Bulletin of Electrical Engineering and Informatics
Vol. 7, No. 4, December 2018, pp. 580~586
ISSN: 2302-9285, DOI: 10.11591/eei.v7i4.1230

Local Entropy and Standard Deviation for Facial Expressions Recognition in Thermal Imaging

Yomna M. Elbarawy[1], Rania Salah El-Sayed[2], Neveen I. Ghali[3]
[1,2]Al-Azhar University, Faculty of Sience, Cairo, Egypt
[3]Future University in Egypt, Faculty of Computers & Information Technology, Cairo, Egypt

Article Info

Article history:

Received May 25, 2018
Revised Aug 26, 2018
Accepted Oct 08, 2018

Keywords:

Facial expressions
Local Entropy
Local Standard Deviation
Support Vector Machine
Thermal Images

ABSTRACT

Emotional reactions are the best way to express human attitude and thermal imaging mainly used to utilize detection of temperature variations as in detecting spatial and temporal variation in the water status of grapevine. By merging the two facts this paper presents the Discrete Cosine Transform (DCT) with Local Entropy (LE) and Local Standard Deviation (LSD) features as an efficient filters for investigating human emotional state in thermal images. Two well known classifiers, K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) were combined with the earlier features and applied over a database with variant illumination, as well as occlusion by glasses and poses to generate a recognition model of facial expressions in thermal images. KNN based on DCT and LE gives the best accuracy compared with other classifier and features results.

Corresponding Author:

Yomna M. Elbarawy,
Al-Azhar University,
Faculty of Sience, Cairo, Egypt
Email: y.elbarawy@azhar.edu.eg

1. INTRODUCTION

Although recognition using thermal images has overcome many challenges of the recognition if compared to visible images as illumination [1], [2]; still thermal images faces its own challenges as temperature [3], aging problem and illumination [4], as well as occlusion by glasses and poses which will be tackled in this research. In 2005 L. Trujillo et al. used Local and Global feature extraction methods using interest point detected by Harris detector clustered by K-means with SVM as classifier over IRIS database achieving 76.6% accuracy [5].

Shangfei Wang et al. in 2012 introduced temperature difference features and voting strategy with KNN as classifier applied over USTC-NVIE database making 61.62% recognition rate [6]. Deep Boltzmann machine DBM model was used by Shangfei Wang in 2014 for emotional recognition with accuracy rate 62.9% over the USTC-NVIE database [7]. 98.2% recognition rate was achieved by M.H. Abd Latif et al. [8] through the use of Gray Level Cooccurrence Matrix (GLCM) as a feature extractor and KNN as a classifier over a new database gathered by the paper team at the International Islamic University in Malaysia.

This paper introduces the use of LE and DCT filters as a feature extractors and KNN as classifier to approach a solution for expression recognition in thermal images. The used dataset is Imaging, Robotics and Intelligent Systems (IRIS) database. Pose variation challenge appears with this database since every object (person) has 11 poses for each expression. Although this variation poses problem the proposed model with the LE still gives accuracy higher than the same model with other different features like Principle Component Analysis PCA and Local Standard Deviation as explained ahead.

The remainder of this paper is ordered as follows. Section 2 gives a brief introduction to feature extraction methods, local standard deviation, local entropy and principle component analysis and discrete cosine transform technique. Classification methods Support vector machine and K-nearest neighbor are also discussed in section 2. Detailed proposed model is given in section 3. Section 4 shows the experimental results and analysis. Conclusions are discussed in section 5.

2. PRELIMINARIES

This section introduces a literature survey on different preliminaries given in this work.

2.1. Feature Extraction Methods

2.1.1. Local Standard Deviation

Statistically SD is the square root of variance which a way to determine unevenness of objects [9]. Here LSD is applied on thermal images to indicate the degree of variability of the intensity values of pixels in an image. LSD acts as a feature extraction method as shown in Figure 1. SD calculation illustrated in algorithm 1 and equation 1. Where n is number of pixels, x represents one pixel in an image and $\bar{x}$ is the mean of the image.

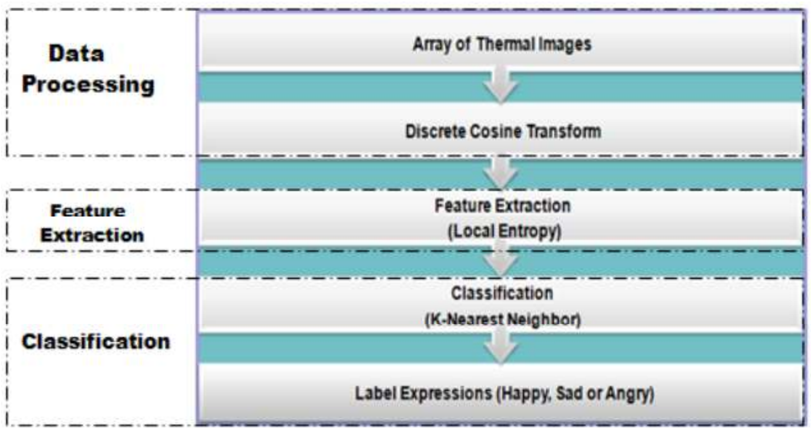

Figure 1. Proposed framework for facial expression recognition

Algorithm 1 : Standard Deviation Algorithm
1-Find the mean of the image.
2- For each pixel, find the square of its distance to the mean.
3- Sum the values from Step 2.
4- Divide step 3 by the number of pixels.
5- Take the square root for step 4

$$SD = \sqrt{\frac{\sum_{i=1}^{n}|x-\bar{x}|^2}{n}} \tag{1}$$

2.1.2. Local Entropy

Local Standard Deviation filters image by replacing every value by the information entropy of the values in its range r neighborhood. The entropy represents the information associated with a single pixel of the image by calculating the probability distribution function of the image [10]. Firstly proposed by Claude Shannon in 1948 and used widely ever since. Local SD represented by equations 2 and 3. Where $E(I)$ is the Shannon entropy of a random pixel I, p_j defined in equation 3 with i_j indicating the jth possible value of I out of n pixels and $p_j$denoting the possibility of $I = i_j$.

$$\mathrm{E(I)} = -\sum_{j=1}^{n} \mathrm{p_j} \log_2 \mathrm{p_j} \tag{2}$$

$$p_j = P_r(I, i_j) \tag{3}$$

2.1.3. Principle Component Analysis

Principle Component Analysis (PCA) technique used widely in recognition field as in [11] and [12]. Its basic steps illustrated in algorithm 2.

Algorithm 2 : PCA Algorithm

1- Input data normalization.
2- Covariance matrix calculation.
3- Finding the eigenvectors of the covariance matrix.
4- Data interpretation into terms of components and compose a feature vector.

2.1.4. Discrete Cosine Transform

First introduced in 1974 and evolved over time [13]. DCT used in various image compression and recognition schemes [14], [15]. DCT matrix U is invertible and orthogonal so that U^T=U^{-1}. In Equation 4 $C(u,v)$ computes the u,v^{th} entry of the DCT of the image. $f(x,y)$ is the x,y^{th} element of the image represented by matrix f. N is the size of the DCT block.

u,v = 0,1,2,…,N - 1 and $\alpha(u),\alpha(v)$ calculated as in Equation 5.

$$C(u,v) = \propto(u) \propto(v) \sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y) . \cos\left[\frac{\pi u(2x+1)}{2N}\right] . \cos\left[\frac{\pi v(2y+1)}{2N}\right] \quad (4)$$

$$\propto(u) = \sqrt{\frac{1}{N}} \ \rightarrow u = 0 \ and \ \propto(u) = \sqrt{\frac{2}{N}} \ \rightarrow u \neq 0 \quad (5)$$

2.2. Classifiers

2.2.1. K-Nearest Neighbor

KNN is a non parametric lazy learning algorithm [16]. Most of the training data is needed during the testing phase and usually makes decision based on the entire training data set in contrast to other techniques like SVM where it can discard all non support vectors. In the general model KNN used as a classifier of testing images in classification frame Figure 1. The major steps of KNN illustrated in algorithm 3.

Algorithm 3 : KNN Classification Algorithm

1- Input dataset as:
- Training (labeled data).
- Test (unlabeled data).
2- Calculate distances of all training vectors to test vector according to Euclidean metric.
3- Pick k closest vectors and predict class by majority vote.
4- Calculate average/majority by inverse distance.

2.2.2. Support Vector Machine

SVM is a supervised machine learning algorithm [17] earlier used for two-class grouping problems. Introduced firstly by Vapnik in 1992 by proposing a non-linear classifier and using kernel trick in 1995 [18].

SVM used a kernel function to classify a set of data into a two class groups. SVM can classify data into multiple classes [19] by training it for each possible pair of classes and classify an unknown point p by applying each of the classifiers and count how many times point p was assigned to a certain class label. Finally, the unknown point assigned to the class label with highest count.

3. PROPOSED METHOD

The proposed framework is shown in Figure 1 which involves three main frames: data processing, feature extraction and classification. The first frame, data processing which covers image acquisition, pre processing. Second frame (feature extraction) applied here using two techniques: Local Entropy and image transformation using Discrete Cosine Transform. Finally, the last frame (data classification) used K-nearest neighbor as classifier.

Image acquisition: The used data was selected from IRIS database, since it has a different poses for each subject. Only poses less than 45° rotation were used here. Further information about the database is illustrated later.

Image transformation was done by using discrete cosine transform filter as in Figure 2. Then, a texture based feature extraction method (Local Entropy) applied over the processed images.

Two *classification* methods were applied in order to find a suitable class label for each test image. First, multiclassification using SVM by training two-class SVM for each pair classes (Surprise and Happy, Happy and Angry and Angry and Surprise) then by counting how many times each image assigned to a certain label class. Classified image belongs to the label class (Surprise, Happy or Angry) with the highest count. Second, KNN applied by calculating distance between each test image and training images using Euclidean distance and predict image class label by the majority voting of the closest training images.

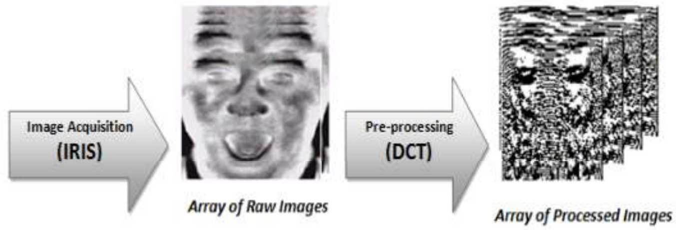

Figure 2. Image acquisition and pre-processing

4. EXPERIMENTAL RESULTS

4.1. Data Set

IRIS dataset in the OCTBVS database which contains images in bitmap RGB format. The database contains approximately 3500 thermal and visible images with size 320 x 240, collected by the long wave IR Camera (Thermal-Raythoen Palm IR-Pro) at the University of Tennessee having uneven illuminations and different poses.

This work used 60 (30 for training and 30 for testing) images since the database has variant illumination, as well as occlusion by glasses and poses. Only poses less than 45° rotation were selected. Each subject has three different expressions Surprise, Happy and Angry. Table 1 shows the IRIS database details [20] and Figure 3 has samples of the used data thermal images.

Table 1. IRIS Database Information

No. of Population	Status		Resolution	No. of Images	
30	Expressions	3	320*240	Thermal	~1529
	Illumination	6			
	Poses	11		Visible	~1529

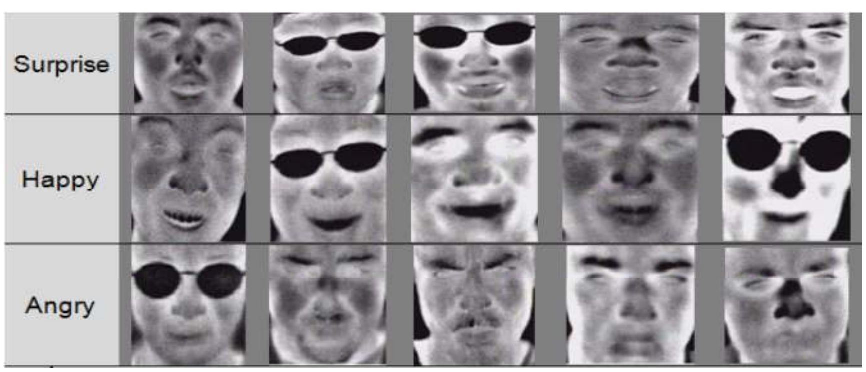

Figure 3. Samples of used data images

4.2. Results and Discussion

The experimental results have two main approaches. The first, applied KNN and SVM with multiple features directly over the selected data. Second approach, applied KNN and SVM with multiple features based on DCT over the tested thermal images.

Experimental results of the first approach indicate that using local standard deviation as an extractor under KNN or SVM has the highest recognition rate 63.33% and 73.33% respectively. Tables 2 and 3 show the detailed confusion matrix for both previous cases and the values of True Positive (TP) and False Negative (FN) rates for each expression. Table 4 has the recognition rates for multiple features (LE, LSD and PCA) which are used directly with the KNN and SVM classifiers.

Table 2. Confution Matrix of KNN Based LSD

	Surprise	Happy	Angry	TPR	FNR
Surprise	6	3	0	60%	40%
Happy	4	7	4	70%	30%
Angry	0	0	6	60%	40%

Table 3. Confution Matrix of SVM Based LSD

	Surprise	Happy	Angry	TPR	FNR
Surprise	6	3	0	60%	40%
Happy	4	7	4	70%	30%
Angry	0	0	6	60%	40%

Table 4. Classifiers and Features Performance Comparison without DCT

Classifier	Feature	Accuracy (%)
	PCA	60
KNN	LSD	63.33
	LE	33.33
	PCA	43.33
SVM	LSD	73.33
	LE	33.33

The second approach of the results that uses KNN and SVM with multiple features based on DCT shows that: local standard deviation has higher recognition rate 83.33% than other features (LE and PCA) under the SVM classifier. While LE has the highest recognition rate with 90% under the KNN classifier based on the DCT filter. Table 5 shows the confusion matrix of KNN based on LE and DCT. Table 6 shows the confusion matrix of SVM based on LSD and DCT with the TP and FN rates. Tables 7 and 8 results illustrate the detailed accuracy by class (Surprise, Happy and Angry) of KNN based on DCT and the detailed accuracy by class of SVM Based on DCT. Table 9 has the recognition rates for multiple features (LE, LSD and PCA) that used under the KNN and SVM classifiers based on the DCT technique which show that KNN has recognition rate 90% with LE feature and 80% with LSD feature while, SVM 73.33% recognition rate with the LE and 83.33% with LSD feature.

Table 5. Confution Matrix of KNN Based LE+DCT

	Surprise	Happy	Angry	TPR	FNR
Surprise	9	1	0	90%	10%
Happy	0	8	0	80%	20%
Angry	1	1	10	100%	0%

Table 6. Confution Matrix of SVM Based LSD+DCT

	Surprise	Happy	Angry	TPR	FNR
Surprise	9	0	0	90%	10%
Happy	0	6	0	60%	40%
Angry	1	4	10	100%	0%

Table 7. Detailed Accuracy by Class of KNN based DCT

Classifier	Class	Feature+DCT	Accuracy (%)
		LSD	100
	Surprise	LE	90
		PCA	70
		LSD	40
KNN	Happy	LE	80
		PCA	50
		LSD	100
	Angry	LE	100
		PCA	60

Table 8. Detailed Accuracy by Class of SVM based DCT

Classifier	Class	Feature+DCT	Accuracy (%)
		LSD	90
	Surprise	LE	70
		PCA	90
		LSD	60
KNN	Happy	LE	70
		PCA	10
		LSD	100
	Angry	LE	80
		PCA	10

Table 9. Classifiers and Features Performance Comparison Based DCT

Classifier	Feature	Accuracy (%)
	PCA	60
KNN	LSD	80
	LE	90
	PCA	36.33
SVM	LSD	83.33
	LE	73.33

Overall results implicate that the highest expression recognition rate was made by the Local Entropy feature under the K-Nearest Neighbor classifier based on DCT filter with 90%. Figure 4 shows visualization of the classification results made using NodeXL tool. This work uses mostly the learner classifiers and features of MATLAB R2014a.

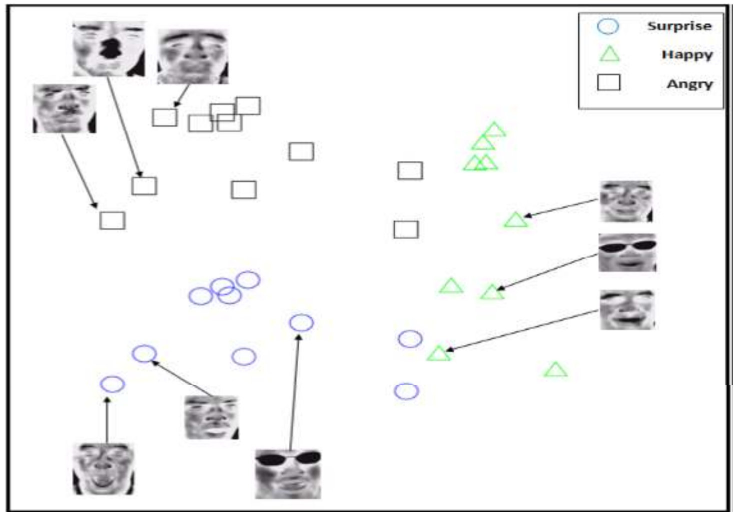

Figure 4. Classification result visualization of using KNN+LE based DCT

5. CONCLUSION

This paper holds two main approaches of conducted results, one based on applying DCT filter over the selected dataset and other without using it in order to present more efficient model for expression recognition in thermal images. The first approach uses K-Nearest Neighbor and Support Victor Machine classifiers with multiple features extraction (LE, LSD and PCA) without applying DCT and experimental results show that the Local Standard Deviation gives high accuracies with both KNN and SVM but higher with SVM with 73.33% recognition rate.

The second approach applies KNN and SVM with multiple features extraction based on applying DCT filter. Experimental results show that generally applying the DCT improves the recognition rates over most used features extraction (LE and LSD) as shown in Table 9. But the most efficient model with the highest accuracy 90% uses the Local Entropy as feature extractor and KNN as classifier based on the DCT as appears in Figure 5. Experimental results on IRIS database demonstrate that the proposed model gives feasible recognition rates although the occlusion by glasses and pose variation of the expressions.

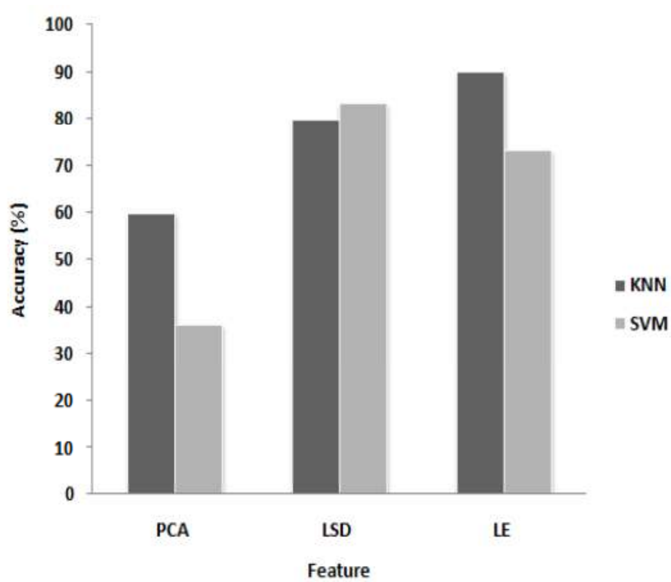

Figure 5. Accuracy rate for feature extractors using classification methods based DCT

REFERENCES

[1] F. Z. Salmam, A. Madani and M. Kissi, "Emotion Recognition From Facial Expression Based on Fiducial Points Detection and Using Neural Network", International Journal of Electrical and Computer Engineering (IJECE), 2018, Vol. 8(1), pp. 52-59.
[2] Y. Wang, X. Yang and J. Zou, "Research of Emotion Recognition Based on Speech and Facial Expression", TELKOMNIKA (Telecommunication, Computing, Electronics and Control), 2013, Vol. 11(1), pp. 83-90.
[3] C. M. Naveen Kumar and G. Shivakumar,"A Survey on Human Emotion Analysis Using Thermal Imaging and Physiological Variables", International Journal Of Current Engineering And Scientific Research (IJCESR), 2017, Vol. 4(4), pp. 122-126.
[4] O. Arandjelovic et al., "Thermal and Reflectance Based Personal Identification Methodology Under Variable Illumination", Pattern Recognition, 2010, Vol. 43, pp. 1801-1813.
[5] L. Trujillo, G. Olague, R. Hammoud, B. Hernandez, "Automatic Feature Localization in Thermal Images for Facial Expression Recognition", Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2005: 14.
[6] S. Wang. P. Shen, and Z. Liu, "Facial Expression Recognition From Infrared Thermal Images Using Temperature Difference By Voting", Proceeding of IEEE 2nd International Conference on Cloud Computing and Intelligent Systems (CCIS), Hangzhou, 2012, Vol. 1, pp. 94-98.
[7] S. Wang, Menghua He, Z. Gao, Shan HE and Qiang JI, "Emotion Recognition from Thermal Infrared Images Using Deep Boltzmann Machine", Front. Comput. Sci., 2014, Vol. 8(4), pp. 609-618.
[8] M.H. Abd Latif et al., "Implementation of GLCM Features In Thermal Imaging for Human Affective State Detection", Procedia Computer Science, 2015, Vol. 76, pp. 308-315.
[9] S. W. Smith,"The Scientist and Engineer's Guide to Digital Signal Processing", 3rd ed. California Technical Publishing, 1997.
[10] Y. Wu et al.,"Local Shannon Entropy Measure with Statistical Tests for Image Randomness", Information Sciences, 2013, Vol.(222), pp. 323-342.
[11] P. Kumar and P. Sehgal,"Face Detection Using Principle Component Analysis", International Journal of Computer Engineering & Technology (IJCET), 2016, Vol. 7(3), pp. 174-178.
[12] S. Kaur and R. SinghVirk, "DCT Based Fast Face Recognition Using PCA and ANN", International Journal of Advanced Research in Computer and Communication Engineering, 2013, Vol. 2(5), pp. 2125-2131.
[13] M. Deboraj et al., "A Simplified Design Approach for Efficient Computation of DCT", Fifth IEEE Computer Society conference on Communicatio System and Network Technologies, Gwalior, 2015, Vol. 134, pp. 483-487.
[14] S. K. Tyagi and P. Khanna,"Face Recognition Using Discrete Cosine Transform and Nearest Neighbor Discriminant Analysis", International Journal of Engineering and Technology (IACSIT), 2012, Vol. 4(3), pp. 311-314.
[15] G. Strang,"The Discrete Cosine Transform", Journal of Society for Industrial and Applied Mathematics (SIAM), 1999, Vol. 41(1), pp. 135-147.
[16] E.K. Garcia, S. Feldman, M.R. Gupta, S. Srivastava, "Completely Lazy Learning", IEEE Trans. Knowledge and Data Eng., 2010, Vol. 22(9), pp. 1274-1285.
[17] J. Xu et al., "Multi-Class Support Vector Machine via Maximizing Multi-Class Margins", Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI), Melbourne, 2017: 3154-3160.
[18] C. Cortes and V. Vapnik,"Support-vector networks", Machine learning, 1995, Vol. 20(3), pp. 273-297.
[19] S. Cheong, S. H. Oh and S.-Y. Lee,"Support Vector Machines With Binary Tree Architecture for Multi-Class Classification", Neural Information Processing-Letters and Reviews, 2004, Vol. 2(3), pp. 47-51.
[20] University of Tennessee: IRIS thermal/visible face database: http://vciplokstate.org/pbvs/bench/, last accessed Jul. 2017.