Presented at the ISG 2006, Shah Alam, Malaysia
19-21 September 2006

STATISTICAL SHAPE ANALYSIS USING LANDMARK FOR CRANIOFACIAL CLASSIFICATION

Mohd Bakery Md Hussin, Halim Setan, Zulkepli Majid & Deni Suwardhi

Medical Imaging Research Group
Faculty of Geoinformation Science and Engineering
Universiti Teknologi Malaysia
81310 UTM Skudai
Johor, Malaysia
+607-5530380
bakery@fksg.utm.my

KEYWORDS: Shape Analysis, Generalized Procrustes Analysis, Classification

Abstract

This paper describes craniofacial medical application, with the emphasis on the craniofacial classification by using statistical shape analysis method to classify the craniofacial datasets. The craniofacial datasets are classified by gender of the datasets. Craniofacial classification involves the construction and description of the craniofacial dataset classes, and then craniofacial classification involves assigning the craniofacial dataset into the known classes. The data consists of 3D laser scans of normal human faces with respect to the normal craniofacial dataset. Craniofacial datasets using 8 landmarks with 3D coordinate on the facial surface. The craniofacial landmarks coordinate are used to analyze the datasets, and classify the datasets into the classes that has created.

INTRODUCTION

In this study, the classification of the craniofacial data uses the facial landmark data. There have 35 craniofacial data, which have 8 landmarks for every data. Landmark is a point of correspondence on each object that matches between and within populations. There are three basic types of landmarks are generally used; anatomical landmarks, geometrical landmarks and pseudo-landmarks. In this study, we concentrate on the geometrical landmarks. Geometrical landmarks are points located on an object according to some mathematical or geometrical property of the object.

Classification is the activity of allocating individuals to one of a set of existing groups. Classification involves assigning an observation to one of a set of previously known classes. The classification of the craniofacial data is based on shape of every craniofacial data. The shape of craniofacial data is created from the coordinate of the craniofacial geometrical landmark data. The geometrical landmarks, also called mathematical landmarks are points on the object according to some geometrical or mathematical property.

The 35 craniofacial data is used to classify the gender of the craniofacial data based on the shape. The coordinate of the 35 craniofacial data is taken by using the 3D laser scans of normal human faces. All the coordinates of the craniofacial data must be free from the scale, rotation and location factor because the shape must be free from the three factors. The shape is all the geometrical information that free from scale, rotation and location factors.

To filter out the three factors of shape, generalized Procrustes analysis is used. Procrustes analysis is used to compare the shape of datasets. The generalized Procrustes analysis being used when there are more than two datasets involve in the analysis. Generalized Procrustes method will estimate the average of shape and filtered out all the three factors. This method involves the translating, rescaling and rotating the configurations relative to each other. This procedure will minimize the total of sum of squares.

The classification method that used is minimum distance, one of the supervised classification methods in principal component analysis. There are two types of classification; supervised classification and unsupervised classification. The datasets will classify into the gender class based on the shape of the datasets.

Principal component analysis (PCA) is a way of identifying patterns in data and expressing the data in such a way as to highlight their similarities and differences.

METHOD

The data consists of 3D scans taken from 35 students, which 18 students are male and 17 students are female. The ages of the datasets are 15 and 17 years old. There are 8 landmarks on each datasets, as show in the Table 1.

Table 1: Landmark, name and landmark definition

No	Landmark	Name	Definition
1	en-r	Endocanthion (right)	The inner corner of the eye fissure where the eyelids meet, not the caruncles (the red eminences at the medial angles of the eyes) (right).
2	en-l	Endocanthion (left)	The inner corner of the eye fissure where the eyelids meet, not the caruncles (the red eminences at the medial angles of the eyes) (left).
3	ex-r	Exocanthion (right)	The outer corner of the eye fissure where the eyelids meet (right).
4	ex-l	Exocanthion (left)	The outer corner of the eye fissure where the eyelids meet (left).
5	ac-r	Alar curvature point (right)	The most posterolateral point of the curvature of the base of the nasal alae, the lateral flaring walls of the nostrils (right)
6	ac-l	Alar curvature point (left)	The most posterolateral point of the curvature of the base of the nasal alae, the lateral flaring walls of the nostrils (left)
7	ch-r	Cheilion (right)	The outer corner of the mouth where the outer edges of the upper and lower vermilions meet (right)
8	ch-l	Cheilion (left)	The outer corner of the mouth where the outer edges of the upper and lower vermilions meet (left)

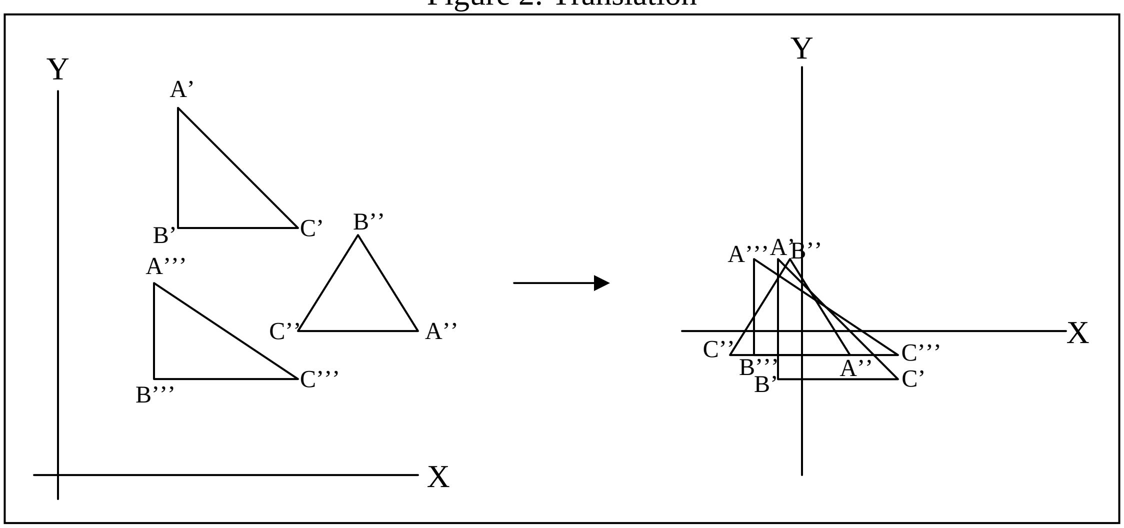
In the Table 1 above, show the 8 landmarks one of the datasets. The point of the landmarks is picked digitally by using the RapidForm software. Every landmark has the 3 dimensional coordinate (x y z). The coordinate origin of the raw data is on the lens of the laser.

Figure 1: Datasets image with 8 landmarks

Figure 1 above show the image one of the craniofacial datasets with 8 geometrical landmarks that used for classification analysis. The geometrical landmarks are located by using the RapidForm software. The coordinate of the geometrical landmarks is getting from the RapidForm software. We need to normalize the geometrical landmark coordinate to create the shape of the craniofacial datasets.

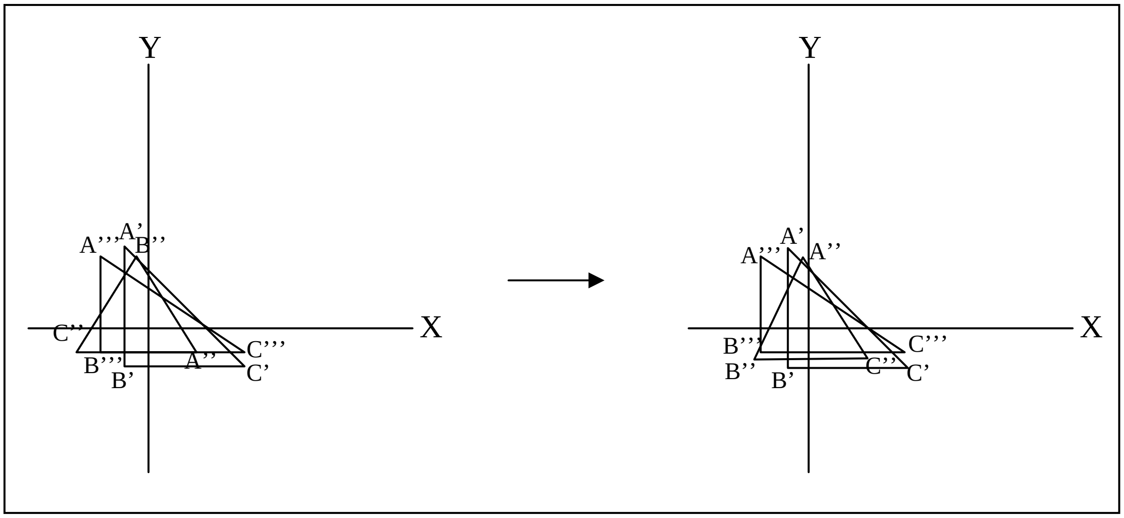
Normalization is needed to translate the datasets into the same coordinate system. Generalized Procrustes analysis is used to normalize the datasets. Generalized Procrustes analysis will filter out the scale, rotation and location of the datasets to create the shape of craniofacial datasets. The configuration is the set of landmarks on a particular dataset and the configuration matrix X is the k×m matrix of Cartesian coordinates of the k landmarks in m dimensions (m = 3).

Figure 2: Translation



Translation is obtained by adding a constant to each entry in X. Translation will use the centre of the configuration to remove the location factor. All the datasets will translate to the origin of the coordinate system based on centroid point. After translation, all the centroid points will be at the origin of the coordinate system.

Figure 3: Rotation



Rotation of the configuration will rotate a matrix to maximum similarity with a target matrix minimizing sum of squared differences. Procrustes rotation is a general term for the mathematical technique to match the datasets. All the data sets is rotated until the distance between A’, A’’ and A’’’, between B’, B’’ and B’’’, and between C’, C’’ and C’’’ are minimum. The datasets are rotated based on the centroid point.

Scaling is applied to remove the variation of size. Scaling will not affect the shape of the datasets and will not remove the allometric effects of size on shape. The datasets is scaled also based on the centroid point.

After removing the location, rotation and scaling factor, the shape of the craniofacial datasets is created. Principal Component Analysis is used to classify the shape of the craniofacial data. The craniofacial shape datasets are in 3 dimensional matrix configurations. The datasets of craniofacial shape data is used for data classification analysis by using principal component analysis.

The datasets is classified into groups of gender; male and female. The scatter plot of the classification result would be in two groups, male and female.

RESULT

Figure 1: Principal component scatter plot between first principal component and second principal component

Figure 2: Principal component scatter plot between first principal component and third principal component

Figure 3: Principal component scatter plot between first principal component and fourth principal component

Figure 4: Principal component scatter plot between first principal component and fifth principal component

Figure 1, 2, 3 and 4 show the result of the craniofacial data classification. The blue scatter plot represents the male craniofacial datasets, and the red scatter plot represents the female craniofacial datasets. From the scatter plot, we can see the difference of the scatter plot between male and female datasets.

CONCLUSION

In this paper, the datasets of craniofacial data is classified to gender class based on the shape of datasets. The classes of the datasets can be used to classify the new datasets either the new datasets are male or female by using the minimum distance method. The plot of the new datasets will be classify either male or female dataset based on the minimum distance between the new dataset and the scatter plot of classified datasets. If the new dataset is near to male classified datasets, so the new datasets is classified as a male craniofacial dataset.

ACKNOWLEDGEMENTS

This research is part of a Prioritized Research (PR) IRPA Vote 74537 sponsored by Ministry of Science, Technology and Innovation Malaysia (MOSTI). The PR involves Universiti Teknologi Malaysia (UTM), Standard & Industrial Research Industrial Malaysia (SIRIM), and Universiti Sains Malaysia (USM).

REFERENCE

Akca, M. D. (2003). Generalized Procrustes Analysis and Its Application in Photogrammetry. Swiss Federal Institut of Technology.

Fred L. Bookstein (1991). Morphometric Tools for Landmark Data. New York: Camridge University Press.

Ian L. Dryden & Kanti V. Mardia (1998). Statistical Shape Analysis. University of Leeds, UK: John Wiley & Sons Ltd.

John Shlens (2003). A Tutorial on Principal Component Analysis: Derivation, Discussion and Singular Value Decomposition. University of California, San Diego.

Lele, S. R. & Richtsmeier, J. T. (2001). An Invariant Approach to Statistical Analysis of Shapes; Interdisciplinary Statistics. Florida, USA: CRC Press.

Leslie G. Farkas (1994). Anthropometry of the Head and Face. New York: Raven Press Ltd.

Shlens, J. (2003). A Tutorial on Principal Component Analysis.

Smith, L. I. (2002). A Tutorial on Principal Component Analysis.