35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

The Use of Pseudo-landmarks for Craniofacial Analysis: A Comparative Study with L_1-Regularized Logistic Regression

Ezgi Mercan[1], Linda G. Shapiro[1], Seth M. Weinberg[2], Su-In Lee[1]

***Abstract*— Morphometrics, the quantitative analysis of shape, is used by craniofacial researchers to study abnormalities in human face shapes. Most of the work in craniofacial morphometrics uses landmark points that are manually marked on 3D face data and processed via a generalized Procrustes analysis. For large data sets this manual process is very time-consuming. Dense sets of pseudo-landmarks have also been proposed and successfully used for classification and clustering, but the two main methods in the literature are very computationally intensive. We have developed a computationally simple method that can compute pseudo-landmark points at different resolutions from 3D meshes of human faces. In this paper, we perform a comparative study employing L_1-regularized logistic regression to train a classifier that predicts the sex of 500 normal adult face meshes in order to compare our method to two alternative pseudo-landmark methods and a distance matrix approach.Our results show that our method, which is fully automatic, achieved similar results to the best-scoring methods with no manual landmarking and with much lower computation time. Use of the distance matrix did not improve classification results.**

I. INTRODUCTION

The FaceBase Consortium[1] is a group of investigators who study craniofacial abnormalities in humans and other species and who are providing data and software tools to be shared by craniofacial researchers [6]. One project, the 3D Facial Norms Database, is studying the variation of facial forms in humans by constructing a repository of 3D facial data (in the form of 3D meshes) and genetic data. Samples of healthy Caucasian individuals are being collected in order to analyze the variation in morphology of the face. Other FaceBase projects deal with abnormalities, including cleft lip, cleft palate, and midface hypoplasia (flattening). The 3D Facial Norms Database is expected to be useful for providing control data for these and other projects studying abnormalities.

Most of the work on morphometrics in the craniofacial research community uses standard hand-marked landmarks to characterize the data. Usually, the data are aligned via these landmarks using the well-known Procrustes algorithm [8] and can then be compared using the related Procrustes distance from the mean or between individuals. For large databases, hand landmarking is very tedious, and automatically determined landmarks still have to be checked by humans. This makes the use of pseudo-landmarks an attractive alternative. There are two pseudo-landmarking methods in the medical literature: 1) the well-known dense correspondence method of Hammond and Hutton [5] and 2) the more recent anthropometric mask approach of Claes [2], both of which require an initial set of hand-marked labels to compute correspondences and have fairly heavy computation.

We have developed a very simple, but effective, method that computes pseudo-landmarks by cutting through each 3D head mesh with a set of horizontal planes and extracting a set of points from each plane. Correspondences among heads are not required, and the user does no hand marking. For classification purposes, we use L_1-regularized logistic regression, which deals effectively with the high dimensionality of the data and modest number of samples. This method also allows us to easily detect the points that most contributed to any given classification task.

We compare our methodology to 1) the use of the 24 standard hand-marked landmark points, 2) the Hammond method, 3) the Claes method, and 4) the use of a distance matrix in terms of accuracy of classification and computation time. While our intended use of this methodology is for general analyses, content-based image retrieval, and grouping within the normal population, we used classification experiments for comparison because of the availability of ground truth data. Although the data came with both sex and age attributes, the task of age discrimination was very difficult, both for the computer and the human experimentors, since our data set consists of adults between 18 and 40 years of age, which are not easily discriminable based on shape alone. Thus our experiments in this paper are all on the sex attribute, which provided a very good test vehicle.

II. RELATED LITERATURE

Classic morphometrics research uses a set of standard landmarks and employs generalized Procrustes analysis [8] for comparison, clustering, or classification. Each head is represented by its configuration of landmark points, which are translated to have the same centroid and scaled to have the same size. Then, each object is rotated to best fit an arbitrarily chosen consensus head according to Procrustes distance, the Euclidean distance between the two sets of coordinates. If the total Procrustes distance for the set is small enough, the procedure terminates. Otherwise, the consensus is set to the average of the newly rotated configurations and the procedure iterates. Principal components analysis (PCA) and other methods can be applied to the dataset that is now represented in its Procrustes shape coordinates.

Hammond *et al.* extended classical morphometrics to their dense surface models [5] in which they obtain a corre-

[1]Department of Computer Science and Engineering, University of Washington, Seattle, Washington
[2]Department of Oral Biology, Center for Craniofacial and Dental Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania

[1]https://www.facebase.org/

spondence between a base mesh and all other individual meshes, starting with an initial hand-marked correspondence and using Bookstein's thin-plate spline technique [1]. Claes *et al.* [2] used an anthropometric mask of spatially dense, uniformly distributed points calculated on an average face from a healthy population and fit onto each other face.

Another methodolgy that used pseudo-landmarks was developed by Ruiz *et al.* [9] for classification of skulls with different types of craniosynostosis, a disorder in which the sutures of the skull fuse prematurely. They manually selected three CT slice planes defined by internal brain landmarks and lying parallel to the skull base plane and extracted a sequence of equally spaced points from the oriented outlines of the CT bone image at each plane. The distance matrix derived from all pairs of these points was the feature vector used in their work. Lin *et al* [4] derived a set of symbolic descriptors from the distance matrix by probabilistic latent semantic analysis (PSLA) [7] and used them for both classifying and clustering the data. Yang *et al* [13] took this approach further by fully automating the process of constructing the distance matrix from a set of planes parallel to a base plane required to pass through the nasion and opisthion points and to be perpendicular to the plane of symmetry of the head. She used different variants of logistic regression for classification.

There has also been a great deal of work in finding facial landmarks and some work on sex classification in the computer vision literature. For the most part, this work uses 2D color or gray tone photographs, rather than 3D meshes. One recent work that is worthy of mention is the classification work of Wu, Smith, and Hancock [14] which obtains a 97% accuracy on sex classification. However, their data is 2D gray tone photographs, which they go to some trouble to convert to an estimation of 3D shape from surface normal vectors. Because their representation, their data set, and their goals are completely different from our own, our results cannot be directly compared to theirs.

III. METHOD

Our method starts with 3D head meshes that have been pose-normalized to face front and whose plane of symmetry (midsagittal plane) has been computed. We compute two landmark points, the sellion and chin tip, and construct planes through these points perpendicular to the plane of symmetry of the face. Using these two planes as base planes, we construct m parallel planes through the head and from each of them sample a set of n points from the face region, where the parameters n and m are selected by the user. The set of all $m \times n$ 3D points becomes the feature vector of pseudo-landmark points representing the face, which can be used in description, quantification, and the classification experiments used for comparison results in this paper.

A. Finding Sellion and Chin Tip

Each head mesh has been pose-normalized to face straight forward and the plane of symmetry detected. The intersection of the plane of symmetry with the surface of the face is a contour. The tip of the nose is found as that point of the contour with the largest z-value. Following the contour up from the tip of the nose, the sellion is located at a local minimum for z. Following the contour down, the chin point is located at a local maximum for z. The sellion was selected as a landmark from which to begin the construction of pseudo-landmarks, because it can be reliably detected in the head meshes and is an approximate upper bound on the midface area studied by many of the researchers in the FaceBase Consortium. While the corners of the mouth would usually mark the bottom of the midface region, the tip of the chin was selected for this work, because the chin area was deemed important in sex classification.

B. Constructing Planes and Extracting Points

A plane PU through the sellion and a second plane PL through the tip of the chin, both perpendicular to the plane of symmetry of the face, are constructed. In this study, the user is asked to select how many equally-spaced y-z planes he/she would like to use between PL and PU and how many above PU, leading to the selection of m horizontal planes. The user also selects the number of points n to be extracted from the face on each plane. A horizontal slice of the head mesh is extracted at each of the m planes and restricted to the facial area by eliminating the back side of the head. Then at n equally-spaced x-positions including that of the plane of symmetry, the y and z values are sampled to obtain n 3D points on each plane. Figure 2a shows a 3D head mesh with two planes above the sellion, one at the sellion, one at the tip of the chin, and eleven in between.

C. L_1-Regularized Logistic Regression

In order to compare our pseudo-landmark methodology to the other methods, we ran a large set of classification experiments, using the sex ground truth of the data. Motivated by the successful work of Yang *et al* [13], we used L_1-regularized logistic regression [10] to learn a model to predict the sex classes of the samples in our dataset. Logistic regression learns the model parameters w_0 and $\mathbf{w}$ in the probability function $p(y|\mathbf{x}, \mathbf{w})$ that a data sample $\mathbf{x}$ belongs to a certain class y.

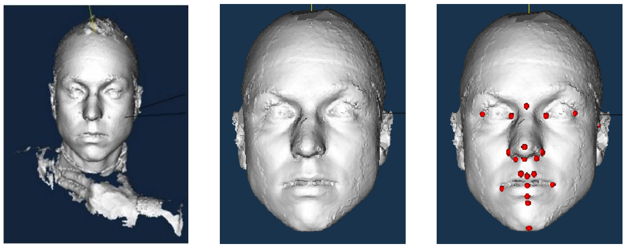
$$p(y|\mathbf{x}, \mathbf{w}) = \frac{1}{1 + exp(-y(\mathbf{w}^T\mathbf{x} + w_0))} \tag{1}$$

Due to the high-dimensionality of the data, learning the unregularized logistic regression [4] will result in overfitting. To avoid overfitting, we applied L_1 regularization that induces sparsity in the solution $\mathbf{w}$ such that many of the coefficients in $\mathbf{w}$ are set to exactly zero. L_1-regularization has been rigorously proven to be effective in selecting relevant features when there are exponentially many irrelevant ones [3]. The parameters w_0 and $\mathbf{w}$ minimize the following loss function called the *lasso*:

$$l(w_0, \mathbf{w}) = \sum_{i=1}^{n} \log(1+\exp(-y_i(\mathbf{w}^T\mathbf{x_i}+w_0)))+\lambda \sum_{i=1}^{m} |w_i| \tag{2}$$

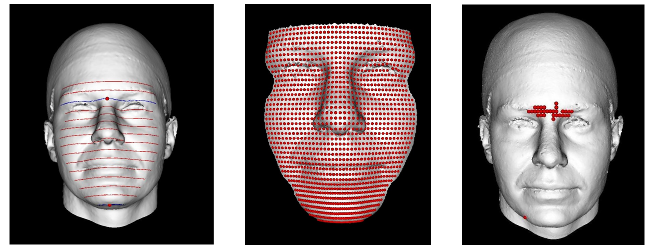
$$\{w_0, \mathbf{w}\} = \min_{w_0, \mathbf{w}} l(w_0, \mathbf{w}) \tag{3}$$

where λ is a regularization parameter for the L_1-norm. L_1 regularization has been widely used in many applications in



(a) (b) (c)

Fig. 1. (a) Uncleaned mesh obtained from the 3dMD imaging system (b) Cleaned and pose-normalized version of the mesh for automated analysis (c) 20 landmark points on the cleaned mesh.



(a) (b) (c)

Fig. 2. (a) Head mesh showing planes at the sellion and the tip of the chin with 11 planes between and 2 planes above them. (b) Face mesh with pseudo-landmarks in 35×35 resolution. (c) Head mesh with pseudo-landmarks that were most useful for sex classification at 35×35 resolution.

which a small number of features must be selected out of many.

Additionally, when we use the original pseudo-landmark coordinates (as opposed to their principal components), the separate x, y, and z features that correspond to the same pseudo-landmark are likely to be either predictive or not predictive together. Therefore, we use a variation of L_1-regularization, the *group lasso* [11], to induce bias such that the features corresponding to a pseudo-landmark tend to be selected together. The loss function of the *group lasso* is,

$$l(w_0, \mathbf{w}) = \sum_{i=1}^{n} \log(1 + \exp(-y_i(\mathbf{w}^T \mathbf{x_i} + w_0))) + \lambda \sum_{i=1}^{m} |w_i| + \mu \sum_{i=1}^{m} \sqrt{w_{i1}^2 + w_{i1}^2 + w_{i1}^2} \tag{4}$$

where μ is a regularization parameter for the new penalty term, and w_{i1}, w_{i2} and w_{i3} are weights that correspond to the ith pseudo-landmark.

IV. DATA

Our head data consists of 3D meshes of 30-40,000 points obtained from several 3dMD® digital stereophotogrammetry imaging systems, which were used to capture the facial surfaces of subjects in the 3D Facial Norms Database. These systems are outfitted with multiple CCD cameras mounted at fixed angles and distances, to capture overlapping views of the face and head. The entire capture process occurs in less than 2 milliseconds and results in a dense 3D connected mesh that conforms to the geometry of the face.

Prior to 3D image capture, scalp hair obscuring the subject's face was cleared away and subjects were positioned so that their heads were centered between the imaging pods. The heads were positioned slightly upward in order to ensure adequate coverage of the subnasal region. Meshes were cleaned to remove extraneous parts of the body and pose-normalized using a method described in [12]. Figure 1 shows a surface mesh before cleaning (a) and the cleaned and pose-normalized surface mesh produced for our analyses (b).

Human experts marked 24 landmarks on each mesh for use in our experiments. The 24 landmarks collected include the nasion, pronasale, subnasale, labiale superius, stomion, labiale inferius, sublabiale, ganthion, endocanthion (right), endocanthion (left), exocanthion (right), exocanthion (left), alare (right), alare (left), alar curvature point (right), alar curvature point (left), subalare (right), subalare (left), crista philtri (right), crista philtri (left), chelion (right), chelion (left), traion (right), and tragion (left). Figure 1c shows the 24 landmark points on the head of Figure 1b.

V. EXPERIMENTS AND RESULTS

Our experiments use a preliminary dataset of 500 adult heads from the 3D Facial Norms Database. We have only 3D meshes and no color data. The distribution is as follows: 250 males, 250 females, with ages between 18 and 40. In these experiments, feature extractions were run on a 64 bit Microsoft Windows 8 machine with an Intel Core i7-2600 3.40 GHz CPU and 4GB RAM, using VTK library functions for 3D processing. Classifications were run on a 64 bit Microsoft Windows Server 2008 R2 machine with Intel Xeon CPU E5649 @ 2.53 GHz (2 processors) and 32GB RAM, using MATLAB 2012b and the SLEP package.

After the processing described above, we obtain a feature vector of landmark points for each head and compute feature vectors of pseudo-landmark points at multiple different resolutions, one of which is shown in Figure 2b. We also apply the Hammond method and the Claes method to achieve feature vectors of pseudo-landmark points from those methods. Finally, we produce a distance matrix from our 35 × 35 resolution pseudo-landmarks for further comparison.

Figure 3 shows the results for sex classification using 10-fold cross validation and running each experiment 10 times with the average accuracy reported. In all experiments, 500 samples were used: 250 male and 250 female. Experiments were run both with original points (if possible) and with principal components analysis (PCA). In the PCA experiments, all principal components were used, since using less tended to reduce the accuracy.

The results show that pseudo-landmarks of all categories achieved higher sex classification accuracy than the standard 24 landmarks both when the original landmark points were used for classification and when the principal components were used. All classification scores improved with PCA. All pseudo-landmark methods–our own, Hammond, and Claes–achieved about 95% accuracy using PCA, with the 35 × 35 distance matrix scoring slightly lower. Since our lowest resolution (35 × 35) feature vector worked just as well

	PTS (num)	PTS (acc)	PCA (acc)	Feat (sec)	Pts (sec)	PCA (sec)
LMK	24	90.6	92.5	1743	45	50
P35	1225	93.4	95.2	2163	4995	142
P45	2025	93.2	95.2	3548	9960	160
P55	3025	93.4	95.2	5166	14200	185
P65	4225	93.3	95.3	7132	19500	207
Ham	200000	NA	95.1	7067	NA	4368
Cla	15000	93.9	95.6	8394	65280	398
Dis	1225	NA	94.6	2223	NA	390

Fig. 3. Sex classification experiments. LMK = 24 Landmarks, P35 = 35 $\times$ 35, P45 = 45 $\times$ 45, P55 = 55 $\times$ 55, P65 = 65 $\times$ pseudo-landmarks, Ham = Hammond, Cla = Claes, Dis = 35 $\times$ 35 distance matrix, num = number, acc = accuracy. Feat (sec) = feature extraction time, PTS (sec) = classification time with points, PCA (sec) = classification time with PCA. Number of points for Hammond and Claes is approximate.

as our highest resolution (65 $\times$ 65) feature vector, it is our method of choice. The Hammond method takes an order of magnitude longer to run than either our method or the Claes method, which takes about 4 times as long as our method. Furthermore, both the Hammond method and the Claes method require a set of hand-marked landmarks for initialization of their correspondence-finding procedures, while our method is fully automatic. Using the 35 $\times$ 35 distance matrix instead of the original points did not gain any improvement in accuracy, and at higher resolutions the matrices were too large to use. Thus in sex classification, our pseudo-landmarking method, using the principal components of the original points, is superior to both hand-landmarking and the other well-known pseudo-landmarking methods.

One advantage of using logistic regression is its ability to determine the most important features (here points of the face) for a given classification task. For this work, we defined the most useful features as those for which logistic regression calculates a weight higher than an empirically determined threshold (here 0.01). Figure 2c shows the points that were most useful for sex classification at 35 $\times$ 35 resolution. As can be seen, the central areas at the level of the eyebrows are most different between adult men and women.

VI. CONCLUSIONS AND FUTURE WORK

We have described a simple, but effective methodology for extracting pseudo-landmarks from 3D craniofacial data in the form of 3D meshes. Our method, unlike the competing methods, is fully automatic; it requires no hand-landmarking by humans, making it especially suitable for large data sets. Due to the availability of ground-truth data, we have used sex classification tasks to compare our methodology to the standard 24-landmark method, the Hammond dense correspondence method, and the Claes method. All three of these methods require prior hand-landmarking. Our experiments show that our method beats the standard 24-landmark method and is equivalent to the Hammond and Claes methods in accuracy, but much faster. Use of a distance matrix in place of the original points did not improve the accuracy of classification.

The purpose of running classification experiments in this paper was to test our features for validity and then adapt them for quantifying various conditions, either abnormal or, in the case of this dataset, normal variations. The methodology developed here for pseudo-landmark extraction is general in that pseudo-landmarks can be extracted from any area of the face and then used to describe the shape in that area. A content-based retrieval system is being developed that will use this methodology, among others, to retrieve heads of similar shape from the large 3D Facial Norms Database that is being collected. The methodology will also be used to group the heads in the database according to similar facial features.

VII. ACKNOWLEDGMENTS

This work was supported by the National Institute of Dental and Craniofacial Research under Grant No. U01-DE020050 and Grant No. U01-DE020078.

REFERENCES

[1] F. L. Bookstein, “Landmark Methods for Forms without Landmarks: Morphometrics of Group Differences in Outline Shape,” /it Mecial Image Analysis, Vol. 1, No. 3, 1997, pp. 225-243.

[2] P. Claes, M. Walters, D. Vandermeulen, J. G. Clement, “Spatially-Dense 3D Facial Asymmetry assessment in Both Typical and Disordered Growth,” *Journal of Anatomy*, Vol 219, Issue 4, 2011, pp. 444-455.

[3] S.-I. Lee, H. Lee, P. Abbeel, and A. Y. Ng, “Efficient L1 Regularized Logistic Regression, *Proceedings of the 21st National Conference on Artificial Intelligence*, 2006.

[4] H. J. Lin, S. Ruiz-Correa, L. G. Shapiro, D. Gatica-Perez, A. V. Hings, M. I. Cunningham, M. L. Speltz and R. W. Sze, “Symbolic Shape Descriptors for Classifying Craniosynostosis Deformations From Skull Imaging,” *IEEE Engineering in Medicine and Biology Society*, 2005, pp. 6325-6331.

[5] P. Hammond, T. J. Hutton, J. E. Allanson, L. E. Campbell, R. C. M. Hennekam, S. Holden, M. A. Patton, A. Shaw, I. K. Temple, M. Trotter, K. C. Murphy and R. M. Winter, “3D Analysis of Facial Morphology,” *American Journal of Medical Genetics*, Vol. 126A, 2004, pp. 339,348.

[6] H. Hochheiser et. al. “The FaceBase Consortium: A Comprehensive Program to Facilitate Craniofacial Research,” *Developmental Biology*, Vol. 355, No. 2, 2011, pp. 175-182.

[7] T. Hofmann, “Unsupervised learning by Probabilistic Latent Semantic Analysis,” *Machine Learning,* Vol. 42, 2001, pp. 177-196.

[8] F. Rohlf and D. Slice, “Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks,” *Systems Zoology*, Vol. 39, pp. 40-59.

[9] S. Ruiz-Correa, R. W. Sze, H. J. Lin, L. G. Shapiro, M. L. Speltz, M. L. Cunningham, “Classifying Craniosynostosis Deformations by Skull Shape Imaging,” *IEEE Computer-Based Medical Systems*, 2005, pp. 335-340.

[10] R. Tibshirani, “Regression Shrinkage and Selection Via the Lasso,” *Journal of the Royal Statistical Society B*, Vol. 58, 1994, pp. 267-288.

[11] M. Yuan and Y. Lin, “Model selection and estimation in regression with grouped variables,” *Journal of the Royal Statistical Society B*, Methodological, 68, 2006, pp. 49-67.

[12] K. Wilamowska, L. G. Shapiro, and C. L. Heike, “Classification of 3D Face Shape in 22q11.2 Deletion Syndrome,” *International Symposium on Biomedical Imaging*, 2009, pp. 537?537.

[13] S. Yang, L. G. Shapiro, M. L. Cunningham, M. Speltz, S.- I. Lee, “Classification and Feature Selection for Craniosynostosis,” *ACM Conference on Bioinformatics, Computational Biology and Biomedicine*, 2011, pp. 340-344.

[14] J. Wu, W. A. P. Smith, and E. R. Hancock, “Gender Discrimination Models from Facial Surface Normals,” *Pattern Recognition*, Vol. 44, 2011, pp. 2871-2886.