New Geometric Based Features for Facial Expression Recognition

Nuri Özbey and Mehmet Bilginer Gülmezoğlu

January 29, 2021

New Geometric Based Features for Facial Expression Recognition

Nuri ÖZBEY
R&D Department
Arçelik A.Ş. Refrigerator Plant
Eskisehir, Turkey
nuri.ozbey@arcelik.com
ORCID: 0000-0002-5980-9985

M. Bilginer GÜLMEZOĞLU
Electrical and Electronics Engineering
Eskişehir Osmangazi University
Eskisehir, Turkey
bgulmez@ogu.edu.tr
ORCID: 0000-0002-1570-4989

***Abstract*—Facial expression recognition has significant benefit due to its possible applications in computer vision. Facial expression recognition include machine and human interaction, security, psychology, and changing appearance in social media applications. In this paper, new geometric based features are proposed for the recognition of seven facial expressions. The number of features is reduced by using feature selection methods. Obtained features are applied to Support Vector Machines (SVM) classifier. In the experimental studies, the extended Cohn-Kanade (CK+) dataset is used and facial expressions was classified by the 10-fold cross-validation method. Classification accuracy of 93.5% was achieved in CK+ data with the features selected by the Sequential Backward Feature Selection method.**

***Index Terms*—facial expression recognition, geometric based features, feature selection, support vector machine**

I. INTRODUCTION

Facial expression is one of the most effective, natural and important tools for people to communicate their emotions and objectives [1]. Facial expressions are commonly used in behavioral understanding of emotions, cerebral science, and social collaborations [2]. There are two common methods to obtain facial features: geometric feature-based methods and appearance-based methods [1]. The geometric features introduce the shape and positions of facial components such as mouth, nose, eyes, and eyebrows that are extracted to create a feature vector representing facial geometry. In appearance-based methods, image filters such as Gabor wavelets are applied to the entire face or specific facial areas to remove facial appearance differences, including wrinkles, bumps and grooves

In the last decade, automatic facial expression recognition has been gaining more and more attention and has become a crucial topic in scientific society because facial expressions are one of the most robust, natural, and immediate tools for people to correspond their emotions and intentions [3]. The recognition rate of individual facial expressions is not high, which limits real-time facial expression recognition. The main complexity of independent facial expression recognition is that the appearance of the human face affects the correct acquisition of expressive features [4]. In recent years, researchers have been realized many enhancements to the expression recognition system to improve the recognition rate of facial expression

By taking the median of the landmark tracking results from facial expression sequences training, the classic expression sequence is created for each facial expression class. [2]. Bartlett et al. presented results on an user independent fully automatic system for real time recognition of basic facial expressions from video [5]. They presented an approach for further speed advantage by combining feature selection based on Adaboost with feature integration based on SVM.

Statistical local features, local binary pattern (LBP), are widely used in facial expression recognition especially with the SVM [3], [6]. Appearance and geometric based features are also used with SVM in facial expression recognition [7], [8]. Li et al. proposed a recognition algorithm of person-independent facial expression based on improved LBP (Local Binary Pattern) and HOSVD (Higher-Order Singular Value Decomposition) [4]. In the phase of facial expression classification and recognition, the conventional nearest neighbor classification is modified into k- nearest neighbor pre-classification, and the local energy obtained by HOSVD is utilized to specify the resemblance of two images for secondary classification.

The Convolutional Neural Network (CNN) is another popular method used for facial expression recognition [5], [9]. Liu et al. proposed a FER model based on improved CNN for Sobel edge detection and fused SVM [10]. To solve the problem of facial recognition in face closure, Feng and Shao proposed a human eye facial expression recognition model for transfer learning [11]. Bartlett et al. introduced a methodical comparison of machine learning methods employed to the automatic recognition problem of facial expressions [12]. They stated results on a sequence of experiments comparing recognition methods, including AdaBoost, SVM, linear discriminant analysis.

In this paper, seven facial expressions in the CK+ dataset are classified. For this purpose new geometrical-based features are extracted from the landmark points on the facial images. These features and the features selected by three selection methods are applied to the SVM classifier.

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

Fig. 1: (a) and (b) denotes feature landmark points, (c) and (d) denotes selected 28 features from SBFS.

The paper is organized as follows: the feature extraction and selection are given in Sections II and III. The classifier and experimental study are presented in Sections IV and V. Finally, Section VI includes the conclusion.

II. Feature Extraction

Feature extraction is the first phase of facial expression recognition that is mainly composed of three stages: face detection, facial landmark tracking, extracting features from the landmark tracking result. We used the Viola-Jones face detection algorithm in the face detection phase [13]. Facial landmarks are taken from the dlib's facial landmark detector [14]. 68 facial landmarks are obtained by using this detector. In literature, most of the researchers use both neutral state and expressions state in the feature extraction. However, in our work, the first group of extracted features corresponds to the Euclidean distances between two odd-numbered landmarks from 1 to 67, 3 to 67, … etc. The second group of extracted features corresponds to the Euclidean distances between two even-numbered landmarks from 2 to 68, 4 to 68, … etc. Feature vectors were created by combining these two groups and a total of 1122 features were extracted.

III. Feature Selection

Most of the feature selection methods can be divided into 3 main categories. These are filter methods, wrapper methods, and tree-based models. In this study, three feature selection methods are used to reduce the size of feature vectors. These methods are two Wrapper-Methods(Sequential Forward Feature Selection and Sequential Backward Feature Selection) and Principle Component Analysis (PCA), which is very common in dimension reduction. They are briefly described in the subsections below.

A. Sequential Forward Feature Selection (SFFS)

In SFFS, first, the best single feature is selected, then two pairs of features are formed using one of the remaining features and this best feature, and the best feature pair is selected. In the third step, triplets of features are taken using one of the remaining features, and these two best features and the best triplet is selected. These steps continue until the best representative subset of features is selected. After the application of SFFS, we have the best 60 features.

B. Sequential Backward Feature Selection (SBFS)

The second feature selection method is SBFS. In this method, firstly, the classification function is computed for all features. Then, each feature is deleted once at a time, the classification function is computed for all subsets with all minus deleted one, and one feature having the worst classification rate is discarded. These steps continue until the best subset of features is selected. After the application of SBFS, we have the best 28 features

C. Feature Selection with Principal Component Analysis

Principal Component Analysis (PCA) is a technique for, especially dimension reduction. The main purpose of PCA is to keep the data set with the highest variance in high dimensional data but to provide dimension reduction while doing this. It provides a lower dimension by finding the general properties of the given dimension. Certain features will be lost with size reduction; but intended, these disappearing have little informational characteristics about the classification. In this study, we applied PCA as a feature selection method. In the CK+ dataset, the number of eigenvalues is tuned from 2 features to all features. Best recognition results are obtained by the selection of 268 features.

IV. SVM Classifier

SVM are a couple of supervised learning methods used for classification, regression, and outliers detection [15]. In this classifier, a datum item is indicated as a point into the n-dimensional space along with the value of each feature corresponding to a specific coordinate. Also, the classification is realized by finding the hyper-plane that discriminates the classes. In the SVM, if there is no linear hyper-plane between two or more classes, a method called the kernel trick is applied. In this study, the RBF kernel is applied and classified 7 facial expression classes.

V. Experimental Study

In the experimental study, the dataset used, and test results are mentioned in the following subsections.

A. Dataset

In the Facial Expression Recognition study, the Extended Cohn-Kanade (CK+) dataset [16] was used. This dataset was widely used by many researchers [1]–[4], [7], [8]. The dataset contains facial images from 123 subjects with 7 emotions (i.e., anger, contempt, disgust, fear, happiness, sadness, and surprise). In our study, the first of frame sequence belonging to each emotion is chosen as a neutral frame. Also, the last two or three emotional frames have been chosen to increase the number of samples on each emotion in the dataset. This leads to a total of 989 images of 8 classes. For a fair comparison with other researchers, we use 7 facial expression classes without a neutral class in the recognition process. Finally, we have 866 images whose distributions according to the classes are given in Table-I.

TABLE I: The number of samples for each classes in CK+ dataset

Anger	Contempt	Disgust	Fear	Happiness	Sadness	Suprise	**Total**
90	82	158	121	138	111	166	**866**

B. Results and Discussions

During the classification process, SVM is used together with the RBF kernel. 10-fold cross-validation is used to obtain average accuracy results. In this method, the dataset is divided into 3 parts as training, testing, and verification. 88.2% recognition rate is obtained from the RBF kernel of SVM using features without selection. The confusion matrix for this case is given in Figure 2. 89.5% and 93.5% recognition rates are obtained using the RBF kernel of SVM for the features selected by SFFS (60 features) and SBFS (28 features) respectively. The confusion matrices with these cases are given in Figures 3 and 4 respectively. When the 268 features selected

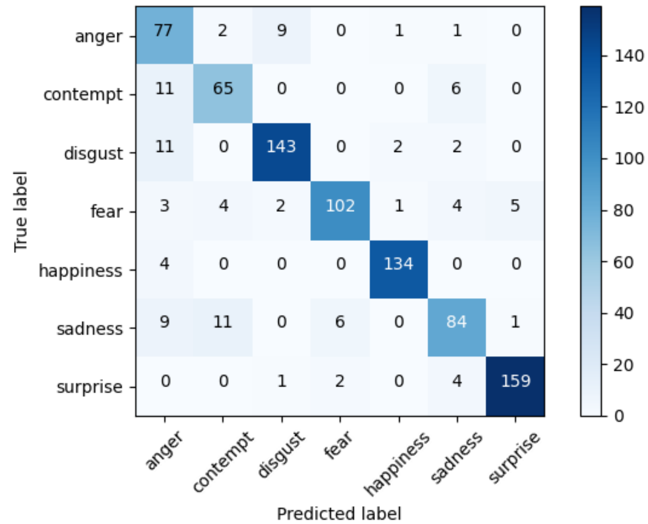

Fig. 2: Confusion matrix obtained for the features without selection.

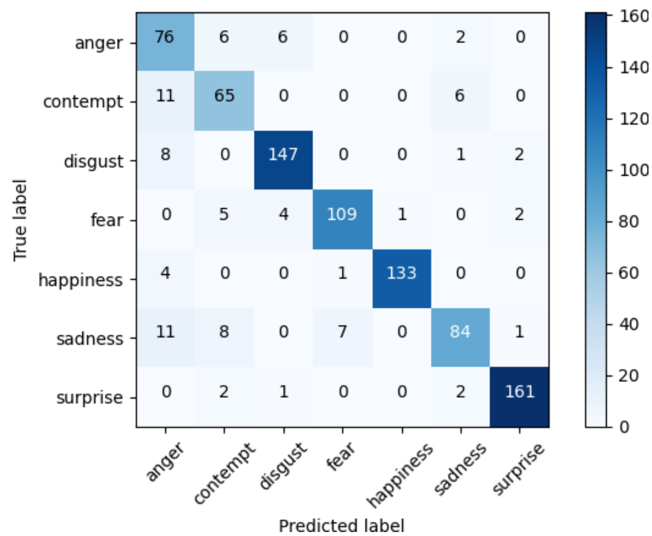

Fig. 3: Confusion matrix obtained for the features selected by SFFS.

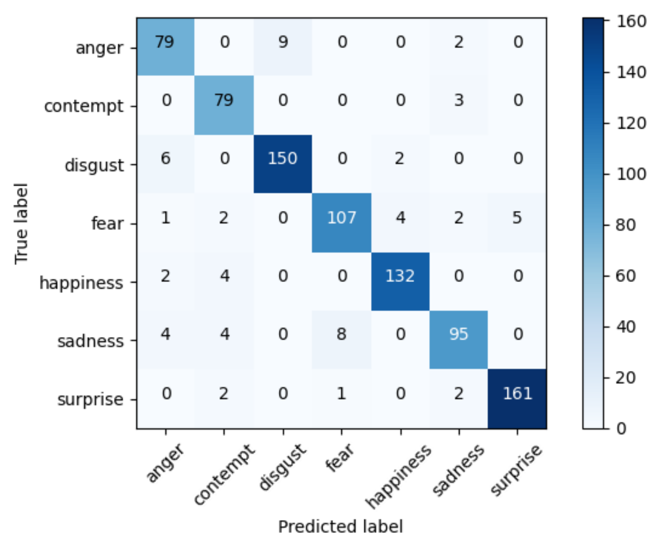

Fig. 4: Confusion matrix obtained for the features selected by SBFS.

TABLE II: The Classification Results on the CK+ Dataset

	Accuracy on CK+ Dataset in Percentage				
Feature Selection Method	Proposed Method	RBF Kernel SVM [7]	Linear Kernel SVM [12]	*LDA [12]	RBF Kernel SVM [17]
None	**88.2**	84.7	88.0	44.4	81.1
SFFS	**89.5**	88.7	–	–	–
SBFS	**93.5**	–	–	–	–
PCA	**90.8**	–	75.5	80.1	–

LDA : Linear Discriminant Analysis

by PCA are used RBF Kernel of SVM gives 90.8% recognition rate. The confusion matrix of RBF kernel SVM classifier for feature selection with PCA is given in Figure 5.

The feature vectors which are obtained by Euclidean distances are classified by 10-fold cross-validation with RBF kernel SVM. The results of the feature selection methods are given in Table-II. The best classification accuracy of 93.5% is achieved in the CK+ dataset with 28 features selected by SBFS in the proposed method. This result is also higher than the results given in references [7] and [17] when the RBF kernel SVM is used as a classifier.

When linear kernel SVM and Linear Discriminant Analysis (LDA) are used as in [12], the classification accuracies are less than those of the proposed method. On the other hand, the study in [12] has inspired our study in order to apply feature selection methods on the geometric features. It has been proven that facial expression recognition accuracy will increase by using PCA and Adaboost when selecting features. The proposed method has also achieved more successful facial expression recognition when compared with feature selections are given in [7] and [12]. The results show better classification accuracy than [17] when only the features are used.

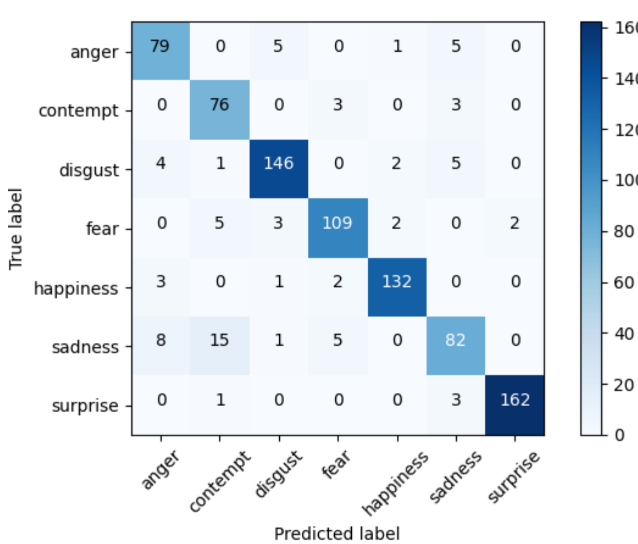

Fig. 5: Confusion matrix obtained for the features selected by PCA.

VI. Conclusion

In this study, different geometric features of landmark points on the facial images are proposed in facial expression recognition instead of usual landmarks. The Euclidean distances between two of the landmark points by skipping one landmark point are used when obtaining geometric features. When these features are used to classify seven facial expressions with the SVM classifier,the best results are obtained with the proposed method compared with the results in the literature.

Three different feature selection methods are applied during the facial expression classification stages due to the negative impact of some features. When SBFS is used 28 features are selected and the classification accuracy of 93.5% is obtained. When SFFS and PCA are used 60 and 268 features are selected and the classification accuracies of 89.5% and 90.8% are obtained respectively. These results are also higher than the results given in [7], [12], [17].

In feature studies, it is aimed to achieve facial expression recognition with a higher success by applying geometric and appearance-based features together. In addition, it is planned to expand the work with deep learning methods to increase the classification performance.

Acknowledgment

We would like to thank Prof. Jeffery Cohn and Megan Ritter for the use of the Extended Cohn–Kanade dataset.

References

[1] 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, pp. 803–816, 2009.

[2] D. Ghimire and J. Lee, “Geometric Feature-Based Facial Expression Recognition in Image Sequences Using Multi-Class AdaBoost and Support Vector Machines”, Sensors, Vol. 13, pp. 7714-7734, 2013; doi:10.3390/s130607714.

[3] X. Zhao and S. Zhang,” Facial Expression Recognition Based on Local Binary Patterns and Kernel Discriminant Isomap”, Sensors, Vol. 11, pp. 9573-9588, 2011; doi:10.3390/s111009573.

[4] Y. He and S. Chen, “Person-Independent Facial Expression Recognition Based on Improved Local Binary Pattern and Higher-Order Singular Value Decomposition”, IEEE Access, Vol. (, pp. 190184-190193, 2020; doi: 10.1109/ACCESS.2020.3032406.

[5] M. S. Bartlett, G. Littlewort, I. Fasel and J. R. Movellan, “Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction”, Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop (CVPRW03), 16-22 June, Madison, Wisconsin, USA.

[6] Caifeng Shan, Shaogang Gong and P. W. McOwan, ”Robust facial expression recognition using local binary patterns,”IEEE International Conference on Image Processing 2005, Genova, 2005, pp. II-370, doi: 10.1109/ICIP.2005.1530069.

[7] C. Gacav, B. Benligiray and C. Topal, ”Greedy search for descriptive spatial face features,”2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, 2017, pp. 1497-1501, doi: 10.1109/ICASSP.2017.7952406.

[8] T. Mohammad and M. L. Ali, ”Robust facial expression recognition based on Local Monotonic Pattern (LMP),”14th International Conference on Computer and Information Technology (ICCIT 2011), Dhaka, 2011, pp. 572-576, doi: 10.1109/ICCITechn.2011.6164854.

[9] L. S. Videla and P. M. A. Kumar, ”Facial Expression Classification Using Vanilla Convolution Neural Network,” 2020 7th International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 2020, pp. 1-5, doi: 10.1109/ICSSS49621.2020.9202053.

[10] S. Liu, X. Tang and D. Wang, ”Facial Expression Recognition Based on Sobel Operator and Improved CNN-SVM,” 2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP), Shanghai, China, 2020, pp. 236-240, doi: 10.1109/ICICSP50920.2020.9232063.

[11] H. Feng and J. Shao, ”Facial Expression Recognition Based on Local Features of Transfer Learning,”2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqing, China, 2020, pp. 71-76, doi: 10.1109/ITNEC48623.2020.9084794.

[12] M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel and J. Movellan, ”Recognizing facial expression machine learning and application to spontaneous behavior,” 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), San Diego, CA, USA, 2005, pp. 568-573 vol. 2, doi: 10.1109/CVPR.2005.297.

[13] P.A. Viola and M.J. Jones, ”Robust real-time face detection”, IEEE International Conference on Computer Vision (ICCV), 2001.

[14] Davis E. King. Dlib-ml: A Machine Learning Toolkit. Journal of Machine Learning Research 10, pp. 1755-1758, 2009

[15] M.A. Hearst ”Support Vector Machines,” in IEEE Intelligent Systems and their Applications , 1998.

[16] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z. and Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101

[17] D. Ghimire, S. Jeong, S. Yoon, J. Choi and J. Lee, ”Facial expression recognition based on region specific appearance and geometric features,” 2015 Tenth International Conference on Digital Information Management (ICDIM), Jeju, 2015, pp. 142-147, doi: 10.1109/ICDIM.2015.7381857.