Analysis of Surfaces Using Constrained Regression Models

Sune Darkner[1,2], Mert R. Sabuncu[3], Polina Golland[3], Rasmus R. Paulsen[2], and Rasmus Larsen[1]

[1] Department of Informatics and Mathematical Modelling, Technical University of Denmark, Denmark
sda@imm.dtu.dk
[2] Oticon A/S, Denmark
[3] Computer Science and Artificial Intelligence Laboratory, MIT, USA

Abstract. We present a study of the relationship between the changes in the shape of the human ear due to jaw movement and acoustical feedback (AF) in hearing aids. In particular, we analyze the deformation field of the outer ear associated with the movement of the mandible (jaw bone) to understand its effect on AF and identify local regions that play a significant role. Our data contains ear impressions of 42 hearing aid users, in two different positions: open and closed mouth, and survey data including information about experienced discomfort due to AF. We use weighted support vector machines (WSVM) to investigate the separation between the presence and lack of AF and achieve classification accuracy of 80% based on the deformation field. To robustly localize the regions of the deformation field that significantly contribute to AF we employ logistic regression penalized with elastic net (EN). By visualizing the selected variables on the mean surface, we provide clinical interpretations of the results.

1 Introduction

One of the big challenges for hearing aid users is acoustical feedback (AF). When a customized hearing aid is produced, the ventilation size and gain are adjusted accordingly. However, when the ear changes shape due to movement of the mandible, false leaks and feedback can occur. Modern feedback cancellation algorithms exist, but they rely on the detection of feedback. The time lag between detection and cancellation causes a squeaking sound when a person talks or chews. Identifying and characterizing the main causes of AF can improve hearing aid designs to minimize AF risk. In this paper, we investigate the relationship between the deformation of the outer ear and AF. We are interested in localizing regions that play a significant role in this phenomenon. In our experiments, we work with surface scans of ear impressions from 42 subjects under two different conditions: open and closed mouth and questionnaire data that includes information about AF-related experience. The ear impressions are co-registered using a group-wise registration algorithm via a kernel-based nonlinear deformation model. We analyze the intra subject deformation fields using a classification method to illustrate the differences between the two groups: subjects who experience AF and subjects who do

not. Using Weighted Support Vector Machines, we achieve 80% cross-validation accuracy on our data. Additionally, we employ logistic elastic net regression (logistic EN) to identify the surface points that consistently explain the difference between the two populations. Generalization performance and statistical significance of the fitted model are used to determine the parameters of the regression algorithm. We compute statistical significance based on a standard likelihood ratio test and the effective degrees of freedom for the model similar to [1]. The model is shown to be significant with $p < 0.001$ on the whole data set.

2   Prior Work

Previous studies investigated the deformation of the ear canal and concha using calipers[2], and deformable shape models [3,4]. Yet, the relationship between deformations of the ear and clinical observations have so far not been explored. In other medical contexts, such as neuroimaging, the relationship between image-derived features and clinical data has been extensively studied [5]. A popular method used to explore differences between two groups in a population is Support Vector Machines (SVM) [6]. The discriminative direction of an SVM can be used to illustrate the differences between classes [7]. However, an important challenge in such approaches, is the interpretation of discriminant features. Another problem commonly presented by medical data is its unbalanced nature: there are typically more negative samples than positive examples. To handle this, we use weighted SVMs [8]. Medical imaging provides further challenges: samples are high dimensional and few. Moreover, we expect that some of these dimensions exhibit significant correlations. Our goal is ideally to discover *all* these dimensions. Ridge regression [9] (or, in general Tikhonov ℓ_2 regularization [10]) takes this type of underlying structure into account. Additionally, we expect that only a small number of dimensions are related to the clinical outcome of interest. This prior knowledge can be formalized using a constraint on the ℓ_1 norm of the regression coefficients [11]. Elastic Net (EN) [12] combines these two approaches to achieve a sparse and correlated set of predictors. A Bayesian interpretation of this method yields an efficient implementation [12]. The method has been extended to fit the generalized linear model framework enabling various types of regressions through the canonical link functions including logistic regression. Analyzing multi-subject medical data requires spatial correspondence, usually determined via image registration. Motivated by group-wise registration methods [13], we use a co-registration formulation that simultaneously aligns all surfaces. We employ a kernel-based nonlinear deformation model [14][15] to achieve a dense, diffeomorphic correspondence within and across subjects.

3   Data

The data consists of 84 impressions from 42 hearing aid users. Two impressions were obtained from each individual in different positions. (ɪ) Normal position, chosen as reference, (ɪɪ) mouth opened. A spacer was used to ensure consistency

with respect to the angle of the mouth opening. The impressions were all acquired from the subjects' right ear. Each impression consists of ≈ 5500 points in 3D i.e. vertices, corresponding to a 16500 dimensional feature vector. In addition to the shape data, the subjects filled out a questionnaire, including questions regarding acoustical feedback (AF). All subjects that experienced frequent AF problems (once a week or more) were grouped together if the annoyance was related to ear deformation, i.e. jaw movement and facial expressions. The latter is included because it often involves jaw movement. Thus, we obtained two groups: subjects who experience AF and subjects who do not.

4   Methods

4.1   Co-registration and Preprocessing

The ear impressions were scanned to obtain surfaces, which were then manu-ally preprocessed to remove artifacts. We represent each surface as a triangular mesh S, with vertices denoted by $\boldsymbol{x}_S \in S \subset \mathbb{R}^3$. The registration framework is based on the method described in [16], which uses the difference between signed distance maps of the two surfaces S_1 and S_2 to compute a similarity metric. For computational efficiency, the distance is only computed on a narrow band $Q \subset \mathbb{R}^3$ that covers both surfaces, i.e., $Q \supset S_1, S_2$. The distance between S_1 and S_2 is then defined as:

$$f(S_1, S_2) = \frac{1}{\|Q\|} \sum_{\boldsymbol{x} \in Q} (d_{S_1}(\boldsymbol{x}) - d_{S_2}(\boldsymbol{x}))^2, \tag{1}$$

where d_S denotes the signed distance map of the surface S and $\|Q\|$ denotes the volume of Q. The pairwise registration problem is formulated as the minimization of Eq. 1 and solved using Newton's method. We parameterize the deformations using a kernel-based approach [15] defined on a control grid in a coarse-to-fine fashion, from $2^3 - 40^3$ control points. The gradient of the objective function with respect to this parametrization can be easily computed. We extend this approach to a multi-subject setting by defining a mean surface $S_\mu = \frac{1}{N} \sum_i S_i$, where N is the number of subjects. The mean surface is updated at every iteration and a set of transformation parameters is computed for each subject based on its distance to the mean surface. To anchor the deformations, we constrain the average deformation across all subjects to be identity. In other words, the mean of the deformation parameters is zero.

4.2   Classification and Regression

Our goal is to analyze the deformation fields obtained from the co-registration step to show differences between subjects who experience Acoustical Feedback (AF) and subjects who do not. Ground truth labels were based on the question-naire data, as described in Section 3. The localized nature of the deformation parametrization allows us to identify regions that influence AF. We use the

popular SVM method with no kernel to classify the data based on the whole deformation field. Since our data contained 18 positive samples and 24 negative samples, we investigated weighted SVM (WSVM) [8]. In contrast with SVM where the penalty of misclassification is universal, WSVM sets a different penalty for each class and the the ratio between class penalties are inversely proportional to the class size ratio. In our experiments, this approach improves results by about 5% when compared with the uniform penalty on errors. Due to the challenges of interpreting the discriminant features, i.e., support vectors in the SVM experiment, we explored a logistic regression approach. Logistic regression models the probability P of a certain outcome, in our case subjects who experience AF and is estimated by the ratio of occurrences. This is accomplished through the canonical link $\log(\frac{p_i}{1-p_i})$, known as the log-odds or logit. Standard logistic regression can thus be formulated as:

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \boldsymbol{\beta}\mathbf{x_i}, \forall i \tag{2}$$

where β_0 is the intercept, $\boldsymbol{\beta}$ is the regression coefficients of size $1\times$ number of dimensions (16,500 in our case) and $\mathbf{x_i}$ is a column vector that represents an observation. Logistic elastic net extends standard logistic regression by penalizing the parameters $\boldsymbol{\beta}$ with the ℓ_1 and ℓ_2 penalty.

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \boldsymbol{\beta}\mathbf{x_i} + \epsilon_i, \quad \text{s.t. } \|\boldsymbol{\beta}\|^2 < \rho, |\boldsymbol{\beta}| < \xi \tag{3}$$

for some $\rho, \xi > 0$ and for all i. The solution can be found via a MAP approach where the following prior on the parameters is used:

$$P_{\lambda,\alpha}(\boldsymbol{\beta}) = C(\lambda,\alpha)e^{-\lambda(\alpha\|\boldsymbol{\beta}\|^2 + (1-\alpha)\|\boldsymbol{\beta}\|)}, \tag{4}$$

where λ is the hyper-parameter that determines the total amount of regularization, α determines the trade-off between the two penalty terms and C is the normalization. Logistic EN has the properties of ridge regression and yields sparse solutions thanks to the ℓ_1 constraint. Thus the model takes covariations into account while only consisting of a small subset of significant regression coefficients. This yields dimensionality reduction conditioned on the information in the data set.

5   Model Validation and Selection

We use cross-validation to select the hyper-parameters α and λ. Furthermore, by using the likelihood ratio test, we can ensure that the selected parameters yield statistically significant models, i.e., explain the data well. The likelihood ratio computes $\Lambda = \frac{L(\boldsymbol{\beta}|\mathbf{X})}{L(\beta_0|\mathbf{X})}$, where the numerator $L(\boldsymbol{\beta})$ is the log-likelihood of the fitted model, defined as:

$$L(\boldsymbol{\beta}|\mathbf{X}) = \sum_{i=1}^{N}\left[y_i\boldsymbol{\beta}^t\mathbf{x_i} - \log(1+\exp(\boldsymbol{\beta}^t\mathbf{x_i}))\right], \tag{5}$$

where $y_i \in \{0, 1\}$ is the ground truth label, β_0 is included in $\boldsymbol{\beta}$ and $\mathbf{x_i}$ is the i'th observation in $\mathbf{X}$, the data to which the model is fitted. $L(\beta_0)$ is the null-hypothesis likelihood, computed with $\boldsymbol{\beta} = 0$. Note that $(-2 \log \Lambda)$ is approximately χ^2 distributed with a parameter equal to the effective degrees of freedom of the model minus 1. Following [17], we can compute the effective degrees of freedom as $\sum_{j=1}^{P} \frac{\alpha_i}{\alpha_i + \lambda}$, where P is the number of dimensions and $\{\alpha_i\}$ are the eigenvalues of the data matrix $\mathbf{X}^T \mathbf{X}$.

6   Analysis

The analysis is divided into two parts, classification and regression. Performing classification demonstrates the difference between the two classes. To reveal how the deformation field relates to AF problem regression is performed. We fit a logistic regression model *to each variable i.e. vertex* independently and selects the significant models as determined by the likelihood test to identify *significant variables*. This approach does not take covariances into account since each variable is treated independently. We use logistic EN on all the data to perform variable selection while accounting for covariances. Logistic EN provides better and more robust localization, which yields interpretable results. Yet, SVM is typically superior with respect to classification accuracy.

6.1   Classification

To investigate if the data set size is sufficient for generalization, we use SVM to classify and perform cross validation (100 random trials) with an increasing number of samples included in the training set. The purpose is to investigate whether including further samples will improve the classification. The test error for each class vs. the number of samples included in the model is plotted in Fig. 1(c). From the figure it is clear that better results could be obtained with more data, however, the test error is around 22-25%, clearly showing separation in the data. Using WSVM reduces the error further to 20%. Fig. 2 illustrates the variable weights on a mean surface in four representative random trials. Comparing these four trials, one sees that even though the general patterns seem to be consistent, there is significant noise and the localization quality is not sufficient. Fig. 2(e)shows the frequency of each variable appearing in the top 20% quantile over all 100 trials. We notice that variables robustly (red regions) appearing in the top 20% quantile are sparse.

6.2   Regression

To get a base line, we perform a logistic regression of each parameter against the response variable. The significant variables are shown in Fig. 4(a), where red indicates a significant model. As can be seen significant variables cluster nicely. However, we expect significance in and around the canal as well, which is not present in this figure. This is due to the fact that each variable is treated
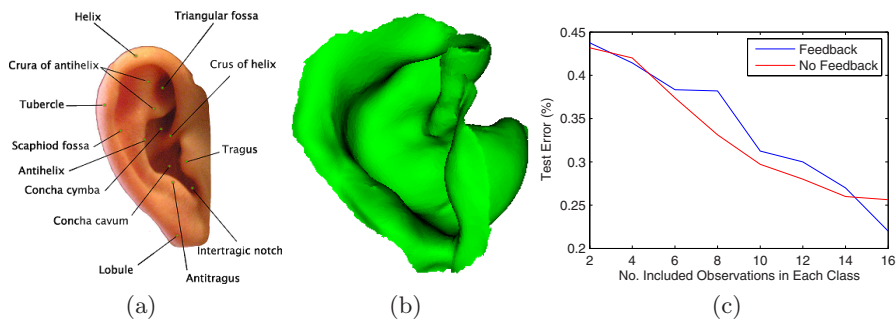
Fig. 1. (a) An anatomical atlas of the ear. (b) A typical ear impression. (c) Number of parameters for each class included in the model vs. the test error.
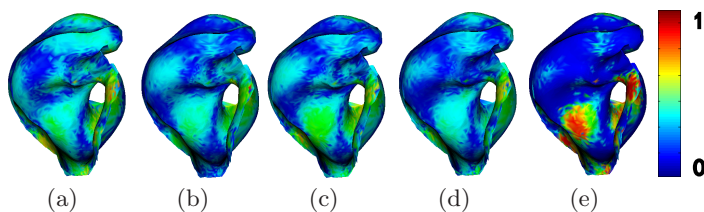

Fig. 2. The normalized coefficients from the SVM solution for 4 random cross validation iterations mapped to the mean surface.(e) the frequency of each vertex appearing in the to 20% quantile.
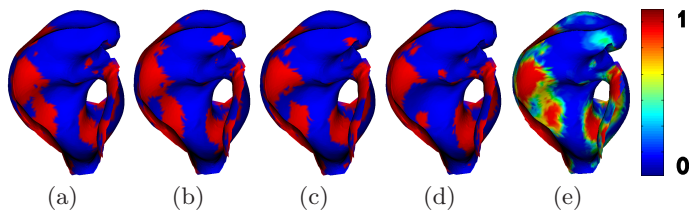

Fig. 3. The selected coefficients from the penalized logistic regression solution for 4 different cross validation iterations mapped to the mean surface. Red indicate a selected variable.3(e) Is the cross-validated probability of a variable being selected.
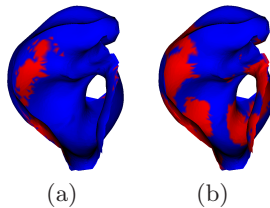

Fig. 4. (a) The significant logistic regression models on each variable mapped to the mean surface. (b) The selected variables of the full logistic EN regression model mapped to the mean surface. Red indicate a selected model/variable/vertex.

independently. To recover from this, we use the Logistic EN with the hyper-parameters determined by cross validation and the likelihood ratio test. From the formula for effective degrees of freedom we can make an educated guess of the initial regularization parameter for the ℓ_2-penalty, which needs to be around 1000 for the model to be statistical significant. An interval around this value is searched to find a suitable combination where the likelihood ratio rejects the null hypothesis. This yields $\lambda = 1000$ and $\alpha = 0.997$. The resulting models contain between 2000-2500 variables but are heavily constrained with 2-3 free parameters. To validate the model, 10-fold cross validation is performed with an equal number of observations from each class used in training, which yields 73% accuracy for both classes. The regularization in the model increases test error to 5-10%. A representative selection of the resulting cross validation models can be seen in Fig. 3. We note that the models are consistent over different runs. Having estimated the regularization parameters we build a model on the full data set. The model is shown in Fig. 4(b). The estimated degrees of freedom for the model is 3.88, the model is significantly better than H_0 ($p < 0.001$). The resulting model is in very good correspondence with the individual models from the cross validation and includes far more of the surface in the model compared to the logistic regression in Fig. 4(a). Moreover, the logistic EN gives better and more consistent localization compared to the SVM, and the results are less noisy (Fig. 2 and Fig. 3). In addition the results are in good correspondence with clinical observations.

6.3   Clinical interpretation

The results are consistent with how hearing aids are situated in the ear. It is interesting that not only the entrance to the canal is important, but also the lower part of the outer ear, tragus and anti tragus (see Fig. 1(a)). The reason is that its often one or both of these that hold the hearing aid in place. The small In The Ear devices (ITE) are held in place by the opening of the canal and Tragus, where as Behind The Ear aids (BTE) are molded to the entire concha (most of the impression). Also the deformation, which we know occur deeper inside the canal seems to have little influence on the acoustical feedback. Only the bottom part of the inner canal seems to have a small influence, which corresponds to clinical observations made during normal practice. The results in Fig. 4(b) lead to the possibility of improving the fit with respect to feedback.

7   Conclusion

By using constrained logistic regression, we find parts of the ear canal surface that explain the feedback problems experienced by users. The regression model gives good localization and the outcome is easy to interpret. Furthermore, by using an extended framework for estimation of the free parameters in the model, a cross validation scheme based on significant models can be used. We show that the final model is significant ($p < 0.001$). In addition, we show that a classifier

based on the WSVM can achieve a classification accuracy of 80% for both classes, with the possibility of improvement if more data were available.

References

1. Zou, H., Hastie, T., Tibshirani, R.: On the "degrees of freedom" of the lasso. Annals of Statistics 35(5), 2173–2192 (2007)
2. Oliviera, R., Hammer, B., Stillman, A., Holm, J., Jons, C., Margolis, R.: A look at ear canal changes with jaw motion. Ear and Hearing 13(6), 464–466 (1992)
3. Paulsen, R.R., Larsen, R., Laugesen, S., Nielsen, C., Ersbøll, B.K.: Building and testing a statistical shape model of the human ear canal. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002. LNCS, vol. 2489. Springer, Heidelberg (2002)
4. Darkner, S., Paulsen, R.R., Larsen, R.: Analysis of deformation of the human ear and canal caused by mandibular movement. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 801–808. Springer, Heidelberg (2007)
5. Sjoestrand, K., Cardenas, V.A., Larsen, R., Studholme, C.: A generalization of voxel-wise procedures for high-dimensional statistical inference using ridge regression. In: SPIE Medical Imaging (2008)
6. Boser, B., Guyon, I., Vapnik, V.: A training algorithm for optimal margin classifiers. In: Fifth Annual Workshop on Computational Learning Theory, pp. 144–152 (1992)
7. Golland, P.: Discriminative direction for kernel classifiers. In: NIPS, pp. 745–752 (2001)
8. Liu, S., Jia, C.Y., Ma, H.: A new weighted support vector machine with ga-based parameter selection, vol. 7, pp. 4351–4355 (2005)
9. Hoerl, A.E., Kennard, R.W.: Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12, 55–67 (1970)
10. Tikhonov, A.N.: Solutions of incorrectly formulated problems and the regularization method. Soviet Math. Dokl. 4, 1035–1038 (1963)
11. Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Statist. Soc. B 58(1), 267–288 (1996)
12. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. J. R. Statist. Soc. B 67(part 2), 301–320 (2005)
13. Zöllei, L., Learned-Miller, E., Grimson, E., Wells, W.M.: Efficient population registration of 3d data. In: ICCV (2005)
14. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: Application to breast mr images. IEEE Transactions on Medical Imaging 18(8), 712–721 (1999)
15. Cootes, T., Twining, C., Taylor, C.: Diffeomorphic statistical shape models. In: Proc. British Machine Vision Conference, vol. 1, pp. 447–456 (2004)
16. Dakner, S., Vester-Christensen, M., Paulsen, R.R., Larsen, R.: Non-rigid surface registration of 2D manifolds in 3D euclidian space. In: SPIE Medical Imaging (2008)
17. Malthouse, E.C.: Shrinkage estimation and direct marketing scoring models 13(4), 10–23 (1999)