THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS INC.

SIGNAL PROCESSING
ALGORITHMS, ARCHITECTURES, ARRANGEMENTS, AND APPLICATIONS
SPA 2016
September 21-23rd, 2016, Poznań, POLAND

DIAST Variability Illuminated Thermal and Visible Ear Images Datasets

Syed Mohd Zahid Syed Zainal Ariffin, Nursuriati Jamil
Digital Image, Audio and Speech (DIAST) Lab, Faculty of Computer and Mathematical Sciences
Universiti Teknologi MARA
Shah Alam, Malaysia
syed.mzahid@gmail.com, liza@tmsk.uitm.edu.my

Puteri Norhashimah Megat Abdul Rahman
Faculty of Computer and Mathematical Sciences
Universiti Teknologi MARA Perak
Tapah, Malaysia
phashi655@perak.uitm.edu.my

***Abstract*— As the research in ear biometrics grows, there is a need to have a publicly available ear images database for researchers to test and validate their methods. There are many available ear databases captured in visible light domain. However, images in those databases are mostly acquired in a controlled environment, not in naturally illuminated environments. Up to the point of writing, there is no thermal-based ear image dataset available for research usage. In this paper, we introduce our ear image datasets that were captured in both visible and thermal domain. We present the detailed process of image acquisition, post-processing, and preparation of the ear datasets. The ear dataset consists of 2,200 images captured from 55 subjects in five different illumination conditions, ranging from 2 lux to 10,700 lux, measured using lux meter.**

Keywords- Illumination, Luminance, Thermal image, Ear image dataset section

I. INTRODUCTION

Studies on ear recognition or ear biometrics have been gaining attention in the last 15 years. All thanks to Iannarelli's [1] that has done studies to prove that the uniqueness of human ear, even for twins. Iannarelli's investigation was done by manually measuring the ears based on 12 measurements as shown in Fig. 1. However, the approach is not feasible enough to be implemented in the real world scenario. Therefore, many studies have been done ever since to automate the ear recognition process by extracting new feature and adapting new recognition methods in order to come out with an effective ear recognition system.

Besides the uniqueness of the ear that qualifies it as a good candidate for biometrics application, there are other advantages of the ear that add to its value. First of all, the ear image can be obtained remotely compared to conventional biometrics mode (i.e. fingerprint and iris) that require subject to interact with the sensor. In addition, an ear is also regarded as an organ that is passively used thus reducing the possibility of tampering its structure in contrast to fingerprint that is prone to cuts and injury, thus changing its structure.

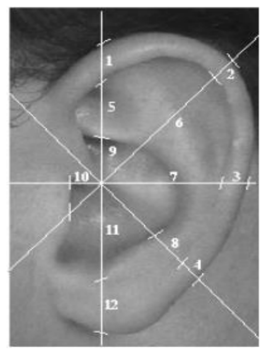


Figure 1. Iannarelli's 12 points of measuring ears.

There are several ear image datasets currently available for research purposes. Some are privately owned, while others can be publicly accessed. All publicly accessible databases consist of visible images, while thermal images are not present. What follows next are descriptions of few visible image datasets and a collection of thermal images of previous works.

A. Visible Image Dataset

USTB (University of Science and Technology, Beijing) Databases

The University of Science and Technology Beijing (USTB) databases are available for academic research [2]. There are three collections in this database comprising image with some pose and illumination variations. Illumination variations of this database are not adequately explained. However, there are yaw pose variations where images were taken rotated to the left at 0°, 5°, 10°, 15°, 20°, 25°, 30°, 40° and 45°; rotated to the right at 0°, 5°, 10°, 15°, 20°, 25°, 30°, 40°, 45°, 50° and 60°.

WPUT Database

The West Pomeranian University of Technology ear databases was collected from 254 and 247 female and male subjects, respectively which age ranging from 0 to more than 50 years old [3]. There are a total of 2,071 ear images which some have occlusions of hair and earrings. Images in the database do show illumination variations. Those variations, however, are not structural. All images from most subjects are all acquired in bright illumination conditions which are in contrast from our dataset where every subject has their ear images taken in five illumination conditions ranging from dark to bright.

IIT Delhi Database

The IIT Delhi Database is contributed by the Hong Kong Polytechnic University consisting of two collections with 125 and 221 subjects, respectively [4]. There are pitch-angle pose variations reflected by the images in this database. However, it does not specifically mention the angle. Similar to WPUT database, the illumination variations in this database are not structural. There are subjects that all images were captured in dark environment or bright environment.

B. Thermal Image Data Collection

Currently, there is no publicly available thermal ear image database for the use of researchers. However, Abaza and Bourlai [5] in their study collected a dataset of side facial profile images from 57 subjects. The datasets consist of 228 images (2 images for each side of the ear) taken indoors over a period of 20 days. Nevertheless, the main intention of this data collection was for face recognition, not ear detection and recognition. Therefore, the live face profile capture may not be suitable for recognition or detection of ears.

To achieve robust ear recognition, researchers must overcome the three main problems: namely pose variations, occlusion and illumination variations. Solving one of the problems is challenging, let alone all three. However, we would like to introduce our ear images dataset that was collected for the purpose of solving illumination problem in ear recognition. The uniqueness of our datasets is that the ear images are acquired in the indoor and outdoor environments, and illuminations are measured using lux meter at actual locations. Furthermore, both visible and thermal images are captured by the imaging device simultaneously ensuring equal luminance value.

Section 2 of this paper discusses basic concept of illumination and measurement method. Details of the ear images collection for the datasets are explained in Section 3, while post-processing and preparations of the ear datasets are discussed in Section 4. Section 5 concludes this paper with further recommendations of the dataset application domains.

II. ILLUMINATION

Illumination can be described as the amount of light present in an area [6]. Illumination is measured in lux (lx) or lumen (lm) per square meter. Lux is a unit of measurement of brightness, or more accurately, illuminance. One lux equals to one lumen of light spread across a surface of one square meter. Equation (1) shows the relation between lux and lumen:

$$Ev = lx = lm/m^2 \quad (1)$$

Lux can be measured using lux meter as seen in Fig. 4. A lux meter is a handheld device for measuring brightness.

III. IMAGE COLLECTION METHODOLOGY

There are three main stages of gathering the ear images for the dataset. Fig. 2 shows the block diagram of the stages involved, which are image acquisition, post-processing and image registration.

A. Thermal and Visible Image Data Collection

Ear images were acquired based on the setup as shown in Fig. 3 adopted from [5] and [7]. The camera used for image acquisition is FLIR E60 that is capable of taking both thermal and visible images simultaneously. With a sensitivity of 0.05°C, this camera can operate in a condition with temperature ranging between -20°C and 650°C. However, the temperature range for the data collection is set between 23°C and 37°C (the average body temperature). The resolution for the visible image is 2048 x 1056 pixels, while the thermal image is 360 x 240 pixels. The distance between the subject and the camera was set at 1.5 feet as the thermal image's resolution is quite small. If the camera is too far away from the subject, the captured ear image is too small. Therefore, the image acquired may not have the detail information needed. However, if the camera is too close to the subject, it may not be feasible to operate.

The acquisition process started with the subject seated on a provided chair. The subject's ear is faced towards the camera, and the camera is set to be at the same level with the ears. Then, the illumination is measured using lux meter as shown in Fig. 4. The lux meter sensor is placed directly on the ear to ensure an accurate reading. Then the image is captured. Both thermal and visible images are captured simultaneously in one snap. Two snaps were taken for each ear in the same illumination condition. The process is repeated in all different illumination conditions.

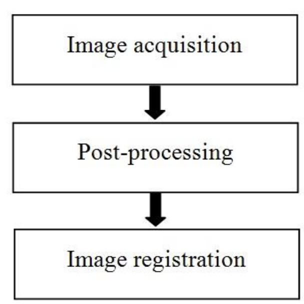


Figure 2. Block diagram of the ear collection methodology.

The illumination is manipulated by controlling the light source, the position of the subject and the location of the site. For example, the light source is controlled by moving the site

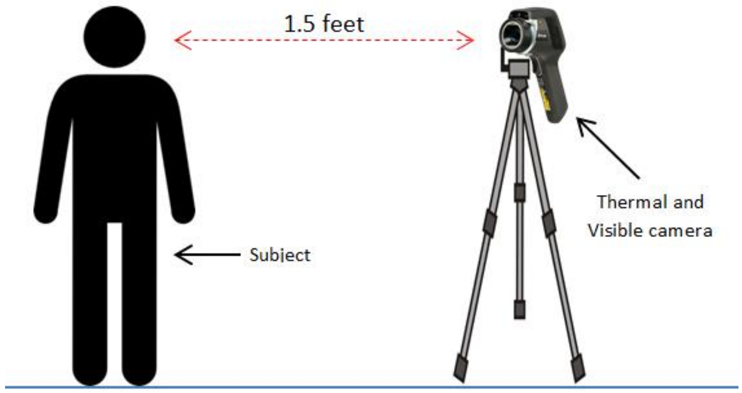


Figure 3. Image acquisition set-up.

Figure 4. Illumination measured using lux meter.

location from indoors to outdoors, by drawing the curtains of the room and by turning off the lights in the room. The subjects are sometimes asked to sit near of far from the window, outside in a field during daytime with bright sunlight. For this study, five different illumination conditions ranging from bright to dark are chosen.

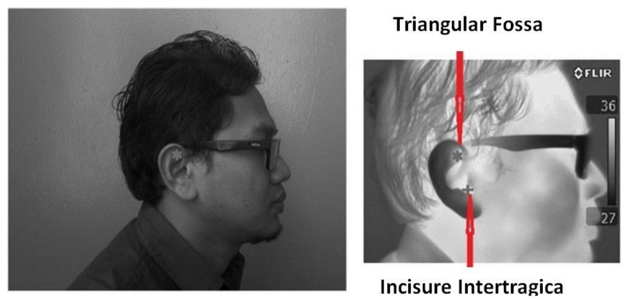
Our image collections are acquired from 55 male volunteers aged 18 to 25. The age mean and standard deviation are 19.49 and 1.63, respectively. This image acquisition process was done within 30 days timeframe.

B. Image Post-processing

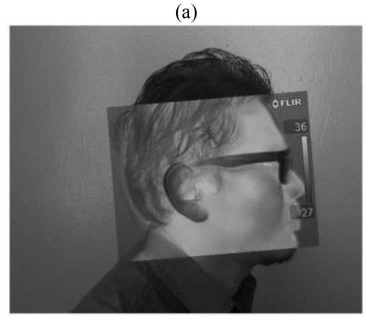
The acquired images are further processed to prepare them as ear datasets. The RGB images of both thermal and visible images were first transformed into grayscale images. This step helped in reducing the processing time for later stages in recognition. As the resolution of visible images is significantly higher than thermal images, they were resized from 2048 x 1056 pixels to 800 x 600 pixels.

C. Image Registration

Even though both thermal and visible images were taken using the same image acquisition device, they, however, are acquired by two different sensors. The deviation of the sensors position caused slight misalignment between the thermal and visible image. Therefore, image registration needs to be done to align both thermal and visible images. For the image registration, Triangular Fossa and the Incisure Intertragica are used as the reference points for both images as these features are stable over time for a particular ear [8].



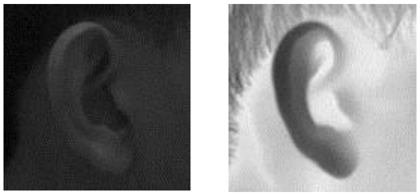
(a)



(b)

Figure 5. (a) Unregistered images with the position of Triangular Fossa and Incisure Intertragica, (b) Registered image.

The registration process starts by defining the fixed image and the moving image. The visible image is set to be fixed image as the resolution is higher than thermal image. Then the control points (i.e. Triangular Fossa and Incisure Intertragica) were defined for both images. Fig. 5 (a) shows the location of Triangular Fossa and Incisure Intertragica. Ear transformation was done by translating, rotating and scaling the moving image to align the control points of both images. The result of the registration process is shown in Fig. 5 (b). Once the registration is done, the area region was cropped to the size of 125 x 125 pixels. Fig. 6 shows the cropped ear region.



(a) (b)

Figure 6. An example of the cropped registered image where (a) thermal image and (b) visible image.

IV. DIAST Ear Image Dataset

A total of 2,200 ear images were collected from the 55 male volunteers. The image datasets comprise 1,100 visible images (550 left ears, 550 right ears); the remaining 1,100 are thermal images (550 left ears, 550 right ears). The illumination values range from 2 lux to 10,700 lux. Fig. 7 and Fig. 8 illustrate few examples of the visible and thermal ear images while Table 1 describes the overview of the data collected.

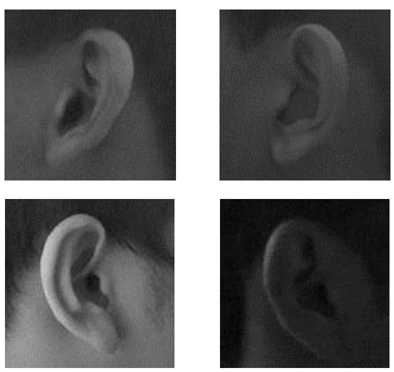


Figure 7. Examples of the visible images acquired.

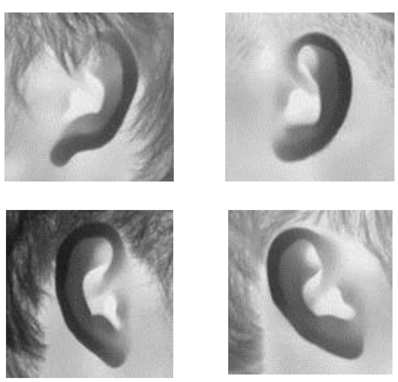


Figure 8. Examples of the thermal images acquired.

Currently, there is no established classification of illumination condition. The determination whether the environment illumination is dark or bright is very subjective and is done without quantitative measure. After analyzing the acquired images, we, therefore, proposed one classification rule for determining the illumination condition based on lux as shown in Table 2. This classification may help in training and testing illumination invariant ear recognition. Fig. 9 shows the example of images acquired in different illumination condition.

Table 1. Overview of the ear image datasets.

Attribute	**Range**
No. of Subjects	55 male subjects
No. of Images	2200 ear images - 1100 left ears (550 visible images, 550 thermal images - 1100 right ears (550 visible images, 550 thermal images)
Subjects age range	18 to 25 years old (Mean = 19.49, SD = 1.63)
Illumination range	2 lux to 10700 lux

Table 2. Illumination condition based on Lux.

Lux	**Condition**
0-20	Dark
21 – 100	Average
> 100	Bright

In general, the environment for dark illumination condition was set up in a room without any light source. Then the illumination was controlled by manipulating the curtains that block the light from the fluorescent lamp outside the room. For average illumination condition, the light source is from the fluorescent lamps situated at the room's ceiling. The variations from moderate to bright were manipulated by controlling how many lamps were switched on and the position of subjects compared to the light source. Images taken outdoor during a bright sunny day are grouped under bright illumination condition.

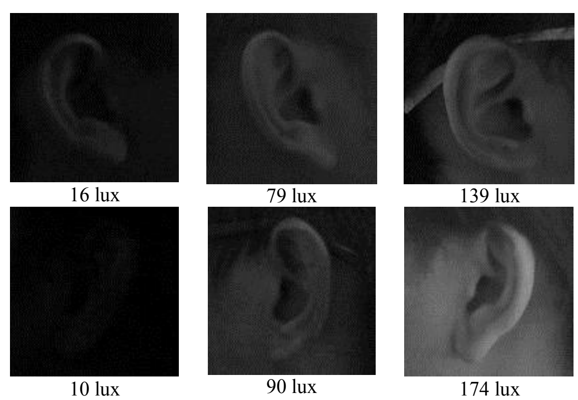


Figure 9. Example of the right (*top row*) and left (*bottom row*) ear images with different illumination condition, dark (*left*), average (*middle*) and bright (*right*).

The naming convention of the image file is based on the format describe in Table 3.

Table 3. The format for image naming.

Code	Attribute	Description
X	Ear side	L - left; R - right
#	Subject ID	Numbered from 01 to 55
#		
x	[redacted]	
#	Lux value	Numbered from 00001 to 10700
#		
#		
#		
#		
x	Image modality	v - visible; t - thermal
#	Image count	Count of the image of the same subject with the same illumination condition. Number 1 or 2

Fig. 10 shows several images with name from the datasets. For examples, Fig. 10 (a) shows the left ear visible image of subject 48 taken in an environment with illumination of 60 lux while Fig. 10 (b) is the thermal image that was simultaneously acquired. Meanwhile, the image in Fig. 10 (c) is the left ear image of subject 48 with 60 lux illumination but taken several seconds later from Fig. 10 (a).

L48x00060v1 (a) L48x00060t1 (b) L48x00060v2 (c)

Figure 10. Images with name in the datasets.

V. EAR RECOGNITION

The thermal and visible ear images datasets can be used to validate any ear recognition methods developed. For example, we did an experiment to compare two ear recognition methods. The two recognition methods tested are based on eigenfaces (E) [9] and local binary pattern eigenfaces (LE) [10]. Three images per subject (one from each illumination condition) were used as training while one image per subject was used for testing. Table 4 describes the recognition results for both methods.

Table 4. Ear recognition rate.

Image Modality	Left (%)		Right (%)		Overall (%)	
	E	*LE*	*E*	*LE*	*E*	*LE*
Thermal	80.00	90.91	78.18	92.73	79.09	91.82
Visible	23.64	30.91	20.00	23.64	21.82	27.27

Based on the result in Table 4, the recognition rate for left and right ears (and subsequently overall recognition rate) using local binary pattern is higher than eigenfaces. This result confirmed the finding in [10]. In addition, the recognition rates for thermal images are always higher compared to visible images. The performance of visible images suffers due to the illumination variations whereas the thermals images are not affected.

VI. CONCLUSION

The establishment of this visible and thermal ear datasets is meant to boost the studies in ear biometrics. Besides the available visible ear database, our visible ear dataset can be an additional dataset for researchers to validate their methods. Furthermore, it adds to the scarce collections of thermal-based ear images. The illumination variations and lux measures make our datasets unique from others; thus it may trigger more studies to solve the illumination problem in ear recognition. The lack of available thermal ear datasets also has been reduced by our thermal ear datasets. The datasets may prompt new studies on thermal-based ear recognition and concurrently proceed with fusion of thermal and visible ear recognition as both types of images have been registered. Researchers who are interested in getting this dataset for their studies are invited to contact the corresponding author through email.

ACKNOWLEDGMENT

Due acknowledgement is accorded to the Ministry of Education, Malaysia for the funding received through the Fundamental Research Grant Scheme, (FRGS/2/2013/ICT07/UITM/02/2), Research Management Center and Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Malaysia.

REFERENCES

[1] A. V. Iannarelli, Ear Identification in Forensic Identification Series. Paramont Publishing, Freemont, 1989.

[2] University of Science and Technology Beijing, USTB database. http://www1.ustb.edu.cn/resb/ en/index.htm.

[3] D. Frejlichowski and N. Tyszkiewicz, "The West Pomeranian University of Technology Ear Database - A tool for testing biometric algorithm" in *Image Analysis and Recognition,* Berlin, Springer-Verlag Berlin Heidelberg, 2010, pp. 227-234.

[4] A. Kumar and C. Wu, "Automated human identification using ear imaging," Pattern Recognition, vol. 45(3), pp.956--968, 2012.

[5] A. Abaza and T. Bourlai, "On Ear-based Human Identification in The Mid-wave Infrared Spectrum", Computing vol. 31(9), pp. 640-648, 2013.

[6] K. Devlin, A. Chalmers and E. Reinhard, "Visual calibration and correction for ambient illumination," ACM Trans. on Applied Perception, vol. 3(4), pp. 429-452, 2006.

[7] D.M. Bulanon, T.F. Burks and V. Alchanatis, V, "Image fusion of visible and thermal images for fruit detection," Biosystems Engineering vol. 103(1), pp. 12-22, 2009.

[8] L. Nanni and A. Lumini, "A multi-matcher for ear authentication," Pattern Recognition Letters, vol. 28(16), pp. 2219-2226, 2007

[9] M. Turk and A. Pentland, Eigenfaces for Recognition, Journal of Cognitive Neuroscience, Vol. 3, No. 1, 1991, pp. 71-86.

[10] L. Lei, D.-H Kim, W.-J Park and S.-J Ko, Face recognition using LBP Eigenfaces, IEICE Transactions on Information and Systems, Vol. E97D, Issue 7, 2014, pp. 1930-1932.