Optimal Step Nonrigid ICP Algorithms for Surface Registration

Brian Amberg
University of Basel
brian.amberg@uni-basel.ch

Sami Romdhani
University of Basel
sami.romdhani@uni-basel.ch

Thomas Vetter
University of Basel
thomas.vetter@uni-basel.ch

Abstract

We show how to extend the ICP framework to nonrigid registration, while retaining the convergence properties of the original algorithm. The resulting optimal step nonrigid ICP framework allows the use of different regularisations, as long as they have an adjustable stiffness parameter.

The registration loops over a series of decreasing stiffness weights, and incrementally deforms the template towards the target, recovering the whole range of global and local deformations.

To find the optimal deformation for a given stiffness, optimal iterative closest point steps are used. Preliminary correspondences are estimated by a nearest-point search. Then the optimal deformation of the template for these fixed correspondences and the active stiffness is calculated. Afterwards the process continues with new correspondences found by searching from the displaced template vertices.

We present an algorithm using a locally affine regularisation which assigns an affine transformation to each vertex and minimises the difference in the transformation of neighbouring vertices. It is shown that for this regularisation the optimal deformation for fixed correspondences and fixed stiffness can be determined exactly and efficiently.

The method succeeds for a wide range of initial conditions, and handles missing data robustly. It is compared qualitatively and quantitatively to other algorithms using synthetic examples and real world data.

1. Introduction

Registering two surfaces means finding a mapping between a template surface and a target surface that describes the position of semantically corresponding points. This can be considered as warping the template onto the target. To choose the "correct" deformation from all possible warps, a registration algorithm has to impose constraints on the deformation. In this context, this is called regularisation of the deformation field. We use regularisation based on minimizing the difference between transformations acting on neighbouring vertices of a mesh. The "correct" registration is not defined uniquely for all points. The correspondence for a point on the cheek might not be uniquely identifiable even by a human, while a point like the corner of an eye can be determined even across subjects.

This work was supported in part by Microsoft Research through the European PhD Scholarship Programme.

Figure 1. Caricatures help to judge the registration quality by exaggerating flaws and artifacts. The marked left and right images are registered, hole-filled scans that have been extrapolated and interpolated from the mean head (center). Extrapolated positions are automatic caricatures.

Dense registration methods find a mapping from each point in the template onto the target while sparse methods find correspondence only for selected feature points. We present a dense registration method.

Registration can be used to build a morphable model, as introduced by Blanz and Vetter [3]. This is done by registering the same template onto multiple targets, yielding a consistent parametrization over all scans, which then form a linear object class.

For noiseless and complete data, a correct registration should be one-to-one. In practice the surfaces contain holes and artefacts resulting from the scanning process. Therefore a useful registration method needs to be robust against outliers and has to fill in missing data in a sensible way. Our input data is acquired with a structured light 3D scanning system. An example of a typical scan, the template mesh used, and the registration result is given in Figure 2.

2. Prior work

Numerous registration algorithms exist, each applicable to different scenarios and having advantages and drawbacks. Our specific requirement is to match tessellated 3D-meshes of similar shapes.

Our scans contain large missing regions, generally more

1-4244-1180-7/07/$25.00 ©2007 IEEE

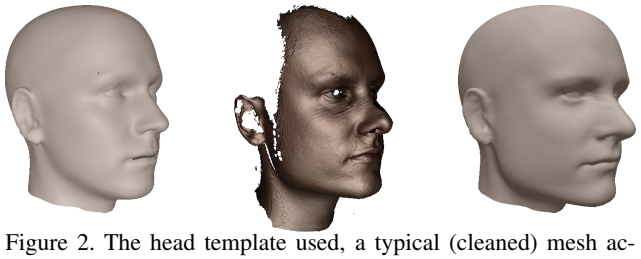

Figure 2. The head template used, a typical (cleaned) mesh acquired by our scanning process, and the registration result.

than half of the surface is missing (Figure 2). These regions have to be filled in a sensible way, incorporating knowledge encoded in the template mesh. This is different from hole-filling methods as described e.g. by Davis et al. [5], which are able to close holes by interpolating between the borders. Instead we fill in details by smoothly deforming the template mesh. Methods which extend the deformation field into regions without correspondence are found in Kähler et al. [12] or Szeliski and Lavallée [14].

Ferrant et al. [9] proposed to regularise based on material properties. We do not focus on this approach as our surfaces are acquired from different subjects and the difference between them does not correspond to a physical deformation.

When registering parts of the face which may move independently, like the lower and upper lip, smoothing isotropically throughout the volume as in [12, 14] has the unwanted effect of tying together unconnected parts. In this paper, this is addressed by using anisotropic regularisation which propagates along the surface, as proposed by Allen et al. [1].

To choose a mapping out of all possible registrations the allowable deformations are constrained by regularisation of the deformation field. A common approach is to smooth the deformation field by minimizing the squared norm of its gradient, effectively allowing locally smooth translations. An efficient solution for volumetric datasets has been proposed by Fisher and Modersitzki [10]. As typical deformations in faces are mixtures of rotations and translations – e.g. lowering the jaw – we search for a regulariser which favours rigid deformations. Curvature based registration as described in Fisher and Modersitzki [11] minimises the Laplacian of the deformation field, allowing for locally smooth affine transformations, a superset of rigid deformations. For surfaces, Feldmar and Ayache [8], and more recently Allen et al. [1] describe a regularisation which uses locally different affine transformations. While different regularisers can be incorporated into our framework, the main algorithm presented here uses a regularisation similar to [1].

Iterative Closest Point methods find a preliminary set of correspondences by searching for the closest points to the template vertices on the target surface, find a transformation which aligns the template to these correspondences and start again with a new set of correspondences found by searching from the vertices of the displaced template. ICP methods are distinguished by the type of deformation they recover, and by the way in which these deformations are found. Chen and Medioni [4] and Besl and MacKay [2] estimate a global rigid transformation. In these methods each step of the iteration is optimal with respect to the fixed correspondences. The difference in the two methods is that in [4] preliminary correspondences are found only along the surface normal, while [2] uses the closest point. Feldmar and Ayache [8] introduce locally different transformations. They choose spherical subsets of the template volume and determine an affine transformation within this region. To get a smooth deformation field the resulting affine transformations are blended linearly between the sphere centers. As the affine transformations are determined independently per sphere, the resulting (blended) deformation is no longer optimal with respect to the fixed correspondences.

When determining an affine transformation for a set of vertices as proposed by Feldmar and Ayache [8], the affine transformation is uniquely determined by the correspondences. Allen et al. [1] propose to determine one affine transformation per vertex. This is surprising, as a single correspondence does not uniquely determine an affine transformation. To further constrain the problem, a regulariser, which can be interpreted as a stiffness term, is added. It forces neighboring vertices to undergo similar transformations. The open question whether or not this is enough to constrain the problem is addressed in our paper.

3. Contributions

We define the optimal step nonrigid ICP framework, which extends ICP methods to nonrigid deformations while retaining the convergence properties of ICP. Different regularisations can be used in this framework. A version based on the locally affine regularisation from [1], and an algorithm using locally smooth translations is developed. Our algorithm improves upon the method in [1] in multiple aspects. Already in the seminal paper of Besl and MacKay [2] it was observed that finding the optimal deformation for a given correspondence is more accurate and more efficient than handing the distance cost to a general optimization method. This was abandoned in [1]. Instead a general Newton-type optimiser (L-BFGS-B, Zhu [15]) was used on the cost function. We show how for the regulariser used the optimal transformation for a given set of correspondences and a fixed stiffness can be determined, allowing us to make optimal steps. This overcomes the numerical instabilities introduced by using a general optimiser with a localised stiffness term, leading to a better registration. It is shown that a method using a general "black-box" optimiser has a very small convergence basin and fails on incomplete surfaces, if the initialization is not already nearly perfect. Our algorithm is robust against bad initialization and handles incomplete surfaces very well.

4. Method

In this section the regularisation used in this paper is introduced, and nonrigid optimal step ICP algorithms are defined. The template $\mathcal{S} = (\mathcal{V}, \mathcal{E})$ is given as a set of n vertices $\mathcal{V}$ and a set of m edges $\mathcal{E}$. The target surface $\mathcal{T}$ can be given in any representation that allows to find the closest point on the surface for any point in 3D-space. We use a triangulated mesh. Registration means finding parameters $\boldsymbol{X}$ describing a set of displaced source vertices $\mathcal{V}(\boldsymbol{X})$. After registration $\mathcal{V}(\boldsymbol{X})$ is projected onto the target surface along the normals of the deformed template to give the final correspondences. The projected vertices define – together with the original topology of the template mesh – a reparametrised version of the original scan.

Locally affine regularisation The cost function used in this paper to determine the warping is similar to the one defined in [1]. The difference is, that by expressing the algorithm in the nonrigid optimal step ICP framework we are able to simplify it for fixed correspondences such that it is a quadratic function which can be solved directly and exactly. Additionally a slightly different norm, which includes the norm from [1] as a special case, is used.

The proposed parametrization of the mapping is one affine 3×4 transformation matrix $\boldsymbol{X}_i$ *per template vertex*. The unknowns are organised in a $4n \times 3$ matrix

$$\boldsymbol{X} := \begin{bmatrix} \boldsymbol{X}_1 & \cdots & \boldsymbol{X}_n \end{bmatrix}^T \quad . \tag{1}$$

Naturally the distance between the deformed template and the target should be small. This is expressed by the first term of the cost function used in this paper:

$$E_d(\boldsymbol{X}) := \sum_{\boldsymbol{v}_i \in \mathcal{V}} w_i \operatorname{dist}^2(\mathcal{T}, \boldsymbol{X}_i \boldsymbol{v}_i) \tag{2}$$

To improve readability, we assume that template vertices are given in homogeneous coordinates $\boldsymbol{v}_i = \begin{bmatrix} x, y, z, 1 \end{bmatrix}^T$. The distance between a point $\boldsymbol{v}$ and its closest point on the target surface is denoted as $\operatorname{dist}(\mathcal{T}, \boldsymbol{v})$. A hierarchical bounding spheres structure is used to speed up nearest point search to $O(\log_2 t)$ in the number of target triangles. The reliability of the match is weighted by w_i. The weights are set to zero for vertices where no corresponding vertex could be found. For the other vertices the weight is set to one. If an additional estimate of the reliability is available, e.g. from the scanner, the weights can be set accordingly.

An additional stiffness term is used to regularise the deformation. We penalise the weighted difference of the transformations of neighbouring vertices under the Frobenius norm $\|\cdot\|_F$ using a weighting matrix $\boldsymbol{G} := \operatorname{diag}(1, 1, 1, \gamma)$.

$$E_s(\boldsymbol{X}) := \sum_{\{i,j\} \in \mathcal{E}} \|(\boldsymbol{X}_i - \boldsymbol{X}_j)\boldsymbol{G}\|_F^2 \tag{3}$$

γ can be used to weight differences in the rotational and skew part of the deformation against the translational part of the deformation. The choice of γ depends on the units of the data, and on the type of deformation that shall be expressed. In the experiments presented here γ was set to one and the data was scaled into the $[-1, 1]^3$ cube.

While the displacement of a vertex can be described by only three parameters, we use twelve parameters per vertex. This will allow us to write the cost as a quadratic function. Constructing a cost function for this regularisation term with only three parameters per vertex results in a minimization problem which can not be solved directly.

The third contributor to the cost function is a simple landmark term, used for initialization and guidance of the registration. Given a set of landmarks $\mathcal{L} = \{(\boldsymbol{v}_{i_1}, \boldsymbol{l}_1), \ldots, (\boldsymbol{v}_{i_l}, \boldsymbol{l}_l)\}$ mapping template vertices into the target surface the landmark cost is defined as

$$E_l(\boldsymbol{X}) := \sum_{(\boldsymbol{v}_i, \boldsymbol{l}) \in \mathcal{L}} \|\boldsymbol{X}_i \boldsymbol{v}_i - \boldsymbol{l}\|^2 \quad . \tag{4}$$

As demonstrated in Section 5, the correct registration can be found even without landmarks. Without landmarks the cost function has global minima where the template is collapsed onto a point on the target surface, but the local minimum corresponding to the correct registration can be found for a wide range of initial conditions.

The full cost function is a weighted sum of these terms

$$E(\boldsymbol{X}) := E_d(\boldsymbol{X}) + \alpha E_s(\boldsymbol{X}) + \beta E_l(\boldsymbol{X}) \quad . \tag{5}$$

The stiffness weight α influences the flexibility of the template, while the landmark weight β is used to fade out the importance of the potentially noisy landmarks towards the end of the registration process.

Contrary to [1] we use a modified ICP algorithm to efficiently and accurately minimise Equation (5).

4.1. Nonrigid optimal step ICP algorithms

The following steps constitute a nonrigid optimal step ICP algorithm:

- Initialise $\boldsymbol{X}^0$.
- For each stiffness $\alpha^i \in \{\alpha^1, \ldots, \alpha^n\}$, $\alpha^i > \alpha^{i+1}$
  - Until $\|\boldsymbol{X}^j - \boldsymbol{X}^{j-1}\| < \varepsilon$
    - Find preliminary correspondences for $\mathcal{V}(\boldsymbol{X}^{j-1})$.
    - Determine $\boldsymbol{X}^j$ as the optimal deformation for the preliminary correspondences and α^i.

It consists of two loops. The outer loop finds a series of deformations of the template that bring the template ever closer to the target. Starting with a strongly regularised (stiff) deformation global alignment is recovered and then successively lower stiffness weights are used, allowing for more localised deformations.

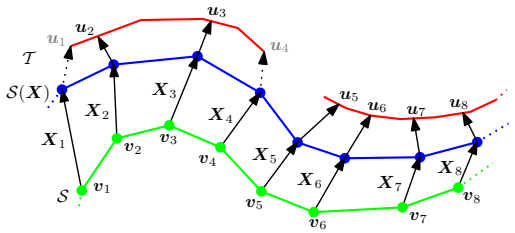

Figure 3. The template surface $\mathcal{S}$ (green) is deformed by locally affine transformations ($\boldsymbol{X}_i$) onto the target surface $\mathcal{T}$ (red). The algorithm determines closest points ($\boldsymbol{u}_i$) for each displaced source vertex ($\boldsymbol{X}_i\boldsymbol{v}_i$) and finds the optimal deformation for the stiffness used in this iteration. This is repeated until a stable state is found. The process then continues with a lower stiffness. Due to the stiffness constraint the vertices do not move directly towards the target surface, but may move parallel along it. The correspondences $\boldsymbol{u}_1$ and $\boldsymbol{u}_4$ are dropped as they lie on the border of the target.

Finding a deformation for a fixed stiffness is done in the inner loop. Preliminary correspondences are found by a nearest point search. Then the optimal deformation of the template for these correspondences and the fixed stiffness is determined. Due to the stiffness of the template the points do not move directly towards their preliminary correspondences, but may move parallel to the target surface. The new template position gives rise to a new set of preliminary correspondences, which are used in the next iteration. This is repeated until the process converges. The stiffness is then lowered and the search continued.

Figure 3 shows a detail of an intermediate step of the registration. The template has moved towards the target, but due to the stiffness constraint it is not yet inside the target surface. The correspondence points $\boldsymbol{u}_i$ of the deformed template $\mathcal{S}(\boldsymbol{X})$ change in each iteration.

Nonrigid optimal step ICP algorithms can be constructed for different regularisers, as long as the regulariser has a stiffness parameter determining the amount of acceptable local deformation. For the locally affine regularisation introduced in the first part of this section we show now how to determine the optimal deformation for fixed stiffness and correspondence. In Section 5 another regulariser is introduced and their performance is compared.

4.2. Efficient and robust minimization

Contrary to the approach of [1] we find in each step the optimal deformation for fixed stiffness and fixed correspondence. This is superior to using a "black-box"-optimiser directly on Equation (5), as the stiffness term is defined only over neighbouring vertices, and propagation of stiffness information over multiple edges is slow and fragile. While a general optimiser uses the fixed correspondences only to determine approximate gradients which influence the new search direction, we show how to make full use of the correspondence information by finding the optimal deformation in each step of the algorithm.

When correspondences are fixed, the cost function becomes a sparse quadratic system which can be minimised exactly. To adapt the cost function to fixed correspondences, only the first term has to be changed. In this section Equation (5) is rewritten for fixed correspondences using matrix expressions.

$$\bar{E}(\boldsymbol{X}) := \bar{E}_d(\boldsymbol{X}) + \alpha E_s(\boldsymbol{X}) + \beta E_l(\boldsymbol{X}) \tag{6}$$

By repeated minimization of $\bar{E}$ a local minimum of E is found.

Distance Term Assuming fixed correspondences $(\boldsymbol{v}_i, \boldsymbol{u}_i)$, the distance term defined in Equation (2) becomes

$$\begin{aligned}\bar{E}_d(\boldsymbol{X}) &:= \sum_{\boldsymbol{v}_i \in \mathcal{V}} w_i \|\boldsymbol{X}_i\boldsymbol{v}_i - \boldsymbol{u}_i\|^2 \qquad (7)\\ &= \left\| (\boldsymbol{W} \otimes \boldsymbol{I}_3) \left(\begin{bmatrix} \boldsymbol{X}_1 & & \\ & \ddots & \\ & & \boldsymbol{X}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_n \end{bmatrix} - \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_n \end{bmatrix} \right) \right\|^2\end{aligned}$$

where $\boldsymbol{W} := \operatorname{diag}(w_1, \ldots, w_n)$. (The $n \times n$ identity matrix is denoted by $\boldsymbol{I}_n$ and the Kronecker product by $\otimes$). Recall that the unknowns are the submatrices $\boldsymbol{X}_i$. As this formulation does not lend itself to being differentiated easily we bring the equation into canonical form by swapping the positions of the unknowns and the fixed vertices $\boldsymbol{v}_i$. We define the sparse matrix $\boldsymbol{D}$ mapping the $4n \times 3$ matrix of unknowns $\boldsymbol{X}$ onto displaced source vertices as

$$\boldsymbol{D} := \begin{bmatrix} \boldsymbol{v}_1^T & & & \\ & \boldsymbol{v}_2^T & & \\ & & \ddots & \\ & & & \boldsymbol{v}_n^T \end{bmatrix} \quad . \tag{8}$$

Arranging the correspondence points in a matrix $\boldsymbol{U} := \left[\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\right]^T$, the distance term can be written as:

$$\bar{E}_d(\boldsymbol{X}) = \|\boldsymbol{W}\left(\boldsymbol{D}\boldsymbol{X} - \boldsymbol{U}\right)\|_F^2 \quad . \tag{9}$$

The identity of these expressions comes directly from the definition of the matrix product.

Stiffness Term The stiffness term penalises differences between the transformation matrices assigned to neighbouring vertices. To express this in matrix notation, we use the node-arc incidence matrix $\boldsymbol{M}$ (e.g. Dekker [7]) of the template mesh topology. This matrix is defined for directed graphs. It contains one row for each arc (edge) of the graph and one column per node (vertex). To construct a node-arc incidence matrix from the source topology, the edges

and vertices of the mesh are numbered and its edges are directed from the lower numbered vertex to the higher numbered. If edge r connects the vertices (i, j) the nonzero entries of $\boldsymbol{M}$ in row r are $\boldsymbol{M}_{ri} = -1$ and $\boldsymbol{M}_{rj} = 1$. With $\boldsymbol{G} := \mathrm{diag}(1, 1, 1, \gamma)$ the stiffness term can be written as

$$E_s(\boldsymbol{X}) = \|(\boldsymbol{M} \otimes \boldsymbol{G})\boldsymbol{X}\|_F^2 \quad . \tag{10}$$

Landmark Term The landmark term is similar to the distance term. We take the rows out of $\boldsymbol{D}$ that correspond to the landmark vertices, denote those by $\boldsymbol{D}_L$ and use the correspondence from the landmarks, $\boldsymbol{U}_L = \left[\boldsymbol{l}_1, \ldots, \boldsymbol{l}_l\right]^T$.

$$E_l(\boldsymbol{X}) = \|\boldsymbol{D}_L\boldsymbol{X} - \boldsymbol{U}_L\|_F^2 \quad . \tag{11}$$

Complete cost function The resulting quadratic function

$$\begin{aligned} \bar{E}(\boldsymbol{X}) &= \left\| \begin{bmatrix} \alpha\boldsymbol{M} \otimes \boldsymbol{G} \\ \boldsymbol{W}\boldsymbol{D} \\ \beta\boldsymbol{D}_L \end{bmatrix} \boldsymbol{X} - \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{W}\boldsymbol{U} \\ \boldsymbol{U}_L \end{bmatrix} \right\|_F^2 \\ &= \|\boldsymbol{A}\boldsymbol{X} - \boldsymbol{B}\|_F^2 \end{aligned} \tag{12}$$

can be minimised directly and exactly by setting its derivative to zero and solving the resulting system of linear equations. $\bar{E}(\boldsymbol{X})$ takes on its minimum at $\boldsymbol{X} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{B}$. This allows us to determine in each step of the algorithm the deformation which is optimal in the sense that it exactly minimises the cost function for fixed stiffness and correspondences.

If $\boldsymbol{A}$ does not have rank $4n$, the Hessian $\boldsymbol{A}^T\boldsymbol{A}$ is not invertible and the problem is ill posed. It is not obvious that the system is fully determined, as only three constraints per vertex result from the correspondence, but twelve free parameters per vertex are used. None of the submatrices $\boldsymbol{M} \otimes \boldsymbol{G}$, $\boldsymbol{D}$, and $\boldsymbol{D}_L$ of $\boldsymbol{A}$ corresponding to the three terms of the cost function has full rank on its own. The only matrix with enough rows is $\boldsymbol{M} \otimes \boldsymbol{G}$, but a node-arc incidence matrix for n nodes, such as $\boldsymbol{M}$, has only rank $n - 1$ [7]. Accordingly $\boldsymbol{M} \otimes \boldsymbol{G}$ has rank $4n - 4$ and a more detailed analysis is needed.

4.3. Well-Posedness

We prove that $\boldsymbol{A}$ has rank $4n$, such that $\boldsymbol{A}^T\boldsymbol{A}$ is invertible. (The landmark term is ignored, as it only adds unnecessary clutter and does not influence the result.) Partition $\boldsymbol{A}$ into n blocks corresponding to the vertices of the template. Each block $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n$ consists of four columns.

$$\begin{aligned} \boldsymbol{A} &= \left[\begin{array}{c|c|c} \alpha\boldsymbol{M}_1 \otimes \boldsymbol{G} & \ldots & \alpha\boldsymbol{M}_n \otimes \boldsymbol{G} \\ \boldsymbol{v}_1^T & & \\ & \ddots & \\ & & \boldsymbol{v}_n^T \end{array}\right] \\ &= \left[\begin{array}{c|c|c} \boldsymbol{A}_1 & \ldots & \boldsymbol{A}_n \end{array}\right] \end{aligned} \tag{13}$$

Denote the p'th column in the i'th block of columns by $\boldsymbol{A}_i^p$.

$$\boldsymbol{A}_i = \left[\begin{array}{c|c|c|c} \boldsymbol{A}_i^1 & \boldsymbol{A}_i^2 & \boldsymbol{A}_i^3 & \boldsymbol{A}_i^4 \end{array}\right] \tag{14}$$

If $\boldsymbol{A}$ had not full rank, we could find a set of nonzero coefficients a_i^p satisfying

$$\sum_p \sum_i a_i^p \boldsymbol{A}_i^p = 0 \quad . \tag{15}$$

The upper part of $\boldsymbol{A}$ is constructed from a node-arc-incidence matrix Kronecker multiplied with a diagonal matrix and scaled by α. The row $4(r-1)+p$ of $\boldsymbol{A}$ corresponding to an edge r which connects the vertices i and j has nonzero values only in the columns $4(i-1)+p$ and $4(j-1)+p$ for $p = 1, \ldots, 4$. The values of these entries differ only in the sign. This implies, that the coefficient a_i^p of column $\boldsymbol{A}_i^p$ is equal to the coefficient a_j^p if vertex j is connected to vertex i in the mesh. As this holds for any i with a nonzero a_i^p it implies that for a connected graph *all* a_i^p for fixed p and any i are equal. In consequence, there are only at most four different a^p values and the linear combination takes the form

$$\sum_p a^p \sum_i \boldsymbol{A}_i^p = 0 \tag{16}$$

The lower part of $\boldsymbol{A}$ representing the vertex correspondences adds more constraints. In row $4n + i$ corresponding to vertex i there are nonzero entries only in $\boldsymbol{A}_{4n+i,4(i-1)+p}$ for $p = 1, \ldots, 4$. All other values in this row are zero. This implies that to cancel out column $\boldsymbol{A}_i^p$ other columns pertaining to the same vertex are needed. As a result every vertex puts a constraint on the ratio between the coefficients.

$$\forall i = 1, \ldots, n: \quad \boldsymbol{v}_i^T \begin{bmatrix} a^1 \\ a^2 \\ a^3 \\ a^4 \end{bmatrix} = 0 \tag{17}$$

When taking all vertices into account this leads to the overdetermined system of equations

$$\begin{bmatrix} \boldsymbol{v}_1^T \\ \vdots \\ \boldsymbol{v}_n^T \end{bmatrix} \begin{bmatrix} a^1 \\ a^2 \\ a^3 \\ a^4 \end{bmatrix} = \boldsymbol{0} \quad . \tag{18}$$

Except for degenerate cases the only solution of this system is the null-vector which implies that all a^p are null, contradicting the assumption. This shows that $\boldsymbol{A}$ has full column rank, the Hessian is invertible, and the problem well posed.

4.4. Missing data and robustness

Template vertices $\boldsymbol{v}_i$ which correspond to missing data in the target are handled by setting the weight w_i from Equation (2) to zero. To detect which vertices have no correspondence three tests are used. A correspondence $(\boldsymbol{X}_i\boldsymbol{v}_i, \boldsymbol{u}_i)$ is

dropped if 1) $\boldsymbol{u}_i$ lies on a border of the target mesh, 2) the angle between the normals of the meshes at $\boldsymbol{X}_i\boldsymbol{v}_i$ and $\boldsymbol{u}_i$ is larger than a fixed threshold, or 3) the line segment $\boldsymbol{X}_i\boldsymbol{v}_i$ to $\boldsymbol{u}_i$ intersects the deformed template. The last test removes wrong correspondences introduced in regions where multiple surface layers are stacked – e.g. the ear – and only the outermost layer has been measured by the scanner.

Smooth infilling of missing data is achieved because of the template stiffness. Vertices without correspondences are moved such that the overall deformation is smooth.

If all vertices have correspondences and no reliability weighting is used, the residual in an optimal step ICP algorithm will always decrease. This is because neither finding a new deformation, nor finding new closest points can increase the residual. The new deformation decreases the residual with respect to the preliminary points, and searching for new closest points can never yield correspondences further away than the closest points from the previous iteration. A formal proof for the rigid case which applies equally to the nonrigid case can be found in [2]. When handling missing data by ignoring the contribution of vertices without correspondence to the distance term, this is no longer true. While the template aligns itself with the target, new correspondences are found and reliability weights w_i are increased. Accordingly, the residual may increase even though the correspondence improves.

The increasing residual inhibits a standard "black-box" optimiser as used in [1] from finding the correct correspondence, as these methods assume that the function has to decrease monotonically towards the optimum. In contrast, our optimization scheme is robust against this effect, as we do not impose the constraint of having a decreasing cost function. Addressing this issue by setting the distance contribution for vertices without correspondences to some arbitrary large value would not improve the situation, as this would push the template completely onto the scanned data, and the hole-filling capabilities of the method would be lost. A plot of the residual during the registration of the nonlinear synthetic example from Figure 5 is shown in Figure 4.

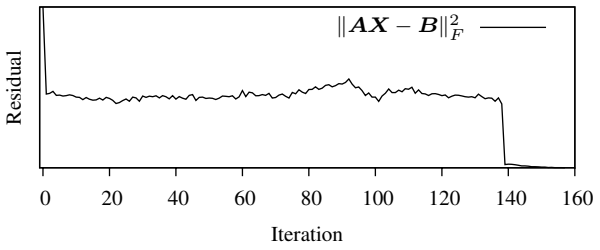

Figure 4. The no-monotonic decrease of the residual prohibits the use of a black-box optimiser. The figure shows the residual versus iteration during a registration. The residual increases between some steps, as the reliability weights increase when the template aligns itself with the target and more points find a correspondence. A general optimiser can not escape from the local minima, while our method is robust against this behaviour. Convergence is determined by checking if a fix-point has been reached.

5. Evaluation

We evaluate the method on a synthetic example and on a real world dataset of facial scans comparing three different registration algorithms.

As "nonrigid optimal step ICP algorithm with a stiffness term based on minimizing the difference of affine matrices assigned to neighbouring vertices" is an awkward name, we introduce the acronym *N-ICP-A* for our method. We will refer to the algorithm in [1] as *Allen'03*.

To demonstrate the advantage of penalizing the difference between affine transformations over penalizing the difference between translations we introduce *N-ICP-T* – a method similar to *N-ICP-A* – which regularises by minimizing the difference between the translations of neighbouring vertices. This corresponds to minimizing the gradient of the deformation field and models locally smooth translations. The unknowns for this method are organised in a $n \times 3$ matrix $\boldsymbol{X} = \left[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right]^T$. Denoting the Frobenius norm as $\|\cdot\|_F$, the cost function is given by

$$E(\boldsymbol{x}) = \left\| \begin{bmatrix} \alpha \boldsymbol{M} \\ \boldsymbol{W}\boldsymbol{I}_n \end{bmatrix} \boldsymbol{X} - \begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{W}(\boldsymbol{U} - \boldsymbol{V}) \end{bmatrix} \right\|_F^2 \quad . \tag{19}$$

A landmark term can be added in the obvious way by taking the corresponding rows from the distance term. Using only three parameters per vertex leads to a significantly faster optimization. On the other hand the deformations modelled by this cost function are more restricted.

Setup For all methods a gradual relaxation of the stiffness constraint lowering α in 100 equally distributed steps was used. The absolute α values depend on the template shape and resolution, and should be chosen such that at the beginning of the algorithm only global deformations are recovered. The lowest possible alpha depends also on the type of data. If α becomes too low the conditioning of $\boldsymbol{A}$ suffers and the solution becomes instable. All experiments were done with the same minimal alpha, staying far away from an unstable system. We start with an excessively high alpha, because starting with a high alpha cannot decrease the quality of the results, it may only lead to more steps being necessary.

The system of linear equations that arises in each step of the *N-ICP-A* and *N-ICP-T* methods was solved using the UMFPACK library [6]. For *Allen'03* we used the L-BFGS-B optimiser [15], providing approximate analytic gradients $\boldsymbol{A}^T(\boldsymbol{AX} - \boldsymbol{B})$.

The stiffness weight is lowered when the norm of the difference of the parameter vectors from two successive iterations is smaller than a threshold.

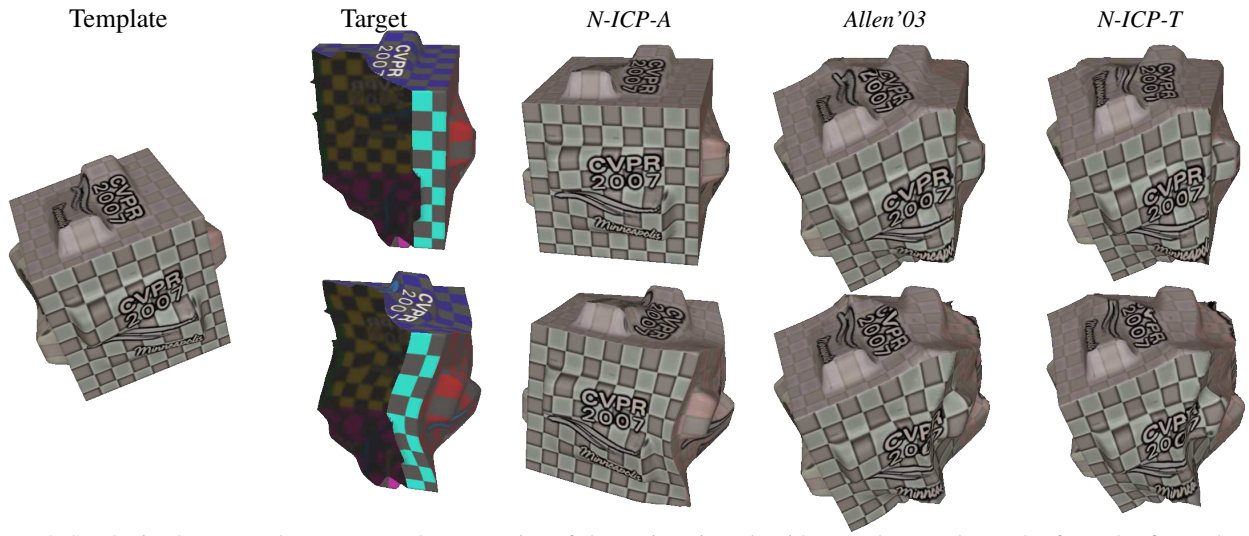

Figure 5. Synthetic shapes used to compare the properties of the registration algorithms. The template cube from the first column is registered onto the two targets in the second column. Backfacing regions are shown darkened. In the top row the target is a rigidly transformed version of the template with a large missing region. The shape in the bottom row is additionally deformed nonlinearly. The texture is only used to visualise the deformation within the surface, it is not used in the registration process. *N-ICP-A* manages to accurately align the borders and smoothly fill the hole, while *Allen'03* fails to find the optimum. *N-ICP-T* uses a different deformation model, which does not include rigid deformations. Therefore it fails already when recovering the global rigid deformation.

Synthetic Data We show results on two synthetic datasets. The template surface of both experiments is an embossed cube. The first target surface is a rigidly transformed version of the template with a large missing region. This deformation should be recovered perfectly by the locally affine registration methods, as it is a globally affine deformation. The second target surface is a nonlinearly deformed version of the first.

The synthetic datasets were registered without landmarks. The texture is only used to allow a better visualization of the deformations inside the surface, the registration method is completely shape based. The synthetic datasets and the registration results are shown in Figure 5.

The *N-ICP-A* method perfectly recovers the rigid deformation and smoothly fills in the missing region in the linear and nonlinear example. Trying to recover the deformation with the third method which models only locally smooth translations fails, as the rigid deformation is not modelled by the algorithm. When new correspondences are found the reliability weights and the residual increase. The *Allen'03* method fails as the L-BFGS-B optimizer tries to monotonically decrease the residual. A plot of the residual during the run of the *N-ICP-A* algorithm is shown in Figure 4.

Real Data Facial scans of 24 subjects of different gender and age ranging from 20 to 40 years were acquired with a structured light system. The scans were roughly cleaned to remove artefacts resulting from the hair and background. Fourteen landmarks as shown in Figure 6 were placed manually. Eight of these landmarks are needed to align the ears, as our scanner produces fragmented ear scans due to the acquisition angle and the concavity of the ear. The faces were first rigidly aligned to the template using the landmark points, then the full rigid and non-rigid deformations were recovered with the three methods presented. The last approximation is projected into the scan along the normals of the deformed template mesh to exactly represent the surface.

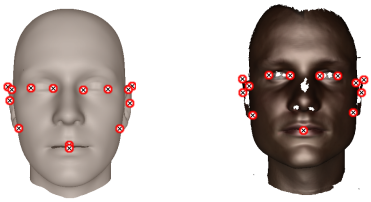

Figure 6. Landmarks used for head registration. As the scans of the ears are extremely fragmented we use four landmarks per ear to determine the overall shape. Additional landmarks at the eyes and at the mouth are used for the initial alignment of the two surfaces.

Measures of registration quality have to be domain specific. For the head scans we propose the following two measures for objective comparison of the registration results. For the first measure, we removed the global rigid deformation between all registered heads by aligning them against the template head. We then average over the angle between the normals of corresponding points on the surfaces of all pairs of registered scans. The intuition behind this is that, when the rigid part is removed, nor-

mals of corresponding points should have approximately the same direction. Only if points are mapped to different parts of the face (i.e. the side of the nose onto the cheek), the normals should deviate strongly. The average angle between normals over all pairs of registered scans is

N-ICP-A	*N-ICP-T*	*Allen'03*
5.6°	6.8°	7.9°

As a second quality measure the generalization capability of a convex model built with the registered faces is used. We reconstruct each mesh by a convex combination of all remaining meshes and measure the average l_2 distance between corresponding vertices of the reconstructed and the original mesh. We denote the stacked vertex positions of a face i as a vector $\boldsymbol{f}_i$. Finding the nearest mesh is a constrained quadratic problem of the form:

$$\boldsymbol{\alpha} = \arg\min_{\boldsymbol{\alpha}} \left\| \left[\boldsymbol{f}_1 \cdots \boldsymbol{f}_{i-1} \quad \boldsymbol{f}_{i+1} \cdots \boldsymbol{f}_n\right] \boldsymbol{\alpha} - \boldsymbol{f}_i \right\|^2$$
$$\forall_i \boldsymbol{\alpha}_i \geq 0 \qquad \sum_i \boldsymbol{\alpha}_i = 1$$

The reconstruction error measured by averaging over the l_2 distance of corresponding points of the reconstructed and the original mesh was:

N-ICP-A	*N-ICP-T*	*Allen'03*
1710μm	1740μm	2040μm

6. Conclusion and Future Work

The class of optimal step nonrigid ICP algorithms was introduced, and two instances were compared. They differ in the modelled deformation space and accordingly regularisation method. It was shown that locally smooth affine deformations can be used to register shapes with significant variation. The *N-ICP-A* method is well suited to register surfaces with missing data and is robust for a wide range of initial conditions. The algorithm is distinguished from previous nonrigid registration methods by taking optimal greedy steps within each iteration.

We proved that the regularisation introduced in [1] and used in *N-ICP-A* leads to a well posed problem.

The algorithm can be expanded naturally to incorporate additional constraints in the nearest point search. Feldmar and Ayache [8] proposed to search for the nearest *compatible* point, taking into account distance and normal direction. Another promising candidate is the shape index (Koenderink and van Doorn [13]), which separates shape and scale of a surface into two orthogonal dimensions.

We did not make use of texture information, though this might prove valuable for certain edges like the corner of the eyes. A problem with texture measure in faces is, that they are not necessarily connected to the underlying shape. For example the position of the eyebrows relative to the underlying bone structure varies strongly between subjects.

Within the same framework a host of different algorithms can be developed. For registration of facial expressions of a single subject, a physically based regularisation might prove useful, and the difference between isotropic and anisotropic regularisation needs to be explored in more detail.

References

[1] B. Allen, B. Curless, and Z. Popović. The space of human body shapes: reconstruction and parameterization from range scans. *ACM Trans. on Graphics (TOG)*, 22(3):587–594, 2003.

[2] P. J. Besl and N. D. Mckay. A method for registration of 3-d shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.

[3] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *SIGGRAPH '99: Proceedings*, pages 187–194, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[4] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. In *IEEE Int. Conference on Robotics and Automation*, volume 3, pages 2724–2729, 1991.

[5] J. Davis, S. R. Marschner, M. Garr, and M. Levoy. Filling holes in complex surfaces using volumetric diffusion. In *3D Data Processing Visualization and Transmission*, pages 428–861, 2002.

[6] T. A. Davis. Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method. *ACM Trans. On Mathematical Software*, 30(2):196–199, June 2004.

[7] M. Dekker. *Mathematical Programming*. CRC, May 1986.

[8] J. Feldmar and N. Ayache. Rigid, affine and locally affine registration of free-form surfaces. *Int. Journal of Computer Vision*, 18(2):99–119, May 1996.

[9] M. Ferrant, S. K. Warfield, C. R. G. Guttmann, R. V. Mulkern, F. A. Jolesz, and R. Kikinis. 3d image matching using a finite element based elastic deformation model. In *MICCAI '99*, pages 202–209, London, UK, 1999. Springer-Verlag.

[10] B. Fischer and J. Modersitzki. Fast diffusion registration. *Contemporary Mathematics*, 313:117–129, 2002.

[11] B. Fischer and J. Modersitzki. Curvature based image registration. *Journal of Mathematical Imaging and Vision*, 18(1):81–85, January 2003.

[12] K. Kähler, J. Haber, H. Yamauchi, and H.-P. Seidel. Head shop: generating animated head models with anatomical structure. In *SCA '2: Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 55–63, New York, NY, USA, 2002. ACM Press.

[13] J. J. Koenderink and A. J. van Doorn. Surface shape and curvature scales. *Image and Vision Computing*, 10(8):557–564, October 1992.

[14] R. Szeliski and S. Lavalle. Matching 3-d anatomical surfaces with non-rigid deformations using octree-splines. *Int. Journal of Computer Vision*, V18(2):171–186, May 1996.

[15] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans. On Mathematical Software*, 23(4):550–560, December 1997.