Construction of Virtual Assistant Based on Basic Emotions Theory

Zhiliang Wang, Ning Cheng, Yumei Fan, Jiwei Liu, and Changsheng Zhu

University of Science and Technology, Beijing 100083
kinchengning.student@sina.com
kinchengning@263.net

Abstract. The purpose of this paper is to construct a virtual assistant. Basic emotions theory points out that compound emotion consists of eight prototype basic emotions and "drives" which reflects people's will. According to this theory, we construct a psychology model. By adjusting parameters in it, we can simulate different human psychologies. Based on this model, combining real-time facial expression and voice recognition and synthesizing technology, we construct a virtual assistant. Proved by experiment, our system obeys human emotion rules.

1 Introduction

Artificial psychology is the machine realization of human psychology actives using research methods in information science. The purpose of artificial psychology is to meet users' requirements for life quality and it has many applying fields. For example, harmonious human-computer interaction needs the comfortable interface which can change automatically according to user's feelings. The virtual assistants of automatic commodity selecting may have the ability of automatically supplying commodities which consumers are interesting in. Computer games require virtual roles to have emotion for attracting users, etc. Currently, active research is being conducted on a user-friendly human-machine interface, such as Virtual characters found in the Oz Project [6] and the Virtual Theatre [7], as well as the humanoid robot WR-4 [8], all includes emotion as a fundamental component. Although some progress has been made in creating agents that display believable emotions, there are still many problems need to resolve. Two important problems in them are how to construct model for emotion [9], and affection model how to be applied in agent for improve agent's performance [10]. So our current goal is to try to resolve these problems and to construct a virtual assistant which have emotions and can interact with users. In Fig.1, we show our system structure chart.

Perceptual input layer gets and processes the sensor information which is associated with the emotion parameters in basic emotion equations by voice and facial expression recognition. Basic emotion layer computes the basic emotion intensity according to emotion equation. Motivation layer describes "drives" of virtual assistant which represents its will. Compound emotion layer computes compound emotion using output of basic emotion layer and motivation layer. Behavior layer responds to the user by voice and facial expression synthesizing. Calm layer denotes the calm state of psychology.

J. Tao, T. Tan, and R.W. Picard (Eds.): ACII 2005, LNCS 3784, pp. 574 – 581, 2005.

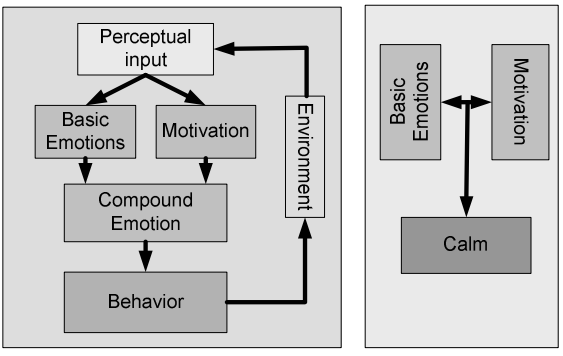

Fig. 1. System structure chart

2 Emotion Model

Psychology model is the core of the system, so we introduce it in detail.

2.1 Emotion Theory

Our emotion theory foundation is basic emotion theory. It points out that emotion has prototypes, i.e. for human, there are many basic emotion types each of which has unique characters, biologic waken mode and display mode. The different combination of these basic emotions form all the human emotions[1]. From the view of individual development, the generation of basic emotions is the result of organism natural mature but not natural study. Plutchik uses one single word to describe each of the eight basic emotions and makes a table as follow[3]:

Table 1. Basic emotions classification and their intensities

Intensity / type	None	Little	middle	strong	Very strong
Happiness	1	2	3	4	5
Acceptance	1	2	3	4	5
Surprise	1	2	3	4	5
Fear	1	2	3	4	5
Sadness	1	2	3	4	5
Disgust	1	2	3	4	5
Interesting	1	2	3	4	5
Anger	1	2	3	4	5

In his book, Frend usually uses the concept of “drives” to represent an inside incentive which is used to affect individual behaviors by adjusting activity contents and types[3]. Emotion has many compound modes. Basic emotions combined with “drives” form the affective experiences such as: content, urgent, eager, etc[3].

2.2 Mathematic Model

Our mathematic model has two parallel parts: motivation driving part and emotion calming down part.

2.2.1 Motivation Driving Part

According to psychology theory mentioned above, we analyze the relations among basic emotions, "drives" and environment information now.

1) Basic Emotion Layer

First, look at basic emotions. We think that people's psychology is a self organization procedure. Self organized natural things have these characters: each part of them exists both relying on other parts and for other parts and these parts interact with each other to integrate as a whole[4]. Think about one basic emotion. Intensity is the most obvious property of emotion and used as order parameter. Let $P(t)$ to be intensity. $P(t)$ changes according to outer stimulations and outer stimulation A is in direct proportion to current emotion intensity, so we induce item $AP(t)$. Because emotion goes to calm over time, we induce another item $-BP^2(t)$. At last, we get the equation:

$$\frac{\mathrm{dP(t)}}{\mathrm{dt}} = AP(t) - BP^2(t) \tag{1}$$

In which, $A \geq 0$, $B > 0$. Initial value is P_0 . When $t \to +\infty$, the solution $P(t) \to \frac{A}{B}$. It means that emotion intensity goes to stable value at last. A is the outer stimulation parameter and B is a predefined value which reflects emotion states. Eight basic emotions correspond to eight such equations:

$$\frac{\mathrm{dP_m(t)}}{\mathrm{dt}} = A_m P_m(t) - B_m P_m^2(t), m = 1,2,\cdots 8 \tag{2}$$

2) Driver Layer

People's motivation ("drives") reflects their will and behavior mode, so it has important position in our model. How to describe motivation precisely is the next problem we need to discuss. Use a vector to denote "drives": $\delta(\delta_1, \delta_2, \delta_3, \delta_4, \delta_5, \delta_6, \delta_7, \delta_8)$. "drives" reflects people's will, so it relates to people's attitudes about outer stimulations. Generally speaking, if outer stimulation is big at one component, "drives" is also big at this component but not in direct proportion to it. Let:

$$\tilde{A} = \frac{A}{\|A\|} = (\tilde{A}_1, \tilde{A}_2, \tilde{A}_3, \tilde{A}_4, \tilde{A}_5, \tilde{A}_6, \tilde{A}_7, \tilde{A}_8),$$

In which, $\|A\| = \sum_{m=1}^{8} A_m$. δ_i *is the component of "drives" vector. When* $0 \le a_i \le a$ *, let* $\delta_i = 0$*, when* $a < a_i \le 1$*, let* $\delta_i = 1 - e^{-(A_i - a)^2}$.

3) Compound Emotion Layer

When we get the expression of "drives", we think about the relation between "drives" and compound emotion. When emotion rising, the drive also rise, so emotion strengthens the "drives". However, emotion and "drives" both have decay tendency. According to these, we describe their relation as:

$$\begin{cases} \dfrac{du}{dt} = -\alpha u + s \\ \dfrac{ds}{dt} = -\beta s + u^2 \end{cases} \tag{3}$$

In which, s is one basic emotion intensity in compound emotion and u is corresponding component's intensity of "drives". α is the decay coefficient of "drives", so α >0. β is the decay coefficient of emotion, so β>0. $s(t)$ can be expressed as:

$$s(t) = (\frac{1}{\beta} + \frac{2}{\beta^2}\alpha + \frac{4}{\beta^3}\alpha^2)u^2(t) \tag{4}$$

Considered the impact of basic emotions to compound emotion, compound emotion is decided as follow:

$$P(t) = (\frac{1}{\beta} + \frac{2}{\beta^2}\alpha + \frac{4}{\beta^3}\alpha^2)u^2(t) \times p(t) \tag{5}$$

In which, $P(t)$ is a component of compound emotion vector, $u(t)$ is the corresponding component of "drives" vector, $p(t)$ is corresponding basic emotion. α, β are constants which relate to personality and sensitivity and are used to control emotion's fluctuation range. The vector of compound emotion consists of eight components $P_i(t)$.

2.2.2 Emotion Calming Down Parts

After rising states, emotion will go to calm automatically over time, i.e. basic emotion goes to calm under the impact of "drives". In this section, we describe this process. We give a method to describe the relation between "drives" and basic emotions. Generally speaking, people's attitude for one thing is in direct proportion to emotion, but the drive to do thing will also decay over time, so we think that "drives" has self-decay tendency and also has the rising tendency under the impact of basic emotion. Basic emotion is decaying under the impact of "drives" and the decay rate is

associated with itself. The bigger its intensity is, the faster it decays. Now, we give the mathematic expressions as follow:

$$\begin{cases} \dfrac{dX}{dt} = k_1 X - k_2 XY = G_x(X, Y, k_1, k_2) \\ \dfrac{dY}{dt} = k_2 XY - k_3 Y = G_y(X, Y, k_2, k_3) \end{cases} \tag{6}$$

In which, X is intensity of one basic emotion and $\frac{dX}{dt}$ is intensity's change rate. Y is intensity of corresponding component of "drives" and $\frac{dY}{dt}$ is this component's change rate. k_1 is basic emotion's rise rate. k_2 is basic emotion's decay rate and is also the rise rate of corresponding component of "drives". k_3 is decay rate of corresponding component of "drives".

We solve this equation set and get the static solution as follow:

$$X_0 = \frac{k_3}{k_2} \quad Y_0 = \frac{k_1}{k_2} \tag{7}$$

Because emotion always has some little disturbance, we analyze the static solution for its stability. Let:

$$X = X_0 + x \quad Y = Y_0 + y \tag{8}$$

In which, $|x| << X_0, |y| << Y_0$.

The result is as follow:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \exp[\pm i\omega t] \tag{9}$$

In which, $\omega = \sqrt{k_1 k_3}$. Expression (9) indicates that if emotion is at some initial state of $\begin{bmatrix} X_0 + x_0 \\ Y_0 + y_0 \end{bmatrix}$ which depart static state of $\begin{bmatrix} X_0 \\ Y_0 \end{bmatrix}$, when t=0, over time, emotion neither to go away infinitely, nor to go to the static state infinitely, but fluctuate around the static state of $\begin{bmatrix} X_0 \\ Y_0 \end{bmatrix}$. The fluctuant frequency is $\omega = \sqrt{k_1 k_3}$. This result describes the phenomenon that emotion fluctuates around a static state when it is calm.

3 Construction of Virtual Assistant

Based on this psychology model, combining real-time facial expression and voice recognition and synthesizing technology, we construct a virtual assistant.

3.1 Psychology Model Realization

We program to realize our psychology model and show it in Fig. 2. The upper two curve graphs describe the changing process of basic emotion and compound emotion in motivation driving part and the third curve graph describes the emotion calming down process.

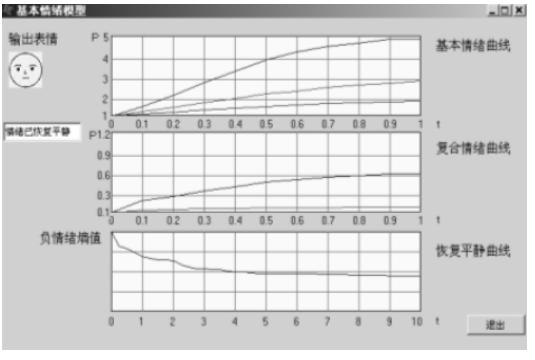

Fig. 2. Realization of psychology model

3.2 Behaviors Database

We construct a rule controlled behaviors database in which we define many usual behaviors and speech such as: can I help you, ok, thank you, etc. For each behavior and speech, we define the applying range rank, so virtual assistant's behavior directly associates with emotion.

3.3 Real Time Expression Recognition and Synthesizing

Real time expression recognition result is used as system input. It is based on exactly character point location. First, we localize the key points of eye, eyebrow and mouth exactly and extract their contours by curve fitting. Second, we analyze and record facial components' movement information, then use our HMM expression model to get recognition result. In Fig. 3, we show a row of recognition results.

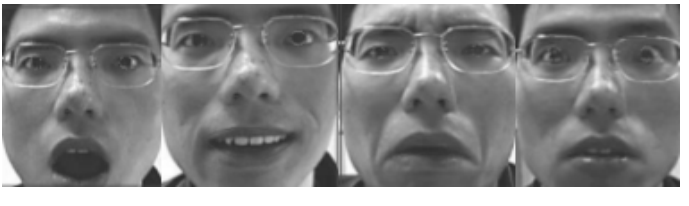

Fig. 3. Real-time recognition results of four facial expressions as: surprise, happiness, sadness and anger

We synthesize virtual facial expression to output. We use character polygons to render 3D face. In the polygon mesh, we define the feature points of eyes, nostril, eyebrows, mouth, etc. These feature points are used to describe the shape and position of facial feature. We adopt two methods for face modeling: one is linear flexible model used to generate face, another is muscle model used for animation and expression transformation. In Fig. 4, we show some of our output facial expressions.

Fig. 4. Face expressions as: joy, surprise, sorrow, and calm

3.4 Real Time Voice Recognition and Synthesizing

We use the ASR sdk software to realize voice recognition as system input. We constructed simple word and speech database, store data such as: good, bad, I need you help, etc, to lead virtual assistant's behaviors. We established corresponding relations between these data and emotion parameters, so voice input can directly impact virtual assistant's emotions.

We use Interphonic (Text To Speech) software to realize voice synthesizing to output. We also established database to store word and speech such as: good morning, nice to meet you, can I help you, etc, as the response of the virtual assistant to users. Of course, there are corresponding relations between data and emotion parameters, so virtual assistant can choose word according to his emotion states.

4 Conclusions

Based on psychology model, we use many technologies to construct a virtual assistant who can first recognize user's expressions and voice in real time, then judge by psychology model and last give the response. Proved by experiments, our system obeys human emotion rules. Of course, we also met some problems such as: expression recognition is not exact enough sometimes and the speech's emotion rank is difficult to decide, etc. Our next work is to continue our research and solve the problem we met. We believe that we will design better system and our system will serve people in near future.

References

1. Ekman, P. An argument for basic emotions, Cognition and Emotion[J], 1992b,6,
2. Zhiliang Wang, Lun Xie, Artificial Psychology—an Attainable Scientific Research on the Human Brain[J]. IPMM'99(KEYNOTE PAPER), honolulu, USA, July 10-15,1999(ISTP)
3. Meng zhaolan, Human emotion[M], Shanghai press, 1989

4. Wu tong, Self-organization methods research[M], Tsinghua press, 2001
5. Weimin Xue, Zhehua Wei, Zhiliang Wang, The Research of Artificial Psychology Model[J], IEEE TENCON'02., Vol. I, p691-693, Beijing, China. October 30_31, 2002.
6. J.Bates, The role of emotion in believable agents. Communications Of The ACM,37(7):122-125. 1994.
7. B.Hayes-Roth, and R..van Gent, Story-making with improvisational puppets. in Proc. 1st Int. Conf. On Autonomous Agents, (Marina del Rey, CA, 1997), 1-7.
8. H. Miwa, T.Okuchi, K. Itoh, H. Takanobu, A.Takanishi: A new mental model for humanoid robots for human friendly communication - introduction of learning system, mood vector and second order equations of emotion, Proceedings of The 2003 IEEE International Conference on Robotics and Automation, pp.3588-3593, 2003
9. C.Calhoun, R.Solomon. What is an emotion[M] . London : Oxford University Press, 1984.
10. Camurri, A. Coglio. An architecture for emotional agents[J] . IEEE Multimedia , 5 (4) : 24 – 33,1998.

Appendix: Support Foundation

Foundational project: Supported by the key Lab named Advanced Information Science and Network Technology of Beijing (No.TDXX0503) and the key foundation of UST of Beijing.