

© 2012 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Efficacy of Biophysiological Measurements at FTFPs for Facial Expression Classification: A Validation

Ihsan A. Oz and Masood Mehmood Khan

Abstract—Recent works [1,2] suggest that thermal intensity values (TIVs) measured around the facial thermal feature points (FTFPs) can help distinguish between the facial expression of affective states. This work investigates if the average pixel grey-levels, instead of TIVs, measured in sub-image masks around the FTFPs allow classifying facial expressions. Thermal infrared images from the IEEE OTCBVS database were used to distinguish between three expressions. The pixel grey-levels measured in sub-image masks were used to measure, for each individual, the Euclidean distance between images of different facial expressions. Linear discriminant analysis was performed to obtain the hyper-planes for separating the clusters of sample images. Significant pixel grey-level differences were observed at FTFPs between the three facial expressions. More than 96 of the original images in a three-expression Gaussian mixture model were separable and clustered around distant centroids in a discriminant space.

I. INTRODUCTION

THERMAL infrared imaging (TIRI) has been employed for illumination invariant, non-invasive and psycho-physiology based facial expression classification [1-5]. Though recent literature cites studies claiming various levels of successful facial expression classification and recognition using TIRI, not many of the reported facial expression classification approaches have been validated yet.

In [1,2], a new feature extraction approach was introduced that for measuring affect-induced thermal variance in thermal infrared images (TIRIs) of facial expression. Starting with 16 square segments selected along the facial muscles, (each of 36×36 pixels), highest thermal intensity values (TIVs) in each square segment were recorded and analyzed to get the $\text{Correlation}_{\min}$ and Variance_{\max} within the image TIVs. The TIRIs were repeatedly divided into an increasing number of square segments along the major facial muscles. The process was repeated until significant thermal variations were discovered at 75 square segments located along the major facial muscles. The TIV data gathered from these 75 FTFPs resulted in significance $\text{Correlation}_{\min}$ and Variance_{\max} in

the TIVs in TIRIs. Figure 1 shows these 75 FTFPs on a neutral human face, exhibits a muscular map of a human face, and represents the geometric profile of the facial thermal feature points [1,2]. The 75 sites shown in Figure 1 consistently experienced significant variation in the TIVs with a change in facial expression.

The major objective of this work is to validate the effectiveness of measurements taken at FTFPs using pixel grey-levels measurements. In order to examine the validity and effectiveness of FTFPs in facial expression classification, it was necessary to select an appropriate database of thermal infrared facial images. Then, using a set of appropriate image processing techniques including; image normalization, registration and subtraction, comparison between various facial expressions was made. Finally appropriate analytical and algorithmic approaches were needed to determine if the pixel grey-levels measurements would allow distinguishing between the three facial expressions.

Section II presents the thermal infrared images selected from the IEEE OTCBVS database. Section III presents the employed analytic and algorithmic approaches. Section IV reports results and the work is concluded in section V.

II. THE THERMAL INFRARED IMAGE DATABASE

Automated facial expression classification (AFEC) is still a new and less exploited domain. Though visible-spectrum images of various affective states are available, very few databases of TIRIs portraying various affective states are available in public domain. For example, the Japanese Female Facial Expression (JAFPE) Database [6] is considered a very good collection of facial expression images but this database contains only visible-spectrum images of 10 female Japanese models posing in 7 facial expressions [6]. Such a database cannot be used for this work.



Manuscript received November 04, 2011. This work was partially supported by Department of Mechanical Engineering, Curtin University, Perth, Western Australia 6012.

Ihsan A. Oz is with the Department of Mechanical Engineering, Curtin University, Perth, Western Australia 6012 (phone: +61 4 78023541; e-mail: ihsan.oz@student.curtin.edu.au).

Masood Mehmood Khan is with Curtin University Department of Mechanical Engineering, Perth, Western Australia 6012.

Fig. 1. Left- Geometric profile of FTFPs; Middle- FTFPs on a facial muscle map; and Right- FTFPs on a human face.



Fig. 2. Images of four individuals from the IEEE OTCBVS database. Expressions shown: Left – neutral; Middle – happy; Right – angry.

The USTC-NVIE (Natural Visible and Infrared facial Expression) database has started getting a lot of attention recently. It contains high quality images with both spontaneous and posed expressions of more than 100 subjects, recorded simultaneously by a visible and an infrared thermal camera [7]. However access to the database requires formal correspondence and approval from the curators.

The publicly available IEEE OTCBVS Benchmark [8] contains two useful TIRI databases; Dataset 02 - IRIS Thermal/Visible Face Database and Dataset 04 - Terravic Facial IR Database [9]. However the Terravic Facial IR database does not contain facial images with varying expressions. The IRIS Thermal/Visible Face database contains 8-bit grey-scale versions of JPEG thermal infrared images. These images do not contain any temperature/thermal information or heat values. Rather, the images provide only pixel grey-level information and show how facial skin temperature would vary with a change in affective state. The quality of images in this database is considered high. So it is a popular and widely used TIRI database. As the scope of this work was limited to verification of the 75 FTFPs approach, only 10 best TIRIs from the IEEE OTCBVS database were selected for this investigation. Figure 2 shows neutral, happy and angry expressions of four individuals used in this work.

III. ALGORITHMIC APPROACH

A. Image Normalization

The adaptive single scale retinex (ASR) method of normalization was invoked on the TIRIs for this work [10].

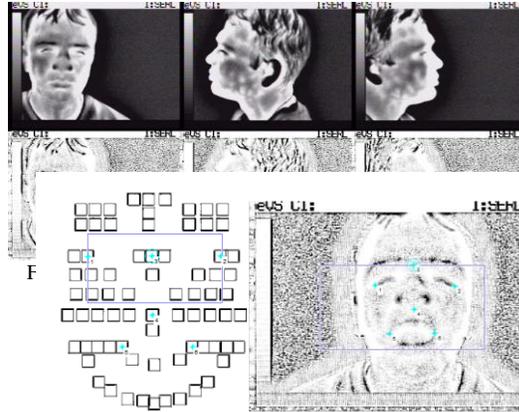


Fig. 4. Left- FTFP Template image showing the 6 control points; Right - 6 control points on a facial image

Figure 3 exhibits image normalization results when ASR method was applied on a set of images exhibiting the neutral



Fig. 5. Three individuals' images after registration and subtraction. Left – neutral; Middle – happy expression; Right - the difference between neutral and happy expressions.

expressions.

B. Initial FTFPs Registration and image subtraction

A FTFP template image was taken as the base image and the preceding input difference images of happiness and anger were registered to it using 6 pairs of control points on the eyes, mouth, and nose, illustrated in Figure 4. An image subtraction operation was performed simultaneously,

between the FTFP template image and the input image. Some of the resulting images are shown in Figure 5.

This image registration and subtraction approach wasn't very efficient and resulted in several problems. All the images were forced-scaled to the size of the FTFP template image. The images in these faces were stretched and transformed accordingly to fit this window size. The 75 FTFP windows, though present in each image, were difficult to realize. The resulting images were quite darker than the original images after image subtraction. Also, the FTFP windows were difficult to align in some images. Especially the windows at the extremes of the mouth and the bottom of the chin were problematic. Hence this process of image registration and subtraction was abandoned as it made it difficult to obtain the average pixel value within each FTFP window.

A major drawback of registering with the FTFP windows was that extracting consistent and reliable pixel information from the resulting images was difficult. The windows surrounding the sides of the mouth were noticeably smaller than the windows above the eyes. This was due to the stretching and squashing of the window images during the alignment process. Furthermore, it was difficult to automate this process for calculating the average pixel value in each window as these windows could not be defined clearly. Hence a more appropriate solution was needed.

C. FTFPs registration using point masks

Using a 'point mask' constructed with the points at the centre of each FTFP window resulted in a pseudo FTFP template, shown in Figure 6. Using this approach, no direct conflict between any two points was noticeable. All points were of equal size (one pixel). The mask could be produced to any scale and would allow accurate alignment possible. If a scale too large was selected compared to the target image, the pixels would be lost when scaling down. This was experienced when using a mask with dimensions of 272 (width) by 320 (height) pixels. However, if the reverse occurs, and a scale smaller than the target image was selected, no pixels were lost when the mask was scaled up, although some pixels would be 'doubled'. Therefore, the most optimal solution was to select a point mask slightly smaller than the face in a target image. This would allow stretching and manipulation, and help determine an arbitrary and consistent method of selecting the 'doubled' pixels which could be removed manually. Hence, a mask of dimensions of 100 (width) by 117 (height) pixels was obtained. The dimensions of the target images were 320 by

240 pixels.

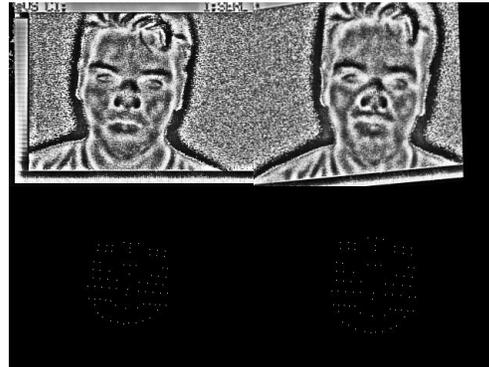


Fig. 7. The resulting FTFP masks after alignment. Neutral image and corresponding point mask is on the left, happy is on the right.

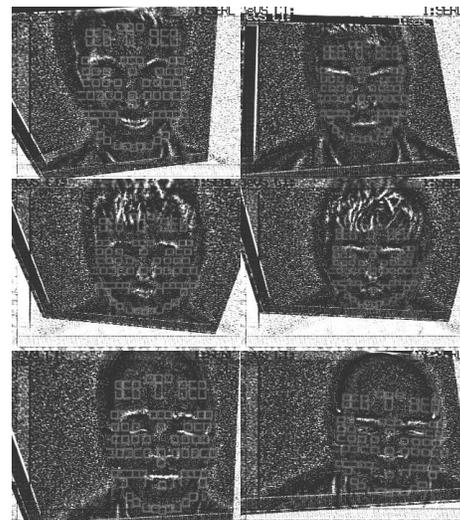


Fig. 8. Improved FTFP window registration for 3 individuals. Images on the left show happiness, images on the right show anger

The FTFP mask allowed for an efficient calculation of the average pixel value in each sub-mask (window) that it represented. Firstly, the mask was inverted, such that it became a set of white pixels (pixel value = 255), and then manipulated such that it was aligned to the face of a target image. Following this, only the layer containing the mask was saved, which was simply an image of a pure black background of zero pixel value with 75 white pixels. This layer was compared with the original image. Both images had the same dimensions. From this comparison a 16x16 'virtual window' centered at each pixel. This was constructed as simply an array of pixel values. The average pixel value in each window was easily calculated and the entire process was easily automated. Figure 6 shows an inverted FTFP mask containing 75 pure white pixels (pixel value = 255) on a pure black background (pixel value = 0). Figure 7 shows the FTFP mask after alignment. Improved image registration

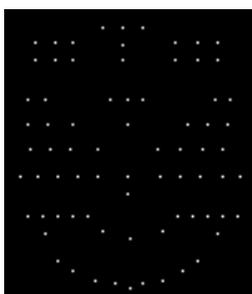


Fig. 6. The inverted FTFP mask

results are shown in Figure 8.

TABLE I
EUCLIDEAN DISTANCE BETWEEN NEUTRAL, HAPPY, AND ANGRY EXPRESSIONS

Individual's Name	Neutral-Happy	Neutral-Angry	Happy-Angry
Bernard	246	120	250
Brad	198	207	194
David	242	251	230
Faysal	217	177	219
Gribok	158	152	95
Justin	307	221	293
Nash	110	133	146
Priya	364	91	357
Rangan	248	126	207
Tak	269	262	152

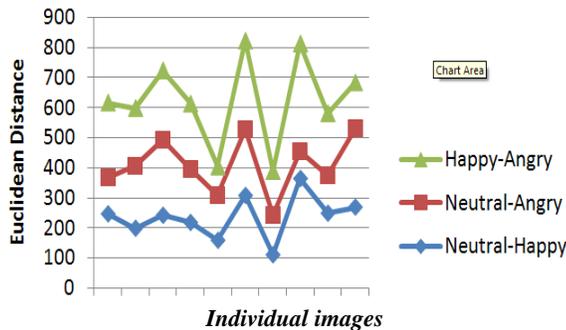


Fig. 9. Euclidean distances between neutral, happy, and angry expressions

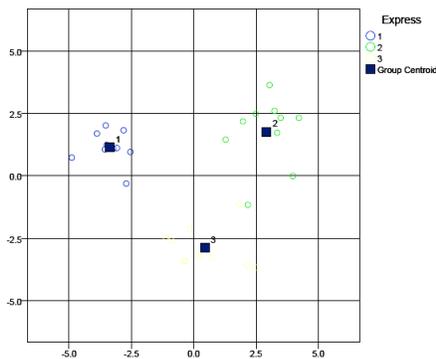


Fig. 10. Separation between the three facial expressions in a discriminant space

Classification Results^a

Original	Express	Count	Predicted Group Membership			Total
			1	2	3	
	1	10	10	0	0	10
	2	0	0	10	0	10
	3	0	0	1	9	10
	%	1	100.0	.0	.0	100.0
		2	.0	100.0	.0	100.0
		3	.0	10.0	90.0	100.0

a. 96.7% of original grouped cases correctly classified.

Fig. 11 Classification of original images using linear discriminant analysis.

IV. RESULTS

The Euclidean distances between neutral, happy and angry expressions of 10 individuals are reported in Tables I. Figure 9 exhibits how the measured Euclidean distance remains consistent in spite of the inter and intra personal differences between the individuals. Figure 10 shows how the three facial expressions were separated in a discriminant space. The linear discriminant analysis results are presented in Figure 11 showing more than 96% successful classification of the original images. In Fig. 11, 1, 2 and 3 respectively represent neutral, happy and angry expressions. As obvious from these results, pixel grey-levels measured in the 75 FTFP windows would suffice classification of the three facial expressions even in the absence of thermal/temperature variation data.

V. CONCLUSION

This work demonstrates that pixel grey-level values measured at FTFPs would allow distinguishing between the neutral, positive and negative facial expressions of affective states. Thus the effectiveness of FTFP approach for automated facial expression classification was verified. However, a larger database of TIRIs will be required to build a robust classifier.

REFERENCES

- [1] M.M. Khan, M. Ingleby and R.D. Ward, "Automated facial expression classification and affect interpretation using infrared measurement of facial skin temperature variation," *ACM Transactions on Autonomous and Adaptive Systems*, vol. 1, no. 1, 2006, pp. 91-113.
- [2] M.M. Khan, R.D. Ward and M. Ingleby, "Classifying pretended and evoked facial expression of positive and negative affective states using infrared measurement of facial skin temperature," *ACM Transactions on Applied Perception*, vol. 6, no. 1, 2009, pp. 6:1-6:22.
- [3] G. Koukiou, G. Panagopoulos and V. Anastassopoulos, "Drunken person identification using thermal infrared images," in *Proc. Digital Signal Processing*, Sentorini-Hellas, 2009, pp. 1-4.
- [4] S. Wang, Z. Liu, S. Lv, G. Wu, P. Peng, F. Chen, and X. Wang "A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference," *IEEE Transactions on Multimedia*, vol. 12, no. 7, 2010, pp. 682-690.
- [5] C. Puri, L. Olson, I. Pavlidis, J. Levine, and J. Starren., "StressCam: Non-contact measurement of users' emotional states through thermal imaging," in *Proc. Computer-Human Interaction*, Portland, Oregon, 2005, pp. 1725-1728.
- [6] M.J. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba. 1998. "Coding Facial Expressions with Gabor Wavelets," *Proc. 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, Nara Japan, April 14-16 1998, pp. 200-205.
- [7] S. Wang, Z. Liu, S. Lv, Y. Lv, G. Wu, P. Peng, F. Chen, and X. Wang, "A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference," *IEEE Transactions on Multimedia*, vol. 12, No. 7, Nov. 2010, pp. 682-691.
- [8] OTCBVS. 2007. Otcvbs Benchmark Dataset Collection. Object Tracking and Classification in and Beyond the Visible Spectrum, Oklahoma State University. Accessed July 3, <http://www.cse.ohio-state.edu/otcbvs-bench/>.
- [9] R. Miezianko, Terravic Research Infrared Database. Available: <http://www.terravic.com/research/facial.htm>.
- [10] Y.K. Park, S.L. Park, and J.K. Kim, "Retinex Method Based on Adaptive Smoothing for Illumination Invariant Face Recognition," *Signal Processing* vol. 88, No. 8, 2008, pp. 1929-1945.