TITLE:
Enumerated sparse extraction of important surgical planning features for mandibular reconstruction

AUTHOR(S):
Nagai, Kazuki; Nakao, Megumi; Ueda, Nobuhiro; Imai, Yuichiro; Kirita, Tadaaki; Matsuda, Tetsuya

CITATION:

ISSUE DATE:
2020

URL:
http://hdl.handle.net/2433/265385

RIGHT:
© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. This is not the published version. Please cite only the published version. この論文は出版社版ではありません。引用の際には出版社版をご確認ご利用ください。

Enumerated sparse extraction of important surgical planning features for mandibular reconstruction

Kazuki Nagai¹, Student Member, IEEE, Megumi Nakao¹, Member, IEEE, Nobuhiro Ueda², Yuichiro Imai³, Tadaaki Kirita², and Tetsuya Matsuda¹, Member, IEEE

Abstract—Because implicit medical knowledge and experience are used to perform medical treatment, such decisions must be clarified when systematizing surgical procedures. We propose an algorithm that extracts low-dimensional features that are important for determining the number of fibular segments in mandibular reconstruction using the enumeration of Lasso solutions (eLasso). To perform the multi-class classification, we extend the eLasso using an importance evaluation criterion that quantifies the contribution of the extracted features. Experiment results show that the extracted 7-dimensional feature set has the same estimation performance as the set using all 49-dimensional features.

I. INTRODUCTION

Clinical treatment is performed by surgeons and medical staff using a large amount of medical knowledge and experience. For instance, in mandibular reconstruction surgery, in which the patient’s own fibula is transplanted [1] [2], the number of fibular segments and the placement in the mandible must be determined. Moreover, various reconstruction plans can be considered depending on the variety in the shape of the mandibles. The quality such plans is difficult to guarantee because manual planning largely depends on the success of past cases and because the planning policy depends on the subjectivity and experience of surgeons. Hence, the systematization of these surgical procedures and a more objective and low-cost surgical planning method are required. In recent years, an automatic planning system for mandibular reconstruction was proposed [3]. However, designing indicators based on the insights of researchers is time consuming and lacks objectivity [4]. However, the high-dimensional deep features that are automatically extracted from medical images are not based on clinically used anatomical features, and it is not possible to understand them from an expert’s viewpoint. To obtain a reliable surgical plan, it is necessary to use uniquely calculated features that can be interpreted by humans.

This study focuses on analysis methods for extracting essential anatomical knowledge from surgical planning databases that have been manually annotated by surgeons for past surgeries. In the last decade, sparse modeling, which can extract essential information using the sparseness that is inherently universal in data, has attracted attention. Kawasaki et al. showed that patient-specific surgical plans can be estimated from past plans using sparse modeling [5]. One of the most common feature selection methods is the least absolute shrinkage and selection operator (Lasso) [6]. In the Lasso, an important feature set is obtained. However, this approach can overlook possibly relevant features not selected by the Lasso. Recently, the enumeration of Lasso solutions (eLasso) for finding multiple models using different feature sets has been proposed [7]. However, this algorithm is intended for the binary classification and cannot classify multiple classes.

In this study, we hypothesized that surgical planning, specifically the number of fibular segments used in mandibular reconstruction, can be determined by a low-dimensional feature set $T_0$. When the Lasso is used, only one feature set is extracted. Therefore, this study extends the concept of eLasso for selecting surgical procedures in fibular-segment based mandibular reconstruction. Our aim is to enumerate multiple sets of low-dimensional features that are important for determining the number of fibular segments. In this paper, we verify the effectiveness of the proposed method using 232 reconstruction plans given by an oral surgeon.

II. METHODS

A. Surgical planning datasets

In this study, we use 3D-CT datasets obtained from 29 patients [3] [8]. The mandible region was extracted from head and neck images, and the fibular segments were extracted from foot images. Then, to evaluate the differences in the resection areas, as shown in Fig. 1(a), we defined six cutting planes. Each cutting plane was defined based...
on anatomical distinctions, specifically $C_0$: the mandibular ramus, $C_1$: the midpoint of the chin and $C_0$, $C_2$: the symmetry point about the midline of $C_3$, $C_3$: the midpoint of the chin and $C_5$, $C_4$: the midpoint of $C_3$ and $C_5$, and $C_5$: the mental foramen. Using these cutting planes, we defined eight resection areas for each dataset, specifically $(C_0, C_2)$, $(C_0, C_3)$, $(C_0, C_4)$, $(C_0, C_5)$, $(C_1, C_2)$, $(C_1, C_3)$, $(C_1, C_4)$, and $(C_1, C_5)$. Assuming that the number of fibular segments was between one and three, the appropriate number of fibular segments and their placements were given by an oral surgeon for these 232 cases. Figures 1(b), (c), and (d) show examples of such reconstruction plans.

In this experiment, features were defined with an anatomical name or medical terms that have been used in the clinical literature for the past 10 years. To eliminate the dependence on the coordinate system, each feature is a distance or angle based on the mandibular condyle and the mental tubercles. The posterior angle of the mandible was defined as the angle formed by the line connecting the left and right mandibular condyles. The most point of the bottom edge of the mandible and the condyle was defined as the distance between the forward-right and forward-left anatomical feature points:

- the distance from the chin to the line of the mandibular condyle: 1 dimension;
- the anterior angle of the mandible: 1 dimension;
- the posterior angle of the mandible: 2 dimensions;
- the angle based on the mandibular condyle and the mental foramen: 1 dimension;
- the distance between anatomical feature points: 28 dimensions;
- the distance between the left and right cutting points and anatomical feature points: 16 dimensions.

The distance from the chin to the line of the mandibular condyle was defined as the distance between the forward-most point of the bottom edge of the mandible and the line connecting the left and right mandibular condyles. The anterior angle of mandible was defined as the angle formed by the line connecting the left and right mandibular angles and the mental tubercles. The posterior angle of the mandible was defined as the angle formed by the line connecting the mandibular condyle and the mental foramen, and the line connecting the mandibular angle and the mental tubercle. The angle based on the mandibular condyle and the mandibular angle was defined as the angle formed by the lines connecting the left and right mandibular condyles and mandibular angles. To incorporate more information about the mandibular morphology as features, the distance of all combinations of anatomical feature points ($8 \times C_2 = 28$ dimensions) was used.

To represent cutting planes by relative position, the distance between two cutting points on the left and right and the anatomical feature points ($2 \times 8 = 16$ dimensions) was used. Then, we normalized each feature between 0 and 1 because the values of different unit systems were used as features.

### B. Outline of the methods

In this study, based on the hypothesis that the number of fibular segments used in mandibular reconstruction can be determined by a low-dimensional feature set $T_0$, we extended the concept of eLasso [7] and applied it to mandibular reconstruction planning. Using a dataset of plans created by an oral surgeon as the training data, we performed the multi-class classification of one, two, and three fibular segment plans. To extend eLasso from binary classification to multi-class classification, we introduce the importance evaluation criterion, which scores the effect of each feature on the classification based on the weight of the feature calculated by the Lasso.

Figure 3 shows the processing flow of the proposed method. First, for case $j$, 49-dimensional features are calculated and a feature vector $d^{(j)} = [d_1, d_2, \ldots, d_{49}]'$ is created. Furthermore, feature matrix $D_f = [d^{(1)} d^{(2)} \ldots d^{(232)}]'$ is created by concatenating the feature vectors. Next, to perform the multi-class classification of the number of fibular segments, we use the one-versus-rest [9]. When performing three-class classification with one, two, and three fibular segments, case $j$ is labeled using a three-element vector $y^{(j)} = (y_1, y_2, y_3)$ as follows.

- Class 1 (one fibular segment): $y^{(j)} = (1, 0, 0)$
- Class 2 (two fibular segments): $y^{(j)} = (0, 1, 0)$
- Class 3 (three fibular segments): $y^{(j)} = (0, 0, 1)$

Each case is labeled, and then correct answer labels $y_i = [y_i^{(1)} y_i^{(2)} \ldots y_i^{(232)}]' (i = 1, 2, 3)$ are created by arranging every $y_i$. The correct answer labels $y_i$ and feature matrix $D_f$ are used as input data. Binary classification of $y_i$ using the Lasso yields regression coefficient $\beta_i$. In this study, feature set $T$ is extracted from these using the importance evaluation criterion $E$ described later. When the Lasso is applied once, only one feature set can be obtained. Therefore, feature matrix $D_f$ is updated based on the extracted $T$, and the Lasso is performed again. This makes it possible to enumerate multiple versions of $T$.

### C. Feature extraction using eLasso

First, we explain how we perform the binary classification of $y_i$. We represent the correct answer labels $y_i$ by a weighted sparse linear combination, where $\beta_i = [\beta_1^{(i)} \beta_2^{(i)} \ldots \beta_n^{(i)}]' (i = 1, 2, 3)$ (the number of features) denotes the weight vectors, which are calculated by minimizing the following objective function.

$$L(\beta_i) = \|y_i - D_f \beta_i\|_2^2 + \lambda \|\beta_i\|_1$$  (1)
In Figs. 4(a), (b), and (c), three weight vectors feature calculated by the Lasso. However, in the multi-class al., feature sets are extracted based on the weight of each feature by the Lasso are enumerated. In the method by Hara et al., feature sets including features that could not be obtained to quantify the effect of each feature on the multi-class classification. Thus, features can be compared. Using the example of the importance evaluation criterion \(E\) calculated, and the class corresponding to the largest value \(\hat{y}_i\) is the estimated number of fibular segments. For example, the calculation for case \(j\) yields \(\hat{y}_j = (0.1, 0.7, 0.3)\), and the estimated number of fibular segments is two because of the maximum \(\hat{y}_2\).

In the proposed algorithm, by introducing eLasso, multiple feature sets including features that could not be obtained by the Lasso are enumerated. In the method by Hara et al., feature sets are extracted based on the weight of each feature calculated by the Lasso. However, in the multi-class classification targeted by this study, as shown in the examples in Figs. 4(a), (b), and (c), three weight vectors \(\beta_1, \beta_2,\) and \(\beta_3\) with different scales were obtained, so the weights of each feature cannot be compared. Therefore, the importance evaluation criterion \(E\) is defined by the following equation to quantify the effect of each feature on the multi-class classification of the number of fibular segments.

\[
E = \sum_{i=1}^{3} \frac{|\beta_i|}{\|\beta_i\|}
\]

All features are evaluated by Eq. (2). Figure 4(d) shows an example of the importance evaluation criterion \(E\) calculated from Figs. 4(a), (b), and (c). It is assumed that a feature with a larger \(E\) is a more important feature in the multi-class classification. Thus, features can be compared. Using the importance evaluation criterion \(E\) as a guideline for feature extraction extends eLasso to multi-class classification, and a framework for enumerating the feature sets with the top-N importance evaluation criterion is provided below.

**STEP 1** For feature matrix \(D_f\), weight vector \(\beta_i\) and the sum \(Q\) of the three objective function values \(L(\beta_i)\) are obtained from Eq. (1).

**STEP 2** All features are evaluated by Eq. (2), and the feature set \(T\) containing the features with the top-N values of \(E\) is obtained. Then, \((T, Q, D_f)\) is retained as a solution candidate.

**STEP 3** Of the retained solution candidates, the one with the smallest sum of the three objective function values is output as the \(k\)th solution.

**STEP 4** The following operation is performed for all features \(t\) of the output solution in STEP 3.

a) The column corresponding to feature \(t\) is removed from feature matrix \(D_f\), and a new feature matrix \(D_f\) is created.

b) For feature matrix \(D_f\), weight vector \(\beta_i\) and the sum \(Q\) of the three objective function values \(L(\beta_i)\) are obtained from Eq. (1).

c) All features are evaluated by Eq. (2), and the feature set \(T\) containing the features with the top-N values of \(E\) is obtained. Then \((T, Q, D_f)\) is retained as a solution candidate.

**STEP 5** Repeat STEP 3 and STEP 4.

In the proposed method, the Lasso and feature extraction are performed with a new feature matrix that is updated by removing the information of the feature extracted by the Lasso from the feature matrix. By repeating this operation, several important feature sets can be output.

Finally, we describe the estimation method using the obtained \(N\)-dimensional feature set \(T\). Creating a new feature matrix \(\hat{D}_f\) that contains only the extracted \(N\)-dimensional feature set, we calculate \(\hat{\beta}\) by performing the Lasso again. The number of fibular segments is estimated by calculating the inner product of these weight vectors and the feature vector of the reconstruction plan. The performance of feature set \(T\) is estimated by the proportion of estimation results that match the number of fibular segments determined by an oral surgeon, that is, by the accuracy rate.

**III. EXPERIMENTS**

The purpose of this experiment was to use the proposed method to obtain multiple low-dimensional feature sets with high estimation performance for classifying the number of fibular segments. To evaluate performance, the number of fibular segments was estimated using only the \(N\)-dimensional feature sets extracted according to \(E\) using the conventional (Lasso-based extraction) method and the proposed (eLasso-based extraction) method. The accuracy rates were then obtained. In the conventional method, the Lasso was applied once, and the top-\(N\) feature set was obtained based on the importance evaluation criterion \(E\). Then, the Lasso was performed again using only the obtained feature...
had the same estimation performance as one in which all
extracted dataset of plans specified by an oral surgeon showed that the
important for determining the number of fibular segments
extracting multiple sets of low-dimensional features that are
using all
The method achieved the same accuracy rate as that obtained
the
proposed method obtained feature sets with higher estimation
N
was used. Specifically,
λ
was used for extraction and
50%
when the estimation was performed using
49-dimensional features defined using traditional anatomical features were used.

**TABLE I**

<table>
<thead>
<tr>
<th>N</th>
<th>Conventional method</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58.62% (2 sets)</td>
<td>60.34% (2 sets)</td>
</tr>
<tr>
<td>2</td>
<td>45.69% (1 set)</td>
<td>75.86% (1 set)</td>
</tr>
<tr>
<td>3</td>
<td>64.66% (2 sets)</td>
<td>79.31% (2 sets)</td>
</tr>
<tr>
<td>4</td>
<td>79.31% (3 sets)</td>
<td>83.62% (3 sets)</td>
</tr>
<tr>
<td>5</td>
<td>79.31% (7 sets)</td>
<td>84.48% (7 sets)</td>
</tr>
<tr>
<td>6</td>
<td>82.33% (6 sets)</td>
<td>85.78% (6 sets)</td>
</tr>
<tr>
<td>7</td>
<td>82.33% (1 set)</td>
<td>87.50% (1 set)</td>
</tr>
</tbody>
</table>

**IV. CONCLUSIONS**

This paper proposed an algorithm based on eLasso for
extracting multiple sets of low-dimensional features that are
important for determining the number of fibular segments
in the mandibular reconstruction. The experiments using a
dataset of plans specified by an oral surgeon showed that the
extracted 7-dimensional feature set obtained by this method
had the same estimation performance as one in which all

**ACKNOWLEDGMENT**

This research was supported by JSPS Grant-in-Aid
for Scientific Research (B) (19H04484) and Grant-in-Aid
for Challenging Exploratory Research (18K19918).
We thank Kimberly Moravec, PhD, from Edanz Group
(www.edanzediting.com/ac) for editing a draft of this
manuscript.

**REFERENCES**

[1] A. F. Flemming, M. D. Brough, N. D. Evans, H. R. Grant, M. Harris,
D. R. James, M. Lawlor, and I. M. Laws, "Mandibular Reconstruction
43, No. 4, pp. 403–409, Jul 1990.
17, No. 1, pp. e23–e30, Jan 2016.
Krita, and T. Matsuda, "Automated Planning with Multivariate Shape
Descriptors for Fibular Transfer in Mandibular Reconstruction," *IEEE
Transactions on Biomedical Engineering*, Vol. 64, No. 8, pp. 1772–
and T. Matsuda, "Volumetric Fibular Transfer Planning with Shape-based
Indicators in Mandibular Reconstruction," *IEEE Journal of Biomedical
T. Krita, and T. Matsuda, "Sparse Shape Model for Fibular Transfer
Planning in Mandibular Reconstruction," in *38th Annual International
Conference of the IEEE Engineering in Medicine and Biology Society
(EMBC)*, 2016, pp. 2508–2511.
Krita, and T. Matsuda, "Statistical Analysis of Interactive Surgical Planning
with support vector data description," in *IEEE Conference on Granular
Computing (GrC2010)*, San Jose, CA, USA, 2010, pp. 817–821.
Optimization and Statistical Learning via the Alternating Direction