RECOGNITION OF OCCLUDED FACIAL EXPRESSIONS BASED ON CENTRIST FEATURES

Jadisha Yarif Ramírez Cornejo, Helio Pedrini

Institute of Computing, University of Campinas
Campinas, SP, Brazil, 13083-852

ABSTRACT

Emotion recognition based on facial expressions plays an important role in numerous applications, such as affective computing, behavior prediction, human-computer interactions, psychological health services, interpersonal relations, and social monitoring. In this work, we describe and analyze an emotion recognition system based on facial expressions robust to occlusions through Census Transform Histogram (CENTRIST) features. Initially, occluded facial regions are reconstructed by applying Robust Principal Component Analysis (RPCA). CENTRIST features are extracted from the facial expression representation, as well as Local Binary Patterns (LBP), Local Gradient Coding (LGC) and an extended Local Gradient Coding (LGC-HD). Then, the feature vector is reduced through Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). For facial expression recognition, K-nearest neighbor (K-NN) and Support Vector Machine (SVM) classifiers are applied and tested. Experimental results on two public data sets demonstrated that the CENTRIST representation achieved competitive accuracy rates for occluded and non-occluded facial expressions compared to other state-of-the-art approaches available in the literature.

Index Terms— emotion recognition, facial expression, occlusion, fiducial landmarks, feature descriptors

1. INTRODUCTION

Although human emotion has been recently investigated in several knowledge areas, interest in emotions has its roots in Charles Darwin's pioneering studies [1]. He conjectured that emotional expression is universal, that is, independent on culture, race or gender.

Emotion can be defined as a subjective experience or physiological reaction of human beings, which can be expressed through several forms, such as body movement, voice intonation, facial expressions, and cardiac rhythm [2]. A facial expression is comprised of one or more facial musculature movements, which is functionally the same for newborns and adults. Facial expressions [3] are considered a universal and non-verbal communication mode that exhibit emotions in all human beings, allowing emotional information to be conveyed in a simple and natural way. There is strong evidence of universal facial expressions for seven emotions [4]: anger, contempt, disgust, fear, happiness, sadness and surprise.

A challenging and interesting task is to automatically classify emotions through computer vision analysis [5, 6, 7, 8, 9, 10]. Over the past few years, there has been increasing attention to the development of robust devices that can help understand emotions and moods of human beings [11, 12, 13, 2]. Furthermore, research has been conducted to apply these devices in the development of tools for video surveillance systems, airport security, behavioral research, aggression detector for closed-circuit television, on-board emotion detector for drivers, among other applications. Therefore, facial expression recognition is an important issue for affective computing research [14].

Three main components are usually distinguished in automatic facial expression recognition systems: facial detection, feature extraction and representation of facial expression, and expression recognition. It is important to mention that most of the available facial expression recognition systems are based on data sets that do not reflect real and natural scenes. Furthermore, the majority of them do not address occlusions caused by hats, beard, sunglasses or scarves. Consequently, the omission of these factors during the training stage might affect the accuracy of the facial expression recognition process.

A novel and effective facial expression recognition methodology robust to occlusions is proposed and analyzed in this work. The method consists in five main stages. Initially, occluded facial regions are reconstructed through the Dual Approach [15], which is based on RPCA principles. Then, facial fiducial points are automatically detected. A set of features is extracted from the facial expressions, including the Census Transform Histogram. To reduce the dimensionality of the extracted features, the resulting descriptor is transformed to a lower dimensional space. Finally, the occluded facial expressions are recognized.

Experiments are conducted on two public data sets. The results obtained with the proposed method were compared to other approaches available in the literature. Our approach achieved high recognition accuracy rates for occluded and non-occluded images without demanding high computational resources.

The remainder of the paper is organized as follows. Section 2 briefly describes some important concepts related to the topic under investigation. Section 3 presents the methodology proposed in this work, describing the preprocessing, the facial expression reconstruction, the facial feature extraction, the feature reduction, as well as the classification process. Section 4 describes and analyzes the experimental results. Section 5 concludes the paper with final remarks and directions for future work.

2. BACKGROUND

This section briefly describes some concepts related to the facial expression recognition problem addressed in our work.

Robust Principal Component Analysis (RPCA) [16] is a variant of the Principal Component Analysis (PCA) [17], which allows to recover a low-rank matrix A from a corrupted data matrix $D = A + E$, with gross but sparse errors E, through the solution of the following convex optimization problem

$$\min_{A,E} \|A\|_* + \lambda |E|_1 \qquad \text{such that } D = A + E \tag{1}$$

where $\|A\|_*$ is the nuclear norm of a matrix A, $|E|_1$ represents the sum of the absolute values of E, and λ denotes a positive weighting parameter. For the facial expression reconstruction task, our work uses the Dual Approach [15] over the samples of the training set, which solves the RPCA problem (Equation 1) via its dual

$$\max_{Y} \langle D, Y \rangle \qquad \text{such that } J(Y) \leq 1 \tag{2}$$

where

$$\langle A, B \rangle = tr(A^T B) \qquad J(Y) = \max(\|Y\|_2, \lambda^{-1}|Y|_\infty) \tag{3}$$

such that $\|Y\|_2$ implies the spectral norm of a matrix Y, and $|Y|_\infty$ is the maximum absolute value of the matrix entries [15, 18].

Local Binary Patterns (LBP) [19] operator is a robust texture descriptor known by its discriminative power and computational simplicity, allowing real-time image processing. It is invariant to monotonic gray-scales transformations, e.g., caused by illumination variations. LBP encodes the difference between the central pixel value and its $n \times n$ neighborhood pixels sequentially, considering the result as a binary number.

CENsus TRansform hISTogram (CENTRIST) [20] is a visual descriptor, initially proposed for topological place and scene category recognition. This operator is based on Census Transform (CT), a non-parametric local transform. It is characterized for being a holistic representation, which captures the structural properties and has high generalization for categorization, suppressing detailed textural information. CENTRIST compares the central pixel intensity value with its 8-neighborhood. If the central pixel intensity value is higher than or equal to one of its neighbors, bit 1 is set in the corresponding location, otherwise bit 0 is set. The obtained 8 bits can be put together in any order, converting the resulting stream to a base-10 number, as the CT value of the current central pixel. Thus, CENTRIST is a histogram vector with 256 bins that represents the appearance frequency of CT.

Local Gradient Coding (LGC) [21], unlike LBP, considers the graylevel relationship between the central pixel and its neighbors. This operator describes graylevel distribution. The LGC algorithm, using a 3×3 neighborhood template as shown in Figure 1, is defined as

$$\begin{aligned} \text{LGC} \; &= \; s(g_1 - g_3)2^7 + s(g_4 - g_5)2^6 + s(g_6 - g_8)2^5 \\ &+ \; s(g_1 - g_6)2^4 + s(g_2 - g_7)2^3 + s(g_3 - g_8)2^2 \\ &+ \; s(g_1 - g_8)2^1 + s(g_3 - g_6)2^0, \end{aligned} \tag{4}$$

where

$$s(x) = \begin{cases} 1, & \text{if } x > 0. \\ 0, & \text{otherwise.} \end{cases}$$

g_1	g_2	g_3
g_4	g_c	g_5
g_6	g_7	g_8

Fig. 1: A 3×3 neighborhood template of LGC operator.

The LGC algorithm compares the vertical, horizontal and diagonal gradients of its 8-neighbors and converts the binary stream into a base-10 number.

Local Gradient Coding of Horizontal and Diagonal gradient priority (LGC-HD) [21] is an extension of the LGC operator. The LGC-HD operator is capable of decreasing the feature space, reducing the processing time and improving the recognition accuracy. LGC-HD is defined as

$$\begin{aligned} \text{LGC-HD} \; &= \; s(g_1 - g_3)2^4 + s(g_4 - g_5)2^3 + s(g_6 - g_8)2^2 \\ &+ \; s(g_1 - g_8)2^1 + s(g_3 - g_6)2^0. \end{aligned} \tag{5}$$

Principal Component Analysis (PCA) [17] is a common technique for dimensionality reduction, which searches for a smaller set of new composite dimensions that represents a multidimensional feature space with minimum loss of information. PCA models the variance-covariance structure in a set of linearly uncorrelated variables, called principal components, through linear combinations from the original variables. A t-dimensional feature vector of N samples of the training set is projected onto a f-dimensional feature space, through PCA, resulting the new feature vector defined as

$$y_i = W_{PCA}^T x_i \qquad i = 1, \ldots, N \tag{6}$$

where W_{PCA}^T represents the linear transformations matrix, whose columns represents the eigenvectors associated with the largest eigenvalues of the scatter matrix S_T, which is expressed as

$$S_T = \sum_{i=1}^{N} (x_i - \mu)(x_i - \mu)^T \tag{7}$$

where μ is the mean of all the samples of the training set [22].

Linear Discriminant Analysis (LDA) [23] is a classification method for producing models with high accuracy. LDA searches for a linear combination of features that best separates two or more classes, preserving as much of the discriminatory information of the class as possible. Two measures are defined for all the samples of the training set, one is the within-class scatter matrix S_W and the other is the between-class scatter matrix S_b

$$S_W = \sum_{j=1}^{C} \sum_{i=1}^{N_j} (y_i^j - \mu_j)(y_i^j - \mu_j)^T \tag{8}$$

$$S_b = \sum_{j=1}^{C} (\mu_j - \mu)(\mu_j - \mu)^T \tag{9}$$

where y_i^j is the ith sample of class j, μ_j denotes the mean of class j, μ represents the mean of all samples of all classes, C is the number of classes, and N_j is the number of samples per class j [22].

It is known that the application of PCA followed by LDA can achieve higher recognition accuracy [22] than using only the individual approaches. The original t-dimensional space for the training set samples is projected onto an intermediate f-dimensional space using PCA. Then, the latter is projected onto a final g-dimensional space using LDA.

3. METHODOLOGY

The proposed methodology for facial expression recognition under occlusion presence is composed of five main stages: preprocessing, facial expression reconstruction, facial feature extraction, feature reduction and classification. These steps are illustrated in Figure 2 and described as follows.

Initially, a preprocessing step is applied to the images in order to provide aligned faces, uniform size and shape, and randomized occluded facial regions. This preprocessing task consists in the following seven steps: (i) automatic fiducial point detection through Chehra Face and Eyes Tracking Software [24]; (ii) extraction of eye

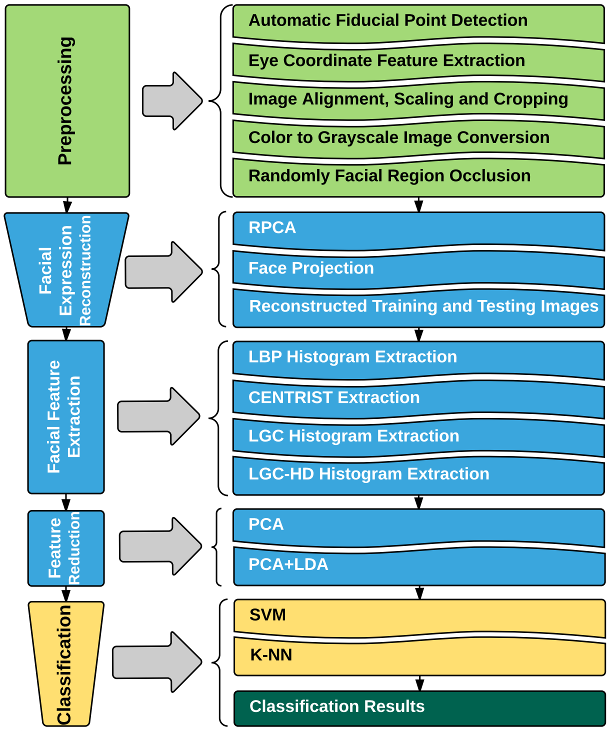


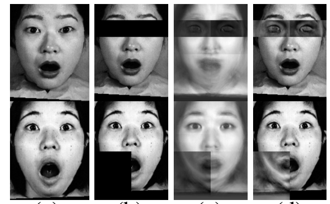
Fig. 2: Diagram with the main steps of the facial expression recognition methodology.

coordinate features; (iii) image rotation to align the eye coordinates; (iv) image scaling proportionally to the minimum distance between the eyes; (v) face region cropping using a proper bounding rectangle; (vi) conversion of the color images to grayscale; (vii) addition of randomized black rectangles to simulate facial occlusion, including left eye, right eye, two eyes, bottom left side of the face, bottom right side of the face or bottom side of the face, as illustrated in Figure 3.

Fig. 3: Cropped images with occluded facial regions from the JAFFE data set.

Although PCA is widely used as a technique for reducing high-dimensional feature subspaces, it does not work well with grossly corrupted observations, e.g., occluded faces, variations of facial expressions, illumination conditions, image noise. On the other hand, RPCA [15] performs in a more effective way with missing data and outliers. RPCA is an extension of the classical PCA procedure and it has been demonstrated to be more robust, among other applications, for the reconstruction of occluded facial expressions [25] and to contribute in achieving better facial expression recognition accuracy [26]. As suggested in [27], RPCA approach is performed for facial expression reconstruction using 150 iterations and a parameter regularization $\lambda = \frac{1}{\sqrt{\max(m,n)}}$ [28], where m and n are the size of matrix D.

After the facial expression reconstruction step, we projected all samples of the testing set onto the space generated by RPCA, such that all occluded facial regions set from the reconstructed faces, for both training and testing sets, were filled. Then, we applied the contrast-limited adaptive histogram equalization (CLAHE) over the reconstructed facial regions to enhance the image contrast. This process is illustrated in Figure 4.



(a) (b) (c) (d)

Fig. 4: (a) Cropped images without occlusions from the JAFFE data set; (b) faces with occluded regions; (c) reconstructed faces; (d) filling the occluded facial regions from (c).

Four visual descriptor types were used in the facial expression recognition: Local Binary Patterns (LBP), Census Transform Histogram (CENTRIST), Local Gradient Coding (LGC) and an extension of the Local Gradient Coding based on the principle of the horizontal and diagonal gradients (LGC-HD).

The LBP descriptor was applied over the entire image to extract the LBP code from each pixel. After obtaining an LBP labeled image and conducting several experiments, we divided the image into 63 (=7×9) regions, as shown in Figure 5. LBP histograms were extracted from each generated region and concatenated all of them into one feature vector of length 16128 (=256×63), which describes local texture and global shape information of the image.

(a) (b) (c)

Fig. 5: (a) Cropped image from the JAFFE data set; (b) LBP image from (a); (c) LBP image is divided into 63 regions.

The images were also divided into 63 regions to extract the CENTRIST features from each region. The resulting vectors are concatenated, forming a vector of 16128 (=256×63). CENTRIST is able to capture local structures of the image. Figure 6 shows a sample of a census transformed facial expression image.

(a) (b)

Fig. 6: (a) Cropped image from CK+ database; (b) census transformed image from (a).

In a similar way to LBP feature extraction, LGC and LGC-HD were applied separately over the entire image. Figure 7 shows the obtained images after applying these operators. The images were then divided into 63 regions to extract their histograms. Hence, the resulting feature vectors have also 16128 (=256×63) dimensions.

(a) (b) (c)

Fig. 7: (a) Cropped image from CK+ database; (b) LGC image from (a); (c) LGC-HD image from (a).

Two techniques for feature reduction, PCA and LDA, were used sequentially, that is, each approach was applied individually for each feature vector. SVM and KNN classifiers were used to compare the recognition accuracy rates.

4. EXPERIMENTAL RESULTS

Our method has been tested on the Cohn-Kanade (CK+) [29] data set and the Japanese Female Facial Expression (JAFFE) [30] data set. The CK data set is available in two versions. We used the second one (CK+) that contains 593 sequential images of posed and non-posed expressions from 123 subjects, categorized into one of seven facial expressions: anger, contempt, disgust, fear, happy, sadness and surprise. The CK+ data set also includes some metadata, such as 68 facial landmarks [29]. The JAFFE data set is a collection of 213 images from 10 Japanese female models who perform seven facial expressions: anger, disgust, fear, happiness, neutral, sadness and surprise [30].

For each data set, we randomly select 80% of samples of each class for the training set and the remaining 20% for the testing set. Then, 50% of the training set samples of each class were occluded and a similar procedure was applied to the testing set. We set 20 different randomized collections of occluded and non-occluded data to perform experiments for both data sets.

From these image collections, we conducted experiments using LBP, CENTRIST, LGC and LGC-HD operators through four methods: PCA+K-NN, PCA+LDA+K-NN, PCA+SVM and PCA+LDA+SVM. The results are shown in Tables 1 and 2, whose values represent the average facial expression recognition accuracy rate from the performed experiments. It is relevant to clarify that RPCA is always applied independently of the feature reduction and classification methods applied.

Table 1: Average accuracy rates (%) for non-occluded facial images from CK+ and JAFFE data sets.

	Method	LBP	CENTRIST	LGC	LGC-HD
CK+	PCA + K-NN	43.74	**55.82**	37.70	36.81
	PCA + LDA + K-NN	92.62	**93.66**	83.44	87.17
	PCA + SVM	77.17	**81.27**	71.80	69.78
	PCA + LDA + SVM	92.84	**94.10**	83.89	85.82
JAFFE	PCA + K-NN	64.41	**87.50**	66.91	73.10
	PCA + LDA + K-NN	**93.00**	91.60	83.10	88.34
	PCA + SVM	84.18	83.40	85.84	**86.31**
	PCA + LDA + SVM	**92.50**	92.00	82.03	87.51

Table 2: Average accuracy rates (%) for occluded facial images from CK+ and JAFFE data sets.

	Method	LBP	CENTRIST	LGC	LGC-HD
CK+	PCA + K-NN	42.06	**54.04**	36.35	33.81
	PCA + LDA + K-NN	88.06	**90.30**	78.06	80.18
	PCA + SVM	75.01	**78.51**	67.97	66.65
	PCA + LDA + SVM	88.44	**90.01**	78.51	79.33
JAFFE	PCA + K-NN	45.84	**87.50**	66.91	74.06
	PCA + LDA + K-NN	83.10	**91.60**	83.10	88.57
	PCA + SVM	70.60	83.40	85.84	**88.67**
	PCA + LDA + SVM	81.44	**92.00**	82.03	87.47

From our experiments, we can observe that the CENTRIST operator allows to achieve more than 90% of accuracy rate for non-occluded image collections, even for occluded ones. Results from CK+ data set showed that the CENTRIST operator is always superior to the other feature extraction methods. From the JAFFE data set results, we can see that the LBP operator is slightly superior to CENTRIST operator for non-occluded images, whereas far better than the others. However, the CENTRIST operator is much better among the other texture operators for occluded images.

Furthermore, despite lower accuracy rates achieved by LGC and LGC-HD operators, we can notice that LGC-HD provides better results than using LGC operator. From the experiments conducted over occluded and non-occluded collections, we can see that following PCA+LDA method always provides higher accuracy rates than only applying PCA.

There are only few similar works available in the literature that consider random partial facial occlusions, especially on both training and testing stages. Then, we compared our approach against other state-of-the-art methods. Table 3 summarizes the best results reached by our approach, including the other methods available in the literature, on both data sets. It is possible to see that the proposed approach (CENTRIST+PCA+LDA+SVM), unlike other visual descriptors, achieves the best results for CK+ and JAFFE data sets under occlusion presence. Table 3 is sorted in descending order by occluded recognition accuracy rate.

Table 3: Comparison of average accuracy rates (%) for non-occluded (non-oc) and occluded facial expression images (oc) on CK+ and JAFFE data sets.

	Approach	Strategy	non-oc	oc
CK+	Proposed method	CENTRIST+PCA+LDA+SVM	94.10	90.01
	Proposed method	LBP+PCA+LDA+SVM	92.62	88.44
	Ramírez et al. [27]	Gabor+PCA+LDA+SVM	94.03	85.68
	Liu et al. [31]	Maximum Likelihood Estimation Sparse Representation	94.29	85.24
JAFFE	Proposed method	CENTRIST+PCA+LDA+SVM	92.00	92.00
	Liu et al. [31]	Maximum Likelihood Estimation Sparse Representation	93.42	86.73
	Proposed method	LBP+PCA+LDA+KNN	93.00	83.10
	Ramírez et al. [27]	Gabor+PCA+LDA+SVM	95.12	82.86
	Zhang et al. [32]	Gabor template and SVM	81.20	48.80

5. CONCLUSIONS AND FUTURE WORK

In this work, we introduced the CENTRIST operator as a potential visual descriptor for emotion recognition. CENTRIST has proven to be robust to occluded and non-occluded facial expressions. Some advantage of this operator include its simple implementation, good performance and high computational speed. Furthermore, experimental results have shown that the use of the PCA+LDA approach increases the recognition rates significantly.

Despite the fact that the LGC and LGC-HD did not provide high accuracy for facial expression recognition, it is important to remark that LGC-HD is much superior than the LGC operator due to its improvement in the classifier recognition accuracy and its computational speed.

As directions for future work, we intend to explore new feature extraction methods, as well as improve the CENTRIST approach. We also plan to conduct experiments by considering real facial occlusions, e.g., sunglasses, scarves, facial hair, caps and beard, to be sure that the proposed approach is robust in real scenes. In addition, we pretend performing experiments for facial expression recognition in video scenes. Finally, we consider crucial the research on the development of an automatic occlusion detector for emotion recognition systems robust to occlusions.

6. REFERENCES

[1] C. Darwin, New York: Oxford University Press, 1872-1998.

[2] S. Zhang, X. Zhao, and B. Lei, "Facial Expression Recognition Using Sparse Representation," *WSEAS Transactions on Systems and Control*, vol. 11, no. 8, pp. 440–441, Aug. 2012.

[3] M. Pantic, "Facial Expression Recognition," in *The Encyclopedia of Biometrics*, pp. 1–8. Springer, 2nd edition, 2014.

[4] P. Ekman, "Facial Expressions," *Handbook of Cognition and Emotion*, vol. 16, pp. 301–320, 1999.

[5] V. Gay, P. Leijdekkers, J. Agcanas, F. Wong, and Q. Wu, "CaptureMyEmotion: Helping Autistic Children Understand their Emotions Using Facial Expression Recognition and Mobile Technologies," *Studies in Health Technology and Informatics*, vol. 189, pp. 71–76, 2013.

[6] A. Majumder, L. Behera, and V. K. Subramanian, "Emotion Recognition from Geometric Facial Features using Self-Organizing Map," *Pattern Recognition*, vol. 47, no. 3, pp. 1282–1293, 2014.

[7] S. Ryoo and J.-K. Chang, "Emotion Affective Color Transfer Using Feature Based Facial Expression Recognition," *Advanced Science and Technology Letters*, vol. 39, pp. 131–135, 2013.

[8] J. A. Rojas, A. Ramirez, and O. Chae, "Facial Expression Recognition Based On Local Sign Directional Pattern," in *19th IEEE International Conference on Image Processing*, 2012, pp. 2613–2616.

[9] C. Shan, S. Gong, and P. W. McOwan, "Facial Expression Recognition based on Local Binary Patterns: A Comprehensive Study," *Image and Vision Computing*, vol. 27, no. 6, pp. 803–816, 2009.

[10] T. Jabid, Md. H. Kabir, and O. Chae, "Robust Facial Expression Recognition Based on Local Directional Pattern," in *27th Conference on Image and Vision Computing*, New Zealand, 2012, pp. 464–468.

[11] A. Sanin, C. Sanderson, M. T. Harandi, and Brian C. Lovell, "Spatio-Temporal Covariance Descriptors for Action and Gesture Recognition," in *IEEE Workshop on Applications of Computer Vision*, 2013, pp. 103–110.

[12] Z. Wang, S. Wang, and Q. Ji, "Capturing Complex Spatio-Temporal Relations among Facial Muscles for Facial Expression Recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 3422–3429.

[13] H. Xue and I. Gertner, "Automatic Recognition of Emotions from Facial Expressions," in *Proceedings of SPIE - 9090, Automatic Target Recognition XXIV*, Baltimore, MD, USA, 2014, pp. 90900O–90900O–12.

[14] K. C. Lin, T.-C. Huang, J. C. Hung, N. Y. Yen, and S. J. Chen, "Facial Emotion Recognition towards Affective Computing-based Learning," *Library Hi Tech*, vol. 31, no. 2, pp. 294–307, 2013.

[15] Z. Lin, A. Ganesh, J. Wright, L. Wu, M. Chen, and Y. Ma, "Fast Algorithms for Recovering a Corrupted Low-Rank Matrix," in *Computational Advances in Multi-Sensor Adaptive Processing*, Dec. 2009, pp. 213–216.

[16] E. J. Candés, X. Li, Y. Ma, and J. Wright, "Robust Principal Component Analysis?," *Journal of the ACM*, vol. 58, no. 3, pp. 1–37, May 2011.

[17] I. T. Jolliffe, *Principal Component Analysis*, Springer-Verlag, New York, 2002.

[18] Z. Lin, M. Chen, L. Wu, and Y. Ma, "The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices," UIUC Technical Report UILU-ENG-09-2215, University of Illinois at Urbana-Champaign, Oct. 2009.

[19] T. Ojala, M. Pietikhenl, and D. Harwood, "Performance Evaluation of Texture Measures with Classification based on Kullback Discrimination of Distributions," *International Conference on Pattern Recognition*, pp. 582–585, 1994.

[20] J. Wu and J. M. Rehg, "CENTRIST: A Visual Descriptor for Scene Categorization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1489–1501, 2011.

[21] Y. Tong, R. Chen, and Y. Cheng, "Facial Expression Recognition Algorithm using LGC based on Horizontal and Diagonal Prior Principle," *Optik - International Journal for Light and Electron Optics*, vol. 125, no. 16, pp. 4186–4189, 2014.

[22] H. Deng, L. Jin, L. Zhen, and J. Huang, "A New Facial Expression Recognition Method Based on Local Gabor Filter Bank and PCA plus LDA," *International Journal of Information Technology*, vol. 11, no. 5, pp. 86–96, 2005.

[23] R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics*, vol. 7, no. 2, pp. 179–188, 1936.

[24] A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic, "Incremental Face Alignment in the Wild," in *IEEE International Conference on Computer Vision and Pattern Recognition*, 2014.

[25] B. Jiang and K. Jia, "Research of Robust Facial Expression Recognition under Facial Occlusion Condition," in *Active Media Technology, Lecture Notes in Computer Science*, vol. 6890, pp. 92–100. Springer-Verlag Berlin Heidelberg, 2011.

[26] X. Mao, Y. Li Xue, Z. Li, K. Huang, and S. Lv, "Robust Facial Expression Recognition based on RPCA and AdaBoost," *10th Workshop on Image Analysis for Multimedia Interactive Services*, pp. 113–116, May 2009.

[27] J. Y. Ramírez C., H. Pedrini, and F. Flórez-Revuelta, "Facial Expression Recognition with Occlusions based on Geometric Representation," in *20th Iberoamerican Congress on Pattern Recognition*, Montevideo, Uruguay, Nov. 2015.

[28] C. Guyon, T. Bouwmans, and E. Zahzah, "Robust Principal Component Analysis for Background Subtraction: Systematic Evaluation and Comparative Analysis," in *Principal Component Analysis*, pp. 223–238. INTECH, Mar. 2012.

[29] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," 2010.

[30] M. J. Lyons, M. Kamachi, and J. Gyoba, "Japanese Female Facial Expressions (JAFFE)," 1997, Database of Digital Images.

[31] S. S. Liu, Y. Zhang, and K. P. Liu, "Facial Expression Recognition under Random Block Occlusion based on Maximum Likelihood Estimation Sparse Representation," in *International Joint Conference on Neural Networks*, July 2014, pp. 1285–1290.

[32] L. Zhang, D. Tjondronegoro, and V. Chandran, "Random Gabor based Templates for Facial Expression Recognition in Images with Facial Occlusion," *Neurocomputing*, vol. 145, pp. 451–464, 2014.