Sensors & Transducers, Vol. 163 , Issue 1, January 2014, pp. 262-269

Sensors & Transducers

http://www.sensorsportal.com

Thermal and Visible Sensor Application: Physiological Thermal Moment Invariant Analysis for Infrared-Based Face Identification

Khairul Hamimah ABAS, Herlina ABDUL RAHIM
Department of Control and Mechatronics Engineering, Faculty of Electrical Engineering,
Universiti Teknologi Malaysia,
UTM Skudai, 81310 Johor, Malaysia
Tel.: 012 700 5316, fax: +607-5566272
E-mail: khairulhamimah@fke.utm.my, herlina@fke.utm.my

Received: 27 June 2013 /Accepted: 9 November 2013 /Published: 31 January 2014

Abstract: Up to date methods and approaches in face identification field heavily depend on facial characteristics; such as location of eyes, length of nose and mouth. Ambient lightings have much influenced the visibility of these facial characteristics; where the visibility varies significantly with inconsistent external light source. In this paper, we present an extended framework for face identification based on thermal information extracted from facial images acquired from a Raytheon Palm-IR-Pro and Raytheon L-3 Thermal-Eye 2000AS sensored lense for themal images and Panasonic WV-CP234 for visible images. The inspiration initiating to this research is to engage in extracting significant facial characteristics from the acquired bio-thermal distribution information within a face, which differs from current facial characteristics that are visible over the skin. Encouraging results are produced which demonstrates the high capability of Hu's classical moment invariants as a feature in thermal based face identification and introducing new ways for classical methods to be further utilized in theoretical and empirical research area. *Copyright © 2014 IFSA Publishing, S. L.*

Keywords: Biometrics, Face identification, Infrared, Moment invariants, Centroid.

1. Introduction

Analysis and identification of facial images acquired from a real and non-ideal imaging system within the visible spectrum still holds many complications since the appearance of faces in concern varies dramatically due to incident angle and light variation, facial expressions, head pose, and image quality. Despite the fact that many advance research have shown dramatic progress in visible spectrum imagery, problems due to the nature of the approach used in visible spectrum is still an open issue. Images acquired from visible spectrum are formed primarily due to reflection and due to this; it is difficult to process because of the high dependency on incident angle and lighting variation from external light source.

As an option to avoid dependency on incident angle and lighting variation from external light source, a review done in [1] suggests the employment of thermal spectrum imagery. Thermal infrared (IR) imagery is basically based on heat emission. Since thermal IR imagery is independent to external light source, problems encountered in visible spectrum-based systems do not exists in thermal IR imagery. This was proven by works done in [2] and [3]. Due to

 Article number P_RP_0062

this, thermal imagery is commonly used for vision in dark and poorly lit environments. Recently, multimodal fusion (thermal and visible spectrum imagery) emerges as paramount method employed for facial identification. Fusion between thermal and visible spectrum imagery was primarily done with one common objective: to create a robust system with supports from thermal images for visible images obtain in a poorly lit environment. The utilization of thermal images in this context neglects the true potential of thermal information distributed within a given image. One of the advantages that were overlooked in fusion systems is the distribution of heat energy or thermal energy within a given face, which differs between individuals, even for identical twins. In thermal spectrum imagery, heat radiation captured through an IR imagery camera depicts a heat map based on heat radiated from blood vessels under the face skin. Thus, individuals having different distribution of blood vessels under the face skin produce different heat maps. Therefore, thermal energy distributed within a face has high potential as features that can be used for identification purposes. This has been proven by works done in [4].

In [4], thermal energy had been interpreted as values that represent blood vessels under the face skin. The obtained thermal distribution is closely related to the distribution of blood vessels under the face skin. The features extracted are used to form a unique thermal faceprint. Distance transform is used to obtain an invariant representation for face identification. This work tremendously achieved 86 percent of rank 1 recognition for the University of Houston database. Motivated by the potential shown in work [4], research on the capabilities of thermal energy distribution as features for identification had been conducted.

In [5], geometric moment invariants were utilized for object identification. To the best of our knowledge, moment invariants are mostly used for objects and hand-writing identifications. Therefore, we ventured the use of Hu's moment invariants for face identification. The Canny and Sobel edge detector was employed in [5] for segmentation prior to the calculation of geometric moment invariants. The first four of Hu's moment invariants were extracted as features from each image. This work [5] achieved 90 % in correct identification rate with Sobel edge detector, while, by employing the Canny edge detector approach, it could only achieved up to 70 % in correct identification rate.

Motivated by works done in [4] and also inspired by the implementation of Hu's classical moment invariants for object classification done in [5], this paper presents the extension of the implementation of moment invariants (with respect to centroid point obtained from each pose of the registered datasets) for thermal-based face identification system. Due to independency towards facial features; such as length of eyes, nose, and mouth, the proposed approach has the potential to be implemented for identification of other body parts; such as arms, abdomens, and legs. The proposed system initially filters the background scenery by implementing seeded region growing method. Later, before the decomposition stage (as done in previous work [6]), anisotropic diffusion is employed to decrease the effect of noise within the face image. Following this, it decomposes a filtered infrared image into 4 thermal regions via 3-valued threshold method. Nevertheless, region with the lowest thermal value is omitted from further processes. This is to ensure that the background scenery is not taken into consideration. Thereafter, the first moment invariant, I_1, from Hu's classical moment invariants were computed from each decomposed layers (with respect to centroid point obtained from each pose of the registered datasets) which constituted the feature database. Due to the increasing amount of registered images (frontal profiles, mid-profiles and left and right profiles), we introduce a customized yet simple pose estimator. The pose estimator utilizes information obtain previously in the pre-processes stage. With this, additional steps or algorithm for common pose estimation is avoided. Minimum distance measurement is employed for classification between Hu's first moment invariants for stored and tests images. This is similar to template matching method. Further throughout this paper, it can be seen that the proposed approach did not involve any training phase at the time this paper was written. This is an advantage within the proposed system, where, for example, a system needs to be constructed with insufficient amount of data provided by the prospect user. Thus, the amount of data received is incompetent to construct a system with training phase that requires a lot of data for training.

The outline of this paper is as follows: Section 2 demonstrates the application of our proposed approach. Experimental results are discussed in section 3. In section 4, conclusions are drawn. Finally, references are presented at the end of this paper.

2. Methodology

Acquired facial image normally contains background scenery. If the entire image is taken into consideration for feature extraction, it may affect the performance of the system. Therefore, we employed seeded region growing method to remove the background scenery. This is done alongside with other conventional image pre-processes; such as histogram equalization and image normalization, prior to the proposed method. Fig. 1 shows the flowchart for methods employed in the proposed system.

The following subsections demonstrate the application of anisotropic diffusion for noise reduction, followed by a brief overview on 3-valued threshold for thermal region decomposition and Hu's classical moment invariants (with respect to centroid point obtained from each pose of the registered datasets) which was employed in previous work. Later, we introduce a simple yet customized pose estimator based on centroid point's natural behavior. Lastly, a brief overview on minimum distance measurement for classification will be given.

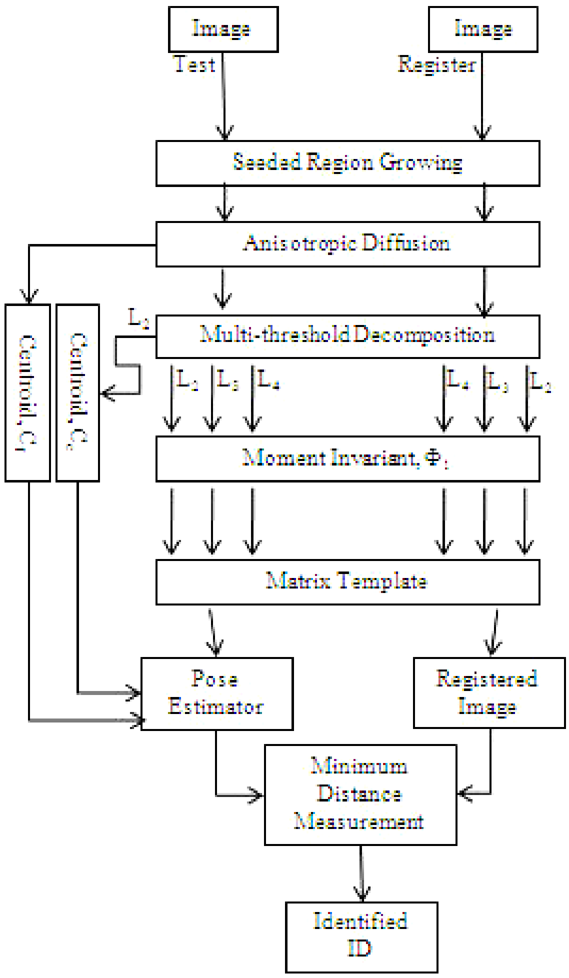


Fig. 1. Flowchart of the proposed system.

2.1. Anisotropic Diffusion

Anisotropic diffusion filter is formulated as a process that enhances object boundaries by performing intra-region as opposed to inter-region smoothing. It is normally used to remove noise from digital images without blurring edges. This is ideal for removing noise and indiscriminate blur edges. For face images, it helps to reduce the noise formed due to facial hairs. This process can be described with the following equation:

$$\frac{\partial I(\bar{x},t)}{\partial t} = \Delta(c(x,y,t)\nabla I(\bar{x},t)), \tag{1}$$

In our case, I is the thermal IR image, $c(x,y,t)$ is called the diffusion function. This diffusion function controls the rate of diffusion and is usually chosen as a function of the image gradient so as to preserve edges in the image.

Anisotropic diffusion filter for discrete cases can be expressed as follows:

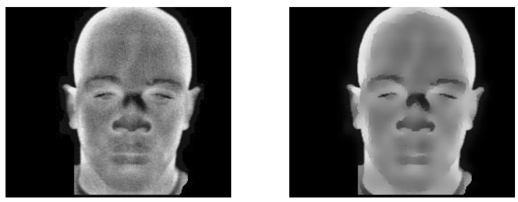
$$\begin{aligned} It+1(x,y) = It + \frac{1}{4} * [&cN,t(x,y)\nabla IN,t(x,y) \\ &+cS,t(x,y)\nabla IS,t(x,y)+cE,t(x,y)\nabla IE,t(x,y) \\ &+cW,t(x,y)\nabla IW,t(x,y)] \end{aligned}, \tag{2}$$

The four diffusion coefficients and gradients in equation (2) represent four directions; north, east, south, and west, with respect to the pixel location *(x,y)*. Each diffusion coefficient and the corresponding gradient are calculated in the same manner as shown in the following equation:

$$cN,t(x,y) = \exp\left(\frac{-\nabla I^2 N,t(x,y)}{k^2}\right), \tag{3}$$

where $I_N,t(x,y) = I_t(x,y+1) - I_t(x,y)$.

Fig. 2(b) shows the result of applying anisotropic diffusion to the segmented facial region shown in Fig. 2(a).



(a) (b)

Fig. 2. (a) Segmented facial region (seeded region growing); (b) Image after anisotropic diffusion.

2.2. Overview On Image Decomposition

In previous work [6], multi-threshold method (in our case is 3-valued threshold method) was employed to decompose input image into several input images (with respect to specified thermal range) for a non-holistically analysis approach.

As aforementioned, heat radiation captured through an IR imagery camera depicts a heat map based on heat radiated from blood vessels under the face skin. Fig. 3 shows an example of blood vessels and veins under the face skin (courtesy of Visuals Unlimited Inc.), which contributed to heat radiation captured through an IR imagery camera. Lesser veins and blood vessels are visible in most convex area of the face; such as the forehead, the cheekbones, the nose area, and also the chin area. These areas are expected to radiate less heat, thus being the coldest area on a face. The inner corner of the eye sockets has a very dense blood vessels and veins connection. Furthermore, in an actual scene, an eyeball too has very dense blood vessels connectivity; therefore, a high heat radiation is to be expected in these areas. Elsewhere is considered to have a mid heat radiation due to the sparsely distributed blood vessels and veins connections. Naturally, the distribution of blood vessels and veins differ between individuals, even for identical twins, thus producing different heat maps.

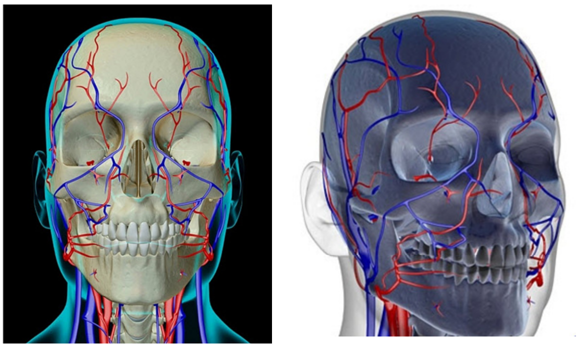


Fig. 3. Graphical demonstration of blood vessels and veins network under the face skin (Courtesy of Visuals Unlimited, Inc.).

As previously clarified, each heat maps captured (IR images) can be grouped into three categories; high heat, mid heat, and low heat radiation regions. For this purpose, we have employed the multi-threshold method for image decomposition to decompose the input image into several images, which consists of a range of heat radiation per image.

The general definition of threshold is represented by the following equation:

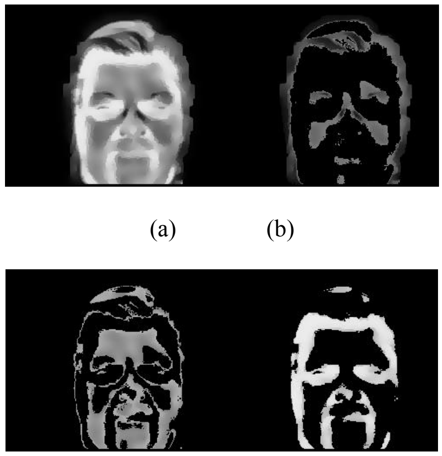
$$g(x,y)=\begin{cases}255, & f(x,y)\geq T\\ 0, & f(x,y)<T\end{cases}, \tag{4}$$

where *f(x,y)* represents the input pixel, *g(x,y)* represents the output pixel, and *T* represents the threshold value. By inserting three threshold values rather than one threshold value, the 3-valued threshold equation can be derived from equation (4) as follows:

$$g(x,y)=\begin{cases}L4, & f(x,y)\geq T3\\ L3, & T2<f(x,y)\leq T3\\ L2, & T1<f(x,y)\leq T2\\ L1, & f(x,y)\leq T1\end{cases}, \tag{5}$$

where T_1, T_2, and T_3 represent the three threshold values whereas L_1, L_2, L_3, and L_4 represent the label for each generated thermal regions. As a result, four thermal regions (four binary formatted images) are generated. In order to obtain these thermal regions, values for T_1, T_2, and T_3 are selected based on the results acquired from the preliminary experiments conducted in [7]. Referring to works done in [7], an initial value for T_1, T_2, and T_3 is randomly selected within the range stated in [8] where the temperatures at all pixels are mapped between 0 and 255. Mapped temperatures between 200 and 225 is said to be common temperature on face, and mapped temperature between 175 – 200 and 225 – 255 are said to be normal temperature on cheeks and maximum temperature on face, respectively. Since the area for maximum mapped temperature on face is small and sparsely located within a face, this would cause the system to identify these areas as noise. Therefore, we selected T_3 to have the initial value of 200; the minimum value for the combination of mapped temperatures for common and maximum temperature on face.

As aforementioned, mapped temperatures between 175 and 200 is said to be normal temperatures on cheeks. With manual tuning done in [7], we discovered that the mapped temperatures between approximately 140 and 200 comprehend temperatures on convex surfaces of a face; such as nose, cheeks, and forehead. Hence, the initial value for T_2 is set to 140. By employing the same manual tuning technique used in [7], T_1 is initially set to approximately 80, where this value affirms with values stated in [8] (mapped temperature value between 0 and 100 normally indicates the background scenery). For maximum assurance that the background scenery is not taken into consideration, the lowest valued region (coldest region, L_1) is omitted from further processes. This makes L_2 the coldest region given within a face. An example of the resulted decomposed image is shown in Fig. 4.



(a) (b)

(c) (d)

Fig. 4. Images produced by the decomposition algorithm. (a) Source image. (b) Low heat region. (c) Mid heat region. (d) High heat region.

2.3. Overview on Hu's Moment Invariants

Originally, Hu's set of classical moment invariants consists of the famous seven rotation invariants from second and third order moments.

Followings are examples of the first four of Hu's moment invariants:

$$\begin{aligned} \Phi 1 &= \mu_{20} + \mu_{02}, \\ \Phi 2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2, \\ \Phi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, \\ \Phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \end{aligned} , \quad (6)$$

Having in mind that thermal energy distributed within a face region are physical densities distributed throughout the face area, we calculated the moment of inertia or impact of the distributed physical density with respect to centroid point obtained from each pose of the registered datasets. For this we have implemented and utilize only the first set of Hu's classical moment invariants, where it is said to be proportional to moment of inertia around the image's centroid. As aforementioned, in our case, image's centroid in this context refers to the centroid point obtained from each pose of the registered datasets.

Followings are the derivation of Hu's first moment invariant corresponding to our proposed method:

$$\phi 1 = \eta 20 + \eta 02 , \quad (7)$$

where

$$\begin{aligned} \eta 20 &= \frac{\mu 20}{\mu 00^2} \\ \eta 02 &= \frac{\mu 02}{\mu 00^2} \end{aligned} , \quad (8)$$

and

$$\begin{aligned} \mu 20 &= M20 - \bar{x}(M10) \\ \mu 02 &= M02 - \bar{y}(M01), \\ \mu 00 &= M00 \end{aligned} \quad (9)$$

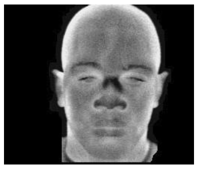
where $\bar{x}$ and $\bar{y}$ are centroid point coordinate (obtained from each pose in the registered datasets) for x-axis and y-axis, respectively, while raw image moment, M_{ij} with pixel intensity, *I(x,y)* are calculated as follows:

$$Mij = \sum_x \sum_y x^i y^j I(x, y), \quad (10)$$

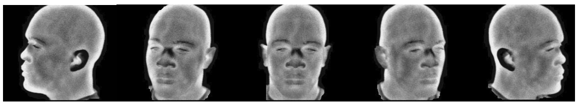
for i,j=0, 1, and 2.

2.4. Pose Estimator

Since our previous approach [6] encountered identification problems for images containing faces with angular deviation more than 45 degrees to the left and right, we extended the capacity of registered images from one registered image to 5 registered images (front profile, mid profiles, left and right profiles). This is shown in Fig. 5. Due to manifold poses available in the registered dataset, a pose estimator is needed prior to classifications process. Up to date, many advance pose estimator are available for employment, nevertheless, we constructed our own pose estimator that utilizes previously obtained information to avoid additional module for the pose estimator itself.



(a)



(b)

Fig. 5. (a) Image used for registered database in [6]. (b) Images used for registered database in current proposed approach.

Based on centroid point behavior, we have constructed a customized yet simple pose estimator to be employed together with our proposed approach. Generally, centroid point in a discrete mass can be defined as follows:

$$C = \frac{\sum m_i r_i}{\sum m_i}, \quad (11)$$

where r_i and m_i are particle positions and mass, respectively. For a binary formatted image, the numerator's particle position, r_i, is substituted with pixel's coordinate and mass, m_i, is substituted with pixel's intensity. For the denominator, $\sum m_i$ is substituted with the total number of pixels with the intensity of 1. The derived equation is as follows:

$$C(x) = \frac{\sum I(x_i)}{\sum pixel}, \quad (12)$$

$$C(y) = \frac{\sum I(y_i)}{\sum pixel}, \quad (13)$$

where *C(x)* and *C(y)* are coordinates for x-axis and y-axis, respectively. Therefore, the actual location of centroid point for a binary formatted image emerges as (*C(x), C(y)*). Note that, intensity, I, holds a value of 1 or 0 for binary formatted images and 0 – 255 for a gray-scaled image.

In physics, centroid point can be defined as an imaginary point in a body of concern where, for convenience in certain calculations, the total weight of the body may be thought to be concentrated. In other words, centroid point is the mean location of all mass density in a substance. By interpreting intensities in images as physical densities, the location of centroid point tends to be allocated near towards or in the high-density region. This principle is shown in Fig. 6.

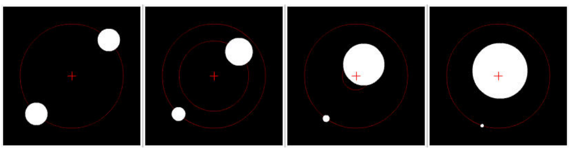


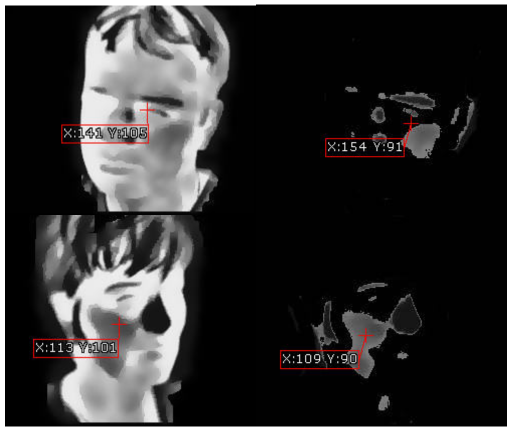
Fig. 6. Characteristics of centroid point.

The amount of radiation emitted by an object increases with temperature; therefore, thermography allows one to see variations in temperature within a face. Temperatures within a face can easily be clustered into 3 major cluster; convex surfaces, non-convex surfaces, and normal flat surfaces. The convex area (consists of cheeks, nose, and sometimes the forehead) plays an important role in our proposed pose estimator. If an image of a face is to be considered as a flat plane with different weights which represents each grayscale valued pixels, the convex surface will encompasses low-weighted weights, whereas the combination of all region will contain a bigger total of weights since more weights are assigned for higher temperature region. Referring to Fig. 6, it can clearly be seen that the area with more weights affects the location of centroid point. As a result from centroid point behavior analysis, we have utilized the centroid point from the convex area (coldest region) in the proposed pose estimator. Pose is estimated by comparing the location of centroid point obtained from whole image to centroid point obtained from the convex area, as shown in Fig. 7. Noted that only the x-axis values are taken into consideration. Since pose is estimated for left and right profiles, the y-axis values are considered a constant value, where changes in the y-axis value do not affect the pose alignment.

2.5. Overview on Minimum Distance Measurement for Classification

Sequentially after pose is estimated, we employ the minimum distance measurement method between the stored and test values of Hu's first moment invariants obtained from each corresponding thermal region for classification purposes. The general definition of minimum distance measurement, x, via Euclidean Distance between two points, P and Q, is shown in equation (14).

$$x = \sqrt{\left(p_1 - q_1\right)^2 + \left(p_2 - q_2\right)^2}\,, \tag{14}$$



(a) (b)

Fig. 7. Samples of centroid point used for pose estimation. (a) This column contains centroid point from whole image. (b) This column contains centroid point from convex region (low heat region).

In our case, since only two values are being compared; therefore equation (14) is redefined as follows:

$$x = \sqrt{\left(r - t\right)^2}\,, \tag{15}$$

where r and t are values of Hu's first moment invariants for registered and test images, respectively.

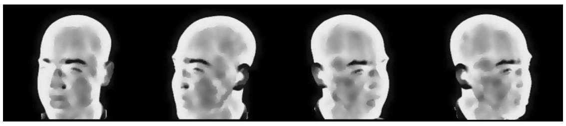
3. Empirical Results

OTCBVS IRIS IR facial database were used to validate the effectiveness of our proposed approach. Fig. 8 shows examples of registered and test images selected from this dataset. All calculations were done with Matlab 7.0 Student Version on a 1.8 GHz Centrino Duo processor with 1 GB RAM.

We conducted two experiments on the OTCBVS IRIS IR facial database to evaluate the performance of the proposed face identification method. We have used 4 test images (2 images with angular deviation within -45 degrees to 45 degrees, and 2 images with angular deviation exceed the previous angular degree range) for 5 registered images (front profile, mid profiles, and left and right profiles). Each image is decomposed into 4 thermal regions where the lowest (coldest) thermal region is not taken into consideration.

In the first experiment, we implemented the original Hu's first moment invariant, Φ1, in the proposed system. In the second experiment, we employed the proposed approach; the first moment invariant (with respect to centroid point obtained from each pose of the registered datasets) in the system. The performance for both experiments is shown in Fig. 9. The CMC curve in Fig. 9 shows that rank 1 identification for the first experiment is approximately 79.16 percent, and rank 2 identification is approximately 95.83 percent. Rank 1 identification for the second experiment achieved approximately 87.5 percent, and rank 2 identification reached approximately 95.83 percent. This is to be expected, since our customized pose estimator only achieved approximately 90 percent of correct pose estimation.

(a)



(b)

Fig. 8. (a) Samples of registered image dataset. (b) Samples of test image dataset.

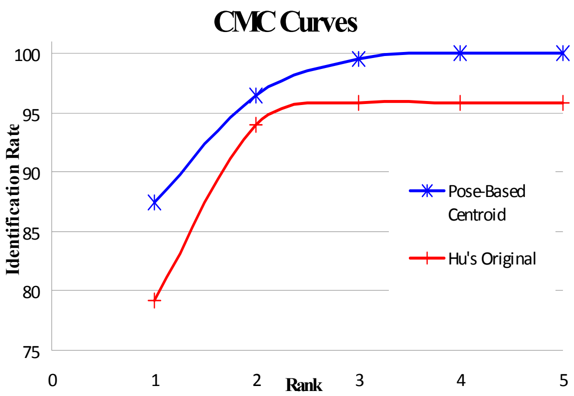


Fig. 9. Comparison between Hu's original moment invariants and our proposed approach.

We have compared the identification performance of our approach with works done in [4]. The comparison of performance between these two approaches is shown in Fig. 10. Referring to Fig. 10, the CMC curves shows that rank 1 identification for our approach is approximately 87.5 percent, which exceeds the performance for work done in [4] (approximately 83.5 %). Albeit our approach demonstrates encouraging performance, the robustness of this approach degrades when health issues are being addressed (fever, flu, etc…). At the moment, this matter is considered as the operational limit for this approach. Furthermore, it is suggested that further refined tuning should be made to enhance our customized pose estimator.

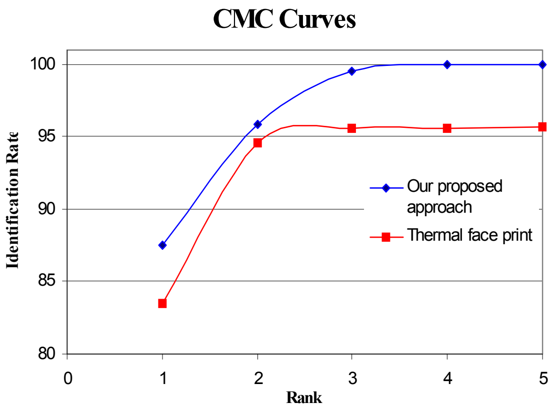


Fig. 10. CMC curves of our approach and works done in (4).

4. Conclusion

In this paper, we extended our previous approach done in [6] by employing anisotropic diffusion prior to the decomposition process and expanded the capacity of registered dataset by regarding frontal profiles, mid profiles, and left and right profiles as registered images. Due to the extension of registered dataset, we introduced a customized yet simple pose estimator primarily formed from thermal distribution analysis and centroid point characteristics. In this research, we are exploring the identification capability within the physiological thermal spectrum imagery. In other words, analysis is done based on the information obtained from distribution of bio-thermal energy physically, without facial features dependency. Our approach was tested on OTCBVS IRIS IR facial database which is publicly available for download at www.cse.ohio-state.edu/otcbvs-bench/. Classifications are done by employing a minimum distance measurement method between the acquired moment invariant from test and registered IR images. As with most methods and approaches, this approach also has some operational limitations. As such, images used in this approach are obtained from individuals with no illnesses since illnesses; such as fevers and influenzas, may affect the thermal distribution. Empirical results obtained shows encouraging performance, where possibilities exist for future improvement by more intensive analysis.

Acknowledgements

The author would like to acknowledge the University Grant for New Academic Staff (4P043), Universiti Teknologi Malaysia.

References

[1]. S. Kong, et al., Recent advantages in visual and infrared face recognition – a review, *Journal of Computer Vision and Image Understanding*, Vol. 97, No. 1, 2005, pp. 103-135.

[2]. Diego A. Socolinsky, Lawrence B. Wolff, Joshua D. Neuheisel, Christopher K. Eveland, Illumination invariant face recognition using thermal infrared imagery, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001),* 8-14 December 2001, Kauai, HI, USA, Vol. 1, pp. 527–534.

[3]. G. Friedrich and Y. Yeshurun, Seeing people in the dark: face recognition in infrared images, in *Proceedings of the 2nd International Workshop, Biologically Motivated Computer Vision* (*BMCV'02)*, 2002, *Lecture Notes in Computer Science*, Vol. 2525, pp. 348-359.

[4]. P. Buddharaju, I. Pavlidis, P. Tsiamyrtzis, and M. Bazakos, Physiology-based face recognition in the thermal infrared spectrum, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, Issue 4, 2007, pp. 613-626.

[5]. Mohamed Rizon et al., Object detection using geometric invariant moment, *American Journal of Applied Sciences*, Vol. 2, Issue 6, 2006, pp. 1876-1878.

[6]. K. H. Abas and O. Ono, Implementation of frontal-centroid moment invariants in thermal-based face identification system, in *Proceeding of the 5th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS'09)*, 2009, pp. 36-41.

[7]. K. H. Abas, O. Ono, K. Suzuki, S. Yoshida, Thermal segmented-based face identification system, in *Proceeding of the Asia Simulation Conference ASC'09*, 2009, Article ID 007.

[8]. P. Buddharaju, I. Pavlidis, and I. Kakadiaris, Face recognition in the thermal infrared spectrum, in *Proceeding of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)*, 2004, pp. 133.
