*J*ournal of ***AI*** and ***D***ata ***M***ining
Vol 7, No 1, 2019, 17-25

DOI: 10.22044/JADM.2017.4676.1554

Improving LNMF Performance of Facial Expression Recognition via Significant Parts Extraction using Shapley Value

M. Rezaei[1, 2] and V. Derhami[1*]

1. Computer Engineering Department, Faculty of Engineering, Yazd University, Yazd, Iran.

2. Faculty of Electrical and Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran.

Received 28 August 2016; Revised 31 May 2017; Accepted 07 November 2017
**Corresponding author: Vderhami@yazd.ac.ir (V. Derhami).*

Abstract

The non-negative Matrix Factorization (NMF) algorithms have been utilized in a wide range of real applications. NMF has been used by several researchers for its part-based representation property, especially in the facial expression recognition problems. It decomposes a face image into its essential parts (e.g. nose, lips). However, in all the previous attempts, it has been neglected that all features achieved by NMF are not required for recognition problems. For example, some facial parts do not have any useful information regarding the facial expression recognition. In this work, addressing the challenge of defining and calculating the contributions of each part, the Shapley value is used. It is applied for identifying the contribution of each feature in the classification problem, and then the effectless features are removed. Experiments performed on the JAFFE and MUG facial expression databases, as benchmarks of facial expression datasets, demonstrate the effectiveness of our approach.

Keywords: *Non-negative Matrix Factorization (NMF), Shapley Value, Game Theory.*

1. Introduction

Facial expressions play an important role in the face-to-face human communications. During human communications, information transmitted by language, tone, and expressions accounts for 7%, 38%, and 55% of personal communication, respectively. This statistic exhibits the importance of facial expressions in mutual understandings [1]. Facial expression recognition is widely used in the actor training software to train inexperienced actors. It gives them a feedback about how well they can mimic different facial states or gesture such as fear, happiness, astonishment, neutral, and sadness. The mentioned facial expressions are produced by certain voluntary contractions of the facial muscles. Therefore, each face expression is considered as an information class that conveys a certain non-verbal message to the audience.

In order to design a facial expression recognition system, a standard dataset should be provided. Then the discriminative component-based features should be extracted from the images for quantitatively describing the state of the face components such as the lips, eyes, and eyebrows, and also the flexion or extension of the facial muscles. A number of papers have addressed the problem of facial expression recognition. They include methods that use Local Binary Patterns (LBP) [2], static topographic modeling [3], texture and shape information fusion [4], Principal Component Analysis (PCA) [5], appearance modeling [6], integration of facial expression with facial appearance models [7], Local gradient pattern (LGP)[8] ,and Non-negative Matrix Factorization (NMF) [9].

Among the mentioned techniques, NMF is the only one that can represent the input data in terms of its sub-structures. For instance, a face image by NMF decomposition can be expressed as multiplication of two low-rank matrices in terms of the basis matrix (W) and its corresponding weight matrix (H). The columns of matrix W can be restructured in the form of low-size images, each exhibiting a face component such as the lips, eyes, eyebrow, nose, and chin. Since each face gesture affects just

certain face components, part-based decomposition (e.g. NMF) is naturally matched for analyzing the facial expression. This issue provides discriminant features to enhance the recognition performance. Moreover, its non-negativity property enables NMF to demonstrate its features, while other methods do not have the ability elicit demonstrable features (positive values) from an image.
A formal analytical framework with a set of mathematical tools is provided by the game theory to assay the complex interactions between the players [10]. This theory is used in a large number of fields including engineering and economics [11-14]. The Shapley value, one of the most important solution concepts of the cooperative games, was introduced by Shapley [10]. It is used to divide surplus according to contribution. Hence, in the proposed method, a contribution of an element is calculated using the Shapley value.
As NMF is a non-convex problem, all the papers published until now are about some issues such as good initialization, cost functions, and some constraints. For example, sparse NMF, introduced in 2004, extends non-negative matrix factorization by adding a constraint to control sparsity of the matrix *H* [15].
In this work, high impact features are determined; the Shapley value is used for this reason. It is neglected that all the features achieved by NMF or its variants are not necessary for recognition problems. Some facial parts do not have any useful information regarding the facial expression recognition. Addressing the challenge of defining the contributions of each part, we use the Shapley value. It determines the contribution of each feature in the classification problem; then the effectless features are removed. As NMF is an unsupervised problem, in this paper, a semi-supervised version of NMF is introduced, which uses a logical combination of LNMF and the Shapley value. In the proposed approach, the parts of faces are elicited with LNMF, and then the low-impact features are discovered and removed.
The rest of this paper is organized as what follows. In Section 2, the NMF and Shapley value methods are described. In Section 3, the proposed method is explained in detail. In Section 4, we describe two benchmark datasets that are used in this paper. Section 5 illustrates the experimental results obtained and discusses the advantages of the proposed method compared to the PCA and NMF methods. Finally, the paper concludes in Section 6.

2. Primary Concepts

In this section, the methods along with the proposed method are illustrated. It includes three sub-sections in terms of NMF, Local NMF (LNMF), and the Shapley value, respectively.

2.1. Non-negative matrix factorization (NMF)

NMF is a low-rank approximation technique for unsupervised multivariate data decomposition, which was introduced by Lee and Seung in 1999 [16]. It acts similar to PCA and Independent Component Analysis (ICA) but with different constraints and interpretations [17].
In the last decade, many research works have been reported on the analysis of the extensions and applications of the NMF algorithm in image processing [18-24], signal processing [25-28], and data mining [29, 30].
NMF decomposes a given non-negative data matrix (e.g. Image, document) $A \in R^{m\times n}$ into a multiplication of the two non-negative matrices $W \in R^{m\times k}$ and $H \in R^{k\times n}$ such that these matrices minimize the following criterion:

$$f(W,H) = \frac{1}{2}\|A - WH\|_F^2 \quad (1)$$

where, *k<<min (m, n)* is a positive integer that determines the rank of the NMF and *F* is the Frobenius norm. In some works, different similarity metrics have been used as the cost function [20, 26].
In the image processing field, $A_{m\times n}$ contains the whole image database, in which *m* is the pixel number of each vectorized image and *n* is the number of the images. Hence, applying NMF leads to obtain *k* dimensional features (i.e. *k<<m*, where *m* is the original dimension of features). Advantages of NMF in comparison with the other factorization methods such as PCA, SVD, ICA, and QR are as follow:

- Elements of *W* are non-negative; consequently, basis columns can be visualized [9].
- Non-negative elements of *H* make a non-subtractive combination of basis components compared to the other methods that allow for additive and subtractive combinations of basis components. This property gives NMF a part-based representation in contrast to the other whole-based representation methods. This is a desirable property in the fields such as image and document processing. For example, if a part of an image is damaged, the defected part affects just a small number of features; however, in other factorization methods, almost all features are affected. Therefore, NMF is

applicable in the fields of image and document processing when only a part of an image or document is damaged. In addition, in most applications, only a part of an image is processed. For instance, in facial expression recognition, face expression is shown by a few parts of a face.

- NMF causes a sparse representation of the input data in terms of the W and H matrices, leading to a reduction in the data storage [21].

2.2. Local non-negative matrix factorization (LNMF)

Although the features found by NMF are part-based, they are not fully localized. In order to solve this problem, LNMF was introduced in 2001 [31]. It identifies the W and H components such that these matrices minimize the following criterion [19]:

$$f_{LNMF(W,H)} = \sum_{i=1}^{m}\sum_{j=1}^{n}(A_{ij}\log\frac{A_{ij}}{[WH]_{ij}} - A_{ij} + [WH]_{ij} + \alpha U_{ij}) - \beta\sum_{i} V_{ii} \quad (2)$$

where, f_{LNMF} is the LNMF cost function and $\alpha, \beta > 0$ are constants, where $U = W^T W$ and $V = HH^T$. Although the rate of convergence for LNMF is slower than NMF, the features obtained by LNMF are highly localized. Consequently, it is expected that LNMF decomposes a face image into its essential parts (e.g. nose, lips), where each face component is arranged to the W columns. The LNMF algorithm updates the factorized matrices of W and H according to the following formula in an iterative manner:

$$H_{kl} = \sqrt{H_{kl}\sum_{i} A_{il}\frac{W_{ik}}{\sum_{k} W_{ik}H_{kl}}} \quad (3)$$

$$W_{kl} = \frac{W_{kl}\sum_{j} A_{kj}\frac{H_{lj}}{\sum_{k} W_{kl}H_{lj}}}{\sum_{j} H_{lj}}$$

$$W_{kl} = \frac{W_{kl}}{\sum_{i} W_{il}}$$

where, A_{ij}, W_{ij}, and H_{ij} are the elements located in the i'th row and j'th column of A, W, and H matrices, respectively. According to the chosen cost function, three additional constraints are imposed on the NMF basis vectors in terms of maximum sparsity in H, maximum expressiveness of W, and maximum orthogonality of W [19]. Thus it is expected to achieve better part-based features than standard NMF.

2.3. Shapley value

A coalitional game [32] is defined as the pair (N, v), where $N = \{1,2,\ldots,n\}$ is the set of players and $v : 2^N \to R$ is a real-valued mapping called the characteristic function. It assigns to every possible coalition a numeric value. It represents the payoff that may be distributed among the members of that coalition. In the game theory, the Shapley value [10] is a solution concept in the cooperative game theory. For each cooperative game, it allocates a unique distribution (among the players) of a total surplus created by the coalition of all players. It demonstrates an efficient approach to the fair division of the gains achieved by cooperation between players of a cooperative game. Fairly distribution is an important issue because some players may contribute more than the others. Consequently, the more each player contributes, the more he gains benefit. According to the Shapley value, the amount that the i^{th} player obtains in the coalitional game (N, v) is:

$$sh_i = \frac{1}{|N|!}\sum_{o\in\Pi(N)} \mu_i(C_i(o)) \quad (4)$$

where, $\Pi(N)$ is the set of all possible orderings of the agents, $C_i(o)$ is the set of agents preceding i in ordering o, and μ_i (C) is the marginal contribution of i to C that is the amount that i adds by joining $C \subseteq Ag/\{i\}$. It is calculated as below:

$$\mu_i(C) = v(C \cup i) - v(C) \quad (5)$$

In fact, the Shapley value of a player is achieved by the weighted mean of its marginal value in all possible permutations of the players.

3. Proposed method

In this paper, we introduce a novel method for eliciting the important parts of face. The proposed method is based upon the extracting parts of face with LNMF and then identifying the contribution of each part with the Shapley value.

As described in Section 2.2, $A_{m\times n}$ contains the whole image database, in which m is the pixel number of each vectorized image and n is the number of images. In the first step, LNMF is applied, which leads to obtain k, low dimensional features.

According to the literature, all features achieved by LNMF are used for classification. However, in fact, all parts of the faces do not influence the recognition. Here, we propose an approach to eliminate the low-impact parts using the Shapley value. In this algorithm, the Shapley values of features are calculated, and the features with values lower than the defined threshold are removed. In order to calculate the Shapley value, all permutations (for n features, $n!$ permutations) should be assessed. In fact, it can be proved that the problem of computing the Shapley value is an NP-complete problem [33, 34].

For addressing this problem, estimation of the Shapley value is used. In this work, the ApproShapley algorithm [35] is applied, which has a polynomial calculation of the Shapley value based on sampling. For this reason, S samples are randomly selected with a probability of $1/_{N!}$. The contribution of each part is evaluated with its Shapley value. Finally, we remove the parts whose Shapley values are lower than the threshold. Figure 1 shows the block diagram of the proposed method; moreover, the pseudo-code is given bellow:

Inputs: A matrix that is achieved by pre-processing phase.

Outputs: The extracted features.

Set columns of matrix A to vectors A_i *(i=1... n)*
Determine k
W and H are initialized randomly,
For i=1: maximum iteration for LNMF
  Update H and W matrices trough Eq. (3)
End
Determine S
Cont=0 and $Sh_i = 0 \ \forall i \in N$
While Cont<S
  $Take\ O\epsilon\pi(N) with\ probablity\ 1/_{N!}$
  For i=1:k
    $Sh_i = Sh_i + \mu_i(C_i(O))$
  Cont=Cont+1;
$Sh_i = {Sh_i}/_{S}$
For i=1:k
  If Sh_i*>T*
    i'th features is added to output features.

4. Datasets

In this section, we describe two publicly available benchmark datasets.

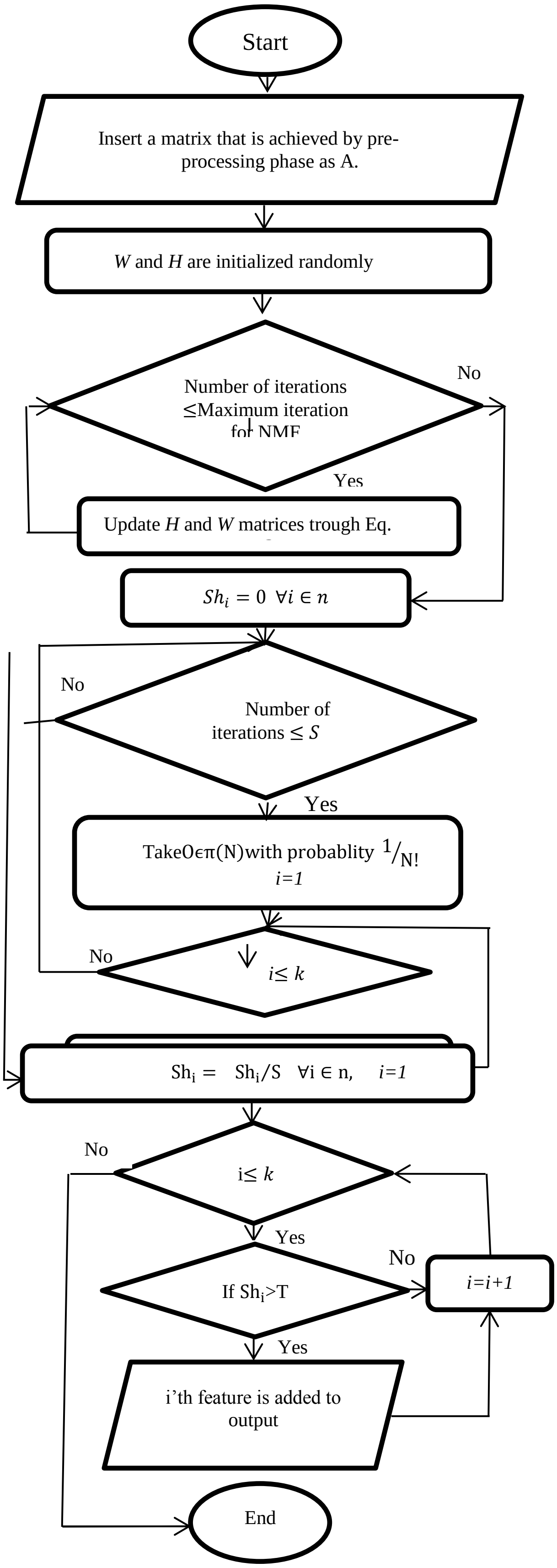

Figure 1. Block diagram of the proposed method.

4.1. JAFFE dataset

The Japanese Female Facial Expressions dataset [36] consists of 213 grayscale images presenting seven facial expressions (happiness, sadness, fear, anger, surprise, disgust, and neutrality) that were posed by 10 Japanese female models. Each image size is of 256 × 256 pixels, and each model has 2–4 samples for each expression. Example images from this dataset are shown in figure 2.

4.2. MUG facial expression dataset

The MUG-FED database [37]contains 401 grayscale images of 25 people (18 men and 7 women between 20 and 35 years of age), each with seven different facial expressions. There are 1-3 samples for each expression type per person. Example images from this dataset are shown in figure 2.

5. Experimental results and discussion

In the pre-processing stage, the facial parts of the images are cut from both sides of the eyes and from top of the eyebrows to the bottom of the chin and aligned into a fixed size (Figure 3). For each aligned image, histogram equalization is applied, and the intensity of pixels at each image is normalized. In the second step, the rank of LNMF is obtained. We apply different k and calculate the classification rates. As shown in figure 4, the highest classification rates are 22 and 24 for the JAFFE and MUG datasets, respectively.

Then LNMF is applied; figures 5 and 6 show the images that were vectorized in W columns for the JAFFE and MUG datasets, respectively. They are considered as primary extracted features. As expected, the extracted parts are local and consist of the significant components of face. In order to clarify that why LNMF is used for the extracting features, the parts achieved by Sparse NMF and NMF are also shown in figures 7-10. The results obtained show that the features achieved by LNMF are more local; it is the main contribution of the LNMF algorithm.

As stated for the proposed method, one can see that all parts in figure 5 (or Figure 6, for the MUG dataset) are not required for facial expression recognition. The Shapley value is used for estimating the contribution of each feature in facial expression recognition; the face parts are mapped to the features, and the payoff is represented by the Shapley value. The Shapley value for each feature is calculated by sampling in space of k! possible permutations (22! in the JAFFE dataset and 24! in the MUG dataset) to show the performance of each part.

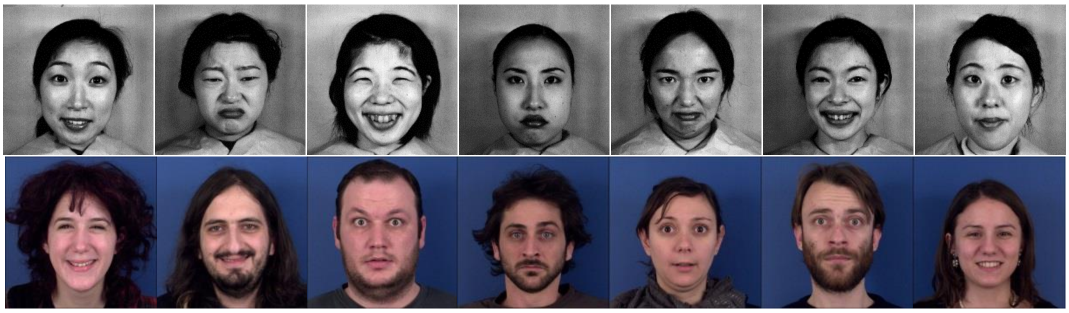

Figure 2. Images in the first row show some examples of JAFFE dataset and images in the second row are from the MUG dataset.

Figure 3. Images in the first row show some examples of JAFFE dataset and images in the second row are from the MUG dataset after pre-processing.

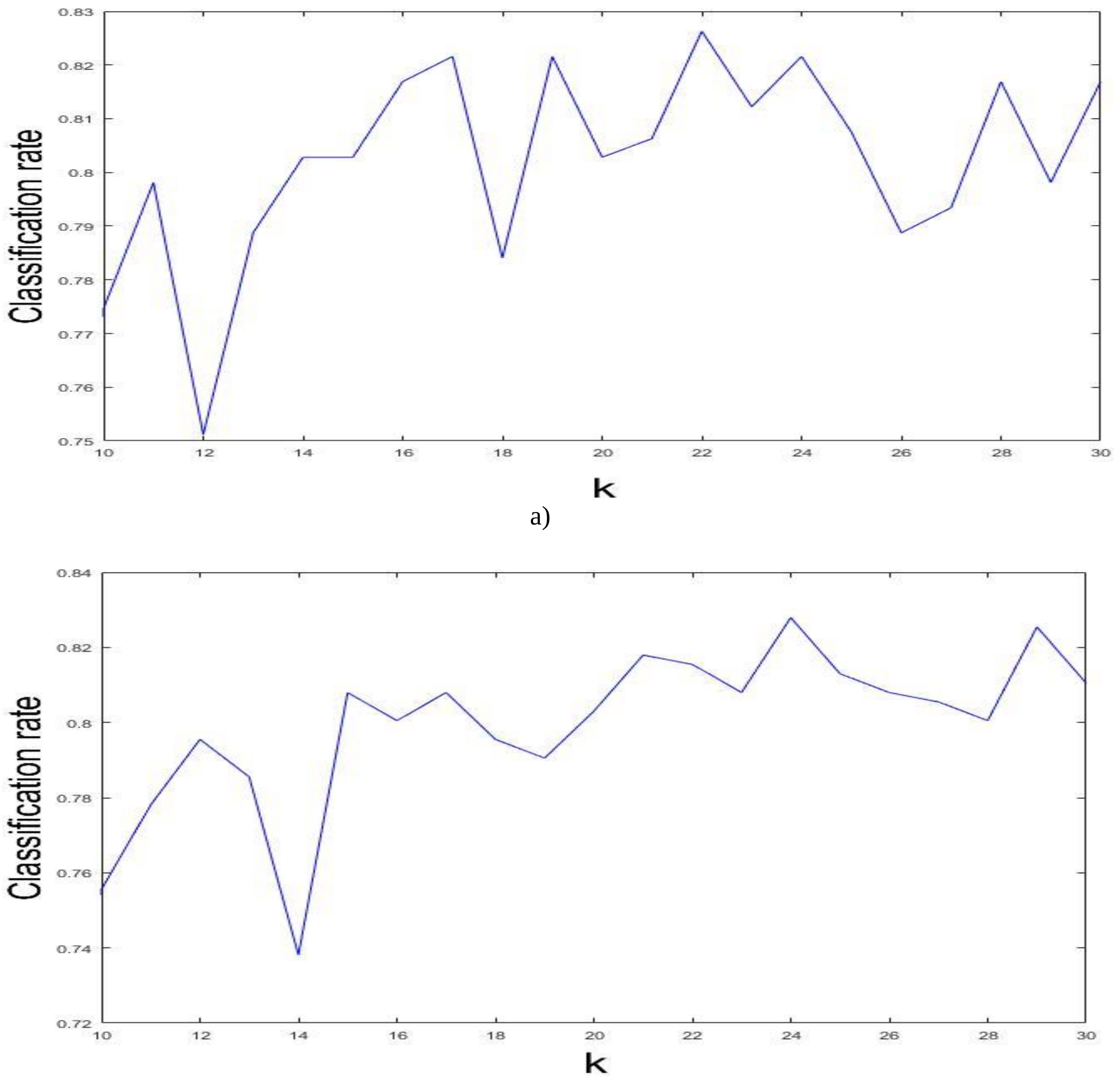

Figure 4. Classification rates achieved by different ranks in (a) JAFFE dataset and (b) MUG dataset.

In our experiment, *S* (number of random selected samples) is set to 5000. In the determination of the amount of *S*, there is a trade-off between accuracy and complexity. If S is selected to be high, the results are more accurate and complex, and vice versa. Here, based on our experiment, S is set to 5000. Threshold of selecting features (*T*) controls the number of output features. Based on our experiment, the number of features that are obtained by setting the variable S equal to 3 for the JAFFE dataset and 5 for the MUG dataset are enough.

The features are sorted and the ones with Shapley values lower than *T* are removed. The left-side images in figure 11 illustrate the final extracted features after removing the effectless ones. The right images of this figure are achieved by adding the values of the left side to show the most important parts of face for expression recognition. They are added for subjective representation, where the extracted features with the proposed method are parts of face with the highest effect on the facial expression recognition, and the goal is to achieve these parts.

In other words, the lighter parts of the right image in figure 11 are more effective in recognition rather than the darker parts. In principle, the number of features in the JAFFE dataset is decreased from 1089 to 6, and in a similar manner, the number of features in the MUG dataset is decreased to 9. After extracting the features using the proposed method, the KNN algorithm with *K*=1 is applied, and the classification rate is calculated. In this work, Leave-One-Out (LOO) was used for cross-validation. LOO applies one sample as the test set and all the remaining samples as the training set, and this step is repeated for all observations. The most important advantage of LOO is selection of all samples once for testing.

The experimental results achieved by applying the proposed method along with PCA, NMF, LNMF, and sparse NMF are depicted in table 1. These results show that the proposed method outperforms other methods in term of accuracy. It is meaningful because there are features that not only have no effect on facial expression recognition but also increase the error rate.

When additional features are deleted, their destructive effect in creating distance in two identical emotional states is removed and the accuracy is increased.

5. Conclusion

In this paper, we have proposed a feature extraction algorithm based on non-negative matrix factorization (NMF). NMF is a part-based representation, and has been applied to many applications. The problem is that all face parts achieved by NMF are not necessary for the recognition problem.

The effectless features have destructive effects on creating distance, and they cause more error. In order to solve this shortcoming, we employed the Shapley value, a well-known approach in game theory, in order to calculate the contribution of each part to facial expression recognition. The experiments performed on the JAFFE and MUG datasets demonstrated the effectiveness of our approach.

Figure 5. Features achieved by LNMF for JAFFE dataset.

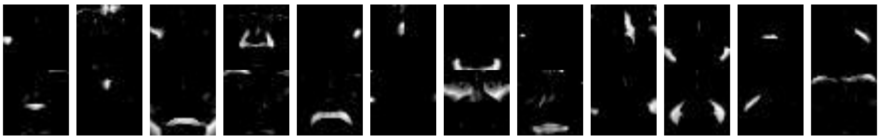

Figure 6. Features achieved by LNMF for MUG dataset.

Figure 7. Features achieved by sparse NMF for JAFFE dataset.

Figure 8. Features achieved by sparse NMF for MUG dataset.

Figure 9. Features achieved by NMF for JAFFE dataset.

Figure 10. Features achieved by NMF for MUG dataset.

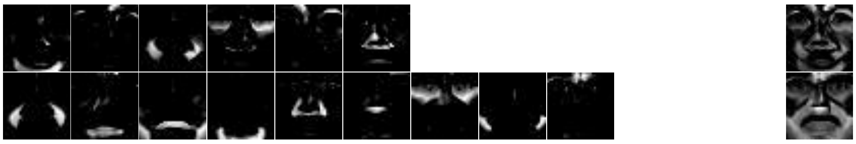

Figure 11. The left side parts are final extracted features t achieved by the proposed method. The right side image is achieved by adding the values of the left side. Top images are related to JAFFE and bottom images to MUG dataset.

Table 1. Comparison of PCA, NMF, LNMF, sparse NMF, and the proposed method in term of classification rate.

	PCA	LNMF	NMF	Sparse NMF	Proposed method
JAFFE	85.91	79.81	84.03	81.22	86.85
MUG	82.79	81.04	82.79	80.54	84.62

References

[1] Zhao, L., Zhang, G. & Xu, X. (2008) Facial Expression Recognition Based on PCA and NMF.IEEE Int. Conf., intelligent control and automation, pp. 6826 -6829.

[2] Shan, C., Gong, S. & Mcowan, P. W. (2009) Facial expression recognition based on local binary patterns: a comprehensive study. Elsevier, Image and Vision Computing, vol. 27, no. 6, pp. 803-816.

[3] Wang, J. & Yin, L. (2007) Static topographic modelling for facial expression recognition and analysis. Elsevier, Computer Vision and Image Understanding, vol. 108, no. (1–2), pp. 19-34.

[4] Kotsia, I, Zafeiriou, S., Nikolaidis, N. & Pitas, I. (2008) Texture and shape information fusion for facial expression and facial action unit recognition, IEEETrans, on Pattern Recognition, vol. 41, no. 30, pp. 833-851.

[5] Calder, A. J. & Burton, A. M., Miller, P. & Young, A. W., Akamatsu S. (2001). A principal component analysis of facial expressions. Elsevier, Vision Research, vol. 41, no. 9, pp. 1179-1208.

[6] Abboud, B., Davoine, F. & Dang, M. (2004). Facial expression recognition and synthesis based on appearance model. Elsevier, Signal Processing, vol. 19, no. 8, pp. 723-740.

[7] Tsai, P., Cao, L., Hintz, T. & Jan, T. (2001). A bi-modal face recognition framework integrating facial expression with facial appearance, Elsevier, Pattern Recognition Letters, vol. 30, no. 12, pp. 1096-1109.

[8] Shahidoleslam, M. (2014). Local gradient pattern - A novel feature representation for facial expression recognition. Journal of AI and Data Mining, vol. 2, no. 1, pp. 33-38.

[9] Wild, S., Curry, j. & Dougherty, A. (2004). Improving nonnegative matrix factorizations through structured initialization. IEEETrans.Pattern Recognition, vol. 37, no. 11, pp. 2217-2232.

[10] Shapley, L. S. (1953). A Value for n-person Games. In Contributions to the Theory of Games, volume II, by H.W. Kuhn and A.W. Tucker, editors. Annals of Mathematical Studies, vol. 28, pp. 307–317. Princeton University Press.

[11] Khanizad, R. & Montazer, G. (2017). Participation against Competition in Banking Markets based on Cooperative Game Theory. The Journal of Finance and Data Science, In Press.

[12] Rezazadeh, A., Zhang, L., Reniers, G., Khakzad, N. & Cozzani, V. (2017). Optimal patrol scheduling of hazardous pipelines using game theory. Process Safety and Environmental Protection, vol. 109, pp.242-256.

[13] Kose, E., Erbas, M. & Ersen, E. (2017). An integrated approach based on game theory and geographical information systems to solve decision problems. Applied Mathematics and Computation, vol. 308, pp.105-114.

[14] Hannesson, R. (2017) Shared stocks, game theory and the zonal attachment principle. Fisheries Research. In Press.

[15] Hoyer, P. O. (2004). Non-negative matrix factorization with sparseness constraints, Journal of Machine Learning Research, vol. 5, pp. 1457-1469.

[16] Lee, D. & Seung, H. (1999). Learning the Parts of Objects by Nonnegative Matrix Factorization. Nature 401. pp. 788–791.

[17] Hyvärinen, A., Karhunen, J. & Oja, E. (2001) Independent Component Analysis, New York: Wiley.

[18] Dou, J., Li, J., Qin, Q. & Tu, Z. (2015) Robust visual tracking based on incremental discriminative projective non-negative matrix factorization. Neurocomputing, vol. 166, pp.210-228.

[19] Feng, T., Li, S. Z., Shum, H. Y. & Zhang, H. (2002). Local Non-Negative Matrix Factorization as a Visual Representation.IEEE Int. Conf., Development and Learning, pp. 178-183.

[20] Finesso, L. & Spreij, L. (2006). Nonnegative matrix factorization and I-divergence alternating

minimization Elsevier.Linear Algebra and its Applications, vol. 416, no. 2c-3, pp.270-287.

[21] Langville, A., Meyer, C., Albright, R., Cox, J. & Duling, D. (2006) Algorithms, initializations, and convergence for the nonnegative matrix factorization, Preprint.

[22] Liao, Q. & Zhang, Q. (2016). Local coordinate based graph-regularized NMF for image representation, Signal Processing, vol. 124, pp.103-114.

[23] Zheng, Z., Yang, J. & Zhu, Y. (2007) Initialization enhancer for nonnegative matrix factorization. Engineering Applications of Artificial Intelligence, vol. 20, no. 1, pp.101-110.

[24] Ying, Z. & Zhang, G. (2009). Facial Expression Recognition Based on NMF and SVM. IEEE Int. Conf., International Forum on Information Technology, pp. 458-462.

[25] Chen, Z. & Cichocki, A. (2005). Nonnegative matrix factorization with temporal smoothness and/or spatial decorrelation constraints. Technical Report, Laboratory for Advanced Brain Signal Processing, RIKEN, Tokyo, Japan.

[26] Cichocki, A., Zdunek, R. & Amari, S. (2006) Csiszar's divergences for Nonnegative Matrix Factorization: Family of new algorithms, Conference on Independent Component Analysis and BlindSignal Separation, vol. 3889, pp.32-39.

[27] Lee, S., Han, D. & Ko, H. (2016). Single-channel speech enhancement method using reconstructive NMF with spectrotemporal speech presence probabilities, Applied Acoustics, In Press.

[28] Mirzaei, S.,Van hamme, H. & Norouzi, Y. (2015). Blind audio source counting and separation of anechoic mixtures using the multichannel complex NMFframework, Signal Processing, vol. 115, pp. 27-37.

[29] Berry, M. W., Browne, M., Langville, A. N., Pauca, V. & Plemmons, R., (2006). Algorithms and Applications for Approximate Nonnegative Matrix Factorization. Elsevier. Computational statistics and data analysis, vol. 52, no. 1, pp. 155-173.

[30] Li, Z., Peng, H. & Wu, X. (2008). A new Descriptive clustering algorithm based on nonnegative matrix factorization. IEEE Int. Conf., Granular Computing, pp. 407-412.

[31] Li, S. Z., Hou, X. W. & Zhang, H. J. (2001). Learning spatially localized, parts-based representation, IEEE Int. Conf., Computer Vision and Pattern Recognition, pp. 207-212.

[32] Wooldridge, M. (2002). An Introduction to Multiagent Systems. Wiley, New York.

[33] Deng, X. & Papadimitriou, CH. (1994) on the complexity of cooperative solution concepts. Mathematics of Operations Research, vol. 19, no. 2, pp. 257-66.

[34] Fernandez, J. R., Algaba, E., Bilbao, J. M., Jimenez, A., Jimenez, N. & Lopez, J. J. (2002). Generating functions for computing the Myerson value. Annals of Operations Research, vol. 58, pp.109-143.

[35] Castroa, J., Gomeza, D. & Tejadab, J. (2009). Polynomial calculation of the Shapley value based on sampling. Computers & Operations Research, vol. 36, no. 5, pp.1726-1730.

[36] Lyons, M. J., Kamachi, M., Gyoba, J. & Akamatsu, S. (1998). Coding facial expressions with Gabor waveletsProc.IEEE Aut. Face and Gesture Recog. pp. 200-205.

[37] Aifanti, N., Papachristou , C., & Delopoulos, A. (2010). The MUG Facial Expression Database. 11th Int. Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Desenzano, Italy.