Anatomically Correct Surface Recovery: A Statistical Approach

Rasmus R. Jensen[(✉)], Jannik B. Nielsen,
Rasmus Larsen, and Rasmus R. Paulsen

DTU Compute, Technical University of Denmark, Richard Petersens Plads,
Building 324, 2800 Kgs. Lyngby, Denmark
{raje,jbol,rlar,rapa}@dtu.dk
http://www.compute.dtu.dk

Abstract. We present a method for 3D surface recovery in partial surface scans. The method is based on an Active Shape Model, which is used to predict missing data. The model is constructed using a bootstrap framework, where an initially small collection of hand-annotated samples is used to fit to and register unknown samples, resulting in an extensive statistical model. The statistical recovery uses a multivariate point prediction, where the distribution of the points is given by the Active Shape Model. We show how missing data in a partial scan, once point correspondence is achieved, can be predicted using the learned statistics. A quantitative evaluation is performed on a data set of 10 laser scans of ear canal impressions with minimal noise and artificial holes. We also present a qualitative evaluation on authentic partial scans from an actual direct in ear scanner prototype. Compared to a state-of-the-art surface reconstruction algorithm, the presented method gives matching prediction results for the synthetic evaluation samples and superior results for the direct scanner data.

Keywords: Surface recovery · Hole closing · Multivariate statistics · Shape modeling · In ear scanning · Active shape model

1 Introduction

Direct surface scanning of humans is an increasingly used modality where the applications include model creation in the entertainment industry, plastic surgery planning and evaluation, craniofacial syndrome evaluation [10,14], and in particular hearing aid production. In this paper, we are concerned with a particular surface shape namely that of the ear canal. Ear canal surface scans are used in custom hearing aid fitting. This is a very large industry that probably makes the ear the most scanned part of the human anatomy. A standard hearing aid producer generates more than a thousand scans per week. When producing custom in-the-ear devices like hearing aids and monitors, the standard routine is to inject silicone rubber into the patients ear and then laser scan this impression. While this technique normally creates complete surfaces, direct ear scanners are

emerging and it is expected that probe scans with these devices will require handling of missing data due to occlusion in the complex anatomy of the human ear and the limited space for the scanner probe. In this paper we present a method for predicting missing data based on the information in the partial scan. Hole filling and missing data recovery is a well studied problem, in particular for 2D images. In 3D, data recovery is sometimes considered a by-product of the surface reconstruction algorithm. The algorithms used to generate triangulated surfaces from point clouds will usually try to cover missing areas using some mathematical or physical assumptions. One series of approaches uses Delaunay triangulation of border points [12]. Such methods are obviously susceptible to noise in the border points and will typically require some form of smoothing. An alternative strategy is to interpolate implicit (signed distance) functions locally or globally under various forms of regularisation [11,18]. Other methods, inspired from 2D inpainting approaches have also been investigated [6,7,20]. These are typically based on a variational definition of the behavior of the surface where the holes are.

In our method, we predict the missing points based on the existing points in the scan. Instead of using variational formulations or physical assumptions on the behaviour of the surface, we utilise population statistics of the given class of surfaces learned from an annotated and co-registered training set. We base our population statistics of the ear canal on an extensive statistical shape model of the ear canal constructed in a bootstrap framework. The method is general and is applicable to all surface scans, where a statistical shape distribution can be estimated. The 3D morphable models introduced for the analysis and synthesis of 3D faces [5] can also be used to recover missing data in surface scans [4]. In [5] a 3D statistical shape and texture model is built based on a set of registered training samples and from this a principal component analysis is performed giving a set of eigenvectors and values. To recover missing data the set of known points are found in a pre-processing step and the missing data points are found by computing the optimal linear combination of eigenvectors fitting the known data. This is combined with a ridge regression regularisation to avoid non-plausible shapes. The approach described in [5] is similar to our prediction step, but in contrast we also include the steps needed to identify the missing points in the described framework. Furthermore, we also weight the geodesic distance from the missing points to the known points in the prediction.

1.1   Data and Preprocessing

The data consists of 310 scanned left-ear impressions. The scans have been acquired from a traditional 3D scanner, resulting in meshes of arbitrary tri-angulation. From this collection, 12 representatives are chosen and from these point correspondence over the selected impressions is created using the method initially described in [17]. Furthermore, the Markov Random Field regulariza-tion of the correspondence field described in [16] was used to further optimize the dense correspondence. This small subset of impressions with point correspon-dence form the basis for the bootstrapping framework. This is used to encompass

the entire collection of ear impressions, with the goal of constructing a statistical shape model as for example described in [8,17]. The method described in [17] requires manual annotation of anatomical landmarks which is non-trivial and therefore an automated method is preferred. A small collection of scans have been acquired by a prototype in-ear 3D scanner[1]. They are partial in the sense that some areas of the surface are missing due to noise and/or occlusion. Finally, a small set of scanned ear-impressions, not part of the original 310 samples, have had holes cut in them to mimic the nature of the partial scans. We denote these manually created partial scans as synthesized partial scans. This set is used for controlled evaluation of our method. In the following, some parameters have been assigned fixed values manually chosen for our data. These parameters should be validated for other uses of the framework.

2 Bootstrapped Active Shape Model

In order to accurately recover missing information in a partial scan we construct a statistical shape model [8]. For this, point correspondence is needed over the training set. Initially a small subset of samples is manually annotated and registered using the approach described in [16,17]. Using this subset, an Active Shape Model (ASM) is constructed as described in [8,17]. The statistical model is aligned and fitted to each unknown sample. This is done iteratively, allowing co-registration to and inclusion in the ASM, thereby expanding the model sample by sample. The ASM thus grows in size as the bootstrapping procedure processes unknown samples, allowing it to explain an increasing amount of shape variation from the dataset. Intuitively this leads to the expectation that the algorithm will become increasingly better at fitting to unknown shapes and that later samples are better registered than former, wherefore a revisit of early registrations may be chosen as a finalising step.

Assuming a collection of m aligned shapes, each consisting of p 3D points $\mathbf{v}_i = (x_1, y_1, z_1, \ldots, x_p, y_p, z_p)^T \in \mathbb{R}^n$. These shapes can be interpreted as being points in an $n = 3p$-dimensional space. The average shape is thus $\bar{\mathbf{v}} = \frac{1}{m} \sum_{i=1}^{m} \mathbf{v}_i$ and the shape deviation from mean $\mathbf{x}_i = \mathbf{v}_i - \bar{\mathbf{v}}$. In order to investigate the variation of the data, an observation matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_m) \in \mathbb{R}^{n \times m}$ can be constructed. The covariance matrix, $\mathbf{\Sigma}$, of $\mathbf{X}$ is found by

$$\mathbf{\Sigma} = \frac{1}{m} \mathbf{X}\mathbf{X}^T \in \mathbb{R}^{n \times n}. \tag{1}$$

Performing an Eigenvalue decomposition of this covariance matrix, thus provides insight in the primary modes of variation within the dataset $\mathbf{\Sigma} = \mathbf{P}\Lambda\mathbf{P}^T$, where $\mathbf{P} = (\mathbf{p}_1, \ldots, \mathbf{p}_m)$ is a matrix consisting of columns of Eigenvectors and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ is a diagonal matrix holding the Eigenvalues. These Eigenvalues corresponds to the variation expressed of the respective Eigenvector directions, i.e. $\lambda_i = \sigma_i^2$. In scenarios where $m < n$, only a subset of the Eigenvalues will be non-zero, the size of this subset will be denoted m'.

Given the collection of non-zero Eigenvalues and corresponding Eigenvectors described above, these can be used as a basis. Any shape $\tilde{\mathbf{v}}$ can then be synthesised by a linear combination of the Eigenvectors, weighted by their Eigenvalues:

$$\mathcal{M}(\mathbf{c}) = \tilde{\mathbf{v}} = \sum_{i}^{m'} c_i \lambda_i \mathbf{p}_i = \mathbf{P}\Lambda\mathbf{c}, \tag{2}$$

where $\mathbf{c} = (c_1, \ldots, c_{m'})$ is a vector of weights determining how much the individual Eigenvectors contributes in the synthesis. This constitutes the Active Shape Model and hereby the ASM can be interpreted as a function of the weights in $\mathbf{c}$, i.e. $\mathcal{M}(\mathbf{c})$.

The raw samples to be included in the ASM may not be positioned or oriented correctly relative to each other. Multiple approaches to automatic alignment of shapes exists, we have chosen to use 3D Shape Context Descriptors [9].

The descriptors describes a point on a 3D surface by a histogram of its local neighbourhood, indicating the local geometric distribution of points. Given a point $\mathbf{q}$ on a surface, any neighbouring point's relative position to $\mathbf{q}$ can be expressed in spherical coordinates (r, θ, φ). Here r is the radial distance between $\mathbf{q}$ and a neighbour $\mathbf{q}_n$. The inclination angle θ and the azimuthal angle φ requires a choice of reference-frame in order to be intercomparable between differently aligned samples. In this experiment, 3D data were acquired from a laser scanner using a rotating platform. The 3D representations of the ear impressions thus have a consistent vertical axis. This consistency can be utilised to construct a common frame of reference. In this frame of reference the third basis element is aligned with the normal of the point $\mathbf{q}$. This is formulated as $\mathbf{b}_3 = \mathbf{n_q} = (n_x, n_y, n_z)^T$. The first basis element is aligned with the vertical axis, with the restraint of being orthogonal to $\mathbf{b}_3$. Denoting a vector pointing along the fixed vertical axis $\mathbf{v} = (0, 1, 0)^T$, this is found by $\mathbf{b}_1 = \mathbf{v} - (\mathbf{v} \cdot \hat{\mathbf{b}}_3)\hat{\mathbf{b}}_3$, i.e. a vector rejection of $\mathbf{v}$ on $\mathbf{n_q}$, where $\hat{\mathbf{b}}$ denotes the normalised value of $\mathbf{b}$. As a result of orthogonal basis vectors in a right-handed coordinate-system, the second basis element is thus restrained to being $\mathbf{b}_2 = \mathbf{b}_3 \times \mathbf{b}_1$. From this basis, a rotation matrix, rotating to the local frame of reference can be constructed $\mathbf{R} = \begin{bmatrix} \hat{\mathbf{b}}_1 & \hat{\mathbf{b}}_2 & \hat{\mathbf{b}}_3 \end{bmatrix}$. Any neighbouring point, $\mathbf{q}_n$, can thus be described in $\mathbf{q}$'s local frame of reference by $\tilde{\mathbf{q}}_n = \mathbf{R}(\mathbf{q}_n - \mathbf{q})$. Within this frame of reference, the inclination angle and the azimuthal angle of the point is given by $\theta = \arccos(\tilde{q}_{n,z}/r)$ and $\varphi = \arctan(\tilde{q}_{n,y}/\tilde{q}_{n,x})$.

Based on the coordinates (r, θ, φ), points in the proximity of $\mathbf{q}$ can be grouped in a discrete set of bins. Hereby a histogram over the 3-dimensional distribution of points surrounding $\mathbf{q}$ can be constructed and used as a feature vector. In our experiment, (r, θ, φ) of points within a radius of 10 mm were divided into $(8, 13, 4)$ bins respectively, yielding a 416-dimensional feature vector or Shape Context Descriptor. The choice of utilising the vertical axis to construct a common frame of reference poses a constraint on the geometry as points having normals parallel to the vertical axis cannot be used. In practice this means that perfectly horizontal surfaces cannot be evaluated. Through the Hungarian method [13], point-descriptors

are matched and based on this matching a corresponding transformation can be computed in a least squares sense. The standard χ^2 histogram distance is used as cost function [3]. An illustration is shown in Fig 1. The resulting set of matched points is used to compute the optimal translation and rotation in a least squares sense [9].

Using this method, unknown samples are aligned to the mean of the ASM, $\bar{\mathbf{v}}$, thus supplying a plausible pre-alignment. Failed pre-alignments are easily detected by evaluating the average Euclidean point to point distance between the mean shape and the aligned shape. In our dataset, alignments with average point-to-point distances above 5 mm are rejected.
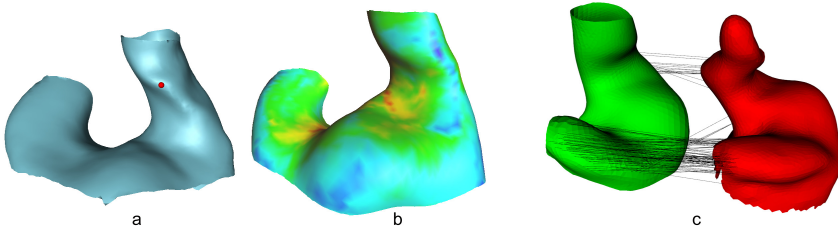

Fig. 1. The descriptor in the point marked by the red marker in a) is compared to the descriptors in all points of b). Colouring in b) corresponds to the χ^2 distance between the descriptors. c) Lines indicating the 100 most significant Shape Context matches between the two samples.

Given a roughly aligned unknown sample, $\mathbf{v}_a$, the alignment is refined and the ASM, $\mathcal{M}(\mathbf{c})$, is fitted. This is done in an iterative manner, where an Iterative Closest Point (ICP) [19] alignment of the sample is followed by an ASM-fitting of the model, and repeated upon until convergence is met. In our ICP implementation the points from the ASM surface are matched to their nearest surface neighbours on the new sample, with the constraint that points connected to a border should be ignored.

For the fitting, we seek to find a deformation of the Active Shape Model that minimises the error between the model and the unknown sample. An ASM constructed from the shape analysis of m samples, each consisting of n points, will be parametrised in an m'-dimensional space and thus have m' modes of variation.

Let $\mathbf{q}_i \in \mathbb{R}^3$ be a point belonging to the ASM, $\mathcal{M}(\mathbf{c})$, and let $\mathbf{q}_i' \in \mathbb{R}^3$ be the closest point on the target sample surface $\mathbf{v}_a$. We seek to find the set of weights $\mathbf{c}^*$ that minimises the sum of distances between $\mathbf{q}$ and $\mathbf{q}'$:

$$\arg \min_{\mathbf{c}} \|\mathcal{M}(\mathbf{c}) - \mathbf{v}_a\| = \arg \min_{\mathbf{c}} \frac{1}{m'} \sum_{i=1}^{m'} \|\mathbf{q}_i - \mathbf{q}_i'\| \tag{3}$$

We solve this optimisation problem by utilising an implementation of the Nelder-Mead method [15]. We reduced the number of parameters used to a number

corresponding to 99% explained variance af the training data. As the model-fitting is basically a synthesisation from a k-dimensional (assumed) normal-distribution, a confidence level for an obtained set of parameters $\mathbf{c}^*$ can thus be computed by utilising the Mahalanobis distance M between the parameter set and the ASM distribution since $M^2 \sim \chi_n^2$. This allows validation of fittings by setting a reasonable confidence limit. In our implementation, a confidence level of 99.9% was used.

Having determined $\mathbf{c}^*$ for $\mathcal{M}$ and $\mathbf{v}_a$, the model mesh $\mathcal{M}(\mathbf{c}^*) = \tilde{\mathbf{v}}$ is propagated to the sample shape $\mathbf{v}_a$ in order to perform a point-wise registration using the procedure described in [16]. The result is a mesh of p points, following the shape of $\mathbf{v}_a$, all with correspondence to the model, $\mathcal{M}$. The quality of the registration is determined by computing the 20th percentile of the minimum angle in the projected triangles and rejecting registration where this is below $15°$. This measure is valid since the model mesh has been optimised and has near equilateral triangles. Secondly, the normals of the projection are compared to the normals of $\tilde{\mathbf{v}}$. Registrations having an average dot-product between the normals of $\tilde{\mathbf{v}}$ and the projection below 0.75 are rejected.

When an alignment and a registration is obtained using the iterative scheme above, they are both refined iteratively. During each iteration the registration-mesh is smoothed using simple meaning of the nearest neighbours. The surface normals of the smoothed mesh are found and regularised using local averaging of directions. A new set of correspondence points are found in the sample scan in the direction of the regularised normals and the alignment, $\mathbf{v}_a$, is adjusted accordingly. This process is repeated until convergence. As the sample input scan is expected to be more densely sampled than $\mathcal{M}(\mathbf{c}^*)$ the iterative update ensures a regular mesh with evenly distributed vertices. Each sample that is successfully fitted is added to the ASM. ASM is hereby improved to cover additional shape variation.

3   Surface Recovery on Partial Scans

As described, a crucial, and not easily solved, part of recovering missing data is to co-register a partial mesh with the ASM. This is required in order to obtain point-correspondence between model and surface, creating a partial scan with a mesh structure identical to that of the model. The process of co-registering an unknown scan to the ASM is basically addressed in section 2. In the case of reconstructing partial scans, however, the exact same approach may not suffice. This is mainly due to the fact that an automatic alignment between a partial scan and model may prove to be difficult for Shape Context features. The difficulties arise in scenarios where the key shape features of the model are not present on the scan or vice-versa. We limit ourselves to the already existing Shape Context alignment approach, and where this failed, manual alignment was used. The result of registering and fitting the ASM to a partial scan is that the ASM template mesh is deformed and propagated to the partial scans in areas where there are valid data. The template mesh vertices are marked as missing when

the corresponding point or area in the partial scan is not present or valid. If an area is missing in the partial scan, the point projection will often result in that the project point is placed on a boundary in the partial scans, thus enabling detection of missing point correspondences. Given a registration, we aim to recover missing surface data in a partial scan such that the recovered data are anatomically correct. We approach this by using a statistical model and define the set of known and unknown data in a partial scan as follows:

$$\text{missing vertices: } \mathbf{s}_1^T = (x_{11}, y_{11}, z_{11}, x_{12}, y_{12}, z_{12}, \ldots)$$
$$\text{known vertices: } \mathbf{s}_2^T = (x_{21}, y_{21}, z_{21}, x_{22}, y_{22}, z_{22}, \ldots)$$

The correspondence allows for differentiation between known vertices and missing vertices in the partial scan. We will determine how the unknown data $\mathbf{s}_1$ are predicted from known vertices in $\mathbf{s}_2$. Without any prior knowledge of the distribution of data, we consider a shape $\mathbf{s}$ consisting of $\mathbf{s}_1$ and $\mathbf{s}_2$ as belonging to the normal distribution:

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} \in N \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix} \right) \quad, \quad \mathbf{\Sigma}_{12}^T = \mathbf{\Sigma}_{21} \tag{4}$$

The expected value of $\mathbf{s}_1$ given $\mathbf{s}_2$ is $\text{E}\{\mathbf{s}_1|\mathbf{s}_2\} = \mu_1 + \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}(\mathbf{s}_2 - \mu_2)$. With the variance $\text{V}\{\mathbf{s}_1|\mathbf{s}_2\} = \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}$ From the ASM we get an aligned set of shapes. This training set is denoted $X_{aligned}$. From the training set the covariances $\mathbf{\Sigma}_{11}$ and $\mathbf{\Sigma}_{12}$ as well as the means (μ_1, μ_2) are learned. As there are far less shapes than points, $\mathbf{\Sigma}_{22}$ will be singular. Let $\mathbf{\Sigma}_{22} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^T$ be the Eigenvalue decomposition. We restrict $\mathbf{\Sigma}_{22}$ to its affine support, i.e. the dimensions spanned by the Eigenvectors corresponding to the k positive Eigenvalues, such that $\mathbf{\Lambda}^* = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$ and $\mathbf{P}^* = [\mathbf{p}_1 \; \mathbf{p}_2 \; \ldots \; \mathbf{p}_k]$. The projection of $\mathbf{s}_2$ using the k selected Eigenvectors $\mathbf{P}^*$: $\mathbf{y}_2 = \mathbf{P}^{*T}\mathbf{s}_2$ has affine support for $\mathbf{s}_2$ and the variance:

$$\text{V}\{\mathbf{y}_2\} = \text{V}\{\mathbf{P}^*\mathbf{s}_2\} = \mathbf{P}^{*T}\mathbf{\Sigma}_{22}\mathbf{P}^* = \mathbf{\Lambda}^* \tag{5}$$

The covariance of $\mathbf{s}_1$ and $\mathbf{y}_2$ is:

$$\text{C}\{\mathbf{s}_1, \mathbf{y}_2\} = \text{C}\{\mathbf{s}_1, \mathbf{P}^{*T}\mathbf{s}_2\} = \text{C}\{\mathbf{s}_1, \mathbf{s}_2\}\mathbf{P}^* = \mathbf{\Sigma}_{12}\mathbf{P}^* \tag{6}$$

Finally, the prediction of the unknown data $\mathbf{s}_1$ can be done using the projection $\mathbf{y}_2$:

$$\text{E}\{\mathbf{s}_1|\mathbf{y}_2\} = \mu_1 + \mathbf{\Sigma}_{12}\mathbf{P}^*\mathbf{\Lambda}^{*-1}\mathbf{P}^{*T}(\mathbf{s}_2 - \mu_2) \tag{7}$$

This expected value can be used for any unknown set of vertices $\mathbf{s}_1$ given a partial scan $\mathbf{s}_2$, be that a single missing vertex or all the missing data. If every unknown vertex is predicted according to the described method the known triangulation from the training set can be propagated to the predicted data set and will then constitute a full surface reconstruction. The method can also be used to filter data for noise if the known scan data are also recovered. By varying the number

k values of Eigenvectors used in the projection the fraction of described variance can be controlled.

Practically, the set of known vertices found during the registration of the partial scan is $\mathbf{s}_2$ and the full scan as provided by the scanner is $\mathbf{s}_{scan}$. Let $\mathbf{s}_1^*$ be the predicted missing data, $\mathbf{s}_2^*$ the prediction of the partial scan and $\mathbf{s}^*$ be the full reconstructed shape. The full average shape is denoted μ. With an initial registration the algorithm works as follows: We repeat the loop body with two

---

Algorithm 1. Anatomical surface recovery

1: **procedure** ANATOMICALSURFACERECOVERY$(X_{aligned}, s_2, s_{scan})$ ▷
2:    $\mathbf{s}^* \leftarrow \begin{bmatrix} \mathbf{0} \\ \mathbf{s}_2 \end{bmatrix}$
3:    **repeat**
4:       Procrustes align $\mathbf{s}^*$ to μ and apply same transformation to $\mathbf{s}_{scan}$
5:       Predict $\mathbf{s}_1^*$ using the described method   ▷ the described variance is increased in each iteration
6:       Predict $\mathbf{s}_2^*$ using the described method   ▷ the described variance is increased in each iteration
7:       $\mathbf{s}^* \leftarrow \begin{bmatrix} \mathbf{s}_1^* \\ \mathbf{s}_2^* \end{bmatrix}$
8:       Find vertex correspondence between $\mathbf{s}^*$ and $\mathbf{s}_{scan}$
9:       Update $\mathbf{s}_2$ and $\mathbf{s}^*$ with the correspondence vertices from $\mathbf{s}_{scan}$
10:   **until** convergence
11:   **return** $\mathbf{s}_1^*$ and $\mathbf{s}_2^*$ ▷
12: **end procedure**

---

different recovery approaches. First $\mathbf{s}_1^*$ and $\mathbf{s}_2^*$ are predicted all at once. In the last few loops the data are predicted vertex by vertex using only the nearest vertices in the prediction. The vertex distances are found as the geodesic distances on the mean shape, so these only have to be calculated once. The geodesic distances are used to ensure topological consistency when selecting a neighbourhood. Our shape model has 3096 vertices and in the local recovery we only use the 10 nearest of these. In the local prediction the recovered data is locally very true to the original scan. We restrict the Eigenvalues in the recovery to the ones describing 30% of the variance and then gradually raise this to 99.9%. Gradually raising the percentage of described variance helps the algorithm produce anatomically correct shapes and prevents the influence of bad correspondences in the initial iterations.

4   Results

Based on the method described in section 2, we were able to construct an extensive Active Shape Model of the left ear based on the available dataset. A total of 310 samples were processed and from these, 241 passed the automatic quality verification. As the Active Shape Model processes new samples, the complexity

of the model increases and thus the fraction of variation explained per principal component must be expected to drop. The fraction of variation explained by the 10 first principal components were computed as the ASM grew in size and it was stabilising, indicating that the shape model eventually captures the true class variability. In the ASM 90% of all variance is contained within the first 37 modes of variation. We do, however, expect that the automatic registration procedure has induced an amount of false variation in form of vertex drifting along the sample surfaces. Such variation of course directly affects the compactness of the ASM in the form of low-variance principal components, assuming that the drifts are uncorrelated. The actual shape variation from the ear is therefore expected to be found within the first principal components. A final manual inspection of mesh distortions resulted in an additional 80 registrations being removed from the ASM. Effectively this resulted in the final ASM consisting of 161 shapes.

In order to compare our approach with existing methods for reconstruction, a collection of 10 scanned ear impressions, not included in the training data, was chosen and all scans had a reasonable sized hole cut in them. The holes were cut between first and second bend of the ear canal, in an area that is known to often be occluded when using experimental optical in-ear scanners. Hereby any reconstruction of these partial scans can be compared to the ground-truth, allowing for a quantitative comparison of methods.

For each mesh in the collection of *synthesised* partial scans the missing data was recovered. This was done using our method, both with and without smoothing, and the Markov Random Field (MRF) surface reconstruction approach [18]. The MRF approach has previously shown to reconstruct anatomical surfaces well. All reconstructions were then compared to the ground truth, by computing a signed distance (based on surface normals) between all reconstructed points and the original surface. In Fig. 2 the reconstructions of synthetic partial scans are shown, where the surface values denotes the signed distance between reconstruction and truth (in millimeters). The average signed point-distances between surface reconstructions can be seen in the table below:

#	MRF	Proposed	Proposed +local smoothing	#	MRF	Proposed	Proposed +local smoothing
1	-0.07	-0.17	-0.27	6	0.07	-0.05	-0.08
2	0.05	0.01	-0.01	7	0.08	-0.001	0
3	-0.04	0.02	0.001	8	0.01	-0.02	-0.03
4	-0.001	-0.015	-0.02	9	0.02	0.03	0.01
5	-0.07	-0.01	-0.013	10	-0.01	0.07	0.07

A significant outlier in the error is observed in sample #1. After inspection, this sample revealed an abnormal cavity in the skin of the ear-canal, explaining the higher error. It should be noted that no prior, neither statistical or physical, would be able to predict such errors. Although this comparison proves high performance of our method, it does not fully illustrate the strength of having a statistically based prior. The MRF approach predicts missing points based on the existing curvature of data in contrast to our method that predicts missing
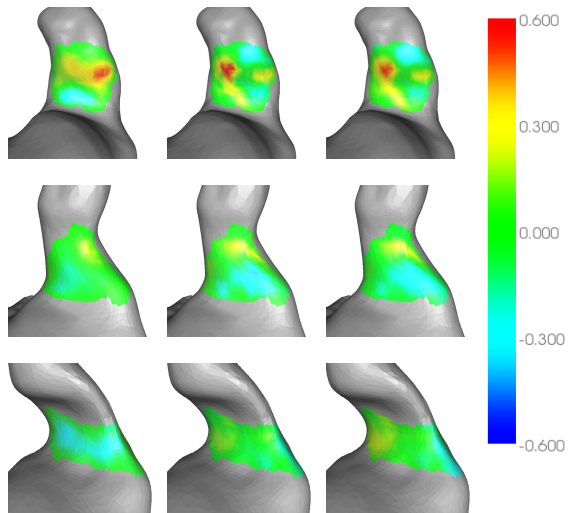
Fig. 2. Reconstruction of missing data for 3 different scans (rows), using Markov Random Field (MRF) reconstruction (column 1), our method (column 2) and the smoothed variant of our method (column 3). Surface values corresponds to the signed distance between reconstruction and ground truth.
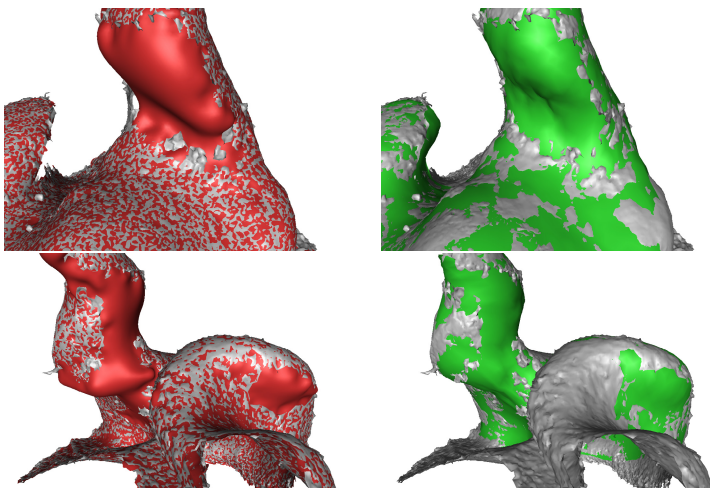

Fig. 3. Partial scans from a prototype direct ear scanner. Raw data is shown with grey. Surfaces reconstructed using the MRF method are red and the surfaces reconstructed using the proposed method are green.

points based on knowledge of the shape variation of an ear population. Effectively this means that where either noisy edges exists or data is sparse, the MRF approach has little chance of estimating the true surface. In this test each hole is surrounded by smooth noiseless surface areas providing an optimum setting for the MRF reconstruction. In the following, we will present a qualitative comparison based on authentic optical 3D scans of the ear, suffering from high noise and sparse point support.

We have tested our algorithm on 12 scans from a prototype direct in ear scanner [1]. In cases with a lot of noise, our strong prior enables our method to produce anatomically correct meshes that are also locally true to the covered areas. Qualitative inspection shows very good hole closing in the 12 scans. In addition all 12 scans were 3D printed as earplugs and tested by the respective test subjects with positive feedback. Fig 3 shows scans with a big part of the ear canal missing. The missing part has been recovered with both the MRF method and the proposed method. As can be seen, our proposed method produces what seems to be a much more plausible surface in the missing part.

5   Conclusion

We have shown that we can predict the missing parts of partial scans using a statistical model. The ability to predict missing data is comparable to state-of-the-art algorithms, when holes are relatively small and the data is fair without too much noise. On scans from a real in-ear scanner probe prototype, the qualitative results produced with the proposed method are much more plausible when visually inspected. The more extensive prior knowledge about the shape to be reconstructed makes the recovery much more robust, when recovering larger holes. The results also seem invariant to the presence of noise, and as such the method can also function as a noise filter. Surface reconstruction algorithms that only use the immediate vicinity in the reconstruction are very sensible to noise on the edges of the area to be recovered.

After using the proposed data recovery method on 12 scans they were 3D printed on a stereolithograpy (SLA) machine [2] and worn by the test subjects for a substantial time. They all proved to be well fitting in the subjects ears even though the hard material from the SLA machine makes the ear plugs very susceptible for non-accurate fitting. We have therefore demonstrated a complete pipeline from direct ear scanning to production of well fitting hearing devices.

References

1. 3Shape, Copenhagen. http://www.3Shape.com
2. Projet MP3000, 3Dsystems. http://www.3DSystems.com/

3. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. on Pattern Analysis and Machine Intelligence **24**(4), 509–522 (2002)
4. Blanz, V., Mehl, A., Vetter, T., Seidel, H.P.: A statistical method for robust 3D surface reconstruction from sparse data. In: Proc. 3D Data Processing, Visualization and Transmission, pp. 293–300 (2004)
5. Blanz, V., Vetter, T.: A morphable model for the synthesis of 3D faces. In: Proc. 26th Annual Conference on Computer Graphics and Interactive Techniques, pp. 187–194 (1999)
6. Caselles, V., Haro, G., Sapiro, G., Verdera, J.: On geometric variational models for inpainting surface holes. Computer Vision and Image Understanding **111**(3), 351–373 (2008)
7. Clarenz, U., Diewald, U., Dziuk, G., Rumpf, M., Rusu, R.: A finite element method for surface restoration with smooth boundary conditions. Computer Aided Geometric Design **21**(5), 427–446 (2004)
8. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models - their training and application. Computer Vision and Image Understanding **61**(1), 38–59 (1995)
9. Frome, A., Huber, D., Kolluri, R., Bülow, T., Malik, J.: Recognizing objects in range data using regional point descriptors. In: Pajdla, T., Matas, J.G. (eds.) ECCV 2004. LNCS, vol. 3023, pp. 224–237. Springer, Heidelberg (2004)
10. Hammond, P., Hutton, T., Allanson, J., Campbell, L., Hennekam, R., Holden, S., Patton, M., Shaw, A., Temple, I., Trotter, M., et al.: 3D analysis of facial morphology. American Journal of Medical Genetics Part A **126**(4), 339–348 (2004)
11. Kazhdan, M., Bolitho, M., Hoppe, H.: Poisson surface reconstruction. In: Proc. Eurographics Symposium on Geometry Processing, pp. 61–70 (2006)
12. Kolluri, R., Shewchuk, J., O'Brien, J.: Spectral surface reconstruction from noisy point clouds. In: Proc. Eurographics Symposium on Geometry Processing, pp. 11–21 (2004)
13. Kuhn, H.W.: The Hungarian method for the assignment problem. Naval Research Logistic Quarterly **2**, 83–97 (1955)
14. Lanche, S., et al.: A statistical model of head asymmetry in infants with deformational plagiocephaly. In: Ersbøll, B.K., Pedersen, K.S. (eds.) SCIA 2007. LNCS, vol. 4522, pp. 898–907. Springer, Heidelberg (2007)
15. Nelder, J.A., Mead, R.: A simplex method for function minimization. The Computer Journal **7**(4), 308–313 (1965)
16. Paulsen, R.R., Hilger, K.B.: Shape modelling using markov random field restoration of point correspondences. In: Taylor, C., Noble, J.A. (eds.) IPMI 2003. LNCS, vol. 2732, pp. 1–12. Springer, Heidelberg (2003)
17. Paulsen, R., Larsen, R., Nielsen, C., Laugesen, S., Ersbøll, B.: Building and testing a statistical shape model of the human ear canal. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002. LNCS, vol. 2489, pp. 373–380. Springer, Heidelberg (2002)
18. Paulsen, R., Bærentzen, J., Larsen, R.: Markov random field surface reconstruction. IEEE Trans. on Visualization and Computer Graphics **16**(4), 636–646 (2010)
19. Rusinkiewicz, S., Levoy, M.: Efficient variants of the icp algorithm. In: Proc. 3D Digital Imaging and Modeling, pp. 145–152 (2001)
20. Verdera, J., Caselles, V., Bertalmio, M., Sapiro, G.: Inpainting surface holes. In: Proc. of International Conference on Image Processing, vol. 2 (2003)