

# Stratified Generalized Procrustes Analysis<sup>1</sup>

Adrien Bartoli  
Adrien.Bartoli@gmail.com

Daniel Pizarro  
Dani.Pizarro@gmail.com

Marco Loog  
Marco.Loog@gmail.com

Clermont Université  
France

Universidad de Alcalá  
Spain

Delft University of Technology  
The Netherlands

---

## Abstract

This paper deals with *generalized procrustes analysis*. This is the problem of registering a set of shape data by finding a reference shape and global rigid transformations given point correspondences. The transformed shape data must align with the reference shape as best possible. This is a difficult problem. The classical approach computes alternatively the reference shape, usually as the average of the transformed shapes, and each transformation in turn.

We propose a *stratified approach* inspired by recent results obtained in Structure-from-Motion. Our stratified approach offers a statistically grounded framework for obtaining both the transformations and the reference shape at once in two steps. First, we compute a reference shape and affine transformations. Second, we upgrade these transformations to the sought after similarity or euclidean transformations. In practice each of these two steps involves solving a non-convex optimization problem. We provide convex approximations and closed-form solutions.

As opposed to the classical alternation approach, our stratified approach processes data in batch. It gracefully deals with missing data. We provide results on synthetic and real data sets. Compared to the alternation schema, our algorithm obtains lower error in both affine and euclidean cases, especially for shapes with high deformations.

## 1 Introduction

In many different problems, data analysis requires one to first compensate for a global transformation between the different datasets of *shape data*. This is known as *procrustes analysis* in the statistics and shape analysis literature [□, □]. More precisely, it is called *generalized procrustes analysis* when more than two shape data are to be registered. In this problem, one global transformation per observed shape has to be computed, so that the shapes are mapped to a common coordinate frame whereby they look as ‘similar’ as possible. This process is called also *rigid registration*.

The global transformation is usually modeled by one of the groups of euclidean transformations ( $\frac{1}{2}d(d+1)$  degrees of freedom), similarity transformations ( $\frac{1}{2}d(d+1) + 1$  degrees of freedom) and affine transformations ( $d(d+1)$  degrees of freedom), with  $d$  the dimension

---

<sup>1</sup>This work has been partly supported by the Spanish Ministry of Science and Innovation under projects VISNU (ref. TIN2009-08984) and SDTEAM-UAH (ref. TIN2008-06856-C05-05)

of the data to be analyzed. The classical approach to generalized procrustes analysis is to select one of the shapes as a *reference shape*, and register each of the other shapes to the reference in turn by solving the absolute orientation problem. It is common to then alternate a re-estimation of the reference shape, as the average of the registered shapes, with shape registration. We call this general paradigm the *alternation approach* to generalized procrustes analysis. Both iterative [10] and algebraic closed-form solutions [6, 8] were proposed for the absolute orientation problem. Integrated solutions based on the alternation approach for multiple shapes have been recently proposed in [6, 8] with a total least squares step. As discussed later in this paper, averaging the registered shapes might not be the best solution to re-estimate the reference shape from a statistical standpoint.

We propose the *stratified approach* to generalized procrustes analysis. This approach draws on recent results and practice from the Structure-from-Motion community [9], whereby one estimates a so-called uncalibrated camera model and then upgrades or self-calibrates it. Here, the uncalibrated model is the affine transformation group, and the calibrated model is one of the similarity and euclidean transformation groups. Our stratified approach has several advantages: it processes data in batch, as opposed to the alternation approach, it gracefully deals with missing data thanks to the reference-space model we propose, and it is statistically grounded as we demonstrate below. Our experimental results show that our approach is very stable against high percentages of missing data, noise and deformations in the shapes.

The paper is organized as follows. We state the problem and give mathematical preliminaries and notation in §2. Our stratified approach is outlined in §3. The affine registration step and the update to similarity-euclidean registration are respectively given in §3.1 and §3.2. We report experimental results on simulated and real data in §4. We conclude in §5.

## 2 Problem Statement and Mathematical Preliminaries

We assume that there exist an unknown reference shape  $\mathbf{S}^\top \stackrel{\text{def}}{=} (\mathbf{S}_1 \ \dots \ \mathbf{S}_m) \in \mathbb{R}^{d \times m}$ ,  $\mathbf{S}_j \in \mathbb{R}^d$ , and unknown global transformations  $\mathcal{T} \stackrel{\text{def}}{=} \{T_1, \dots, T_n\}$ ,  $T_i: \mathbb{R}^d \rightarrow \mathbb{R}^d$ , such that the discrepancy between the model predicted shape points  $T_i(\mathbf{S}_j) \in \mathbb{R}^d$  and the observed shape points  $\mathbf{D}_{i,j} \in \mathbb{R}^d$  follows a gaussian *i.i.d* distribution of unknown variance  $\sigma^2$ . The model likelihood is thus:

$$\mathcal{L}(\mathcal{T}, \mathbf{S}) \stackrel{\text{def}}{=} \frac{1}{\sigma^d (2\pi)^{\frac{d}{2}}} \prod_{i=1}^n \prod_{j=1}^m \exp\left(-\frac{v_{i,j}}{2\sigma^2} \|\mathbf{D}_{i,j} - T_i(\mathbf{S}_j)\|_2^2\right), \quad (1)$$

where  $\|\mathbf{u}\|_2 = \sqrt{\mathbf{u}^\top \mathbf{u}}$  is the vector two-norm. Here we have omitted the conditional dependencies on the observed shape points to leave the notation uncluttered. The variables  $v_{i,j} \in \{0, 1\}$  allow us to model *missing data*, the fact that some points may not be observed in some shapes. Our goal is to maximize the likelihood under a set of constraints  $\mathcal{C}$  ensuring that the solution is non-degenerate. These constraints may also fix some of the *gauge freedoms* of the cost under the global transformation group being considered. It is straightforward to see that maximizing the likelihood  $\mathcal{L}$  amounts to minimizing the negative log likelihood, proportional to the *data-space cost*  $\mathcal{E}$ , defined by:

$$\mathcal{E}(\mathcal{T}, \mathbf{S}) \stackrel{\text{def}}{=} \sum_{i=1}^n \sum_{j=1}^m v_{i,j} \|\mathbf{D}_{i,j} - T_i(\mathbf{S}_j)\|_2^2. \quad (2)$$

The data-space cost is related to a generative modeling of the data, that we hereinafter call the *data-space model*. It matches the definition of [10] of registration based on the notion of average shape. The data-space cost is *gauge invariant* (i.e. the cost does not depend on the coordinate frame in which the reference shape  $S$  and the transformations  $\mathcal{T}$  are expressed) and it is nonlinear and non-convex (See figure 1 for a graphical representation). Consequently, there does not exist a general closed-form solution to the maximum likelihood problem for the data-space model. In the general case, one has to resort to iterative nonlinear least squares optimization. This raises the problem of finding an initial solution lying as close as possible to the global minimum of the data-space cost (2).

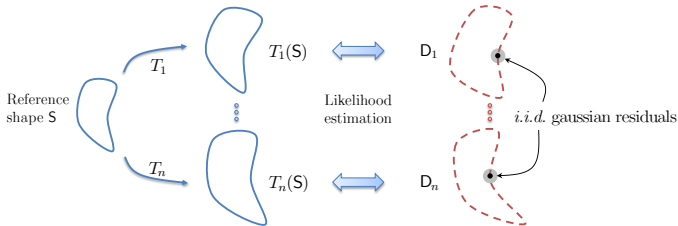


Figure 1: The data-space model.

**Transformations.** A euclidean transformation  $T_E$  is represented by a matrix  $E \in SO(d)$  (the rotation) and a vector  $\mathbf{e} \in \mathbb{R}^d$  (the translation). It is applied to a point  $\mathbf{S}$  as  $\mathbf{S} \mapsto T_E(\mathbf{S}) \stackrel{\text{def}}{=} E\mathbf{S} + \mathbf{e}$ . A similarity transformation  $T_S$  has the same representation with a scalar  $\zeta > 0$  (the scale factor). It is applied to a point  $\mathbf{S}$  as  $\mathbf{S} \mapsto T_S(\mathbf{S}) \stackrel{\text{def}}{=} \zeta E\mathbf{S} + \mathbf{e}$ . Finally, an affine transformation  $T_A$  is represented by a matrix  $A \in \mathbb{R}^{d \times d}$  (the ‘rotational’ part) and a vector  $\mathbf{a} \in \mathbb{R}^d$  (the translation). It is applied to a point  $\mathbf{S}$  as:

$$\mathbf{S} \mapsto T_A(\mathbf{S}) \stackrel{\text{def}}{=} A\mathbf{S} + \mathbf{a}. \quad (3)$$

An affine transformation factors in a purely affine transformation and a euclidean transformation using QR decomposition of its rotational part  $A \rightarrow QZ$  where  $Q \in O(d)$  and  $Z \in \mathbb{R}^{d \times d}$  is an upper triangular matrix. If  $\det(Q) = -1$  switching the sign of the last column of  $Q$  and of  $Z_{d,d}$  ensures  $Q \in SO(d)$ , i.e., to be a proper rotation matrix. In this decomposition,  $Z$  represents the ‘purely affine’ part of the transformation and  $Q$  the euclidean part, leading to  $T_A(\mathbf{S}) = QZ\mathbf{S} + \mathbf{a}$ . Abusing notation, we use a letter such as  $T$  for the generic  $\mathbb{R}^d \rightarrow \mathbb{R}^d$  transformation function, and use the same letter for the transformation parameters over which a minimization should be conducted. Sets of transformations are written using calligraphic fonts such as  $\mathcal{T} = \{T_1, \dots, T_n\}$ . For each type of registration, a minimum number of data point correspondences per shape (in general configuration) is required. While these minimal numbers can be analytically derived, we prefer to numerically check the conditioning of the design matrices to ensure that (i) the minimal number of data points is reached and (ii) they do not form a singular configuration.

### 3 The Stratified Approach

Our contribution in this paper is the definition of a stratified approach to generalized procrustes analysis and a set of models and tools that implements it. This makes it possible

to efficiently register multiple shapes, in any dimension and for the three aforementioned transformations. Our framework has four main steps. The first two steps perform affine registration (initialization and nonlinear refinement, respectively), and are needed in all cases, except registration without missing data that is dealt with matrix factorization. The last two steps perform similarity-euclidean registration (initialization from the affine registration and nonlinear refinement, respectively).

### 3.1 Affine Registration

We define  $\mathcal{A} \stackrel{\text{def}}{=} \{A_1, \dots, A_n\}$  to be the set of unknown affine transformations. An affine transformation  $A_i$  is represented as a pair  $A_i = (A_i; \mathbf{a}_i) \in \mathbb{R}^{d \times d} \times \mathbb{R}^d$ . Substituting the affine transformation model (3) into the general data-space cost (2) gives:

$$\mathcal{E}(\mathcal{A}, S) = \sum_{i=1}^n \sum_{j=1}^m v_{i,j} \|\mathbf{D}_{i,j} - A_i \mathbf{S}_j - \mathbf{a}_i\|_2^2. \quad (4)$$

Finding a minimizer of this cost is what we aim at. A closed-form solution based on rank- $d$  factorization of a measurement matrix containing all the shape data exist only if all points are observed in all shape data. We do not give the details of this solution. For the general case our strategy is to get an initial (approximate) closed-form solution based on the *reference-space model* that we propose and derive below.

#### 3.1.1 Initial Solution

So as to approximate problem (4), we introduce the inverse  $B_i = (B_i; \mathbf{b}_i) \stackrel{\text{def}}{=} (A_i^{-1}; -A_i^{-1} \mathbf{a}_i)$  of the sought affine transformations (and the set  $\mathcal{B} \stackrel{\text{def}}{=} \{B_1, \dots, B_n\}$ ), and approximate the data-space cost  $\mathcal{E}$  by what we call the reference-space cost  $\tilde{\mathcal{E}}$ :

$$\mathcal{E}(\mathcal{A}, S) \approx \tilde{\mathcal{E}}(\mathcal{B}, S) \stackrel{\text{def}}{=} \sum_{i=1}^n \sum_{j=1}^m v_{i,j} \|B_i \mathbf{D}_{i,j} + \mathbf{b}_i - \mathbf{S}_j\|_2^2. \quad (5)$$

The major advantage of the reference-space cost is that it is a sum of squares linear in the adjustable parameters, and thus leads to a linear least squares optimization problem. Even though  $\tilde{\mathcal{E}}$  is an approximation of  $\mathcal{E}$ , it can also be interpreted as a negative log likelihood under the hypothesis that the residuals are gaussian *i.i.d.* in the registered shape points. We now set the optimization problem to be solved in the initialization step:

$$\min_{\mathcal{B}, S} \tilde{\mathcal{E}}(\mathcal{B}, S) \quad \text{s.t.} \quad S^\top S = \mathbf{I}_{(d \times d)} \quad \text{and} \quad S^\top \mathbf{1}_{(m \times 1)} = \mathbf{0}_{(d \times 1)}. \quad (6)$$

The non-degeneracy constraint we choose here is  $S^\top S = \mathbf{I}$ . It is necessary to prevent the reference shape to collapse *i.e.*, to avoid the trivial zero-cost solution  $S = 0$  and  $\mathcal{A} = \{(0; \mathbf{0}), \dots, (0; \mathbf{0})\}$ . For pure convenience in our optimization algorithm, we choose to fix the origin of the coordinate frame at the centroid of the reference shape using the constraint  $S^\top \mathbf{1}_{(m \times 1)} = \mathbf{0}_{(d \times 1)}$ . We rewrite problem (6) as:

$$\min_{\tilde{X} \in \mathbb{R}^{p \times d}} \|\mathbf{K}_B \tilde{X} + \mathbf{K}_S S\|_{\mathcal{F}}^2 \quad \text{s.t.} \quad S^\top S = \mathbf{I}_{(d \times d)} \quad \text{and} \quad S^\top \mathbf{1}_{(m \times 1)} = \mathbf{0}_{(d \times 1)}, \quad (7)$$

where  $\|U\|_{\mathcal{F}} \stackrel{\text{def}}{=} \sqrt{\text{tr}(U^T U)}$  is the matrix Frobenius norm and  $K_B$  and  $K_S$  are:

$$K_B \stackrel{\text{def}}{=} \begin{pmatrix} K_{B,1} & & \\ & \ddots & \\ & & K_{B,n} \end{pmatrix} \in \mathbb{R}^{\alpha \times n(d+1)} \quad \text{and} \quad K_S \stackrel{\text{def}}{=} \begin{pmatrix} K_{S,1} \\ \vdots \\ K_{S,n} \end{pmatrix} \in \mathbb{R}^{\alpha \times m}, \quad (8)$$

with:

$$K_{B,i} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{D}_{i,1} & 1 \\ \vdots & \vdots \\ \mathbf{D}_{i,m} & 1 \end{pmatrix} \in \mathbb{R}^{\beta_i \times d+1} \quad \text{and} \quad K_{S,i} \stackrel{\text{def}}{=} \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \in \mathbb{R}^{\beta_i \times m}. \quad (9)$$

Only the ‘rows’ for which  $v_{i,j} = 1$  are instantiated in  $K_{B,i}$  and  $K_{S,i}$ . It should be noted that both matrices  $K_B$  and  $K_S$  are very sparse. Temporarily fixing  $S$ , we get the minimizer for  $B$  as  $B^* \stackrel{\text{def}}{=} -K_B^\dagger K_S S$ . Thanks to the structure of matrix  $K_B$ , the pseudo-inverse  $K_B^\dagger$  can be computed very efficiently. Substituting  $B^*$  in the problem formulation (7), we get:

$$\min_{S \in \mathbb{R}^{m \times d}} \| -K_B K_B^\dagger K_S S + K_S S \|_{\mathcal{F}}^2 \quad \text{s.t.} \quad S^T S = I_{(d \times d)} \quad \text{and} \quad S^T \mathbf{1}_{(m \times 1)} = \mathbf{0}_{(d \times 1)}. \quad (10)$$

Matrix  $\hat{K}_B \stackrel{\text{def}}{=} K_B K_B^\dagger$  is called the *hat matrix*. It allows us to rewrite the problem as:

$$\min_{S \in \mathbb{R}^{m \times d}} \|(I - \hat{K}_B) K_S S\|_{\mathcal{F}}^2 \quad \text{s.t.} \quad S^T S = I_{(d \times d)} \quad \text{and} \quad S^T \mathbf{1}_{(m \times 1)} = \mathbf{0}_{(d \times 1)}. \quad (11)$$

The reference-space cost is invariant to similarity transformations. Therefore, the position constraint  $S^T \mathbf{1} = \mathbf{0}$  can be simply enforced by adding it to the cost, giving the problem:

$$\min_{S \in \mathbb{R}^{m \times d}} \|LS\|_{\mathcal{F}}^2 \quad \text{s.t.} \quad S^T S = I_{(d \times d)} \quad \text{with} \quad L \stackrel{\text{def}}{=} \begin{pmatrix} (I - \hat{K}_B) K_S \\ \mathbf{1}^T \end{pmatrix} \in \mathbb{R}^{\alpha+1 \times m}. \quad (12)$$

One could now find the constrained minimizer  $S^*$  using a simple SVD (Singular Value Decomposition) of matrix  $L$ . However, We propose a solution that goes deeper into exploiting the problem structure using regular full matrices. We first note that the right singular vectors of  $L$  and of  $W \stackrel{\text{def}}{=} L^T L$  are identical. We thus rewrite the problem as:

$$\min_{S \in \mathbb{R}^{m \times d}} \|WS\|_{\mathcal{F}}^2 \quad \text{s.t.} \quad S^T S = I_{(d \times d)} \quad \text{with} \quad W \stackrel{\text{def}}{=} L^T L \in \mathbb{R}^{m \times m}. \quad (13)$$

The complexity thus depends on the number of model points  $m$  but not on the number of shape data  $n$ . Matrix  $\hat{K}_B$  can be formed block-wise since  $K_B = \text{diag}(K_{B,i})$  implies  $\hat{K}_B = \text{diag}(\hat{K}_{B,i})$ . The complexity of computing  $\hat{K}_B$  is thus proportional to the inversion of the symmetric matrices  $K_{B,i}^T K_{B,i} \in \mathbb{R}^{d+1 \times d+1}$ . Matrix  $W \in \mathbb{R}^{m \times m}$  is then given by  $W = \mathbf{1}\mathbf{1}^T + \sum_{i=1}^n K_{S,i}^T (I - \hat{K}_{B,i}) K_{S,i}$  since  $(\hat{K}_{B,i} - I)^T (\hat{K}_{B,i} - I) = I - \hat{K}_{B,i}$ . Given that matrix  $K_{S,i}$  just insert rows and columns of zeros a very efficient algorithm can be implemented to compute  $W$ . The solution  $S^*$  from (13) is simply obtained as the last  $d$  columns of the matrix  $V$  from the SVD  $W \xrightarrow{\text{SVD}} U \Sigma V^T$ .

### 3.1.2 Iterative Refinement

Given the initial suboptimal but closed-form solution in the reference-space, we propose to make iterative refinements through nonlinear least squares algorithms such as Gauss-Newton, Levenberg-Marquardt and other variants. Two main issues arise at this point.

First it must be noticed that the problem has a very sparse representation. The residuals in the cost function (4), for those points on shape  $i$  depend only on the  $i$ -th transformation  $A_i = (A_i; \mathbf{a}_i)$ , and not on the others. Similarly, the residuals for the instances of point  $j$  over the shape data depend only on the  $j$ -th reference shape point  $\mathbf{S}_j$ . Therefore, the cost Jacobian used in a Gauss-Newton algorithm has a strong block structure that allows one to solve the normal equations in a highly efficient manner. Since this is a very standard technique, we shall not describe it in details. The interested reader can refer to [9], for instance.

The second issue to deal with is the gauge freedoms of the data-space cost. There is a  $d(d+1)$ -dimensional linear subspace of equivalent solutions. The normal equations derived from the Gauss-Newton method are equally affected. The damping strategy in the Levenberg-Marquardt algorithm gracefully deals with this kind of gauge freedoms.

## 3.2 Similarity-Euclidean Registration

In this section we deal with the problem of initializing a similarity or a euclidean registration from an affine registration  $A_i = (A_i; \mathbf{a}_i)$ ,  $i = 1, \dots, n$ . In the euclidean case, this requires one to find the registration  $\mathcal{E}_i = (E_i, \mathbf{e}_i)$  such that  $E_i \in SO(3)$ ,  $i = 1, \dots, n$ . A trivial solution that comes to mind is to find each  $E_i$  by projecting  $A_i$  to the Stiefel manifold of orthonormal matrices. This however cannot be done as directly. Indeed, one has to take into account the nature of the gauge.

In the data-space model framework, the affine registration we compute is up to an unknown global affine transformation  $G = (G; \mathbf{g})$ . The rotational part of this affine transformation factors in the product of a euclidean  $Q$  and a purely affine transformation  $Z$ :  $G = QZ$ . The upgrading transformation is represented by  $Z$ . The projection to Stiefel manifold must take place after upgrading.

### 3.2.1 Computing the Upgrading Transform

What we know is that there exist an upgrading transform  $Z$  and orthonormal transformations  $Q_i$  such that:

$$\mathbf{D}_{i,j} \approx Q_i Z \mathbf{S}_j + \mathbf{a}_i. \quad (14)$$

Let  $\mathcal{X} \in \mathbb{R}^{nd \times m}$  be the centred measurement matrix, with entries  $\mathbf{D}_{i,j} - \mathbf{a}_i$ .<sup>1</sup> We may rewrite the previous equation in matrix form as:

$$\mathcal{X} \approx \mathcal{Q} Z \mathbf{S}^\top \quad \text{with} \quad \mathcal{X} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{D}_{1,1} - \mathbf{a}_1 & \cdots & \mathbf{D}_{1,m} - \mathbf{a}_1 \\ \vdots & \ddots & \vdots \\ \mathbf{D}_{n,1} - \mathbf{a}_n & \cdots & \mathbf{D}_{n,m} - \mathbf{a}_n \end{pmatrix} \quad \text{and} \quad \mathcal{Q} \stackrel{\text{def}}{=} \begin{pmatrix} Q_1 \\ \vdots \\ Q_n \end{pmatrix}. \quad (15)$$

Given that  $\mathbf{S}$  is column-orthonormal,  $\mathbf{S}^\top \mathbf{S} = \mathbf{I}$ . We multiply each side of the equation to the right by  $\mathbf{S}$  and get:

$$\mathcal{X} \mathbf{S} \approx \mathcal{Q} Z. \quad (16)$$

<sup>1</sup>Missing data points (for which  $v_{i,j} = 0$ ) are simply approximated by their prediction by the affine registration  $\mathbf{D}_{i,j} \approx A_i \mathbf{S}_j + \mathbf{a}_i$ .

Since  $\mathcal{Q}$  is made of orthonormal matrices, we have that  $\mathcal{Q}^\top \mathcal{Q} = n\mathbf{I}$ . Multiplying each side of the equation to the left by its transpose gives:

$$\mathbf{S}^\top \mathcal{X}^\top \mathcal{X} \mathbf{S} \approx n\mathbf{Z}^\top \mathbf{Z} = \mathbf{Y}, \quad (17)$$

from which we obtain  $\mathbf{Y}$  that is guaranteed to be SPD (Symmetric Positive Definite). Cholesky factorization is used to recover an upper triangular  $\mathbf{Z}$ .

In the scaled-euclidean case,  $\mathcal{Q}$  is made of scaled orthonormal matrices, and therefore  $\mathcal{Q}^\top \mathcal{Q} = \gamma \mathbf{I}$  for some  $\gamma \in \mathbb{R}$ ,  $\gamma > 0$ . As above, multiplying each side of the equation to the left by its transpose gives:

$$\mathbf{S}^\top \mathcal{X}^\top \mathcal{X} \mathbf{S} \approx \gamma \mathbf{Z}^\top \mathbf{Z} = \gamma \mathbf{Y}, \quad (18)$$

Again, since  $\gamma > 0$ , the recovered matrix is guaranteed to be SPD. Matrix  $\mathbf{Z}$  can only be recovered up to scale, which is not important since it corresponds to the ambiguous global scale.

Given that we have computed the upgrading transformation  $\mathbf{Z}$ , we now sought orthonormal transformations  $\mathbf{E}_i$  close to  $\tilde{\mathbf{E}}_i \stackrel{\text{def}}{=} \mathbf{A}_i \mathbf{Z}^{-1}$ . We propose to find  $\mathbf{E}_i \in SO(d)$  such that  $\|\mathbf{E}_i - \tilde{\mathbf{E}}_i\|_{\mathcal{F}}^2 \rightarrow \min$ . The solution is given from the SVD  $\tilde{\mathbf{E}}_i \xrightarrow{\text{SVD}} \mathbf{U} \Sigma \mathbf{V}^\top$ : we get  $\mathbf{E}_i \leftarrow \mathbf{U} \mathbf{V}^\top$  in the euclidean case and  $\mathbf{E}_i \leftarrow \frac{1}{d} (\sum_{k=1}^d \sigma_k) \mathbf{U} \mathbf{V}^\top$  in the scaled-euclidean case, where the  $\sigma_k$  are the singular values in  $\Sigma$ .

### 3.2.2 Iterative Refinement

The iterative refinement in this case is similar to the affine case. The difference lies in the update rule and the parameterization of the transformations being used. We use a non-singular redundant global parameterization for rotations that we update via a singular but locally well-behaved minimal parameterization. We represent a rotation by a matrix in  $SO(d)$  (redundent global parameterization with  $d^2$  parameters), that is locally updated by an angle-based parameterization with the minimal  $\frac{d(d-1)}{2}$  parameters, corresponding to the exponential map. Define  $\mathbf{R}(\theta)$  the function that maps the  $\frac{d(d-1)}{2}$  vector  $\theta$  to the corresponding rotation matrix. We use the update rule  $\mathbf{E}_i \leftarrow \mathbf{E}_i \mathbf{R}(\theta_i)$  and perform optimization over  $\theta_i$  around  $\mathbf{0}$  at each iteration since  $\mathbf{R}(\mathbf{0}) = \mathbf{I}_{(d \times d)}$ .

## 4 Experimental Results

### 4.1 Simulated Data

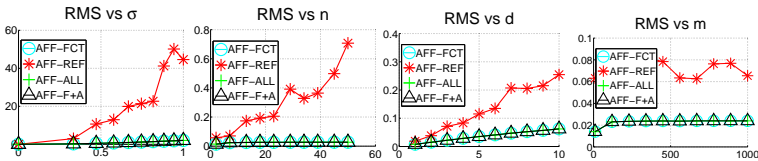
The following methods are compared. AFF-FCT: affine factorization (complete datasets only) based on rank- $d$  factorization of the measurement matrix. AFF-REF: the proposed closed-form affine registration using the reference-space method. AFF-ALL: the whole proposed affine registration part including iterative refinement. AFF-F+A: the whole affine registration method with AFF-FCT used as initialization. EUC-UPG: affine registration with euclidean upgrading. EUC-ALL: the whole proposed euclidean registration. ALT: the following version of the alternation algorithm is considered: first, the reference shape is obtained as the mean of the transformed data-shapes, using the inverse of the transformations obtained in the last iteration. Second, we find the set of transformations that minimize the

data-space cost between the reference shape and the data-shapes in turn, using the closed-form solution of [9]. The algorithm stops when the reference shape does not change between alternations.

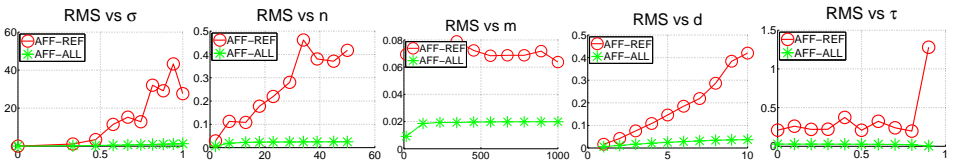
**Data generation.** The reference shape is generated by drawing  $m = 50$  points in an origin-centred hyper-sphere of unit radius in dimension  $d = 3$ . Each of the  $n = 5$  affine transformations is randomly generated by selecting  $d + 1$  control points into the unit hypersphere. The euclidean transformations are generated as the orthonormal part of the QR factorization of the above affine transformations and ensuring positiveness of the determinant. The generated shapes are obtained by applying the  $n$  transformations to the  $m$  reference points. Non-rigidity and noise are both modeled by an additive gaussian process with variance  $\sigma^2 = 0.01$ . Missing data are obtained by erasing points with  $\tau = 0.5$  (50% missing data).

**Experiments.** We use as quality criterion the data-space cost. Each experiment is run with 100 trials and we report RMS values. The ranges are the following:  $\sigma^2 = 0 \dots 1$ ;  $n = 2, \dots, 50$ ;  $m = 10, \dots, 1000$ ;  $d = 1, \dots, 10$ ;  $\tau = 0 \dots 0.9$ .

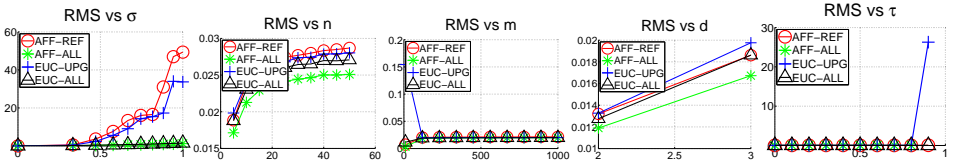
Experiment 1: affine transformation, full-data ( $\tau = 0$ ); varying  $\sigma^2$ ,  $n$ ,  $m$  and  $d$



Experiment 2: affine transformations, varying  $\sigma^2$ ,  $n$ ,  $m$ ,  $d$  and  $\tau$



Experiment 3: euclidean transformations, varying  $\sigma^2$ ,  $n$ ,  $m$ ,  $d$  and  $\tau$



Experiment 4: euclidean transformations with high deformations, varying  $\sigma^2 = 0 \dots 20$  and  $d = 2, 3$

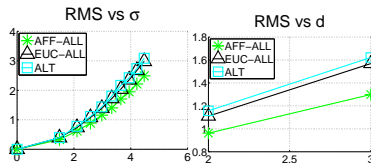


Figure 2: Experimental results (see main text for details).

The results clearly show that AFF-REF is suboptimal compared to the maximum likelihood solution offered by the other methods. However it is very useful to initialize the iterative refinement AFF-ALL in case of missing data. In the context of (scaled-)euclidean



registration it can be clearly seen that EUC-UPG is prone to errors and needs the iterative refinement performed by EUC-ALL. Much of this error is caused by the use AFF-REF for the affine solution that is being upgraded.

Overall, our methods are very efficient and resistant to various perturbation (extreme noise and amounts of missing data). The best combination is clearly the reference-space affine closed-form solution, followed by data-space affine nonlinear refinement, (scaled-)euclidean upgrading and nonlinear refinement. Our results show that avoiding the affine nonlinear refinement may strongly harm the result of (scaled-)euclidean upgrading. With high deformations, the EUC-ALL and AFF-REF performs better than the alternation algorithm in terms of data-space error.

## 4.2 Real Data

We tried our algorithms on two real datasets. The first dataset contains  $n = 10$  2D shapes with  $m = 40$  2D points that represent a human face with different poses in front of a camera. This dataset includes missing data ( $\tau \approx 0.1$ ) when the face turns, as it suffers from self-occlusion. The second dataset corresponds to 3D point coordinates given by Motion Capture (MOCAP) sensors, available from the HumanEVA [24] database. The dataset has  $n = 7$  shapes corresponding from a walking person.

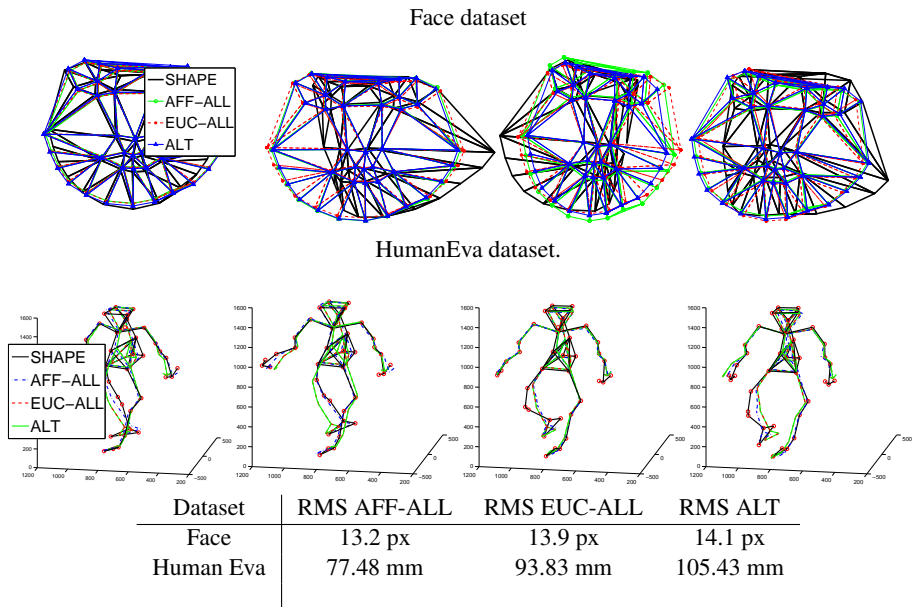


Figure 3: Visual comparison showing the shapes from the datasets and the corresponding predicted shapes after registration given by  $\mathcal{T}_i(S)$ . The table gives the data-space registration error.

As can be seen from the results in figure 3, AFF-ALL reaches a lower data-space error than EUC-ALL or ALT, which was expected since the affine transformations have more flexibility than the scaled-euclidean ones. For both types of transformations, however, we get convincing results. It can be seen that EUC-ALL reaches lower error in the data-space than the alternation algorithm in both datasets.

## 5 Conclusion

This paper presented a stratified approach to the generalized procrustes analysis problem for rigid registrations. It is similar in spirit to the stratified Structure-from-Motion paradigm, but brings specific methods well adapted to registration in arbitrary dimensional spaces. The proposed framework first solves the problem for the affine transformation group, for which a closed-form solution and an iterative method were developed. Similarity or euclidean transformations are then obtained through an upgrade of the affine solution. We show that the solution in the so-called data-space corresponds to the maximum likelihood estimate assuming *i.i.d.* gaussian noise in the measured shape points. The paper gives insights about how to exploit the problem sparsity and proposes algorithms to efficiently solve each step.

At this point, we emphasize that our stratified approach to generalized procrustes analysis provides a novel handle in that problem, which is generally solved by alternation algorithms. Experimental results showed that our stratified approach is very accurate and stable with respect to various factors such as data noise and missing data, while being able to gracefully handle missing data. Further experiments assessing in particular the computational runtime and comparison with various alternation-based methods on extensive datasets are however required. They will be provided in an extended version of the paper.

## References

- [1] I. Dryden and K. Mardia. *Statistical shape analysis*. Wiley New York, 1998.
- [2] D. Eggert, A. Lorusso, and R. Fisher. Estimating 3-D rigid body transformations: a comparison of four major algorithms. *Machine Vision and Applications*, 9(5), 1997.
- [3] J. Gower and G. Dijkstra. *Procrustes problems*. Oxford University Press, 2004.
- [4] R. Hartley and A. Zisserman. *Multiple view geometry in Computer Vision*. Cambridge university press Cambridge, UK, second edition edition, 2003.
- [5] B. Horn et al. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A*, 4(4):629–642, 1987.
- [6] S. Krishnan, P. Lee, J. Moore, and S. Venkatasubramanian. Global registration of multiple 3D point sets via optimization-on-a-manifold. *Eurographics symposium on Geometry processing*, 2005.
- [7] L. Sigal and M. Black. Humaneva: Synchronized video and motion capture dataset for evaluation of articulated human motion. *Brown University TR*, 2006.
- [8] M. Walker, L. Shao, and R. Volz. Estimating 3-D location parameters using dual number quaternions. *Computer Vision, Graphics and Image Processing: Image Understanding*, 54(3):358–367, 1991.
- [9] G. Wen, D. Zhu, S. Xia, and Z. Wang. Total least squares fitting of point sets in mD. In *Computer Graphics International 2005*, pages 82–86, 2005.
- [10] A. Yezzi and S. Soatto. Deformation: Deforming motion, shape average and the joint registration and approximation of structures in images. *International Journal of Computer Vision*, 53(2):153–167, 2003.