

On Matching Point Configurations*

DAVID JIMÉNEZ AND GUERGANA PETROVA

We present an algorithm that verifies if two unlabeled configurations of N points in \mathbb{R}^d are or are not an orthogonal transformation of one another, and if applicable, explicitly compute that transformation. We also give a formula for an orthogonal transformation in the case of noisy measurements.

Keywords and Phrases: Matching point configurations, orthogonal Procrustes problem, perturbation bounds.

Mathematics Subject Classification 2010: 65F15, 65F35, 65G99.

1. Introduction

In computer vision applications, it is often necessary to match an unidentified image to an image from a library of known images such as fingerprints, faces and others. This process is often done by identifying points (landmarks) on the incoming image and checking whether they match the point configuration from an already indexed image in an existing collection.

If we denote by $O(d)$ the group of the $d \times d$ orthogonal matrices, and by \mathcal{S}_N the group of all permutations of $\{1, 2, \dots, N\}$, a rigid motion \mathcal{R} in \mathbb{R}^d is defined by

$$\mathcal{R}\mathbf{x} = A\mathbf{x} + \mathbf{b}, \quad \mathbf{x} \in \mathbb{R}^d, \quad \text{with } A \in O(d), \quad \mathbf{b} \in \mathbb{R}^d. \quad (1)$$

The problem of matching two images can be formulated in the following way. Given two collections of N points $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ and $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ in \mathbb{R}^d , is there an orthogonal matrix $A \in O(d)$, a vector $\mathbf{b} \in \mathbb{R}^d$, and a permutation $\pi \in \mathcal{S}_N$, such that in the Euclidean norm $\|\cdot\|$, the rigid motion defined by (1), satisfies

$$\|\mathcal{R}\mathbf{p}_i - \mathbf{q}_{\pi(i)}\| \leq \varepsilon, \quad i = 1, \dots, N, \quad (2)$$

*Research supported by the Office of Naval Research Contract ONR N00014-11-1-0712, by the NSF Grant DMS 1222715, and by the Bulgarian Science Fund Grant DFNI-T01/0001.

for a sufficiently small ε ? A positive answer to this question would mean that we have found a match. The problem of matching two collections of points has been vastly studied using different approaches. One of them is based on the Gramians of the point configurations and has the advantage that it retains the information about the labeling (indexing) of the points. For example, it is a well known fact that given two point configurations $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ and $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ in \mathbb{R}^d , there exists a rigid motion \mathcal{R} such that $\mathcal{R}\mathbf{p}_i = \mathbf{q}_i$, $i = 1, \dots, N$, if and only if the Gramians $P^T P = Q^T Q$, where P and Q are the matrices with columns $\mathbf{p}_i - \bar{\mathbf{p}}$ and $\mathbf{q}_i - \bar{\mathbf{q}}$, respectively, with $\bar{\mathbf{q}} = \frac{1}{N} \sum_{i=1}^N \mathbf{q}_i$ and $\bar{\mathbf{p}} = \frac{1}{N} \sum_{i=1}^N \mathbf{p}_i$. Another well studied related problem is the orthogonal Procrustes problem for finding a matrix $A \in O(d)$, such that the Frobenius norm $\|AP - Q\|_F$ is minimized. Its solution, see [4, 5, 6, 9] and the references therein, is given by the matrix $A = UV^T$, where U and V come from the singular value decomposition of $Z = QP^T$, $QP^T = U\Sigma_Z V^T$.

In this paper, we first investigate the matching and registration of unlabeled point configuration using the Gramian approach. We propose and test a new algorithm that verifies whether two unlabeled configurations of N points in \mathbb{R}^d are or are not an orthogonal transformation of one another, and if applicable, explicitly compute that transformation. The algorithm is based on ideas used in variable decorrelation, which is routinely solved by principal component analysis (PCA). Existing algorithms for matching unlabeled point clouds are based on iterative closest point methods, see [8, 11], and deal with the registration of unlabeled point clouds of different sizes in the presence of noise. Although quite useful in practice, these methods often assume certain additional information about the point cloud. For example, they assume that the rigid motion \mathcal{R} is a small perturbation or it is roughly known, or that the nature of the cloud is such that there is a fast procedure to label a reasonably large subcloud, and thus compute the rigid motion based on that labeled subcollection. In contrast to these techniques, our algorithm does not require any information about the geometry of the cloud or the the nature of the rigid motion. Note that, compared to the $\binom{N}{2}$ values used in some other approaches (such as the distance distribution approach), our algorithm uses at most $d(1 + d + 2N)$ values to process a point configuration, thus reducing the memory cost and the data access time.

Here, we also present and test a stability result, see Theorem 3, where we explicitly compute a matrix $A \in O(d)$ for the rigid motion \mathcal{R} , see (1), in the case of noisy labeled point clouds. The question how noise in the data affects the computed rigid body motion is an important issue in practical applications. We show that if ε is small enough, and $P^T P = Q^T Q + \varepsilon M$, where the entries of the matrix M are bounded, there is an orthogonal matrix $A = A(\varepsilon)$ (which we construct), such that $\|AP - Q\| \leq \varepsilon \tilde{c} N^2$, with an explicitly computed constant \tilde{c} . The converse statement is not true and we show it by constructing a counterexample. Note that in [2], an explicit expression of the error in A , to first order, is given in terms of the errors in P and Q when $d = 3$. While

this result is much more specific than general error bounds that have been established before, it requires the exact values of the matrices P , Q and A . Our result from Theorem 3 does not require such knowledge and is in the spirit of the work in [10], where supremum bounds for the perturbation error in the solution A of the orthogonal Procrustes problem with the additional restriction that A has a positive determinant are derived.

2. Preliminaries

This section contains well known facts that will be used throughout the paper, as well as certain Procrustes analysis results, stated for self containment. Some of the proofs are included for clarity, while the basic results are only stated.

2.1. Matrices

Let $\pi \in \mathcal{S}_N$ be a permutation of $\{1, 2, \dots, N\}$. Then the $N \times N$ matrix E_π with columns $\{\mathbf{e}_{\pi(1)}, \dots, \mathbf{e}_{\pi(N)}\}$, where \mathbf{e}_j is the j -th element of the canonical basis of \mathbb{R}^N is called the permutation matrix associated to π . Multiplying a matrix A on the right by E_π permutes the columns of A by π .

Lemma 1. *Let E_π be the permutation matrix associated to $\pi \in \mathcal{S}_N$. Let A and B be two positive semidefinite symmetric matrices such that $B = E_\pi A E_\pi^T$. Then the set of eigenvalues of A and B are identical, including their algebraic multiplicity, and there exist eigenvalue decompositions of $A = U_A \Lambda U_A^T$ and $B = U_B \Lambda U_B^T$, such that $E_\pi = U_B U_A^T$.*

Note that, given arbitrary decompositions $A = \tilde{U}_A \Lambda \tilde{U}_A^T$ and $B = \tilde{U}_B \Lambda \tilde{U}_B^T$ for the positive semidefinite symmetric matrices A and B , the matrix $\tilde{U}_B \tilde{U}_A^T$ need not be a permutation matrix.

Lemma 2. *Let P and Q be two $d \times N$ matrices. Then $P^T P = Q^T Q$ if and only if there is a $d \times d$ orthogonal matrix A such that $AP = Q$. We call such a matrix A an equivalence matrix.*

Proof. If $AP = Q$, where A is an orthogonal matrix, we have $Q^T Q = (AP)^T AP = P^T (A^T A) P = P^T P$. Conversely, if $P^T P = Q^T Q$, then P and Q have the same singular values and the same right singular vectors. Then, using the singular value decomposition for P and Q , there are $d \times d$ orthogonal matrices U_P and U_Q such that $P = U_P \Sigma V^T$ and $Q = U_Q \Sigma V^T$, where V is the $N \times N$ orthogonal matrix containing the eigenvectors of $P^T P$. Therefore $Q = U_Q \Sigma V^T = (U_Q U_P^T) U_P \Sigma V^T = AP$, where $A = U_Q U_P^T$. \square

Finally, in this paper, unless stated otherwise, we use the Euclidean norm of a vector $\mathbf{x} \in \mathbb{R}^N$ and the corresponding induced matrix norm of a $d \times N$ matrix A , $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$. Note that the induced matrix Euclidean norm is also the spectral norm of A , namely

$$\|A\| = \sqrt{\lambda_{\max}(A^T A)} = \sqrt{\lambda_{\max}(A A^T)} = \|A^T\|,$$

where $\lambda_{\max}(A^T A)$ is the maximal eigenvalue of $A^T A$, and it is submultiplicative, that is $\|AB\| \leq \|A\| \cdot \|B\|$.

2.2. Labeled Point Configurations

We fix a coordinate system in \mathbb{R}^d and denote by $\mathcal{P} := \{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ and $\mathcal{Q} := \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ two collections of N points in \mathbb{R}^d , where \mathbf{p}_i is the coordinate vector of the i -th point from \mathcal{P} with respect to this coordinate system. Let $\bar{\mathbf{p}} = \frac{1}{N}(\mathbf{p}_1 + \dots + \mathbf{p}_N)$ and $\bar{\mathbf{q}} = \frac{1}{N}(\mathbf{q}_1 + \dots + \mathbf{q}_N)$ be the center of mass of \mathcal{P} and \mathcal{Q} , respectively, $\bar{\mathcal{P}}$ and $\bar{\mathcal{Q}}$ be the new (centered) collections $\bar{\mathcal{P}} := \{\mathbf{p}_1 - \bar{\mathbf{p}}, \dots, \mathbf{p}_N - \bar{\mathbf{p}}\}$ and $\bar{\mathcal{Q}} := \{\mathbf{q}_1 - \bar{\mathbf{q}}, \dots, \mathbf{q}_N - \bar{\mathbf{q}}\}$, and P and Q be the $d \times N$ matrices with columns $\mathbf{p}_i - \bar{\mathbf{p}}$ and $\mathbf{q}_i - \bar{\mathbf{q}}$, respectively.

If there is a rigid motion \mathcal{R} such that $\mathcal{R}\mathbf{p}_i = \mathbf{q}_i$, $i = 1, \dots, N$, we say that \mathcal{P} and \mathcal{Q} are *identically equivalent*. The following theorem, based on techniques more extensively discussed in [7], provides a tool to find \mathcal{R} if it exists.

Theorem 1. *The following statements are equivalent.*

- (i) \mathcal{P} and \mathcal{Q} are identically equivalent.
- (ii) $P^T P = Q^T Q$.
- (iii) There is an orthogonal matrix A such that $AP = Q$.

Proof. The equivalence for (ii) and (iii) has been proven in Lemma 2. If there is a rigid motion \mathcal{R} , $\mathcal{R}\mathbf{x} = A\mathbf{x} + \mathbf{b}$ with an orthogonal matrix A and a vector \mathbf{b} , such that $\mathcal{R}\mathbf{p}_i = \mathbf{q}_i$, $i = 1, \dots, N$, it is easy to verify that $\mathcal{R}\bar{\mathbf{p}} = \bar{\mathbf{q}}$. Therefore, $\mathbf{q}_i - \bar{\mathbf{q}} = \mathcal{R}\mathbf{p}_i - \mathcal{R}\bar{\mathbf{p}} = A(\mathbf{p}_i - \bar{\mathbf{p}})$ for every $i = 1, \dots, N$, and thus $AP = Q$. Conversely, if $AP = Q$ for an orthogonal matrix A , then $\mathbf{q}_i - \bar{\mathbf{q}} = A(\mathbf{p}_i - \bar{\mathbf{p}})$, and therefore $\mathbf{q}_i = A\mathbf{p}_i + (\bar{\mathbf{q}} - A\bar{\mathbf{p}}) = \mathcal{R}\mathbf{p}_i$, with $\mathbf{b} = \bar{\mathbf{q}} - A\bar{\mathbf{p}}$, $i = 1, \dots, N$. This establishes the equivalence between (i) and (iii). \square

Note that given $P^T P = Q^T Q$, the matrix A can be computed directly if P has rank d . In this case PP^T is invertible, and it can be shown that $A = QP^T(PP^T)^{-1}$.

3. Unlabeled Point Configurations

Often, when landmarks are extracted from an image to generate a point configuration \mathcal{P} , it is not possible to a priori enumerate the points in a manner consistent with the enumeration of an existing point collection \mathcal{Q} in our library.

In this case, if there exist a rigid motion \mathcal{R} and a permutation $\pi \in S_N$, such that $\mathcal{R}\mathbf{p}_i = \mathbf{q}_{\pi(i)}$, $i = 1, \dots, N$, we call the collections \mathcal{P} and \mathcal{Q} *equivalent*. This section explores how, given two point configurations, we decide whether they are equivalent or not.

Theorem 2. *Let $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ and $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ be two collections of N points in \mathbb{R}^d . The following statements are equivalent.*

- (i) \mathcal{P} and \mathcal{Q} are equivalent.
- (ii) There is a permutation matrix E_π , such that $P^T P = E_\pi^T Q^T Q E_\pi$.
- (iii) There is an orthogonal matrix A , completely determined in Lemma 2 and a permutation matrix E_π , such that $AP = QE_\pi$. Moreover, $E_\pi = U_\mathcal{Q} U_\mathcal{P}^T$, where $U_\mathcal{P}$ and $U_\mathcal{Q}$ are orthogonal matrices from an eigenvalue decomposition of $P^T P$ and $Q^T Q$, respectively. (Note that one does not know which particular eigenvalue decomposition will provide the matrices $U_\mathcal{P}$ and $U_\mathcal{Q}$.)

Proof. Let $\pi \in S_N$ be the permutation from the definition of equivalence of the two collections \mathcal{P} and \mathcal{Q} . We consider the permutation matrix E_π , associated with π . We denote by \mathcal{Q}_π the re-enumerated collection of points \mathcal{Q} with corresponding matrix Q_π with columns $\mathbf{q}_{\pi(i)} - \bar{\mathbf{q}}$. Note that $Q_\pi = QE_\pi$. Then, \mathcal{P} and \mathcal{Q} are equivalent if and only if \mathcal{P} and \mathcal{Q}_π are identically equivalent, which using Theorem 1, is true if and only if

$$P^T P = E_\pi^T Q^T Q E_\pi = E_{\pi^{-1}} Q^T Q E_{\pi^{-1}}^{-1}.$$

By Lemma 1, there are orthogonal matrices $U_\mathcal{P}$ and $U_\mathcal{Q}$ with $P^T P = U_\mathcal{P} \Lambda U_\mathcal{P}^T$ and $Q^T Q = U_\mathcal{Q} \Lambda U_\mathcal{Q}^T$, such that $E_{\pi^{-1}} = U_\mathcal{P} U_\mathcal{Q}^T$, and therefore $E_\pi = U_\mathcal{Q} U_\mathcal{P}^T$.

Relation (ii) can be written as $P^T P = (QE_\pi)^T QE_\pi$. By Lemma 2 this is true if and only if there is an orthogonal matrix A , described in this lemma, such that $AP = QE_\pi$. \square

Notice that the matrix QQ^T does not depend on the permutation of the columns of Q , since $Q_\pi Q_\pi^T = QE_\pi E_\pi^T Q^T = QQ^T$. The next lemma provides some insight on whether a suitable matrix A , related to the rigid motion \mathcal{R} exists and if it does, gives another way of its explicit construction.

Lemma 3. *Let $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ and $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ be two equivalent collections of N points in \mathbb{R}^d . Then the $d \times d$ matrices PP^T and QQ^T have the same eigenvalues $0 \leq \lambda_1 \leq \dots \leq \lambda_d$, including their algebraic multiplicities. Let \mathbf{v}_i and \mathbf{w}_i be the orthonormal eigenvectors of PP^T and QQ^T corresponding to λ_i , $i = 1, \dots, d$, respectively. Let V and W be the orthogonal matrices with columns $\{\mathbf{v}_i\}$ and $\{\mathbf{w}_i\}$. If the eigenvalues $\{\lambda_i\}$ are distinct, then there are integers $\epsilon_i = \pm 1$, $i = 1, \dots, d$, and a permutation $\pi \in S_N$, determined from $Q_\pi = AP$, such that*

$$\langle \mathbf{p}_k - \bar{\mathbf{p}}, \mathbf{v}_i \rangle = \epsilon_i \langle \mathbf{q}_{\pi(k)} - \bar{\mathbf{q}}, \mathbf{w}_i \rangle, \quad k = 1, \dots, N, \quad i = 1, \dots, d. \quad (3)$$

Moreover, if $E := \text{diag}(\epsilon_1, \dots, \epsilon_d)$, then the equivalence matrix A can be written as $A = WEV^T$.

Proof. It follows from Theorem 2 that if \mathcal{P} and \mathcal{Q} are equivalent, then there is an orthogonal matrix A , such that $AP = QE_\pi$ for some permutation $\pi \in S_N$, with $\pi = id$ if the points are identically equivalent. Then we have $QQ^T = APE_\pi^T E_\pi P^T A^T = A(PP^T)A^T$. Since PP^T is a real symmetric matrix, and \mathbf{v}_i , $i = 1, \dots, d$, are orthonormal eigenvectors of PP^T corresponding to the eigenvalues $\lambda_1 \leq \dots \leq \lambda_d$, we have $PP^T = V\Lambda_P V^T$, with $\Lambda_P = \text{diag}(\lambda_1, \dots, \lambda_d)$, and therefore $QQ^T = (AV)\Lambda_P(AV)^T$. Clearly, AV is an orthogonal matrix as a product of two orthogonal matrices. Also, QQ^T and PP^T have the same eigenvalues, including their algebraic multiplicities, and AV is a matrix whose columns $\{A\mathbf{v}_i\}$ are eigenvectors of QQ^T .

Let us consider now the case when PP^T and QQ^T have d distinct eigenvalues λ_i . Then the dimension of the corresponding eigenspaces $\text{Ker}(PP^T - \lambda_i I)$ will be one, and therefore if $\{\mathbf{w}_i\}$ is an orthonormal system of eigenvectors for QQ^T , then $\epsilon_i \mathbf{w}_i = A\mathbf{v}_i$ with $\epsilon_i = \pm 1$. The latter can be written as $WE = AV$, namely $A = WEV^T$. Since $Q_\pi = AP$ and A is orthogonal matrix, we have $\langle \mathbf{p}_k - \bar{\mathbf{p}}, \mathbf{v}_i \rangle = \langle A(\mathbf{p}_k - \bar{\mathbf{p}}), A\mathbf{v}_i \rangle = \epsilon_i \langle \mathbf{q}_{\pi(k)} - \bar{\mathbf{q}}, \mathbf{w}_i \rangle$, and the proof is completed. \square

Note that if $\langle \mathbf{p}_k - \bar{\mathbf{p}}, \mathbf{v}_i \rangle = 0$ for every $k = 1, \dots, N$, ϵ_i cannot be determined from (3). If this happens, then $P^T \mathbf{v}_i = \mathbf{0}$, and therefore $PP^T \mathbf{v}_i = \mathbf{0}$. This means that \mathbf{v}_i is the eigenvector that corresponds to the eigenvalue 0, namely $i = 1$ and $\lambda_1 = 0$. Thus, if $0 < \lambda_1 < \dots < \lambda_d$, which happens if $\text{rank}(P) = d$, then there is at least one k which may depend on i , such that $\langle \mathbf{p}_k - \bar{\mathbf{p}}, \mathbf{v}_i \rangle \neq 0$, and we have $\epsilon_i = \langle \mathbf{p}_k - \bar{\mathbf{p}}, \mathbf{v}_i \rangle / \langle \mathbf{q}_{\pi(k)} - \bar{\mathbf{q}}, \mathbf{w}_i \rangle$. In this case the matrix $E = \text{diag}(\epsilon_1, \dots, \epsilon_d)$, and $A = WEV^T$ is completely determined if π is known.

Since the point collections are not labeled, we do not know π and could not use the above formula unless we go through all possible $N!$ choices for π . But that would be just an application of the well known PCA for each of the $N!$ choices of π , which is not computationally efficient. Our goal is to find E , and therefore A , without an apriori knowledge of the permutation π , under the assumption that all eigenvalues of PP^T are distinct.

Let $L_i^-(\mathcal{P}) := \{|\langle \mathbf{p}_k - \bar{\mathbf{p}}, \mathbf{v}_i \rangle| : 1 \leq k \leq N, \langle \mathbf{p}_k - \bar{\mathbf{p}}, \mathbf{v}_i \rangle < 0\}$, $i = 1, \dots, d$, be the collection of the absolute values of all negative scalar products, and $L_i^+(\mathcal{P}) := \{\langle \mathbf{p}_k - \bar{\mathbf{p}}, \mathbf{v}_i \rangle : 1 \leq k \leq N, \langle \mathbf{p}_k - \bar{\mathbf{p}}, \mathbf{v}_i \rangle > 0\}$, be the collection of all positive scalar products, including their repetitions. We similarly define $L_i^-(\mathcal{Q})$ and $L_i^+(\mathcal{Q})$. If $\text{rank}(P) = d$, at least one of $L_i^-(\mathcal{P})$ or $L_i^+(\mathcal{P})$ will have at least one element. The same holds for $L_i^-(\mathcal{Q})$ and $L_i^+(\mathcal{Q})$. Let \mathcal{P} be equivalent to \mathcal{Q} . Then for any fixed $i = 1, \dots, d$, if $L_i^+(\mathcal{P}) \neq L_i^-(\mathcal{P})$, it follows from (3) that only one of the following two cases happens:

- either $L_i^+(\mathcal{P}) = L_i^+(\mathcal{Q})$ and $L_i^-(\mathcal{P}) = L_i^-(\mathcal{Q})$, and thus $\epsilon_i = 1$, or
- $L_i^+(\mathcal{P}) = L_i^-(\mathcal{Q})$ and $L_i^-(\mathcal{P}) = L_i^+(\mathcal{Q})$, and thus $\epsilon_i = -1$.

If there is i_0 such that $L_{i_0}^+(\mathcal{P}) = L_{i_0}^-(\mathcal{P})$, we cannot make the decision whether $L_{i_0}^+(\mathcal{P}) = L_{i_0}^+(\mathcal{Q})$ or $L_{i_0}^+(\mathcal{P}) = L_{i_0}^-(\mathcal{Q})$, and therefore determine whether $\epsilon_{i_0} = 1$

or $\epsilon_{i_0} = -1$. In this case, we should consider both cases. In general, we can have $m \leq d$ indices i, i_1, \dots, i_m , for which $L_{i_\ell}^+(\mathcal{P}) = L_{i_\ell}^-(\mathcal{P})$, $\ell = 1, \dots, m$, and we have to consider 2^m matrices E, E_1, \dots, E_{2^m} , corresponding to the various cases of ± 1 located at the positions described by these m indices. If $Q \neq WE_k V^T P$ for $k = 1, \dots, 2^m$, then \mathcal{Q} is not equivalent to \mathcal{P} . Otherwise, if there is k , such that $Q = WE_k V^T P$, they are equivalent, and $A = WE_k V^T$ is the matrix of equivalence.

Let us denote by $\mathcal{L} := \{\mathcal{P}\}$ a library of collections of N points in \mathbb{R}^d . Let \mathcal{M} be the subset of \mathcal{L} that contains all collections \mathcal{P} for which the eigenvalues $0 \leq \lambda_1 \leq \dots \leq \lambda_d$ of PP^T are distinct. Let \mathcal{Q} be a configuration of N points in \mathbb{R}^d that we need to match to a collection from the library \mathcal{L} . The algorithm, described below, is based on the above observations and always determines whether \mathcal{Q} is equivalent to a collection from \mathcal{M} and may or may not determine whether \mathcal{Q} is equivalent to a collection from $\mathcal{L} \setminus \mathcal{M}$.

We have performed several numerical experiments to test our algorithm. In our implementation, as it is usually done in practice, the equality in lines 1, 7, 9, 11 and 23 in Algorithm 1 has been substituted by ϵ -distance. For example, $Q = WEV^T P$ has been substituted by $\|Q - WEV^T P\| \leq \epsilon$, with ϵ ranging from 10^{-6} to 10^{-10} .

Test 1: For each pair (d, N) , $d \in \{2, 3, 4\}$, $N \in \{2^n : 3 \leq n \leq 10\}$, we have generated in random a library $\mathcal{L} = \mathcal{L}(d, N)$ of 4000 collections of N points in \mathbb{R}^d uniformly distributed inside the unit sphere. We next build a set $\mathcal{T} = \mathcal{T}(d, N)$ of point collections by first choosing (in random) 2000 collections from \mathcal{L} , each of which is subsequently shuffled and rotated (in random). For each collection $\mathcal{Q} \in \mathcal{T}$, we apply Algorithm 1 with \mathcal{P} exhausting all elements from \mathcal{L} until a match is found. The algorithm was able to match each \mathcal{Q} from \mathcal{T} to its respective collection in \mathcal{L} .

Test 2: For each pair (d, N) , $d \in \{2, 3, 4\}$, $N \in \{2^n : 3 \leq n \leq 10\}$, we generate in random a library $\mathcal{L} = \mathcal{L}(d, N)$ of 4000 collections of N points in \mathbb{R}^d uniformly distributed inside the unit sphere. We next generate the same way a set $\mathcal{T} = \mathcal{T}(d, N)$ of 2000 point collections, and for each collection $\mathcal{Q} \in \mathcal{T}$ apply Algorithm 1 with \mathcal{P} exhausting all elements from \mathcal{L} until a match is found. As expected, the algorithm was not able to find a match.

Note that the eigenvalues of the matrix PP^T cannot be computed exactly, as they are roots of a degree d polynomial. However, there are high precision algorithms with complexity $O(d^3)$ to compute the eigenvalue decomposition for Gramians, see [1]. Our algorithm requires the computation of PP^T , (complexity $O(d^2 N)$), its eigenvalue decomposition (complexity $O(d^3)$), the computation of the $2d$ vectors in \mathbb{R}^d , ($\langle \mathbf{p}_1 - \bar{\mathbf{p}}, \mathbf{v}_i \rangle, \dots, \langle \mathbf{p}_N - \bar{\mathbf{p}}, \mathbf{v}_i \rangle$), $i = 1, \dots, d$, and ($\langle \mathbf{q}_1 - \bar{\mathbf{q}}, \mathbf{w}_i \rangle, \dots, \langle \mathbf{q}_N - \bar{\mathbf{q}}, \mathbf{w}_i \rangle$), $i = 1, \dots, d$ (complexity $O(dN^2)$), the computation of at most 2^m matrices A , each with complexity at most $O(d^3)$, and the computation of at most 2^m matrices AP , each with complexity $O(d^2 N)$. Therefore, for large values of $N \approx d2^d$, our algorithm has complexity of $O(dN^2)$. Note that for every collection $\mathcal{P} \in \mathcal{L}$ the algorithm needs only the d

Algorithm 1 Decision and Orthogonal Matrix Computation

Require:
 $P, \{\lambda_i\}, \{\mathbf{v}_i\}, \{L_i^+(\mathcal{P})\}, \{L_i^-(\mathcal{P})\}.$
 $Q, \{\gamma_i\}, \{\mathbf{w}_i\}, \{L_i^+(\mathcal{Q})\}, \{L_i^-(\mathcal{Q})\}.$

% The eigenvalues should be given in increasing order.

Ensure:

res % Decision value. It may be true, false or inconclusive.

A % Orthogonal transformation, if res is true.

```

1: if  $\{\lambda_i\} \neq \{\gamma_i\}$  then
2:   return res  $\leftarrow$  false
3: else if  $\lambda_1 < \dots < \lambda_d$  then
4:    $f \leftarrow 1, i \leftarrow 0$ 
5:   while  $i < d$  and  $f = 1$  do
6:      $i \leftarrow i + 1$ 
7:     if  $\{L_i^+(\mathcal{P})\} = \{L_i^-(\mathcal{P})\} = \{L_i^+(\mathcal{Q})\} = \{L_i^-(\mathcal{Q})\}$  then
8:        $\epsilon_i = \pm 1$ 
9:     else if  $\{L_i^+(\mathcal{P})\} = \{L_i^+(\mathcal{Q})\}$  and  $\{L_i^-(\mathcal{P})\} = \{L_i^-(\mathcal{Q})\}$  then
10:       $\epsilon_i = +1$ 
11:    else if  $\{L_i^+(\mathcal{P})\} = \{L_i^-(\mathcal{Q})\}$  and  $\{L_i^-(\mathcal{P})\} = \{L_i^+(\mathcal{Q})\}$  then
12:      $\epsilon_i = -1$ 
13:    else
14:      $f \leftarrow 0$ 
15:    end if
16:  end while
17:  if  $f = 0$  then
18:    return res  $\leftarrow$  false
19:  else
20:     $W \leftarrow [\mathbf{w}_1, \dots, \mathbf{w}_d]$ 
21:     $V \leftarrow [\mathbf{v}_1, \dots, \mathbf{v}_d]$ 
22:     $E \leftarrow \text{diag}(\epsilon_1, \dots, \epsilon_d)$ 
23:    if  $Q = WEV^T P$  then
24:      return res  $\leftarrow$  true,  $A \leftarrow WEV^T$ 
25:    else
26:      return res  $\leftarrow$  false
27:    end if
28:  end if
29: else
30:   return res  $\leftarrow$  inconclusive
31: end if

```

eigenvalues of PP^T , and when they are distinct, the $d \times N$ matrix P , the $d \times d$ matrix V , and the d vectors $(\langle \mathbf{p}_1 - \bar{\mathbf{p}}, \mathbf{v}_i \rangle, \dots, \langle \mathbf{p}_N - \bar{\mathbf{p}}, \mathbf{v}_i \rangle)$, $i = 1, \dots, d$, which in total is at most $d(1 + d + 2N)$ numbers.

4. Robustness

In this section we investigate the problem of matching two labeled point configurations \mathcal{P} and \mathcal{Q} in \mathbb{R}^d in the presence of noise. The collections $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ and $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ are said to be ε -*identically equivalent*, if $|\|\mathbf{p}_i - \mathbf{p}_j\|^2 - \|\mathbf{q}_i - \mathbf{q}_j\|^2| \leq \varepsilon$ for all $i, j = 1, \dots, N$. The following statement, whose proof we omit, holds.

Lemma 4. (i) *If \mathcal{P} and \mathcal{Q} are ε -identically equivalent, then*

$$|\langle \mathbf{p}_i - \bar{\mathbf{p}}, \mathbf{p}_j - \bar{\mathbf{p}} \rangle - \langle \mathbf{q}_i - \bar{\mathbf{q}}, \mathbf{q}_j - \bar{\mathbf{q}} \rangle| \leq 2\varepsilon, \quad \forall 1 \leq i, j \leq N. \quad (4)$$

(ii) *If the Gramians for \mathcal{P} and \mathcal{Q} satisfy*

$$|\langle \mathbf{p}_i - \bar{\mathbf{p}}, \mathbf{p}_j - \bar{\mathbf{p}} \rangle - \langle \mathbf{q}_i - \bar{\mathbf{q}}, \mathbf{q}_j - \bar{\mathbf{q}} \rangle| \leq \varepsilon, \quad \forall 1 \leq i, j \leq N, \quad (5)$$

then \mathcal{P} and \mathcal{Q} are 4ε -identically equivalent.

Lemma 4 shows that if \mathcal{P} and \mathcal{Q} are $C\varepsilon$ -identically equivalent with C being a fixed constant, then their Gramians are close, namely $P^T P = Q^T Q + \varepsilon M$, where the entries m_{ij} of M are bounded by some positive constant c_0 , $|m_{ij}| \leq c_0$, and vice versa.

Next, we investigate whether a statement similar to Lemma 2 holds in the presence of noise, that is whether two point configurations \mathcal{P} and \mathcal{Q} are $C\varepsilon$ -equivalent if and only if there is an orthogonal matrix A for which AP is close to Q . The answer to this question is given in Theorem 3 and Theorem 4.

Theorem 3. *Let $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ and $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ be two collections of N points in \mathbb{R}^d , such that $\text{rank}(P) = d$, $\|\mathbf{p}_i - \bar{\mathbf{p}}\| \leq c$ and $\|\mathbf{q}_i - \bar{\mathbf{q}}\| \leq c$, $i = 1, \dots, N$, $c = \text{const}$. Let*

$$P^T P = Q^T Q + \varepsilon M, \quad (6)$$

where $M = (m_{ij})$ is a matrix with bounded entries, $|m_{ij}| \leq c_0$, $c_0 = \text{const}$, and

$$0 < \varepsilon \leq (1 - \delta)(\|(PP^T)^{-1}\|Nc_0)^{-1} \quad (7)$$

for some $0 < \delta < 1$. Then there exists an orthogonal matrix $A = A(\varepsilon)$, such that $\|AP - Q\| \leq \varepsilon \tilde{c} N^2$, with a constant \tilde{c} ,

$$\tilde{c} = cc_0 \|(PP^T)^{-1}\| (2 + c \|(PP^T)^{-1}\|^{1/2} \delta^{-1/2}).$$

Proof. First, we will derive estimates for the norms of some of the matrices we consider. It is easily seen that $\|P\| \leq c\sqrt{N}$ (and similarly, $\|Q\| \leq c\sqrt{N}$), since

$$\|P\|^2 = \lambda_{\max}(P^T P) \leq \text{trace}(P^T P) = \sum_{i=1}^N \|\mathbf{p}_i - \bar{\mathbf{p}}\|^2 \leq c^2 N,$$

and that $\|M\| \leq c_0 N$. Note that PP^T is invertible since $d = \text{rank}(P) = \text{rank}(PP^T)$, and we can consider the matrices $P^T(PP^T)^{-1}$ and $(PP^T)^{-1}P$. We have $(PP^T)^{-1}P = (P^T(PP^T)^{-1})^T$, thus $\|(PP^T)^{-1}P\| = \|P^T(PP^T)^{-1}\|$. Clearly, $(P^T(PP^T)^{-1})^T(P^T(PP^T)^{-1}) = (PP^T)^{-1}$, and we have

$$\|P^T(PP^T)^{-1}\|^2 = \lambda_{\max}[(PP^T)^{-1}].$$

This result, combined with the fact that

$$\begin{aligned} \|(PP^T)^{-1}\|^2 &= \lambda_{\max} [((PP^T)^{-1})^T(PP^T)^{-1}] \\ &= \lambda_{\max} [((PP^T)^{-1})^2] = (\lambda_{\max}[(PP^T)^{-1}])^2, \end{aligned}$$

gives

$$\|(PP^T)^{-1}P\| = \|P^T(PP^T)^{-1}\| = \sqrt{\|(PP^T)^{-1}\|}. \quad (8)$$

It follows from (6) that

$$I = (PP^T)^{-1}PQ^TQP^T(PP^T)^{-1} + \varepsilon(PP^T)^{-1}PMP^T(PP^T)^{-1}.$$

As it was done in Lemma 2, we construct the matrix $B := QP^T(PP^T)^{-1}$, set $L := (PP^T)^{-1}PMP^T(PP^T)^{-1}$, and rewrite the above representation of I as

$$I = B^T B + \varepsilon L. \quad (9)$$

Using the bounds for the norms of Q and M and (8), we obtain

$$\|B\| \leq \|Q\| \cdot \|P^T(PP^T)^{-1}\| \leq c\sqrt{N}\|(PP^T)^{-1}\|^{1/2} \quad (10)$$

and

$$\|L\| \leq \|P^T(PP^T)^{-1}\|^2 \cdot \|M\| \leq c_0 N \|(PP^T)^{-1}\|. \quad (11)$$

Notice that B is not an orthogonal matrix, and thus cannot be a candidate for an equivalence matrix. However, we can modify B in order to obtain an orthogonal matrix. Let $0 \leq \beta_1(\varepsilon) \leq \dots \leq \beta_d(\varepsilon)$ be the eigenvalues of $B^T B$ and U_B be the orthogonal matrix such that

$$U_B^T B^T B U_B = \text{diag}(\beta_1(\varepsilon), \dots, \beta_d(\varepsilon)).$$

If $\mathbf{x}_i \in \mathbb{R}^d$, $\|\mathbf{x}_i\| = 1$ is the eigenvector corresponding to $\beta_i(\varepsilon)$, then

$$\begin{aligned} \beta_i(\varepsilon) &= \|B^T B \mathbf{x}_i\| = \|\mathbf{x}_i - \varepsilon L \mathbf{x}_i\| \geq \|\mathbf{x}_i\| - \|\varepsilon L \mathbf{x}_i\| \geq \|\mathbf{x}_i\| - \varepsilon \|L\| \cdot \|\mathbf{x}_i\| \\ &= 1 - \varepsilon \|L\| \geq 1 - (1 - \delta) (c_0 N \|(PP^T)^{-1}\|)^{-1} \cdot c_0 N \|(PP^T)^{-1}\| = \delta, \end{aligned}$$

where we have used (11) and the inequality for ε .

Next, we give an explicit construction of the equivalence matrix A . First, we set

$$\Lambda = \Lambda(\varepsilon) := \text{diag}(\beta_1(\varepsilon)^{-1/2}, \dots, \beta_d(\varepsilon)^{-1/2}),$$

and consider the matrix

$$A = A(\varepsilon) := BU_B \Lambda U_B^T. \quad (12)$$

Note that A is orthogonal since

$$A^T A = U_B \Lambda U_B^T B^T B U_B \Lambda U_B^T = U_B \Lambda \text{diag}(\beta_1(\varepsilon), \dots, \beta_d(\varepsilon)) \Lambda U_B^T = I.$$

Moreover, we will show that $\|AP - Q\| \leq \varepsilon \tilde{c} N^2$ for the constant \tilde{c} defined in the theorem. We write

$$AP - Q = (A - B)P + (BP - Q) = BU_B(\Lambda - I)U_B^T P + (BP - Q), \quad (13)$$

and compute $BP - Q$. Multiplication of (6) on the right by P^T leads to $P^T P P^T = Q^T Q P^T + \varepsilon M P^T$, which gives $P^T = Q^T B + \varepsilon M P^T (P P^T)^{-1}$. We have

$$P = B^T Q + \varepsilon (P P^T)^{-1} P M^T,$$

and therefore

$$BP - Q = (B B^T - I)Q + \varepsilon B (P P^T)^{-1} P M^T = \varepsilon (B (P P^T)^{-1} P M^T - LQ), \quad (14)$$

where in the last equality we have used (9). It follows from (13) and (14) that

$$\|AP - Q\| \leq \|BU_B(\Lambda - I)U_B^T P\| + \varepsilon \|B (P P^T)^{-1} P M^T - LQ\|. \quad (15)$$

We next estimate each of the norms on the right-hand side. Note that using the lower bound for $\beta_i(\varepsilon)$ and (11), we have

$$\begin{aligned} \|\Lambda - I\| &= \max_{i=1, \dots, d} |\beta_i(\varepsilon)^{-1/2} - 1| \\ &\leq \delta^{-1/2} \cdot \max_{i=1, \dots, d} |1 - \beta_i(\varepsilon)| \leq \delta^{-1/2} \|I - B^T B\| \\ &= \varepsilon \|L\| \delta^{-1/2} \leq \varepsilon c_0 N \|(P P^T)^{-1}\| \delta^{-1/2}, \end{aligned}$$

and therefore

$$\|BU_B(\Lambda - I)U_B^T P\| \leq \|B\| \cdot \|\Lambda - I\| \cdot \|P\| \leq \varepsilon \delta^{-1/2} c_0 c^2 N^2 \|(P P^T)^{-1}\|^{3/2}.$$

The second norm in (15) is evaluated, using (8), (10), (11) and the bounds for the norms of M and Q , as follows:

$$\|B (P P^T)^{-1} P M^T - LQ\| \leq \|B (P P^T)^{-1} P M^T\| + \|LQ\| \leq 2c c_0 N^2 \|(P P^T)^{-1}\|.$$

The last two inequalities result in

$$\|AP - Q\| \leq \varepsilon c c_0 \|(PP^T)^{-1}\| (c\delta^{-1/2} \|(PP^T)^{-1}\|^{1/2} + 2) N^2,$$

and the proof is completed. \square

Unfortunately, the converse of this theorem is in general false. The following theorem holds.

Theorem 4. *For any positive constants c, c_0, \tilde{c} , and for any $0 < \varepsilon \leq \frac{1}{2}c^2c_0^{-1}$, we can find d, N , two collections of N points $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ and $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ in \mathbb{R}^d and an orthogonal matrix A , such that $\text{rank}(P) = d$, $\|\mathbf{p}_i - \bar{\mathbf{p}}\| \leq c$ and $\|\mathbf{q}_i - \bar{\mathbf{q}}\| \leq c$ for $i = 1, \dots, N$, $\|AP - Q\| \leq \varepsilon \tilde{c} N^2$, but $P^T P = Q^T Q + \varepsilon M$, where M is a matrix for which for at least one entry $|m_{ij}| \geq c_0$.*

Proof. Consider any constants c, c_0, \tilde{c}, d a positive integer, and $N = 2d$. Let $\{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ be the canonical basis for \mathbb{R}^d , and let $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_N\}$ be the collections of points, where $\mathbf{p}_{2i-1} = \frac{c}{\sqrt{2}}\mathbf{e}_i$ and $\mathbf{p}_{2i} = -\frac{c}{\sqrt{2}}\mathbf{e}_i$ for $i = 1, \dots, d$, and $\mathcal{Q} = \{\mathbf{q}_1, \dots, \mathbf{q}_N\}$ be such that $\mathbf{q}_1 = \mathbf{q}_2 = \mathbf{0}$ and $\mathbf{q}_j = \mathbf{p}_j$, $j = 2, \dots, N$. A simple computation yields $\bar{\mathbf{p}} = \bar{\mathbf{q}} = \mathbf{0}$, and therefore $\|\mathbf{p}_i - \bar{\mathbf{p}}\| = \|\mathbf{p}_i\| = c/\sqrt{2} < c$. Similarly, $\|\mathbf{q}_i - \bar{\mathbf{q}}\| = \|\mathbf{q}_i\| < c$. If $A = I$, then $AP - Q = P - Q$,

$$(P - Q)^T(P - Q) = \frac{c^2}{2} \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

and therefore $\lambda_{\max}((P - Q)^T(P - Q)) = c^2$, which gives that $\|P - Q\| = c$. We compute directly that

$$P^T P - Q^T Q = \frac{c^2}{2} \begin{pmatrix} 1 & -1 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} = \varepsilon M.$$

For any $\varepsilon \leq \frac{1}{2}c^2c_0^{-1}$, we have that $|m_{11}| = \frac{c^2}{2\varepsilon} \geq c_0$, but if $2d = N$ is large enough, namely $2d = N \geq c^{1/2}(\varepsilon\tilde{c})^{-1/2}$, we have that $\varepsilon\tilde{c}N^2 \geq c = \|AP - Q\|$, and the proof is completed. \square

Finally, we verify the theoretical results from Theorem 3 by performing a series of numerical experiments. For each pair (d, N) , $d \in \{2, 3, 4\}$, $N \in \{2^n : 3 \leq n \leq 10\}$, we choose in random 1000 collections \mathcal{P} of N uniformly

N	d = 2	d = 3	d = 4
8	0.9468	0.7066	0.5333
16	0.8403	0.6470	0.4670
32	0.6114	0.5323	0.3984
64	0.4622	0.4595	0.3724
128	0.3829	0.3626	0.2839
256	0.2709	0.2327	0.2128
512	0.2002	0.1760	0.1566
1024	0.1415	0.1295	0.1182

Table 1. Simulation results.

distributed points inside the unit sphere and 1000 collections \mathcal{P} of N uniformly distributed points on the unit sphere. We select $\delta = 0.9$, $c_0 = 1$, $\varepsilon = \frac{1}{2}(1 - \delta)(\|(PP^T)^{-1}\|Nc_0)^{-1} = 0.05(\|(PP^T)^{-1}\|N)^{-1}$. The matrix Q is then generated so that (6) holds, the matrix A is computed according to (12), and $c := \max_{i=1, \dots, N} \{\|\mathbf{p}_i - \bar{\mathbf{p}}\|, \|\mathbf{q}_i - \bar{\mathbf{q}}\|\}$. For each pair (d, N) the biggest ratio $\|AP - Q\|/(\varepsilon\tilde{c}N^2)$ (among the 2000 choices) is recorded in Table 1. It is clearly seen that the empirical results confirm our theoretical bound.

Bibliography

- [1] J. DONGARRA AND D. SORENSSEN, A fully parallel algorithm for the symmetric eigenvalue problem, *SIAM J. Sci. and Stat. Comput.* **8** (1987), 139–154.
- [2] L. DORST, First order error propagation of the Procrustes method for 3D attitude estimation, *IEEE Trans. Pattern Anal. and Machine Intell.* **27** (2005), no. 2, 221–229.
- [3] J. GOWER, Generalized Procrustes analysis, *Psychometrika* **40** (1975), 33–51.
- [4] J. GOWER AND G. DIJKSTERHUIS, Procrustes problems, *Oxford Statistical Science Series*, Oxford University Press, 2004.
- [5] R. HANSON AND M. NORRIS, Analysis of measurements based on the singular value decomposition, *SIAM J. Sci. Stat. Comput.* **2** (1981) 363–373.
- [6] N. HIGHAM, Functions of matrices: theory and computation, *SIAM*, 2008.
- [7] B. HORN, H. HILDEN AND S. NEGAHDARIPOUR, Closed-form solution of absolute orientation using orthonormal matrices, *J. Optical Soc. of America* **5** (1988), 1127–1135.
- [8] S. RUSINKIEWICZ AND M. LEVOY, Efficient variants of the ICP algorithm, *Proceedings International Conference on 3-D Digital Imaging and Modeling*, **3** (2001), 145–152.
- [9] P. SCHÖNEMANN, A generalized solution of the orthogonal Procrustes problem, *Psychometrika* **31** (1966) 1–10.

- [10] I. SÖDERKVIST, Perturbation analysis of the orthogonal Procrustes problem, *BIT* **33** (1993), no. 4, 687–694.
- [11] Z. ZHANG, Iterative point matching for registration of free-form curves and surfaces, *Int. J. Comp. Vision* **13** (1994), no. 2, 119–152.

DAVID JIMÉNEZ

Department of Mathematics

Universidad de Costa Rica

San Jose

COSTA RICA

E-mail: david.jimenezlopez@ucr.ac.cr

GUERGANA PETROVA

Department of Mathematics

Texas A&M University

College Station, Texas 77843

USA

E-mail: gpetrova@math.tamu.edu