International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.7, No.5 (2014), pp.349-360
http://dx.doi.org/10.14257/ijsip.2014.7.5.30

Puzzlement Detection from Facial Expression Using Active Appearance Models and Support Vector Machines

Jinwei Wang[1,2], Xirong Ma[2], Jizhou Sun[1], Ziping Zhao[2] and Yuanping Zhu[2]

[1] *School of Computer Science and Technology, Tianjin University, Tianjin, China.*
[2] *College of Computer and Information Engineering, Tianjin Normal University, Tianjin, China*
wangjinwei@tju.edu.cn

Abstract

Affective state detection, as an emerging field of artificial intelligence, is the key to designing effective natural human-computer interaction, especially for e-learning. It will be helpful to make the computer understand learners' perceptions and provide appropriate guidance, just like teachers in traditional face-to-face classroom learning. Puzzlement is the most frequent non-neutral affective state in learning, and it is usually a sign that learners need more information and guidance. In this paper, we explore a machine learning approach for puzzlement detection from natural facial expression. We use active appearance models (AAMs) to decouple shape and appearance parameters from the face video sequences. Support vector machines (SVMs) are utilized to classify puzzlement and non-puzzlement with several features derived from AAMs. Using a 10-fold cross validation, we achieve the highest recognition rate of 98.9%. Experimental results indicate the feasibility of automatic frame-level puzzlement detection.

Keywords: *Human-computer interaction, Emotion recognition, Active appearance models, Support vector machines, Facial expression, E-learning*

1. Introduction

Human-computer interaction is an important part of success for developing interactive systems. It influences the efficiency with which they are used and provides a means of communication between the user and the virtual environment [1, 2].

As a growing trend of educational interactive systems, e-learning has numerous advantages in resource-share, interaction and multimedia technique compared with the traditional classroom learning. However, the absence of emotional communication in human-computer interaction makes e-learning difficult to provide learners appropriate guidance as teachers do in traditional face to face learning. Psychological research has shown that affective states are central to learning [3, 4]. Learning-centered affective states are inextricably linked with the cognitive of learning. Positive affective states may facilitate learning, whereas the negatives may block learning. Therefore, detecting and understanding the learner affective states to promote positive states and eliminate negatives has become an important research problem and the key to designing effective natural human-computer interaction of e-learning systems.

Recent researches show that the most frequent non-neutral affective state of learners is puzzlement [5, 6], which is the negative affective state. It is usually a sign that learners are blocked and need more information and guidance. Therefore, it is very desirable to design an approach that can automatically recognize puzzlement to give learners timely help in e-learning systems.

ISSN: 2005-4254 IJSIP

To identify learners' affective states, researchers often investigate nonverbal behavior, *i.e.*, the eye gaze and head pose [7]. Even some physiological signals, such as heart rate, skin conductance and EEG brainwaves are used to detect learners' affective states [8]. Actually, the most powerful, natural and direct nonverbal channel used by humans to communicate and understand each other's affective state is facial expression [9], which has been intensely studied for decades. However, most of these studies [10-14] focused on only six universal facial expressions triggered by basic emotions. The computational models that can recognize subtle facial expressions (*i.e.,* fatigue [15, 16]) are limited. Facial expressions triggered by learning-centered affective states, i.e. puzzlement, are also subtle. There is still a need to more fully explore facial expression features to recognize such affective state [3].

In the past several years, significant progress has been made in machine learning to automatically recognize facial expressions related to emotion [17, 18]. In machine learning, the choice of features is known to influence recognition performance [19]. Both texture and shape features have been investigated. Examples of texture based representations are raw pixels and Gabor filters [20, 21]. Active shape models [22] are typical shape-based representations. The active appearance models (AAMs) [23], first proposed by Cootes et al. [24], and is generative and parametric models. Given a pre-defined linear shape model with linear texture variation, AAMs have been demonstrated to be one of the most powerful model-based object detecting and tracking algorithms. They decouple shape and texture and perform well in the task of expression recognition, especially non-rigid face motion.

In this paper, we propose a novel machine learning approach that can detect puzzlement by analyzing shape and texture information of the face. This approach is based on AAMs which are used to extract shape and texture features from the face video frames. Support vector machines (SVMs) are used to classify puzzlement and non-puzzlement with several features derived from AAMs. Through training, we obtain a series of frame-level puzzlement detection classifiers.

The rest of the paper is organized as follows. Section 2 describes the dataset we use. Section 3 briefly describes AAMs. Section 4 presents features which are extracted from AAMs. Section 5 introduces the basic principle of SVMs. Section 6 describes the experiments and results. Section 7 concludes the paper and presents some future work.

2. Dataset

We use the Video Database of Moving Faces and People [25] to test the performance of the proposed approach. The total size of this database is very large (160-gigabyte). It contains various still images and video clips of the facial speech, facial expression, parallel gait and conversation from a large number of individuals (284 subjects, Males=76, Females=208), who have different color of skin and different ages, taken in various contexts.

For the purpose of our study -- to detect puzzlement, we use the facial expression clips that contain puzzlement as training and testing dataset. These clips record common non-rigid movements of the faces. Most of them are 5 seconds long. Database designers made every subject watch a 10-minute video which contained various scenes intended to elicit different emotions. This method that triggers puzzlement affective state is similar with that in learning or e-learning. Some clips contain more than puzzled expression, which turns to surprise or disbelief, and finally laughter. Compared with other facial expression databases, the expressions in this database are more natural and subtle.

There are 191 subjects whose clips contain puzzlement. Because we adopt the basic AAM algorithm which does not take into consideration any complex situations, of these 191, 40 are excluded for reasons including glasses, hair covering eye and unobvious expression. The reason for unobvious expression means the movement of the subject's face is too subtle to be recognized his expression. The final dataset consisted of 151 subjects, including 30 males and 121 females.

Figure 1. Example Frames from the Dataset (a) and (c): Neutral, (b) and (d) Puzzlement

The clips are stored in DV stream format at a resolution of 720x480 with 24-bit color and 29.97 frames per second. From these clips, totals 22,936 frame images were obtained and converted to grayscale. In each frame image, the face area spanned an average of approximately 180x200 (36,000) pixels. Figure 1 shows some example frames from the dataset.

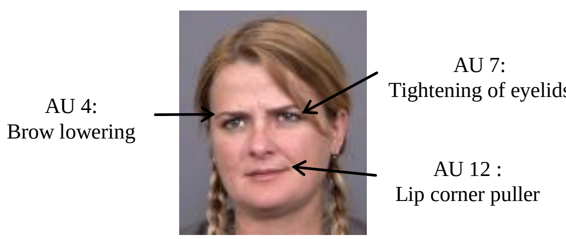


Figure 2. An Example of Facial Actions Associated when a Person is Puzzled

According to the recent psychology study [6], puzzlement can be manifested by the Ekman’s Facial Action Coding System [26]. That is lowered brow (AU 4), the tightening of the eyelids (AU 7), and lip corner puller (AU 12). Figure 2 presents an example of puzzlement in which these AUs are prominent. We used this criterion to divide the 22,936 frames in two parts. The first part contained frames that show subjects with puzzlement, while the second part contained frames without puzzlement. In this way, 8933 frames (38.95%) are categorized in the first part and 14003 frames (61.05%) are categorized in the second part.

3. Active Appearance Model

In this section, we briefly describe the AAM algorithm, which has two procedures: modeling and fitting, which mostly follow the work of [23] and [24].

3.1. AAM Modeling

The AAM decouples and models the shape and the texture of the deformable object to generate a variety of instant photos realistically. The shape is a vector that is described by a set of concatenation position elements of the labeled landmarks, while the texture describes the measure of pixels, which is usually represented by intensities or colors. For modeling, a training set of labeled images, in which key landmark points are marked on each example object, is required. A sample-labeled image is shown in Figure 3(a). As described in [23] and [24], the shape and the texture are modeled, respectively. In two-dimensional cases, the shape of the AAM is described by a triangulation mesh. In particular, the coordinates of the mesh vertices define the shape, which can be represented by a vector s consisting of the coordinates of n landmark points. The shapes are then normalized by the Procrustes analysis and projected onto the shape subspace created by Principal Component Analysis (PCA).

(a) (b)

Figure 3. (a) Shape, (b) The Shape-free Patch

For texture modeling, each image in the training set is warped to match the mean shape to produce "shape-free patches". A sample of the shape-free patches is shown in Figure 3(b). Then the texture can be raster scanned into a vector, g. To minimize the effect of global lighting variation, vector, g, is linearly normalized by applying offsetting (changing brightness) and scaling (changing the contrast) of the entire shape-free patch. The texture is ultimately projected onto the texture subspace based on PCA.

Finally, the coupled relationship between the shape and the texture is analyzed by PCA, and the appearance model is created, which can be described as follows:

$$s \leftarrow s_0 + Q_s c \quad (1)$$

$$g \leftarrow g_0 + Q_g c \quad (2)$$

where c is a vector of appearance parameters controlling both the shape and the texture, s_0 is the mean shape, g_0 is the mean texture, and Q_s and Q_g are matrices describing the modes of variation derived from the training set.

3.2. AAM Fitting

In general, the AAM fit their shape and texture components through a gradient descent search. Let p denote the parameter vector of AAM, which is given by $p^T = (c^T | t^T | u^T)$. It is the combination of the appearance parameter vector, c, the pose parameter vector, t, and the texture transformation parameter vector, u, where u defined as 0 is the identity transformation. The appearance parameter, c, and the pose parameter, t, define the shape of the image patch, S, which can be calculated as follows using equation (1):

$$S \leftarrow S_t (s_0 + Q_s c) \quad (3)$$

where S_t is a similarity transformation described by the pose parameter, t.

During fitting, let g_s denote the texture vector of the new image, which is generated by sampling the pixels in S and then projecting these pixels into the texture model frame, and let g_m denote the texture vector generated by the model using equation (2). It can be supposed that there is a linear relationship between g_s and g_m. Then, the variance of p, Δp, and the texture difference between an image and the model, $r(p)$, can be described as follows:

$$\Delta\, p = R\, r(p) \tag{4}$$

$$r(p) = g_s - g_m \tag{5}$$

where R is the linear relationship (or gradient matrix) between Δp and r. In the AAM, R is assumed to be fixed and is estimated once from a training set by multivariate linear regression techniques. By equations (4) and (5), the AAM fitting can be solved by an iterative procedure.

In our research, the key frames within each video sequence are manually labeled. Then, we built a person specific model for each subject using the modeling process described in section 3.1. Thus, the remaining frames are automatically aligned based on person specific models and the gradient descent fitting process as described above. We also implemented a preprocessing step to improve the robustness of the AAM illumination variations. Indeed we carried out the histogram equalization as a preprocessing step on images before the AAM modeling and fitting. This method provides a good image sequence analysis of faces.

4. Feature Extraction

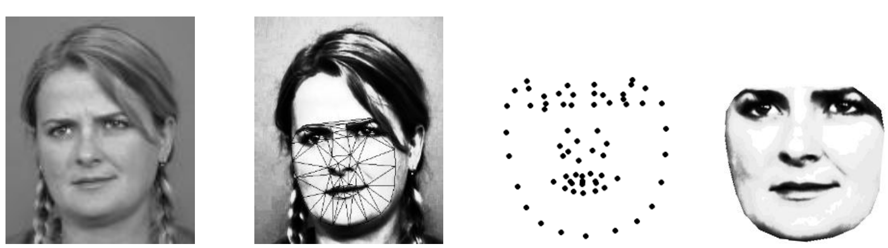
AAMs identify facial feature points, and they provide an approximation of the facial texture, but they do not provide direct information about characteristics of the face that are of relevance to its facial expression. Thus, we investigate two types of features, which are discussed separately in the next two subsections.

4.1. Pose Normalized Shape

On the basis of section 3.2, the shape S is decided by two parts in the AAM parameter p: the appearance parameter c and the pose parameter t. t is the rigid transformation parameter which is associated with the geometric similarity transform, *i.e.*, scale, in-plane rotation and translation. c is the object-specific parameter that is the residual parameter representing geometric variations associated with the actual object shape, *e.g.*, eyes blinking, mouth opening and closing.

Obviously, object-specific variations are directly of relevance to the facial expression. Therefore, we define the pose normalized shape as the first type of feature, which gives the locations of shape mesh vertices after all rigid transformations (scale, in-plane rotation and translation) have been removed.

Let s_n denote the pose normalized shape, which can be obtained by synthesizing a shape instance of s, using equation (3), that ignores the similarity transformation, S_t, described by the pose parameter t. Examples of this shape mesh are shown in the third column of Figure 4. In our research 68 vertex points, which are defined like [27], are used in s_n for both x-coordinate and y-coordinate. Then we obtained a raw 136 dimensional feature vector.



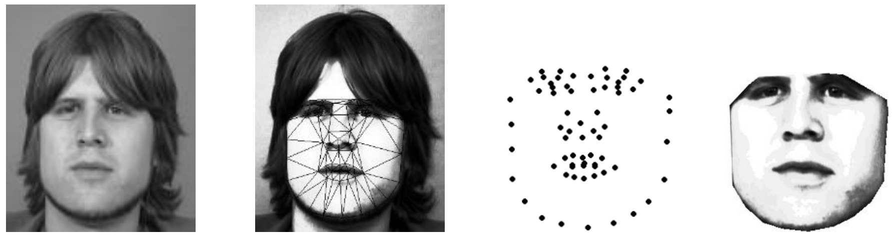


Figure 4. Examples of the Shape and Texture Features Extracted from AAM

4.2. Length Normalized Texture

Pose normalized shape features do not possess the information on the facial texture, such as the presence of wrinkles in the face. Thus, we need a type of texture feature that captures such information.

According to section 3.2, during fitting the most direct texture feature is g_s, which is the sampled texture vector of the current image and projected to the texture model frame. We use the final value of g_s in the fitting iterative procedure as the texture feature. Because the AAM models that we built are person specific models, if not normalized process is used, the length of g_s varies with different subject. Therefore, when building the specific model for each subject we normalize the texture vector g of each model to the same length by scaling the image patch shape of each subject. Then during fitting the g_s of all subjects have the same length.

In our research, a 32,768 dimensional raw feature vector of g_s is used. We chose 32,768, because it is approximately the size of face area (36,000 pixels) in the frame, and it is integer multiples of 4,096, which can make our parallel algorithm of the AAM fitting on GPUs [28] achieves better real time performance than other length which is not integer multiples of 4,096. The example of such normalized textures is shown in the fourth column of Fig. 4. The first column is the original frame. The second column is the histogram equalized and AAM tracked frame. The third column is the pose normalized shape features (PNS). The fourth column is the length normalized texture features (LNT).

5. SVM Classifiers

Support vector machines (SVMs) are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. Because SVMs are binary classifiers and have been proven very effective in numerous pattern recognition tasks including face and facial expression recognition [10, 11, 29], they are well suited to our purpose of classifying puzzlement and non-puzzlement.

Originally, SVMs were defined for the classification problem of linearly separable classes of objects. They try to find the hyper-plane that maximizes the margin between positive and negative observations for a specified class.

For any particular set of two-class objects, an SVM finds the unique hyper-plane having the maximum margin. As shown in Figure 5, the hyper-plane *H1* defines the border with class +1 objects, whereas the hyper-plane *H2* defines the border with class -1 objects. There are two objects from class +1 that define the *H1*, and there are three objects from class -1 that define the *H2*. These objects, which are represented inside circles in Figure 5, are called support vectors. A special characteristic of SVMs is the support vectors that determine the maximum margin hyper-plane represent the solution to a classification problem.

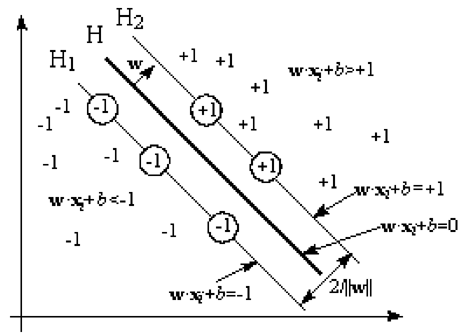


Figure 5. Schematic Diagram of SVMs in Classification

A linear SVM classification decision is made for an unlabeled test observation x_i:

$$\begin{cases} y_i = +1 \ \ if \ \ w \cdot x_i + b > +1 \\ y_i = -1 \ \ if \ \ w \cdot x_i + b < -1 \end{cases} \tag{6}$$

where w is the vector normal to the separating hyper-plane (H: $w\ x_i + b = 0$) and b is the bias. Both w and b are estimated so that they minimize the structural risk of a train-set, thus avoiding the possibility of over fitting to the training data. Typically, w is not defined explicitly, but through a linear sum of support vectors.

SVM can also be used to classify classes that cannot be separated with a linear classifier. In such cases, the coordinates of the objects are mapped into a feature space using nonlinear functions, such as the radial basis function (RBF), the polynomial function and the sigmoid function.

As suggested in [30], if the number of features is large, one may not need to map data to a higher dimensional space. That is, nonlinear mapping does not improve the performance. Using the linear SVM is good enough. Therefore, in our experiments a linear SVM is used due to its ability to generalize well to unseen data in many pattern recognition tasks [30].

6. Experiment

For evaluating the utility of various AAM representations, we trained different SVM classifiers by using features described as Section 4 in the following single or combination:

1) PNS: pose normalized shape, s_n

2) LNT: length normalized texture, g_s

3) PNS + LNT: pose normalized shape, s_n, combined with length normalized texture, g_s

All the three single or combination features are evaluated on the complete dataset described in Section 2 that includes 8933 frames for puzzlement examples and 14003 frames for non-puzzlement examples. The obtained results are presented in Table 1. These values are calculated using 10-fold cross validation. Table 1 shows recognition rates and equal-error rates using each of the three features. This result intuitively highlights the importance of texture features for puzzlement recognition. Only using the feature of pose normalized shape (PNS), we achieve relatively low recognition accuracies (Recognition Rate = 84.13%, Equal-Error Rate = 15.87%), whereas, only using length normalized texture (LNT), we achieve very high recognition accuracies (Recognition Rate = 97.89%, Equal-Error Rate = 2.11%). The best results (Recognition Rate = 98.91%, Equal-Error Rate = 1.09%) are for pose normalized shape combined with length normalized texture (PNS+LNT). This combination may add additional dimensions that aid in finding an optimal separating hyper-plane between two classes. However, it is just 1% higher than the LNT feature.

Table 1. Performance of the Three Features Described in Section 6

Features	Recognition Rate	Equal-Error Rate
PNS	84.13%	15.87%
LNT	97.89%	2.11%
PNS + LNT	98.91%	1.09%

We also studied the influence of the size of the training set on the performance of the three single or combination features used in our experiment. The result is shown in Figure 6. We have calculated the average recognition rates using different number of folds (*k*) for the *k*-fold cross validation technique and plotted them in Figure 6. Figure 6 also shows the performance of the LNT feature is quite closed to that of the PNS + LNT feature combination with every size of training set. It is also observable from the figure that both LNT feature and PNS + LNT combination feature achieved a very high recognition rate even with a relatively small training set (*i.e.*, 2-folds). This indicates that how well our proposed features are suitable for the puzzlement classification task.

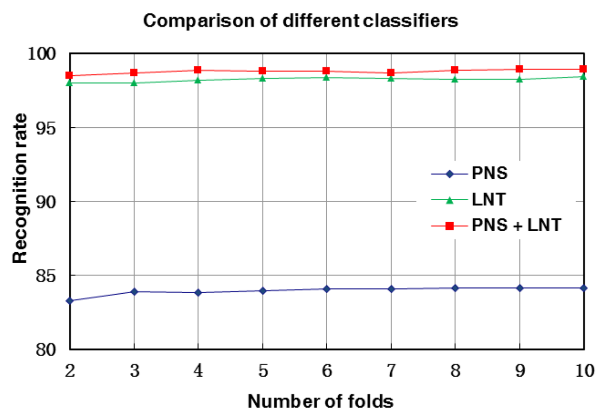


Figure 6. The Evolution of the Achieved Average Recognition Rate for Puzzlement with the Increasing Number of Folds for the *k*-fold Cross Validation Technique

By integrating the two results from both Table 1 and Figure 6, the LNT feature plays a decisive role in the task of puzzlement detection in our experiment, and only using this feature alone to recognize puzzlement is feasible and will not lose much accuracy.

7. Conclusion

In this paper, we presented a novel machine learning approach based on AAM algorithms and SVM classifiers for automatically detecting puzzlement affective state from natural and subtle facial expression. We explored two facial features derived from AAMs, and demonstrated the utility of these features for the task of puzzlement detection. After a series of experiments, we obtain some conclusions: (1) using the combination of pose normalized shape and length normalized texture (PNS + LNT) is superior to either pose normalized shape (PNS) or length normalized texture (LNT) alone. (2) The texture feature has a decisive role to play in puzzlement detection particularly. Only using the LNT alone, we can achieve high recognition rates. (3) Automatic puzzlement detection through video at the level of individual frames appears to be a feasible task. This suggests an opportunity to build interactive systems that can give help information by detecting learners' puzzlement in e-learning. This approach may be also useful for developing puzzlement-triggered intelligent help systems for software. One limitation of our method is the exclusion of subjects wearing glasses or having hair covering eye which limits its generalizability. Our method can also be used for real-time applications since we designed a parallel algorithm of the AAM fitting on GPUs.

In future research, firstly, we will improve the robustness of our approach to make it generalizable. Secondly, we plan to extend this approach so that it can recognize other affective states from natural and subtle facial expression, such as boredom and frustration.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China under No. 61203259 and No. 61103074; the Natural Science Foundation of Tianjin, China under No. 11JCYBJC00600.

References

[1] R. Gupta, "Human Computer Interaction--A Modern Overview," International Journal of Computer Technology and Applications, vol. 3, no. 5, **(2012)**, pp. 1736-1740.
[2] J. Grudin, "Human-computer interaction", Annual review of information science and technology, vol. 45, no. 1, **(2011)**, pp. 367-430.
[3] D'Mello, K. Sidney and A. C. Rafael, "Significant accomplishments, new challenges, and new perspectives," New Perspectives on Affect and Learning Technologies, Springer Publishing, New York **(2011),** pp. 255-271.
[4] O'Brien, L. Heather and E. G. Toms. "The development and evaluation of a survey to measure user engagement," Journal of the American Society for Information Science and Technology , vol. 61, no. 1, **(2010)**, pp. 50-69.
[5] B. McDaniel, D. Sidney, B. King, P. Chipman, K. Tapp and A. Graesser, "Facial features for affective state detection in learning environments," Proceedings of the 29th Annual Cognitive Science Society, **(2007)**August 1-4, pp. 467-472, Nashville, Tennessee, USA.
[6] S. K. D'Mello, S. D. Craig, J. Sullins, and A. C. Graesser, "Predicting affective states through an emote-aloud procedure from AutoTutor's mixed-initiative dialogue," International Journal of Artificial Intelligence in Education, vol. 16, no. 1, **(2006)**, pp. 3-28,.
[7] S. Asteriadis, P. Tzouveli, K. Karpouzis and S. Kollias, "Estimation of behavioral user state based on eye gaze and head pose—application in an e-learning environment," Multimedia Tools and Applications, vol. 41, no. 3, **(2009)**, pp. 469-493.
[8] L. Shen, M. Wang and R. Shen, "Affective e-learning: using" emotional" data to improve learning in pervasive learning environment," Educational Technology and Society, vol. 12, no. 2, **(2009)**, pp. 176-189.
[9] D. Keltner and P. Ekman, "Facial expression of emotion," Handbook of Emotions 3rd, The Guilford Press, New York, **(2010)**, pp. 211-234.
[10] C. Shan, S. Gong and P. W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study", Image and Vision Computing, vol. 27, no. 6, **(2009)**, pp. 803-816.
[11] S. B. Kazmi and M. A. Jaffar, "Wavelets-based facial expression recognition using a bank of support vector machines", Soft Computing, vol. 16, no. 3, **(2012)**, pp. 369-379.
[12] T. Jabid, M. H. Kabir and O. Chae, "Robust facial expression recognition based on local directional pattern," ETRI journal, vol. 32, no. 5, **(2010)**, pp. 784-794.
[13] R. A. Khan, A. Meyer, H. Konik and S. Bouakaz, "Human vision inspired framework for facial expressions recognition," Proceedings of 2012 IEEE International Conference on Image Processing, **(2012)** September 30-October 3, pp. 2593-2596, Orlando, Florida, USA.
[14] R. A. Khan, M. Alexandre, H. Konik and S. Bouakaz, "Framework for reliable, real-time facial expression recognition for low resolution images," Pattern Recognition Letters, vol. 34, **(2013)**, pp. 1159-1168.
[15] M. H. Sigari, M. Fathy and M. Soryani, "A Driver Face Monitoring System for Fatigue and Distraction Detection", International Journal of Vehicular Technology, article ID 263983, **(2013)**.
[16] M. Patel, S. Lal, D. Kavanagh and P. Rossiter, "Fatigue detection using computer vision", International Journal of Electronics and Telecommunications, vol. 56, no. 4, **(2010)**, pp. 457-461.
[17] M. Pantic, A. Pentland, A. Niholt and T. S. Huang, "Human computing and machine understanding of human behavior: a survey," Proceedings of the 8th International Conference on Multimodal Interfaces, **(2006)** November 2-4, pp. 47-71, Banff, Canada.

[18] Y. Tian, J. F. Cohn and T. Kanade, "Facial expression analysis," Handbook of face recognition, Springer Publishing, New York, **(2005)**, pp. 247-276.
[19] R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification 2nd, John Wiley and Sons Inc., New York, **(2001)**.
[20] J. Ou, X. B. Bai, Y. Pei, L. Ma and W. Liu, "Automatic facial expression recognition using Gabor filter and expression analysis", Proceedings of 2nd IEEE International Conference on Computer Modeling and Simulation, **(2010)** January 22-24, vol. 2, pp. 215-218, Sanya, China.
[21] T. Wu, S. B. Marian and R. M. Javier, "Facial expression recognition using gabor motion energy filters," Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, **(2010)** June 13-14, pp. 41-47, San Francisco, California, USA.
[22] T. F. Cootes, D. H. Cooper, C. J. Taylor and J. Graham, "Active shape models—their training and application," Computer vision and image understanding, vol. 61, no. 1, **(1995)**, pp. 38-59,.
[23] T. F. Cootes, G. J. E. Cootes and C. J. Taylor, "Active appearance models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, **(2001)**, pp. 681-685.
[24] T. F. Cootes, G. J. E. Cootes, C. J. Taylor, "Active appearance models," Proceedings of 5th European Conference on Computer Vision, **(1998)** June 2-6, pp. 484-498, Freiburg, Germany.
[25] A. J. O'Toole, J. Harms, S. L. Snow, D. R. Hurst, M. R. Pappas, J. H. Ayyad and H. Abdi, "A video database of moving faces and people," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 5, **(2005)**, pp. 812-816.
[26] P. Ekman, W. Friesen, and J. Hager, "Facial Action Coding System: Research Nexus," Network Research Information, Salt Lake City, **(2002)**.
[27] T. Cootes, "Talking Face Video," **(2013),** http://www.isbe.man.ac.uk/~bim/data/talking_face/talking_face.html.
[28] J. W. Wang, X. R. Ma, Y. P. Zhu, and J. Z. Sun, "Efficient Parallel Implementation of Active Appearance Model Fitting Algorithm on GPU," The Scientific World Journal, article ID 528080, **(2014)**.
[29] E. Gumus, N. Kilic, A. Sertbas and O. N. Ucan, "Evaluation of face recognition techniques using PCA, wavelets and SVM," Expert Systems with Applications, vol. 37, no. 9, **(2010)**, pp. 6404-6408.
[30] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification," Technical Report, **(2005)**.

Authors

Jinwei Wang, he received the M.S. degrees in Computer Software and Theory from Institute of Computing Technology, Chinese Academy of Sciences. Now he is a lecturer at the College of Computer and Information Engineering, Tianjin Normal University, and pursuing his Ph.D. degree in the School of Computer Science and Technology, Tianjin University. His research work focuses on affective computing and pattern recognition.

Xirong Ma , she received the Ph.D. degrees in Computer Science and Technology from College of Information Technical Science, Nankai University. Now she is a professor at the College of Computer and Information Engineering, Tianjin Normal University, and the Council Member of the Chinese Association for Artificial Intelligence. Her research work focuses on affective computing and pattern recognition.

Jizhou Sun, he received the Ph.D. degrees in Computer Science from College of Engineering, the University of Sussex. Now he is a professor at the School of Computer Science and Technology, the Executive Director of China Society of Image and Graphics, and the Executive Director of China Society of Microcomputer. His research work focuses on image processing and parallel computing.

Ziping Zhao, he received the Ph.D. degrees in Computer Science and Technology from College of Information Technical Science, Nankai University. Now he is an associate professor at the College of Computer and Information Engineering, Tianjin Normal University. His research work focuses on pattern recognition and image processing.

Yuanping Zhu, he received the Ph.D. degrees in Computer Science and Technology from Institute of Automation, Chinese Academy of Sciences. Now he is an associate professor at the College of Computer and Information Engineering, Tianjin Normal University. Before joining TJNU he was with Fujitsu as a research fellow. His research work focuses on pattern recognition and image processing.