A Novel Face Detection Algorithm Using Thermal Imaging

*Yuen Kiat Cheong, Vooi Voon Yap, Humaira Nisar
Faculty of Engineering and Green Technology
Universiti Tunku Abdul Rahman
Kampar, Malaysia
*cheongyuenkiat@gmail.com

***Abstract*— Face detection is an important step for successful face recognition system. Variation in lighting conditions result in degradation in performance of face detection. Thermal images are not affected by variation in illumination. In this paper, we have proposed a novel face detection algorithm by using thermal imaging. In our proposed algorithm, Otsu's thresholding method is used for converting the thermal images into binary form. Horizontal projection of the image is calculated to identify the global minimum that helps in identifying height and width of the head region. We have developed a UTAR-YK Thermal database to verify the performance of the proposed algorithm. Experimental results show that our proposed algorithm has an average accuracy of 92.16%, that is measured by finding the overlapping ratio between the automatically-detected face region and manually drawn face region. An average 1.13 bounding box ratio, which is the comparison of the size of the box between the automatically-detected face region and manually drawn face region.**

***Index Terms*— Face detection, image segmentation, thresholding, thermal imaging**

I. Introduction

Face detection is an important research topic that belongs to the field of pattern recognition and computer vision. It offers a non-intrusive way of identifying a person. Hence, it is widely used in security applications. For example, access control, in which the user's face must be verified via face detection application in the camera. The most important application of face detection and face recognition [11] belongs to the class of surveillance systems, which are commonly used by the law enforcing agencies to find criminals and terrorists. In the past decade, face recognition research is widely conducted in the visual spectrum band. The captured images are highly dependent on the lighting conditions of the environment. The problem of the visual imaging systems increases when the lighting conditions are not proper. e.g. low light as in the case of evening sun or irregular illumination. As a consequence, face detection and face recognition using digital images in visual spectrum can lead to performance degradation. Therefore, thermal imaging systems are introduced and these are not dependent on lighting conditions.

In addition, the chances of error increase a lot when the video surveillance operators have to sit in front of surveillance systems for long periods of time. Hence, an automated face detection system can lead to better performance in terms of accuracy and time saving. Thermal images are widely used in biometric applications [1, 2]. Some of the advantages of the thermal images [2] are:

1) These are not affected by illumination variation, therefore, works well in dark condition.
2) These are hard to fake than visible images. Because the vein under the skin is directly related to heat emission.
3) These contain complementary information [3] to visible images, which opens a large amount of possibilities of data fusion for enhancing pattern recognition applications.

There are numerous techniques available for face detection using thermal images and normal images (knowledge based methods, feature invariant facial approaches, template matching methods and appearance based methods [4, 10]). One of the famous methods for face detection, is proposed by Viola and Jones [5]. Y. Zheng's proposed a thermal image based face detection method [6]. In this technique, a projection profile analysis algorithm is proposed, where region growing and morphology operations are used to segment the body of a subject; then the derivatives of two projections (horizontal and vertical) are calculated and analyzed to locate a minimal rectangle containing the face area. The problem of this method is the time consumption in the region growing process. Their approach to segmentation examines neighboring pixels of initial "seed points" and determines whether the pixel neighbors should be added to the region. Hence, it consumes a long processing time for the system.

In J. Mekyska's [7] method, the thermal image is binarized and segmented by Otsu's method. Vertical and horizontal projections are calculated to find the rectangle containing face area. The limitation of this method is that it cannot detect the face when a person stands in different poses; at different distances. Based on their method, half of the vertical projection's length is calculated after the image segmentation. After that the half of the vertical projection's length is used to calculate the width of the head. Hence, the

978-1-4673-5159-1/14/$26.00 ©2014 IEEE

poses of the person and the distance between the subject and camera leads to the failure of image segmentation and the result of face detection.

In this paper, the proposed method has taken advantage of above two algorithms to solve the limitation of these two methods in term of the processing time and accuracy rate. In our proposed algorithm, the thermal image is segmented using Otsu's method. Then the horizontal projection is calculated and analyzed to locate the global minimum area. The height of the head is calculated from the global minimum. The width of the head is calculated using horizontal projection as in Eq. 1. The face of the person is cropped and detected using horizontal width information.

The rest of the paper is organized as follows: Section II describes experimental setup. The proposed algorithm is discussed in section III. Experimental results are presented in section IV. Finally, paper in concluded in section V followed by references.

II. Experimental Setup

The proposed thermal imaging based face detection system employs a thermal camera to acquire thermal images. A database of thermal image has been populated which will be discussed in section II B.

A. Face detection system using thermal imaging

For thermal image acquisition, FLIR I7 is used. It has a resolution of 140 x 140 pixels, range of temperature is -20^{o} to 250^{o}C and frame rate is 9 Hz. Images are stored on a SD card in JPEG format. The computer equipped with Intel core I5 2450 and 6 GB Ram. It receives the thermal images from the thermal camera and processes with proposed algorithm using MATLAB R2008b.

B. Face detection system using thermal imaging

The UTAR-YK thermal database consists of 300 images which are acquired from 30 subjects. The images are captured under 2 poses which are 1) hand hanging on sides, and 2) hands folded in chest from five distances. All face images were divided into five subsets; based on the distance between the thermal camera and the subjects. Fig. 1 shows the experimental setup for acquisition of thermal image for 5 different subsets. These five subsets are categorized on the basic of the distances between camera and object which are shown in TABLE 1. The reason for developing this database is to study the variation of detection rate between difference possess and different distances within the same pose. Fig. 2 and Fig. 3 shows the thermal images of 30 persons for subset 1 only. Table 1 shows the distance of the subject and the camera in subset 1.

III. Proposed Algorithm

In this paper, we have proposed an illumination invariant face detection system by using thermal imaging. The purpose of using thermal imaging is because it is not affected by variance of illumination. In our proposed algorithm, thermal image is converted into the gray scale image and is binarized based on Otsu's thresholding method. After that, the horizontal projection of the image is calculated to identify the global minimum. The height and width of the head region is identified by using the global minimum point. The block diagram of the proposed algorithm is shown in Fig. 4. Following are the steps involved in the proposed algorithm.

TABLE 1: Distance for each subset

	Distance between camera and subjects	Hand hanging on sides	Hand hanging on sides
Subset 1	471 cm	30 subjects	30 subjects
Subset 2	411cm	30 subjects	30 subjects
Subset 3	351cm	30 subjects	30 subjects
Subset 4	291cm	30 subjects	30 subjects
Subset 5	231cm	30 subjects	30 subjects

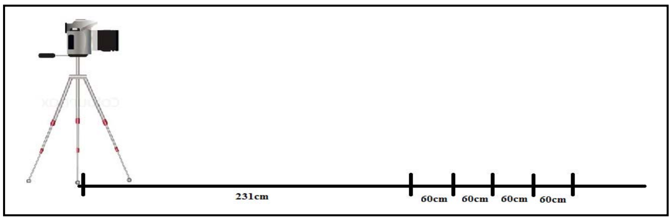


Figure 1: Experimental setup for thermal image acquisition

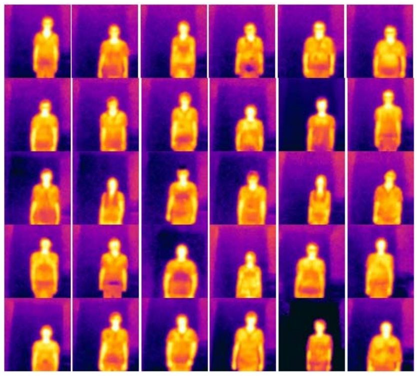


Figure 2: Thermal images for 30 persons (Subset 1 with hand hanging on sides)

Figure 3: Thermal images for 30 persons (Subset 1 with hand folded on chest)

A. Steps of proposed algorithm

1. Convert colour thermal image to gray scale image. Binarize the gray scale image by using the Otsu's thresholding method [8]. Otsu's method is used to automatically perform histogram shape-based image thresholding.

2. The horizontal projection (summation of the pixels of each row of the image) is calculated. The curve smoothing technique is used to smooth the data in the column vector by using a moving average filter which is shown in Fig. 5.

3. Global minimum points are calculated from the horizontal projection curve which is shown in Fig. 5 [9].

4. Sort all the local minimums coordinate in ascending order and save the second local minimum coordinate as Y2.

5. Find the starting point of the curve and save the coordinate of it as Y1.

6. Find the height between Y1 and Y2 which is H = Y2-Y1.

7. The part of face from Y1 to H/2 is selected and then positions X1 and X2 are estimated as the left and right limits of this portion.

8. The lower part of the face is detected by means of the following formula [2].

$$Y2\_new=Y1+(X2-X1)*(13/9) \qquad (1)$$

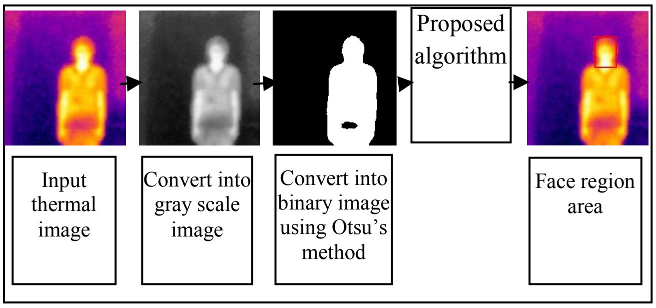


Figure 4: Block diagram of proposed algorithm

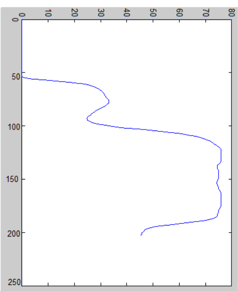


Figure 5 a. Horizontal projection curve

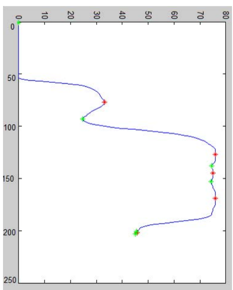


Figure 5b.Horizontal projection curve with minimum points

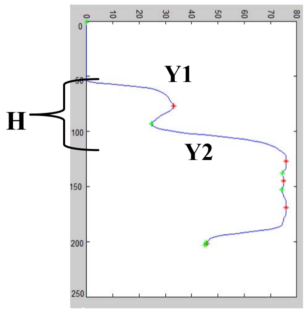


Figure 6a. Head region in projection curve

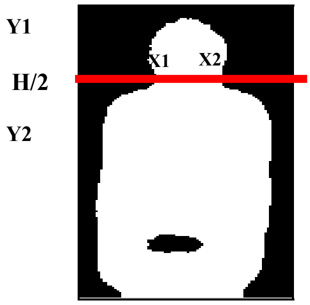


Figure 6b. Head region

IV. EXPERIMENTAL RESULTS

The result of the proposed face detection algorithm has been evaluated in this section. To test this face detection system, UTAR-YK thermal database is used. Fig. 7 shows the subjects in thermal images with five different distances. All of the subjects have hand hanging and the size of the subject/person is proportional to the distance of the object/person from the camera. Fig. 8 shows the subjects in thermal images with five different distances with hands folded in chest and the size of the subject/person is proportional to the distance of the object/person from the camera.

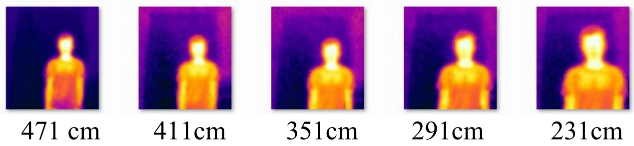


Figure 7: Thermal image with first pose

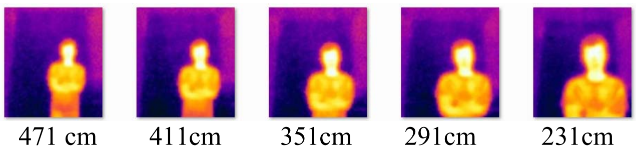


Figure 8: Thermal image with second pose

A. Result comparison for UTAR-YK thermal testing set

The proposed method has been compared with a two methods and the result is shown in TABLE 2 and TABLE 3. Based on the detection result, it is obviously shown that our proposed method has improved the face detection rate are compared to J. Mekyska [7] and Y. Zheng [6].

TABLE 2: Comparison of methods for standing pose

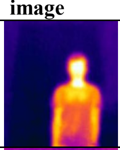
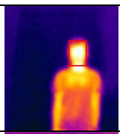
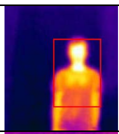
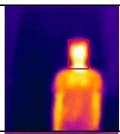
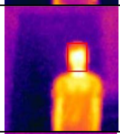
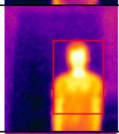
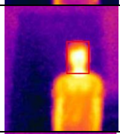
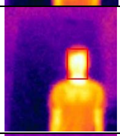
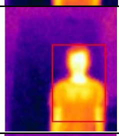
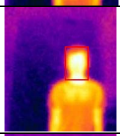
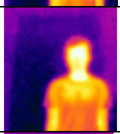
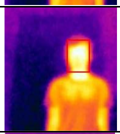
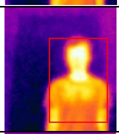
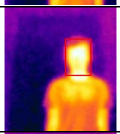
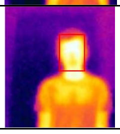
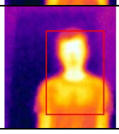
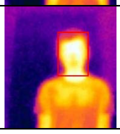
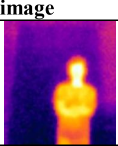
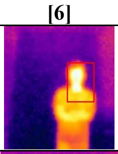
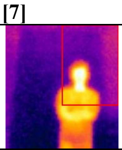
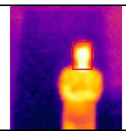
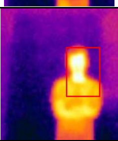
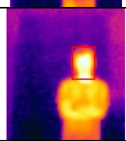
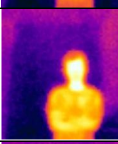
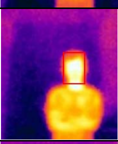
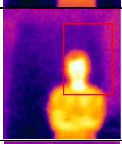
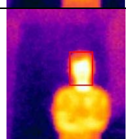
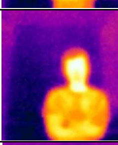
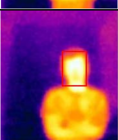
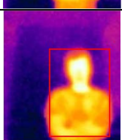
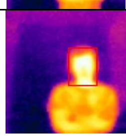
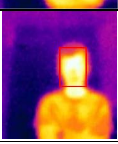
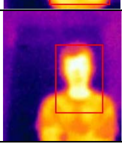
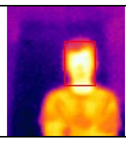
	Original image	Y. Zheng [6]	J. Mekyska [7]	Proposed method
471 cm				
411 cm				
351 cm				
291 cm				
231 cm				

TABLE 3: Comparison of methods for crossing hand pose

	Original image	Y. Zheng [6]	J. Mekyska [7]	Proposed method
471 cm				
411 cm				
351 cm				
291 cm				
231 cm				

B. Overlapping ratio

The accuracy of our algorithm can be measured by calculating the overlapping ratio [7] between the automatically-detected face region and manually drawn face region.

$$r_m = \frac{2\|M_{AUTO} \cap M_{MAN}\|}{\|M_{AUTO}\|+\|M_{MAN}\|} \quad (2)$$

Where M_{AUTO} is binary mask of auto detected face region, while M_{MAN} is the binary mask of manually drawn face region. Symbol '$\cap$' denotes the Boolean AND operator and '$\|\ \|$' denotes the norm operator. $r_m \in [0\ 1]$; $r_m = 1$ is perfect face detection.

The experimental results for the testing set with hand hanging pose is shown in TABLE 4. Our proposed method has achieved average accuracy 92.16% for overlapping ratio. Whereas, Y. Zheng has obtained average accuracy of 82.31% and J. Mekyska has obtained average accuracy of 71.25%. Our result has achieved 10.30% higher accuracy than Y. Zheng and 21.36% higher accuracy than J. Mekyska.

The experimental results for the testing set with hand crossing pose are shown in TABLE 5. Our proposed method has achieved average accuracy 92.61% for overlapping ratio. Whereas, J. Y. Zheng has obtained average of accuracy 75.83% and J. Mekyska has obtained average accuracy 68.48%. Our result has achieved 16.77% higher accuracy than Y. Zheng and 24.13% higher than J. Mekyska

TABLE 4: Comparison of overlapping ratio for 150 Images (Hand hanging pose)

Overlapping ratio	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	Average
Y. Zheng [6]	0.8768	0.8560	0.8320	0.7634	0.7875	0.8231
J. Mekyska [7]	0.5873	0.6811	0.7527	0.7595	0.7822	0.7125
Proposed method	0.9241	0.9243	0.9237	0.9186	0.9175	0.9216

TABLE 5: Comparison of overlapping ratio for 150 images (Hand crossing pose)

Overlapping ratio	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	Average
Y. Zheng	0.6715	0.7611	0.7726	0.8206	0.7657	0.7583
J. Mekyska	0.4705	0.6036	0.7648	0.7949	0.7906	0.6848
Proposed method	0.9298	0.9304	0.9272	0.9254	0.9180	0.9261

C. Bounding box ratio

The bounding box ratio of the face detection region can be measured between the bounding box area of automatically-detected face region and bounding box area of manually drawn face region.

$$r_b = \frac{B_{AUTO}}{B_{MAN}} \quad (3)$$

Where B_{AUTO} is bounding box area of auto detected face region, while B_{MAN} is the bounding box area of manually drawn face region. $r_b > 1$ is bounding box of manually drawn face region larger than auto detected face region. As we know that, the detected faces are used for the face recognition process. Auto detected face region larger than ground truth image can take longer processing time for the face recognition system. Whereas, the auto detected face region smaller than ground truth image can influence the face recognition result due to less information contained in the image. Therefore, $r_b = 1$ or close to 1 is good for face detection.

The experimental results of the proposed method are shown in TABLE 6 and 7. Our proposed method has achieved 1.1369 and 1.1324 for average bounding box ratio. These two results are close to 1 which showing good performance as compared to algorithm proposed by J. Mekyska and Y. Zheng .

TABLE 6: Comparison of bounding box ratio for 150 images (Hand hanging pose)

Bounding box ratio	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	Average
Y. Zheng	1.1407	1.3051	1.1546	1.9592	2.2414	1.5602
J. Mekyska	18.3850	10.4382	7.8847	7.2294	5.3999	9.8674
Proposed method	1.1178	1.1100	1.1379	1.1950	1.1240	1.1369

TABLE 7: Comparison of bounding box ratio for 150 images (Hand crossing pose)

Bounding box ratio	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	Average
Y. Zheng	3.6416	3.3759	2.7072	2.0898	1.9423	2.7513
J. Mekyska	23.3835	13.6912	7.1872	6.5990	5.6072	11.2936
Proposed method	1.1334	1.1424	1.1060	1.1400	1.1405	1.1324

D. Processing time

The processing time of face detection can be measured by the TOC (timer) function on MATLAB. TOC function is a stop clock to measure performance of the programme. TABLE 8 and TABLE 9 are shows the total processing time for the 150 images in hand hanging pose and folded hand pose using three different methods. The total processing time of three methods are shown in TABLE 8 and TABLE 9. Our proposed method has achieved 0.2528s and 0.2461s in total processing times for 300 images. Whereas, Y. Zheng has achieved 23.2648s and 30.2699s and J. Mekyska has achieved 0.1272s and 0.1171s. The proposed algorithm has achieved average of 22.75% higher detection rate than J. Mekyska algorithm and average of 1.13 bounding box ratio. However the processing time is 0.1s different between both algorithms. In addition, our proposed algorithm has 15 times faster than Y. Zheng algorithm and almost same detection rate with each other.

TABLE 8: Comparison of total processing times for 150 images (Hand hanging pose)

Time (s)	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	Average
Y. Zheng	14.729	15.247	30.993	29.075	26.28	23.2648
J. Mekyska	0.1129	0.1279	0.123	0.1461	0.1263	0.1272
Proposed method	0.2398	0.2474	0.2598	0.2608	0.2562	0.2528

TABLE 9: Comparison of total processing times for 150 images (Hand crossing pose)

Time (s)	Subset 1	Subset 2	Subset 3	Subset 4	Subset 5	Average
Y. Zheng	33.4795	35.5511	29.6358	29.5029	23.1805	30.2699
J. Mekyska	0.1195	0.1173	0.1207	0.1148	0.1134	0.1171
Proposed method	0.2295	0.2401	0.2582	0.2528	0.2503	0.2461

V. CONCLUSION

In this paper, we have proposed an illumination invariant face detection algorithm using thermal imaging. In our proposed method, we have improved the face detection rate and the speed of detection. In this algorithm, thermal images binarized using Otsu's thresholding method. After that, the horizontal projection of the image is calculated to identify the global minimum. The height and width of the head region is identified by using the global minimum point. We have also developed a database (UTAR-YK thermal database) to evaluate the performance of the proposed algorithm. The experimental results for our proposed algorithm in terms of face detection rate and processing time shows that the proposed algorithm has achieved good results.

ACKNOWLEDGEMENT

We would like to thank for Mr. Lam Hee Long from Accutac (M) Sdn Bhd, Penang, Malaysia for providing equipment for this research. This work is supported by the UTAR research fund grant (UTAR-RF) funded by UTAR, Malaysia (Project No. IPSR/RMC/UTARRF/2012-C2/Y03)

REFERENCES

[1] Aamir Saeed Malik, Tae-Sun Choi, Humaira Nisar, "Depth map and 3D Imaging applications: Algorithms and Technologies," IGI Global, Nov 2011.

[2] J.Mekyska, V.Espinosa –Duro, and M.Faundez-Zanuy "Face segmentation: A comparison between visible and thermal images", Security Technology (ICCST), 2010.

[3] L. H. Liang, A. H. Zhou,G. Y. Xu, B. Zhang and L. H. Zhao, "A survey of human face detection," Chinese Journal of Computers, Vol. 25, No 5, pp. 450-458, 2002.

[4] M. Faundez-Zanuy "Biometric Security Technology" IEEE Aerospace and Electronic Systems Magazine, Vol.21 nº 6, pp.15-26, June 2006.

[5] M. H. Yang, D. J. Kriegman, N. Ahuja, "Detecting faces in images: a survey," IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (1) (2002) 34-58.

[6] N. Otsu, "A threshold selection method from gray-level histograms." IEEE Transactions on Systems, Man and Cybernetics, Vol. 9, No. 1. 1979, pp. 62-66.

[7] P. Viola, M. Jones, “Robust real-time object detection.” Technical Report CRL 2001/01, Cambridge Research Laboratory, 2001

[8] S. Kong, J. Heo., B. Abidi, J. Paik, and M. Abidi, “Recent advances invisual and infrared face recognition: a review”, Computer Vision & Image Understanding, vol. 97, p. 103-135, 2005.

[9] V. Espinosa-Duró, M. Faundez-Zanuy, J. Mekyska “Beyond cognitive signals”. Cognitive computation, 3(2), 374-381.

[10] YK Cheong, VV Yap, H Nisar, “A robust face recognition algorithm under varying illumination using adaptive retina modeling, ICMLS 2013, - AIP Conf. Proc. 1559, 155 (2013).

[11] Y.Zheng “Face detection and eyeglasses detection for thermal face recognition “, SPIE, 2010.

.