Micro-Expression Recognition based on 2D Gabor Filter and Sparse Representation

Hao Zheng[1,2,3]

[1] Key Laboratory of Trusted Cloud Computing and Big Data Analysis, School of Information Engineering, Nanjing XiaoZhuang University, Nanjing, China
[2] MOE Key Laboratory of Computer Network and Information Integration, School of Computer Science and Engineering, Southeast University, Nanjing, China
[3] State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China

E-mail: zhh710@163.com

Abstract. Micro-expression recognition is always a challenging problem for its quick facial expression. This paper proposed a novel method named 2D Gabor filter and Sparse Representation (2DGSR) to deal with the recognition of micro-expression. In our method, 2D Gabor filter is used for enhancing the robustness of the variations due to increasing the discrimination power. While the sparse representation is applied to deal with the subtlety, and cast recognition as a sparse approximation problem. We compare our method to other popular methods in three spontaneous micro-expression recognition databases. The results show that our method has more excellent performance than other methods.

1. Introduction

Expression recognition has been widely studied in the pattern recognition. However, micro-expressions recognition has three characteristics with poor performance. Firstly, micro-expressions are rapid facial movements with less than 0.55 second. Secondly, the intensity of the micro-expressions is very low according to facial muscles' movement. Thirdly, micro-expressions involve a fragment of the facial region. Therefore micro-expressions are difficult to recognize in the aspect of its short duration and occurrence. The flourishing of expression recognition methods depend on the well-established facial expression databases, such as JAFFE, CK+, BU-3DFE, Oulu-CASIA and Multi-PIE. However, there are few well-established micro-expression databases due to the difficulty of eliciting micro-expressions. In current literature, there are only three spontaneous micro-expression databases, i.e. SMIC [1], CASME [2], CASME II [3], while there are some works to date on automatic recognition of spontaneous micro-expressions.

Among all the available work, Li published the SMIC dataset and released a baseline performance of up to 48.78% accuracy for 3 classes, adopting LBP-TOP for feature extraction and SVM polynomial kernel for classification with cross validation method. Yan also reported a baseline performance of up to 63.41% accuracy for a 5-class classification task on CASME II database, adopting LBP-TOP and SVM for feature extraction and classification respectively. Then Wang et al. [4] proposed LBP-SIP method, which incorporated a multi-resolution Gaussian pyramid by concatenating the feature histograms of all four pyramid levels, and obtained the best recognition accuracy of 67.21%. In addition, Polikovsky et al. [5] used 3D-gradient descriptor for micro-expression recognition. Wang et al.[6] treated a micro-expression gray-scale video clip as a 3rd-order tensor and used Discriminant Tensor Subspace Analysis and Extreme Learning Machine to recognize micro-expressions.

But above methods have not achieved ideal results, we need to further improve the performance. 2D Gabor functions and wavelets have achieved impressive results when used for pattern recognitions. Many researchers have used Gabor wavelets in order to improve recognition performance [7]. Gabor wavelets show desirable characteristics of spatial locality and orientation selectivity, and are localized in the space and frequency domains [8]. In addition, sparse representation has been proved to improve the robustness of recognition, and wildly used in face recognition, expression recognition, etc. Motivated by above idea, in this paper, a novel method named 2D Gabor filter and sparse representation (2DGSR) is proposed to improve robustness for spontaneous micro-expression recognition. In the proposed method, 2D Gabor filter is used for enhancing the robustness of the variations due to increasing the discrimination power. While the sparse representation is applied to deal with the subtlety, and cast recognition as a sparse approximation problem. Our extensive experiments on spontaneous micro-expression databases show that the proposed 2DGSR method has very competitive performance with state-of-the-arts.

The rest of the paper is organized as follows. Section 2 presents the proposed 2DGSR method. Section 3 conducts experiments. Section 4 concludes the paper.

2. 2D Gabor filter and sparse representation (2DGSR)

The Gabor transform adds the Gaussian window to Fourier transform in order to realize the local analysis in spatial domain and spatial frequency domain. 2D Gabor filter can achieve optimal joint localization properties in the spatial domain and in the spatial frequency domain. So 2D Gabor filter is introduced in the proposed method.

2.1. 2D Gabor filter

When 2D Gabor filters are used for micro-expression recognition, the face image should be filtered by 2D Gabor filters firstly, then feature extraction is employed, finally downsample can be performed. The 2D Gabor filtering for face image is that an image $I(x,y)$ is convolved with a 2-D Gabor function $g(x,y)$ to obtain a Gabor feature image $u(x,y)$ as follows:

$$u(x,y) = \iint_{\Phi} I(\xi,\eta) g(x-\xi, y-\eta) d\xi d\eta \tag{1}$$

where $(x,y)\in\Phi$, Φ denotes the set of image points.

So the 2D Gabor function $\psi(x,y)$ can be obtained by a product of a Gaussian and a cosine function as following:

$$\psi_{\lambda,\theta,\phi}(x,y) = e^{-((x'^2+\gamma^2 y'^2)/2\sigma^2)} \cos(2\pi \frac{x'}{\lambda} + \phi) \tag{2}$$

where $x' = x\cos\theta + y\sin\theta, y' = -x\sin\theta + y\cos\theta, \sigma = 0.56\lambda, \gamma = 0.5$. σ represents the scale parameter, the small σ means high spatial resolution, the image filtering coefficients reflect local properties in fine scale, while the large σ means low spatial resolution , the coefficients reflect local properties in coarse scale. θ specifies the orientations of 2D Gabor filter, λ denotes the wavelength of the cosine factor, γ is the spatial aspect ratio and specifies the ellipticity of Gaussian factor, ϕ specifies the phase offset of the cosine factor.

IOP Conf. Series: Journal of Physics: Conf. Series **787** (2017) 012013 doi:10.1088/1742-6596/787/1/012013

Suppose $I(x,y)$ is the gray level distribution of an image. The Gabor representation is generated from convoluting the image $I(x,y)$ and the Gabor kernel $\psi_{\lambda,\theta,\phi}(x,y)$. The Gabor representation is defined as：

$$\Theta_{\lambda,\theta,\phi}(x,y) = I(z) * \psi_{\lambda,\theta,\phi}(x,y) \tag{3}$$

To reduce the space dimension, we downsample each $\Theta_{\lambda,\theta,\phi}(x,y)$ by a factor ρ and normalize it to zero mean and unit variance.

$$\Theta^{\rho}{}_{\lambda,\theta,\phi}(x,y) = \frac{\Theta_{\lambda,\theta,\phi}(x,y) - m_{\lambda,\theta,\phi}}{\sigma_{\lambda,\theta,\phi}} \tag{4}$$

where $m_{\lambda,\theta,\phi}$, $\sigma_{\lambda,\theta,\phi}$ denote the mean and the standard deviation of the sums, respectively. So the new Gabor representation $\Theta^{\rho}{}_{\lambda,\theta,\phi}(x,y)$ is obtained. Then we concatenate the 40 Gabor representations together and get an augmented feature vector

$$\tau^{\rho}{}_{\Xi,\Xi} = \underset{\mu=0}{\overset{7}{\Xi}}\underset{v=0}{\overset{4}{\Xi}}\Theta^{\rho}{}_{\lambda,\theta,\phi} \tag{5}$$

Where Ξ denotes the junction symbol. So the argument Gabor feature vector which encompasses all the elements of the Gabor wavelets representation set is obtained.

2.2. 2D Gabor filter and sparse representation

The augmented Gabor face feature vector τ, which is a local feature descriptor, can not only enhance the face feature but also tolerate face image local deformation to some extent. Suppose we denote

$$T(A) = [T(A_1)\ T(A_2)\ \cdots\ T(A_K)], T(A_i) = [\tau(s_{i,1})\ \ \tau,(s_{i,2})\ \cdots\ \tau(s_{i,n_i})],$$
$$x = [x_1; x_2; \cdots; x_m] \tag{6}$$

The 2D Gabor feature sparse representation is as following:

$$\tau(y) = T(A_1)x_1 + T(A_2)x_2 + \cdots + T(A_m)x_2 = T(A)x \tag{7}$$

So the object function of 2DGSR is

$$< \hat{x} >= \arg\min_{\hat{x}}(\| \tau(y) \text{-} T(A)x \|_2^2 + \lambda \|x\|_1) \tag{8}$$

By using 2D Gabor features for micro-expression recognition, the feature dictionary A will be transformed into 2D Gabor feature dictionary.

Thus the proposed 2DGSR method is summarized in Algorithm 1.

Algorithm 1 (**2D Gabor filter and Sparse Representation** (2DGSR))

Step 1 obtain 2D Gabor feature dictionary $T(A)$

Step 2:Normalize the columns of $T(A)$

Step 3: Solve the l_1 minimization problem:

$$< \hat{x} >= \arg\min_{\hat{x}}(\| \tau(y) \text{-} T(A)x \|_2^2 + \lambda \|x\|_1)$$

Step4:Compute label of the test image:

$$r_i(y) = \| y - T(A)\delta_i(\hat{x}) - \hat{\varepsilon} \|_2,\ for\ i = 1,...,k.$$

3. Experiments and discussions

In this section, we compare our proposed 2D Gabor feature with two baseline features in micro-expression recognition, i.e., LBP-TOP[9],[10] and HOOF[11]. We use two versions of the HOOF feature such as HOOF-whole, HOOF-ROIs. HOOF-whole is the original HOOF feature applied to the entire facial region, four to ten bins are used in the HOOF histogram. HOOF-ROIs is a combinatorial HOOF feature which created by applying the HOOF feature in each of the 36 ROIs, four to ten bins are used in each ROI's HOOF histogram. We test our proposed method on three micro-expression database, including SMIC, CASE, and CASME II to verify its effectiveness for micro-expression recognition. These databases consist of spontaneous micro-expressions which appear in real life. In all experiments we applied the SRC to classify and employed Leave-one-subject-out (LOSO) cross validation.

3.1 CASME database:

The Chinese Academy of Sciences Micro-Expression (CASME) database includes 195 spontaneous facial micro-expression recorded by two different 60 fps cameras. Two expert coders were recruited in the work to code the duration and AU combination in these micro-expressions. The selected micro-expressions either have a total duration less than 500ms or onset duration less than 250ms. These samples are coded with the onset, apex and offset frames, furthermore tagged with AUs. In this database, micro-expressions are classified into 7 categories (happiness, surprise, disgust, fear, sadness, repression and tense).

CASME database is divided into two classes: Class A and Class B. The samples in Class A were recorded by the BenQ M31 consumer, camera with 60fps, with the resolution set at 1280*720 pixels. The participants were recorded in natural light. The samples in Class B were recorded by the Point Grey GRAS-03K2C industrial camera with 60 fps, with the resolution set to 640*480 pixels. The participants were recorded in a room with two LED lights. In the experiments, we merged the 7 categories into 4 classes: Positive, Negative, Surprise, Others. We selected 97 samples from Class B.

From Table 1, it can be seen that 2DGSR outperform than other methods, especially 2DGSR is almost 20% higher than HOOF-whole feature extraction. This is because 2DGSR makes full use of gabor feature information about spatial and orientation, and enhance the robustness of recognition. Meanwhile from Table 2, we can see that the results of 2DGSR had a good recognition rates in all four classes.

Table 1: The best recognition rates of four features for micro-expression recognition in CASME database.

Feature	LBP-TOP	HOOF-whole	HOOF-ROIs	2DGSR
Accuracy	66.34%	51.89%	57.92%	**71.19%**

Table 2: Confusion matrix of micro-expression recognition obtained by 2DGSR on CASME database

	Positive	Negative	Surprise	Others
Positive	48.6%	4.1%	3.6%	6.6%
Negative	5.4%	57.6%	5.2%	13.2%
Surprise	3.9%	4.8%	69.1%	2.5%
Others	42.1%	33.5%	22.1%	77.7%

3.2 CASME II database

The CASME II database includes 246 spontaneous facial micro-expressions recorded by a 200 fps camera. These samples were selected from more than 2,500 facial expressions. The selected micro-expressions in this database either had a total duration less than 500 ms or an onset duration (time from onset frame to apex frame) less than 250 ms. In CASME II, the samples in which the facial feature points in the first frame cannot be correctly detected by using the DRMF method. We used 236 samples from 26 subjects, categorized into four classes: positive (31 samples), Negative(65 samples), Surprise(21sample), and Others(119 samples).

Table 3 show that 2DGSR outperformed other three methods, and also achieved the best recognition rates in CASME II database. We can see that the recognition rates in CASME II database is worse than those in CASME database, because compared with CASME, the database is improved in increased sample size, fixed illumination, and higher resolution(both temporal and spatial).

Table 3. The best recognition rates of four features for micro-expression recognition in CASME II database.

Feature	LBP-TOP	HOOF-whole	HOOF-ROIs	2DGSR
Accuracy	59.12%	44.36%	54.65%	**64.88%**

4. Conclusion

In the paper, we apply the 2D Gabor features and sparse representation to micro-expression recognition. 2D Gabor feature and sparse representation make our robust to its subtlety. Experimental results show that our method has better performances than other three popular methods. Although our method performs well, recognition speed is not fit for real time. In the future, we will improve the method to speed up Gabor wavelet transformations.

References

[1] Li, X., Pster, T., Huang, X., Zhao, G., Pietikainen, M.: A spontaneous micro-expression database: inducement, collection and baseline. In: 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), pp.1-6 (2013)
[2] Yan, W.J., Wu, Q., Liu, Y.J., Wang, S.J., Fu, X.: CASME database: a dataset of spontaneous micro-expressions collected from neutralized faces. In: 10th IEEE In-ternational Conference and Workshops on Automatic Face and Gesture Recognition (FG), pp. 1-7 (2013)
[3] Yan, W.J., Li, X., Wang, S.J., Zhao, G., Liu, Y.J., Chen, Y.H., Fu, X.: CASME II: an improved spontaneous micro-expression database and the baseline evaluation. PLoS One 9(1), e86041 (2014)
[4] Wang Y., See J, Phan RC-W, Oh Y-H.: Efficient Spatio-Temporal Local Binary Patterns for Spontaneous Facial Micro-Expression Recognition. PLoS ONE 10(5): e0124674 (2015)
[5] Polikovsky S., Kameda Y., and Ohta Y.: Facial micro-expressions recognition using high speed camera and 3D-gradient descriptor. In 3rd International Conference on Crime Detection and Prevention. IET, pp. 1-6 (2009)
[6] S.J. Wang, H.L. Chen, W.J. Yan, Y.H. Chen, and X. Fu: Face recognition and micro-expression based on discriminant tensor subspace analysis plus extreme learning machine. Neural Processing Letters (2013)

[7] Ilkka Autio, Tapio Elomaa, Flexible View Recognition for Indoor Navigation Based On Gabor Flters and Support Vector Machines, Pattern Recognition, vol. 36, 2003, pp.2769-2779.
[8] Chengjun Liu, and Harry Wechsler, Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition, IEEE Transactions on Image Processing, vol. 11, 2002, pp.467-476.
[9] T. Pfister, X. Li, G. Zhao, and M. Pietik¨ainen, "Recognising spontaneous facial micro-expressions," in 2011 IEEE International Conference on Computer Vision (ICCV). IEEE, 2011, pp. 1449–1456.
[10] G. Zhao and M. Pietik¨ainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 915–928, 2007.
[11] R. Chaudhry, A. Ravichandran, G. Hager, and R. Vidal, "Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions,"in 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2009, pp. 1932–1939.