# Detecting Hidden Affective States: A Review of Micro-expressions Analysis

牛雪松 201618013229049
Institute of Computing Technology, Chinese Academy of Sciences
nxs583966@163.com

## Abstract

*Mirco-expressions are the subtle facial expressions happened in the high-stake situation. It is very import for reading the hidden emotion of the subjects. The recognition of micro-expressions contains two steps: micro-expression spotting and micro-expression recognition. In this paper, I divide the micro-expression spotting process into face preprocess and tracking, feature generation, feature difference analysis, and review the methods used for each part. At the same time, I review the existing methods of preprocessing methods, features and classifiers used for micro-expression recognition. Finally, public available micro-expression databases are also summarized.*

## 1. Introduction

Facial expressions are very important for detecting the emotional states of a person, and affective states recognition has drawn many attentions over the past few decades. Instead of traditional facial expression recognition, which mainly focus on detecting primary emotion expressions including disgust, fear, joy, surprise, sadness, anger [2], a lot of studies have also paid attention to micro-expression spotting and recognition in recent years. The first study on micro-expression is reported by Haggard and Issacs in 1966 [4], and micro-expressions are defined as brief and subtle facial expressions, which usually appear in high-stakes situation, and they are uncontrollable to people themselves. Micro-expressions lasts for 1/25 to 1/3 second [37], and has low intensity, which makes it hard to detect and recognize, and an automatic micro-expression recognition system is required.

The main need for micro-expression recognition is that it is an important clue for real affective states detection, especially when the subject tries to conceal his feelings. In high-stake situations like suspects being interrogated, the micro-expressions across the face could help the police to judge whether the subject is lying or not. In order to help the police to judge whether someone is lying or being honest, Ekman developed a micro-expression training tool to help to improve the abilities to recognize micro-expressions [3]. Another situation micro-expressions recognition system may be used is in psychotherapy, the patients' real emotions may be understood by detecting the micro-expressions.

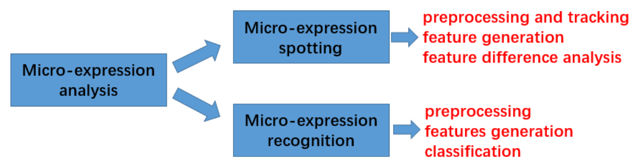


Figure 1: An Overview of the key steps for micro-expression analysis

Although the micro-expressions are important to understand the genuine feelings of subjects, it is hard to spot and recognize even for trained people because of the low intensity and the short duration. Therefore, an automatic analysis of micro-expressions is required using computer vision methods. In recent years, a few micro-expression databases have been released for researches, and some effective methods have been put forward. In this paper, I present the approaches for micro-expression spotting and recognition in Section 2, and introduce the databases for micro-expressions in Section 3. The final conclusion is drawn in Section 4.

## 2. Micro-expression Analysis

As shown in Figure 1, the analysis of micro-expressions contains two steps, spotting and recognition. Spotting means detecting when the micro-expression occurs in a video stream and recognition refers to deciding what expression the micro-expression represents. A successful analysis system of micro-expressions needs effective ways both in spotting and in recognition.

### 2.1. Micro-expression Spotting

The target of micro-expression spotting is to detect the temporal interval of a micro-expression in a frame sequences. The same problem for ordinary facial expressions or eye blink has been studied, and a few solutions have been proposed. However, the problem of spotting micro-expressions is rarely studied and only a few papers focus on solving this problem. In general, this task is a two class's classification, and the methods proposed mainly deal with the difference between frames. Most of the methods can be divided into three steps: face preprocess and tracking, feature generation, feature difference analysis.

The first preprocessing step of micro-expression spotting mainly includes face alignment and block division. There are a lot of fast and accuracy methods for face alignment that can be used in micro-expression analysis [39] [41]. After the face area is aligned into the same size, the face area is usually divided into blocks for detail analysis, and tracking for key points are usually applied. Most of the existing methods use Kanade-Lucas-Tomasi algorithm [26] for tracking, while in [29], Su-Jing Wang et.al employed a robust local optical flow [23] [28] [29] which is more adaptive for different region sizes and illumination situations with a little increase in computational complexity.

The second step of micro-expression spotting is feature generation. Many different descriptors have been used for detecting micro-expressions. Optical flow is the first descriptor used for differentiating micro-expressions and macro-expressions in a video stream [25] [24], and latter, a robust local optical flow feature is also used for micro-expressions spotting [23] [28] [29].The spatiotemporal integration of optical flow has also proven effective [18]. 3D-Gradient orientation histogram descriptor is proposed by Polikovsky et.al [20] [21] and Gabor features are used in [34] by Wu et.al. Local Binary Patterns (LBP) histogram features is proposed by Moilanen et.al for detecting both the temporal and spatial location of micro-expressions [15],and geometric model is used in [35] for classification of micro-expressions and macro-expressions. Recently, Hog features [1] and Histogram of Optical Flow(HOOF) [11] are also be used for micro-expression spotting.

After generating the expressive features, the key steps for micro-expression spotting is to calculate the difference between frames, and decide whether micro-expressions appear. On the one hand, some works use a threshold and peak detection method of the scores calculated from features, such as the optical flow and the 3D-Gradient orientation histogram descriptor. In [24], scores are calculated as the average of optical flow, and in [21], scores are computed from the gradient of facial points. In [28] and [29], the maximal difference values in the main direction of the optical flows (MDMD feature) is proposed and outperforms the former methods based on optical flow. In [34], a boosting method applied to the features is applied for classification. While on the other hands, some recent methods take the dissimilarity and the correlations between frames into consideration. In [15], [1] and [11], Chi-Squared distance is successfully used for calculating the difference between frames, and an auto-calculated threshold is applied to decide whether micro-expression occurs or not. In [35], the deformation correlation between frames is given to an Adaboost model and a random walk model for classification.

## 2.2. Micro-expressions Recognition

After spotting the micro-expressions, effective ways to recognize micro-expressions are needed. There are several ways transferred from traditional video-based face recognition or facial expressions recognition methods, while at the same time, some works take the sparsity of micro-expressions into consideration and proposed specific methods. The same as other recognition task, the micro-expression recognition problem can be divided into three steps: preprocessing, feature generation and classification.

The same as micro-expression spotting, face alignment technology is very important for recognition. At the same time, since the micro-expressions have low intensity and short duration, the preprocessing step focuses on motion magnification and temporal interpolation. Motion magnification is firstly used for micro-expression recognition by Le Ngo et.al [8], and Euler Motion Magnification methods [33] are proven effective. Temporal interpolation is introduced by Pfister et.al [19] and shows effectiveness on improving performance of recognition.

After preprocessing the input video stream, features generated from the video are the key for recognition. Some features for micro-expression spotting are also used for recognition, such as 3D-Gradient orientation histogram descriptor [20] [21], while some other appearance features are specifically designed. Local Binary Pattern with Three Orthogonal Planes (LBP-TOP) is an effective feature for video-based facial expression analysis and has been successfully applied in micro-expressions recognition [19] [9], and an external version of LBP-TOP feature, SpatioTemporal Local Binary Pattern with Integral Projection (STLBP-IP), is proposed by Huang et.al [6] in 2015. Oh et al. proposed a multi-scale Riesz wavelet representation based on LBP feature also achieves good performance on recognition [16]. Another statistical spatiotemporal textural features used for micro-expressions recognition is Local Spatiotemporal Directional Features (LSTD) introduced by Wang et.al [32] which is comparable to LBP-TOP. Besides of those appearance features, features based on optical flow also play important roles in micro-expression recognition. Optical strain, a derivative of optical flow, is firstly utilized by Liong et.al [12]. Another optical flow based feature is the Main Directional Mean Optical-flow (MDMO) feature introduced by Liu et.al [13]. Facial Dynamics Map for optical flow is also used for recognition [36].

Another kind of features for recognition is based on learning methods. Tensor is an effective analysis tools for

Table 1: Posed and spontaneous micro-expression databases

| Database | Elicitation | Resolution | Frame-Rate | Subjects | Emotions | Video Chips | Tagging |
|---|---|---|---|---|---|---|---|
| USF-HD [24] | posed | $1280 \times 720$ | 29.7fps | - | - | 100 | mirco/non-mirco |
| Polikovsky [20] | posed | $640 \times 480$ | 200fps | 10 | 7 | - | FACS |
| SMIC-HS [10] | Spontaneous | $640 \times 480$ | 100fps | 16 | 3 | 164 | Emotion category |
| SMIC-VIS [10] | Spontaneous | $640 \times 480$ | 25fps | 8 | 3 | 71 | Emotion category |
| SMIC-NIR [10] | Spontaneous | $640 \times 480$ | 25fps | 8 | 3 | 71 | Emotion category |
| CASME-A [38] | Spontaneous | $1280 \times 720$ | 60fps | 7 | 8 | 100 | Emotion category/FACS |
| CASME-B [38] | Spontaneous | $640 \times 480$ | 60fps | 12 | 8 | 95 | Emotion category/FACS |
| CASMEII [22] | Spontaneous | $640 \times 480$ | 200fps | 26 | 4 | 247 | Emotion category/FACS |

recognition and it is firstly introduced for micro-expression recognition by Wang et.al [27]. In 2015, they extended this method to Tensor Independent Color Space(TICS) [30]. Latter, they take the sparsity of micro-expressions into consideration and proposed Sparse Tensor Canonical Correlation Analysis (STCCA) [31]. Besides of tensor subspace methods, a multi-task mid-level feature learning mechanism is proposed in [5]. Although deep learning has achieved great success in some recognition task, the feature representations based on deep learning for micro-expression recognition is limited by the scale of data, several works have tried using CNN [14] [17] and LSTM [7] to find representative features for micro-expression recognition.

The final step for a recognition task is to choose a powerful learning strategy. SVM is the most commonly used classifier for micro-expression recognition [19] [16] [8] [13] [31] [5] [30], while Multiple Kernel Learning (MKL) and Random Forest (RF) are also been tried for micro-expressions recognition [19]. Besides of those typical machine learning methods, extreme learning machine (ELM) [27] and relaxed K-SVD [40] are also been introduced to achieve better performance.

## 3. Micro-expressions Databases

There are two kinds of databases for micro-expressions, posed and spontaneous micro-expressions databases. Since the micro-expressions are in low intensity and short duration, it is hard to collect data for micro-expression analysis as well as label them. There are only five public available databases for micro-expression analysis. An overview of the existing databases are shown in Table 1.

### 3.1. Posed Micro-expression Databases

In the early works on micro-expressions recognition, posed micro-expression databases are used to bypass the difficulty of getting spontaneous micro-expressions. There are two posed micro-expression databases: USF-HD and Polikovsky's database.

The USF-HD is constructed by Shreve et.al [24]. The database includes 100 chips of micro-expressions at a resolution of $1280\times720$ and frame-rate of 29.7 fps. The subjects are asked to pose micro-expressions of example videos. The emotion categories of USF-HD is not clear according to [24] and only a few works have used this database.

Polikovsky et.al collected a micro-expressions database in 2009 [20]. 10 subjects from different countries participated in the data collection, and all the records are at the resolution of $640 \times 480$ and at the frame-rate of 200 fps. The micro-expressions of Polikovsky's database are posed by asking the participates to perform 7 basic emotions with low intensity as fast as possible. The motion-spotting label of this database includes 'Constrict', 'InAction' and 'Release' and categories are labelled using the Facial Action Coding System (FACS).

### 3.2. Spontaneous Micro-expression Databases

The problem of posed micro-expression databases is that they are different from natural micro-expressions. Different from posed micro-expressions, naturally micro-expressions usually occurs when the subject not even realize it, and they often appear differently in spatial and temporal properties. Thus, spontaneous micro-expressions databases are need, and there are three existing spontaneous micro-expression databases: SMIC, CASME, and CASMEII.

The first spontaneous micro-expression database is SMIC [10]. In this database, 20 subjects are involved, and all videos are recorded at the resolution of $640 \times 480$. There are three kinds of videos included in SMIC database: High-speed camera (HS), normal visual camera (VIS) and near-infrared(NIR). High-speed videos are recorded at the frame-rate of 100 fps while the normal visual camera and near infrared videos are at 25 fps. 16 subjects participated the High-speed part of experiment and 8 subjects participates the normal visual camera and near-infrared experiment. All the emotions are performed by ask the participants watch 16 carefully selected movie clips, and three kinds of emotions (positive, negative and surprise) are involved. 164 HS videos, 71 VIS videos and 71 NIR videos are recorded.

In 2013, Yan et.al collected another spontaneous micro-expression database CASME [38]. There are two parts of the database: CASME-A and CASME-B. 19 participants are involved and the CASME-A part of the database is record at the resolution of $1280 \times 720$ and at the frame-rate of 60 fps, while the CASME-B part is at the resolution of $640 \times 480$ and at the frame-rate of 60 fps. 7 subjects participate the CASME-A experiment and 100 chips are recorded. At the same time, 12 participants are involved in CASME-B, and 95 chips are recorded. 8 emotions (amusement, sadness, disgust, surprise, contempt, fear, repression and tense) are performed in this experiment and the Facial Action Coding System is also used for labelling.

Another micro-expression database, CASMEII, are also record by Yan et.al [22], which provides more chips with resolutions. This database includes 247 samples from 26 subjects. All the chips are recorded at the resolution of $640 \times 480$ and at the frame-rate of 200fps. There are 4 kinds of labels of micro-expressions (positive, negative, surprise and other) and the the Facial Action Coding System is also used.

## 4. Conclusion

The studies of micro-expressions are increasing in recent years. Various methods are proposed for micro-expression spotting and recognition. In this paper, I survey the existing methods for micro-expression spotting and recognition. Moreover, at the same time, I present the existing databases for micro-expression recognition. There has been great progress in micro-expression spotting and recognition. However, the problems are not perfectly solved yet. For the spotting task, features that are more suitable for spotting task are needed, and better analysis of the difference between adjacent frames or of a sequence of frames, even of the whole video is needed. For the recognition task, features that are more expressive are needed, and maybe deep learning features will bring us surprise. The databases for micro-expressions are also limited, effective way to get spontaneous micro-expressions and larger scale of databases are required for further study.

## References


[1] A. K. Davison, M. H. Yap, and C. Lansley. Micro-facial movement detection using individualised baselines and histogram-based descriptors. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, pages 1864–1869. IEEE, 2015.

[2] P. Ekman. Strong evidence for universals in facial expressions: a reply to russell's mistaken critique. 1994.

[3] P. Ekman. Microexpression training tool (mett). university of california, san francisco, 2002.

[4] E. A. Haggard and K. S. Isaacs. Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy. In *Methods of research in psychotherapy*, pages 154–165. Springer, 1966.

[5] J. He, J.-F. Hu, X. Lu, and W.-S. Zheng. Multi-task mid-level feature learning for micro-expression recognition. *Pattern Recognition*, 66:44–52, 2017.

[6] X. Huang, S.-J. Wang, G. Zhao, and M. Piteikainen. Facial micro-expression recognition using spatiotemporal local binary pattern with integral projection. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 1–9, 2015.

[7] D. H. Kim, W. J. Baddar, and Y. M. Ro. Micro-expression recognition with expression-state constrained spatio-temporal feature representations. In *Proceedings of the 2016 ACM on Multimedia Conference*, pages 382–386. ACM, 2016.

[8] A. C. Le Ngo, Y.-H. Oh, R. C.-W. Phan, and J. See. Eulerian emotion magnification for subtle expression recognition. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pages 1243–1247. IEEE, 2016.

[9] A. C. Le Ngo, R. C.-W. Phan, and J. See. Spontaneous subtle expression recognition: Imbalanced databases and solutions. In *Asian conference on computer vision*, pages 33–48. Springer, 2014.

[10] X. Li, T. Pfister, X. Huang, G. Zhao, and M. Pietikäinen. A spontaneous micro-expression database: Inducement, collection and baseline. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*, pages 1–6. IEEE, 2013.

[11] X. Li, H. Xiaopeng, A. Moilanen, X. Huang, T. Pfister, G. Zhao, and M. Pietikainen. Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods. *IEEE Transactions on Affective Computing*, 2017.

[12] S.-T. Liong, J. See, R. C.-W. Phan, A. C. Le Ngo, Y.-H. Oh, and K. Wong. Subtle expression recognition using optical strain weighted features. In *Asian Conference on Computer Vision*, pages 644–657. Springer, 2014.

[13] 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*, 7(4):299–310, 2016.

[14] V. Mayya, R. M. Pai, and M. M. Pai. Combining temporal interpolation and dcnn for faster recognition of micro-expressions in video sequences. In *Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on*, pages 699–703. IEEE, 2016.

[15] A. Moilanen, G. Zhao, and M. Pietikäinen. Spotting rapid facial movements from videos using appearance-based feature difference analysis. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 1722–1727. IEEE, 2014.

[16] Y.-H. Oh, A. C. Le Ngo, J. See, S.-T. Liong, R. C.-W. Phan, and H.-C. Ling. Monogenic riesz wavelet representation for

micro-expression recognition. In *Digital Signal Processing (DSP), 2015 IEEE International Conference on*, pages 1237–1241. IEEE, 2015.

[17] D. Patel, X. Hong, and G. Zhao. Selective deep features for micro-expression recognition. In *Pattern Recognition (ICPR), 2016 23rd International Conference on*, pages 2258–2263. IEEE, 2016.

[18] D. Patel, G. Zhao, and M. Pietikäinen. Spatiotemporal integration of optical flow vectors for micro-expression detection. In *International Conference on Advanced Concepts for Intelligent Vision Systems*, pages 369–380. Springer, 2015.

[19] T. Pfister, X. Li, G. Zhao, and M. Pietikäinen. Recognising spontaneous facial micro-expressions. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 1449–1456. IEEE, 2011.

[20] S. Polikovsky, Y. Kameda, and Y. Ohta. Facial micro-expressions recognition using high speed camera and 3d-gradient descriptor. In *Crime Detection and Prevention (ICDP 2009), 3rd International Conference on*, pages 1–6. IET, 2009.

[21] S. Polikovsky, Y. Kameda, and Y. Ohta. Detection and measurement of facial micro-expression characteristics for psychological analysis. *Kameda's Publication*, 110:57–64, 2010.

[22] F. Qu, S.-J. Wang, W.-J. Yan, and X. Fu. Cas (me) 2: A database of spontaneous macro-expressions and micro-expressions. In *International Conference on Human-Computer Interaction*, pages 48–59. Springer, 2016.

[23] T. Senst, V. Eiselein, and T. Sikora. Robust local optical flow for feature tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(9):1377–1387, 2012.

[24] M. Shreve, S. Godavarthy, D. Goldgof, and S. Sarkar. Macro- and micro-expression spotting in long videos using spatio-temporal strain. In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, pages 51–56. IEEE, 2011.

[25] M. Shreve, S. Godavarthy, V. Manohar, D. Goldgof, and S. Sarkar. Towards macro-and micro-expression spotting in video using strain patterns. In *Applications of Computer Vision (WACV), 2009 Workshop on*, pages 1–6. IEEE, 2009.

[26] C. Tomasi and T. Kanade. Detection and tracking of point features. 1991.

[27] 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*, 39(1):25–43, 2014.

[28] S.-J. Wang, S. Wu, and X. Fu. A main directional maximal difference analysis for spotting micro-expressions. In *Asian Conference on Computer Vision*, pages 449–461. Springer, 2016.

[29] S.-J. Wang, S. Wu, X. Qian, J. Li, and X. Fu. A main directional maximal difference analysis for spotting facial movements from long-term videos. *Neurocomputing*, 230:382–389, 2017.

[30] 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 Transactions on Image Processing*, 24(12):6034–6047, 2015.

[31] S.-J. Wang, W.-J. Yan, T. Sun, G. Zhao, and X. Fu. Sparse tensor canonical correlation analysis for micro-expression recognition. *Neurocomputing*, 214:218–232, 2016.

[32] 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. In *Workshop at the European Conference on Computer Vision*, pages 325–338. Springer, 2014.

[33] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. 2012.

[34] Q. Wu, X. Shen, and X. Fu. The machine knows what you are hiding: an automatic micro-expression recognition system. *Affective Computing and Intelligent Interaction*, pages 152–162, 2011.

[35] Z. Xia, X. Feng, J. Peng, X. Peng, and G. Zhao. Spontaneous micro-expression spotting via geometric deformation modeling. *Computer Vision and Image Understanding*, 147:87–94, 2016.

[36] F. Xu, J. Zhang, and J. Wang. Microexpression identification and categorization using a facial dynamics map. *IEEE Transactions on Affective Computing*, 2016.

[37] W.-J. Yan, Q. Wu, J. Liang, Y.-H. Chen, and X. Fu. How fast are the leaked facial expressions: The duration of micro-expressions. *Journal of Nonverbal Behavior*, 37(4):217–230, 2013.

[38] W.-J. Yan, Q. Wu, Y.-J. Liu, S.-J. Wang, and X. Fu. Casme database: A dataset of spontaneous micro-expressions collected from neutralized faces. In *Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on*, pages 1–7. IEEE, 2013.

[39] J. Zhang, S. Shan, M. Kan, and X. Chen. Coarse-to-fine auto-encoder networks (cfan) for real-time face alignment. In *European Conference on Computer Vision*, pages 1–16. Springer, 2014.

[40] H. Zheng, X. Geng, and Z. Yang. A relaxed k-svd algorithm for spontaneous micro-expression recognition. In *Pacific Rim International Conference on Artificial Intelligence*, pages 692–699. Springer, 2016.

[41] X. Zhu, Z. Lei, X. Liu, H. Shi, and S. Z. Li. Face alignment across large poses: A 3d solution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 146–155, 2016.