Sex Determination of Skull Based on Fuzzy Decision Tree

Yang Wen[1, a], Reziwanguli Xiamixiding[2, b], Xu Jiachen[1, c], Wang Piao[1, d], Hu Jiabei[1, e], Liu Xiaoning[1, f]

[1]College of information science and technology, Northwestern University, Xi'an 710127, China

[2]College of computer science and technology, Xinjiang Normal University, Urumqi 830054, China

[a]920759301@qq.com,
[b]2367757800@qq.com, [c]443987531@qq.com, [d]1603884625@qq.com, [e]m18502967030@163.com, [f]nwu_yw @126.com

Abstract: The skeleton plays an important role in sex determination in forensic anthropology. The skull bone is considered as the second best after the pelvic bone in sex determination due to its better retention of morphological features. Different populations have varying skeletal characteristics, making population specific analysis for sex determination essential. However, most previous studies have used the traditional discriminant function method, which is highly dependent on the population, and does not consider the uncertainty in the process of sex identification. In view of the above problems, this paper proposes a method of sex identification based on fuzzy decision tree. The proposed method is to improve the uncertainty, by constructing the fuzzy function of statistical information based on the cranial measurement item, as the basis of the fuzzy decision process, the classification factor is determined at the end of the training process, and a fuzzy decision binary tree is established. The experimental results show that the proposed method is effective and can provide reference for the practical application of gender discrimination.

1. Introduction

The skeleton plays an important role in sex determination in the field of forensic anthropology. Though, entire skeleton is essential for accurate results, it is rarely available completely and in good condition. Therefore, it is important to apply appropriate methods for gender determination from skeletal remnants [1,2]. Previous studies [3-6] indicate that pelvis is the most reliable indicator of sex assessment, and skull is the second one. However, not all forensic cases provide a complete skeleton due to breakage or postmortem destruction, while the skull can be well preserved in most cases since it is composed of hard tissue. So the skull is the most commonly used skeleton part in forensic anthropological analysis. Morphology discriminant method and measurement discriminant method are the two main techniques employed by forensic anthropolgists for sex determination [7]. The earliest studies using morphological assessment were reported in [8-10]. The morphological method is greatly influenced by the observer and needs a certain degree of professional experience. At present, the method of sex discrimination based on skull is still mainly used in measurement discrimination method [11-15]. Compared with the morphological discriminant method, the measurement discrimination method is more objective, but the accuracy rate is not enough, because there is a overlap between the male value and the female value in many measurement indexes, which makes the uncertainty in the sex identification process.

In order to solve the uncertainty in the process of skull sex identification, we apply the fuzzy method to gender identification, and propose a method of skull sex identification based on fuzzy decision tree. Combining the decisional impact of more measurements for sex determination through the construction of a Fuzzy Decision Tree with Fuzzy Sets taylored for each measurement based on the statistical distribution of measurement values from the raw data set on each sex class.

The distribution of skull measurements from can be used to devise fuzzy functions in order to model the measurement's appartenance to either the male or female class and such an approach will

take better advantage of the uncertain values than a classical crisp model. Since a Fuzzy Set is used to describe the membership degree of an instance to a given class and a very popular Fuzzy Set model follows a Gaussian curve to model such membership, Fuzzy Sets can be efortlessly generated for each skull measurement when provided with the mean and standard deviation values for either the male and female classes.

This article is arranged as follows: the second section is about the materials and the methods; the third part is about the discussion; the last part is conclusion.

2. Materials and Methods

2.1 Materials

It is carried out on a database of 267 whole skull CT scans (153 females and 114 males) on voluntary persons that mostly come from Uighur ethnic group in North of China, age 18-88 years for females and 20-84 years for males. The images of each subject are restored in DICOM format with a size of approximately $512 \times 512 \times 250$. Each 3D skull surface is extracted from the CT images and represented as a triangle mesh including about 220,000 vertices. All the skulls are substantially complete; that is, each skull contains all the bones from calvarias to jaw and has the full mouth of teeth. In this paper, all skull samples are unified and normalized using reference paper [16].

2.2 Fuzzy Decision Tree Methods

2.2.1 Fuzzy Set Construction

Sex determination is better conducted within a given population and therefore a data set for this problem will contain only information characterizing the members of that population. The vaster the data set, the better chances of exploiting the patterns inside it for sex determination. Fuzzy Functions for sex class (male or female) membership evaluation can be constructed based either on the statistical study of each of the attributes in the data set that usually reveals a normal trend of the data and provides values for the mean and standard deviation for each sex. In the vast majority of data sets providing skull measurement information on archaeological or forensic data, the male and female classes are characterized by Gaussian shaped normal membership functions that can be parameterized entirely using the mean and standard deviation values. On the other hand, when working on a population that was previously published, the Fuzzy Functions may be constructed based on the statistical values for mean and standard deviation published in the literature for that population.

2.2.2 Fuzzy ID3 Approach for Skull Sex Determination

Previous studies [17] have applied decision trees to bone sex identification and achieved good results. However, due to the uncertainty of some measurement values of the skeleton, there will be misjudgement. If a wrong categorical crisp decision was to be replaced with a decision based on fuzzy sets that copes better with uncertainty the impact of a choice on the problematic measurement will not be as drastic with respect to the overall final classification. With this idea in mind, the proposed Fuzzy ID3 approach was elaborated always taking into account the requirements of the Skull Sex Determination problem. Conceptually,the proposed decision tree variant follows the classical ID3 algorithm but transforms the operations used in construction so that they are based on the fuzzy membership values for each sex class. The steps of the algorithm are as follows.

ID3(Example,Target,Attributes)

- The Root node that creates the tree
- If Examples are all positive, return label =+ single-node tree Root
- If Examples are all negative, return label =+ single-node tree Root
- If Attributes is empty, then return the single-node Root, the most common Target_attribute value in label=Examples
- Otherwise start
    - The best attributes of the ability to classify examples in A← Attributes
    - Root's decision attribute ← A
    - For each possible value of A, v_i
        - Add a new branch to test A= v_i under Root
        - Let $Examples_{v_i}$ be a subset of Examples that satisfies A attribute value v_i
        - If $Examples_{v_i}$ is empty
            - Add a leaf node to the new branch, the most common Target_attribute value in the label=Examples of the node
            - Otherwise, add a subtree ID3 under the new branch($Examples_{v_i}$,Target_attribute,Attributes-{A})
- End
- Return Root

The proposed method follows the classical process used for classical decision trees. The main differences occur only when aspects of fuzziness and uncertainty are met. The algorithm is recursive. Like in the case of classic decision trees, the root node starts by having all the instances in the data set whose fuzzy membership function is initialized to 1 and the attribute set contains all the measurements in the initial data set.

For each measurement in the attribute set the fuzzy information gain is computed and the measurement yielding the best information gain is selected for the splitting of the data set characterizing the current node into the children data sets. Since there are two fuzzy functions to split on, each node will have two children, one for each sex class, thus leading to a binary fuzzy decision tree.

3. Result

3.1 The result of the measurement term for build fuzzy set

The sex of the adult preserved skull was initially determined by standard morphological traits. A total of 16 craniometric parameters were included in the study. All craniometric parameters are measured automatically by computer software. The parameters studied were orbital breadth (dacryon-ectoconchion), interorbital breadth (dacryon-dacryon), upper facial height (nasion-prosthion), Protruding of eyebrow, frontal chord (nasion-bregma), inter-mastoid width, nasal breadth (alare-alare), frontal arc (nasion-bregma), occipital protruding angle, mandibular angle width, mandibular height, maximum cranial length (glabella-opisthocranion), maximum cranial breadth (euryon-euryon), basion-bregma height (basion-bregma), cranial base length (basion-nasion), maximum cranial circumference.

All the measurements were taken to the nearest millimeter and were converted to centimeter and the range of angles is [0,180]. In order to reduce the measurement error, all skulls are measured three times, and then averaged as the final measurement data. The result of the measurement is shown in Table 1.

Table 1 The mean and standard deviation of the measurement term

index	parameters	male		female	
		Mean	SD	Mean	SD
X_1	Orbital breadth	98.7	3.1	95.2	3.5
X_2	Inter-orbital breadth	24.9	2.69	23.8	2.1
X_3	Upper facial height	106.3	5.1	103.8	4.6
X_4	Protruding of eyebrow	32.3	8.9	29.4	5.6
X_5	Frontal chord	113.5	4.9	110.1	5.6
X_6	Inter-mastoid width	110.9	5.9	107.9	4.9
X_7	Nasal breadth	29.9	2.7	29.4	8.9
X_8	Frontal arc	26.5	2.1	25.7	1.7
X_9	Occipital protruding angle	167.0	11.1	177.8	3.1
X_{10}	Mandibular angle width	102.3	6.4	97.7	5.8
X_{11}	Mandibular height	45.7	3.9	43.6	3.3
X_{12}	Maximum cranial length	178.1	7.4	168.6	5.7
X_{13}	Maximum cranial breadth	138.7	6.3	136.4	5.5
X_{14}	Basion-bregMa height	13.4	3.6	12.9	2.9
X_{15}	Cranial base length	9.6	2.6	9.2	3.1
X_{16}	Maximum cranial circumference	52.3	4.4	50.2	4.1

Notes: The unit of measurement for distance is mm; the angle range is [0,180].

The mean and standard deviation of each measurement are given in the table. These values are used to construct a male fuzzy set and a female fuzzy set for each measurement item. The fuzzy decision tree used for gender identification is a binary tree.

3.2 Results of sex identification

The 267 skulls of Xinjiang Uygur in China collected by the Institute of Visualization Technology of Northwest University were used as experimental subjects, 70% of which were used as training samples, and the remaining samples were used as test samples, and verified by using a leave-one-cross validation strategy.

Based on the measurement results of the 3.1 section, the fuzzy decision tree is constructed by using the method in this paper and the method in the literature [17], and then the discriminant function is established by the method of literature [18]. Three classification models are tested respectively to compare their performance. The results are shown in Table 2.

Table 2 Results of sex identification

	This paper method	Literature [18] method	Literature [19] method
Measurement term	**89.4%**	87.6%	86.3%

In order to verify the universality of the method, we have also verified on two public datasets: the Coimbra Identified Skeletons Collection (University of Coimbra, Portugal), composed of Portuguese individuals deceased during the first half of the twentieth century [19, 20], and the Thai Skeletal Collection hosted by the Department of Anatomy of the University of Chiang-Mai (Thailand), composed of Thai individuals deceased at the end of the twentieth century [21]. The method of this paper, the method of document [17] and the method of document [18] are applied to two datasets respectively. The results of sex identification are shown in Table 3.

Table 3 Results of sex identification for different data sets

	This paper method	Literature [18] method	Literature [19] method
Coimbra Skulls Collection	**88.7%**	86.5%	85.8%
Thai Skulls Collection	**89.0%**	88.3%	86.2%

4. Discussion

On the Chinese Uygur data set, the accuracy rate of the fuzzy decision tree proposed in this paper is 89.4%, and the classification performance is better than the other two methods. The method proposed in this paper is also optimal when validates on two public data sets. The results show that: the fuzzy decision tree approach adapt better to classifying the uncertain cases by not committing to only one of two branches at each split point in the decision tree but by selecting each option to different extents and drawing the discriminatory conclusion at the end as it was present in the previous sections.

The sex determination process was tackled in various forms in the literature, most of the approaches treating the matter in a traditional mathematical manner rather than providing computational solutions. While usually performant, such methods tend to overfit with respect to the data set they were constructed on, a case that the fuzzy decision tree approach tries to avoid. Although the three methods are based on cranial measurements, the other two methods [17,18], do not consider the overlap and uncertainty, and this method is specially considered, so their data sets are the same as the measurement items, but the results are different.

Other researchers have also used the decision tree method to skull sex determination [22,23]. The data sets are different from us and can't be compared accurately, but if we only pay attention to the accuracy we get, our method based on fuzzy decision tree is better than their method.

In view of the sex identification problem of skull, this paper presents an effective solution, especially the problem of gender identification with uncertainty. In recent years, some computer vision and deep learning methods have been applied to bone sex research [24-27]. Therefore, the next step is to combine traditional methods with these methods, or to improve traditional methods with a new idea.

5. Conclusion

Fuzzy Decision Trees prove to be good sex classifiers when skull measurements are given, provided accurate fuzzy membership functions can be provided for each of the sex classes. The drawbacks of such an approach occur when the distributions of the sex classes are very close together and thus the uncertain area that enables the fuzzy decision process is reduced. Also, the classifier is efficient only when provided with sufficient, preferably balanced, training data. Even if in its current status the fuzzy decision tree approach achieves significant accuracy rates in practice on real archaeological and forensic data sets.

We are currently collecting a larger sample to establish a gender discriminant model based on the fuzzy decision tree, which will be used in the practical application of the unknown bone sex assessment in forensic cases. Also, we hope our study can contribute to further population data studies, or even to a craniometric meta-analysis on other ethnic groups.

Acknowledgments

REZIWANGULI Xia-mi-xi-ding, College of computer science and technology, Xinjiang Normal University.

References

[1] Naikmasur V G, Shrivastava R, Mutalik S. Determination of sex in South Indians and immigrant Tibetans from cephalometric analysis and discriminant functions[J]. Forensic Science International, 2010, 197(1):122.e1-122.e6.

[2] Gonzalez P N, Bernal V, Perez S I. Analysis of sexual dimorphism of craniofacial traits using geometric morphometric techniques [J]. International Journal of Osteoarchaeology, 2011, 21(1):82-91.

[3] Spradley M K, Jantz R L. Sex estimation in forensic anthropology: skull versus postcranial elements. [J]. Journal of Forensic Sciences, 2015, 56(2):289-296.

[4] Williams B A, Rogers T. Evaluating the accuracy and precision of cranial morphological traits for sex determination. [J]. Journal of Forensic Sciences, 2010, 51(4):729-735.

[5] Ubelaker D H, Volk C G. A test of the phenice method for the estimation of sex [J]. Journal of Forensic Sciences, 2002, 47(1):19-24.

[6] Pickering R B, Bachman D C. The use of forensic anthropology [M]. CRC Press, 2009.

[7] Franklin D, Cardini A, Flavel A, et al. The application of traditional and geometric morphometric analyses for forensic quantification of sexual dimorphism: preliminary investigations in a Western Australian population [J]. International Journal of Legal Medicine, 2012, 126(4):549-558.

[8] Konigsberg L W, Hens S M. Use of ordinal categorical variables in skeletal assessment of sex from the cranium [J]. American Journal of Physical Anthropology, 1998, 107(1):97-112.

[9] Walker P. Sexing skulls using discriminant function analysis of visually assessed traits [J]. American Journal of Physical Anthropology, 2010, 136(1):39-50.

[10] Iscan M Y, Steyn M. The Human Skeleton in Forensic Medicine [M]. Royal Society of Medicine Press, 2013.

[11] Devadas P, Janardhan R M, Yesender M, et al. STUDY ON THE MORPHOMETRIC ANALYSIS OF THE FORAMEN MAGNUM AS AN INDICATOR FOR SEX DETERMINATION[J]. International Journal of Anatomy & Research, 2017, 5(3.1):4163-4167.

[12] Butt S, Ahmed I. Reliability of Cephalogram in Determining Skull Gender Dimorphism [J]. 2016, In Press (In Press).

[13] Zaafrane M, Ben K M, Naccache I, et al. Sex determination of a Tunisian population by CT scan analysis of the skull[J]. International Journal of Legal Medicine, 2017:1-10.

[14] Shah T, Patel M N, Nath S, et al. Determination of sex using cephalo-facial dimensions by discriminant function and logistic regression equations [J]. Egyptian Journal of Forensic Sciences, 2016, 6(2):114-119.

[15] Kruger G C, Labbe E N, Stull K E, et al. Sexual dimorphism in cranial morphology among modern South Africans[J]. International Journal of Legal Medicine, 2015, 129(6):1285-1285.

[16] Liu XN, Zhu LP, Lu YN, et al. Hierarchical Skull Registration Method with a Bounded Rotation Angle[C]. Intelligent Computing Methodologies, 2017:561-571.

[17] Mircea IG, Czibula G, Petrusel MR. Sex identification in archaeological remains using decision tree learning [J]. Studia Universitatis Babes-Bolyai Series Informatica, 2015:91–103.

[18] Shui WY, Yin RC, Zhou MQ, et al. Sex determination from digital skull model for the Han

people in China [J]. Chinese Journal of Forensic Medicine, 2013, 28(6):461-464.

[19] Rocha M A. Les collections ostéologiques humaines identifiées du Musée Anthropologique de l'Université de Coimbra [J]. Antropologia Portuguesa, 1995, 13:7-38.

[20] Wasterlain S N, Hillson S, Cunha E. Dental caries in a Portuguese identified skeletal sample from the late 19th and early 20th centuries.[J]. American Journal of Physical Anthropology, 2010, 140(1):64-79.

[21] Mahakkanukrauh P, Khanpetch P, Prasitwattanseree S, et al. Stature estimation from long bone lengths in a Thai population.[J]. Forensic Science International, 2011, 210(1):279.e1-279.e7.

[22] Kumar A, Hanmandlu M, Gupta H M. Fuzzy binary decision tree for biometric based personal authentication [J]. Neurocomputing, 2013, 99(1):87-97.

[23] Shirley R. Document Title: AGE AND SEX ESTIMATION FROM THE HUMAN CLAVICLE: AN INVESTIGATION OF TRADITIONAL AND NOVEL METHODS [J]. 2009.

[24] Miholca D L, Czibula G, Mircea I G, et al. Machine Learning Based Approaches for Sex Identification in Bioarchaeology[C]// International Symposium on Symbolic and Numeric Algorithms for Scientific Computing. IEEE, 2017:311-314.

[25] Arigbabu O A, Liao I Y, Abdullah N, et al. Computer vision methods for cranial sex estimation [J]. IPSJ Transactions on Computer Vision & Applications, 2017, 9(1):19.

[26] Pinto S C D, Urbanova P, Cesar R M. Two-Dimensional Wavelet Analysis of Supraorbital Margins of the Human Skull for Characterizing Sexual Dimorphism[J]. IEEE Transactions on Information Forensics & Security, 2016, 11(7):1542-1548.

[27] Afrianty I, Nasien D, Kadir M R A, et al. Back-Propagation neural network for gender determination in forensic anthropology[C]// Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2015:1524-1527.