Journal of the Korean Statistical Society 44 (2015) 477–489

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

journal homepage: www.elsevier.com/locate/jkss

Finding standard dental arch forms from a nationwide standard occlusion study using a Gaussian functional mixture model

Kyeong Eun Lee[a], Johan Lim[b], Joong-Ho Won[b], Sungim Lee[c,*], Shin-Jae Lee[d]

[a] Department of Statistics, Kyungpook National University, Republic of Korea
[b] Department of Statistics, Seoul National University, Republic of Korea
[c] Department of Applied Statistics, Dankook University, Republic of Korea
[d] School of Dentistry, Seoul National University, Republic of Korea

ARTICLE INFO

Article history:
Received 8 September 2014
Accepted 21 April 2015
Available online 13 May 2015

AMS 2000 subject classifications:
primary 62H30
secondary 62P10

Keywords:
Dental arch form
Functional clustering
Korean Standard Occlusion Study
Integrated M-spline
Mixture model

ABSTRACT

Orthodontists are interested in finding a set of standard arch forms for clinical orthodontic practice. In this paper, we propose a functional clustering method for the dental arches based on a mixture of U-shaped curves. We decide the number of clusters (equivalently, mixture components) using the Bayesian information criterion and the jump criterion based on a given distortion function. We apply our method to clustering the dental arch data from the nationwide standard occlusion study conducted in Korea from 1997 to 2005. The data are composed of dental arches of 306 subjects with normal occlusion selected from 15,836 young adults. We also provide the comparison of the proposed method to other existing methods.

1. Introduction

The aim of this paper is to introduce a functional method for finding a few number of standard dental arch forms. Finding standard dental arch forms is essential when fabricating orthodontic archwires. In clinical orthodontic practice, applying an archwire that can minimize the distance from the archwire to each tooth is an important issue when straightening irregular teeth. It is also beneficial for the clinicians to be equipped with as small as possible number of standard archwires. If there are variable number of archwires, it interferes with the practical use, i.e., clinicians selection, inventory management, and so on (Lee, Lee, Lim, Park, & Wheeler, 2011; Lim, Lee, Park, Lee, & Lee, 2014).

Orthodontic practice has relied on a single ideal standard template to which the teeth of most of the patients with malocclusion are aligned. The form of a template is predicated on the clinical experiences or the preferred forms of certain clinicians or archwire manufacturing companies. Researchers focus on describing and fitting *average* arch forms in a more accurate manner geometrically or mathematically (Mutinelli, Cozzani, Manfredi, Bee, & Siciliani, 2008; Noroozi, Nik, & Saeeda, 2001; Taner, Ciger, El, Germec, & Es, 2004; Valenzuela, Pardo, & Yezioro, 2002). However, recent studies (*e.g.*, Ronay, Miner, Will, & Arai, 2008, Trivino, Siqueira, & Scanavini, 2008), raise a question on the appropriateness of a single standard

* Corresponding author.
E-mail address: silee@dankook.ac.kr (S. Lee).

http://dx.doi.org/10.1016/j.jkss.2015.04.003

template. Indeed, some authors (Henrikson, Persson, & Thilander, 2001; Lee et al., 2011; Lim et al., 2014) argue that dental arch forms have significant anatomical variations even among the subjects with normal occlusion. Hence, identifying a small number of standard forms that capture most of the variation may significantly reduce the total duration of individual treatments, because less alignment will be required compared to alignment based on typical treatment, which will increase the effectiveness of the orthodontic practice.

Therefore, this study attempts to cluster the observed arches into a small number of representative groups. Arch form data are typically obtained at the locations of 14 pairs of the teeth on a two-dimensional plane as displayed in Fig. 1. To cluster the measurements, Gu, Shibata, Fujita, and Takada (2002) use the Lloyd algorithm that maps each arch to a point in the Euclidean space. It iteratively (i) estimates the standard arch form and (ii) minimizes the Euclidean distance of each arch to the estimated standard arch forms. Recently, Lee et al. (2011) revisit this problem and propose to use the partition around medoids (PAM) algorithm based on L_1-type distance between two arches. They interpolate an observed arch using a piecewise linear function and find the translation and the rotation that minimize the L_1-type distance between two interpolated arches. Their proposed L_1-type distance is designed to be invariant to the translation and the rotation by nature. In comparison, Gu et al. (2002) assume the locations of the teeth in an arch to be independent, whereas Lee et al. (2011) consider the functional relationship among those using linear interpolation.

In this paper, we propose an improved functional method to cluster dental arch forms that can ensure pragmatic clinical application. The arch of each individual is assumed to be from a mixture of Gaussian curves, each of which has a smooth symmetric U-shaped mean curve and correlated Gaussian errors. The mean curve is modeled as an integrated M-spline (I-spline) proposed by Ramsay (1988). The model is estimated using the expectation and maximization (EM) algorithm. The number of clusters (mixture components) is decided by the Bayesian information criterion (BIC) and the jump criterion (JC) (Sugar & James, 2003). We demonstrate the superiority of the proposed method over the L_1 method by Lee et al. (2011) by comparing them in two simulation scenarios: Scenario 1 is based on two quadratic functions, and scenario 2 is based on two I-spline basis functions. The comparison is evaluated in terms of several indexes including the mis-clustering error rate (MER), the Rand Index (RI), and the adjusted Rand index (ARI).

We apply the proposed method to clustering the arches of a total of 306 subjects with normal occlusion. The subjects are selected from 15,836 young Korean adults, recruited in a community dental health survey conducted from 1997 to 2005 in Korea. Results show that two clusters are the most appropriate from both the data analysis and the dental practice points of view. These estimated standard arch forms are then compared with those of Lee et al. (2011). Their study finds three clusters using the PAM method with a newly defined L_1-type distance and estimates the representative arch forms of each cluster by fitting a cubic spline function.

This paper is organized as follows. In Section 2, we introduce raw data and pre-processing steps to adjust uncontrolled error in measurements. In Section 3, we discuss the functional clustering method of the paper. In Section 3.1, we present the details of how to model the symmetric U-shaped mean curves using I-splines. We then specify the functional clustering model as a mixture of Gaussian curves. We further propose the EM procedure to estimate the model in Section 3.2. The method to find the number of clusters follows in Section 3.3. In Section 4, we numerically compare the proposed functional approach to the PAM clustering with L_1-type distance by Lee et al. (2011). In Section 5, we apply the proposed method to clustering the arch data with 306 subjects. We conclude the paper in Section 6.

2. Data and preprocessing

2.1. Data acquisition

A total of 306 subjects with normal occlusion are selected from 15,836 young Korean adults who responded to a community dental health survey conducted in Seoul, Korea from 1997 to 2005. They include 186 men and 120 women with mean age of 20.0 years (a range from 15.1 to 31 years). The selection criteria are: (1) class I molar and canine relationship with normal occlusal interdigitation, (2) complete permanent dentition erupted except for the third molars, (3) normal overjet and overbite (about 2–4 mm), (4) minimal crowding (<2 mm) and spacing (<1 mm), and (5) no history of previous orthodontic or prosthodontic treatment. In addition, subjects with proximal caries or fillings that affected a tooth's size and shape, gross restorations, significant attrition, congenital defects, or deformed teeth are excluded. The absence of tooth anomalies of structure and development is also considered.

From each subject, 14 reference points from the central incisors to the second molars are digitized (Intuos 2 Graphic Tablet, Wacom Company, Japan) to Cartesian coordinates from the photoscan of the dental stone casts made from alginate impressions using custom software written in the Delphi (CodeGear, Scotts Valley, CA) programming language. The resulting data points are shown in Fig. 1. The figure shows that the dental arches are not relevant with gender; see Section 6 for further discussion.

2.2. Registration

Although designed to have the same origin and rotational direction, the data acquisition process inevitably exhibits locational shift and random rotation in the scanned data points, see Fig. 1. Moreover, subjects with normal occlusion