3rd International Conference on Advanced Technologies
for Signal and Image Processing - ATSIP'2017,
May 22-24, 2017, Fez, Morroco.

Facial Expression Recognition Using Face-Regions

Khadija Lekdioui[1,2], Yassine Ruichek[1], Rochdi Messoussi[2], Youness Chaabi[2], Raja Touahni[2]

[1]Le2i FRE2005, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté, UTBM, F-90010 Belfort, France
Email: {khadija.lekdioui,yassine.ruichek}@utbm.fr
[2]Laboratoire des Systèmes de Télécommunication et Ingénierie de la Décision (LASTID)
Université Ibn Tofail BP 133, Kenitra 14000, Maroc

***Abstract*—This paper proposes a facial expression recognition method based on a novel facial decomposition. First, seven regions of interest (ROI), representing the main components of face (left eyebrow, right eyebrow, left eye, right eye, between eyebrows, nose and mouth), are extracted using facial landmarks detected by IntraFace algorithm. Then, different local descriptors, such as LBP, CLBP, LTP and Dynamic LTP, are used to extract features. Finally, feature vector, representing face image, is fed into a multiclass support vector machine to achieve the recognition task. Experimental results on two public datasets show that the proposed method outperforms state of the art methods based on other facial decompositions.**

Keywords*—*Facial expression recognition, facial landmarks, Facial decomposition, feature descriptor, SVM

I. Introduction

Face is one of the most important means of human communication. It plays a central role in all social interactions. Facial expressions are non-verbal clues to emotions. Indeed, some facial muscles are specifically associated with certain emotional states and allow, according to Ekman [8] the expression of primary emotions (Sadness, Anger, Fear, Joy, Disgust and Surprise). These external signals express the internal emotional state of an individual, and therefore his intentions. In fact, 7% of the communication relies on verbal interaction, 38% represent tone and sound of voice, 55% are articulated around gestures and expressions of the face according to Meharabian [18]. Automatic recognition of facial expressions is an interesting problem which finds its interest in several fields such as eLearning and affective computing [20], [15], [21]. When designing an automatic facial expression recognition system, three problems are considered: face detection, facial feature extraction, and classification of expressions. First, face acquisition is a processing stage to automatically locate the face region in the input images. The next step is to extract and represent facial changes caused by facial expressions. Finally, the classification task allows to infer the facial expressions. According to the existing types of feature extraction, facial expression recognition process can be commonly divided into appearance-based and geometric-based methods. For instance, the methods addressed in [23], [3], [1] extract textures from image to characterize facial appearance changes and the methods in [14], [24] measure geometric displacement, distances and angles between facial landmarks to extract geometric information. Recently, several works, based on appearance-based features, used local descriptors such as LBP [23], [9], its variants like LTP [9], MTP [2], CLBP [1]. Similarly, HOG [3], PCA [17], LDA and Wavelet [16] were used for appearance-based feature extraction. Shan et al. [23] carried out extensive experiments where they evaluated LBP features with different classification techniques and they showed that LBP features are effective and efficient for facial expression recognition, even in low-resolution video sequences. Gritti et al. [9] also extensively investigated different local features, such as LBP, LTP, HOG and Gabor and demonstrated that LBP outperforms all the tested features. In [3], Caragní et al. largely studied HOG parameter settings (cell size and number of orientation bins) and concluded that the choice of a proper set of HOG parameters can make HOG descriptor one of the powerful techniques for facial expression recognition. Ahmed et al. [1] proposed an LBP variant called Compound Local Binary Pattern (CLBP) that combines magnitude information of the difference between two gray values and the basic LBP. Another new local texture pattern called Median Ternary Pattern (MTP) is proposed in [2]. It combines the advantages of median filter and quantization of gray-scale values into 3 value codes. Geometric-based approaches use facial landmarks to represent the whole face shape. As in many research works [10], [24], the movement and positions of facial landmarks are calculated to extract geometric information. In [10], the authors detect facial expressions by using Active Appearance Models (AAM) to extract key features and observing changes of their value using Fuzzy Logic. Shbib and Zhou [24] applied Active Shape Model (ASM) fitting technique to extract facial feature points. Then geometric displacement of projected ASM feature points, and the mean shape of ASM were analyzed to recognize facial expressions. Two feature extraction approaches Correlation Features Selection (CFS) and Empirical Normalized Distances (END) are applied in [6]. These techniques are both based on geometric feature extraction using Point Distribution Model (PDM) tracker to localize landmark positions. As reported in the literature [3], [7], [11], appearance feature extraction techniques can be applied to the whole face, specific face-regions, patches around some facial landmarks to extract appearance textures of the face. Lately, many studies are interested in recognizing facial expressions using specific face-regions. In [27], the authors extracted facial regions using AAM, then extracted facial features from the defined regions applying Gabor wavelet transformation. The authors in [5], [7] defined facial components from which they extracted HOG feature descriptors. Almost all the studies stated above used Support Vector Machine (SVM) as a classifier for facial expression recognition. The objective of this work is to recognize automatically the six basic emotions as well as the neutral state by applying appearance-feature methods like LBP and its variants to new specific face-regions. Two main contributions are reported in this paper: (1) the decomposition of the face into seven regions of interest (ROI) where each ROI contains one face component, this novel decomposition improves facial expression recognition performance compared to state of the art methods based on specific face-regions or whole face. (2) A comprehensive study is carried out with different descriptors, especially a dynamic LTP where different formulas are adopted to calculate its threshold. The evaluation is performed on two databases using SVM.

II. Method

Our proposed methodology aims to recognize the six universal facial expressions (Joy, Fear, Disgust, Surprise, Sadness, Anger) and the neutral state. It consists of three main steps. First, we extract 7 face-regions, called regions of interest (ROI), with the help of some facial landmarks detected using IntraFace (IF). Then, we conduct a large experiment using LBP, CLBP, LTP and Dynamic LTP to extract features in each ROI. Finally, the feature vector is fed into SVM classifier to recognize facial expression.

A. Face-regions Extraction

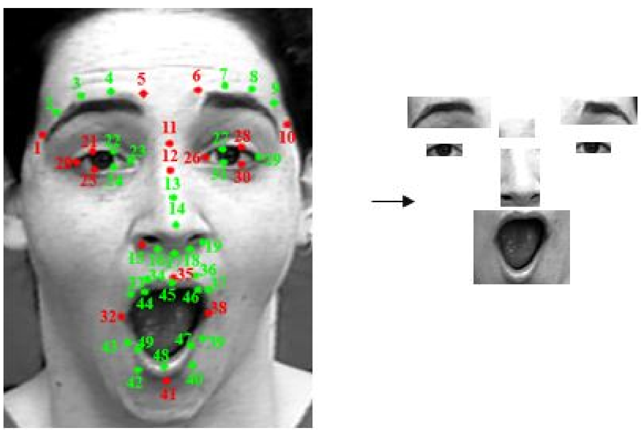


Fig. 1. The left image represents 49 Landmarks detected by IF algorithm and the right image represents our facial decomposition extracted by using the position of the red landmarks.

To achieve the ROI extraction step, we start by detecting facial landmarks using IF framework. The IF algorithm can detect 49 landmarks around the regions of eyebrows, eyes, nose, and mouth (see Fig. 1) using the Supervised Descent Method (SDM) [29]. SDM is a supervised method that learns to optimize non-linear least squares (NLS) problems. It is a non-parametric shape model which allows to generalize better to untrained situations. During training, SDM learns a sequence of generic descent directions. In testing, SDM uses the learned descent directions to minimize the NLS objective function. Consider an image where landmarks are manually allocated and which are appointed as x_* and x_0 an initial configuration of landmarks that corresponds to an average shape. Face alignment can be represented by minimizing the NLS function given as follow:

$$f(x_0 + \Delta x) = ||h(d(x_0 + \Delta x)) - \phi_*|| \tag{1}$$

where $\phi_* = h(d(x_*))$ represents SIFT values. By extracting SIFT features from patches around landmarks, a robust representation against illumination can be achieved. For details on SDM, the reader is referred to [29].
Once the detection of facial landmarks is done, 7 main ROIs, in which information is concentrated about facial expression, are extracted (left eyebrow, right eyebrow, left eye, right eye, between eyebrows, nose and mouth) (see Fig. 1). These ROIs undergo an important changes when expressing different emotions. For example, using the region between eyebrows allows to distinguish between sadness, anger and fear expressions where vertical wrinkles are significant and other expressions where these wrinkles are absent.

This step is accomplished by using the location of some particular facial landmarks like the ones labeled 1, 5, 6, 10, 35, and 41, represented by the red color in Fig. 1, and which delineate the face components.
These face-regions contain more changes in texture information, whence the interest of applying different feature texture descriptors. In the following, we present the different descriptors used in this work (LBP, CLBP, LTP, and Dynamic LTP) as well as the SVM classifier.

B. Local Binary Pattern

LBP operator was introduced by Ojala et al. [22]. Although very simple to operate, LBP demonstrated very good performances in texture classification. The method analyzes the textures by encoding each pixel of the image into a string of binary values. This is performed by thresholding the differences between its value and the 3x3 neighborhood pixel values. LBP operator can be expressed as follows:

$$LBP_{P,R}(x_c, y_c) = \sum_{p=1}^{P} s(g_p - g_c)2^{p-1} \tag{2}$$

with,

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{3}$$

where (x_c, y_c) are the coordinates of the current pixel (central pixel), P is the number of pixels of the neighborhood, R is the radius of the neighborhood, gc is the gray level of the central pixel, and gp is the gray level of the pixel p.

C. Compound Local Binary Pattern

To overcome the limitations of LBP descriptor, Ahmed et al. [1] proposed CLBP which combines the sign and the magnitude of the differences between the center and the neighbor gray values. In CLBP, the indicator function $s(x)$ is represented as:

$$s(x) = \begin{cases} 00 & \text{if } x < 0, |x| \leq M_{avg} \\ 01 & \text{if } x < 0, |x| > M_{avg} \\ 10 & \text{if } x \geq 0, |x| \leq M_{avg} \\ 11 & \text{otherwise} \end{cases} \tag{4}$$

where M_{avg} is the average magnitude of the differences between the gray value of the central pixel and the neighbor gray values. The first bit represents the sign of the difference, like in the basic LBP encoding. The other bit is used to encode the magnitude of the difference with respect to the threshold value M_{avg}. Fig. 2 illustrates the reduction of the dimensionality of CLBP which splits each CLBP pattern into two sub-CLBP codes. The first one consists of the horizontal and vertical CLBP patterns, and the second one consists of the diagonal patterns. At the end, two histograms are concatenated to form CLBP descriptor.

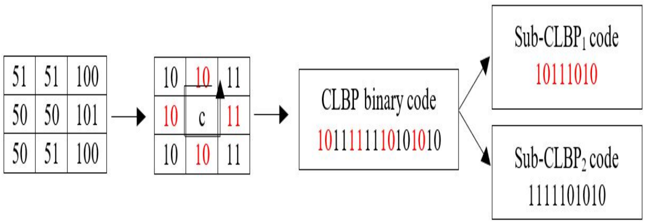


Fig. 2. CLBP operator and the generation of two sub-CLBP.

D. Dynamic LTP

LBP operator is robust to monotonic illumination variation. However, it is sensitive to noise. To deal with this problem, Tan and Triggs [25] extended the original LBP formulation to a variant with three value codes, called LTP. LTP can take three values according to distance of the values of the neighboring pixels to the value of the central pixel. In LTP, the indicator function $s(x)$ can be defined as:

$$s(x) = \begin{cases} 1 & \text{if } x \geq t \\ 0 & \text{if } |x| < t \\ -1 & \text{if } x \leq -t \end{cases} \tag{5}$$

This encoding is based on a fixed threshold t, which is not easy to adjust. For addressing this issue, many works adopted different formulas to calculate a dynamic threshold [28], [19], [12]. In this work, we exploit the LTP at its best by investigating the problem of how to chose an appropriate threshold that allows to reach high facial expression recognition performance. Different techniques for threshold computation were adopted in our experiments. Concerning fixed threshold, different values from 1 to 5 were tested. For dynamic threshold, formula (F1) (see Table I) proposed in [12] is used considering different values for the scaling factor δ (0.02, 0.08, 0.05, 0.1, 0.15 and 0.2). We also propose in this paper other simple formulas to achieve dynamic thresholding (F2, F3, F4 and F5) (see Table I).

TABLE I. FORMULAS BASED LTP OPERATOR (p_c IS THE VALUE OF CENTRAL PIXEL p_i IS THE VALUE OF THE i^{th} NEIGHBOR PIXEL. N IS THE NUMBER OF NEIGHBORING PIXELS (IN THIS WORK N=8)).

Formula 1	Formula 2	Formula 3	Formula 4	Formula 5
$t = p_c \times \delta$	$t = \frac{\sum_{i=0}^{N-1} \sqrt{p_i}}{N}$	$t = \frac{\sum_{i=0}^{N-1} p_i}{N}$	$t = \frac{\sqrt{\sum_{i=0}^{N-1} p_i}}{N}$	$t = \sqrt{\sum_{i=0}^{N-1} \frac{p_i}{N}}$

E. Support Vector Machine

The input of the classifier is the concatenation of all the histograms calculated from each block of each ROI (see Fig. 3). Block decomposition of ROI allows to extract more local information. For the classification task, we employ linear Support Vector Machines (SVM). Given a labeled training data $x_i, y_i, i = 1, \cdots, p, y_i \in \{-1, 1\}, x_i \in \Re^d$, the algorithm builds a hyperplane with largest margin which separates the positive from the negative examples. The support vector classifier solves the following optimization problem:

$$\min_{w} \frac{1}{2} w^T w + C \sum_{i=1}^{p} \xi(w, x_i, y_i) \tag{6}$$

where w is normal to the hyperplane, $C > 0$ is a penalty parameter. $\xi(w, x_i, y_i)$ is a misclassification error. SVM is a binary classifier, but it can solve also a multi-class problem by mitigating the problem to a set of multiple binary classifiers. In this work, the multi-class classification is performed by using the one-against-one approach implemented by LIBSVM [4], which trains binary classifiers to discriminate one class from each other and determines the class label by the binary classifier that gives maximum output values.

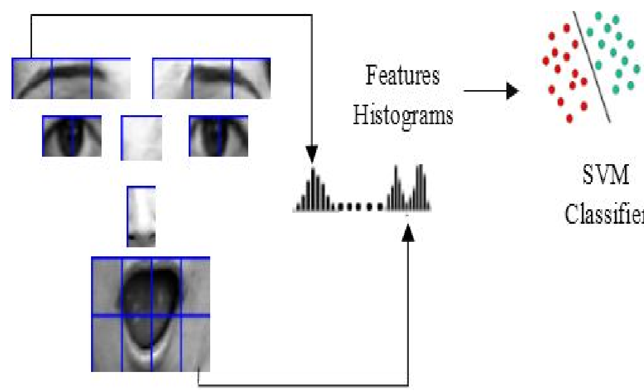


Fig. 3. Each ROI is partitioned into blocks from which feature histograms are extracted and concatenated to form the input of SVM Classifier.

III. EXPERIMENTS

A. Datasets

We evaluate our proposed methodology on two publicly databases: the CK database [13] which represents deliberate facial expressions and the FEED database [26] which contains spontaneous ones. The first database contains frontal image sequences of 97 subjects aged from 18 to 30 years. The lighting conditions are relatively uniform and the images present small variations in poses. The facial expression in each sequence is displayed from neutral to target state. In our study, we selected a dataset of 610 images from 32 subjects representing the 7 basic facial expressions, namely neutral, joy, sadness, surprise, anger, fear and disgust. Fig. 4 shows some images from the CK database. The second database comprises facial expression image sequences of 18 subjects representing the 7 basic facial expressions. In our study, we selected a dataset of the 7 basic facial expressions which comprises a total of 630 images from 16 subjects. Some images from the FEED database are shown in Fig. 5.

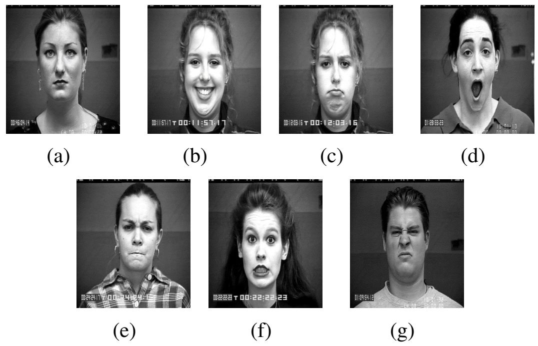


Fig. 4. Examples of different facial expressions from the CK database. (a) neutral (b) joy (c) sadness (d) surprise (e) anger (f) fear (g) disgust.

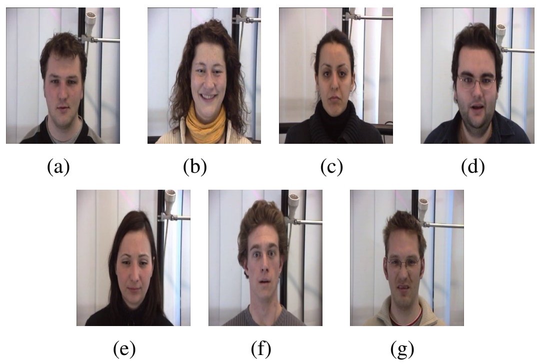


Fig. 5. Examples of different facial expressions from the FEED database. (a) neutral (b) joy (c) sadness (d) surprise (e) anger (f) fear (g) disgust.

B. Evaluation

The facial expression recognition performance is sensitive to ROIs size, number of blocks in ROIs and the descriptor used to extract features. In our experiments, all these parameters as well as the threshold of the LTP descriptor were optimized to get the best recognition performance. In this section, we perform 10-fold cross validation on each dataset to evaluate the proposed method. Each dataset is divided into 10 groups which are used to conduct 10 experiments. In each experiment, 9 groups are used for training and the rest is used for the test. Then, the recognition rate is calculated using F-score. Through large experiments, we compare 3 ROI face decompositions with different sizes and numbers of blocks in ROI (see Tables II, III and IV). The descriptors are also taken into account in the evaluation with their parameters, especially the parameter t of LTP. Here we focus only on the configurations that yield best recognition rate. Tables II, III and IV report the results of 1 ROI, 6 ROIs and 7 ROIs respectively. The first decomposition considers the whole face as a single ROI (see Table II), as proposed in [1], [3]. In the second one, the face is composed into 6 ROIs (see Table III), as proposed in [7]. The third decomposition is the one we proposed, it considers 7 ROIs, as shown in Table IV.

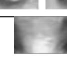
First, we start by considering the whole face as single ROI (Table II). For CK dataset, the higher F-score of 91.46 % is obtained using the first configuration (ROI size and number of blocks) and dynamic LTP with threshold computed from formula 1 (see Table I) and scaling factor δ=0.15. Formula 1 and 5 give close F-scores. As we can see in Table II, the same descriptor with the same parameters (formula 1 and δ) reaches the best recognition rate (89.85%) considering the second configuration. For FEED dataset, the best result is 84.39% when dynamic LTP is applied with threshold based on formula 2. We can also observe that almost the same recognition rate is obtained with fixed threshold (t=1).

TABLE II. RECOGNITION PERFORMANCE FOR DIFFERENT NUMBERS OF BLOCKS, ROI SIZES AND DESCRIPTORS WITH NUMBER OF ROI=1

	Number of ROI = 1		Size	Number of blocks	Size	Number of blocks
			165x154	11x11	104x117	8x9
CK	LBP		89.42		85.67	
	CLBP		87.13		85.39	
	LTP	t=F1	**91.46**		89.85	
		t=F2	91.31		88.17	
		t=F5	91.29		88.81	
FEED	LBP		83.42		84.17	
	CLBP		70.88		75.01	
	LTP	t=F2	78.81		**84.39**	
		t=1	83.6		84.35	
		t=F4	82.48		83.89	

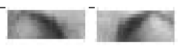
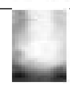
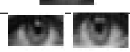
Table III shows the recognition rates when face is decomposed into 6 ROIs. For CK dataset, compared to the results of the whole face as one ROI in Table II, the recognition performance is increased and the best F-score of 92.19% is obtained using dynamic LTP with threshold formula 1 and δ=0.02. For FEED dataset, it can be observed that the recognition rate is significantly decreased in all the used descriptors compared to the previous case (number of ROI=1). However, as we can infer from Table III, LTP descriptor is the best descriptor, which yields the higher recognition rate (77.35%) when fixed threshold (t=1) is used.

TABLE III. RECOGNITION PERFORMANCE FOR DIFFERENT NUMBERS OF BLOCKS, ROI SIZES AND DESCRIPTORS WITH NUMBER OF ROI=6

	Number of ROI = 6		Size	Number of blocks	Size	Number of blocks
	r1		80x40	4x2	120x36	5x2
	r1		30x30	1x1	24x36	1x2
	r3/r4		36x24	2x2	36x24	1x2
	r5		30x30	1x1	24x18	1x1
	r6		80x60	4x3	72x54	3x3
CK	LBP		90.17		89.46	
	CLBP		85.58		86.6	
	LTP	t=F1	89.08		**92.19**	
		t=F2	89.89		91.07	
		t=F5	90.05		90.95	
FEED	LBP		71.25		72.41	
	CLBP		69.08		71.29	
	LTP	t=F1	74.63		76.72	
		t=1	**77.35**		73.02	
		t=F4	75.01		72.63	

The proposed face decomposition (number of ROI=7) achieved the best recognition rate of 94.11% on CK dataset and 87.57% on FEED dataset, as shown in Table IV. The descriptor that provided these performances is the dynamic LTP using threshold formula 1 with δ=0.02 and threshold formula 5, on CK and FEED datasets respectively. From Tables II, III and IV, we can see that our face decomposition (number of ROI = 7) performs better compared to the other decompositions considered in this work (number of ROI = 1 and number of ROI = 6). This is true regardless of the used descriptor and the tested dataset. The main purpose of this work is to compare between different face registration namely whole face and facial decompositions into ROIs in order to show that relevant local regions around facial components, as proposed, outperforms the holistic method which uses the whole face as one ROI and an other facial decomposition. Hence, we applied the same descriptors to all the decompositions considering in this work (1 ROI, 6 ROIs, 7 ROIs) to evaluate our proposed facial decomposition.

TABLE IV. RECOGNITION PERFORMANCE FOR DIFFERENT NUMBERS OF BLOCKS, ROI SIZES AND DESCRIPTORS WITH NUMBER OF ROI=7

	Number of ROI = 7		Size	Number of blocks	Size	Number of blocks
	r1/r2		80x20	4x1	72x18	3x1
	r3		30x30	1x1	18x26	1x1
	r4/r5		36x24	2x2	36x18	2x1
	r6		30x30	1x1	18x26	1x1
	r7		80x60	4x3	72x36	3x2
CK	LBP		91.21		92.09	
CK	CLBP		86.69		90.53	
CK	LTP	t=F1	92.5		**94.11**	
CK	LTP	t=F2	91.45		92.48	
CK	LTP	t=F5	90.14		93.4	
FEED	LBP		86.96		76.73	
FEED	CLBP		81.7		71.51	
FEED	LTP	t=F2	86.96		79.36	
FEED	LTP	t=1	87.22		79.84	
FEED	LTP	t=F5	**87.57**		78.49	

For all the three face decompositions, it is quite evident that the recognition rate on CK dataset is high than on FEED dataset because CK dataset is composed of posed facial expressions which are more exaggerated compared to spontaneous ones (FEED dataset) which are acquired with low intensity, and therefore are more difficult to identify.

IV. CONCLUSION

In this paper, we proposed a new facial decomposition for basic emotion states recognition. Based on facial landmarks detected by IntraFace algorithm, seven regions of interest (ROI), corresponding to the main components of face, are first extracted to represent face image. A preprocessing stage is then applied on these ROIs for resizing and partitioning them into blocks, before performing feature extraction to build face feature descriptor. Finally, a multiclass SVM classifier is utilized to infer emotion state. A comprehensive experimental study, using different local features, is carried out to compare the proposed method to two state of the art methods; one is based on whole face as a single ROI and the other one uses facial decomposition with six ROIs. Experimental results demonstrated the effectiveness of our proposed method on all tested datasets using all tested descriptors. The proposed facial decomposition outperformed the state of the art ones.

ACKNOWLEDGMENT

This research work is part of Volubilis project registered under Volubilis MA/14/302.The authors would like to thank the Franco-Moroccan Volubilis program for its support.

REFERENCES

[1] F. Ahmed, H. Bari, and E. Hossain. Person-independent facial expression recognition based on compound local binary pattern (clbp). *Int. Arab J. Inf. Technol.*, 11(2):195–203, 2014.

[2] F. Bashar, A. Khan, F. Ahmed, and M. H. Kabir. Robust facial expression recognition based on median ternary pattern (mtp). In *Electrical Information and Communication Technology (EICT), 2013 International Conference on*, pages 1–5. IEEE, 2014.

[3] P. Carcagnì, M. Coco, M. Leo, and C. Distante. Facial expression recognition and histograms of oriented gradients: a comprehensive study. *SpringerPlus*, 4(1):1, 2015.

[4] C.-C. Chang and C.-J. Lin. Libsvm: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[5] J. Chen, Z. Chen, Z. Chi, and H. Fu. Facial expression recognition based on facial components detection and hog features. In *International Workshops on Electrical and Computer Engineering Subfields*, pages 884–888, 2014.

[6] J. de Andrade Fernandes, L. N. Matos, and M. G. dos Santos Aragão. Geometrical approaches for facial expression recognition using support vector machines. In *Graphics, Patterns and Images (SIBGRAPI), 2016 29th SIBGRAPI Conference on*, pages 347–354. IEEE, 2016.

[7] M. M. Donia, A. A. Youssif, and A. Hashad. Spontaneous facial expression recognition based on histogram of oriented gradients descriptor. *Computer and Information Science*, 7(3):31, 2014.

[8] P. Ekman. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200, 1992.

[9] T. Gritti, C. Shan, V. Jeanne, and R. Braspenning. Local features based facial expression recognition with face registration errors. In *Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on*, pages 1–8. IEEE, 2008.

[10] A. A. Gunawan et al. Face expression detection on kinect using active appearance model and fuzzy logic. *Procedia Computer Science*, 59:268–274, 2015.

[11] S. Happy and A. Routray. Automatic facial expression recognition using features of salient facial patches. *IEEE transactions on Affective Computing*, 6(1):1–12, 2015.

[12] M. Ibrahim, M. Alam Efat, S. Kayesh, S. M. Khaled, M. Shoyaib, and M. Abdullah-Al-Wadud. Dynamic local ternary pattern for face recognition and verification. In *Proceedings of the International Conference on Computer Engineering and Applications, Tenerife, Spain*, volume 1012, 2014.

[13] T. Kanade, J. F. Cohn, and Y. Tian. Comprehensive database for facial expression analysis. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 46–53. IEEE, 2000.

[14] G. Lei, X.-h. Li, J.-l. Zhou, and X.-g. Gong. Geometric feature based facial expression recognition using multiclass support vector machines. In *Granular Computing, 2009, GRC'09. IEEE International Conference on*, pages 318–321. IEEE, 2009.

[15] K. Lekdioui, R. Messoussi, and Y. Chaabi. Etude et modélisation des comportements sociaux d'apprenants à distance, à travers l'analyse des traits du visage. In *7ème Conférence sur les Environnements Informatiques pour l'Apprentissage Humain (EIAH 2015)*, pages 411–413, 2015.

[16] P. Marasamy and S. Sumathi. Automatic recognition and analysis of human faces and facial expression by lda using wavelet transform. In *Computer Communication and Informatics (ICCCI), 2012 International Conference on*, pages 1–4. IEEE, 2012.

[17] S. S. Meher and P. Maben. Face recognition and facial expression identification using pca. In *Advance Computing Conference (IACC), 2014 IEEE International*, pages 1093–1098. IEEE, 2014.

[18] A. Mehrabian et al. *Silent messages*, volume 8. Wadsworth Belmont, CA, 1971.

[19] A. A. Mohamed and R. V. Yampolskiy. Adaptive extended local ternary pattern (aeltp) for recognizing avatar faces. In *Machine Learning and Applications (ICMLA), 2012 11th International Conference on*, volume 1, pages 57–62. IEEE, 2012.

[20] G. Molinari, C. Bozelle, D. Cereghetti, G. Chanel, M. Bétrancourt, and T. Pun. Feedback émotionnel et collaboration médiatisée par ordinateur: Quand la perception des interactions est liée aux traits émotionnels. In *Environnements Informatiques pour l'apprentissage humain, Actes de la conférence EIAH*, pages 305–326, 2013.

[21] R. Nkambou and V. Heritier. Reconnaissance émotionnelle par l'analyse des expressions faciales dans un tuteur intelligent affectif. In *Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et l'Industrie*, pages 149–155. Université de Technologie de Compiègne, 2004.

[22] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59, 1996.

[23] C. Shan, S. Gong, and P. W. McOwan. Facial expression recognition based on local binary patterns: A comprehensive study. *Image and Vision Computing*, 27(6):803–816, 2009.

[24] R. Shbib and S. Zhou. Facial expression analysis using active shape model. *Int. J. Signal Process. Image Process. Pattern Recognit*, 8(1):9–22, 2015.

[25] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE transactions on image processing*, 19(6):1635–1650, 2010.

[26] F. Wallhoff. Facial expressions and emotion database. *Technische Universität München*, 2006.

[27] L. Wang, R. Li, and K. Wang. A novel automatic facial expression recognition method based on aam. *Journal of Computers*, 9(3):608–617, 2014.

[28] X. Wu, J. Sun, G. Fan, and Z. Wang. Improved local ternary patterns for automatic target recognition in infrared imagery. *Sensors*, 15(3):6399–6418, 2015.

[29] X. Xiong and F. De la Torre. Supervised descent method and its applications to face alignment. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 532–539, 2013.