Towards Thermal Region of Interest for Human Emotion Estimation

Thu Nguyen
Ho Chi Minh City University of Education
Ho Chi Minh City, Vietnam
thunta@hungvuong.sgdbinhduong.edu.vn

Khang Tran
Ho Chi Minh City University of Education
Ho Chi Minh City, Vietnam
khangphysix1997@gmail.com

Hung Nguyen
Ho Chi Minh City University of Education
Ho Chi Minh City, Vietnam
hungnv@hcmue.edu.vn

Abstract – In the recent years, researches in implementing visual-based Automatic Facial Expression Analysis (FEA) and human emotions estimation have been a great interest and had widespread applications. However, vision-based visual systems face challenges such as poor-quality images due to low light conditions, poker-faces, the contrast of expressions and emotions. Because of the advantages of thermal images - unsensitive to ambient light conditions, we focus on developing a human emotion estimating system using thermal-based cues. It overcomes the limits of visible images, however, once a change in emotion occurs, only some areas of the face is affected in terms of temperature. In addition, the loss of thermal information in the glasses areas is a matter of concern. This paper presents our proposed method, thermal region of interest, to fill the gap when using thermal images to estimate seven emotions.

Keywords—Human Emotion Estimation, Affective Computing, Thermal Image, Thermal Region of Interest, PCA, EMC

I. INTRODUCTION

The fourth industrial revolution (also known as the Industry 4.0), which has been mentioned in the past few years, especially in the recent time, has had a strong impact on many fields: the appearance of robots and artificial intelligence bring many applications in society [1-4]. The interaction of humans and robot (machine) has to be improved for humans to receive better services and application. To make human-machine interaction more efficient and accurate, the emotion - recognition system becomes a field that has been attracting the attention and research of scientists. Some biological signals such as ECG [5], blood volume [6], EEG [7], etc, have been studied to predict human's emotions. In Mehrabian's study [8], it was found that in communication, speech only contributes 7%, intonation contributes 38%, while facial expressions contribute 55% to the efficiency of communication. This shows that facial expressions form the primary mode of communication, behaviors, and emotions. In addition, this is also an important factor in Human-Computer Interaction (HCI) systems [9], [10]. Recently, some studies and applications have achieved more than 80% accuracy in facial expressions analysis for emotional identification using visible information [11]. However, vision-based visual systems face challenges such as poor-quality images due to low light conditions, poker faces and the contrast of expressions and emotions.

Due to advancements in infrared (IR) images, we can detect and measure physiological signals, body temperature, through non-invasive and non-contact equipment [12]. When an emotion occurs, the facial temperature rises because the blood flows through the blood vessels in the subcutaneous region [13], this change may be eligible and quantified to be able to research emotions through Infrared Thermal images (IRT). This is one of the solutions that can overcome the limitations of the visible image. M.M.Khan et al. [14] suggested using Facial Thermal Feature Points (FTFP) to accurately measure from 66.3% to 83.8% with five types of emotions. L.Trujillo et al. [15] proposed using facial localization in thermal images to classify three emotions. Ioannou et al. [16] studied a range of emotions, such as fear [17] and stress [18]. Finally, Liu and Wang [19] analyzed facial temperature data sequence, statistical feature calculations, and differential temperature histogram feature calculations. In addition, the Hidden Markov model (HMM) is used to distinguish happiness, disgust and fear with the respective recognition rates of 68.11%, 57.14% and 52.30%. The results also showed that the temperature of the forehead was more useful than the other areas on the face. They used samples from the USTC-NVIE database (facial and natural facial expression) to evaluate the results [20].

All researches show that thermal images are insensitive to light conditions, unaffected by ambient conditions, that can be used to detect facial temperature change [21]. However, there're some disadvantages when using thermal images, as some areas on the face do not respond to emotional changes or infrared ray in the area around the eyes is observed by the eyeglasses. In order to overcome these constraints, we propose to use the regions of Interest (ROIs) for thermal IR data to increase the accuracy of the real emotion. We used Principal Component Analysis (PCA) method, Eigen-space Method based on class-features (EMC), and PCA-EMC on Kotani Thermal Facial Emotions (KTFE) database to estimate seven basic types of emotions (anger, disgust, fear, happiness, neutral, sadness, surprise).

II. METHODS

In this section, we propose a method of extracting regions of interest (ROIs) on thermal IR data. To see how effective it is, we conduct emotional recognition with three canonical approaches in the field of computer vision, such as PCA, EMC and PCA-EMC. Experiment results on benchmark databases demonstrate the effectiveness and superior performance of our proposed method

A. Finding of ROIs

Determining the ROIs is very helpful and necessary. It identifies areas where the temperature rises or decreases dramatically when emotions change and are mostly concentrated in the forehead, eyeholes, cheekbones and maxillary . This is an important factor to make the identification of emotions more accurate. More specifically, it helps us better understand the nature of the relationship between temperature and emotion.

978-1-5386-6113-0/18/$31.00 ©2018 IEEE

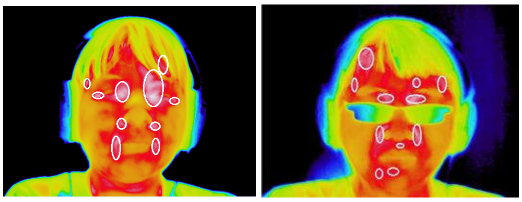

Figure 1. Example of ROIs

From the initial thermal matrix, after preprocessing, ROIs are determined as follows:

Let h, g, f be the mapping functions defined below:

$$\begin{matrix} h: F \to Tc \\ (i,j) \mapsto h(i,j) \end{matrix};\quad \begin{matrix} g: F \to Gr \\ (i,j) \mapsto g(i,j) \end{matrix};\quad \begin{matrix} f: Gr \to Ic \\ (i,j) \mapsto f(i,j) \end{matrix} \qquad (1)$$

Where F is the face space; Tc is the temperature space; Gr is the face space in the grayscale; Ic is the intensity of the pixel in the grayscale.

ROIs are found through the following equations:

$$\Delta T_F = T^F_{Max} - T^F_{Min} \qquad (2)$$

$$\delta T_F = \Delta T_F / T^{Ic}_{mgr} \qquad (3)$$

$$T^{Ic}_{mgr} = \max\left(f\left(i,j\right)\right) \qquad (4)$$

$$L^F_k = \left\{(i,j) \in F \,\middle|\, T^F_{Min} + \delta T_F * g(i,j) \le h(i,j) \le T^F_{Max} - \delta T_F * g(i,j)\right\} \quad (5)$$

$$k \in \overline{1,K}$$

Where T^F_{Max}, T^F_{Min} is the largest and smallest temperature of each surface in frame k; T^{Ic}_{mgr} is the highest intensity in the grayscale, K is the frame number

After finding ROIs, we use to PCA, EMC and PCA-EMC to detect seven base emotions.

B. *Estimate emotion*

Firstly, we use PCA [22], the goal is to build a facial space, which includes basic vectors called the main components, which describe the facial images better. To detect emotions using PCA, we divided the training set into seven classes and calculated the individual space of each class as follows:

Step 1: Each thermal data frame is determined from a two-dimensional matrix to a one-dimensional vector. With the M_f data frame as the training data, we convert these data into corresponding column vectors.

Step 2: Find the covariance matrix representing the average over-dispersion of training data.

$$S = \frac{1}{M_f} \sum_{k=1}^{M_f} (x_k - \bar{x})(x_k - \bar{x})^T \qquad (6)$$

$$\bar{x} = \frac{1}{M_f} \sum_{k=1}^{M_f} x_k \qquad (7)$$

Where x_k is an N-dimensional vector, M_f is the total number of frames per class and S is the covariance matrix.

Step 3: Calculate the eigenvectors and eigenvalues of the covariant matrix.

Step 4: Choose the largest H eigenvalues and eigenvectors.

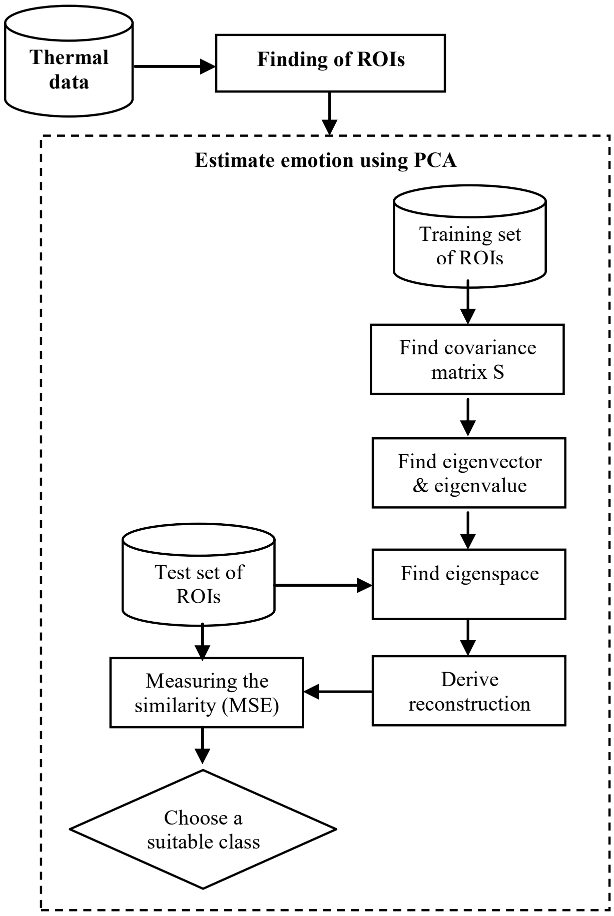

Figure 2. ROIs and PCA for human emotion estimation

For each test data, we project it into the individual space of each class and retrieve the renewable thermal data from each of the individual spaces. Using the variance, measuring the similarity between the input thermal data and the reconstructed thermal data, we can select an appropriate class for the input thermal data as the minimum of the variance.

Secondly, we use EMC [23], we divided the training into seven classes and created a separate space. For each IR thermal test data, we project it into the individual space of each layer and an emotion is selected if it has the largest cosine between the obtained vector after projection and the individual space of each layer.

$$S = S_B - S_W \qquad (8)$$

$$S_W = \frac{1}{M} \sum_{f\in F} \sum_{f\in F}^{M_f} M_f \left(\bar{x}_{fm} - \bar{x}_f\right)\left(\bar{x}_{fm} - \bar{x}_f\right)^T \qquad (9)$$

$$S_B = \frac{1}{M} \sum_{f\in F} M_f \left(\bar{x}_f - \bar{x}\right)\left(\bar{x}_f - \bar{x}\right)^T \qquad (10)$$

$$\bar{x}_f = \frac{1}{M} \sum_{m=1}^{M_f} x_{fm};\quad \bar{x} = \frac{1}{M} \sum_{f\in F} \sum_{m=1}^{M_f} x_{fm} \qquad (11)$$

where F is a set of expression classes, M_f facial-patterns are given for each class $f \in F$ and x_{fm} is an N-dimension vector of the m-th facial pattern, $m = \overline{1, M_f}$

Finally, with PCA-EMC, we use PCA to reduce the size and then apply EMC for testing.

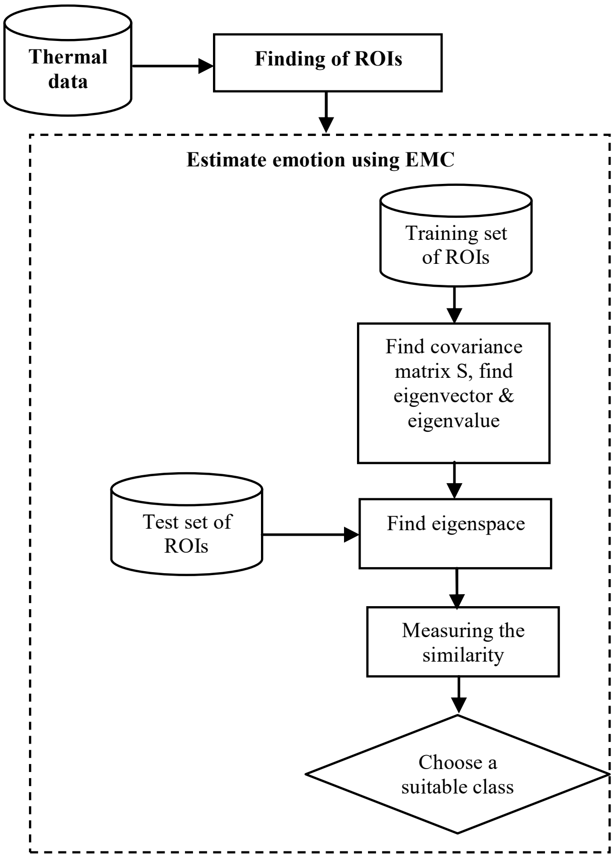

Figure 3. ROIs and EMC for human emotion estimation

III. EXPERIMENTS

For databases, we use the Kotani Thermal Facial Emotions database [24]. This database has 30 subjects: Vietnamese, Japanese, Thai from age 11 to 32 with seven emotions.

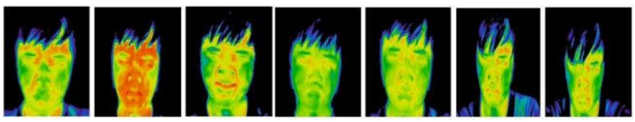

Figure 4. Examples of thermal images of seven expressions (neutral, anger, happiness, sadness, fear, disgust, surprise) [24]

In our experiments, we used only thermal data to estimate human emotions. From 130 GB of thermal data, we extract 27.2 GB of thermal data for seven emotions. We use 70% of the data for training and testing is 30%.

Table 1 presents the PCA facial estimation results for the thermal data using ROI. The results show that the accuracy rate is very high at 90.42%, the error rate is only 9.59%. In particular, anger reached 92.5%, happiness reached 97.44% and neutral reached 98.64%.

Table 2 presents the EMC facial estimation results for the thermal data using ROI. The results show that angry rate reached 93.75%, higher PCA is 1.25%, disgust reached 91.11%, while the PCA only reached 83.25%, increase 7.86%, sadness reached 90.32%, up 4.22% and surprise reached 94.12%, up 5.9% compared to PCA. So, the accuracy rate is 89.99%, the error rate is only 10.01%.

Table 3 presents the PCA-EMC facial estimation results for the thermal data using ROI. Results showed that the accuracy rate is 85.94%, the error rate is only 14.06%. Compared with the PCA, although the average was lower, the fear rating increased by 0.28% and sadness by 2.17%. Compared with the EMC, the happiness rate increased by 0.97% and neutral by 0.39%

Thus, from the results obtained in Tables 1, 2 and 3, there is no surprise in any case of anger, no disgust is recognized as happiness and sadness, no sadness is recognized as a surprise. Based on the results, we believe that thermal data is important additional information to support more accurate emotional identification.

	An	Di	Fe	Ha	Ne	Sa	Su
An	92.5 %	0%	2.81 %	0%	0%	1.2 %	0%
Di	0%	83.25 %	0%	0%	0%	0%	3%
Fe	4.17 %	8.89 %	86.78 %	1.36 %	0.91 %	7.73 %	4.18 %
Ha	0%	2.08 %	1.94 %	97.44 %	0.45 %	0.88 %	2.73 %
Ne	1.67 %	0%	1.91 %	0%	98.64 %	1%	1.88 %
Sa	1.67 %	0%	6.56 %	1.2 %	0%	86.1 %	0%
Su	0%	5.77 %	0%	0%	0%	3.08 %	88.22 %
Avg							**90.42 %**

Table 1. Confusion matrix with PCA. From left to right (or top to bottom): anger (An), disgust (Di), fear (Fe), happiness (Ha),neutral (Ne), sadness (Sa) and surprise (Su).

	An	Di	Fe	Ha	Ne	Sa	Su
An	93.75 %	0%	0%	0%	3.23 %	4.3 %	5.88 %
Di	6.25 %	91.11 %	6.45 %	0%	0%	2.15 %	0%
Fe	0%	0%	87.1 %	5.52 %	3.45 %	0%	0%
Ha	0%	0%	1.08 %	86.87 %	3.45 %	0%	0%
Ne	0%	0%	0%	0%	86.65 %	3.23 %	0%
Sa	0%	2.22 %	5.38 %	7.6 %	3.23 %	90.32 %	0%
Su	0%	6.67 %	0%	0%	0%	0%	94.12 %
Avg							**89.99 %**

Table 2. Confusion matrix with EMC. From left to right (or top to bottom): anger (An), disgust (Di), fear (Fe), happiness (Ha),neutral (Ne), sadness (Sa) and surprise (Su).

	An	Di	Fe	Ha	Ne	Sa	Su
An	87.16 %	2.14 %	1.81 %	1.43 %	2.27 %	0%	1.76 %
Di	5.24 %	78.59 %	0%	0%	4.23 %	0%	3.67 %
Fe	2.86 %	5.18 %	87.06 %	5.44 %	1.42 %	5.15%	3.43 %
Ha	4.74 %	1.22 %	3.86 %	87.64 %	1.8 %	2.86%	1.88 %
Ne	0%	4.17 %	0.71 %	0.37 %	87.04 %	2.75%	3.43 %
Sa	0%	0.67 %	5.49 %	2.97 %	2.85 %	88.27 %	0%
Su	0%	8.03 %	1.07 %	2.15 %	0.38 %	0.97%	85.83 %
Avg							**85.94 %**

Table 3. Confusion matrix with PCA-EMC. From left to right (or top to bottom): anger (An), disgust (Di), fear (Fe), happiness (Ha),neutral (Ne), sadness (Sa) and surprise (Su).

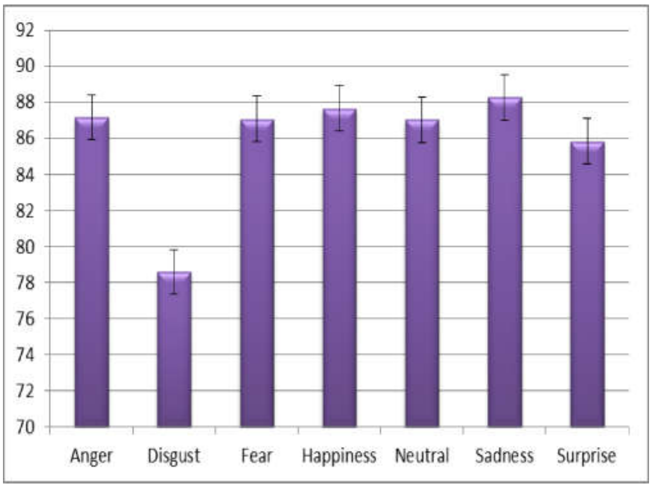

Figure 7. The emotion estimation results of PCA-EMC

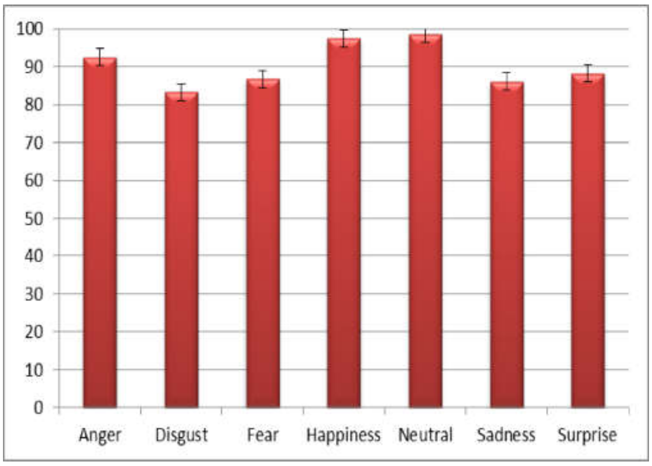

Figure 5. The emotion estimation results of PCA

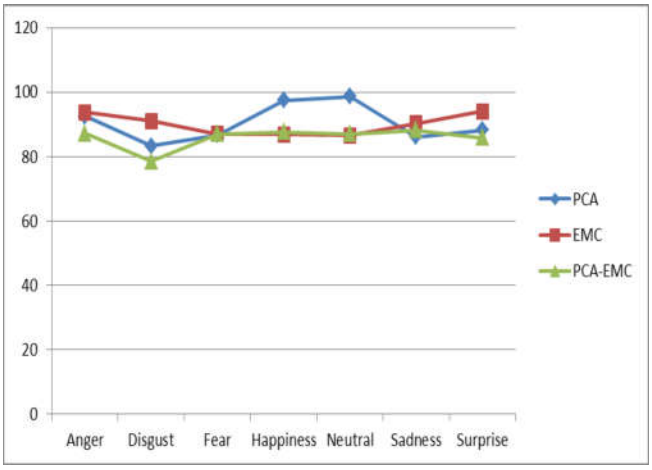

Figure 8. Comparison between PCA, EMC and PCA-EMC

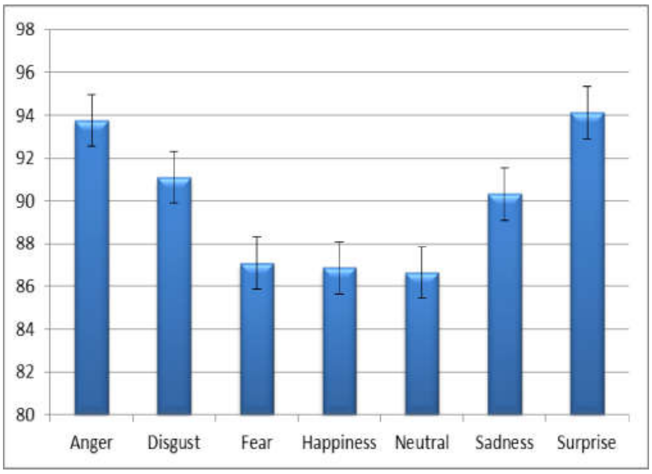

Figure 6. The emotion estimation results of EMC

	Average Recognition Rate	
	Method [25]	**Our Method**
PCA	85.43%	90.42%
EMC	86.00%	89.99%
PCA-EMC	-	85.94%

Table 4. Comparative analysis

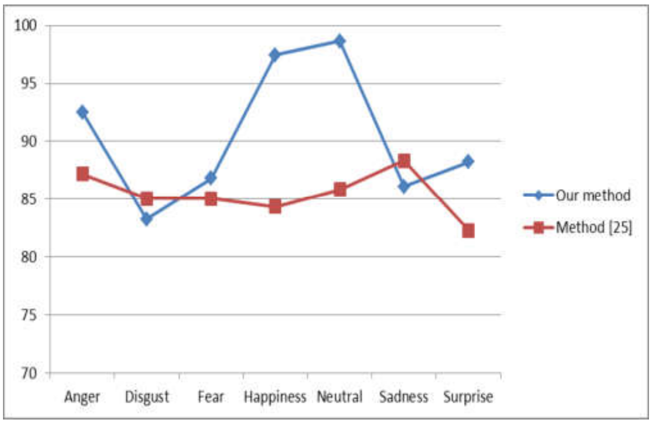

Figure 9. Comparison between our method and method [25] for PCA

The proposed method was compared with the results obtained by performing PCA and EMC methods, implemented on the same KTFE database [25]. With method [25], the authors proposed to extract the ROIs are quite simple. The ROIs found does not provide really useful information when emotions change and had a problem with heating obstacles. Table 4 shows that the performance achieved by the proposed method is better than the other methods. Because we improve ROIs, they contain more useful information, showing a better relationship between temperature and emotions. With PCA, the accuracy rate increased by 4.99%, from 85.43% to 90.42%. While EMC achieved 89.99%, up 3.99%. In addition, we also tested with PCA-EMC. Results reached 85.94%

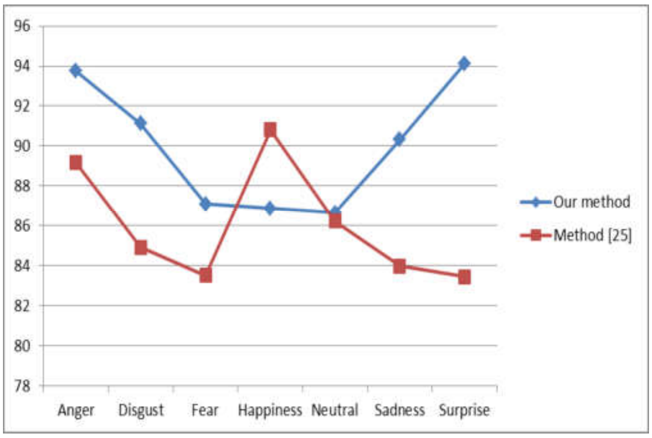

Figure 10. Comparison between our method and method [25] for EMC

IV. CONCLUSIONS

In this paper, we introduce a method of estimating human emotions based on thermal images. Here, describe a way to overcome the influence of environmental light and reduce the impact of eyeglasses by improving the extraction process of the region of interest in the thermal data. The algorithm is tested on the KTFE database. The seven basic emotions, anger, disgust, fear, happiness, sadness, surprise and neutral are estimated with high accuracy of 90.42%, 89.99% and 85.94% respectively for PCA, EMC and PCA-EMC. Results confirm that temperatures and emotions are closely related. Using temperature information will increase the accuracy of the human emotions estimation. In the future, we will continue improving the ROIs on thermal IR data to contribute better results.

REFERENCES

[1] J.A. Russell: “Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies,” Psychological Bulletin, vol. 115, no. 1,pp. 102-141, Jan. 1994.

[2] P. Ekman: “Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique,” Psychological Bulletin, vol. 115, no. 2, pp. 268-287, Mar. 1994.

[3] E. Izard, “Innate and universal facial expressions: Evidence from developmental and cross-cultural research,” Psychological Bulletin, vol. 115, no. 2, pp.288-299, Mar 1994.

[4] B. Scassellati, “Eye finding via face detection for a foveated, active vision system,” in Proc. 15th Nat. Conf. Artificial Intelligence, pp. 969 - 976, 1998.

[5] Xu Ya, and Liu Guang-Yuan, “A Method of Emotion Recognition Based on ECG Signal,” in proc. of International Conference on Computational Intelligence and Natural Computing, pp. 202-205, 2009.

[6] C. Puri, “StressCam : Non-contact Measurement of Users ’ Emotional States through Thermal Imaging,” in proc of Int. Con. on Human Computer Interaction, pp. 1725–1728, 2005.

[7] H. Xu and K. N. K. Plataniotis, “Affect Recognition Using EEG Signal,” IEEE 14th International Workshop on Multimedia Signal Processing, pp. 299-304, 2012.

[8] Maja Pantic, Leon J.M. Rothkrantz, “Automatic Analysis of Facial Expressions: The State of the Art”, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.22, no. 12, pp. 1424-1445, 2000.

[9] C.L.Lisetti and F.Nasoz, “Using non-invasive wearable computers to recognize human emotions from physiological signals", EURASIP Journal of Applied Signal Processing, no. 11, pp. 1672-1687, 2004.

[10] H.G.Hosseini and Z.Krechowec, “Facial expression analysis for estimating patients emotional states in RPMS", In the Proceedings of the 2004 IEEE International Conference on Engineering in Medicine and Biology, IEEE EMBC 2004, vol. 2, pp. 1517- 1520, 2004.

[11] Zhihong Zeng, Maja Pantic, Glenn I. Roisman, Thomas S. Huang, “A Survey of Affect ecognition ethods: Audio, Visual, and Spontaneous Expressions”, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.31, no. 1, pp. 39-58, 2009.

[12] S. Ioannou, V. Gallese, A. Merla, “Thermal infrared imaging inpsychophysiology: potentialities and limits”, Psychophysiology 51 pp.951–963, 2014.

[13] T.J.Phillips, “High performance thermal imaging technology", Advanced Semiconductor Magazine, vol. 15, no. 7, pp. 32-36 , 2002.

[14] M.M.Khan, R.D.Ward, M.Ingleby, “Classifying pretended and evoked facial expressions of positive and negative affective states using infrared measurement of skin temperature", ACM Transactions on Applied Perception, vol. 6, pp. 122 , 2009.

[15] L.Trujillo, G.Olague, R.Hammoud, B.Hernandez, “Automatic feature localization in thermal images for facial expression recognition", IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, CVPR Workshops, p. 14, 2005.

[16] K. Kuraoka, K. Nakamura, “The use of nasal skin temperature measurements in studying emotion in macaque monkeys”, Physiol. Behav. 102, pp.347–355, 2011.

[17] D. Shastri, A. Merla, P. Tsiamyrtzis, I. Pavlidis, “Imaging facial signs of neurophysiological responses”, IEEE Trans. Biomed. Eng. 56 pp.477–484, 2009.

[18] I. Pavlidis, P. Tsiamyrtzis, D. Shastri, a. Wesley, Y. Zhou, P. Lindner, P. Buddharaju, R. Joseph, A. Mandapati, B. Dunkin, B. Bass, “Fast by nature – how stress patterns define human experience and performance in dexterous tasks”, Sci. Rep. 2, pp. 1–9, 2012 .

[19] B.R.Nhan and T.Chau, “Classifying affective states using thermal infrared imaging of the human face", IEEE Transactions on Biomedical Engineering, vol. 57, pp. 979-987, 2010.

[20] N.Powar, ”An approach for the extraction of thermal facial signatures for evaluating threat and challenge emotional states", Doctor of Philosophy in Electrical Engineering, The School of Engineering of the university of Dayton , 2013.

[21] E. Izard, “Innate and universal facial expressions: Evidence from developmental and cross-cultural research,” Psychological Bulletin, vol. 115, no. 2, pp.288-299, Mar 1994.

[22] D.T.LIN, "Facial Expression Classification Using PCA and Hierarchical Radial Basis Function Network", Journal of Information Science and Engineering, vol. 22, pp. 1033-1046 , 2006.

[23] T.Kurozumi, Y.Shinza, Y.Kenmochi, K.Kotani, "Facial Individuality and Expression Analysis by Eigenspace Method Based on Class Features or Multiple Discriminant Analysis", ICIP 1999.

[24] H.Nguyen, K.Kotani, F.Chen, B.Le, "A thermal facial emotion database and its analysis", PSIVT 2013.

[25] H.Nguyen, K.Kotani, F.Chen, B.Le, "Fusion of Visible Images and Thermal Image Sequences for Automated Facial Emotion Estimation", Journal of Mobile Multimedia, Vol. 10, No. 3&4, pp. 294–308, 2014.