Image Processing Technique for Gender Determination from Medical Microscope Image

Wichan Thumthong[1] Hathaichanok Chompoopuen[2] and Pita Jarupunphol[3]

[1] Faculty of Computer Science and Information Technology,
Rambhai Barni Rajabhat University, Chanthaburi, Thailand
[2] Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand
[3] Department of Digital Technology, Phuket Rajabhat University, Phuket, Thailand
wichan.t@rbru.ac.th

Abstract. This article proposes an image processing technique to calculate the histological variables from microscope image of occipital bones for gender determination. Occipital bones are highlighted based on 13 parameters to find out the relationship with the gender of samplings. The samples are 80 images classified into 46 males and 34 females with ages between 25 and 90 years old. Direct and stepwise discriminant functions are two methods for accuracy evaluation. The experiments show demonstrative results of gender determination accuracy based on both direct and stepwise discriminant functions. While the direct discrimination on 5 parameters represents 97.5% for both gender classification and gender prediction, the stepwise discrimination on 3 parameters is also at 97.5% for both gender classification and gender prediction.

Keywords: Forensic science, Gender determination, Image processing

1 Introduction

Biological identification consists of four basic data types, including gender, age-at-death estimation, height and nationality. Since biological identification plays a crucial role in the individual specification, gender determination has become a vital stage in forensic science [1]. The most reliable non-metric based research considers gross morphological features of the pelvis [2] [3]. Initially, a traditional approach for understanding sexually dimorphic features of the pelvis [2] cranium [4] was dependent upon subjective visual assessment. In Thailand, several gender determination studies have applied the gross morphological analysis to determine sex, such as the lumbar vertebrae [5], metacarpals [6] and talus [7]. Different sex estimation metrics were conducted on various skeletal elements, e.g., scapula [8], sternum and 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. In many cases, skeletonremains are unable to be identified by traditional gross morphology features to

P. Boonyopakorn et al. (Eds.): IC2IT 2019, AISC 936, pp. 1–9, 2020.
https://doi.org/10.1007/978-3-030-19861-9_13

estimate a biological profile when they are fragmented, poorly preserved, and incomplete. 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, image processing has been applied in many pathological projects. MATLAB have also become an essential image processing tool in terms of 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]. In the meantime, bio-medical image processing is also considered as the most challenging and upcoming field in the present world [11]. In this research, the image processing technique will be used to calculate the histological parameters from microscope image for automatic gender determination from occipital bones.

2 Research and Methodology

2.1 Graphical Method

The samples consist of 80 images divided into 46 males and 34 females of age between 25 and 90 years with identified age, sex and the cause of death. All the samples were delivered to the Forensic Medicine Department for routine autopsy. 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 digital images of bones were observed under light microscope at 10x magnification and formatted into JPG with 1600 × 1200 pixels.

2.2 Image Processing Method

The image processing technique and the algorithm source code for analyzing histological parameters are written using MATLAB 2016a, which can be described into 3 stages below.

Pre-Processing Stage. The original image of occipital bone is given as the input. The image parameters are separated into four sections with the area of intact osteon and fragmented osteon. While the position of intact osteon is drawn and replaced with red, the fragmented osteon position is replaced with green. Then, the original image background-color is replaced with black color and converted to grayscale.

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

Threshold Segmentation Stage. The pixels are divided according to the intensity value. This method is based on the threshold value in which the gray scale image is converted to the binary image. Local methods are deployed to transform the threshold value of each pixel to the local image characteristics for segmentation [11] as shown in Fig.1.

Post-Processing Stage. The regionprops is used to measure the image selected region in pixel. The regionprops instruction is used to estimate the scaled area, which is the actual number of pixels in the selected region [14]. The pixel count of the proposed image is dependent upon the distance between the camera and the object when the picture is taken. A reference object is an object with known area required to translate the pixel count area [14].

In this study, the photograph is formatted in a digital JPG image at 1,600 pixels width and the object of femur bone width is 1406.36 µm. Therefore, the pixel value calculation is based on the following.

$$1 \text{ pixel value} = \text{width object/width image; } 1 \text{ pixel value} = 0.879\ \mu m^2; \quad (1)$$

Please note that the area of parameters (μm^2) will be X * 1 pixel if the pixel count area is X pixels. Post-processing stage from the original image is represented in Fig.1.

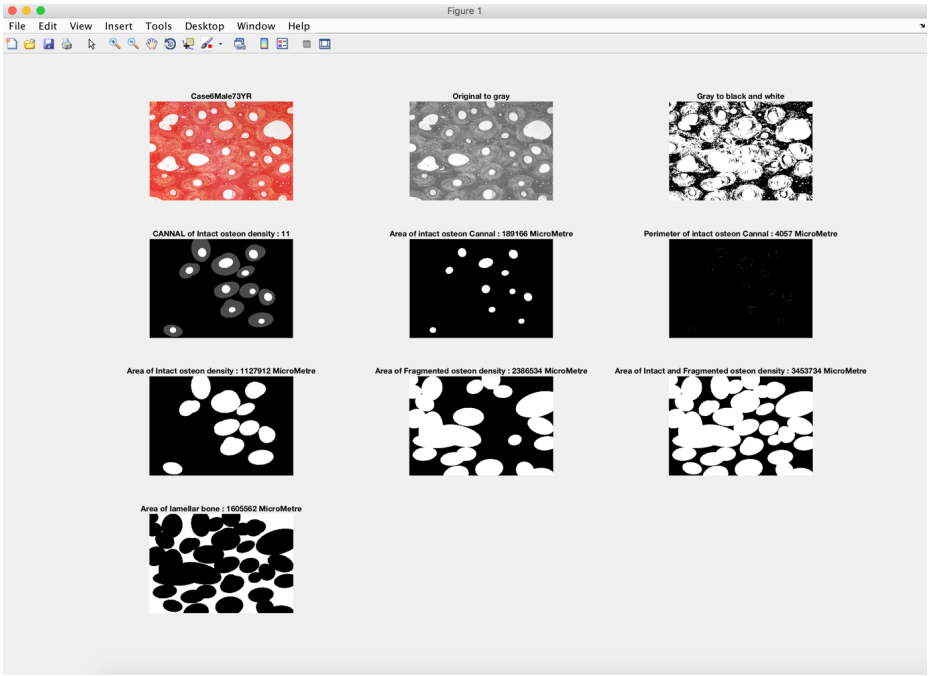

Fig. 1. Results of post-processing stage.

2.3 Histological Parameters

Thirteen histomorphometric parameters with their abbreviations are clarified in Table 1.

Table 1. Abbreviations of histomorphometric parameters.

No.	Histomorphometric Parameters	Abbreviations
1	Secondary osteon area	On.Ar
2	Secondary osteon maximum diameter	On.max
3	Secondary osteon minimum diameter	On.min
4	Perimeter of big secondary osteon	Pm.big.on
5	Big secondary osteon maximum diameter	Big.on.max
6	Big secondary osteon minimum diameter	Big.on.min
7	Perimeter of small secondary osteon	Pm.small.on
8	Small secondary osteon maximum diameter	Small.on.max
9	Small secondary osteon minimum diameter	Small.on.min
10	Haversian canal area	Hc.Ar
11	Haversian canal max diameter	Hc.max
12	Haversian canal min diameter	Hc.min
13	peTrimeter of haversian canal	Pm.Hc

Table 2 represents the histomorphometric parameters calculated using descriptive statistics, such as mean, minimum, maximum, and standard deviation.

Table 2. 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 (μm^2), μm=micrometers

2.4 Statistical Analysis

Descriptive statistics used to summarize and calculate indices of both males and females are illustrated in Table 3. The results show that 6 parameters are significantly different (P< 0.05) between males and females. However, 7 parameters, including On.Min, Big.on.max, Big.on.min, Pm.small.on, Small.on.max, Small.on.min, and Pm.Hc, are not considered. The results show that there is no statistical significance (P> 0.05) between males and females.

Table 3. Comparation of occipital histological parameters between female and male.

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

n: number of samples, [a] Log 10, * significant (P<0.05),
[NS] insignificant (P>0.05)

Direct and stepwise discriminant functions are two evaluation criteria for gender determination accuracy. The discriminant analysis shows that the gender determination accuracy using the discriminant function in the direct method and the classification function formulae for all variables are clarified in Table 4. The results indicate that parameter 1 is a prominent indicator for genders determination. Please noted that Hc.Ar alone is capable of classifying the gender at 98.8 % for both original and cross validated groups. In this case, the accuracy obtained by using direct variable, especially the Hc.Ar variable, can be used for gender determination from histomorphometry. Table 4 illustrates the moderate accuracy of gender classification and prediction observed from 5 variables, followed by On.Ar, On.Max, Pm.big.on, Hc.Max and Hc.min, respectively. For the stepwise discriminant function analysis, however, there are three 3 selected variables, including On.Ar, Pm.big.on and Hc.min, from the total of 5 (also see Table 4).

Table 4. 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

3.1 Results

This study performed 2 methods, as follows.

The results of direct discriminant analysis in Table 5 based on 5 variables, such as On.Ar, On.max, Pm.big.on, Hc.max and Hc.min, show moderate correlations. The discriminant score is calculated by multiplying the unstandardized coefficient with each particular measurement, summarising them, and adding the constant below.

$$\begin{aligned} \textit{Discriminant score} = {} & (6.775 \times \textit{On.Ar}) + (0.758 \times \textit{On.Ar.max}) + \\ & (0.009 \times \textit{Pm.big.on}) + (-1.286 \times \textit{Hc.max}) + \\ & (-8.636 \times \textit{Hc.min}) + (-9.750) \end{aligned} \tag{2}$$

Table 5 also shows 3 selected variables from 5 variables in discriminant function, including On.Ar, Pm.big.on, and Hc.min. The discriminant score was calculated as:

$$\begin{aligned} \textit{Discriminant score} = {} & (6.926 \times \textit{On.Ar}) + (0.009 \times \textit{Pm.big.on}) + \\ & (-9.476 \times \textit{Hc.min}) + (-8.823) \end{aligned} \tag{3}$$

After the discriminant score has been calculated from direct and stepwise discriminant functions, the score is compared with the specified criteria: 1) 0 (a halfway between the female and male centroids); 2) greater than 0 (male); and 3) less than 0 (female).

Table 5. Coefficients for the stepwise and direct discriminant function analysis.

Parameters	Unstandardized Coefficient	Eigen-value	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	0
2. On.max	0.758				F=2.266	
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	0
2. Pm.big.on	0.009				F= 2.245	
3. Hc.min	-9.476					
Constant	-8.823					

An eigenvalue indicates the proportion of variance calculated from between-groups sums of squares divided by within-groups sums of squares. A large eigenvalue is associated with a strong function. Based on the eigenvalue in Table 5, each of the eigenvalue of direct and stepwise methods are 4.309 and 4.230 respectively. The eigen-value must not be less than 1 to confirm the discriminant function accuracy. Please also note that the greater different values between the groups will imply higher discriminative possibility.

The canonical relation is a correlation between the discriminant scores and the dependent variable levels. A higher correlation indicates a better discriminative function. Canonical correlation is the value representing the relation between independent variables and dependent variables. The canonical correlation value from 0.8 to 1.0 is considered high. Consequently, the analyzed values of this study are 0.901 and 0.899, which are very 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. Table 4 show the results of direct and stepwise methods, which are 0.188 and 0.191 respectively. This means that the selected independent variables in the discriminant function of two methods are divided into groups. Furthermore, the value of Wilks' Lambda, which has statistical significance (P-value < 0.005), indicates that the group means are varied.

3.2 Accuracy of Classification

The accuracy results for gender determination of the discriminant function derived from 6 variables in the direct method are illustrated in Table 6. The accuracy results for both classification and prediction are 97.5 %. In the meantime, the results of those in the stepwise method derived from 3 variables for both classification and prediction are also 97.5%.

Table 6. Classification and prediction accuracy using direct and stepwise discriminant models.

Parameters	Classification Accuracy			Prediction Accuracy		
	Male	Female	Total	Male	Female	Total
Direct	100.0	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

The accuracy results for gender determination based on the discriminant function derived from 5 variables in the direct method are 97.5 % for both classification and prediction. When the stepwise is applied on 3 variables, the promising results were also represented in similar manner to the direct method. The accuracy at 97.5% for classification and prediction has been witnessed. In this case, the techniques used for this study can be considered as very effective and suitable for forensic science since the accuracy of gender classification and prediction can be observed at 97.5%.

4 Conclusions

An effective approach for gender determination from medical microscope image using image processing and its methodological steps have been proposed in this article. While direct and stepwise discriminant functions are two key approaches underlying the determination accuracy criteria, the experimental results on both functions are demonstrative. However, further experiments on this proposed technique must continue on different medical microscope images to yield the results validity.

References

1. Angi M. Christensen, Nicholas V. Passalacqua, Eric J. Bartelink. Forensic Anthropology: Current Methods and Practice. 1st edition, Academic Press. (2014).
2. Phenice TW. A newly developed visual method of sexing the Os pubis. Amercan Journal of Physical Anthropology, 30(2), pp. 297-301. (1969).
3. Bruzek J. A method for visual determination of sex, using the human hip bone. Amercan Journal of Physical Anthropology, 117(2), pp. 157-68. (2002).
4. Jane E. Buikstra, Douglas H. Ubelaker. Standards for Data Collection from Human Skeletal Remain. Proceedings of a seminar at The Field Museum of Natural History, Arkansas Archeological Survey,12154th edition. (1994).
5. Kelly Ostrofsky, Steven E. Churchill. Sex Determination by Discriminant Function Analysis of Lumbar Vertebrae. Journal of Forensic Sciences, 60(1). (2014).
6. Pongsak Khanpetcha, Sukon Prasitwattansereeb, D. Troy Casec, Pasuk Mahakkanukrauh. Determination of sex from the metacarpals in a Thai population. Forensic science international, 217(1-3):229.e1-8. (2011).
7. Pasuk Mahakkanukrauha, Sithee Praneatpolgranga, Sitthiporn Ruengdita, Phruksachat Singsuwana, Phuwadon Duangtob, D. Troy Case. Sex estimation from the talus in a Thai population. Forensic science international, 240(152). (2014).

8. Gretchen Dabbs, Peer H. Moore-Jansen. A Method for Estimating Sex Using Metric Analysis of the Scapul. Journal of Forensic Sciences, 55(1), pp. 149-52. (2009).
9. Pasuk Mahakkanukrauh, Sukon Prasitwattanseree, D. Troy Casec. Determination of sex from the proximal hand phalanges in a Thai population. Forensic science international, 226(1-3). (2013).
10. Christian Crowder, Sam Stout. Bone Histology An Anthropological Perspective. CRC Press; 1st edition. (2011).
11. D. Dhilip Kumar, S. Vandhana, K. Sakthi Priya, S. Jeneeth Subashini. Brain Tumour Image Segmentation using MATLAB. International Journal for Innovative Research in Science & Technology, 1(12), pp. 447-451. (2015).
12. Rajesh C. Patil, Dr. A. S. Bhalchandra. Brain Tumour Extraction from MRI Images Using MATLAB. International Journal of Electronics, Communication & Soft Computing Science and Engineering, 2(1). (2015).
13. G. Nithyanandam. Extraction of Cancer Cells from MRI Prostate Image Using MATLAB. International Journal of Engineering Science and Innovative Technology, 1(1), pp. 27-35. (2012).
14. Sanjay B. Patil, Dr. Shrikant K. Bodhe. Betal Leaf Area Measurement Using Image Processing. International Journal on Computer Science and Engineering, 3(7), pp. 2656-2660. (2011).