IEICE TRANS. INF. & SYST., VOL.E99–D, NO.6 JUNE 2016

LETTER

# Micro-Expression Recognition by Regression Model and Group Sparse Spatio-Temporal Feature Learning*

**Ping LU**[†], **Wenming ZHENG**[††a)], **Ziyan WANG**[††], **Qiang LI**[††], **Yuan ZONG**[††], ***Nonmembers***, **Minghai XIN**[††], ***Member***, ***and*** **Lenan WU**[†††], ***Nonmember***

**SUMMARY** In this letter, a micro-expression recognition method is investigated by integrating both spatio-temporal facial features and a regression model. To this end, we first perform a multi-scale facial region division for each facial image and then extract a set of local binary patterns on three orthogonal planes (LBP-TOP) features corresponding to divided facial regions of the micro-expression videos. Furthermore, we use GSLSR model to build the linear regression relationship between the LBP-TOP facial feature vectors and the micro expressions label vectors. Finally, the learned GSLSR model is applied to the prediction of the micro-expression categories for each test micro-expression video. Experiments are conducted on both CASME II and SMIC micro-expression databases to evaluate the performance of the proposed method, and the results demonstrate that the proposed method is better than the baseline micro-expression recognition method.

***key words:*** *micro-expression recognition, local binary patterns on three orthogonal planes (LBP-TOP), group sparse least squares regression (GSLSR)*

## 1. Introduction

Micro expressions are short duration facial expressions that cannot be repressed when a person tries to conceal a genuine emotion [1]. Ekman [2] discovered micro expressions by watching a surveillance video in 1969, in which he saw a patient suffering from depression and trying to conceal his attempt to commit suicide showed an extremely pained expression in two frames out of a long image sequence. Because of the broad practical applications in many areas such as clinic diagnosis or criminal investigation, micro-expression research has attracted increasing attention in the scientific community in recent years. For instance, in clinical diagnosis, micro-expression recognition could be used to help doctors to understand the true feelings of patients. Another example is the application of micro-expression recognition in criminal investigation, in which micro-expression recognition could provide a useful way to detect lying of person since the micro expressions may be unconscious and cannot be independently controlled by human beings.

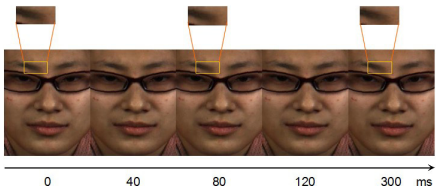



**Fig. 1** Example of micro expression occurrence, where the image sequence is chosen from CASME II database. The apex frame presents at about 80ms from the onset, in which the three rectangles above the images show the small change of AU4.

Basically, micro expression is featured by its short duration, whose upper limit of duration is usually about $\frac{1}{2}$s [3]. Figure 1 shows an example of image sequence with disgust micro expression occurrence in the middle of the image sequence [4], in which the apex frame presents at about 80ms from the onset, the three rectangles above the images show the small change of AU4 (inner brow) [5]. From this figure, we can see that micro expression is just a kind of fleeting facial expressions with low intensity. Accordingly, it takes considerable time and efforts to develop the ability to recognize the micro-expression by the naked eyes. Because of this, it is very challenging but meaningful to develop an automatic micro-expression recognition system.

Although the study of automatic facial expression recognition began as early as 1970's, only a few work available in the literatures addressing the study of micro-expression recognition in recent years [6]–[10]. In [6], Wu et al. developed an automatic micro-expression recognition system, in which both Gabor features and Gentleboost method are utilized, respectively, for feature extraction and selection. Then, a support vector machine (SVM) classifier is used for micro-expression recognition. In [7], Ruiz-Hernandez and Pietikainen encoded the local binary patterns (LBP) to generate more robust histograms as feature for micro-expression recognition. More recently, Wang et al. [8] proposed a method of dealing with micro-expression recognition under the varying lighting conditions by using discriminant tensor subspace analysis (DTSA) and extreme

Manuscript received October 15, 2015.
Manuscript revised February 1, 2016.
Manuscript publicized February 29, 2016.

[†]The author is with the ZTE Corporation, Nanjing 210012, Jiangsu, P.R. China.
[††]The authors are with the Key Laboratory of Child Development and Learning Science of Ministry of Education, Research Center for Learning Science, Southeast University, Nanjing 210096, Jiangsu, P.R. China.
[†††]The author is with the School of Information Science and Engineering, Southeast University, Nanjing 210096, Jiangsu, P.R. China.

*This work was supported in part by the National Basic Research Program of China under Grant 2015CB351704, the National Natural Science Foundation of China (NSFC) under Grants 61231002 and 61572009, and the Natural Science Foundation of Jiangsu Province under Grant BK20130020.

a) E-mail: wenming_zheng@seu.edu.cn (Corresponding author)
DOI: 10.1587/transinf.2015EDL8221

Copyright © 2016 The Institute of Electronics, Information and Communication Engineers

learning machine (ELM).

The aforementioned investigations have significantly advanced automatic micro-expression recognition research. However, there are still limitations existing in these methods. One of the major problems is that most of the micro expressions used in the research are not real spontaneous expressions. Moreover, from the recognition point of view, most of micros-expression recognition approaches focus on the global facial features and ignore the local features. Taking the example shown in Fig. 1, the micro expression involves the extremely minor changes in AU4 (inner brow) and AU9 (bridge of the nose), which indicate that the region of muscle movement in micro expression is smaller than that in basic expressions. Because of this, it would be helpful when dealing with micro-expression recognition to take into consider the important region selection problem, where the important regions are associated with the major muscle motion when micro expression occurs.

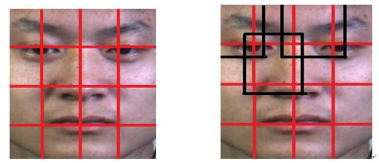
In our preliminary work [11], we have proposed a method based on group sparse learning approach to simultaneously cope with both facial region selection and multi-view expression recognition. Motivated by the work of [11], in this letter we also investigate the group spare learning approach to the major facial region selection in micro-expression recognition. More specifically, we adopt the local binary patterns on three orthogonal planes (LBP-TOP) [12] to extract the facial features from the different facial sub regions of each image sequence and then use group sparse least square regression (GSLSR) model to deal with the micro-expression recognition task. Finally, we conduct extensive experiments on two spontaneous micro-expression databases.

## 2. Micro-Expression Recognition Based on GSLSR

In this section, we will address our micro-expression recognition approach in details. Following the facial expression recognition framework of [11], the micro-expression recognition procedures can be divided into the following major parts: (1) preprocessing and feature extraction part; (2) GSLSR model training part; (3) Micro-expression recognition part.

### 2.1 Micro-Expression Feature Extraction

Micro expression is a kind of expression with very short duration. Consequently, we adopt the spatio-temporal facial feature to describe the micro expressions. Specifically, to extract the micro-expression facial features, we firstly normalize the facial images of each image sequence to a uniform size. Then, we use four different scale sizes to divide each facial image sequence, such that each facial image in the sequence is divided into 1, 4, 16 and 64 facial sub regions, respectively. For instance, in the $i$-th image partition, each facial image is equally divided into $i$ columns and $i$ rows. As a result, we obtain $i^2$ facial regions in total. Figure 2 (a) illustrates an example of the image division, where



(a) (b)

**Fig. 2** Regional segmentation examples. (a) The equal division of $4 \times 4$ areas; (b) overlapping example $4 \times 4$ (black parts).

$i$ is fixed at 4 and hence the total number of image regions is $4^2 = 16$. In this case, we finally obtain $1 + 2^2 + 4^2 + 8^2 = 85$ facial regions in total. Moreover, in order to improve the recognition performance, instead of focusing on each sub region to extract the facial expression features, we extract the facial expression feature in the region with a bit overlapping with its neighbors as shown in Fig. 2 (b).

To extract the spatio-temporal micro-expression features, we apply LBP-TOP operator [12] to each candidate facial region in the micro-expression video, where we denote the LBP-TOP features with the form of $LBP-TOP^{u_2}_{8,8,8,R_X,R_Y,R_T}$. For each region in an image sequence, we obtain a $59 \times 3$ dimensional LBP-TOP feature vector. Finally, for all of the 85 facial regions, we concatenate the feature vectors into a higher feature vector with dimensionality of $59 \times 3 \times 85 = 15045$.

### 2.2 Micro-Expression Recognition Based on GSLSR

In the above section, we have addressed the method of region-based micro-expression feature extraction. However, it is notable that the facial features extracted from different facial regions would contribute differently to the micro-expression recognition. For instance, as what we have shown in Fig. 1, the micro expressions usually occur accompanied with some typical AUs changes. Consequently, it would be helpful to assign a weight for each facial region such that the different contributions of the regions to micro-expression recognition are proportional to the value of weight parameters. If the facial features of a face region contribution less to the micro-expression recognition, then the corresponding weight parameter would be set to be zero. As a result, the corresponding facial features would have no any effect on the recognition of micro expressions.

Now let $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_M]$ denotes the class label matrix corresponding to $M$ facial image sequences, where the $i$-th column of $\mathbf{X}$, denoted by $\mathbf{x}_p = [x_p^1, \cdots, x_p^c]^T$, represents the class label vector of the $p$-th facial image sequences, and $c$ is the number of micro expression classes. The value of $x_p^j$ is set to be 1 if the corresponding facial image sequences belonging to the $j$-th micro-expression category. Otherwise, it would be set to be 0. Let $\mathbf{Y} = [(s_1^1\mathbf{Y}_1^1)^T, (s_2^1\mathbf{Y}_2^1)^T, \cdots, (s_2^4\mathbf{Y}_2^4)^T, (s_4^1\mathbf{Y}_4^1)^T, \cdots, (s_8^{64}\mathbf{Y}_8^{64})^T]^T$ denotes the concatenation of the LBP-TOP facial feature matrices associated with the 85 facial region weight parameters as addressed in above section, where $s_i^j$ denotes the weight parameter. Then the relationship between $\mathbf{X}$ and $\mathbf{Y}$

can be described using the following least squares regression (LSR) problem:

$$\min_{\mathbf{U}} \|\mathbf{X}-\mathbf{U}^T\mathbf{Y}\|_F^2 = \min_{s_i^j\in 0,1,\mathbf{U}_i^j} \|\mathbf{X}-\sum_{i\in\{1,2,4,8\}}\sum_{j=1}^{i^2} s_i^j \mathbf{U}_i^{j^T}\mathbf{Y}_i^j\|_F^2 \tag{1}$$

where $\mathbf{U}$ is the regression coefficient matrix of LSR, and $\|.\|_F$ denotes the Frobenius norm operator.

In order for the facial region selection, we add a regularization with respect to the weight parameters $s_i^j$ onto the regression formula of the LSR model [11], then the LSR model of Eq. (1) can be reformulated into the following one:

$$\min_{s_i^j\in 0,1,\mathbf{U}_i^j} \|\mathbf{X}-\sum_{i\in\{1,2,4,8\}}\sum_{j=1}^{i^2} s_i^j \mathbf{U}_i^{j^T}\mathbf{Y}_i^j\|_F^2 + \lambda \sum_{i\in 1,2,4,8}\sum_{j=1}^{i^2} \|s_i^j\mathbf{U}_i^j\|_F \tag{2}$$

where the coefficient $\lambda > 0$ is served as a trade-off regularization parameter. The minimization problem of Eq. (2) can be solved using alternated direction method (ADM) approach as well as inexact augment Largangrian multiplier (inexact ALM) method, the details of the solving process can be referred to [11].

Denote the multiplication of $s_i^j$ and $\mathbf{U}_i^j$ by $\tilde{\mathbf{U}}_i^j$ such that $\tilde{\mathbf{U}}_i^j = s_i^j\mathbf{U}_i^j$. After obtaining the optimal solution of $\tilde{\mathbf{U}}_i^j$, we can use the regression model to predict the micro-expression label of a image sequence. Specifically, suppose that $\mathbf{y}_t$ is given LBP-TOP feature vector extracted from a testing image sequence and suppose that $\mathbf{x}_t$ is the corresponding expression class label vector. Then, the value of $\mathbf{x}_t$ can be predicted by solving the following minimization problem:

$$\min_{\mathbf{x}_t} \|\mathbf{x}_t - \sum_{i\in\{1,2,4,8\}}\sum_{j=1}^{i^2} \tilde{\mathbf{U}}_i^{j^T}\mathbf{y}_t\| \tag{3}$$

Then, the micro-expression category of the given testing image sequence is assigned as:

$$expression_label = arg\max_k \mathbf{x}_t(k) \tag{4}$$

where $\mathbf{x}_t(k)$ denotes the $k$-th element of $\mathbf{x}_t$.

## 3. Experiments

In this section, we will conduct extensive experiments on two spontaneous micro-expression databases to evaluate the effectiveness of the proposed micro-expression recognition approach. The first database is CASME II built by Yan et al. from Institute of Psychology of Chinese Academy of Sciences [4], and the second one is the SMIC micro-expression database built by Li et al. from Oulu University [13].

### 3.1 CASME II Database

The CASME II database consists of 255 micro-expressions samples of 26 subjects, where the frames of onset and offset of each sequence were coded. Given that the micro-expressions are usually local and in low intensity, the micro-expressions are defined based on the AUs, the contents of the video episodes and participants' self-report. Temporal resolution and the face resolution of the recordings in CASME II database is 200 fps and 280 × 340 pixels, respectively. Table 1 summarizes the data information of CASME II database.

**Table 1** Summary of data information of CASME II database.

| Micro-expressions | Numbers |
|---|---|
| Disgust | 63 |
| Happiness | 32 |
| Repression | 27 |
| Surprise | 25 |
| Others | 99 |

**Table 2** Comparison of micro-expression recognition accuracy on CASME II database.

| Method | Accuracy |
|---|---|
| LBP-TOP+SVM [4] | 63.41% |
| LBP-SIP+RBF kernel [14] | 66.40% |
| OSW-LBP-TOP+RBF kernelSVM [15] | 66.40% |
| LSTD+RPCA+SVM [16] | 65.45% |
| LBP-TOP+GSLSR | **67.89%** |
| LBP-TOP+RPCA+GSLSR | 62.60% |

In our experiments, the cropped images of CASME II were converted to 256 × 256 pixels with gray scale. The radii used in LBP-TOP are empirically set using the following approach: the radius of X axis and Y axis are fixed at 1, respectively, and the one of T axis is fixed at 4. For all the three planes used in LBP-TOP, the neighboring points are fized at 8. To evaluate the recognition performance of the proposed micro-expression recognition method, we use leave-one-out cross-validation strategy to design the experiments. In addition, for comparison purpose, we choose four recently published works as the baseline methods [4], [14]–[16].

The results of all the methods are reported in Table 2. From Table 2, we can see that the proposed LBP-TOP+GSLSR method achieves the best recognition accuracy among all the methods, where we got the best performance 67.89%, which has 4.48% improvement over the LBP-TOP+SVM method. Besides, based on the work of Wang et al. [16], we also integrate robust principal component analysis (RPCA) with our methods (LBP-TOP+RPCA+GSLSR). It is notable that the recognition rate is only 62.60%, which is much less than that of the orignal proposed method. We think that maybe GSLSR discards the important information which is extracted by RPCA, namely the learned weight corresponding to image region containing the subtle motion information of expression is 0.

### 3.2 SMIC Database

The SMIC database consists of 164 micro-expression video

**Table 3** Summary of data information of SMIC database.

| Source | Subjects | Micro-expression sequence | | | |
|---|---|---|---|---|---|
| | | positive | negative | surprise | Total |
| HS | 16/20 | 51 | 70 | 43 | 164 |
| VIS | 8/10 | 28 | 23 | 20 | 71 |
| NIR | 8/10 | 28 | 23 | 20 | 71 |

**Table 4** Comparison of micro-expression recognition accuracy on SMIC database.

| Method | Accuracy |
|---|---|
| LBP-SIP+RBF kernel [14] | 64.02% |
| OSW-LBP-TOP+RBF kernel [15] | 57.71% |
| LBP-TOP+Nearest Neighbor [17] | 65.83% |
| DLSTD+RPCA+SVM [16] | 68.29% |
| LBP-TOP+GSLSR | **70.12**% |
| LBP-TOP+RPCA+GSLSR | 67.07% |

clips with 16 participants. The videos were recorded by a high speed camera (HS) of 100 fps and later, a normal visual camera (VIS) and a near-infrared camera (NIR), both with 25 fps, were used for the 71 micro-expression sequences respectively. There are three classes of micro expressions in this database, i.e., positive (happy), negative (sad, fear, disgust) and surprise. Table 3 summarizes the data information of this database.

In our experiments, we only used the data recorded by the high speed camera (HS) of 100 fps. The original color image sequences were converted into gray sequences with size of 128×128. As we did in the above experiment, we extracted LBP-TOP features and the leave-one-out cross validation was applied. For comparison purpose, four works are used as the baseline methods, i.e., in addition to the research work in [14], [15] and [16], we also choose the recent work of Guo et al. in [17]. Besides, our method plus RPCA scheme proposed in [16] is included in the comparison as well. Table 4 summarizes the experimental results, from which we can see that our LBP-TOP+GSLSR method obtains the best recognition accuracy among the four methods, where the best accuracy is reaching as high as 70.12%. Finally, we notice that although LBP-TOP+RPCA+GSLSR outperforms some methods, its performance is still not satisfactory compared with LBP-TOP+GSLSR.

## 4. Conclusions

In this paper, we have presented a micro-expression recognition method based on LBP-TOP feature and GSLSR model. The LBP-TOP feature is able to make use of the spatio-temporal information that characterizes the facial changes of micro expressions. On the other hand, by using the GSLSR method, we are able to optimize the extracted LBP-TOP features such that features only slightly contributing to expression recognition are removed from the final micro-expression model. The experiments on both CASME II and SMIC micro-expression databases demonstrated the superior recognition performance of the proposed LBP-TOP+GSLSR method. The results showed that the methods in this paper could effectively recognize the micro-expression in the image sequences and they outperformed most recent micro-expression algorithms.

### References

[1] P. Ekman, Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage (Revised Edition), WW Norton & Company, 2009.

[2] P. Ekman and W.V. Friesen, "Nonverbal leakage and clues to deception," Psychiatry, vol.32, no.1, pp.88–106, 1969.

[3] X.-B. Shen, Q. Wu, and X.-L. Fu, "Effects of the duration of expressions on the recognition of microexpressions," Journal of Zhejiang University Science B, vol.13, no.3, pp.221–230, 2012.

[4] W.-J. Yan, X. Li, S.-J. Wang, G. Zhao, Y.-J. Liu, Y.-H. Chen, X. Fu, and K. Guo, "Casme II: An improved spontaneous micro-expression database and the baseline evaluation," PLoS ONE, vol.9, no.1, 2014.

[5] P. Ekman and E.L. Rosenberg, What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS), Oxford University Press, 1997.

[6] Q. Wu, X. Shen, and X. Fu, "The machine knows what you are hiding: An automatic micro-expression recognition system," International Conference on Affective Computing and Intelligent Interaction (ACII2011), pp.152–162, Springer, 2011.

[7] J.A. Ruiz-Hernandez, M. Pietikainen, "Encoding local binary patterns using the re-parametrization of the second order gaussian jet," IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG2013), pp.1–6, IEEE, 2013.

[8] S.-J. Wang, H.-L. Chen, W.-J. Yan, Y.-H. Chen, and X. Fu, "Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine," Neural Processing Letters, vol.39, no.1, pp.25–43, 2014.

[9] S.-J. Wang, W.-J. Yan, X. Li, G. Zhao, C.-G. Zhou, X. Fu, M. Yang, and J. Tao, "Micro-expression recognition using color spaces," IEEE Trans. Image Process., vol.24, no.12, pp.6034–6047, 2015.

[10] Y.-J. Liu, J.-K. Zhang, W.-J. Yan, S.-J. Wang, G. Zhao, and X. Fu, "A main directional mean optical flow feature for spontaneous micro-expression recognition," IEEE Transactions on Affective Computing, 2015.

[11] W. Zheng, "Multi-view facial expression recognition based on group sparse reduced-rank regression," IEEE Trans. Affective Comput., vol.5, no.1, pp.71–85, 2014.

[12] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," IEEE Trans. Pattern Anal. Mach. Intell., vol.29, no.6, pp.915–928, 2007.

[13] X. Li, T. Pfister, X. Huang, G. Zhao, and M. Pietikainen, "A spontaneous micro-expression database: Inducement, collection and baseline," IEEE International Conference on Automatic Face and Gesture Recognition (FG2013), pp.1–6, IEEE, 2013.

[14] Y. Wang, J. See, R.C.-W. Phan, and Y.-H. Oh, "LBP with six intersection points: Reducing redundant information in lbp-top for micro-expression recognition," Asian Conference on Computer Vision (ACCV2014), pp.525–537, Springer, 2014.

[15] S.-T. Liong, J. See, R.C.-W. Phan, A.C.L. Ngo, Y.-H. Oh, and K. Wong, "Subtle expression recognition using optical strain weighted features," Asian Conference on Computer Vision (ACCV2014), pp.644–657, Springer, 2014.

[16] S.J. Wang, W.J. Yan, G. Zhao, X. Fu, and C.G. Zhou, "Micro-expression recognition using robust principal component analysis and local spatiotemporal directional features," Computer Vision-ECCV 2014 Workshops, pp.325–338, Springer, 2014.

[17] Y. Guo, Y. Tian, X. Gao, and X. Zhang, "Micro-expression recognition based on local binary patterns from three orthogonal planes and nearest neighbor method," IEEE International Joint Conference on Neural Networks (IJCNN2014), pp.3473–3479, IEEE, 2014.