Recognizing and Expressing Affect

Xin Chen

Abstract— In this paper, firstly introduce the research status of facial expression recognition and then the different ways to recognize the emotions. After that, an affective model is built, in which the transition of emotion can be viewed as a Markov process. In addition, in this model, the relation of emotion, external incentive and personality are also be described, and then the Dempster-Shafer evidence theory is used to combine the emotion information from vison, speech and other aspects. Then, this affective model will be integrated into a robot system and make it have emotion and personality just like a human.

Keywords —Affective model, Markov model, Dempster-shafer evidence theory

1 INTRODUCTION

IN our daily life, affective competence with normal rational thinking and logical reasoning ability is playing an important role. At present the computer is still based on logical reasoning ability, if humans are able to bring the computer emotional ability, like bringing the computer a higher, comprehensive intelligence, establishing a harmonious environment for human-computer interaction, and giving the computer ability to adapt to human emotions, then to some extent, it will provide a lot of convenience to life. Shaping the artificial character and combining it with facial expression recognition, expression synthesis and artificial intelligence techniques can construct a virtual human proto-system which can continuously express emotions.

Affective computing refers to the calculation which arises from or deliberately influences emotions, and its main purpose is that computers are able to acquire the ability of recognizing, understanding and expressing affect. Affective computing will receive facial expression and signal from the change of people's bodies by various sensor and then recognizing these signals by affective models, which is for understanding human emotions and make an appropriate response.

This project will focus on emotion modeling, emotion recognition and its related applications. Emotional models are created to make computers smarter, more friendly, and more capable. Therefore, we use a variety of methods to establish emotional models, which is for bringing humans and computer better interaction. In addition, collecting some relevant information like reading papers, journals and magazines, can make us understand the current development of the technology and future applications.

2 RESEARCH STATUS OF FACIAL EXPRESSION RECOGNITION

Facial expression recognition is the first step of affect recognition, which plays an important role of results of affect recognition. Nowadays, most of researches are under the circumstance of having clear and discernible faces.

There are two ways for us to locate and detect a face in a picture. One, we can regard a face as a whole to recognize. Two, we can detect some important characters of faces to recognize them.

Pantic and Rothkrantz viewed a face as a whole to recognize. They focused on the shapes and colors of faces. However, there are many disadvantages like they are not able to eliminate the interference of glasses or hairs on faces.

In 2001, Paul Viola held the view that people could apply AdaBoost algorithm to recognizing faces. AdaBoost algorithm would select some vital characters to construct classifiers and rapidly focused on some areas which are like faces. This method not only improved the speed of detection but also increased the detection rate. In addition, most of facial recognition systems are based on AdaBoost algorithm and then improve themselves.

2.1 Feature Extraction

Nowadays, the technology of facial recognition has improved a lot and it also meets needs of researches, so the researches of facial affect recognition now are mostly focus on how to make sure different emotions match with the different correct facial expression characters. In general, facial expressions are classified as surprise, fear, abhor, anger, happy and sad. Therefore, most of researches are focusing on facial feature extraction.

The feature extraction of facial expressions is important, and the feature directly determines the speed and efficiency of recognition. For static images, the deformation features of expressions need to be extracted. Such methods generally only consider the spatial information and the face geometry information of a single frame image, and the calculation will be simple. The image sequence is different. It is necessary not only to consider the characteristics of deformation in each frame, but also to consider the movement characteristics of the image. This method has a high recognition rate, but it is more difficult to calculate. Considering the different ranges of images, features can be further divided into two types, global and local. The global feature extraction method directly extracts the motion features and deformation features of the entire face. Based on the local feature extraction method, the face is divided into several sub-regions. Each sub-region is extracted and then the results are concatenated. The facial deformation feature extraction method refers to the neutral expression face, and the deformation information is obtained by comparing the neutral face with the current face. The combination of information should be reflected in two aspects: facial shape changes and facial lines change. This type of extraction can be roughly divided into three categories:

(1) Method based on geometric features. The human face shows different expressions through changes in the eyes, nose, mouth, etc. It is possible to create a geometric model of a specific area of the face and extract the geometric features of the expression. Pantic M. et al. used multiple detectors to spatially locate human faces and facial organs, and finally extracted ten human face fiducials and 19 human facial organ fiducials to identify 32 muscle motor units.

(2) Based on Gabor wavelet method. Wavelet transform essentially uses a set of filters of different scales to filter the signal and then decompose the result into different frequency bands for processing and analysis. The effect of this method is similar to the multi-channel filtering model in the human visual system. Therefore, in the face expression recognition application, wavelet transform is used to extract the wavelet feature of the image. The Gabor wavelet is insensitive to changes in global brightness and affine transformations. It reflects image structure characteristics. The Gabor wavelet coefficient method with uniform sampling points and the Gabor wavelet transform-based recognition method have achieved good results in facial expression recognition. The Gabor wavelet method is less affected by the change of illumination, it can detect multi-directional and multi-scale texture changes, and it is widely used in feature extraction of face expressions. Donato compares several methods to identify the performance of facial AU, and the results show that wavelet transform is superior to other methods. Yao Wei has proposed a two-layer Gabor feature selection method. Firstly, the high-dimensional vector is filtered by improving the variance ratio as the distinguishing ability of the evaluation feature, and then the features of the obtained feature subset are selected by AdaBoost. Finally, the most discriminative feature is selected to achieve the purpose of reducing the feature dimension.

(3) Model-based approach. A model is used to describe the structure of a human face. A geometric model (point model) is a simple model method. Tim Cootes, a scientist at the Department of Image Science-Biology Engineering at the University of Manchester, UK, first proposed the AAM method, which is also a widely used face feature extraction method. AAM differs from the Active Shape Model (ASM) because it takes into account the surrounding area covered by the object shape, in addition to the information on the edge of the object shape. The AAM method statistically models the textures and shapes of interest objects so that the resulting model contains as many valid objects as possible. The AAM method consists of two parts: building an object model and searching for new objects. Firstly manually locate the feature point object in the image. After the shape of the object is formed, the texture of the object is generated based on the image covered by the shape. A statistical model is created for all the textures and shapes of the objects to be learned, and the Appearance Model is created using the Texture Model and the Shape Model. Create new objects by changing the appearance model's parameters. When the texture of the object to be detected and the object error generated by the model are the minimum, the position and shape information of the new object are obtained, and the parameters of the model are assigned to the new object. The model-based method can obtain more reliable face features, but it has the disadvantages of difficulty in obtaining initial points and large calculation volume. Zuo Kunlong analyzed the feasibility of facial expression recognition (FER) using facial expression features extracted from M. He tried to use the FER based on this feature vector to locate the eye area based on the characteristics of the human face image, and then use AAM's optimization algorithm to obtain the new object's characteristics. This not only improves the accuracy of positioning, but also greatly reduces the time required to locate new objects from the M method.

(4) Subspace-based methods. The dimensions of facial expression images are generally very high. At the same time, the distribution in high-dimensional space is very scattered, which is not conducive to classification, and the computational complexity is also quite large. In order to obtain a more compact distribution of the image, it is necessary to reduce the dimension of the original high-dimensional space, and finally perform classification and recognition of the expression in the low-dimensional subspace. The idea of subspace analysis is to reduce the compression of the original data to a low-dimensional space, to make the distribution of data in the subspace more compact, and to greatly reduce the subsequent calculation. Subspace-based facial expression recognition methods can be divided into two kinds of nonlinear and linear space transformation. The linear subspace methods that have been successfully applied in facial expression recognition include: Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), and Independent Element Analysis (CA). Nuclear-based nonlinear subspace analysis methods include: Kernel Principal Component Analysis (KPCA) and Nuclear Fisher Discriminant Analysis (KFDA).

(5) Based on the local binary mode method. The Local Binary Patterns (LBP) can describe the texture features of the graph well. It obtains local texture features by comparing the gray value of any point in the image with the surrounding points. Due to its good characteristics of anti-rotation and brightness changes, it has been widely used in many fields of pattern recognition in recent years. In terms of precise positioning of faces, face detection, image content recognition, texture recognition, and face recognition, good results have been achieved. Compared with the Gabor wavelet feature, extracting LBP features in face images will be faster, and the dimension will be much smaller, and at the same time, the face information will be effectively preserved. Shan uses LBP to extract the texture features of the face image. Firstly she calculated the LBP histograms of the face region in blocks, and then uses these LBP histograms to string together all of them as facial feature results, which has a better recognition effect. Sun Ning et al. proposed a local feature extraction method based on 2-D PLS (2DPLS) and applied it effectively in facial expression recognition. The method first used the LBP

operator to extract the texture features of all the subregions of an image, and then combined them into a local texture feature matrix. 2DPLS adapted the corresponding relation matrix so that the matrix form can adapt to the sample, and at the same time, it showed the importance of different local information in the face. Cui Jie and others used LBP to extract facial features and perform coarse-to-fine facial expression classification. At the rough classification stage, two expressions were selected as preliminary classification results (candidate expressions). Then in the fine classification stage, the final classification was determined by calculating the weighted chi-squared value. The facial motion feature extraction rule was to compare the changes of face information in a period of time. The extraction of such features includes feature point tracking, optical flow, differential images, and motion models. The optical flow method is an earlier and more widely used method. It uses the optical flow to express the characteristic movement of different expressions, and uses the optical flow of the characteristic point as the expression feature. For example, Lien et al. used a wavelet-based multi-resolution dense optical flow to analyze the whole face motion. However, the optical flow method has a large amount of calculations and is sensitive to the discontinuity of motion. It may take a lot of time to calculate the optical flow in the entire face area. Feature point tracking is a traditional algorithm for extracting facial features. It is different from the optical flow method and does not track every pixel point in the image. It only tracks certain interest points. The marker point method is similar to the feature point tracking method, but more is used in areas where the cheeks do not have much texture information.

2.2 Expression Recognition

Expression classification is the final stage of expression recognition. It is closely related to the extraction methods of facial expression features. The purpose is to classify different expressions into the corresponding expression classes. Some scholars have suggested that how to select and design classifiers depends on whether they use time information. The classification method using time information is called time domain airspace method, and the classification method that does not use time information is called airspace method, similar to the static and dynamic method of division. Typical of the airspace method is the artificial neural network. The neurons in the input layer correspond to the input face data, and the output neurons correspond to expression classes. However, neural network training is difficult, so Kaiser proposed a rule-based neural network. Because the airspace method only uses eigenvectors, the general classification method can be classified as this type. Wang et al. used the adaboost algorithm to construct a classifier, and Buciu et al. used a support vector machine (SVM) as a classifier, and Guo et al. used a linear programming method to construct a classifier. At the same time, the expression is also a dynamic process, and the higher recognition rate can be obtained by considering the use of dynamic information when designing and selecting methods. The time domain airspace method effectively uses dynamic information. Currently Hidde Markov Model (HMM) is often used in facial expression recognition to better describe the dynamic sequence. In addition, recursive neural networks also have many successful applications in facial expression recognition systems.

2.3 Emotion Research and Application

Emotion computing is a new multidisciplinary research area. It includes computer science, sensor technology, psychology, physiology, cognitive science, philosophy, and so on. The ultimate goal of emotional computing is to make computers feel like humans. There are many problems that need to be solved to achieve this goal now. The key technologies studied include human emotion recognition, emotional modeling, and effective expression of emotional perception. Emotional experience includes both psychological and physiological processes. Emotional research attempts to make an overall explanation of the psychological and physiological processes of emotions and the relationships between them. Psychologists put forward a variety of perspectives from different perspectives and research perspectives, thus forming many emotional theories. On the basis of emotional theory, people have further established an emotional model that can explain the process of emotional development. These theories include physiological response theory, stimulus-response theory, motivation theory, facial expression theory, and subjective evaluation theory. At present, the cognitive evaluation theory of emotion is most accepted by people. The construction of emotional subjects is mostly based on certain cognitive evaluation theories, and artificial intelligence scholars have also pay more attention to it. According to the theory of cognitive evaluation, when the subject evaluates an event that is subjectively important, it will produce an emotional experience, which in turn will generate emotions. This evaluation process is more subjective, and it depends on the beliefs of the subject, specific goals and norms. Different subjects have different internal psychological structures. As a result, even the same external stimulus, the interpretation may be different. The ultimate emotions will be produced through their own cognition and subjective evaluation of the stimuli.

The most influential of these evaluation theories is the OCC theory proposed by Ortony, Collins and Clore. The OCC model is the first model designed for computer implementation. They assume that emotions result from a process called cognitive evaluation. Evaluation depends on three components: events, objects, and subjects. According to the subjective goal, the events in the objective world are rated as satisfactory or not. Subjects are rated like or dislike based on the subject's attitude. The behavior of the subject itself or other subject is rated as favored or disapproved according to a set of criteria. These evaluation variables eventually form a hierarchy of 22 emotions.

The synthesis of happy emotions in the OCC model is expressed as follows:

Let D(p,e,t) represent the degree of expectation that person p wants event e at time t. If event e is good, it will return a positive value.

Let P_j(p, e, t) express the possibility of happiness, and I_g(p, e, t) is the total combination of intensity variables (eg, expectation, approximation, reality). An example of a rule that produces happy is

$$D(p, e, t)>0 \tag{1}$$

$$I_j(p, e, t)=F_j(D(p, e, t), I_g(p, e, t)) \tag{2}$$

And $F_j()$ is specified as a function that represents happiness. Other emotions such as sad can also be implemented with similar rules. $P_d()$ can be changed by changing the expectation of negative conditions in D(p, e, t), $I_j(p, e, t)= F_d(D(p, e, t), I_g(p, e, t))$.
The above formula does not cause state happiness, but it can trigger another rule, thus changing the intensity of happiness I_j. After a threshold is given

$$P_j(p, e, t)>T_j(p, t) \tag{3}$$

$$I_j(p, e, t)=P_j(p, e, t)- T_j(p, t) \tag{4}$$

$$I_j(p, e, t)=0 \tag{5}$$

This rule activates happy emotions. When intensity exceeds a happy threshold, the resulting intensity can be mapped into happy emotions.
Happiness is the simplest example. The realization of other emotions in the OCC model is more complicated.
OCC structure as shown:

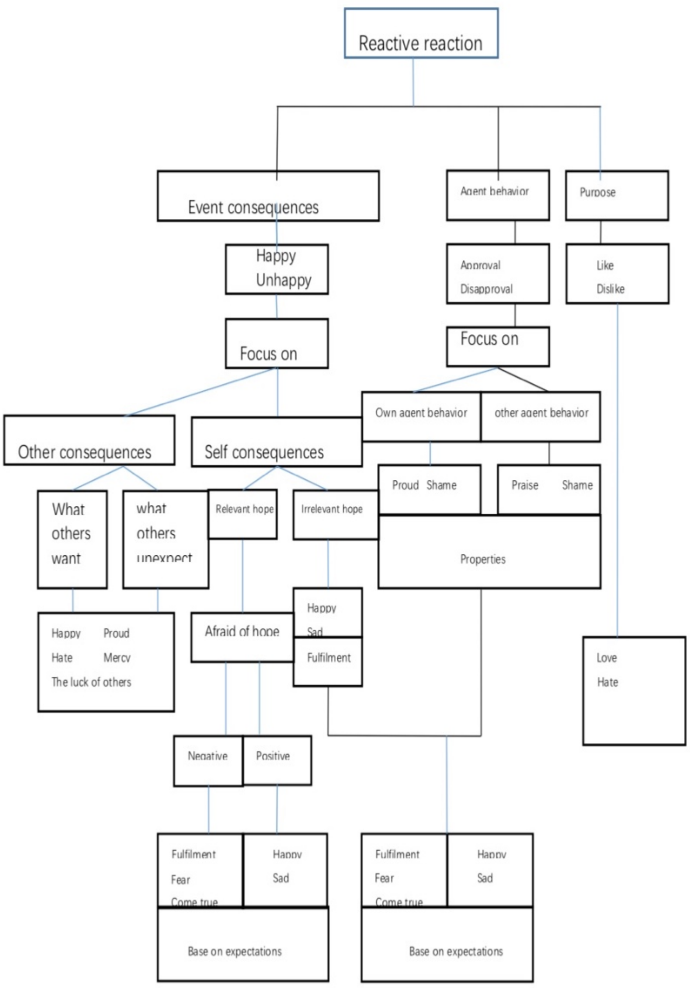


Fig. 1. OCC Model Emotional Cognitive Structure

3 Create an emotional model

3.1 Markov Model

This paper analyzes the nature of the facial expressions and gives a qualitative description of the correspoing facial expression space, and then proposes a new facial-space model with the characters of both crete affective sp-ace model and continuous affective space model.
We will focus on discrete affective space model. Firstly, we will construct a three dimensional model based on fear, anger and happy, and any emotional status will match a certain point on the 3 dimensional space. In order to realize it, we will have to simplify that by discretizing. We can set that every basic emotion only have three intensity, which means that every dimensional will have 0, 0.5, and 1, three different values. Take happy for example, there are three strength, unhappy, a little happy and happy. Therefore, there are 27 discrete emotional status. We can define happy as a, anger as b, fear as c, so, there are $a \epsilon (0, 0.5, 1)$, $b \epsilon (0, 0.5, 1)$, $c \epsilon (0, 0.5, 1)$.

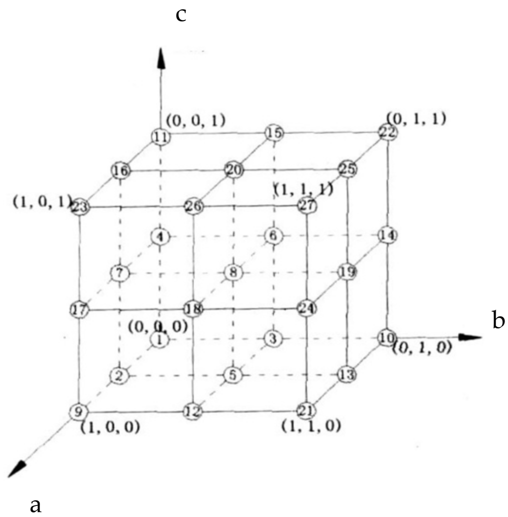


Fig. 2. Emotional Space Model

So the enumerable emotional state has 27 states, as shown in Table 1. The origin (0,0,0) not only represents calm, but also represents other unknown emotional states. In other words, when the living body only has the above three emotions, this is the complete emotional space, and the origin is the point of calm. When the emotion of a living body is not only these three emotions, we call this an incomplete emotional space. The origin may be a calm point, or it may be some other emotional state, but this emotional state is not included in this emotional space. When emotions change in this state space, it has a statistical pattern. This is a Markov process. Therefore, we can use the Markov model to describe the changes in the emotional state. It is very appropriate to describe the human emotions using the discrete state isomorphic Markov model.

TABLE 1
27 STATES IN THE EMOTIONAL STATE SPACE

No.	Status of Emotion
1	Quiet(0,0,0)
2	A little happy(0.5, 0, 0)
3	A little angry(0, 0.5, 0)
4	A little fear(0, 0, 0.5)
5	A little happy, a little angry(0.5, 0.5, 0)
6	A little angry, a little fear(0, 0.5, 0.5)
7	A little happy, a little fear(0.5, 0, 0.5)
8	A little happy, a little fear, a little angry(0.5, 0.5, 0.5)
9	Happy(1, 0, 0)
10	Angry(0, 1, 0)
11	Fear(0, 0, 1)
12	Fear, A little happy, a little angry(0.5, 0.5, 1)
13	Happy, A little angry, a little fear(1, 0.5, 0.5)
14	Angry, A little happy, a little fear(0.5, 1, 0.5)
15	Fear, happy, a little angry(1, 0.5, 1)
16	Happy, angry, a little fear(1, 1, 0.5)
17	Angry, A little happy, fear(0.5, 1, 1)
18	happy, a little angry(1, 0.5, 0)
19	Happy, a little fear(1, 0, 0.5)
20	Angry, A little happy(0.5, 1, 0)
21	A little angry, fear(0, 0.5, 1)
22	a little fear, angry(0, 1, 0.5)
23	A little happy, fear(0.5, 0, 1)
24	Happy, angry(1, 1, 0)
25	Fear, angry(0, 1, 1)
26	Happy, fear(1, 0, 1)
27	Happy, fear, angry(1, 1, 1)

During the 27 points, origin of coordinates means quiet or no emotions. Some vertex also means quiet or some other emotion status. It is clear that this model is not completed. When someone's emotion go through this space, it is easy for us to find its statistical natures. This process we also call it Markov. Therefore, we can use Markov model to describe how emotions change.

There is probability of changing emotion status in Markov model, so in this emotion model, there are 27 $P_{i,j}$ (i,j ∈ [1,2,3 … … ,27]) and they construct 27 dimensional Probability matrix A_p.

$$A_p=\begin{bmatrix} P_{1,1} & \cdots & P_{1,27} \\ \vdots & \ddots & \vdots \\ P_{27,1} & \cdots & P_{27,27} \end{bmatrix} \quad (6)$$

$P_{i,j}$(i,j ∈ [1,2,3 … … ,27]) is the probability of i_{th} status to j_{th} status. Additionally, there is a relationship among them:

$$\sum_{i=1}^{27} p_{i,j} = 1, i \in [1,2,3 \ldots \ldots,27] \quad (7)$$

We can conclude that if there are m emotions and there will be m dimensional emotional space, and for every emotion there are n levels, which means that there will be n^m emotion status. Reagard l=n^m, we can find that:

$$A_p=\begin{bmatrix} P_{1,1} & \cdots & P_{1,l} \\ \vdots & \ddots & \vdots \\ P_{l,1} & \cdots & P_{l,l} \end{bmatrix} \text{ and } \sum_{i=1}^{l} p_{i,j} = 1, i \in [1,2,3 \ldots \ldots, l] \quad (8)$$

In Markov model, the probability of change emotion will be influenced by many factors like personal characters, conscious stimulation. For example, under the positive simulation, the probability of changing a certain emotion to positive emotion will be larger than that of changing a certain emotion to negative emotion.

3.2 Dempster-Shafer evidence theory

The D-S evidence theory was proposed by Dempster in 1967. Its student Shafer developed and organized it into a complete mathematical reasoning theory. D-S evidence theory can be regarded as a general extension of classical probability inference theory over finite fields. Its main features are to support the description of different levels of accuracy and to directly introduce the description of unknown uncertainty. D-S evidence theory can support probabilistic reasoning, diagnosis, risk analysis, decision support, etc., and has been applied in many application fields such as multi-sensor network and medical diagnosis.

The D-S evidence theory is based on the theory of non-empty finite fields Θ. Θ is called frame of discernment or FOD, which represent a finite number of system states {θ_1, θ_2, ⋯, θ_n}, the system state assumes that H_i is a subset of Θ, an element of FOD's power set P(Θ).

The objective of the D-S evidence theory is to estimate the state of the current system based only on observations of the state of the system E_1, E_2, ⋯, and E_n. These observations do not uniquely identify certain system states, but only represent the uncertainty of the system state. As the most basic concept of D-S evidence theory, we first need to define a probability function that supports a system state for an evidence. We call it basic probability assignment or BPA. D-S theory evidence uses the frame of discernment Θ to represent interesting proposition set. It defines a set function: m: P(Θ)→ [0, 1], which satisfies:

$$M(\Phi)=0 \quad (9)$$

$$\sum_{A\subseteq\Theta} m(A) = 1 \quad (10)$$

m is the basic credibility distribution of Θ. The mass m(A) of A, a given member of the power set, expresses the proportion of all relevant and available evidence that supports the claim that the actual state belongs to A but to no particular subset of A. The value of m(A) pertains only to the set A and makes no additional claims about any subsets of A, each of which have, by definition, their own mass.

The function that satisfies the following conditions is the Belief Function:

(1) $\mathrm{Bel}(A)=\sum_{B\subseteq A} m(B), \quad \forall A\in \Omega$ (11)

(2) $\mathrm{Bel}(\Phi)=0$ (12)

(3) $\mathrm{Bel}(\Theta)=1$ (13)

(4) $\mathrm{Bel}(\Theta)=1, \Omega \to [0,1]$ (14)

m_1 and m_2 are two independent evidences for frame of discernment, and Ω is a power set of Θ. In addition, B,C is two elements of a power set. Then the combined evidence obtained after combining these two evidences is

$$m(A) = \sum_{B\cap C=A} m_1(B)m_2(C)/(1-k) \quad (15)$$

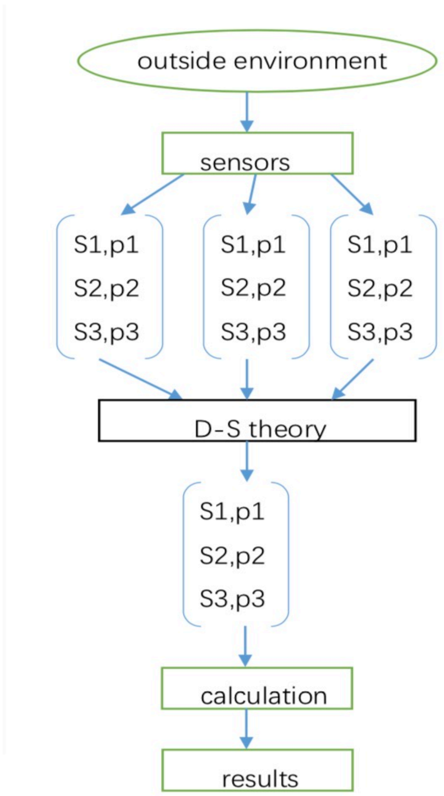
$$k=\sum_{B\cap C=\Phi} m_1(B)m_2(C)$$

This is the famous Dempster-Shafer evidence-combination formula, which provides rules for combining 2 evidences. For a combination of multiple evidences, Equation (15) can be used repeatedly to combine pairs of multiple evidences. For multiple basic probability assignment functions, there is a rule m(A)=$m_1 \oplus m_2 \oplus \cdots \oplus m_n$, The combined probability value after combination is

$$m(A)=\frac{\sum_{\cap A_i=A\ 1\le i\le n} \cap m(A_i)}{1-k} \qquad A \neq \Phi$$
$$k=\sum_{\cap A_i=A\ 1\le i\le n} \cap m(A_i) \qquad (16)$$

4 EMTIONAL MODEL

Firstly, according to emotional model and D-S evidence theory, construct an emotional robot model. We could regard the figure 2 as a emotional space for a robot. Combine emotional space model with d-s evidence theory. Someone's emotion will always stay at a certain point on this space and it will probably change to another point. But this process not only relate to its original position but also simulation from outer space. For example, it will be influenced by sounds, weather, pictures, smell and so on. This model will capture outer simulation by sensor, and then D-S theory will be applied into it for combining outside emotion information, which finally will promote the transfer of emotion status and reach a new state of emotion.



S = the strength of simulation from outside environment

P = basic credibility

Fig. 3. Emotional Information Processing Flow

To begin with, you need to set an initial state, that is, set the node at the initial moment of feeling. Then, what kind of state will the emotion reach in the next moment? This is a probabilistic behavior. We assume that the initial emotion is in a state of calm, that is, the origin of the emotional space. While without the outside interference, the state of the emotions tends to be calm. However, when it is subjected to external stimuli, the emotional state will shift. For the origin (0, 0, 0) in the emotional space, whether it represents calm or other unknown emotional states, as long as it is subjected to external stimuli, it will shift in the direction of maximum probability.

As forexternal stimuli, we define them into three categories: visual, phonetic, and other information. From these three aspects, we extract the environmental incentives. The incentive intensity I takes values within the [0,1] interval and is divided into three levels according to the size of the incentive: mild incentive, moderate incentive, and major incentive. The thresholds β1, β2 are selected. When 0 ≤ I ≤ β1, it is mild excitation; when β1 ≤ I ≤ β2, it is moderate excitation; when β1 ≤ I ≤ 1, it is significant excitation. According to human experience, we assume that: when mildly motivated, the happy state transitions to the direction of happiness intensity with the greatest probability; when moderately motivated, the happy state remains; and when it is heavily motivated, happy state increases in the direction of increasing the intensity of happiness. Similarly, when mildly motivated, the fear state shifts to the greatest reduction in fear. In the medium incentive, the fear state remains the most probable; when severely motivated, the fear state increases in the direction of fear. Also, when mildly motivated, the anger state has the greatest probability of shifting to anger-reducing direction; when moderately motivated, the anger state remains the most probable; and when heavily motivated, the anger state has the greatest probability of transition to anger-increasing direction.

Secondly, the application of D-S evidence theory integrates visual and speech, visual and other, speech and other information factors, respectively. Consider T_l as mild incentives, T_m as moderate incentives, T_h as significant incentives, and U as uncertainty. Plus, consider m_v (A) as a visually determined basic probability assignment, m_s (B) as a basic probability assignment determined by speech, and m_t (C) as a basic probability assignment determined by other information. The basic probability assignments determined by these three kinds of emotional incentive information are shown in Table 2.

TABLE 2
BASIC PROBABILITY ASSIGNMENT DETERMINED BY THREE EMOTIONAL STIMULUS INFORMATION

	Tl	Tm	Th	U
$m_v(A)$	0.40	0.20	0.30	0.10
$m_s(B)$	0.30	0.40	0.20	0.10
$m_t(C)$	0.20	0.30	0.40	0.10

According to the Dempster combination formula, the basic probabilistic assignments of the visual and speech excitation conditions are combined. The combination of m_v (A) and m_s (B) is shown in Table 3, where Φ denotes an empty set. The inconsistent factor K for m_v (A) and m_s (B) is

$$K = 0.06+0.09+0.16+0.12+0.08+0.04=0.55 \tag{17}$$

Then, basic probabilities combined with visual and speech emotion information

$$m_{v\times s}(T_i) = (0.12 + 0.03 + 0.04)/(1 - k)=0.42 \tag{18}$$

$$m_{v\times s}(T_i) = (0.08 + 0.02 + 0.04)/(1 - k)=0.31 \tag{19}$$

$$m_{v\times s}(T_i) = (0.08 + 0.02 + 0.03)/(1 - k)=0.24 \tag{20}$$

$$m_{v\times s}(U) = 0.01/(1 - k)=0.02 \tag{21}$$

TABLE 3
COMBINATION OF M_V(A) AND M_S(B)

$m_s(B)$ \ $m_v(A)$	Tl(0.40)	Tm(0.20)	Th(0.30)	U(0.10)
Tl(0.30)	Tl(0.12)	Φ(0.06)	Φ(0.09)	Tl(0.03)
Tm(0.40)	Φ(0.16)	Tm(0.08)	Φ(0.12)	Tm(0.04)
Th(0.20)	Φ(0.08)	Φ(0.04)	Th(0.06)	Th(0.02)
U(0.10)	Tl(0.04)	Tm(0.02)	Th(0.03)	U(0.01)

In order to make robot more like a human, we should consider the influence of personality on emotional transfer. This project set the robot also had its own character, it was able to change the character itself by parameters. When the robot had different characters, the same external stimulus would produce different degrees of emotional stimulation. This paper set up a personality matrix and a sensitive factor for each stimulus level in each emotional dimension. Dividing the stimuli into n levels, m emotions were able to form a m dimensional emotional space. Then there was n×m dimensional personality matrix as follows

$$R=\begin{bmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{m1} & \cdots & R_{mn} \end{bmatrix} \tag{22}$$

For $R_{ij}(i \in [1, 2, 3, \cdots, m], j \in [1,2,3, \cdots, n])$, it means the i-th emotion is sensitive to the intensity of j-th stimuli and satisfies

$$\sum_{j=1}^{i} R_{ij} = 1, i \in [1, 2, 3, \cdots, n] \tag{23}$$

5 CONCLUSION

In order to verify the feasibility of this sentiment model, some researchers have been able to perform experiments on robotic systems. First of all, set the emotional space is three-dimensional: anger, fear and joy. There are three different incentives: mild incentives, medium incentives and major incentives. That is, n=3, m=3, when the initial state of the emotion is set to (1, 0.5, 0), the character matrix given is

$$R=\begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}$$

The basic emotion probability obtained is assigned

$$\begin{bmatrix} 0.042 & 0.093 & 0.144 \\ 0.093 & 0.124 & 0.072 \\ 0.21 & 0.093 & 0.048 \end{bmatrix}$$

Then the state of fear received the major stimulus T_h, then the fear remains very fearful. In the state of anger, it is moderately stimulated by T_m. The anger state is maintained by a bit of anger. In the happy state, it is mediated by mild incentive T_l. The happy state is kept quiet, so the emotional state is finally changed to (1, 0.5, 0).

By adjusting the personality matrix, it is possible to make robotic emotion changes conform to the emotional changes of a person with a certain personality. Therefore, after the character matrix was added, the influence of human factors on emotion was taken into account, making the model more consistent with human natural emotion transfer rules. Of course, to make the model conform to the emotional rules of real people, we must also conduct large-scale training on the model, and constantly correct the values in the character matrix.

REFERENCES

[1] Eyben F, Wollmer M, Graves A, Schuller B, Douglas-Cowie E, Cowie R. On-Line emotion recognition in a 3-D activation-valence- time continuum using acoustic and linguistic cues. Journal on Multimodal User Interfaces, 2010,3(1-2):7-19 .

[2] Dempster A P. Upper and lower probabilities induced by a muhivalued mapping [J]. Annals of Mathematical Statistics, 1967, 38(4): 325-339.

[3] Zadeh LA. Review of shafers A mathematical theory of evidence [J]. AI Magazine, 1984, 5(3):8l—83.

[4] Misral R B, Srinivaasan S, Mital D P. The Use ofWeb Services Technology in the Design of Complex Software Interfaces: An Educational Perspective [J]. Information Technology Journal, 2006, 5(6): 1127-1130