Microexpression Recognition based on Improved Robust Principal Component Analysis and Texture Feature Extraction

Dong Xiaochen
School of Computer Science and Technology
Qingdao University
Qingdao, China
dxcchen@163.com

Zhao Zhigang
School of Computer Science and Technology
Qingdao University
Qingdao, China
zhaolhx@263.net

Li Qiang
School of Computer Science and Technology
Qingdao University
Qingdao, China
lq.sxt@163.com

ABSTRACT

Micro-expression can reflect the emotional information that humans can hardly conceal. There are three main characteristics: short duration, low intensity and local movement. From these characteristics it can be seen that the motion of the micro-expression is sparse. For sparse micro-expression movement, a robust principal component analysis (RPCA) was proposed to extract subtle micro-expression motion information. Using improved Edge Direction Histogram (EOH) algorithm and Binary Gradient Contours (BGC) algorithm to extract local texture features can solve the problem of spatio-temporal domain and obtain high recognition accuracy. Experiments on the SMIC database show that the proposed algorithm has better performance.

CCS Concepts

Computing methodologies→Computer graphics→Image manipulation→Image processing

Keywords

Robust principal component analysis; EOH algorithm; BGC algorithm; recognition accuracy

1. INTRODUCTION

Expression is the most intuitive way for humans to express emotions. Although it is not expressed in words, it fully demonstrates human inner activities. In 1872, Darwin's article "The Expression of Humans and Animals" made people start to pay attention to the expression field[1].In the real world, facial expressions are rich and varied. Initially, people only paid attention to macro-expressions [2]. In fact, there is more than just this type of expression, and there is also a short duration, low-level expression that is called a micro-expression [3]. Micro-expressions are emotions that humans try to cover up. Their duration is only 1/25s to 1/5s. It is extremely short-lived and the change rate is very low. The identification of micro-expressions is difficulty. The discovery of micro-expressions is tortuous [4].

ICCIP 2018, November 2–4, 2018, Qingdao, China

ACM ISBN 978-1-4503-6534-5/18/11…$15.00
http://doi.org/10.1145/3290420.3290421

Haggard and Isaacs [5] discovered a special expression long ago, but due to various limitations of the research hierarchy at the time, they did not continue to study. Later, in 1969, Paul Ekman and Wallace V. Friesen [6] inadvertently discovered an extremely short-lived painful expression when they studied the patient's expression video. They were soon covered by other expressions. Therefore, they named the expression as a micro-expression. Soon, Ekman and Frisen designed the BART (Brief Affect Recession Test) [7] experiment and found that the micro-expression has a certain relationship with lying. In addition, it has done a lot of different experiments and the conclusion is the same as this conclusion. In 2002, they designed the JACBART (Japanese and Caucasian Brief Affect Recognition Test) experiment [8]. Frank et al.[9] modeled the experimental results to analyze the relationship between lies and micro-expressions. In 2006, Russell, Elvina, and Mary et al.[10] linked clinical research with microexpressions. In 2009, Endres and Laidlaw studied the differences in student's micro-expression recognition [11]. In 2008, Russell experimented with the effects of micro-expression recognition on psychotic patients [12]. Later, people studied the relationship between micro-expression recognition ability and the ability to identify lies [13], and the relationship between micro-expression recognition ability and stimuli [14]. In order to improve people's performance in recognizing micro-expressions, Ekman developed a micro-expression training tool (METT) that enables people to better recognize micro-expressions. In order to better apply micro-expressions as clues, discovering deception in practice, computer scientists try to train computers to automatically recognize micro-expressions.

In recent years, micro-expression research has gradually become a hot topic. Wu Qi, the Chinese Academy of Sciences, has studied the potential applications of micro-expressions. There are some shortcomings in the micro-expression database, there are SMIC and CASME/CASME II, and few other databases automatically identify micro-expressions. The earliest research on micro-expression recognition was made by Pfister et al. This research was extremely representative and laid the foundation for the future micro-expression recognition work. This method uses the Active Shape Model (ASM) to locate the key points on the face. When dealing with dynamic textures, LBP-TOP is still very popular in feature extraction, such as texture recognition, motion recognition and facial expression recognition. LBP-TOP feature is to increase the robustness and rotation-invariant descriptors of view-based changes to textures.

Polkovsky et al. [15] use three-dimensional gradient descriptors for micro-expression recognition. Wang et al. [16] processed gray-scale micro-expression video clips as a third-order tensor and tensor space analysis (DTSA) and extreme learning machine (ELM) micro-expressions. Pfister et al. [17] used a time-interpolation model [18] based on the Laplacian matrix to normalize the number of frames of the micro-expression video segment. In addition, LBP-TOP [19] is used to extract the appearance of micro-expression movements for multi-core learning. Sherve et al. [20] used optical flow to locate and capture micro-expressions. In 2011, Xiaolan Fu [21] and others also developed their own micro-expression database and used the Gabor filter in extracting the features of the micro-expression. Principal component analysis is a kind of data analysis method. It largely reduces high-dimensional data to low dimensions. It has good application in image processing but lacks robustness. This paper proposes robust principal component analysis and it has a good effect. The algorithms for extracting texture features include EOH algorithm and BGC algorithm. In this paper, by improving these two algorithms, it is solved that the problem of micro-expression about space-time domain, the recognition accuracy is improved, and the texture feature extraction effect is better.

2. ROBUST PRINCIPAL COMPONENT ANALYSIS

In the process of feature extraction of images, too many feature dimensions extracted often lead to over complication of feature matching, consuming system resources, and having to adopt feature reduction techniques. The so-called feature dimension reduction uses a low latitude feature to represent high latitude. There are generally two types of methods for feature dimension reduction: feature selection and feature extraction. The feature selection selects one of the subsets from the high latitude as a new feature; and the feature extraction refers to mapping the feature of high latitude to a low latitude as a new feature through a function. A commonly used feature extraction method is principal component analysis.

2.1 Principal Component Analysis

The essence of principal component analysis is to linearly transform and map the original features into low-dimensional spaces with the best possible representation of the original features. The principal component analysis problem can be described as: for n samples in d-dimensional space, consider how they can be better represented in low rank spaces. This can be answered by first assuming that in a 0-dimensional space, a point is used to represent these features and then extended to 1 dimension up to m(m<d) dimensions. PCA is arguably the most widely used statistical tool for data analysis and dimensionality reduction today. Unfortunately, gross errors are now ubiquitous in modern applications such as image processing, web data analysis, and bioinformatics, where some measurements may be arbitrarily corrupted (due to occlusions, malicious tampering, or sensor failures) or simply irrelevant to the low-dimensional structure we seek to identify. A number of natural approaches to robustifying PCA have been explored and proposed in the literature over several decades.

2.2 Robust Principal Component Analysis Model

Robust principal component analysis [22], there is a data matrix, which is the superposition of a low-rank component and a sparse component. Can we recover each component individually? We prove that under some suitable assumptions, it is possible to recover both the low-rank and the sparse components exactly by solving a very convenient convex program called Principal Component Pursuit; among all feasible decompositions, simply minimize a weighted combination of the nuclear norm and of the ℓ_1 norm.

Given a sequence of micro-expressions $V \in R^{h\times w\times f}$, h and w are the height and width of the micro-expression sequence, f is the number of frames, and its data matrix is $D \in R^{h\times w\times f}$ .Since the change of the micro-expression is subtle, the data matrix D can be divided into two parts, namely, static and dynamic, and the dynamic representation of subtle changes in micro-expressions is defined as follows:

$$D = A + E \tag{1}$$

Among them, *A* is the low rank subspace, *E* represents the subtle motion information of the micro-expression.In the following, we will use the improved edge direction histogram (EOH) algorithm or the modified binary contour (BGC) algorithm to extract the dynamic texture features with subtle changes in the image. This may be formulated as follows:

$$\min_{A,E} rank(A) + \| E \|_0 \quad \text{subject to} \quad D = A + E \tag{2}$$

When the constraints are the sum of the minimization of the optimization of the low rank matrix *A* and the zero norm of the matrix *E*, this problem is non-convex and can be transformed into the following convex optimization problem for ease of solution:

$$\min_{A,E} \| A \|_* + \lambda \| E \|_1 \quad \text{subject to} \quad D = A + E \tag{3}$$

Equation 3 introduces the weight parameter λ (>0), which is the kernel norm of the low rank matrix *A* and the sum of the product of λ and the 1 norm of the matrix *E*. Lin et al.[23] proposed the method of augmented Lagrange multipliers (ALM) and added the constraint to equation 3 using augmented Lagrangian multiplier method. The augmented Lagrangian function can be defined as follows:

$$L(A,E,Y,\mu) = \| A \|_* + \lambda \| E \|_1 + \langle Y, D - A - E \rangle + \frac{\mu}{2}$$

$$\| D - A - E \|_F^2 \tag{4}$$

Where *A* is a Lagrangian multiplier, μ is positive. Here, we chose inexact ALM to extract the subtle facial motion information.

2.3 Iterative Threshold Algorithm

The inexact ALM method can solve equation (4) and update *A, E, Y,* alternately. The specific algorithm steps are as follows:

Input: data matrix *D*, parameters λ , maximum number of iterations *T*

Initialize: $Y = 0$, $A = 0$, $E = 0$, $\mu = 10^{-6}$, $T = 10^6$

$\varepsilon = 10^{-8}, \partial = 10^{-6}$

*a)*Fix other parameters to update *A*, that is when

$E = E_k, Y = Y_k, \mu = \mu_k$:

$$A_{k+1} = \arg\min_A L(A, E_k, Y_k, \mu_k) = \arg\min_A \frac{1}{\mu_k} \| A \|_* + \frac{1}{2}$$

b)Fix other parameters to update *E*, that is when $A = A_{k+1}$, $Y = Y_{k+1}$:

$$E_{k+1} = \arg\min_A L(A_{k+1}, E, Y_k, \mu_k) = \arg\min_E \frac{\lambda}{\mu_k} \| E \|_1 + \frac{1}{2}$$

$$\| E - \left(\frac{Y_k}{\mu_k} + D - A_{k+1} \right) \|_F^2$$

c)Fix other parameters to update *Y*,that is when $A = A_{k+1}$, $E = E_{k+1}$:

$Y_{k+1} = Y_k + \partial_k \left(D - A_{k+1} - E_{k+1} \right)$,Where *A* is the step size and is a positive number less than 1.

d)When the number of iterations is reached, the algorithm ends; or when $\| D - A - E \| < \varepsilon$, the algorithm ends.

2.4 Application

Robust Principle Component Analysis (RPCA) arises naturally in time-varying source separations such as video foreground-background separation. So there is a compressive online robust PCA with prior information for recursively separating sequences of frames into sparse and low-rank components from a small set of measurements [24]. In addition, there is a vision-based application for recognizing the chord being played by a guitarist to help guitar learners to practice by themselves [25].

3. TEXTURE EATURE EXTRACTION

The edge describes the shape of the image and represents the part where the image changes. The Edge Direction Histogram (EOH)[26] has a good effect on the subtle changes of the micro-expression image, and has good robustness to light changes, it has better performance on the edge of the image and texture features.Binary Gradient Contours (BGC) is a feature description operator based on the texture frame, which has a good effect on the subtle changes of micro-expressions. The improved edge direction histogram and binary gradient profile can effectively describe all the texture features of the micro-expression video sequence in the space-time domain.

3.1 Edge Direction Histogram (EOH)

In the field of computer vision, sobel operator has simple features and high efficiency in edge detection. Therefore, Sobel operator is used to perform edge detection on the image, and the edge detection result is represented by binary bits, and $\psi_k(x, y)$ is the k-th binary number. The window R is defined in the image and the "integral image"[27] is introduced. The integral image can be expressed as:

$$E_k(R) = \sum_{(x,y)\in R} \psi_k(x, y) \tag{5}$$

According to Formula 5, a set of features A can be defined:

$$A_{k1,k2}(R) = \frac{E_{k1}(R) + \varepsilon}{E_{k2}(R) + \varepsilon} \tag{6}$$

The parameter ε plays a smoothing role for the numerator and the denominator.

3.2 Binary Gradient Contours (BGC)

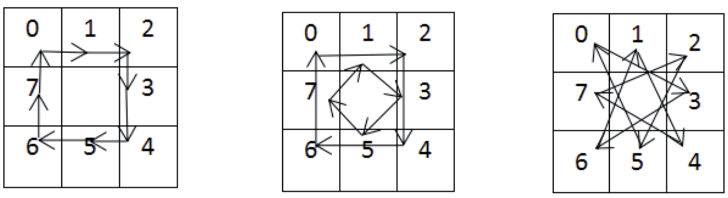
The main idea of the binary gradient contours is to divide a gray scale image into a window with a size of 8 neighborhoods, and then calculate a binary gradient in a certain order. The texture feature can be represented as a binary number. The specific gray distribution is as shown in Fig.1.

I_0	I_1	I_2
I_7	I_c	I_3
I_6	I_5	I_4

Figure 1. Grayscale Distribution

I_c represents the gray value of the center pixel, and I_p $(p = 0,...,7)$represents the pixel value of the neighborhood.

The BGC uses a circular closed path to calculate gradients, and then performs binary calculations. The current BGC mainly has three ways to calculate binary gradients: single-loop calculations, double-loop calculations, and triple-loop calculations. Three ways as shown in Fig.2.



(2a) (2b) (2c)

Figure 2. BGC Three Ring Closed Paths (2a)single-loop (2b)double-loop (2c)triple-loop

Fig.(2a) shows the single-loop binary gradient contours calculation process. The closed path is(0,1,2,3,4,5,6,7),pixel pairs are{(0,1), (1,2),(2,3),(3,4),(4,5),(5,6),(6,7),(7,0)},then $b_1 = \{s(I_1 - I_0), s(I_2 - I_1), s(I_3 - I_2), s(I_4 - I_3), s(I_5 - I_4), s(I_6 - I_5), s(I_7 - I_6), s(I_0 - I_7)\}$.

Where $S(x)$ is a symbolic function:

$$S(x) = \begin{cases} 1, x > 0 \\ 0, otherwise \end{cases} \tag{7}$$

Fig.(2b) shows the double-loop binary contour calculation process,there are two closed paths, one of which is(0,2,4,6),another path is (7,1,3,5),The first path pixel pair is{(0,2),(2,4),(4,6),(6,0)}, then $b_{21} = \{s(I_2 - I_0), s(I_4 - I_2), s(I_6 - I_4), s(I_0 - I_6)\}$ .The second path pixel pair is {(7,1), (1,3),(3,5),(5,7)}, $b_{22} = \{s(I_1 - I_7), s(I_3 - I_1), s(I_5 - I_3), s(I_7 - I_5)\}$, combine two eigenvectors into one eigenvector, $b_2 = \{b_{21}, b_{22}\}$ .Similarly,Fig.(2c) shows the three-loop binary contour calculation process.

3.3 Improved Edge Direction Histograms and Binary Gradient Contours

LBP has a very good effect on local dynamic texture extraction. Zhao Guoying et al.proposed the LBP-TOP operator based on dynamic texture analysis in the space-time domain. The improved edge direction histogram and binary gradient contours are based on LBP-TOP. The basic principle is to give a sequence of micro-expressions T. The sequence is divided into three orthogonal planes XY, XT, and YT. X and Y are the axes of the space domain. T is the time axis. Then three orthogonal planes dynamic texture information are extracted. The three plane pixels are described as a cube, as shown in Fig. 3.

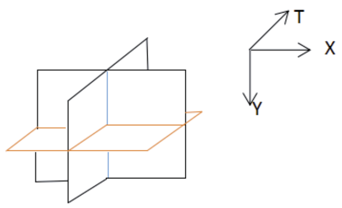
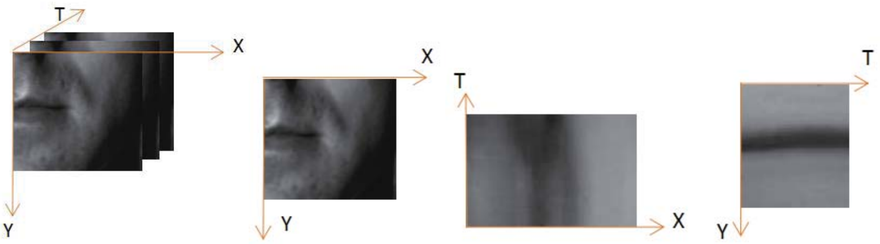


Figure 3. Three Orthogonal Planes

Fig.4 shows how the EOH algorithm can be extracted on three orthogonal planes XY, XT, and YT. First, the image is on the XY plane. The XT and YT planes are the XY planes that change with the time axis T. The EOH algorithm is extraction of texture features from each plane can be represented as XY-EOH, XT-EOH, YT-EOH. Texture feature histogram can be obtained for each plane, so that the dynamic texture feature extraction of the entire micro-expression sequence can obtain 3 histograms. The XY-EOH plane represents the dynamic texture features of the space domain. XT-EOH and YT-EOH represent dynamic texture features of the time domain.

The algorithm performs EOH quantification on each plane, extracts features, and performs feature fusion in the space-time domain to obtain improved EOH features. Similarly, BGC feature extraction is performed on three orthogonal planes, three histograms are obtained, and then feature fusion is performed to obtain improved BGC spatio-temporal domain features.



(4a) (4b) (4c)

Figure 4. EOH algorithm for feature extraction in three orthogonal planes (4a)XY plane(4b)XT plane(4c)YT plane

4. EXPERIMENT

The SMIC database is a spontaneous micro-expression database consisting of 164 samples and 6 emotions: happiness, surprise, disgust, fear, sadness and anger. These samples were captured by three types of cameras: HS (recorded by a high speed camera), VIS (recorded by a normal visual camera), and NIR (recorded by a near-infrared camera).Fig (5a) is a sample in the database, that is, a sequence of micro-expression frames. The RPCA experiment is performed on this sequence to obtain subtle motion information of micro-expression. As shown in Fig. (5b), subtle change information of RPCA extraction is performed. There is almost no identification information, which can further improve the accuracy of subsequent classifications.

(5a)

(5b)

Figure.5 (5a)Original microexpression frame sequence (5b)Extracting subtle facial movement information

In order to verify the effectiveness of the improved EOH algorithm and BGC algorithm in extracting textures, the micro-expression samples in the SMIC database are used, including two micro-expression sequences, as shown in the Fig.6. The sample size of each frame is 350*285, and the block size is 70*70. Using SVM classification, the performance of the two improved algorithms is compared. Experiments are performed on the extraction accuracy, extraction time, and feature dimensions. Experimental data, such as Table1 shows.Half of the samples were trained, half were tested, and after 8-10 repetitions, the average was calculated.

Table1 shows that the improved BGC extraction time and feature dimensions are not as high as the improved EOH algorithm, but the extraction accuracy is similar, so both algorithms have better performance.

Figure 6. Micro-expression series samples

5. CONCLUSION

This paper presents an algorithm for extracting subtle changes in micro-expressions, and also gives an improved method for extracting dynamic texture features based on LBP-TOP algorithm. The calculation is simple and easy to implement, solving the spatial and temporal problems in the extraction of micro-expression sequences. Experiments show that the proposed algorithm has good performance.

Table1.Experimental results

performance / Algorithm	Feature dimensio-n	Extraction time	Extraction accuracy
Improved EOH Algorithm	48	69s	98.66%
Improved BGC Algorithm	765	124s	99.37%

6. ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China No.2017YFB0203102.

7. REFERENCES

[1] Wu Qi, Shen Xunbing, Fu Xiaolan. Research on micro-expression and its application[J].Advances in Psychological Science,2010 (9):1359-1368.

[2] Ekman P, Rosenberg E. What the face reveals : basic and applied studies of spontaneous expression using the facial action coding system (FACS)[J]. Oxford University Press, 2005, 68(1):83-96.

[3] Li X, Pfister T, Huang X, et al. A Spontaneous Micro-expression Database: Inducement, collection and baseline[C]// IEEE International Conference and Workshops on Automatic Face and Gesture Recognition. IEEE, 2013:1-6.

[4] Ekman P. Telling lies:, Clues to deceit in the marketplace, politics, and marriage.[J]. 1992.

[5] Haggard E A, Isaacs K S. Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy [M]// Methods of Research in Psychotherapy. Springer US, 1966:154-165.

[6] Ekman P, Friesen W V. Nonverbal leakage and clues to deception [J]. Psychiatry, 1969, 32(1): 88-106.

[7] Ekman P, Friesen W V. Nonverbal behavior and psychopathology [J]. Contemporary Theory and Research, 1974: 203-232.

[8] Matsumoto D, Le Roux J, Wilson-Cohn C, et al. A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART) [J]. Journal of Nonverbal Behavior, 2000, 24(3): 179-209.

[9] Kallel F, Bertrand M. Tissue elasticity reconstruction using linear perturbation method [J]. Medical Imaging, IEEE Transactions on, 1996, 15(3): 299-313.

[10] Russell T A, Chu E, Phillips M L. A pilot study to investigate the effectiveness of emotion recognition remediation in schizophrenia using the micro-expression training tool [J]. British Journal of Clinical Psychology, 2006, 45(4): 579-583.

[11] Endres J, Laidlaw A. Micro-expression recognition training in medical students: a pilot study [J]. BMC medical education, 2009, 9(1): 47-52.

[12] Russell T A, Green M J, Simpson I, et al. Remediation of facial emotion perception in schizophrenia: concomitant changes in visual attention [J]. Schizophrenia research, 2008, 103(1): 248-256.

[13] Warren G, Schertler E, Bull P. Detecting deception from emotional and unemotional cues [J]. Journal of Nonverbal Behavior, 2009, 33(1): 59-69.

[14] Fellner A N, Matthews G, Funke G J, et al. The effects of emotional intelligence on visual search of emotional stimuli and emotion identification[C]//Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Sage Publications, 2007, 51(14): 845-849.

[15] Polikovsky, S., Kameda, Y., Ohta, Y.: Facial micro-expressions recognition using high speed camera and 3D-gradient descriptor. In: 3rd International Conference on Crime Detection and Prevention. pp. 1-6. IET (2009).

[16] Wang, S.J., Chen, H.L., Yan, W.J., Chen, Y.H., Fu, X.: Face recognition and micro-expression based on discriminant tensor subspace analysis plus extreme learning machine. Neural Processing Letters 39(1), 25–43 (2014).

[17] Pfister, T., Li, X., Zhao, G., Pietikainen, M.: Recognising spontaneous facial micro-expressions. In: 12th IEEE International Conference on Computer Vision. pp.1449–1456. IEEE (2011).

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

[19] Zhao, G., Pietikainen, M.: Dynamic texture recognition using local binary patterns with an application to facial expressions.

IEEE Transactions on Pattern Analysis and Machine Intelligence 29(6), 915–928 (2007).

[20] Shreve M, Godavarthy S, Manohar V, et al. Towards macro- and micro-expression spotting in video using strain patterns[C]//Applications of Computer Vision (WACV),2009 Workshop on. IEEE, 2009: 1-6.

[21] Wu Q, Shen X, Fu X. The machine knows what you are hiding: an automatic micro-expression recognition system [M]//Affective Computing and Intelligent Interaction.Springer Berlin Heidelberg, 2011: 152-162.

[22] Emmanuel J. Candès, Xiaodong Li, Yi Ma, John Wright: Robust principal component analysis. J. ACM 58(3): 11:1-11:37.

[23] Lin, Z., Liu, R., Su, Z.: Linearized alternating direction method with adaptive penalty for low-rank representation. In: Neural Information Processing Systems (NIPS) (2011).

[24] Incorporating Prior Information in Compressive Online Robust Principal Component Analysis, Huynh Van Luong, Nikos Deligiannis, Jurgen Seiler, Soren Forchhammer, Andre Kaup, 2017.

[25] Real-time guitar chord recognition system using stereo cameras for supporting guitarists. Transactions on Electrical Engineering, Electronics, and Communications (ECTI), Pages 147-157.

[26] Wu Y T, Kanade T, Cohn J, et al. Optical flow estimation using wavelet motion model[C]//Computer Vision, 1998. Sixth International Conference on. IEEE, 1998:992-998.

[27] Ekman P, O'Sullivan M. Who can catch a liar[J]. American psychologist, 1991, 46(9), 913-920.