Affective Modeling and Recognition of Learning Emotion: Application to E-learning

Yanwen Wu, Tingting Wang, Xiaonian Chu
Department of Information & Technology, Central China Normal University, Wuhan, 430079, China
ttw@mails.ccnu.edu.cn

***Abstract*—Affective computing has been the focus of artificial intelligence for several years. The amalgamation of affective computing and facial expression recognition technique has lead to the possibility of harmonious human-computer interaction in E-learning. But the fact of E-learning is that emotional absences between computer and E-learner are serious. This paper researches into the learning emotions that E-learner may present. We define three basic learning emotions: absorbed, neuter and fatigue and discuss the features of every defined learning emotion. Layered approach for modeling emotions and facial expression recognition technique are adopted to describe and recognize the defined learning emotions. We also make a demonstrability research, as results indicate, characteristic parameters are exactly identified and face expressions of E-learner are accurately recognized. After E-learner's emotion state is recognized and confirmed, in the future research, it is possible to take corresponding emotion incentive pleasures of E-learners' given emotion state to decrease the emotional absences in E-learning.**

***Index Terms*—learning emotion, affective modeling, facial expression recognition, E-learning, emotion incentive**

I. INTRODUCTION

As one of the inseparable parts of human intelligence, affect and emotion plays a great role in many human activities such as apperception, consequence, decision-making and sociality. Affective computing is computing that relates to, arises from or deliberately influences emotions [1]. Its final aim is to build an intelligent system that can apperceive, recognize and understand human emotion [2]. As the first step to build an intelligent system with the affective ability similarly with human, Affective modeling is to build an exact computing model to simulate human emotion.

As a possible approach to realize information education and education modernization, E-learning has numerous advantages in resource-share, interaction and multimedia technique than traditional educational mode. However, the emotional absences in E-learning have aroused more and more attention. There are some shortages and limitation in educational idea, organization and interaction in E-learning. Especially emotional interaction between teachers and students is few as a result of the separation in space-time. These absences badly influence E-learner's learning effect and mental health when E-learner faces the bland screen with no pleasures and incentive for a long time. The emotional absences in E-learning make the harmonious Human-Computer Interaction (HCI) difficult [3]. This paper takes cognitive psychology and affective computing as academic foundation, uses layered approach for modeling personality, moods and emotions Kshirsagar S. provided [4], and defines three basic learning emotions of E-learner. The characteristics of three learning emotions are researched and facial expression recognition (FER) technique is used to recognize E-learner's facial expressions. After E-learner's emotion state is estimated and confirmed, some corresponding emotion incentive pleasures to their given emotion state would be taken to decrease the emotional absences in E-learning.

The rest of this paper is organized as follows. In Section 2, related work in affective computing and our research contents are presented. We research into the emotion factor in E-learning in Section 3. In Section 4 we research the expressions of E-learner and provide a simple and effective approach to expression modeling. We make an experimental research to recognize E-learner's facial expressions and some experimental results are provided in Section 5. Finally, we conclude with future directions of research.

II. RELATED WORK

A. Psychologic Foundation of Human Emotion

Emotion is a production of human history development. It is in different levels along with the long process of history development. When emotion is used as a professional term to describe different psychological phenomenon by different people, it also means different in contents.

Several emotion models are proposed to explain this psychological phenomenon, and a widely accepted model is "emotion- cognition-activity" response model. In this opinion, emotion is developed along with the life process. It includes three aspects: emotional experience, corresponding activities of brain and neural system and facial expressions. It is believed that in the process of emotional activation, people and environment are interactional, and the inner process of individual cognition plays a significant role. It is emphasized that emotion plays a motivational function in personality integration. It is believed that personality is a complex organization, including six subsystems: aesthesia, cognition, activity, motivation, emotion and inner balance. Emotion is the core part of this organization [5]. The cognitive theories of emotion inherit the opinion that

doi:10.4304/jsw.4.8.859-866

emotion contains biologic components and evolutional value. It is also emphasized that social cultural environment, individual experience and personality structure play important functions on emotion. In summary, emotion is an experience based on whether objective things meet one's demand and a reflection of relations between objective things and one's demand.

B Affective Computing and Affective Modeling

Affective computing is a broad, multifaceted subject in artificial intelligence. The conception of Affective computing was firstly proposed by R.Picard in 1995 and was described in her monograph "Affective Computing". It was defined that Affective computing is computing that relates to, arises from or deliberately influences emotions [1]. This definition received widely acceptation and application. Affective computing is a focus in information science, cognition science and psychology. The main research contents of affective computing include emotion mechanism, affective signal capturing, affective pattern recognition, emotion modeling and comprehending, emotion synthesizing and expressing, emotion transferring and alternating, and so on. The final aim of affective computing is to build an intelligent computing system which can sense, recognize and understand human emotions. It can response intelligently, sensitively and friendly like human. However, making computing system intelligent is a topic full of change because emotion is changeable and influenced by many factors.

As the first step to build an intelligent system with the affective ability similarly with human, the target of affective modeling is to build an exact computing model in mathematics to describe human emotions. At present, many models were proposed. We divide these models into three categories: cognitive models, models based on probability, and other models. These models have solved different problems in different areas but have different disadvantages respectively. Layered approach to affective modeling proposed by Kshirsagar S. in 2002 is the first model that has built the relation between personality, mood and emotion. This paper understands and introduces the layered model in another way in the following, in order to build a relation between E-learner's emotions and expressions.

C Layered Approach to Affective Modeling

Many affective models have been proposed and a well accepted model is cognitive model. There are some other models such as probability-based models and so on. But we would like to introduce layered approach to affective modeling. Layered approach to affective modeling proposed by Kshirsagar S. in 2002 builds a relation between personality, mood and emotion initially. It is well applied into expression synthesis of virtual human and receives widely attention.

We understand and describe this model in another way in order to explain the relation between E-learners' emotions and expressions. E-learners' personality influences mood, mood influences emotion change and facial expression. It is shown as follows Fig. 1.

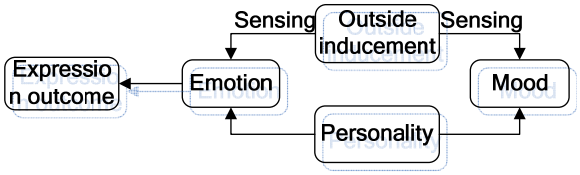


Figure 1. Layered affective model of E-learner

The model is divided into four layers, personality, mood, emotion, expression. The mutual influence between each other is introduced as follows.

1) Personality is on the highest layer, and is considered changeless during the whole life. It is introverted or extroverted and influences the change of moods.
2) Mood is on the second high layer, and is influenced by personality and emotional state.
3) As the third layer, emotion is influenced by personality and mood and influences the facial expressions.
4) The external stimulation is sensed and evaluated into emotional variable by E-learner. Then emotional variable acts on mood space and emotion space under the influence of personality. Moods and emotion states are transferred in the meantime and facial expressions are finally presented.

D Facial Expression Recognition

With the increased interests in human-computer interface, affective computing and facial expression recognition (FER) become a well focus area. If computers could understand and convey emotion just like human to each other, the harmonious human-computer interface is possible. Facial expressions carry rich information of human inner emotion and affect. FER is very useful in many applications such as computer vision, computer graphics, pattern recognition and human-computer interface, etc. We divide an FER system into three steps: detection and location of face, facial feature extraction, and facial expression classification.

Many systems for face detection have been proposed and a well accepted model is neural networks. There are some other systems such as probability -based approach and etc [6] [7]. We have realized our face detection system using AdaBoost algorithm based on skin color. Firstly, skin color segmentation and morphological operators were used to detect skin regions, which were filtered in terms of the statistical characteristics of human face to form candidate face regions. Finally, these regions were scanned by cascade classifier based on AdaBoost for more accurate face detection. Experimental results show that the proposed system results in better

performance than the other methods, in terms of correct detection rate and capacity of coping with the problems of expression, lighting, conditions, and orientation. The aim of facial feature extraction is to find the most appropriate representation of the face images for recognition. We divide all the systems into two approaches: holistic template-matching systems and geometric feature-based systems [8]. There is still a small very successful amount of work on FER. Essa and Pentland build a dynamic parametric model by tracking facial motion [9]. Cottrell and Metcalfeuse holistic representations based on principal components, extracted by feed forward networks [10].

E OurResearch

This paper uses layered approach to modeling personality, moods and emotions of E-learner, and defines three universal learning emotions of E-learner. The characteristics of three learning emotions are researched. Then facial expression recognition technique is used to recognize E-learner's facial expressions. After E-learner's emotion state is estimated and confirmed, some corresponding emotion incentive pleasures to their given emotion state would be taken to decrease the emotional absences in E-learning. The system overview is shown as follows Fig. 2.

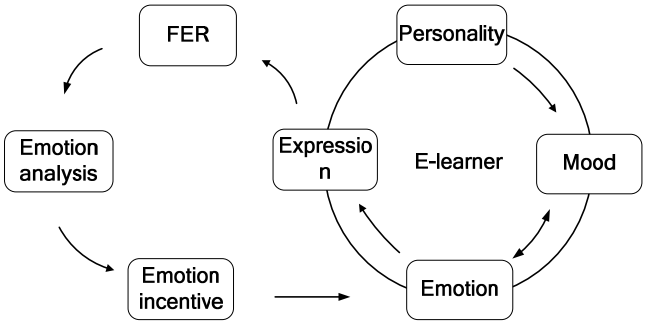


Figure 2. System overview

1) In the emotional model, personality is on the highest layer and considered changeless after birth. Mood is on the second high layer and emotion is on the third layer. They influence each other and are influenced by external stimulation in the meantime. The external stimulation is sensed and evaluated into emotional variable by E-learner. Then emotional variable influences mood space and emotion space. Moods and emotion states are transferred and facial expressions are finally presented.
2) After facial expressions are finally exported, basing the expression model proposed in Section 4, facial expression recognition technique is used to recognize E-learner's facial expressions. Of course, an FER system needs to solve the following problems: detection and location of face, facial feature extraction, and facial expression classification. After the confirmed given expression is analyzed, then some corresponding emotion incentive pleasures are adopted to help the E-learner to maintain good learning emotion.

III. RESEARCH ON AFFECTIVE FACTORS IN E-LEARNING

A Relation betweenEmotion State andLearning Efficiency

Psychology research has shown that emotional feelings can promote or hinder the efficiency of working memory, reasoning operation and solving problems. The long-term psychological background or the temporary psychological status posed by emotion experience can influence and regulate the cognitive process such as perception, memory and thinking [11].

We have taken an English word memory experiment. Testees are 30 students in grade three of high school. Test contents are 40 English words of CET-4 random collected, time is 30 minutes. High school students in grade three generally were very concerned about their academic achievements. Usually, the more progress the have in scores and place, the more pleasant mood they have; on the contrary, the more retrogression the have in scores and place, the more low mood they have. Before the test, these students were informed the last semester final exam scores and it made a great influence on their emotions. Their places change from raising 20 to declining 23 and suppose that their moods change from most pleasant to most low. We numbered testees from1 to 30 according to their emotions in descending order. Results of the test are shown as follows Fig. 3.

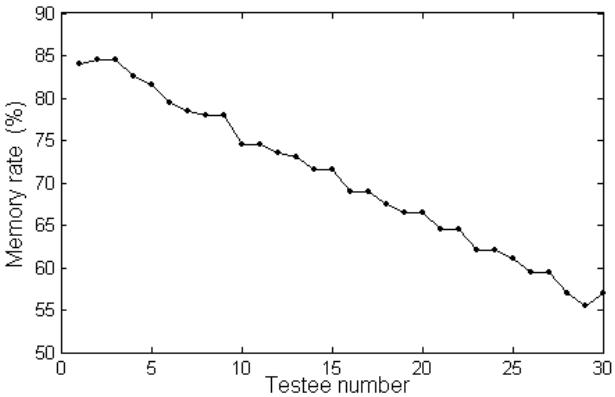


Figure 3. Words memory results

As results indicate, students with pleasant mood remember more words than those with low mood in the same time; the greater difference in mood, the greater difference in the memory efficiency. There is a small number of anomaly, we assume that in the similar emotion status, students' English background is on different levels. Test results further testify the close relationship between emotional status and the efficiency of learning memory.

B Investigation in Emotional Absences of E-learning

The network platform presents a new and wider learning world, and provides learning rights and opportunity to all the learners. E-learning based on multimedia and computer technology is much better than traditional media education at resource sharing and strong interaction .It is a possible educational mode can realize education optimization, but the problem of emotional absence exposed has drawn people's more attention.

We make a comparative research on the phenomenon of the decrease of learning efficiency caused by the emotional absence. Testees are 50 sophomores majored in electronic information engineering, which is named Group 1; 100 registered distance education learners from online school, which is named Group 2. Testee contents are whether students' problems receive satisfactory solution in "Analog electronics" course. "Analog electronics" is a provincial excellent course constructed by CCNU. It has two kinds of teaching methods at present: real-time education in the classroom with the teacher for the undergraduate students at the school and teaching video for the registered distance education learners. Results of the investigation are shown as follows Table I.

TABLE I. PROBLEM-SOLVING RESULTS

	All the problems be solved	Most of the problems be solved	A few problems be solved	Seldom problems be solved
Group 1	12	25	11	2
Rate	24%	50%	22%	4%
Group 2	12	34	40	14
Rate	12%	34%	40%	14%

As table 1 indicates, there are 74% students in Group 1 can get most problems solved, including all the problems solved, but only 46% in Group 2. Excluding the different learning level between two groups, the main reason for the gap is adequate teacher-student motional interaction in the real classroom, but the deficiency of teacher-student motional interaction in E-learning. Therefore, we come to the conclusion that the emotional absences caused the separation of teachers and students which has a bad impact on the learning quality of E-learners.

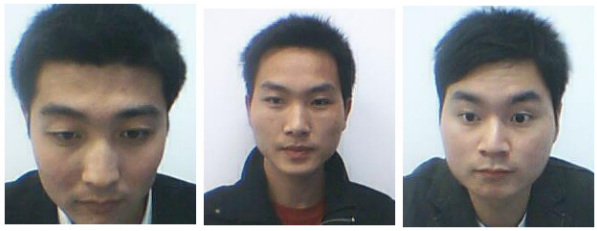
IV. ACIAL EXPRESSION MODELING OF LEARNING EMOTION

A Theoretic Foundation

1) Facial expressions have universal meaning. Facial expressions are an outside behavior of conveying specific information and have whole mankind's intercommunity. It is well known that babies don't need to study but can have a similar facial expression model to adults. And it is also well known that people have similar facial expression model during different races. So it can be concluded that facial expressions have universal meaning among human.
2) Relation between expression and emotion. Facial expression isn't isolated; there are much affiliation between facial expression and emotion. The adaptive value and social function of emotion in the society development are often manifested through expression to a large extent. Facial expressions, voice expressions and body posture constitute the emotional representation; then emotion representation and emotion experience constitute the emotional psychology. Although, the composing of emotional psychology is complicated, there are many factors that influence human emotion. Because of the different social cultural background and other factors, everybody's emotional representation is different. And facial expressions represented outside are usually not simple, but complicated and mixed. But contacts between expression and emotion are still can be found after careful observation. So it is inseparable to each other when emotion and expression is discussed.
3) Emotion can be accurately identified by facial expression recognition. After facial expression is recognized, then it is important to decide the corresponding emotional and affective state. According to experimental results of Facial Action Coding System (FACS) proposed by Ekman and Friesen, if taking no account of other inducements and false signals of testee, the recognition rate from a specific facial expression to a single emotional state is 88% [12]. It can be concluded that emotion can be accurately identified by facial expression recognition.

B Classification and Analysis to Learning Expression

Research on facial expression recognition by Ekman is creative and makes great and significant development in theory and representation, making a detailed and valuable foundation. He defined six kinds of most basic facial expressions: surprise, fear, disgust, anger, happiness and sadness. His definition received widely acceptation and application [13] [14]. It is universal about these six kinds of most basic facial expressions in different races and under different background. Considering the characteristics of E-learning, fortunately, there is no need to research all six kinds of expressions. It is enough to research expressions related to learning emotion. Considering the universality and convenience of modeling, this paper defines three basic learning expressions: absorbed, neuter and fatigue in E-learning. Three kinds of basic learning expressions are collected and shown as follows Fig. 4. We analyze and describe the main characteristics of each basic learning expression in Table II.



(a)

(b)

(c)

Figure 4. Sample of three basic learning expressions. (a) Absorbed expressions. (b) Neuter expressions. (c) Fatigue expressions.

TABLE II. ANALYSIS AND DESCRIPTION OF THREE BASIC LEARNING EXPRESSIONS FEATURE

Expressions	Eyebrow and forehead	Eyes	Mouth
Absorbed	The eyebrow may raises; furrow may exist in the forehead.	Space between upper and lower eyelid may raise; wink frequency may decrease and act of wink is fast.	Mouth is natural closed, and is with no change.
Neuter	Eyebrow and forehead are natural.	Space between upper and lower eyelid is natural; wink frequency is common.	Mouth is natural closed, and is with no change.
Fatigue	When gaping, eyebrow may raises; furrow may exist in the forehead.	Space between upper and lower eyelid may decrease; wink frequency may raise and act of wink is slow.	When gaping, mouth would be opened and the space between upper and lower lip may raise.

C Learning Expression Modeling

Facial Action Coding System (FACS) proposed by Ekman and Friesen adopt 44 independent AU to describe facial activities. This system has defined six kinds of basic expressions: surprise, fear, disgust, anger, happiness and sadness, and 33 kinds of expression tendency [12]. FACIAL ACTION CODING SYSTEM (FACS) proposed by Ekman and Friesen make a good foundation for research on facial expression modeling. Because facial expression and its change is mixed and complicated, we need to simplify the modeling. Three single basic learning expressions model is discussed in following.

We define vector X= (xl，x2，x3，a), xl equals to the highness of eye, x2 equals to the highness of mouth, x3 equals to the width of mouth, a equals to the highness from brow to chin. They are shown as follows Fig. 5.

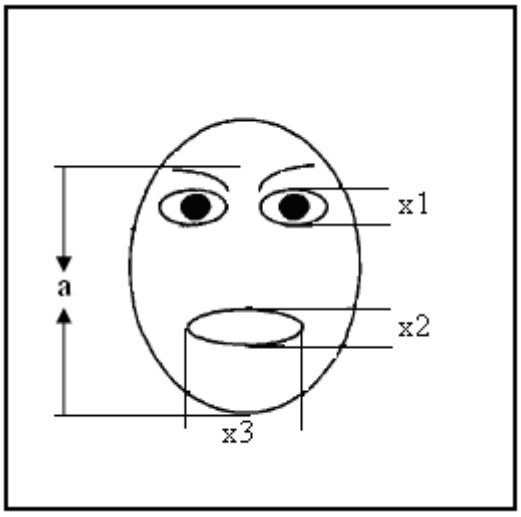


Figure 5. Characteristic parameters of facial expression

It is discovered from collected E-learners' images that the highness from brow to chin is constant under different absorbed, neuter and fatigue learning expressions. So parameter a can be disposed as constant parameter. Vector Y is defined to modeling facial

expression, Y= (yl， y2， y3).

$$yl= xl/a,\ y2= x2/a,\ y3= x3/a \quad (1)$$

Firstly facial expression characteristic vector is defined as follows. Then Mahalanobis distance is used as determinant standard.

$$Y_i=(y_{i1},y_{i2},y_{i3}),\ i=1(\text{absorbed}),\ 2(\text{tired}),\ 3(\text{neuter}) \quad (2)$$

The difference expression vector about three basic learning expressions and neuter learning expression is defined as follows.

$$DifY_i=(y_{i1},y_{i2},y_{i3})-(y_{31},y_{32},y_{33}) \quad (3)$$

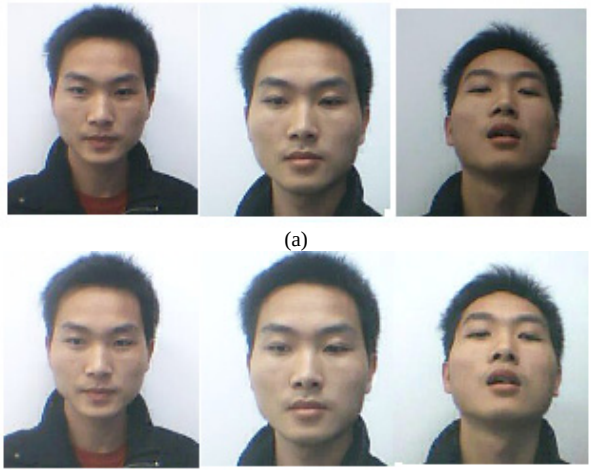
Mahalanobis distance is computed in the following.

$$d^2(Y)=(Y-\mu)^t\Sigma^{-1}(Y-\mu) \quad (4)$$

In above formula, μ equals to the mean of Y, Σ equals to the variance of Y, and Y is difference expression vector DifY .

V. EXPERIMENTAL RESEARCH: LEARNING EXPRESSION RECOGNITION

Experiments were performed to verify the effectiveness of the approach we proposed. The database of E-learners' learning expressions is colleted when E-learner is studying by E-learning. This system has constructed the database for facial expression recognition from video camera; these images contain several kinds of facial expression. Firstly, illumination compensation pleasures are taken to decrease the influence of asymmetrical illumination condition. As is shown in Fig. 6, the influence of asymmetrical illumination is eliminated.

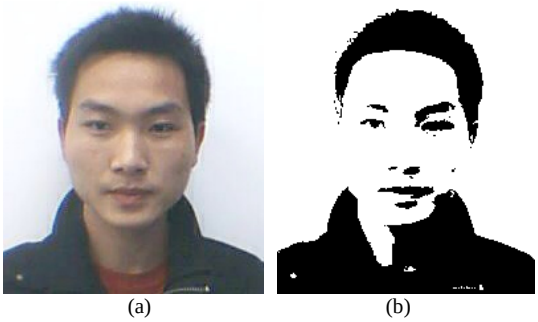


(a)

(b)

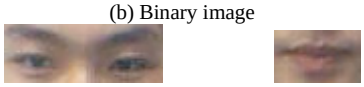
Figure 6. Illumination compensation. (a) Primitive collected images. (b) Illumination compensated images

After the converting from color image to gray image according to the Gauss model, proper threshold is chosen along with gray image is converted to binary image. After the pixel-based classification and proper threshold is chose, a binary image is produced consequently. Then integral calculus projection in the horizon direction is adopted, considering the physiological structure characteristic of eyes, the eyes area is obtained. Considering the physiological structure characteristic of mouth, integral calculus projection in the vertical direction is adopted to accurately confirm mouth area. The process above is shown in Fig. 7 and Fig. 8.



(a) (b)

Figure 7. Image processing. (a) Illumination compensated image. (b) Binary image



(a) (b)

Figure 8. Segmentation results. (a) Eyes area segmentation result. (b) Mouth area segmentation result

Because we define only three kinds of basic learning expressions: absorbed, neuter and fatigue in E-learning, the kinds of facial expressions are less, facial expression classification approach based on regular is adopted in this paper. The classification tree proposed in this paper is shown as follows Fig. 9.

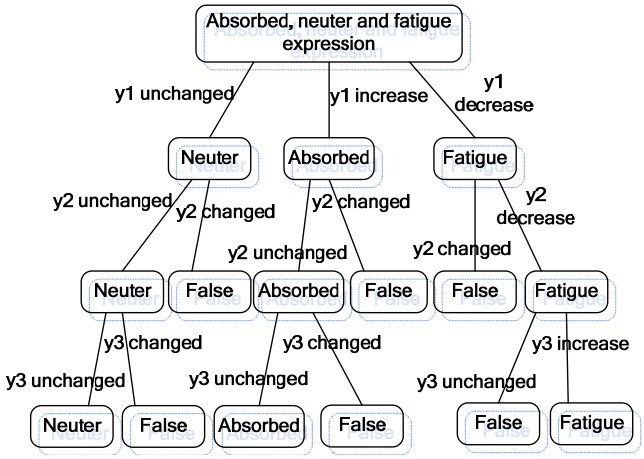


Figure 9. Expression classification tree basing on regular

One of the main purposes of this paper is to detect whether E-learner is fatigue emotional state. In the process of E-learning, the change of eyes and mouth is the most notable characteristic. For absorbed expression, the change of y1 is obvious; for fatigue expression, the change of y2 and y3 is obvious. According to the change of y1, three expression subsets, absorbed subset, neuter subset, and fatigue subset are produced. On the foundation of these three subsets, the change of y2 and y3 are considered. If the result parameters don't match the model parameters system set, system returns false message and go into next recognition. Accordingly, the

recognition of three kinds of basic learning expressions is finally accomplished.

In the experiments, nine expression images of three E-learners are randomly chose; they contain all three kinds of expression images: absorbed, neuter and fatigue expressions and the number of each kind is nine. Experimental results are shown as follows Table III and Table IV.

TABLE III. LEARNING EXPRESSION RECOGNITION RESULTS

	Absorbed	Neuter	Fatigue
Absorbed	7	2	0
Neuter	3	5	1
Fatigue	0	1	8

TABLE IV. LEARNING EXPRESSION RECOGNITION RATE

Expression	Image Num	Correct	Recognition Rate
Absorbed	9	7	77.8%
Neuter	9	5	55.6%
Fatigue	9	8	88.9%
Average recognition rate: 74.1%			

As results indicate, the learning expressions of E-learner can be accurately recognized, and the average recognition rate is 74.1%. For the facial expression characteristics of absorbed expression and fatigue expression are more obvious than neuter expression, so it can be found that the recognition rate of absorbed expression and fatigue expression is higher than neuter expression. After E-learner's expression state is estimated and confirmed, some corresponding emotion incentive pleasures to their given emotion state would be taken to decrease the emotional absences in E-learning.

VI. CONCLUSION

The emotional state of E-learner receives less concern but the emotional state is of great relation to learning efficiency of E-learner. Facing the truth that there are serious emotional absences in E-learning, we propose an approach to learning expression modeling of E-learner, using affective computing theory and facial expression recognition technique. We define three basic learning emotions and research the characteristics of three learning emotions. This paper makes a foundation to taking emotion incentive pleasures to decrease the emotional absences in E-learning. Our experimental results demonstrate the effectiveness of our work.

There are several difficulties in facial expression recognition (FER) due to the variation of facial expression even for the same individual. Because the common expressions of E-learner are usually complex and mixed, and dividing expressions into only several basic kinds may decrease the universal of the system. On the other hand, it is still not very clear what factors contribute to the final expression representation of human. When E-learners' eyes is closed but not be fatigue, it may make a false recognition, so only one image being recognized is not enough. In future work, we plan to extend more universal kinds of learning emotions of E-learner and take emotion incentive pleasures to E-learner. We would also like to research the relation between learning emotions and expressions deeply and try to build an emotion state transformation model.

ACKNOWLEDGMENT

This paper is supported by Hubei Province College Teaching Research Project: Research and Design of Undergraduate's Graduation Design Dissertation Quality Control (20050185).

REFERENCES

[1] Picard, R. W. Affective Computing. MIT Press, Cambridge. 1997.
[2] M.J. Blackand Y. Yacoob. "Recognizing facial expressions in image sequences using local parameterized models of image motion," *Int. J. Conmput. Vis.*, vol. 25, no. 1, pp. 23-48, 1997.
[3] Liu Jingfu, Zhong Zhixian. "Lack of sensation in on-line teaching and countermeasure," *China Distance Education*, June 2001.
[4] Kshirsagar S. "A Multilayer Personality Model," *In: SMARTGRAPH'02: Proceedings of the 2nd International Symposium on Smart Graphics*, New York, pp. 107-115, 2002.
[5] Picard, R. W. "Towards Computers that Recognize and Respond to User Emotions," *IBM System Journal*, vol 39, pp.705-719, 2001.
[6] H. Rowley, S. Baluja, and T. Kanade. "Human face detection in visual scenes," *Technical Report CMUCS-95-158R*, School of Computer Science, Carnegie Mellon University, November 1995.
[7] K.-K. Sung and T. Poggio. "Example-based learning for view-based human face detection," *TechnicalReport A.I. Memo 1521*, MIT, pp. 112, December 1994.
[8] R. Chellappa, C.Wilson, and S. Sirohey. "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, pp. 705–740, May 1995.
[9] I. Essa and A. Pentland. "Coding, analysis, interpretation, and recognition of facial expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 757–763, July 1997.
[10] G. Cottrell and J. Metcalfe. "Face, gender and emotion recognition using holons," *In D. Touretzky, editor, Advances in Neural Information Processing Systems 3*, Morgan and Kaufman, San Mateo, pp. 564–571, 1991.
[11] R. W. Picard and J. Healey, "Affective Wearables," *Personal Technologies*, vol. 1, no. 4, pp. 231-240, 1997.
[12] Paul Ekman, Wallace V Friesen. "Facial Action Coding System," *California: Consulting Psychologists Press Inc*, 1978.
[13] Ekman.P, Friecen, W. "Unmasking the face. A guide to recognizing emotions from facial expressions," *Consulting Psychologists Press*, Palo Alto, CA, 1975.
[14] Ekman.P. "An argument for basic emotions," *Cognition and Emotion*, 6, pp. 169-200. 1992

Yanwen Wu, Female, professor, PhD degree, Master degree candidate's tutor. Engaged in the Artificial Intelligence and its application for a long time, involved the intellectual calculating, the network education, the community education, intellectual tutoring system, web mining, etc. She has already released three specialty books and lots of dissertations in home and abroad.

Tingting Wang, Male, graduate student of Information Technology Department in Huazhong Normal University in Wuhan, engaged in Artificial Intelligence and its application to education.

Xiaonian Chu, Female, graduate student of Information Technology Department in Huazhong Normal University in Wuhan, engaged in computer science and educational technology.