LETTER

Micro-Expression Recognition by Leveraging Color Space Information*

Minghao TANG[†a)], **Yuan ZONG**[††b)], **Wenming ZHENG**[††], **Jisheng DAI**[†], **Jingang SHI**[†††], *Nonmembers,*
and **Peng SONG**[††††], *Member*

SUMMARY   Micro-expression is one type of special facial expressions and usually occurs when people try to hide their true emotions. Therefore, recognizing micro-expressions has potential values in lots of applications, e.g., lie detection. In this letter, we focus on such a meaningful topic and investigate how to make full advantage of the color information provided by the micro-expression samples to deal with the micro-expression recognition (MER) problem. To this end, we propose a novel method called color space fusion learning (CSFL) model to fuse the spatiotemporal features extracted in different color space such that the fused spatiotemporal features would be better at describing micro-expressions. To verify the effectiveness of the proposed CSFL method, extensive MER experiments on a widely-used spatiotemporal micro-expression database SMIC is conducted. The experimental results show that the CSFL can significantly improve the performance of spatiotemporal features in coping with MER tasks.
key words: spontaneous micro-expression recognition, color space, fusiong learning

1.   Introduction

Micro-expressions were first discovered by Haggard and Isaacs in 1966 [1]. Subsequently, Ekman and Friesen [2] found them again in 1969 when they were watching a video of a depressed patient and formally name them micro-expressions. Compared with ordinary facial expressions, the duration of micro-expressions is very short often lasting for between $\frac{1}{25}$ and $\frac{1}{3}$ second [2] and meanwhile the intensity of micro-expressions is extremely low. Due to these two important characteristics, it is difficult for human beings

to directly observe micro-expressions. However, detecting and recognizing micro-expressions has great values in lots of practical applications, e.g., emotion interfaces [3], clinical diagnosis [4], interrogation [5], and security field [6]. As a good example, micro-expressions in interrogation can be used for lie detection. This is because micro-expressions can be viewed as a "leakage" occuring on someone's face when this person tries to conceal his hidden emotions [7]. Hence, it is urgent to investigate automatic micro-expression recognition (MER) and develop a high-quality MER system to assist human beings in recognizing different micro-expressions.

To this end, in recent years resesrchers form various fileds, e.g., affective computing, computer vision, and pattern recognition, have been devoted to MER research and proposed lots of effective MER methods [8]–[16]. For example, in the work of [8], Pfister et al. first used the widely-used spatiotemporal descriptor, local binary pattern from three orthogonal planes (LBP-TOP) [17], to describe micro-expressions and demonstrated its effectiveness in dealing with MER tasks. Subsequently, lots of spatiotemporal temporal descriptors are developed for describing micro-expressions. For example, Wang et al. [9] considered the redundancy of LBP-TOP and design a novel spatiotemporal descriptor called LBP with six intersection points (LBP-SIP). In addition to this, other types of spatiotemporal descriptors have been also designed such as main directional mean optical (MDMO) proposed by Liu et al. [12] and facial dynamics map (FDM) developed by Xu et al. [13].

Besides the design of spatiotemporal descriptors, some important cues existing in micro-expression samples, e.g., color space information and facial local region information, are considered to cope with MER [10]–[12], [15] as well. Specifically, Wang et al. [10] invesigated whether color space information is helpful for MER by proposing a color space decomposition method. Moreover, the facial local region information has been also demonstrated to benefit MER problem. Wang et al. [10] and Liu et al. [12] respectively designed a set of region-of-interests (ROIs) which cover one or more action unit (AU) regions in the face. Their MER experiments showed that spatiotemporal descriptors together with ROIs would achieve superior performance.

In this letter, we will take full advantage of color space information in micro-expression samples to solve MER problem. Following this direction, we propose a novel feature learning method called color space fusion learning

(CSFL) to fuse the spatiotemporal features extracted from different color chanels in the color space such that the fused spatiotemporal features would be better to handle MER tasks. To evaluate the performance of the proposed CSFL, we conduct extensive experiments on a publicly available spontaneous micro-expression database (SMIC)[†] [18]. The results showed that the CSFL can effecitvely improve the performance of spaiotemporal descriptors by leveraging the beneficial information from the RGB color space in dealing with MER tasks.

2. Proposed Method

2.1 Formulation of CSFL

Suppose we have a set of training micro-expression samples and their feature matrices in three different color channels (RGB color space) are denoted by $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N] \in \mathbb{R}^{d \times N}$, $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_N] \in \mathbb{R}^{d \times N}$, and $\mathbf{Z} = [\mathbf{z}_1, \cdots, \mathbf{z}_N] \in \mathbb{R}^{d \times N}$, where d is the feature dimension and N is the sample number. Note that the feature here can be any widely-used spatiotemporal feature used for describing micro-expressions, e.g., LBP-TOP [17]. Let $\mathbf{L} \in \mathbb{R}^{c \times N}$ be the label matrix corresponding to the above feature matrices, where c denotes the number of the micro-expression categories. Each column $\mathbf{l} = [l_1, \cdots, l_c]^T$ in the label matrix is a binary-valued vector. Only the ith element is 1 whereas others are 0 if the corresponding sample of such label vector $\mathbf{l}$ belongs to the the ith micro-expression.

The basic idea of the proposed color space fusion learning (CSFL) model is simple and straightforward, i.e., learning three projection matrices for different color channels to extract the micro-expression discriminative information and meanwhile construct a compact relationship among them. Thus the extracted features by the projection matrices can be efficiently fused. Following this idea, we design the following optimization problem for CSFL model whose objective function consists of two corresponding terms:

$$\min_{\mathbf{C},\mathbf{D},\mathbf{E}} f_1(\mathbf{C},\mathbf{D},\mathbf{E}) + \lambda f_2(\mathbf{C},\mathbf{D},\mathbf{E}), \qquad (1)$$

where $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{E}$ are the target learned projection matrices of our CSFL model and λ is the trade-off parameter controlling the balance between the terms of f_1 and f_2. In what follows, we will introduce the details of these two terms.

The first term f_1 aims to make the learned projection matrices have the ability to extract the discriminative information that can be easily used for distinguishing different micro-expressions from different color channels. To this end, we use the following regression objective function to build the relationship between the feature matrices ($\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$) and the label matrix ($\mathbf{L}$).

$$f_1 = \|\mathbf{L} - \mathbf{C}^T \mathbf{X}\|_F^2 + \|\mathbf{L} - \mathbf{D}^T \mathbf{Y}\|_F^2 + \|\mathbf{L} - \mathbf{E}^T \mathbf{Z}\|_F^2$$
$$+ \mu(\|\mathbf{C}^T\|_{2,1} + \|\mathbf{D}^T\|_{2,1} + \|\mathbf{E}^T\|_{2,1}). \qquad (2)$$

[†]can be obtained from http://www.cse.oulu.fi/SMICDatabase

By minmizing Eq. (2), we are able to learn such satisfactory projection matrices $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{E}$. It is notable that we also add the $L_{2,1}$ norm of $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{E}$ for f_1 to serve as the regularization term, which would result in the spareness of several rows in the $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{E}$ and hence enable them to pick out the important features contributing to MER in each color channel. Similar with λ, μ is also a trade-off parameter which controls the sparsity of model.

As to the second term f_2, its target is to output a compact representation of the projected fefatures in each color channel such that they can be better fused after projection by using the learned $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{E}$. To achieve this goal, we design a chain regularization term as f_2 consisting of three mean projected vector difference between each other in all the color channels. By minimizing f_2, the projected features are hoped to compactly lie together. f_2 is formulated as:

$$f_2 = \|\mathbf{C}^T \bar{\mathbf{x}} - \mathbf{D}^T \bar{\mathbf{y}}\|_F^2 + \|\mathbf{C}^T \bar{\mathbf{x}} - \mathbf{E}^T \bar{\mathbf{z}}\|_F^2$$
$$+ \|\mathbf{D}^T \bar{\mathbf{y}} - \mathbf{E}^T \bar{\mathbf{z}}\|_F^2, \qquad (3)$$

where $\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$, $\bar{\mathbf{y}} = \frac{1}{N} \sum_{i=1}^N \mathbf{y}_i$, and $\bar{\mathbf{z}} = \frac{1}{N} \sum_{i=1}^N \mathbf{z}_i$.

Thus, we are able to arrive at our CSFL model whose final optimization problem can be written as:

$$\min_{\mathbf{C},\mathbf{D},\mathbf{E}} \|\mathbf{L} - \mathbf{C}^T \mathbf{X}\|_F^2 + \|\mathbf{L} - \mathbf{D}^T \mathbf{Y}\|_F^2 + \|\mathbf{L} - \mathbf{E}^T \mathbf{Z}\|_F^2$$
$$+ \lambda(\|\mathbf{C}^T \bar{\mathbf{x}} - \mathbf{D}^T \bar{\mathbf{y}}\|_F^2 + \|\mathbf{C}^T \bar{\mathbf{x}} - \mathbf{E}^T \bar{\mathbf{z}}\|_F^2 + \|\mathbf{D}^T \bar{\mathbf{y}} - \mathbf{E}^T \bar{\mathbf{z}}\|_F^2)$$
$$+ \mu(\|\mathbf{C}^T\|_{2,1} + \|\mathbf{D}^T\|_{2,1} + \|\mathbf{E}^T\|_{2,1}), \qquad (4)$$

2.2 Optimization of CSFL

The proposed CSFL can be efficiently solved by using alternated direction method (ADM) [19], i.e., iteratively updating one of three variables in CSFL while fixing others. More specifically, the updating rule for CSFL consists of the following three major steps: (1) fix $\mathbf{D}$, $\mathbf{E}$ and update $\mathbf{C}$; (2) fix $\mathbf{C}$, $\mathbf{E}$ and update $\mathbf{D}$; (3) fix $\mathbf{C}$, $\mathbf{D}$ and update $\mathbf{E}$. Note that the sub-optimization problems in all three steps are very similar and hence they can be solved by the same method. Herein, we detail the solving procedure for Step (1) as the example. The optimization problem in this step is:

$$\min_{\mathbf{C}} \|\mathbf{L} - \mathbf{C}^T \mathbf{X}\|_F^2 + \lambda(\|\mathbf{C}^T \bar{\mathbf{x}} - \mathbf{D}^T \bar{\mathbf{y}}\|_F^2$$
$$+ \|\mathbf{C}^T \bar{\mathbf{x}} - \mathbf{E}^T \bar{\mathbf{z}}\|_F^2) + \mu\|\mathbf{C}^T\|_{2,1}, \qquad (5)$$

and can be further rewritten as the following succinct formulation:

$$\min_{\mathbf{C}} \|\tilde{\mathbf{L}} - \mathbf{C}^T \tilde{\mathbf{X}}\|_F^2 + \mu\|\mathbf{C}^T\|_{2,1}, \qquad (6)$$

where $\tilde{\mathbf{L}} = [\mathbf{L}, \sqrt{\lambda}\mathbf{D}^T \bar{\mathbf{y}}, \sqrt{\lambda}\mathbf{E}^T \bar{\mathbf{z}}]$, $\tilde{\mathbf{X}} = [\mathbf{X}, \sqrt{\lambda}\bar{\mathbf{x}}, \sqrt{\lambda}\bar{\mathbf{x}}]$.

In this paper, we use an inexact augmented Lagrangian multiplier (IALM) approach [20] to solve Eq. (6). Specifically, we first convert the optimization problem in Eq. (6) to an unconstrained one by introducing an auxiliary variable $\mathbf{P}$ which equals $\mathbf{C}$, which can be formulated as:

$$\min_{\mathbf{C},\mathbf{P}} \|\tilde{\mathbf{L}} - \mathbf{P}^T \tilde{\mathbf{X}}\|_F^2 + \mu\|\mathbf{C}^T\|_{2,1}, \quad \text{s.t. } \mathbf{P} = \mathbf{C}. \qquad (7)$$

Table 1 The experimental results of of using several spatiotemporal descriptors on SMIC database.

Spatiotemoral Descriptor	Parameter Setting	Method	Accuracy / mean F1-Score
LBP-TOP [17]	$R = 1, P = 4$	Method I: Gray + No Fusion	**43.90 / 0.4487**
		Method II: RGB + No Fusion	40.24 / 0.4080
		Proposed: RGB + CSFL	**43.90 / 0.4410**
	$R = 1, P = 8$	Method I: Gray + No Fusion	39.02 / 0.3989
		Method II: RGB + No Fusion	39.63 / 0.3932
		Proposed: RGB + CSFL	**42.07 / 0.4190**
	$R = 3, P = 4$	Method I: Gray + No Fusion	36.56 / 0.3697
		Method II: RGB + No Fusion	34.76 / 0.3477
		Proposed: RGB + CSFL	**42.68 / 0.4007**
	$R = 3, P = 8$	Method I: Gray + No Fusion	39.63 / 0.3942
		Method II: RGB + No Fusion	40.24 / 0.4039
		Proposed: RGB + CSFL	**43.29 / 0.4346**
LBP-SIP [9]	$R = 1$	Method I: Gray + No Fusion	41.46 / 0.4222
		Method II: RGB + No Fusion	40.24 / 0.4097
		Proposed: RGB + CSFL	**42.68 / 0.4283**
	$R = 3$	Method I: Gray + No Fusion	39.02 / 0.3989
		Method II: RGB + No Fusion	34.76 / 0.3515
		Proposed: RGB + CSFL	**47.56 / 0.4840**
HOG-TOP [22]	$P = 4$	Method I: Gray + No Fusion	32.32 / 0.3130
		Method II: RGB + No Fusion	36.59 / 0.3609
		Proposed: RGB + CSFL	**39.02 / 0.3914**
	$P = 8$	Method I: Gray + No Fusion	38.41 / 0.3707
		Method II: RGB + No Fusion	35.98 / 0.3579
		Proposed: RGB + CSFL	**42.68 / 0.4410**

Algorithm 1 The updating rule for solving the optimization problem in Step (1).

Repeating steps 1) to 4) until convergence:
1) Fix $\mathbf{C}$ and $\mathbf{T}$, update $\mathbf{P}$:

$$\min_{\mathbf{P}} \|\tilde{\mathbf{L}} - \mathbf{P}^T \tilde{\mathbf{X}}\|_F^2 + tr[\mathbf{T}^T(\mathbf{P} - \mathbf{C})] + \frac{\kappa}{2}\|\mathbf{P} - \mathbf{C}\|_F^2.$$

The close-form solution is

$$\mathbf{P} = [\frac{2\tilde{\mathbf{X}}\tilde{\mathbf{X}}^T}{\kappa} + \mathbf{I}]^{-1}(\frac{2\tilde{\mathbf{X}}\tilde{\mathbf{L}}^T - \mathbf{T}}{\kappa} + \mathbf{C}),$$

where $\mathbf{I}$ represents the identity matrix.
2) Fix $\mathbf{P}$ and $\mathbf{T}$, update $\mathbf{C}$:

$$\min_{\mathbf{C}} \frac{\mu}{\kappa}\|\mathbf{C}^T\|_{2,1} + \frac{1}{2}\|\mathbf{C}^T - (\mathbf{P}^T + \frac{\mathbf{T}^T}{\kappa})\|_F^2,$$

according to Lemma 4.1 in [20], we will obtain that

$$c_i = \begin{cases} \frac{\|p_i + \frac{t_i}{\kappa}\|_F - \frac{\mu}{\kappa}}{\|P_i + \frac{t_i}{\kappa}\|_F}(p_i + \frac{t_i}{\kappa}) & , \text{if } \frac{\mu}{\kappa} < \|p_i + \frac{t_i}{\kappa}\|_F \\ 0 & , \text{otherwise.} \end{cases}$$

where c_i, p_i, and t_i are the ith row of $\mathbf{C}$, $\mathbf{P}$ and $\mathbf{T}$, respectively;
3) Update $\mathbf{T}$ and κ: $\mathbf{T} = \mathbf{T} + \kappa(\mathbf{P} - \mathbf{C})$, $\kappa = \min(\rho\kappa, \kappa_{max})$, where ρ is a scaled parameter;
4) Check convergence: $\|\mathbf{P} - \mathbf{C}\|_\infty < \epsilon$, where ϵ denotes the machine eplsion and is suggested to fix at 1e-8.

Then, we are able to obtain its corresponding Lagrangian function of (6) which is expressed as follows:

$$L(\mathbf{C}, \mathbf{P}, \mathbf{T}, \kappa) = \|\tilde{\mathbf{L}} - \mathbf{P}^T\tilde{\mathbf{X}}\|_F^2 + tr[\mathbf{T}^T(\mathbf{P} - \mathbf{C})]$$
$$+ \frac{\kappa}{2}\|\mathbf{P} - \mathbf{C}\|_F^2 + \mu\|\mathbf{C}^T\|_{2,1}, \quad (8)$$

where $\mathbf{T}$ is a Lagrangian multiplier, $\kappa > 0$ is a regularization parameter, and $tr(\cdot)$ denotes the trace of the matrix.

Finally, by iteratively minimizing the Lagrangian func-tion in Eq. (8) with respect to one of four variables while fixing others, the optimal $\mathbf{C}$ can be learned. The updating rule is summarized in Algorithm 1.

2.3 Fusing Spatiotemporal Features in Different Color Channels with CSFL

Once the optimal projection matrices $\hat{\mathbf{C}}$, $\hat{\mathbf{D}}$, and $\hat{\mathbf{E}}$ are learned, we can fuse the spatiotemporal features extracted in different color channels by concatenating them one by one to serve as the final micro-expression features, which can be expressed as follows:

$$\mathbf{F} = \begin{bmatrix} \hat{\mathbf{C}}^T\mathbf{X} \\ \hat{\mathbf{D}}^T\mathbf{Y} \\ \hat{\mathbf{E}}^T\mathbf{Z} \end{bmatrix}, \quad (9)$$

where $\mathbf{F}$ is the final fused feature.

3. Experiments

3.1 Databases and Experiment Setting

In this section, we conduct extensive experiments to evalu-ate the proposed CSFL method. A public available spon-taneous micro-expression database SMIC [18] is adopted. SMIC has three datasets (HS, VIS, and NIR) recorded by three different equipments, i.e., high-speed camera, visual camera, and near-infrared camera. All the samples in SMIC are divided into three different micro-expressions (*Positive*, *Negative*, and *Surprise*). We use the HS dataset comprising 164 samples belonging to 16 subjects for the experiments. The face images in the micro-expression video clips are all cropped and then transformed to 112×112 pixels. The tem-poral interpolation model (TIM) [21] is used to normalize

the frame number of the video clips to 16.

In the experiments, leave-one-subject-out (LOSO) strategy is chosen to calculate the *Accuracy* and *mean F*1-*S core* to report the performance of a MER method. To see whether the proposed CSFL is an effective methd to enable the spaitotemporal descriptors to better describe microexpressions, we choose three representative spatiotemporal descriptors including LBP-TOP [17], LBP-SIP [9], and HOG-TOP [22] with various parameter settings together with 8×8 spatial divison grid to extract the features in each channel of RGB color space for CSFL to conduct the MER experiments under the LOSO strategy. As to the parameter settings of spatiotemporal descriptors, for LBP-TOP, the neighboring radius R and the number of the neighboring points P are set as ones of $\{1, 3\}$ and $\{4, 8\}$, respectively. Similarly, the neighboring radius R of LBP-SIP is also set as 4 and 8, respectively. For HOG-TOP, the number of their feature vector bins p is fixed at 4 and 8, respectively. Linear support vector machine (SVM) with fixed parameter $C = 1$ is served as the classiier thoughout the experiments. The trade-off parameters of CSFL, i.e., λ and μ are determined by the cross-validation method. In addition, to offer the comparison, we also conduct the experiments of using the above three spatiotemporal descriptors in the gray color space (Method I) and directly concatenated in the RGB color space without consideration of color information fusion (Method II), respectively. Note that the major difference between Method I and Method II is that Method I only considers to extract the spatiotemporal descriptors from the gray space of micro-expression samples, while Method II considers the color space of micro-expression samples and three color channel information are used for extracting micro-expression features.

3.2 Results and Analysis

The experimental results in terms of both *Accuracy* and *mean F*1-*S core* are given in Table 1. From the results, it is clear to see that the proposed CSFL performs best among all three methods in the experiments of using any one of three spatiotemporal descriptors with fixed parameters as the micro-expression feature. It indicates that the proposed CSFL can effectively make use of the color information and leverage the micro-expression discriminative information in the color space to enhance the performance of spatiotemporal descriptors in dealing with MER tasks. Moreover, although the RGB color space offers more beneficial information for MER than the Gray one, the spatiotemporal descriptors extracted from the RGB color space without fusion operation (Method II) seems mostly perform worse than Gray space (Method I), which can be clearly seen from the results of LBP-TOP (R=1, P=4), LBP-TOP (R=3, P=4), LBP-SIP (P=4), LBP-SIP (P=8) and HOG-TOP (P=8). We think this may attribute to the fact that compared with the Gray space, more redundant information also exist in the RGB color space, which degrades the performance of the spatiotemporal descriptors. However, we can find that the pro-

posed CSFL is able to overcome the shortcoming of Method II (direct concatenation) and benefit from the beneficial information for MER in the RGB color space.

4. Conclusion

In this letter, we have investigated the micro-expression recognition (MER) problem and proposed a novel feature learning method called color space fusing learning (CSFL) model. By resorting to CSFL, existing spatiotemporal descriptors used for describing micro-expressions can benefit from the color space information. The spaiotemporal features corresponding to different color chanels in the color space would be fused togehter resulting in a better feature representation for describing micro-expressions. Meanwhile, we have also given an efficent method solve the optimization problem of the CSFL model. Extensive experiments on SMIC micro-expression database has demonstrated the effectiveness and superior performance in MER.

References

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

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

[3] S. Wang, Z. Liu, S. Lv, Y. Lv, G. Wu, P. Peng, F. Chen, and X. Wang, "A natural visible and infrared facial expression database for expression recognition and emotion inference," IEEE Trans. Multimedia, vol.12, no.7, pp.682–691, 2010.

[4] M. Frank, M. Herbasz, K. Sinuk, A. Keller, and C. Nolan, "I see how you feel: Training laypeople and professionals to recognize fleeting emotions," Annual Meeting of the International Communication Association, Sheraton New York, New York City, 2009.

[5] M.G. Frank, C.J. Maccario, and V. Govindaraju, "Behavior and security," Protecting Airline Passengers in the Age of Terrorism, Greenwood Pub Group, Santa Barbara, California, pp.86–106, 2009.

[6] M. O'Sullivan, M.G. Frank, C.M. Hurley, and J. Tiwana, "Police lie detection accuracy: The effect of lie scenario," Law and Human Behavior, vol.33, no.6, pp.530–538, 2009.

[7] P. Ekman, Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage, revised edition, WW Norton & Company, 2009.

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

[9] 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, vol.9003, pp.525–537, Springer, 2014.

[10] 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.

[11] P. Lu, W. Zheng, Z. Wang, Q. Li, Y. Zong, M. Xin, and L. Wu, "Micro-expression recognition by regression model and group sparse spatio-temporal feature learning," IEICE Trans. Inf. & Syst., vol.99, no.6, pp.1694–1697, 2016.

[12] 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, vol.7, no.4, pp.299–310, 2016.

[13] F. Xu, J. Zhang, and J.Z. Wang, "Microexpression identification and categorization using a facial dynamics map," IEEE Transactions on

Affective Computing, vol.8, no.2, pp.254–267, 2017.

[14] Y. Zong, X. Huang, W. Zheng, Z. Cui, and G. Zhao, "Learning a target sample re-generator for cross-database micro-expression recognition," Proceedings of the 2017 ACM on Multimedia Conference, pp.872–880, ACM, 2017.

[15] Y. Zong, X. Huang, W. Zheng, Z. Cui, and G. Zhao, "Learning from hierarchical spatiotemporal descriptors for micro-expression recognition," IEEE Trans. Multimedia, vol.20, no.11, pp.3160–3172, 2018.

[16] Y. Zong, W. Zheng, X. Huang, J. Shi, Z. Cui, and G. Zhao, "Domain regeneration for cross-database micro-expression recognition," IEEE Trans. Image Process., vol.27, no.5, pp.2484–2498, 2018.

[17] G. Zhao and M. Pietikäinen, "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.

[18] X. Li, T. Pfister, X. Huang, G. Zhao, and M. Pietikäinen, "A spontaneous micro-expression database: Inducement, collection and baseline," 2013 10th IEEE International Conference and Workshops on Automatic face and gesture recognition (fg), pp.1–6, IEEE, 2013.

[19] W. Zheng, "Multichannel eeg-based emotion recognition via group sparse canonical correlation analysis," IEEE Transactions on Cognitive and Developmental Systems, vol.9, no.3, pp.281–290, 2017.

[20] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," IEEE Trans. Pattern Anal. Mach. Intell., vol.35, no.1, pp.171–184, 2013.

[21] Z. Zhou, G. Zhao, and M. Pietikäinen, "Towards a practical lipreading system," 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.137–144, IEEE, 2011.

[22] X. Li, X. Hong, A. Moilanen, X. Huang, T. Pfister, G. Zhao, and M. Pietikäinen, "Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods," IEEE Transactions on Affective Computing, vol.9, no.4, pp.563–577, 2018.