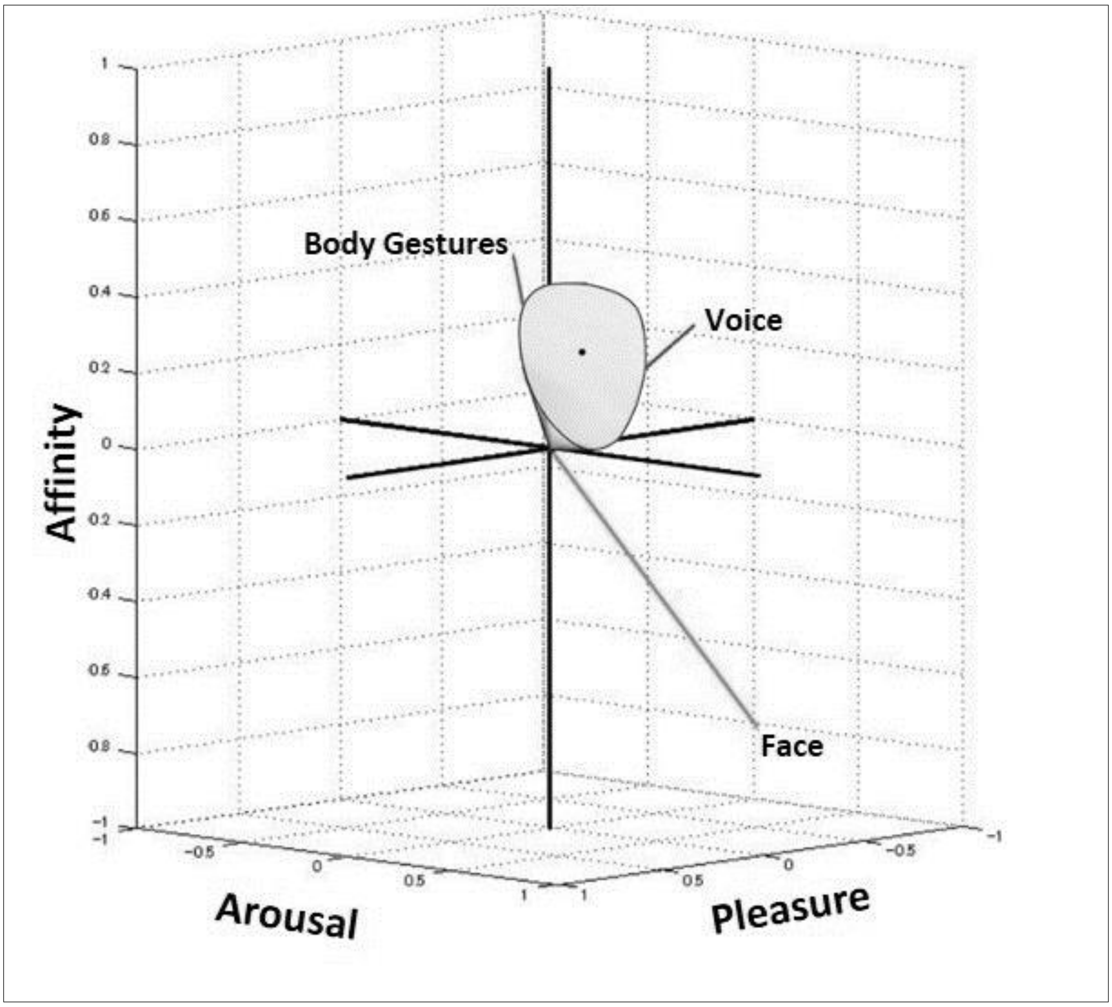


Fig. 2. 13. Affinity Pleasure-Arousal space showing the deep level emotion centroid and the standard deviation cone

The mean square errors are calculated for the three different axis to compare the deep level emotion with surface level emotion, i.e., (0.0805, 0.086, 0.0396) for the deep level emotion, and (0.111, 0.257, 0.056) for the surface emotion. The error for the deep level emotion is smaller than the surface level emotion.

The experimental environment consists of a computer with 32-bit processor Intel(R) Core (TM) i7-2600 CPU 3.40GHz, 4 GB RAM and Kinect for Windows model 1517. For the software requirements, computing device equipped with Microsoft Operating system Windows 7 Enterprise, MATLAB environment where Neural Network Toolbox and Audio Analysis Library are installed, and Microsoft Visual Studio Express 2012 for

Windows Desktop Version, 11.0.60315.01 Update 2 for coding for the Kinect sensor were used.

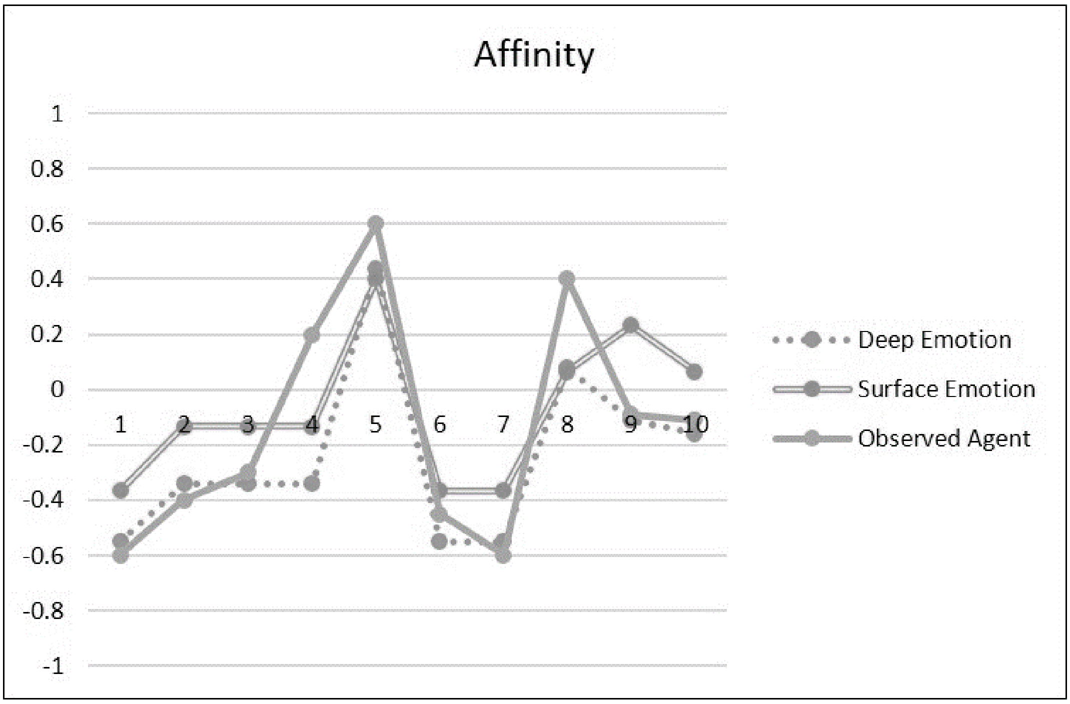


Fig. 2. 14. Comparison between the deep emotion, surface emotion, and the observed reported emotion in affinity axis

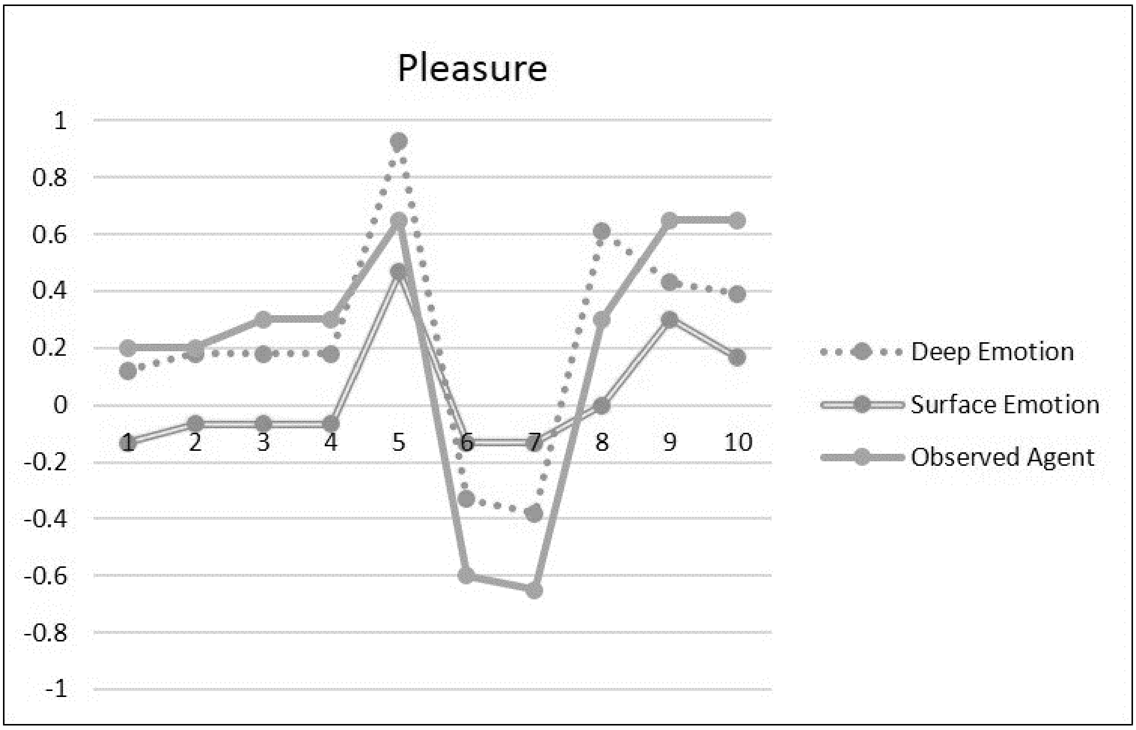


Fig. 2. 15. Comparison between the deep emotion, surface emotion, and the observed reported emotion in pleasure axis

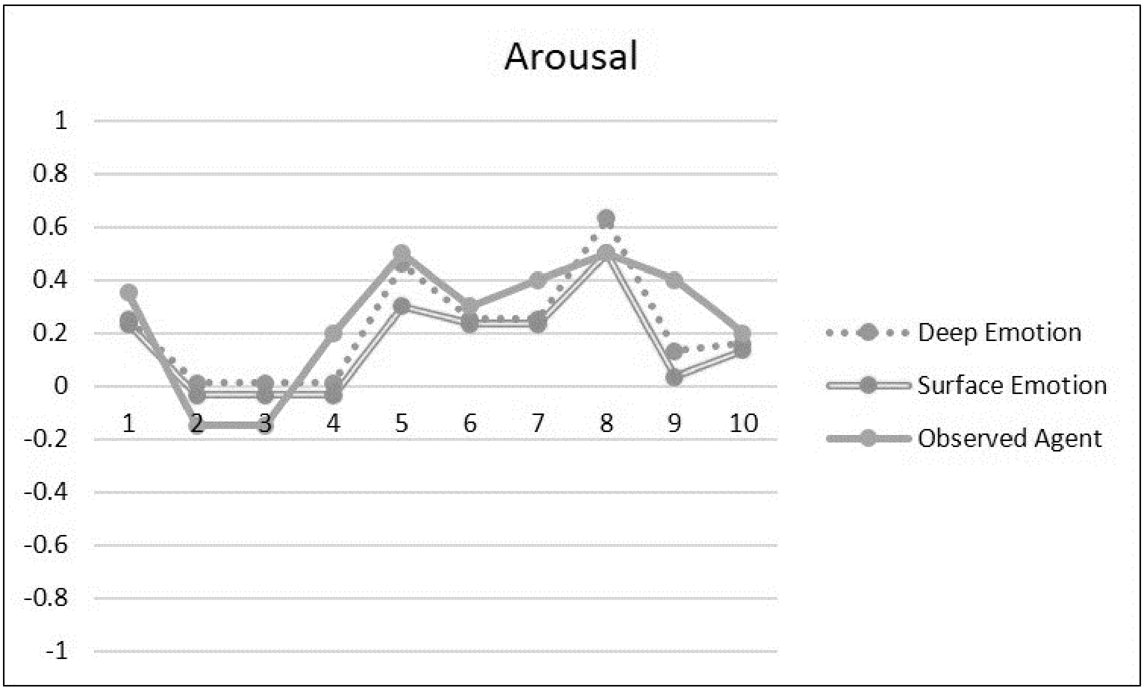


Fig. 2. 16. Comparison between the deep emotion, surface emotion, and the observed reported emotion in arousal axis

2.5. Chapter Summary

A deep level emotion understanding method is proposed for human-robot communication to handle the complex and individual way that every human express their emotions. The inputs are captured by a Kinect sensor and processed in three neural networks to obtain the surface level emotions. Then the surface emotions and customized knowledge are combined in 'fuzzy deep emotion' algorithm to obtain the deep level emotion.

In the experiment, the customized knowledge is introduced as the feature to achieve a deep level emotion understanding for realizing smooth communication between an observed agent (human employee) and an observer agent (robot secretary). In the demonstration scenario there exist two interactions, i.e., in the first interaction, using the customized knowledge about the employee and the employee's situation information, the secretary (robot) makes a meeting room reservation for the human employee based on an understanding the employee's emotions in a normal state. Subsequently, in the second interaction, at the request of the employee, the secretary (robot) changes the earlier meeting room reservation schedule by understanding the employee's mistake and giving sympathized thoughtfulness to the employee. The secretary (robot) understands the change in emotional state of the employee from the first to the second interaction by applying the knowledge about the employee to the surface level observed information from the Kinect sensor device. This confirms that the proposed method can assist to depicting the deep emotion from the surface emotion combination with the customized knowledge. The square mean errors (0.0805, 0.086, 0.0396) show closer relation between the real reported emotion and the deep level emotion.

The proposed method is being planned to be applied to an emotion understanding module in the demonstration experiments of authors' ongoing robotics research project entitled "Multi-Agent Fuzzy Atmosfield" [41], where the script entitled "a routine of beloved employee" is performed by 5 humans (employee, manager, colleague A, bar new guest, and employee’s wife) and 5 robots (secretary, colleague B, bar lady, PARO, kid) is being pursued. It is important to mention the perceived necessity of systems that can learn from normal interactions with humans, and improve their reaction, based on that acquired knowledge to achieve the maximum adaption, and personalization.

In future direction, the need 1) to add more features to the customized knowledge and 2) to improve an identification system to implement the right customized knowledge to the right person requires attention. Similarly, utilizing pattern recognition to create a new input for the system to handle the quick and fast gestures during the conversation could improve emotion understanding in human-computer communication.

Chapter 3

Visualization Method of Emotion Information and its Application to Long Distance Interaction

3.1. Studies on Emotion Visualization and Problems in Long Distance Interaction

In the long distance interaction, telecommuting and distance learning are the main tasks, but the lack of automatic communication between the users and administrator of the system have make the system less appealing for the users. The data visualization is a central issue to confirm achievements not only for expert system administrators but also for non-expert users [30]. A simple and effective emotion visualization method from a view point of affective computing is necessary to be developed for the sake of realizing smooth interaction. A few approaches for emotion visualization have been studied [24][42], where the visualization of the emotion plays an important role in the design and understanding of the system, but those methods are not easy to understand emotions intuitively and/or completely.

A visualization method is proposed to represent users' emotions in long distance interaction and to improve the interaction between users and administrators, where the emotion is characterized by a 3D vector in affinity pleasure-arousal space $[-1, 1]^3$ [34]. In accordance with any non-invasive emotion recognition method, like those based in voice [43], face [28], body gestures [29], or combination of them [22], the visualization is carried out by accepting a newly proposed shape-brightness-size (SBS) model to understand users' emotions easily and intuitively. The shape changes from an X-shape to square and finally to a circle to represent -1, 0, and 1 on pleasure-displeasure axis of affinity pleasure-arousal space, respectively. The brightness stands for the value on the arousal-sleep axis starting from the black (= -1) passing by all the gray's gradients until the white (= 1). The size indicates the position on the affinity- no-affinity axis, where the smallest size means no-affinity and the full size implies complete affinity.

The proposed visualization method provides SBS figure as the information about each user's emotions successively on the user's terminal screen. By observing the easily understandable SBS figure displayed on the screen one after another, each user makes the best use of self-emotional transitional states for his/her further self-improvement. All users' emotional personal histories are recorded in the system storage, and the system administrator puts the information to practical use for establishing business strategies such as brushing up the contents and graphical user interface, and supporting users who are faced with hard tasks.

Two experiments are accomplished; one is to validate the correct correspondence between the fundamental emotions (anger, disgust, fear, neutral, happiness, sadness, and surprise) and the SBS figures, and another is to confirm the availability of users' emotional personal histories. The first experiment consists to display 7 fundamental emotions in the shape-brightness-size figures to 5 non-experts users who are asked to select the best SBS figure that matches the given emotion. The same test is applied to 3 expert users. The second experiment involves giving 5 reading-text-tasks (one page English sentences in the field of chemistry, neuroscience, biology, mathematics, and quantum computing) to the same 8 users, and the system administrator detects the trends of users' preferences to tasks by asking users about his/her emotion after finishing each task.

Emotion visualization concept in long distance interaction system is presented in 3.2. A shape, brightness, and size (SBS) visualization method is proposed in 3.3. Two experiments on long distance interaction are done in 3.4.

3.2. Emotion Visualization Concept in Long Distance Interaction

Simplicity and full information representation of data are key issues for both expert and non-expert users in visualization method [23]. Emotion and its visualization have proved to be important in the human-computer interaction to implement and to improve services [24][42]. While the extraction and recognition of emotion have been widely studied, the visualization of the emotion and its implementation in different areas have not investigated enough.

A 3D vector emotion on the affinity pleasure-arousal space $[-1, 1]^3$ [34] is used to represent emotion. The emotion is obtained by any type of engines or methods, e.g., voice [43], face [28], body gestures [30], or their combinations [22]. The proposed visualization method is designed by introducing three different features depending on each axis in affinity pleasure-arousal space.

For the pleasure-displeasure axis [-1, 1], some meaningful shapes are accepted. Based on Japanese culture, circle represents a positive or good answer, while the X-shape represents a negative or bad answer as shown in Fig. 3.1. That makes the shape a suitable way to represent the pleasure (= 1) or displeasure (= -1). In between displeasure and pleasure, continuous deformed shape from X to circle is used as shown in the upside of Fig. 3.2.

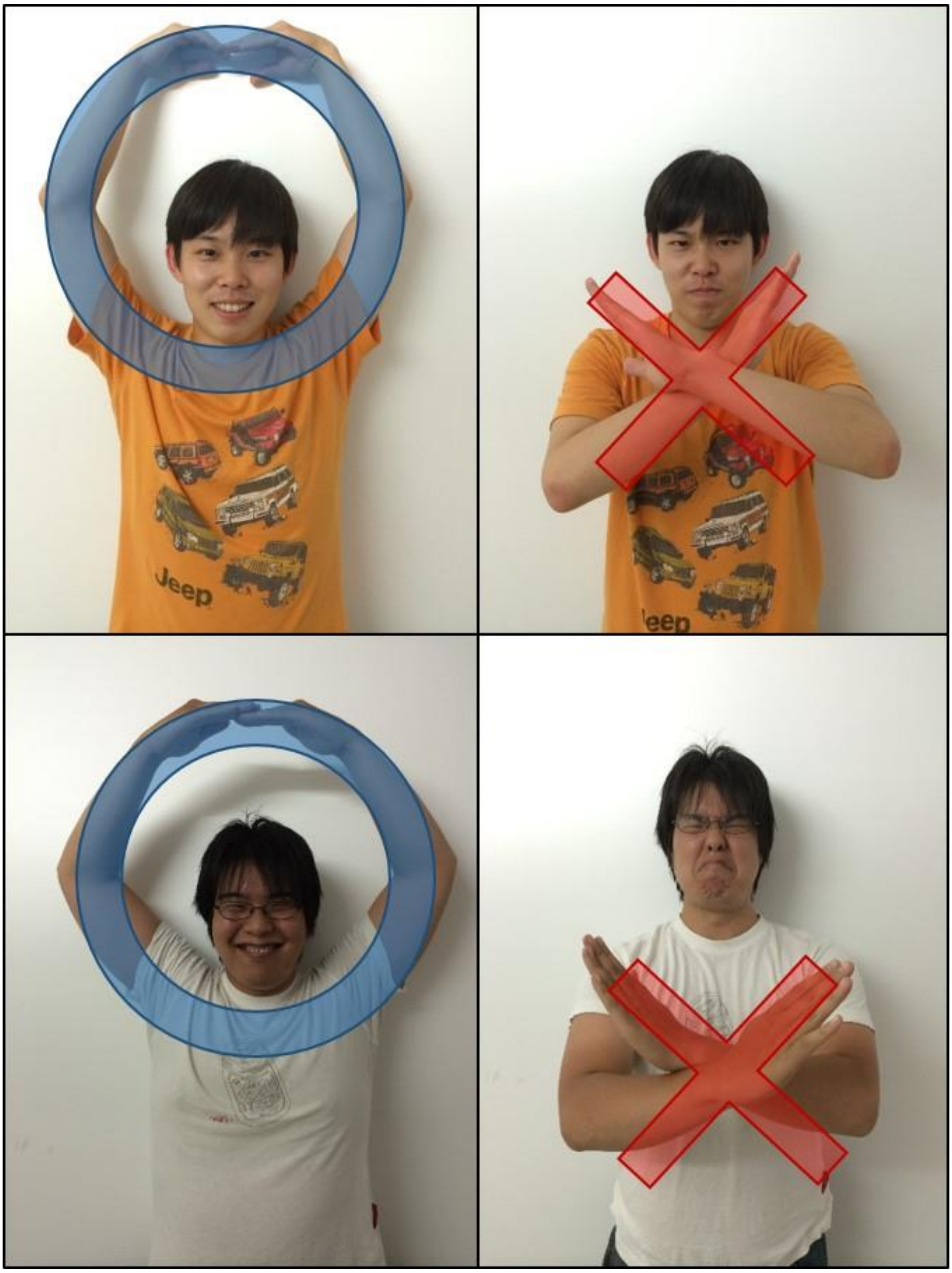


Fig. 3.1. Japanese males making a positive gesture (circle) in the left,
and a negative gesture (X-shape) in the right.

Brightness of inside the shape is adopted to represent the arousal–sleep axis [-1, 1]. White is the brightest color that denotes vivid, activeness, and arousal (= 1), while black is the darkest color that denotes gloom, passiveness, and sleep (= -1). The degree from sleep to arousal is expressed by gray level degree as shown in the middle of Fig. 3.2.

For the affinity- no-affinity axis [-1, 1], the size of the shape is used from the smallest in the case of no-affinity (= -1) to the full size in affinity (= 1) as indicated in the bottom of Fig. 3.2.

The emotion information of users is used to check and to improve the users' performance by the system administrator, when users are facing too difficult tasks in long distance interaction. The interaction concept of the proposed long distance interaction system is depicted in Fig. 3.3.

The emotion information of all users during execution of a certain task are recorded in the system storage. With the information, the administrator localizes and identifies the tasks where the users have difficulties or hard times to accomplish the tasks.

Recognition of difficult part in the task may guide the system administrator to improve contents of the task. Also the identification of users who are having difficulties leads the administrator to focus on and to support them by creating proactive relation between users and system administrator, which will lead to better user–system interaction overall. For each user, a report of his/her own emotion history displayed on the terminal screen may help to detect his/her weak or interesting potions of the task.

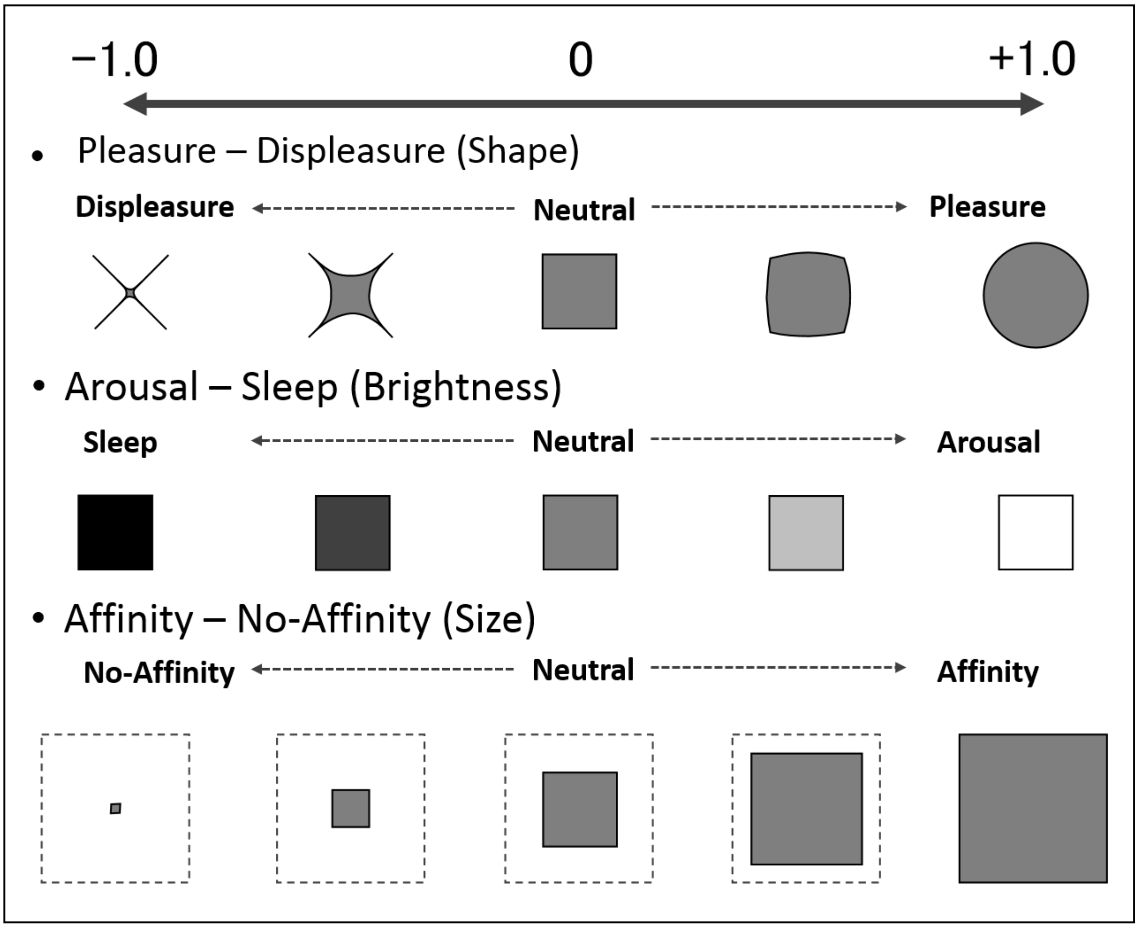


Fig. 3.2. Proposed visualization for the emotion in Affinity Pleasure-Arousal space

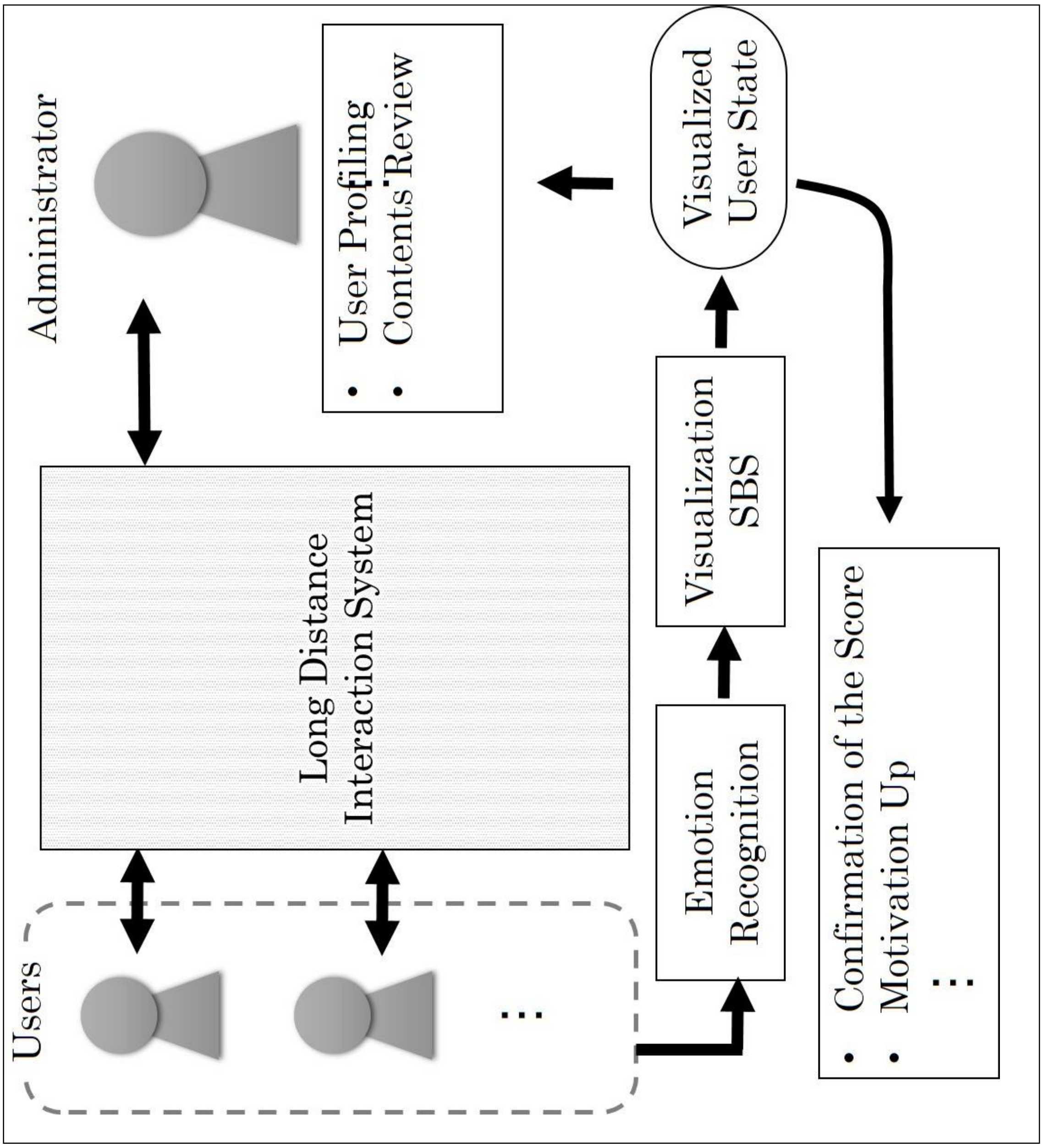


Fig. 3.3. Proposed long distance interaction concept where users and administrator have feedback from the system to improve users' performance.

3.3. Emotion Visualization Method based on Shape, Brightness, and Size (SBS) Model

To create a visualization program of user's emotion, Lamé curve or superellipse [44] rotated 45 degrees is used to generate shapes shown in the upper part of Fig. 3.2. Each value on pleasure-displeasure axis in affinity pleasure-arousal space $[-1, 1]^3$ corresponds to a particular shape of SBS model. The Lame curve [44] is defined by

$$\left|\frac{x}{a}\right|^n + \left|\frac{y}{b}\right|^n = 1, \qquad (1)$$

where *a* and *b* are the same and *n* is given by

$$n = (pleasure - displeasure\ value) + 1, \qquad (2)$$

since the pleasure-displeasure value is in [-1, 1], *n* belongs to [0, 2]. The shape of SBS model is displayed with 45 degrees rotation to create an X-shape (n=0), a square (n=1), and a circle (n=2).

Hence, the final coordinate (*x', y'*) to draw the shape is given by

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \qquad (3)$$

The brightness parameter in SBS model is used to represent the arousal-sleep value, and is given by

$$Brightness = \frac{(arousal - sleep\ value) + 1}{2} \in [0,1], \qquad (4)$$

where the shape is filled with black if the brightness is 0, with white if the brightness is 1, or any corresponding gray scale between 0 to 1.

Finally, the size of SBS model representing the affinity- no-affinity value is given by

$$\text{Size} = \frac{(affinity - no\ affinity\ value) + 1}{2} \times Maximum\ size\ , \quad (5)$$

where the maximum size depends on the implemented environment, e.g., monitor size, output image size, or application window size. It is set by 533px out of 927px for the experiments in IV.

An example of emotion visualization is presented in Fig. 3.4, where the emotion state is (-0.7, -0.5, 0.4), i.e., strongly displeasure, relatively sleepy, and fairly affinity.

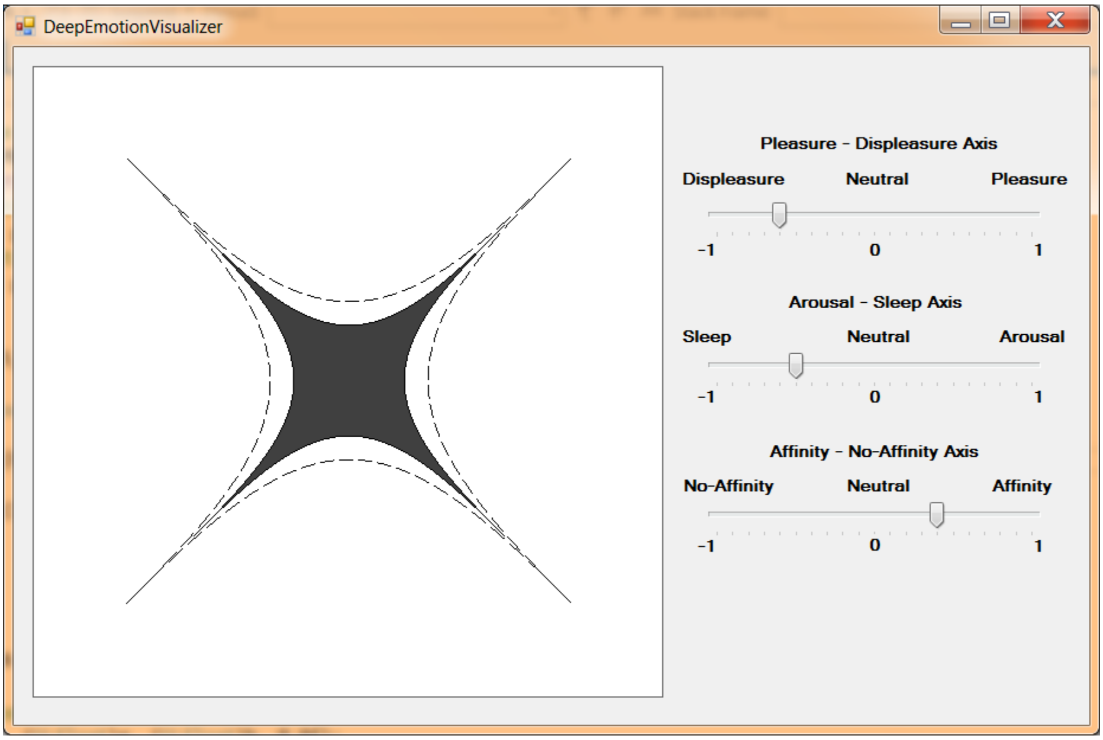


Fig. 3.4. An example of emotion visualization in the emotional state (-0.7, -0.5, 0.4)

The visualization program is coded in C++ language (66 KB) on a computer with 32-bit processor Intel(R) Core (TM) i7-2600 CPU 3.40GHz, and 4 GB RAM. The computer used Microsoft Operating system Windows 7 Enterprise, Microsoft Visual Studio

Express 2012 for Windows Desktop Version, 11.0.60315.01 Update 2 and OpenGLUT 0.6.3.

3.4. Evaluation Experiments on Emotion Visualization in Long Distance

Two experiments are performed, i.e., first to confirm the suitability of the relationship between emotion labels and SBS figures, and the second to confirm the usability of the users' emotion report to the system administrator.

3.4.1. Matching Experiment between Expression by SBS Figures and Users' Perception

To confirm understandability of SBS figures to users, a matching experiment between users' understanding of 7 SBS figures and their impression of 7 fundamental emotion labels (anger, disgust, fear, happiness, neutral, sadness, and surprise) is done by 8 users using a questionnaire form in Fig. 3.5. Users are requested to select which figure would be the best representation of the emotion label by indicating one from three SBS figures. The matching percentage is accepted as an evaluating criterion based on number of correctly matching answers between the expected SBS figures and users' emotion label perception. The total 8 users are divided into expert (3 users) and non-expert (5 users), where the expert users are accustomed to emotional research including affinity pleasure-arousal space [34], and the non-expert users are not expected to have any knowledge about emotional subjects.

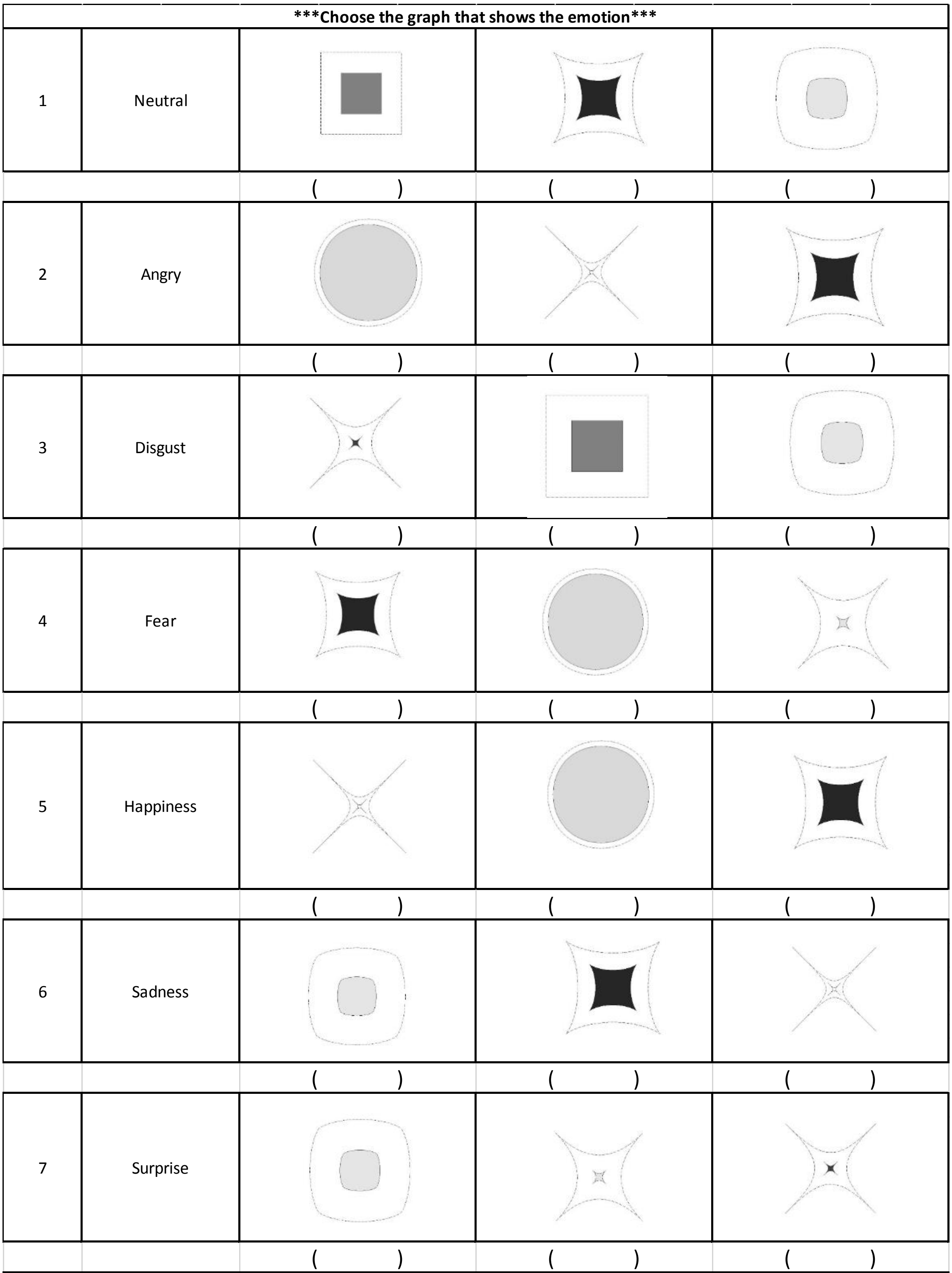


Fig. 3.5. Two Examples of emotion label and its SBS figure-candidates in the questionnaire.

Table 3.1 shows the matching percentage by expert users and non-expert users. In the case of anger, happy, and neutral SBS figures are clearly identified by all users with the score 100%. The lowest case is fear (40%) by non-experts, which indicates that the figure of fear is hard to be recognized by non-expert users because of the closeness of fear emotion to disgust/surprise emotions. Low percentage of non-expert users in sadness emotional label may depend on the users' different perception of the emotion in their cultures (users are from Japan, China, Taiwan, and Mexico). It is inferred that non-expert users will improve the identification of closer related emotions such as fear, sadness, and surprise as they become familiar with the proposed visualization method.

Table 3.1. Results of the Matching Experiment between SBS Figures and Emotional Labels

	Expert (%)	**Non-expert (%)**	**Average (%)**
Anger	100	100	100
Disgust	67	100	88
Fear	100	40	63
Happiness	100	100	100
Neutral	100	100	100
Sadness	100	60	75
Surprise	67	60	63
Average (%)	90.47	80.95	84.14

The results show high understandability of SBS figures because non-expert users matching percentage reaches 80.95% and that of expert users achieves 90.47%. It indicates that the proposed visualization method of emotion is simple and easily understandable for any users.

3.4.2. Questionnaire on Users' Feeling Emotions in Reading-text-tasks

The second experiment consists of obtaining users' feeling emotions by giving 5 reading-text-tasks. The topics of the texts are randomly selected in different fields, i.e., biology, neuroscience, mathematics, quantum computing, and chemistry, which are called as task 1-5, respectively. Each text is printed on one sheet in such a way that task 1 (biology in Fig. 3.6) sentences only, task 2 (neuroscience in Fig. 3.7) sentences with a figure, task 3 (mathematics in Fig. 3.8) sentences with an equation and a figure, task 4 (quantum computing in Fig. 3.9) sentences with a figure, and task 5 (chemistry in Fig. 3.10) also sentences with a figure. The questionnaire (Fig. 3.11) is completed to the same 8 users as in experiment 1. After finishing each task, the users are requested to check their emotion state on three [-1, 1]-scales of affinity pleasure-arousal space [34]. The users' emotions are represented by SBS figures, recorded as "Users' emotion report" in Fig. 3.11, and presented to the system administrator.

1.- Evolution

Evolution is the change in the inherited characteristics of biological populations over successive generations. Evolutionary processes give rise to diversity at every level of biological organisation, including species, individual organisms and molecules such as DNA and proteins.

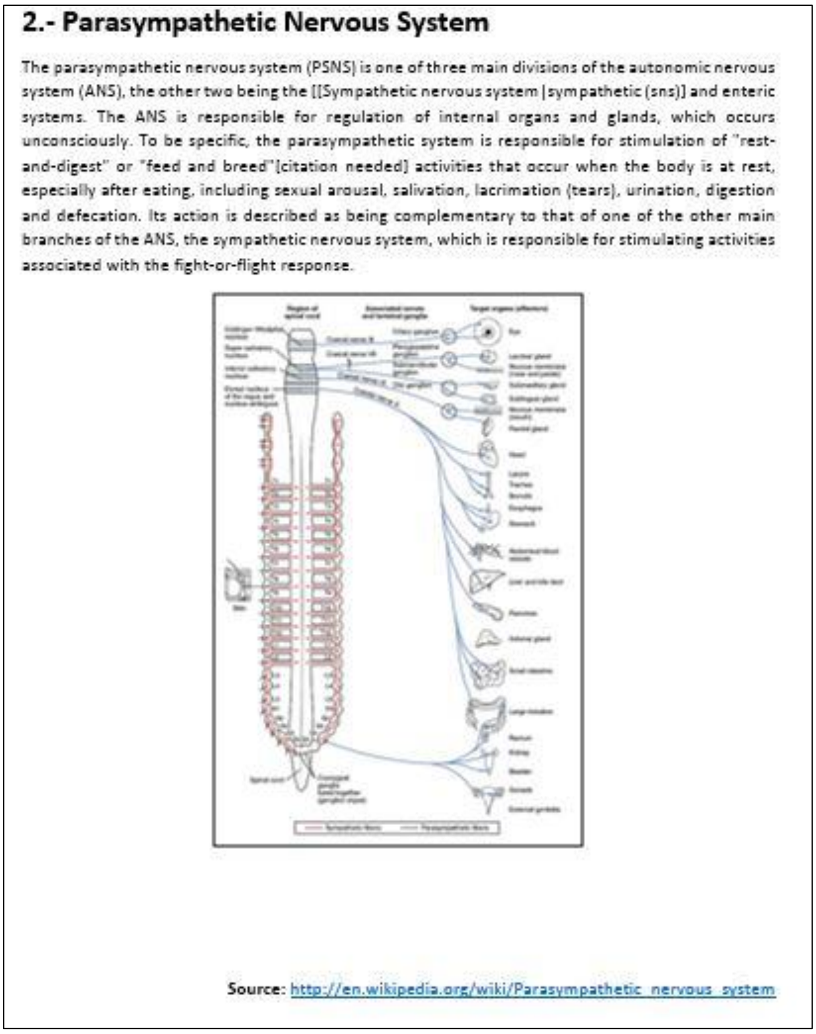
All life on Earth is descended from a last universal ancestor that lived approximately 3.8 billion years ago. Repeated speciation and the divergence of life can be inferred from shared sets of biochemical and morphological traits, or by shared DNA sequences. These homologous traits and sequences are more similar among species that share a more recent common ancestor, and can be used to reconstruct evolutionary histories, using both existing species and the fossil record. Existing patterns of biodiversity have been shaped both by speciation and by extinction.

Charles Darwin was the first to formulate a scientific argument for the theory of evolution by means of natural selection. Evolution by natural selection is a process inferred from three facts about populations: 1) more offspring are produced than can possibly survive, 2) traits vary among individuals, leading to different rates of survival and reproduction, and 3) trait differences are heritable. Thus, when members of a population die they are replaced by the progeny of parents better adapted to survive and reproduce in the environment in which natural selection takes place. This process creates and preserves traits that are seemingly fitted for the functional roles they perform. Natural selection is the only known cause of adaptation, but not the only known cause of evolution. Other, nonadaptive causes of evolution include mutation and genetic drift.

In the early 20th century, genetics was integrated with Darwin's theory of evolution by natural selection through the discipline of population genetics. The importance of natural selection as a cause of evolution was accepted into other branches of biology. Moreover, previously held notions about evolution, such as orthogenesis and "progress" became obsolete. Scientists continue to study various aspects of evolution by forming and testing hypotheses, constructing scientific theories, using observational data, and performing experiments in both the field and the laboratory. Biologists agree that descent with modification is one of the most reliably established facts in science. Discoveries in evolutionary biology have made a significant impact not just within the traditional branches of biology, but also in other academic disciplines (e.g., anthropology and psychology) and on society at large.

Source: http://en.wikipedia.org/wiki/Evolution

Fig. 3.6. Reading task number 1 in biology



2.- Parasympathetic Nervous System

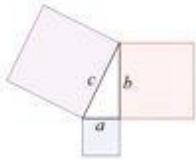
The parasympathetic nervous system (PSNS) is one of three main divisions of the autonomic nervous system (ANS), the other two being the [[Sympathetic nervous system |sympathetic (sns)]] and enteric systems. The ANS is responsible for regulation of internal organs and glands, which occurs unconsciously. To be specific, the parasympathetic system is responsible for stimulation of "rest-and-digest" or "feed and breed"[citation needed] activities that occur when the body is at rest, especially after eating, including sexual arousal, salivation, lacrimation (tears), urination, digestion and defecation. Its action is described as being complementary to that of one of the other main branches of the ANS, the sympathetic nervous system, which is responsible for stimulating activities associated with the fight-or-flight response.

Source: http://en.wikipedia.org/wiki/Parasympathetic_nervous_system

Fig. 3.7. Reading task number 2 in anatomy

3.- Pythagorean Theorem

In mathematics, the Pythagorean Theorem or Pythagoras's theorem—is a relation in Euclidean geometry among the three sides of a right triangle. It states that the square of the hypotenuse (the side opposite the right angle) is equal to the sum of the squares of the other two sides. The theorem can be written as an equation relating the lengths of the sides a, b and c, often called the Pythagorean equation:



$a^2 + b^2 = c^2$,

where c represents the length of the hypotenuse, and a and b represent the lengths of the other two sides.

The Pythagorean theorem is named after the Greek mathematician Pythagoras (ca. 570 BC—ca. 495 BC), who by tradition is credited with its proof,[2][3] although it is often argued that knowledge of the theorem predates him. There is evidence that Babylonian mathematicians understood the formula, although there is little surviving evidence that they used it in a mathematical framework.[4][5] Mesopotamian, Indian and Chinese mathematicians have all been known for independently discovering the result, some even providing proofs of special cases.

The theorem has numerous proofs, possibly the most of any mathematical theorem. These are very diverse, including both geometric proofs and algebraic proofs, with some dating back thousands of years. The theorem can be generalized in various ways, including higher-dimensional spaces, to spaces that are not Euclidean, to objects that are not right triangles, and indeed, to objects that are not triangles at all, but n-dimensional solids. The Pythagorean Theorem has attracted interest outside mathematics as a symbol of mathematical abstruseness, mystique, or intellectual power; popular references in literature, plays, musicals, songs, stamps and cartoons abound.

Source: http://en.wikipedia.org/wiki/Pythagorean_theorem

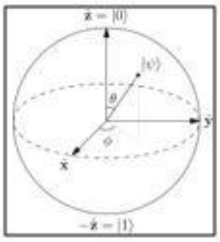
Fig. 3.8. Reading task number 3 in mathematics

4.- Quantum Computer

A quantum computer is a computation device that makes direct use of quantum-mechanical phenomena, such as superposition and entanglement, to perform operations on data.[1] Quantum computers are different from digital computers based on transistors. Whereas digital computers require data to be encoded into binary digits (bits), each of which is always in one of two definite states (0 or 1), quantum computation uses qubits (quantum bits), which can be in superpositions of states. A theoretical model is the quantum Turing machine, also known as the universal quantum computer. Quantum computers share theoretical similarities with non-deterministic and probabilistic computers; one example is the ability to be in more than one state simultaneously. The field of quantum computing was first introduced by Yuri Manin in 1980[2] and Richard Feynman in 1982.[3][4] A quantum computer with spins as quantum bits was also formulated for use as a quantum space–time in 1969.[5]

As of 2014 quantum computing is still in its infancy but experiments have been carried out in which quantum computational operations were executed on a very small number of qubits.[6] Both practical and theoretical research continues, and many national governments and military funding agencies support quantum computing research to develop quantum computers for both civilian and national security purposes, such as cryptanalysis.[7]

Large-scale quantum computers will be able to solve certain problems much quicker than any classical computer using the best currently known algorithms, like integer factorization using Shor's algorithm or the simulation of quantum many-body systems. There exist quantum algorithms, such as Simon's algorithm, which run faster than any possible probabilistic classical algorithm.[8] Given sufficient computational resources, however, a classical computer could be made to simulate any quantum algorithm; quantum computation does not violate the Church–Turing thesis.[9]



Source: http://en.wikipedia.org/wiki/Quantum_computer

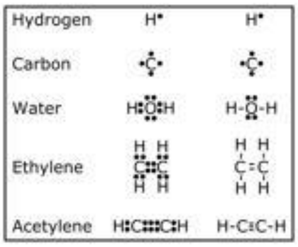
Fig. 3.9. Reading task number 4 in Quantum Information

5.- Chemical Bond

A chemical bond is an attraction between atoms that allows the formation of chemical substances that contain two or more atoms. The bond is caused by the electrostatic force of attraction between opposite charges, either between electrons and nuclei, or as the result of a dipole attraction. The strength of chemical bonds varies considerably; there are "strong bonds" such as covalent or ionic bonds and "weak bonds" such as dipole–dipole interactions, the London dispersion force and hydrogen bonding.

Since opposite charges attract via a simple electromagnetic force, the negatively charged electrons that are orbiting the nucleus and the positively charged protons in the nucleus attract each other. An electron positioned between two nuclei will be attracted to both of them, and the nuclei will be attracted toward electrons in this position. This attraction constitutes the chemical bond. Due to the matter wave nature of electrons and their smaller mass, they must occupy a much larger amount of volume compared with the nuclei, and this volume occupied by the electrons keeps the atomic nuclei relatively far apart, as compared with the size of the nuclei themselves. This phenomenon limits the distance between nuclei and atoms in a bond.

In general, strong chemical bonding is associated with the sharing or transfer of electrons between the participating atoms. The atoms in molecules, crystals, metals and diatomic gases— indeed most of the physical environment around us— are held together by chemical bonds, which dictate the structure and the bulk properties of matter.



Source: http://en.wikipedia.org/wiki/Chemical_bond

Fig. 3.10. Reading task number 5 in Chemistry

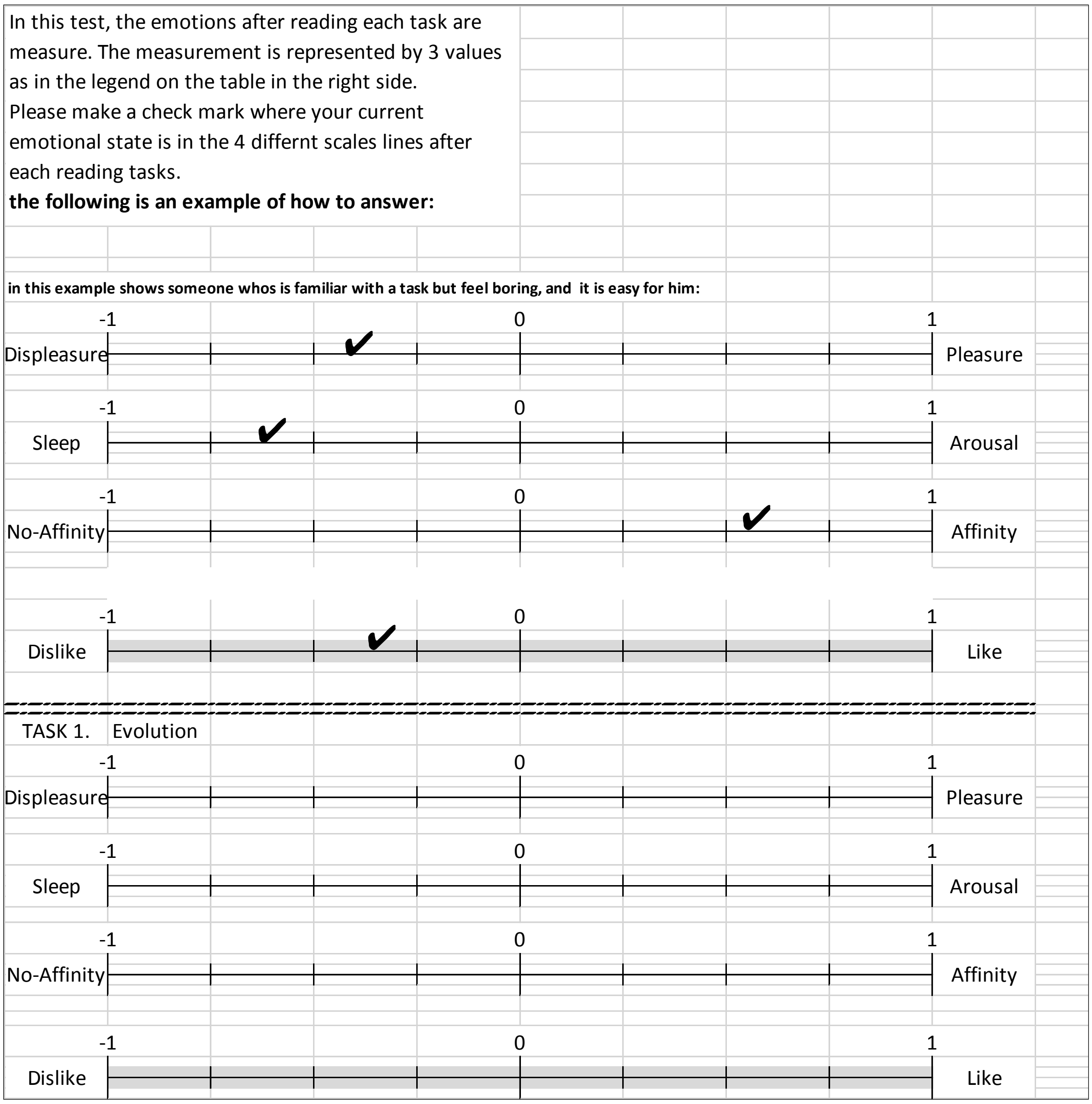


Fig. 3.11. Example of the answer sheet for one of the tasks

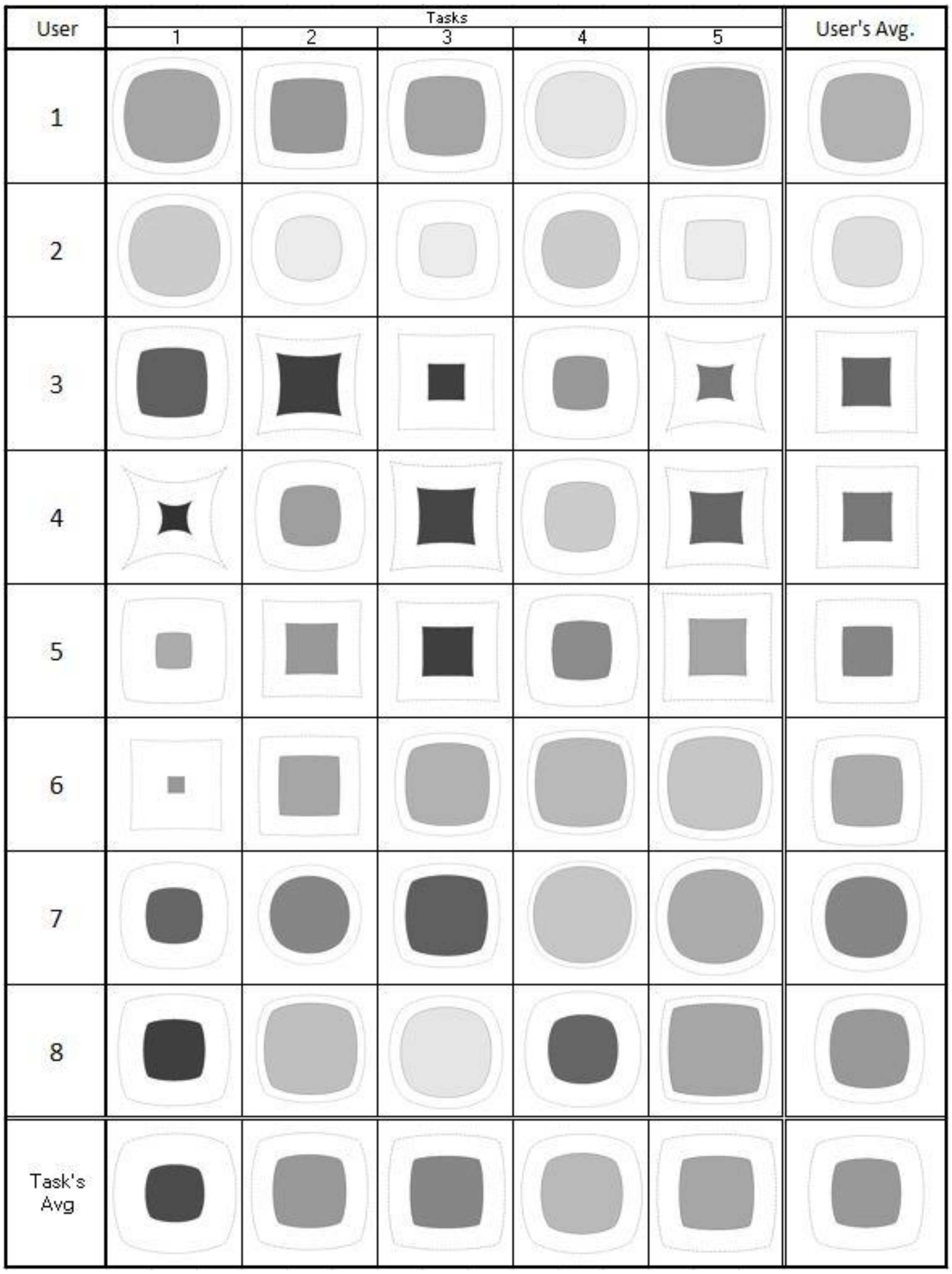


Fig. 3.12. Users' emotion report to the system administrator.

Fig. 3.11 provides the system administrator some trends of tasks and users, e.g., the average of task 1 in comparison with the others shows sleepy (dark) and low affinity (small size), and the closer averages of user 3, 4, and 5 show neutral state (square, gray, and middle size).

To confirm the results observed in the "Users' emotion report", another questionnaire is done to the same 8 users, where the users indicate by [-1, 1] score checking how easy and how prefer each task is, and the results of easiness and preference are shown in Table 3.2 and 3.3, respectively. A strong relationship is confirmed in Table 3.2 and 3.3 between SBS visualization method and users' responses to each task. Data of user 8 in Fig. 3.11 and Table 3.2 show that darker brightness is related with task difficulty, which is reflected by the negative values in Table 3.2.

Users 1, 2, and 7 have interests in all tasks, which is revealed by the closeness to circle shape, lighter brightness, and bigger figure size. It is also corroborated with the big positive values in Table 3.2 and 3.3. Task 1 has closer to X-shape, darker, and smaller size in comparison to other tasks in Fig. 3.11, which is also cross checked with smaller average values of the task 1 in Table 3.2 and 3.3.

Table 2.2. Questionnaire answers about if tasks are easy for users

Users	Task 1	Task 2	Task 3	Task 4	Task 5	Users' Avg.
1	**0.15**	0.75	**0.06**	0.95	0.91	0.56
2	0.86	**0.11**	**0.11**	0.36	0.41	0.37
3	0.48	0.97	0.75	**0.21**	**-0.27**	0.43
4	**-0.90**	0.25	-0.14	0.43	**-0.43**	-0.16
5	**0.36**	0.92	**0.00**	0.46	0.50	0.45
6	**-0.39**	**0.39**	0.57	0.80	0.83	0.44
7	**0.29**	0.95	**0.64**	0.71	0.64	0.64
8	**-0.64**	0.41	0.75	**-0.74**	0.36	0.03
Avg.	0.03	0.59	0.34	0.40	0.37	0.35

Table 3.3. Questionnaire answers if users prefer the task or not

Users	**Task 1**	**Task 2**	**Task 3**	**Task 4**	**Task 5**	**Users' Avg.**
1	0.50	**0.29**	0.40	0.69	**0.17**	0.41
2	0.61	0.37	**0.37**	0.57	**0.36**	0.46
3	-0.01	-0.26	**-0.28**	0.23	**-0.26**	-0.11
4	**-0.36**	0.15	0.15	0.33	**-0.13**	0.03
5	**0.13**	0.25	**0.00**	0.19	0.62	0.24
6	**-0.07**	**0.14**	0.59	0.61	0.69	0.39
7	0.56	0.43	0.70	0.79	0.80	0.66
8	**-0.01**	0.23	0.73	0.50	**-0.50**	0.19
Avg.	0.17	0.20	0.33	0.49	0.22	0.28

3.4.3. Confirmation of advantages of SBS figures

It is confirmed from the results of experiment 1 that the proposed visualization method of emotion is simple and easily understandable for all users because even non-expert users achieve relatively high matching percentage 80.95% in average. It turns out to be more easily acceptable as the user becomes use to the proposed visualization method because the matching percentage of expert users exceeds about 10% to that of non-expert users. Discrepancy exists in the data shown in Table 3.1, but it is mainly caused by lower matching percentage of non-expert users hence the problem may be solved as they become familiar with the proposed visualization method.

The system administrator finds a hint of what tasks need to be improved by checking the “Users’ emotion report” in experiment 2 and selecting X-shape closer, darker, and smaller sized SBS model figures in the report. “Users’ emotion report” suggests various strategic hints to the system administrator, e.g., 1) the result for task 1 (biology) indicates that text is not easy to understand for most users, which provides suggestion of inserting appropriate illustrations, 2) users 3, 4, and 5 report negative responses to all tasks, which makes the system administrator to interview them for asking the reasons of negative responses (users 3 and 5 say the contents are boring, user 4 confesses too difficult).

3.4.4. Ongoing Language Learning Distance Education

The proposed visualization method of emotion is planning to be applied to several practical long distance interaction problems. Among them, a new distance education system is investigated to realize global scale language education through internet. Distance education has various advantages compared with face-to-face education such as everywhere, every time, every subject, and everyone, but the biggest problem is lack of mutual communications among learners. A concept of virtual classroom is introduced in the proposed distance education system, where classroom atmosphere and learners' emotions make important role to solve the problem. Learners' emotion is related to facial expression, gesture/posture, and voice information especially in language learning. Multimodal interface device like Kinect® is supposed to be attached to each learner's terminal device and captures learner's basic information to identify his/her emotion. The proposed visualization method will be applied to provide learner's emotion profile. English-German language learning is focused, where the contents of the course will be based on the learner's weak points and preferences without any manual setup by the learners. The contents and the lectures inside of the course could be reviewed automatically by the learners' emotions during his/her learning time. In the near future, the proposal will be implemented on advanced smart phones for worldwide use.

3.5. Chapter Summary

Two experiments are practiced on long distance interaction, i.e., matching test between 7 emotional labels and 7 SBS figures and questionnaire on 8 users' feeling emotions to 5 reading-text-tasks. The results from the first experiment between 7 fundamental emotions and 7 SBS emotion model figures obtain matching percentage 80.95% by 5 non-expert users and 90.47% by 3 expert users. The second experiment for the same 5 non-expert and 3 expert users results in relatively big individual difference of feeling emotions, which is irrelevant to expert or non-expert.

It is confirmed from the results of the first experiment that the proposed visualization method of emotion is simple and easily understandable for users because even non-expert users achieve 80.95%, and that it turns out to be more easily acceptable as the user becomes use to the visualization method because the matching percentage of expert users exceeds about 10% to that of non-expert users. The system administrator finds a hint of what tasks need to be improved by checking the “Users’ emotion report” in experiment 2 and selecting X-shape closer, darker, and smaller sized SBS model figures in the report.

Although two experiments carried out are elementary trials in long distance interaction, more practical implementation is ongoing in the field of distance education since the authors’ group concentrates on education through internet from global viewpoints. Learners’ emotion is related to facial expression, gesture/posture, and voice information especially in language learning, and such information is easily obtained by using multimodal interface like Kinect. The proposed visualization method is planning to be applied to English-German language learning application as a smartphone application, where the contents of the course will be based on the user’s weak points and preferences without any manual setup by the users. The contents and the lectures inside of the course could be reviewed automatically by the users’ emotions during his/her learning time.

Chapter 4

Deep Level Emotion Understanding

Applied to Distance Learning

4.1. Studies on Emotion Applied to Distance Learning and Concept of Artificial Inspirer

In Distance Learning is taking a very important part in the modern life, big companies and opportunity to business opened a big interest in this area, being able to reach all areas of the world, without having the limitation of space of time, it is just wonderful. But distance learning have a problem is the adaptation to the different users around the world, and how to support their experience actively, while getting static of the content. A study shown that good performance of the online learning is fast, but in the long term, it loses against the regular paper studies [45], mainly because the little adaptability that systems do for the users.

The relation between the learning style of each learner and the online learning application is critical to achieve good learners' performance and to keep the learners motivated to continue learning [46] [47]. Different methods to create adaptive systems are being proposed [48] [49], but there is need for more active and automatic adaptation of the e-learning system. The personality and emotions affects the learning performance [50].

The usage of deep level emotion understanding is proposed as the mechanist to achieve and control the e-learning system to adapt to the user necessities without the direct intervention of the system administrator. A customized knowledge combined with emotions to create the deep level emotion understanding [51], proved to be a good approach to realize the learner deep emotion level state and adapt the system.

The proposal consists of an audio, video, depth capture device like Kinect ®, the inputs are processed to obtain the deep emotions based on the learner customized knowledge, and then the deep personality decision module will select and command the adaptive e-learning system to provide the appropriated content to the learner.

The deep level emotion understanding in language distance learning concept is presented in 4.2. The artificial inspirer in the experiment on English-German learning is proposed in 4.3.

4.2. Deep Level Emotion Understanding in Language Distance Learning

Online learning is fast and effective in short term, but in a long term normal studies achieve better results [45], normally because the lack of inspiration, and the low adaptability of the learning system to the learner needs.

A motivated learner will give better results, adaptive e-learning system had being proposed [46] [47], but an identification of the personality [50] of the learner will improve their performance. Some of the methods like [48] [49], but there is need for more active and automatic adaptation of the e-learning system.

The deep level emotion understanding is proposed as the key to control an e-learning system to adapt to the user necessities without the direct intervention of the system administrator. A customized knowledge combined with emotions to create the deep level emotion understanding [41], the system fusion the deep level emotion understanding theory with an adaptive e-learning system to obtain a “Deep Level Adaptive E-Learning System,” that provided better and more accurate content to the learner, shown in Fig. 4.1.

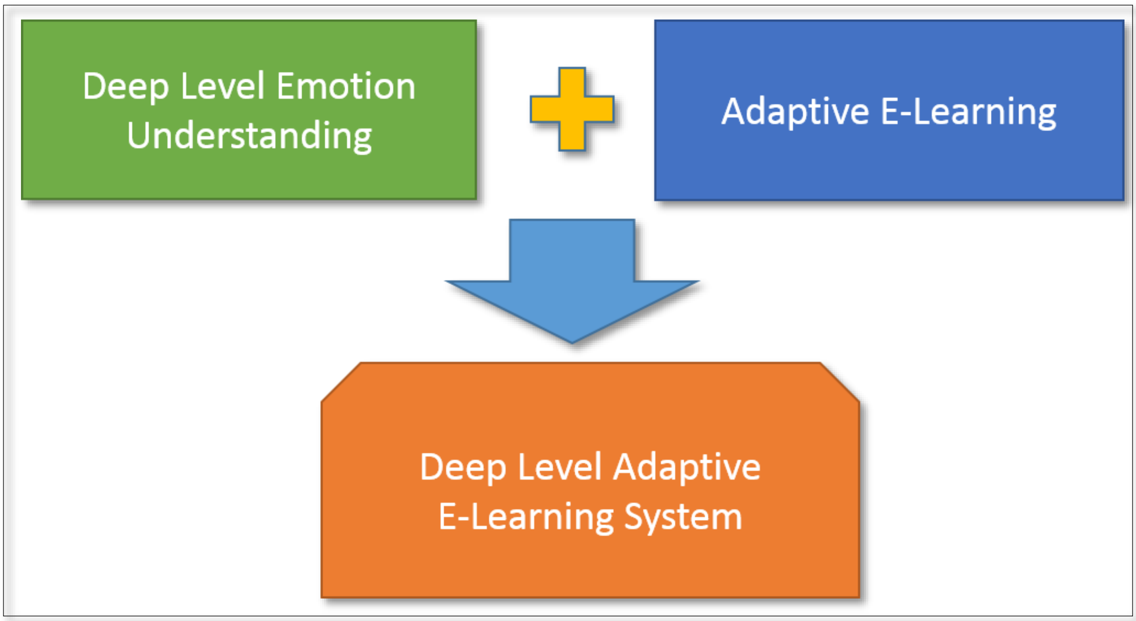


Fig. 4.1.Combination of the Deep Level Emotion Understanding to control the Adaptive E-Learning system to create a new system "Deep Level Adaptive E-Learning System".

4.3. Artificial Inspirer in Experiment on English-German Language Learning

The deep level adaptive system diagram is shown in Fig. 4.2. Where the learner facial, voice and depth information is captured by a Kinect® and send to the deep level understanding engine. The outputs of the Kinect® are combine with the customized knowledge about the learner to create the deep level emotion which is the input to the deep personality decision, here the deep emotion level is processed to create the appropriated command to the e-learning server data, which will send the list of content for that learner based in the deep level adaptive system.

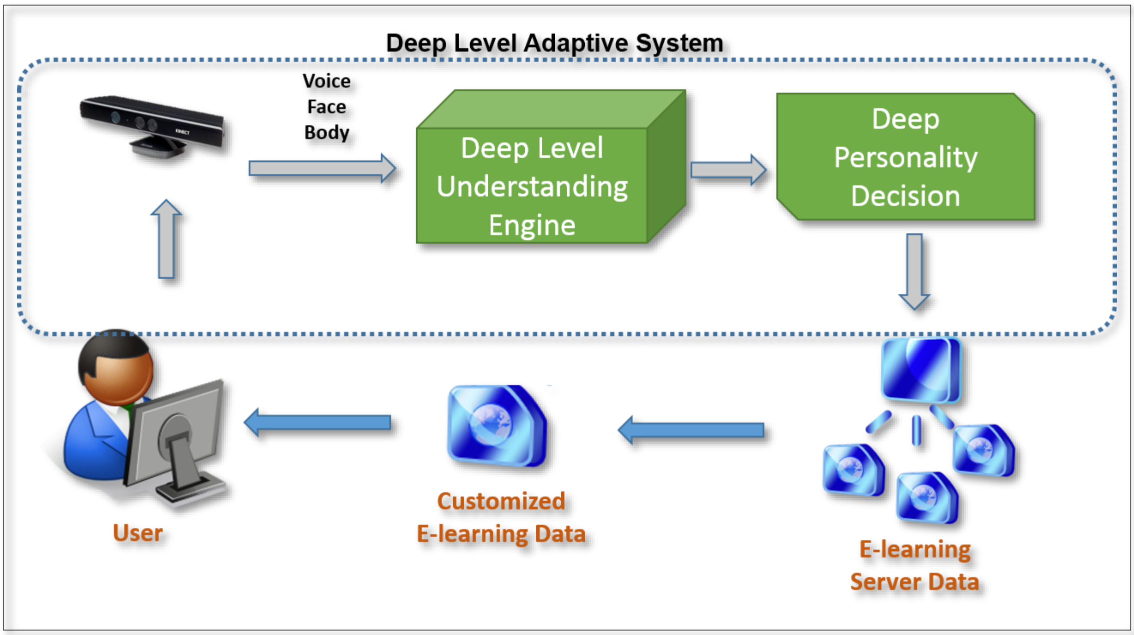


Fig. 4.2.Proposal diagram of "Deep Level Adaptive E-Learning System".

In the matching experiment, the comparison between the normal system and the proposal is done via web application in an English-German learning language, where users used the system while it is adapting to them, also a report page for the administrator localize the contents to improve.

4.5. Chapter Summary

The experiment shown that with the emotion an automatic feedback can be obtained and used to improve the interaction learner – system, system – administrator, and learner – administrator.

Chapter 5

Conclusion and Future Perspective

5.1 Conclusions

Deep level emotion understanding method is proposed for human-robot communication to handle the complex and individual way that every human express their emotions, and in some way emulate the day that human understand the emotions of other human. In the proposed method the inputs are captured by a Kinect sensor and processed in three neural networks to obtain the surface level emotions. Then the surface emotions and customized knowledge are combined in 'fuzzy deep emotion' algorithm to obtain the deep level emotion. In the experiment, the customized knowledge was introduced as the feature to achieve a deep level emotion understanding for realizing smooth communication between an observed agent (human employee) and an observer agent (robot secretary). The deep emotion is the representation of the real emotion of agent, to prove it some demonstration scenarios were proposed, where two interactions were made, i.e., in the first interaction, using the customized knowledge about the employee and the employee's situation information, the secretary (robot) makes a meeting room reservation for the human employee based on an understanding the employee's emotions in a normal state.

In the emotion visualization experiments were performed on long distance interaction, i.e., matching test between 7 emotional labels and 7 SBS figures and questionnaire on 8 users' feeling emotions to 5 reading-text-tasks. The results from the experiments show that the SBS emotion model figures obtain a great matching percentage 80.95% by 5 non-expert users and 90.47% by 3 expert users, making it possible the utilization. The proposed method shown very good assimilation from the users, and also it could be adjusted to show different range of features, like the usage of colors or the use of 3D volumes to increase the representation properties of the method.

The visualization method results confirmed that the proposed visualization method of emotion is simple and easily understandable for users because even non-expert users achieve 80.95%, and that it turns out to be more easily acceptable as the user becomes use to the visualization method because the matching percentage of expert users exceeds about 10% to that of non-expert users.

And finally the implementation of the deep emotion level understanding to distance education shows that the correct emotion using and tracking, combined with an adaptive e-learning system can improve the learning quality of the users, get the best contain to the learner way of studying and also create an automatic channel to review the contain inside of the e-leaning system.

5.2 Future Perspective

In the near future perspective works is to add more features to the customized knowledge to have more accurate deep emotion recognition engine and to implement an identification system to apply the right customized knowledge to the right person, and also exploring other input devices to improve speed and accuracy of the obtained emotions.

In future perspective works is the implementation of more complex adaptive e-learning system that could manage more details about the learners to make more fully usage of the deep emotion level understanding, similarly, utilizing pattern recognition to create a new input for the system to handle the quick and fast gestures during the conversation/interaction.

Bibliography

[1] Panteleimon Ekkekakis, “Affect, Mood, and Emotion,” Measurement in Sport and Exercises Psychology, Human Kinectics Press, ISBN-10: 0-7360-8681-1, pp. 321-331, 2012

[2] Oxford Online Dictionaries: http://www.oxforddictionaries.com/definition/english/emotion

[3] J. G. Taylor, Klaus Scherer, R. Cowie, “'Emotion and brain: understanding emotions and modelling their recognition',” Neural Networks, ELSEVIER, DOI: 10.1016/j.neunet.2005.04.001, pp. 313-316, 2005

[4] Russel J., Niit T., Lewicka M., “A Cross-Cultural Study of a Circumplex Model of Affect,” Journal of Personality and Social Psycology, Volume 57, Number 5, pp. 848-856, 1989

[5] Xiao-Ping Gao, John H. Xin,Tetsuya Sato, Aran Hansuebsai, Marcello Scalzo, Kanji Kajiwara, Shing-Sheng Guan, J. Valldeperas, Manuel Jose´ Lis, Monica Billger, “Analysis of Cross-Cultural Color Emotion,” Color Research & Application, Volume 32, Issue 3, online published, 2007

[6] Kirsten Boehner , Rogério Depaula , Paul Dourish , Phoebe Sengers, “ How emotion is made and measured,” Int. J. Human-Computer Studies, ELSEVIER, Volume 65, pp. 275-291, 2007

[7] Gouizi K., Bereksi Reguig Maaoui, "Emotion recognition from physiological signals," Journal of Medical Engineering and Technology, v 35, n 6-7, p 300-307, 2011

[8] Gouizi K., Reguig, F.B., Maaoui C., "Analysis physiological signals for emotion recognition," Systems, Signal Processing and their Applications (WOSSPA), 2011 7th International Workshop , pp. 147 - 150, 2011

[9] Valenza Gaetano, Citi Luca, Lanata Antonio, Scilingo Enzo Pasquale, Barbieri Riccardo," A nonlinear heartbeat dynamics model approach for personalized emotion recognition," Proceedings of the Annual International Conference of the IEEE

Engineering in Medicine and Biology Society, 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2579-2582, 2013

[10] Henriques Rui, Paiva Ana, Antunes Claudia, "Accessing emotion patterns from affective interactions using electrodermal activity," Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013, p 43-48, 2013

[11] Lee Chung K, YooSun K, "ECG-based biofeedback chair for self-emotion management at home," IEEE International Conference on Consumer Electronics, Digest of Technical Papers International Conference on Consumer Electronics, The Mobile Consumer,2008

[12] Rothkrantz Leon J.M., Horlings Robert, Datcu Dragos, "Emotion recognition using brain activity," Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing, 2008

[13] Fextha Satria, Murayama Nobuki, Igasaki Tomohiko, Hayashida Yuki, "Effects of emotional stimulation on functional connection between brain and muscle in humans," 2012 ICME International Conference on Complex Medical Engineering, CME 2012 Proceedings, pp. 90-95, 2012

[14] Yi-Lin Lin, Gang Wei, "Speech emotion recognition based on HMM and SVM," Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference, IEEE, Volume 8, pp. 4898-4901, 2005

[15] Ashish B. Ingale, D. S. Chaudhari, "Speech Emotion Recognition," International Journal of Soft Computing and Engineering (IJSCE),Volume-2, Issue-1, pp. 235-238, 2012

[16] Munaf Rashid, S. A. R. Abu-Bakar, Musa Mokji, "Human emotion recognition from videos using spatio-temporal and audio features," The Visual Computer, Springer, Volume 29, Issue 12, pp. 1269-1275, 2013

[17] Yoshihiro Miyakoshi, Shohei Kato, "Facial Emotion Detection Considering Partial

Occlusion of Face Using Bayesian Network," IEEE Symposium on Computers & Informatics,DOI:10.1109/ISCI.2011.5958891, pp. 96-101, 2011

[18] Spiros V. Ioannou, Amaryllis T. Raouzaiou, Vasilis A. Tzouvaras, Theofilos P. Mailis,Kostas C. Karpouzis, Stefanos D. Kollias, "Emotion recognition through facial expression analysis based on a neurofuzzy network," Neural Networks, ELSEVIER, Volume 18, Issue 4, pp. 423–435, 2005

[19] Catherine Soladié, Hanan Salam, Catherine Pelachaud, Nicolas Stoiber, Renaud Séguier, "A multimodal fuzzy inference system using a continuous facial expression representation for emotion detection," A multimodal fuzzy inference system using a continuous facial expression representation for emotion detection,doi>10.1145/2388676.2388782, pp. 493-500, 2012

[20] Konrad Schindler, Luc Van Gool, Beatrice de Gelder, "Recognizing emotions expressed by body pose: A biologically inspired neural model," Neural Networks, ELSEVIER, Volume 21, Issue 9, pp. 1238–1246, 2008

[21] Sidney K. D'Mello, Arthur Graesser, "Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features," User Modeling and User-Adapted Interaction, Springer, Volume 20, Issue 2, pp. 147-187, 2010

[22] Kessous L., Castellano G., Caridakis G, "Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic analysis," Journal on Multimodal User Interfaces, Volume 3, Issue 1-2, pp. 33-48 (2010)

[23] P. Gough, C. B. Wall, and T. Bednarz, "Affective and Effective Visualisation: Communicating Science to Non-Expert Users," 2014 IEEE Pacific Visualization Symposium (Yokohama, Japan), pp.335-339, DOI 10.1109/Pacific Vis.2014.39, 2014.

[24] I. Hupont, S. Baldassarri, E. Cerezo, and R. Del-Hoyo, "Advanced Human Affect Visualization," 2013 IEEE International Conference on Systems, Man, and

Cybernetics (Manchester, England), pp.2700-2705, DOI 10.1109/SMC.2013.460, 2013.

[25] Cañamero L., “Emotion understanding from the perspective of autonomous robots research,” Neural Networks, Emotion and Brain, Volume 18, Issue 4, pp. 445–455, 2005

[26] Bethel C.L., “Survery of Psychophysiology Measurements Applied to Human-Robot Interaction,” The 16th IEEE International Symposium on Robot and Human interactive Communication, pp. 732 – 737, 2007

[27] Fontaine J.R., Scherer K.R., Roesch E.B., Ellsworth P.C., “The World of Emotions is not Two-Dimensional,” Psychological Science, Volume 18, Number 12, pp. 1050-1057, 2007

[28] Ilbeygia M., Shah-Hosseinib H., “A novel fuzzy facial expression recognition system based on facial feature extraction from color face images,” Engineering Applications of Artificial Intelligence, Volume 25, Issue 1, pp.30–146, 2012

[29] Ashish B.I., Chaudhari. D.S., “Speech Emotion Recognition,” International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-1, pp. 235-238, 2012

[30] Zhao Y., Wang X., Goubran M., Whalen T., Petriu E. M., “Human emotion and cognition recognition from body language of the head using soft computing techniques,” Journal of Ambient Intelligence and Humanized Computing, Volume 4, Issue 1, pp. 121-140, 2013

[31] Miranda L., Vieira T., Martinez D., Lewiner T., Vieira A. W., Campos F.M., “Real-time gesture recognition from depth data through key poses learning and decision forests,” Brazilian Symposium of Computer Graphic and Image Processing, 25th SIBGRAPI: Conference on Graphics, Patterns and Images, pp. 268-275, 2012

[32] Castellano G., Kessous L., Caridakis G., “Emotion Recognition through Multiple Modalities: Face, Body Gesture, Speech,” Affect and Emotion in Human-Computer

Interaction, Lecture Notes in Computer Science, Volume 4868, pp. 92-103, Springer Berlin Heidelberg, 2008

[33] Kazemifard M., Ghasem-Aghaee N., Koenig B. L., Ören T. I., "An emotion understanding framework for intelligent agents based on episodic and semantic memories," Autonomous Agents and Multi-Agent Systems, Springer US, doi: 10.1007/s10458-012-9214-9, 2013

[34] Yamazaki Y., Hatakeyama Y., Dong F., Nomoto K., and Hirota K, "Fuzzy Inference based Mentality Expression for Eye Robot in Affinity Pleasure-Arousal Space," Journal of Advanced Computational Intelligence and Intelligent Informatics, Volume 12, Number 3, pp. 304-313, 2008

[35] Matsumoto D., "Cultural Similities and Differences in Display Rules," Motivation and Emotion, Volume 14, Number 3, pp. 195-214, 1990

[36] Matsumoto D., "Culture and Emotional Expression. Understanding Culture: Theory, Research, and Application," Psychology Press, ISBN-10: 1848728085, pp. 263-279, 2009

[37] Russel J., "A Circumplex model of affect," Journal of Psychology and Social Psychology, Volume 39, Number 6, pp. 1161-1178, 1980

[38] Livingston M. A., Sebastian J., Ai Z., Decker J., "Performance Measurements for the Microsoft Kinect Skeleton," Conference Proceedings, Virtual Reality Short Papers and Posters IEEE , doi: 10.1109/VR.2012.6180911, pp. 119-120, 2012

[39] Kinect libraries used in the coding: Audio library http://msdn.microsoft.com/en-us/us-en/library/jj131025.aspx, Face Tracking library http://msdn.microsoft.com/en-us/library/jj130970.aspx, Skeletal Tracking library http://msdn.microsoft.com/en-us/us-en/library/jj131025.aspx

[40] Burkhardt F., Paeschke A., Rolfes M., Sendlmeier W., Weiss B., "A Database of German Emotional Speech, Proceedings Interspeech Lissabon," Portugal , pp. 1517-1520, 2005

[41] Liu Z., Wu M., Li D., Chen L., Dong F., Yamazaki Y., Hirota K., "Concept of Fuzzy Atmosfield for Representing Communication Atmosphere and its Application to Humans-Robots Interaction," Journal of Advanced Computational Intelligence and Intelligent Informatics, Volume 17, Number 1, pp. 3-17, 2013

[42] Y. Liu, O. Sourina, and M. K. Nguyen, "Real-time EEG-based Human Emotion Recognition and Visualization," 2010 International Conference on Cyberworlds (Singapore), pp.262-268, DOI 10.1109/CW.2010.37, 2010.

[43] B. I. Ashish, and D. S. Chaudhari, "Speech Emotion Recognition," International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Vol. 2, Issue 1, pp. 235-238, 2012.

[44] M. Gardner, "Piet Hein's Superellipse," Mathematical Carnival: A New Round-Up of Tantalizers and Puzzles from Scientific American, Knopf Doubleday Publishing Group, Chapter 18, pp. 240-254, 1977.

[45] Reina Shimizu, Katsuhiko Ogawa, "Which Is More Effective for Learning German and Japanese Language, Paper or Digital?," Learning and Collaboration Technologies. Technology-Rich Environments for Learning and Collaboration Lecture Notes in Computer Science Volume 8524, pp 309-318, 2014.

[46] Asil Oztekin, Zhenyu James Kong, Ozgur Uysal, "UseLearn: A novel checklist and usability evaluation method for eLearning systems by criticality metric analysis," International Journal of Industrial Ergonomics, Volume 40, Issue 4, pp. 455–469, 2010

[47] Funda Dağ, Aynur Geçer, "Relations between online learning and learning styles," Procedia - Social and Behavioral Sciences, Volume 1, Issue 1, pp. 862–871, 2009

[48] Jia-Jiunn Lo, Ya-Chen Chan, Shiou-Wen Yeh, "Designing an adaptive web-based learning system based on students' cognitive styles identified online," Computers & Education, ELSEVIER, Volume 58, Issue 1, pp. 209–222, 2012

[49] Mieke Vandewaetere, Piet Desmet, Geraldine Clarebout, "The contribution of

learner characteristics in the development of computer-based adaptive learning environments," Computers in Human Behavior, ELSEVIER, Volume 27, Issue 1, pp.118–130, 2011

[50] Jieun Kim, Ahreum Lee, Hokyoung Ryu, "Personality and its effects on learning performance: Design guidelines for an adaptive e-learning system based on a user model," International Journal of Industrial Ergonomics, ELSEVIER, Volume 43, Issue 5, pp. 450–461, 2013

[51] Garcia S. Jesus A., Shibata Atsushi, Fangyan Dong, Kaoru Hirota, "Deep Level Emotion Understanding based on Customized Knowledge for Agent to Agent Communication," (IWACIII2014), pp 7-12, Fukui, Japan, 2014

Related Publications

Journal Papers

[J1] **Jesus A. Garcia Sanchez**, Kazuhiro Ohnishi, Atsushi Shibata, Fangyan Dong, and Kaoru Hirota, "Deep Level Emotion Understanding using Customized Knowledge for Human-Robot Communication," *Journal of Advance Computational Intelligence and Intelligent Informatics (JACIII)*, Vol. 19 No. 1, 2014.

International Conference Papers

[C2] **Jesus A. Garcia Sanchez**, Kazuhiro Ohnishi, Atsushi Shibata, Fangyan Dong, and Kaoru Hirota, "Visualization Method of Emotion Information for Long Distance Interaction ," 7th IEEE International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management, Philippines (IEEE HNICEM – ISCIII 2014),p. 115, 2014.

[C1] **Garcia S. Jesus A.**, Shibata Atsushi, Fangyan Dong, Kaoru Hirota, Deep Level Emotion Understanding based on Customized Knowledge for Agent to Agent Communication. *(IWACIII2014)*, pp 7-12, Fukui, Japan, 2014.

Domestic Conference Papers

[D1] **Jesus A. Garcia Sanchez**, Kazuhiro Ohnishi, Atsushi Shibata, Fangyan Dong, and Kaoru Hirota, "Concept of Deep Level Emotion Understanding and its Visualization Method," 30th Fuzzy System Symposium (Kochi, September 1-3, 2014), WD1-1, pp.796-797, 2014.

Awards

[A1] **Best Paper Award: “Visualization Method of Emotion Information for Long Distance Interaction”**
7th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM 2014), joint with 6th International Conference on Computational Intelligence and Intelligent Information (ISCIII) and co-located with 10th ERDT Conference. Hotel Centro, Puerto Prinsesa, Palawan, Philippines, November, 2014.

Others Publications

Journal Papers

1. Sun Bo, Iliyasu Abdullah M., Yan Fei, **Sanchez Jesus A. Garcia**, Dong Fangyan, Al-Asmari Awad Kh., Hirota Kaoru, “Multi-channel information operations on quantum images,” Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol. 18, Num. 2, pp 140-149, 2014

2. Betancourt Janet Pomares, Fatichah Chastine, Tangel Martin Leonard, Yan Fei, **Sanchez Jesus Adrian Garcia**, Dong Fang-Yan, Hirota Kaoru, “Similarity-based fuzzy classification of ECG and capnogram signals,” Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol. 17, Num. 2, pp 302-310, 2013

3. Iliyasu Abdullah M. , Le Phuc Q., Yan Fei, Sun Bo, **Garcia Jesus A.S.**, Dong Fangyan, Hirota Kaoru, “A two-tier scheme for greyscale quantum image watermarking and recovery,” International Journal of Innovative Computing and Applications, v 5, n 2, p 85-101, 2013

4. Iliyasu Abdullah M., Le Phuc Q., Yan Fei, Sun Bo; **Garcia Jesus A.S.**, Dong Fangyan, Hirota Kaoru, “A two-tier scheme for greyscale quantum image watermarking and recovery,” International Journal of Innovative Computing and Applications, Vol. 5, Num. 2, pp 85-101, 2013

5. Phuc Q. Le, Abdullah M. Iliyasu, **Jesus A. Garcia Sanchez**, Fangyan Dong, and Kaoru Hirota, "Representing Visual Complexity of Images Using a 3D Feature Space Based on Structure, Noise, and Diversity," Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol.16, No.5, pp. 631-640, 2012

International Conference Papers

1. Fei Yan, Le P.Q., Iliyasu A.M., Bo Sun, **Garcia J.A.**, Fangyan Dong, Hirota, K., "Assessing the similarity of quantum images based on probability measurements," WCCI 2012 IEEE World Congress on Computational Intelligence, 2012

2. Phuc Q. Le, Abdullah M. Iliyasu, **Jesus A. Garcia Sanchez**, Fangyan Dong, and Kaoru Hirota, “A 3D feature space for representing visual complexity of images using their structure, noise, and diversity,” 12th International Symposium on Advanced Intelligent Systems (ISIS2011), Sep. 2011.

3. Sun Bo, Le Phuc Q., Iliyasu Abdullah M., Yan Fei, **Garcia, J. Adrian**, Dong Fangyan, Hirota Kaoru, “A multi-channel representation for images on quantum computers using the RGB color space,” WISP 2011 - IEEE International Symposium on Intelligent Signal Processing, Proceedings, pp 160-165, 2011

4. Betancourt Janet Pomares, Fatichah Chastine, Tangel Martin Leonard, Yan Fei, **Sanchez Jesus Adrian Garcia**, Dong Fang-Yan, Hirota Kaoru, "Similarity-based Fuzzy Inference System for ECG and Capnogram Signals Classification," International Symposium on Soft Computing Sponsored by ASPIRE LEAGUE, 2012.