Automatic Genders Determination from Medical Microscope Image by using Image Processing Technique

Wichan Thumthong
Faculty of Computer Science and Information Technology
Rambhai Barni Rajabhat University
Chanthaburi, Thailand
E-Mail: wichan.t@rbru.ac.th

Abstract

This study focused on using of the image processing technique to calculate the histological parameters from microscope image of occipital bones to automatic gender determination in case of there was no evidence for identifying. Occipital bones will be highlighted based on 13 parameters to find out the relation with the gender of samplings. The samplings group consisted of 80 images divided into 46 males and 34 females aged between 25 – 90 years. There were two methods applied for this study. To find out gender accuracy and to predict the gender of skeleton. The first is direct discriminant function analysis method, the results of the percentage accuracy for gender determination from 5 parameters was 97.5% of correctly classified and 97.5% of accuracy prediction. The second is stepwise discriminant function analysis method, the results from 3 parameters was 97.5% of correctly classified and 97.5% of accuracy prediction.

Keywords: Forensic Science, Gender Determination, Image Processing, Occipital Bones

1. Introduction

Biological identification consists of 4 types of basic information including gender, age-at-death estimation, height and nationality. To specify the identity of the individual, these biological identification plays a crucial role. Gender determination is a necessary step in forensic science [1].

The most reliable non-metric based research considers gross morphological features of the pelvis [2] [3]. Traditionally, subjective visual assessment of sexually dimorphic features of the pelvis [2], cranium [4]. In Thailand several parts of studies has evaluated the gross morphological analysis used to determine the sex, such as the lumbar vertebrae [5], metacarpals [6] and talus [7]. Metric analyses studies for sex estimation were conducted on various other skeletal elements such as the scapula [8], sternum, proximal hand phalanges [9].

The disadvantage of gross morphological aging and sexing methods are that they usually require a nearly complete and well preserved skeleton. Oftentimes the remains are fragmented, poorly preserved, incomplete skeleton, unable identify by traditional gross morphology feature to estimate a biological profile. In case where there is no clear indication of individualized traits, fragment bone, anthropologists have turned to bone or dental histology, or the study of microscopic tissues [10].

Nowadays, The image processing is applied in many works of pathology by using MATLAB such as image segmentation [11], cell measurement, cell counting, degraded cell measurement and brain extraction from MRI images [12]. Medical image processing is one of the most challenging and emerging topics in today research field [13] and Bio-medical image processing is the most challenging and upcoming field in the present world [11].

However, this research aims to use the image processing technique to calculate the histological parameters from microscope image for automatic gender determination from occipital bones.

2. Research and Methodology

This research consisted the following parts:

2.1 Graphical Method

The samplings group consisted of 80 images divided into 46 males and 34 females aged between 25 – 90 years with know age, sex and the cause of death. All the samples were

sent for routine autopsy at the Department of Forensic Medicine and the fresh frozen un-embalmed cadavers were received as body donors from the Department of Anatomy for medical surgical training at the Cadaveric Surgical Training Center, Faculty of Medicine, Chiang Mai University, Thailand.

The digitized images of bone sections were observed under light microscope at 10x magnification. The photographed consisting of digital JPG image files (image size: 1600 × 1200 pixels).

2.2 Image Processing Method

The image processing technique and the algorithm source code for analyzing histological parameters is written in MATLAB 2016a. The whole process into 3 stages; follow:

2.2.1 Pre-Processing Stage

The original image of the occipital bone is given as the input. The parameters of original image were separated into four sections which also divided the unit area of intact osteon and fragmented osteon in order to use the program to draw and replace the position of intact osteon with red and fragmented osteon with green. Then replace the background-color of the original image with black color and convert to grayscale image.

The brightness levels of the red (R), green (G) and blue (B) components are each represented as a number from decimal 0 to 255, or binary 00000000 to 11111111. The lightness of the gray is directly proportional to the number representing the brightness levels of the primary colors [13].

2.2.2 Threshold Segmentation Stage

In this method the pixels are divided according to the intensity value and are separated. This method is based on the threshold value where the gray scale image is converted to the binary image. We use local methods which adapt the threshold value on each pixel to the local image characteristics for segmentation [14] shown in Fig.1.

2.2.3 Post-Processing Stage

The function regionprops in MATLAB was used to be the measurements are of selected region of an image in pixel count. Regionprops instruction is used to estimate area closed. The area is the actual number of pixels in the selected region [15]. The pixel count of the proposed image depends on the distance between the camera and the object when the picture is taken. A reference object is an object with known area, needed to translate the pixel count area [15].

In this study of the photographed consisting in digital JPG image files the width is 1,600 pixels and the object of femur bone width is 1406.36 μm. Therefore,

1 pixel value = width object/width image;

1 pixel value = 0.879 μm^2

For parameters, if the pixel count area is X pixels. Therefore, the area of parameters (μm^2) is *X* 1 pixel value*, Post-processing stage from original image by using MATLAB shown in Fig. 1.

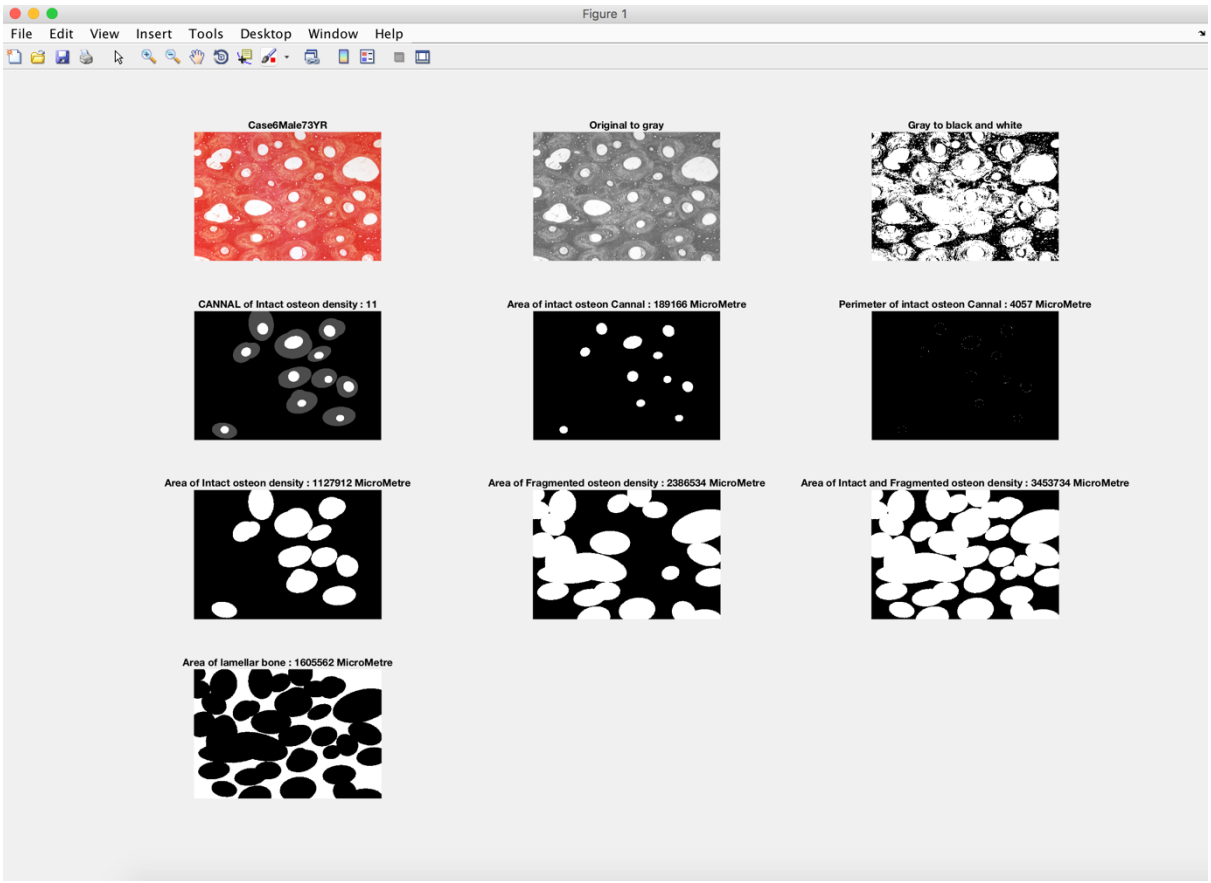

Fig. 1. Post-processing stage from original image by using MATLAB

2.3 Histological Parameters

The sections were observed 13 histomorphometry parameters. Abbreviations of each histomorphometric parameters: secondary osteon area (On.Ar); secondary osteon maximum diameter (On.max); secondary osteon minimum diameter (On.min); perimeter of big secondary osteon (Pm.big.on); big secondary osteon maximum diameter (Big.on.max); big secondary osteon minimum diameter (Big.on.min); perimeter of small secondary osteon (Pm.small.on); small secondary osteon maximum diameter (Small.on.max); small secondary

osteon minimum diameter (Small.on.min) Haversian canal area (Hc.Ar); Haversian canal max diameter (Hc.max); Haversian canal min diameter (Hc.min); peTrimeter of haversian canal (Pm.Hc). The data obtained were analyzed by descriptive statistics for histomorphometric parameters were calculated, such as the mean, minimum, maximum, and standard deviation, as shown in Table 1.

Table 1. Occipital parameters for the pooled genders

Parameters	Mean	S.D.	Min	Max
1.On.Ar [a] (μm^2)	4.883	0.086	4.50	5.00
2.On.max [a] (μm)	4.601	0.093	4.44	4.82
3.On. min [a] (μm)	4.387	0.168	3.79	4.91
4.Pm.big.on (μm)	655.013	108.481	503.00	898.00
5.Big.on.max (μm)	273.850	54.132	201.00	421.00
6.Big.on.min (μm)	177.088	30.779	74.00	260.00
7.Pm.small.on (μm)	98.225	34.991	30.00	229.00
8.Small.on.max (μm)	200.388	48.546	93.00	375.00
9.Small.on.min (μm)	144.350	28.412	75.00	242.00
10.Hc.Ar [a] (μm^2)	3.612	0.140	3.33	3.84
11.Hc.max [a] (μm)	3.404	0.125	3.09	3.66
12.Hc.min [a] (μm)	3.291	0.114	3.02	3.47
13.Pm.Hc (μm)	142.363	36.718	63.00	262.00

[a] Log 10, square micrometers (μm2), μm=micrometers

2.4 Statistical Analysis

Summarizes the descriptive statistics for each measurement was calculated indices of the both genders between males and females are presented in Table 2 such as mean, standard deviations. The results showed that all the measurements were statistically significantly different ($P< 0.05$) between males and females except 7 variable such as On.Min, Big.on.max, Big.on.min, Pm.small.on, Small.on.max, Small.on.min, Pm.Hc the results showed that there were no statistically significant ($P> 0.05$) between genders. There for, 5 variables were not included in the discriminant function analysis.

Table 2. Comparation of occipital histological parameters between female and male

Variable	Female (n=35)		Male (n=45)		t-value	p-value
	Mean	SD	Mean	SD		
1. On.Ar [a]	4.946	0.043	4.832	0.077	-8.409	0.000*
2. On.max [a]	4.663	0.081	4.550	0.068	-6.810	0.000*
3. On.min [a]	4.370	0.175	4.402	0.163	0.841	0.403 [NS]
4. Pm.big.on	746.028	84.082	580.545	57.066	-10.064	0.000*
5. Big.on.max	260.917	52.565	284.432	53.664	1.968	0.053 [NS]
6. Big.on.min	177.806	32.683	176.500	29.499	-0.188	0.852 [NS]
7. Pm.small.on	92.528	33.705	102.886	35.713	1.324	0.190 [NS]
8. Small.on.max	195.25	50.649	204.59	46.921	0.855	0.395 [NS]
9. Small.on.min	139.75	29.657	148.11	27.111	1.316	0.192 [NS]
10. On.Hc [a]	3.487	0.090	3.714	0.074	12.363	0.000*
11. On.Hc.max [a]	3.465	0.086	3.465	0.086	5.721	0.000*
12. On.Hc.min [a]	3.191	0.087	3.373	0.049	11.185	0.000*
13. Pm.Hc	136.000	38.598	147.568	34.679	1.410	0.162 [NS]

n: number of samples, [a] Log 10, * statistically significant at P<0.05,

[NS] statistical not significant P>0.05

This study performed 2 methods, as follows.

2.4.1.1 Direct univariate discriminant function analysis

The discriminant analysis showed that the percentage accuracy for genders determination of the discriminant functions in the direct method and the classification function formulae for all variables were tabulated in Table 3. The results indicated that 1 prarameter are good indicators of genders determination, it was noted that (Hc.Ar) alone can classify the gender in 98.8 %, of both original grouped and cross validated respectively. Therefore, the accuracy obtained by using direct variable could be determined gender determination by histomorphometry, especially the Hc.Ar parameter. While the moderate accuracy classification of gender prediction was observed for 5 variables, followed by On.Ar, On.Max, Pm.big.on, Hc.Max and Hc.min, respectively as shown in Table 3.

2.4.1.2 Stepwise discriminant function analysis

The stepwise discriminant function analysis method found that selected variables in discriminant function have 3 from 5 variables such as, On.Ar, Pm.big.on and Hc.min as shown in Table 3.

Table 3. Direct univariate discriminant function analysis

Parameters	Discriminant function equation	Centroids		Average accuracy (%)	
		M	F	O	C
1. On.Ar*	(7.710×On.Ar)+(-37.020)	-0.377	0.510	87.5	87.5
2. On.Max*	(0.008×Pm.big.on)+(-4.954)	0.217	-0.293	80.0	80.0
3. Pm.big.on*	(0.021×Big.on.max)+(-5.648)	0.444	-0.601	85.0	85.0
4. Hc.Ar	(0.21×Small.on.max)+(-4.603)	-0.220	0.297	98.8	98.8
5. Hc.max*	(8.558×Hc.Ar)+(-34.271)	1.047	-1.416	80.0	80.0
6. Hc.min*	(19.129×Hc.max)+(-66.974)	1.399	-1.893	58.8	58.8

* = observed parameters

O = original group correctly classified, C = cross-validated group correctly

3. Results and Discussion

Result and discussions consisted the following parts:

3.1 Results

This study performed 2 methods, as follows.

3.1.1 The direct discriminant function analysis

The result of direct discriminant analysis using 5 parameters, moderate correlation shown in Table 4, such as, On.Ar, On.max, Pm.big.on, Hc.max and Hc.min. The discriminant score is calculated by multiplying the unstandardized coefficient with each particular measurement, summing them and then adding the constant, as follows.

$$\textit{Discriminant score} = (6.775 \times \textit{On.Ar}) + (0.758 \times \textit{On.Ar.max}) + (0.009 \times \textit{Pm.big.on}) + (-1.286 \times \textit{Hc.max}) + (-8.636 \times \textit{Hc.min}) + (-9.750)$$

3.1.2 The stepwise discriminant function analysis

The analysis found that selected parameters in discriminant function have 3 from 5 parameters shown in Table 4 such as, On.Ar, Pm.big.on and Hc.min. The discriminant score was calculated as:

Discriminant score=(6.926× On.Ar) + (0.009× Pm.big.on) +
(- 9.476×Hc.min)+ (-8.823)

After calculating discriminant score from the direct discriminant function analysis method and the stepwise discriminant function analysis method, it was compared to sectioning point, 0 which is halfway between the female and male centroids. A score greater than 0, signified males, whereas a score less than 0, signified females.

Table 4. Coefficients for the stepwise and direct discriminant function analysis

Parameters	Unstandardized Coefficient	Eigenvalue	Canonical Correlation	Wilks' Lambda	Centroids	Sectioning point
Direct discriminant function analysis						
1. On.Ar	6.775	4.309	0.901	0.188	M=-1.854 F=2.266	0
2. On.max	0.758					
3. Pm.big.on	0.009					
4. Hc.max	-1.286					
5. Hc.min	-8.636					
Constant	-9.750					
Stepwise discriminant function analysis						
1.On.Ar	6.926	4.230	0.899	0.191	M=-1.837 F= 2.245	0
2. Pm.big.on	0.009					
3. Hc.min	-9.476					
Constant	-8.823					

An eigenvalue indicates the proportion of variance explained. (Between-groups sums of squares divided by within-groups sums of squares). A large eigenvalue is associated with a strong function. An eigenvalue as shown in the Table 4, considering each of the

eigenvalue of direct methods and stepwise methods were shown 4.309 and 4.230 respectively, which the eigenvalue must not be under 1 in order to be considered that the discriminant function analysis from the analysis is correct. Moreover, the greater different values between the groups, the more possibly they are to be highly discriminative.

The canonical relation is a correlation between the discriminant scores and the levels of the dependent variable. A high correlation indicates a function that discriminates well. Canonical correlation is the value showing the relation between independent variables and dependent variables which the highest value of canonical correlation from 0.8 up is considered as the high value. The analyzed values of this study are 0.901 and 0.899 respectively, are not extremely high.

Wilks' lambda is a measure of how well each function separates cases into groups. Smaller values of Wilks' lambda indicate greater discriminatory ability of the function and that group means appear to differ. The results of the analysis of direct methods and stepwise methods are 0.188 and 0.191 respectively, shown in (Table 4), which mean the selective independent variables in the discriminant function of two methods are able to divide into groups. Furthermore, considering significant suggests that the value of Wilks' Lambda, which has statistical significance (P-value < 0.005), thus, the group means appear to differ.

3.2 Accuracy of Classification

The results of the percentage accuracy for genders determination of the discriminant function derived from 5 variables in the direct method were presented in Table 5. There was 97.5 % of correctly classified and 97.5 % of accuracy prediction and derived from 3 variables in the stepwise method were presented in Table 5. There was 97.5% of correctly classified and 97.5% of accuracy prediction.

Table 5. Direct and stepwise discriminant model and percent of accuracy

Discriminant model	% corrected classification			% accuracy prediction		
	Male	Female	Total	Male	Female	Total
Occipital bones						
Direct	100	94.4	97.5%	100.0	94.4	97.5%
Stepwise	100.0	94.4	97.5%	100.0	94.4	97.5%

3.3 Discussion

This study performed 2 methods. The results of the percentage accuracy for genders determination of the discriminant function derived from 5 parameters in the direct method was 97.5 % of correctly classified and 97.5 % of accuracy prediction and derived from 3 parameters in the stepwise method was 97.5% of correctly classified and 97.5% of accuracy prediction. The results showed that by using Direct Discriminant Function Analysis and Stepwise Discriminant Function Analysis the accuracy for gender prediction and gender classification were both at 97.5%.

It can conclude that the techniques used for this study are effective and suitable for forensic science because it reflects that the gender accuracy and classification were almost 97.5 percent.

4. References

[1] N. P. Angi Christensen, Eric Bartelink, *Forensic Anthropology: Current Methods and Practice* (Academic Press is an imprint of Elsevier). 2014, p. 426.

[2] P. T.W., "A newly developed visual method of sexing the Os pubis," *Amercan Journal of Physical Anthropology,* vol. 30, pp. 297-302, 1969.

[3] J. B., "A method for visual determination of sex, using the human hip bone," *Amer J Phys Anthrop,* vol. 117(2), pp. 157-68, 2002.

[4] U. D. H. Buikstra J.E, "Standards for Data Collection from Human Skeletal Remain," *Proceedings of a seminar at The Field Museum of Natural History,* 1994.

[5] C. S. Ostrofsky KR, "Sex Determination by Discriminant Function Analysis of Lumbar Vertebrae," *J Forensic Sci,* vol. 60(1), 2015.

[6] P. S. Khanpetch P, D. Troy Case, P. M., "Determination of sex from the metacarpals in a Thai population," *Forensic Sci Int,* vol. 217(1-3) 2012.

[7] P. S. Mahakkanukrauh P, Ruengdit S, Singsuwan P, Duangto P, T. CD., "Sex estimation from the talus in a Thai population," *Forensic Sci Int,* vol. 240:152, 2014.

[8] P.-J. Gretchen R. Dabbs "A Method for Estimating Sex Using Metric Analysis of the Scapula," *journal of ferensic scicences,* vol. 55(1), 2009.

[9] P. K. Pasuk Mahakkanukrauh, Sukon Prasitwattanseree, Troy D, "Determination of sex from the proximal hand phalanges in a Thai population," *Forensic Sci Int,* vol. 226(1-3), 2013.

[10] S. S. Crowder C, "Bone Histology An Anthropological Perspective," *CRC Press Taylor & Francis Group,* vol. 410, 2012.

[11] S. V. D. Dhilip Kumar, K. Sakthi Priya, S. Jeneeth Subashini, "Brain Tumour Image Segmentation using MATLAB," *International Journal for Innovative Research in Science & Technology,* vol. 1, no. 12, 2015.

[12] D. A. S. B. Rajesh C. Patil, "Brain Tumour Extraction from MRI Images Using MATLAB," *International Journal of Electronics, Communication & Soft Computing Science and Engineering,* vol. 2, 2015.

[13] G. Nithyanandam, "Extraction of Cancer Cells from MRI Prostate Image Using MATLAB," *International Journal of Engineering Science and Innovative Technology,* vol. 1, 2012.

[14] S. V. D. Dhilip Kumar, K. Sakthi Priya, S. Jeneeth Subashini, "Brain Tumour Image Segmentation using MATLAB," *International Journal for Innovative Research in Science & Technology,* vol. 1, no. 12, 2015.

[15] D. S. K. B. Sanjay B.Patil, "Betal Leaf Area Measurement Using Image Processing," *International Journal on Computer Science and Engineering,* pp. 2656-2660, 2011.