

Articulated Shape Matching Using Laplacian Eigenfunctions and Unsupervised Point Registration

Diana Mateus* Radu Horaud David Knossow Fabio Cuzzolin† Edmond Boyer
INRIA Rhône-Alpes
655 avenue de l'Europe- Montbonnot
38 334 Saint Ismier Cedex France
diana.mateus@inrialpes.fr

Abstract

Matching articulated shapes represented by voxel-sets reduces to maximal sub-graph isomorphism when each set is described by a weighted graph. Spectral graph theory can be used to map these graphs onto lower dimensional spaces and match shapes by aligning their embeddings in virtue of their invariance to change of pose. Classical graph isomorphism schemes relying on the ordering of the eigenvalues to align the eigenspaces fail when handling large data-sets or noisy data. We derive a new formulation that finds the best alignment between two congruent K -dimensional sets of points by selecting the best subset of eigenfunctions of the Laplacian matrix. The selection is done by matching eigenfunction signatures built with histograms, and the retained set provides a smart initialization for the alignment problem with a considerable impact on the overall performance. Dense shape matching casted into graph matching reduces then, to point registration of embeddings under orthogonal transformations; the registration is solved using the framework of unsupervised clustering and the EM algorithm. Maximal subset matching of non identical shapes is handled by defining an appropriate outlier class. Experimental results on challenging examples show how the algorithm naturally treats changes of topology, shape variations and different sampling densities.

1. Introduction

Shape matching is a central topic in many areas of computational vision, such as object modeling, motion tracking, object and action recognition. One of the main goals of shape matching is to find dense correspondences between the representations of two objects (as illustrated on Figure 1.

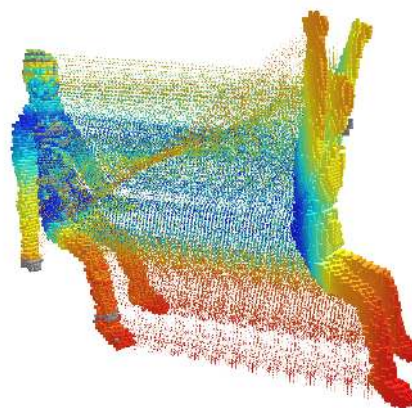


Figure 1. Two articulated shapes, of 12577 and 12346 voxels, matched with our method. The two voxel-sets are embedded in a 7-dimensional space. Using eigenfunction alignment to initialize the 7×7 orthogonal transformation, EM converges after 9 iterations. There are only 90 unmatched voxels (in grey).

In the recent past there has been tremendous interest in both 2-D and 3-D shape matching. In spite of these efforts, *the problem of matching 3-D articulated shapes* remains very difficult, mainly because it is not yet clear how to choose and characterize the group of transformations under which such shapes should be studied [9], [11], [1], [10]. One possible approach, among many others, is to represent shapes by locally connected sets of points, i.e. *sparse* graphs, and to use spectral embedding methods in order to isometrically map these graphs onto a lower dimensional space. As a result, a dense match between shapes can be found through the alignment of their embeddings. Umeyama [19] proposed a solution to the weighted graph matching problem based on eigendecomposition of the graphs' adjacency matrices, and Scott and Longuet-Higgins [16] who used the Gaussian proximity matrix to match sets of rigid points, pioneered this research area. The proximity matrix was also used in [17] while [5, 6] extended the Gaussian kernel to other robust weight functions and proposed a probabilistic

*is a Ph.D fellow of Marie-Curie Early Stage Training action Visitor.

†F.C. is a post-doctoral fellow of Marie-Curie research training network Visiontrain.

framework for point-to-point matching. The main drawback of these embedding techniques is that they cannot be easily extended to articulated shapes. Also they are quite sensitive to noise, to outliers, as well as to discrepancies between the two point-sets (e.g. missing parts, deformations, etc.).

To extend the the use of spectral graph methods to the unsupervised representation and matching of articulated shapes, one can use local isometries as invariants preserved by articulated motion (except in the vicinities of the joint). Methods such as Laplacian Eigenmaps [3], generally used for dimensionality reduction, are explicitly conceived to find a mapping that best preserves such pairwise relationships. Indeed, the Laplacian matrix [7] encodes *local* geometric and topological properties of a graph, and hence it is well suited for describing articulated objects. Belkin and Niyogi [3] provide a theoretical justification for combining the Laplacian operator with the Gaussian (or heat) kernel. Laplacian embedding has strong links with spectral clustering [14] and with other local-based embedding methods such as LLE [15]. In [10], the Laplacian embedding is used to embed 3-D meshes using a global geodesic distance and to match the embeddings using the ICP algorithm [21].

In this paper we propose a new method for matching articulated shapes which combines Laplacian embedding with probabilistic point matching, [8, 12, 20]. We thoroughly justify the choice of the Laplacian for capturing the properties of locally-rigid shapes, for mapping them onto a low-dimensional subspace of the eigenspace, and for performing probabilistic point registration. Problems with the ordering of eigenvalues have been identified before [10] with point-sets of cardinality 10^2 . We offer an interesting alternative to eigenvalue ordering, namely the use of eigenfunction histograms. Our new alignment method is particularly relevant when the cardinality of the point-sets is very large (of the order of 10^4) while the dimension of the embedding space is of the order of 10. Since we conclude that articulated shape matching is equivalent to point-to-point assignment under an orthogonal transformation, we concentrate on the latter and provide an algorithm that handles point registration in the framework of unsupervised clustering. We claim that, by adding a uniform component to a mixture of Gaussians, it is possible to deal with discrepancies between the shapes and with outliers.

The remainder of this paper is organized as follows. Section 2 establishes a link between graph isomorphism and point registration. Section 3 describes a method for aligning eigenfunctions based on their histograms. Section 4 casts the point registration problem in the framework of unsupervised clustering. Section 5 describes experiments with several data sets, and Section 6 draws some conclusions and proposes directions for future work.

2. Problem statement

2.1. Graph Laplacian operator

Given a voxel-set representation of an articulated shape, let $\mathcal{X} = \{X_1, \dots, X_N\}$ be the set of voxel centers ($X_i \in \mathbb{R}^3$). It is possible to define a graph \mathcal{G}_x , where each node is associated with a X_i and where the local geometry of the set is used to define the edges. The voxel connectivity together with a notion of local distance between pairs of nodes determines a weighted adjacency matrix \mathbf{W} with entries:

$$W_{ij} = \begin{cases} \exp(-d^2(i, j)/\nu^2) & \text{if } i \in \mathcal{N}(j) \\ 0 & \text{if } i \notin \mathcal{N}(j) \\ 0 & \text{if } i = j \end{cases} \quad (1)$$

where $d(i, j)$ denotes the *local* shortest-path distance between points associated with nodes i and j , ν is a scalar, and $\mathcal{N}(j)$ is the set of neighbors of j . In order to analyze some of the geometric and topological properties of these graphs (e.g., patterns of connectivity), we can use spectral graph theory [7], and in particular, the normalized graph Laplacian $\mathbf{L} = \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$. Here \mathbf{D} stands for the diagonal degree matrix whose elements verify $D_{ii} = \sum_j W_{ij}$.

2.2. Spectral graph matching

Consider another articulated shape $\mathcal{Y} = \{Y_1, \dots, Y_M\}$ and let for the time being $N = M$. From \mathcal{X} and \mathcal{Y} we build two graphs \mathcal{G}_x and \mathcal{G}_y and two adjacency matrices \mathbf{L}_x and \mathbf{L}_y . Matching these graphs is the problem of finding a $N \times N$ permutation matrix \mathbf{P} minimizing the function:

$$J(\mathbf{P}) = \|\mathbf{L}_x - \mathbf{P}\mathbf{L}_y\mathbf{P}^\top\|^2. \quad (2)$$

According to Umeyama's theorem [19], the solution can be found through the alignment of the graphs' eigenspaces, as follows. Let $\mathbf{L}_x = \mathbf{U}_x\mathbf{\Lambda}_x\mathbf{U}_x^\top$ and $\mathbf{L}_y = \mathbf{U}_y\mathbf{\Lambda}_y\mathbf{U}_y^\top$ be the eigendecompositions of the corresponding adjacency matrices, where \mathbf{U}_x and \mathbf{U}_y are two orthogonal matrices and $\mathbf{\Lambda}_x = \text{Diag}[\lambda_1, \dots, \lambda_N]$, $\mathbf{\Lambda}_y = \text{Diag}[\delta_1, \dots, \delta_M]$. First, the domain of the objective function (2) is extended to the space of all orthogonal matrices \mathbf{Q} . This extension is natural because permutation matrices belong to the group of orthogonal matrices. Umeyama's theorem states that *if the eigenvalues of \mathbf{L}_x and \mathbf{L}_y are distinct and if they can be ordered*, then the minimum of $J(\mathbf{Q})$ is reached for $\mathbf{Q}^* = \mathbf{U}_x\mathbf{S}\mathbf{U}_y^\top$. The diagonal matrix $\mathbf{S} = \text{Diag}[s_1, \dots, s_N]$, $s \in \{+1, -1\}$ accounts for the sign ambiguity in the eigendecomposition. The matrix \mathbf{Q}^* is an alignment of the complete eigenbasis. For two perfectly isometric shapes, \mathbf{Q}^* is also identical to the searched permutation matrix \mathbf{P} and determines a node-to-node assignment function (acting on the indexes of one node sets) $\omega : \{1, \dots, M\} \mapsto \{\omega(1), \dots, \omega(M)\}$.

One of our main *contributions*, is the extension of this theorem in order to solve the maximum sub-graph matching problem through the registration of the graphs' embeddings under an orthogonal transformation.

2.3. Graph matching in a reduced eigenspace

For *large* and *sparse* graphs, Umeyama's theorem holds only *weakly*. Indeed, one cannot guarantee that the eigenvalues of the Laplacian matrix are all distinct. Furthermore, the presence of symmetries in the shape causes some of eigenvalues to have an algebraic multiplicity greater than one. Under these circumstances and due to numerical approximations, it might *not* be possible to properly order the eigenvalues. In the absence of an absolute eigenvalue ordering, Umeyama's theorem is impractical (one would need to find the best out of all possible $N!$ orderings).

One elegant way to overcome this problem, is to reduce the dimension of the eigenspace, along the line of spectral nonlinear reduction techniques. Following the Laplacian eigenmaps scheme [3], one can perform a generalized eigendecomposition of the Laplacian matrix \mathbf{L} and use the K smallest eigenvalues¹ with their associated eigenspace ($K \ll N$). We denote by \mathbf{U}_x^K the $N \times K$ block matrix of \mathbf{U}_x . The columns of \mathbf{U}_x^K correspond to the K eigenvectors associated with the selected eigenvalues, while its rows represent the coordinates of the set $\bar{\mathcal{X}}$ in the embedded K -dimensional space: $\bar{\mathcal{X}} = \{x_1, \dots, x_N\}$. Since $X_i \in \mathbb{R}^3$ and $x_i \in \mathbb{R}^K$, this corresponds to a mapping $\mathbb{R}^3 \mapsto \mathbb{R}^K$. Respectively, for the second shape \mathcal{Y} , we can use the matrix \mathbf{U}_y^K to derive the embedding $\bar{\mathcal{Y}} = \{y_1, \dots, y_M\}$.

Because it preserves local geometry, the Laplacian embedding projects a pair of locally isomorphic shapes (true for articulated shapes) onto two congruent point-sets in \mathbb{R}^K . This condition makes it suitable for Umeyama's framework. If the K eigenvalues of the reduced embeddings were distinct and reliably ordered, one could directly use Umeyama's theorem and the minimizer of $J(\mathbf{Q})$ in the reduced eigenspace:

$$\mathbf{Q}^* = \mathbf{U}_y^K \mathbf{S}^K \mathbf{U}_x^{K\top}. \quad (3)$$

Notice that \mathbf{S}^K is now $K \times K$. If we can not reliably order the eigenvalues: $\{\lambda_1, \dots, \lambda_K\}$, $\{\delta_1, \dots, \delta_K\}$, we need to add a permutation \mathbf{P}^K :

$$\mathbf{Q}^* = \mathbf{U}_y^K \mathbf{S}^K \mathbf{P}^K \mathbf{U}_x^{K\top}. \quad (4)$$

Again, \mathbf{P}^K is only $K \times K$. Let $\mathbf{R}_K = \mathbf{S}^K \mathbf{P}^K$ and extend the domain of \mathbf{R}_K to all possible orthogonal matrices of

¹leaving out the eigenvalue 0 and its unitary eigenvector

size $K \times K$. This is done both, to find a close-form solution and to deal with algebraic multiplicities. As a result:

$$\mathbf{Q}^* = \mathbf{U}_y^K \mathbf{R}^K \mathbf{U}_x^{K\top}, \quad (5)$$

\mathbf{R}^K works as an orthogonal transformation aligning the K -dimensional coordinates of the two point-sets:

$$\tilde{\mathbf{U}}_x^{K\top} = \mathbf{R}^K \mathbf{U}_y^{K\top}. \quad (6)$$

where the column vectors of $\tilde{\mathbf{U}}_x^{K\top} = [x_{\omega(1)} \dots x_{\omega(M)}]$ are assigned to the column vectors of $\mathbf{U}_y^{K\top} = [y_1 \dots y_M]$, with $\omega : \{1, \dots, M\} \mapsto \{\omega(1), \dots, \omega(M)\}$. Under an exact local-isomorphism assumption, the alignment matrix \mathbf{R}^K makes the estimation of ω trivial (it can be performed by a simple nearest-neighbor method). To conclude, we can state the following proposition:

Proposition: Let two articulated shapes be described by two graphs as defined in section 2.1. Consider the Laplacian embeddings of these two graphs onto a K -dimensional space. Solving for the local graph isomorphism problem is equivalent to finding a solution for the point registration problem under the group of orthogonal transformations. Namely, estimate a $K \times K$ orthogonal matrix \mathbf{R}^K that aligns the reduced eigenspaces and find the trivial assignment ω .

Nevertheless, in real scenarios the two shapes are never locally identical and the number of voxels in each set can be different. To be able to recover the assignment ω , the problem needs to be reformulated as a *maximum sub-graph isomorphism*, i.e. finding the largest match between sub-graphs in the two shapes. In practice, this amounts to relaxing the constraints of the assignment ω being strictly one-to-one and the two graphs having the same number of nodes. One possible solution is to devise an algorithm that alternates between estimating the $K \times K$ orthogonal transformation \mathbf{R}^K , which aligns the K -dimensional coordinates of the two points sets, and finding an assignment ω . One may observe that \mathbf{R}^K belongs to the orthogonal group $O(K)$, therefore this framework is different from the classical rigid point registration, where one seeks for a 3-D rotation, i.e. a member of the special orthogonal group $SO(3)$.

Since an alternation approach may lead to local-minima, we propose a solution to the *maximum sub-shape isomorphism* problem in two stages. First, in section 3, we describe a method that estimates the best matrices \mathbf{P}^K and \mathbf{S}^K , relying neither on eigenvalue ordering nor on exhaustive search, but instead in a *comparison of the eigenfunction histograms* of the two embeddings. Then, in section 4, we detail a robust point registration method, based on maximum likelihood with latent variables. The algorithm follows the clustering formulation of EM, and is initialized with the \mathbf{P}^K and \mathbf{S}^K obtained in the previous stage.

3. Alignment using Laplacian eigenfunctions

In the context of embedded representations of graphs, the columns of matrix \mathbf{U}_x^K may be viewed as a set of eigenfunctions, each such eigenfunction maps the graph \mathcal{G}_x onto \mathbb{R} . Similarly a column of \mathbf{U}_y^K maps \mathcal{G}_y onto \mathbb{R} . For two identical graphs, these eigenfunctions should be identical, up to the ordering of the eigenvalues (which implies an ordering of the column vectors of both \mathbf{U}_x^K and \mathbf{U}_y^K) and up to a node-to-node assignment. However, empirical evidence points out that eigenvalue ordering is not reliable (hence the presence of matrix \mathbf{P}^K) and we do not have a node-to-node assignment. Let \mathbf{u}_x^k (respectively \mathbf{u}_y^k) be the k^{th} eigenfunction of the Laplacian associated with the graph \mathcal{G}_x (respectively \mathcal{G}_y). \mathbf{u}_x^k can be seen as a column vector, whose histogram $h(\mathbf{u}_x^k)$ is obviously invariant with respect to the order of the vector's components, and thus to the order in which the graph nodes are considered. The histogram can be regarded then as an *eigenfunction signature*.

The problem of finding an estimate for the matrices \mathbf{P}^K and \mathbf{S}^K can therefore be addressed as the problem of finding a set of assignments $\{\mathbf{u}_x^k \Leftrightarrow \pm \mathbf{u}_y^l, 1 \leq k, l \leq K\}$ based on the comparison of the eigenfunction signatures, namely the histograms. This is an instance of the standard bipartite maximum matching problem whose complexity is $O(K^3)$. Notice however that the eigenvectors are defined up to a sign (hence the presence of matrix \mathbf{S}^K) so two different histograms $h(\mathbf{u})$ and $h(-\mathbf{u})$ can be associated with each eigenfunction and need to be compared. Let $C(h(\mathbf{u}), h(\mathbf{v}))$ be a measure of dissimilarity between two histograms. Computing the dissimilarity of all pairs of eigenfunctions ($\mathbf{u}_x^k, \pm \mathbf{u}_y^l$) we can build a $K \times K$ matrix \mathbf{A} whose entries are defined by:

$$A_{kl} = \min\{C(h(\mathbf{u}_x^k), h(\mathbf{u}_y^l)), C(h(\mathbf{u}_x^k), h(-\mathbf{u}_y^l))\}$$

The Hungarian algorithm provides an optimal solution to the problem of finding an assignment between eigenfunctions of the two graphs (shapes), taking \mathbf{A} as input and producing as output a permutation matrix. Because the sign ambiguity has been explicitly taken into account, this solves as well for \mathbf{S}^K . This method provides an good initialization for the point registration method described in section 4.

As an example, Figure 2 shows the histograms of the first five eigenfunctions associated with the two different poses of the wooden mannequin of Figure 3. One can visually realize the striking similarity of histograms related to extremely separated poses of the same shape: as a matter of fact the topology of the body is different in the two cases due to the presence of the hands' self-contact. The effectiveness of the histograms as eigenfunction signatures is attested by the correct assignments represented as arrows. Notice that, due to the topological differences, it has not been possible to align all the eigenfunctions. This is not an

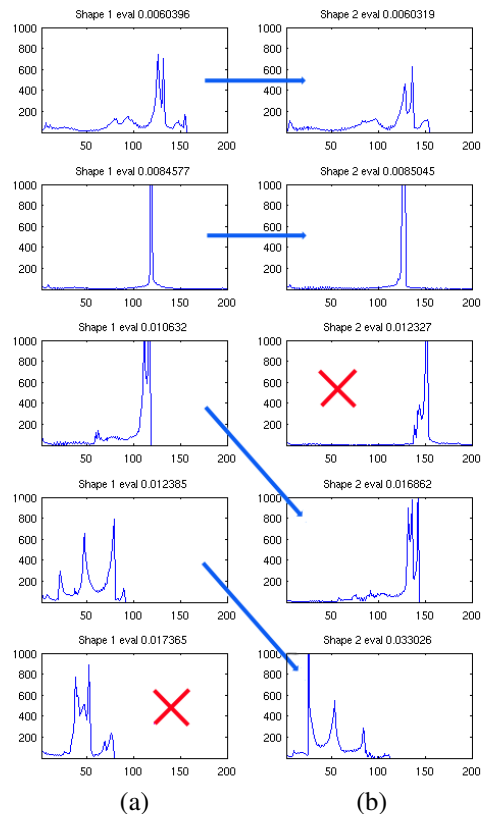


Figure 2. Histograms associated with the embeddings of the wooden-mannequin example. (a) pose 1 (b) pose 2. First 5 histograms out of the 20 compared. Notice the sign flip when the fourth histogram on the left is aligned with the fifth histogram on the right (horizontal reflection).

issue since we only retain the aligned eigenfunctions. These form an eigenspace where finding the assignment is easier. In practice, the algorithm starts with the most significant, e.g. the first 20-25, eigenvalues and eigenfunctions for each shape, but retains only only the pairs of eigenfunctions with the best alignment score. This improves the robustness of the matching algorithm.

4. Point registration and clustering

Points belonging to two poses of the same articulated shape are congruent after graph Laplacian embedding. To find an assignment between them we need to find an optimal alignment between two congruent sets of points in the reduced space of dimension K . An immediate consequence of the result obtained in section 2 is that articulated shape matching casts into a *point registration* problem, which we propose to solve in the framework of *clustering*. The points \mathcal{X} are treated as observations, while the points in the second shape, \mathcal{Y} , are treated as centers of normally distributed clusters. In addition to these Gaussian clusters, we will need to consider an *outlier component* with uniform distribution. We refer then to our method as *unsupervised robust point registration*.

Under these conditions, the likelihood of an observation x_n to belong to a cluster m , $1 \leq m \leq M$, has a Gaussian distribution with mean $^2 \mu_m = \mathbf{R}y_m$ and covariance Σ . We also introduce a $(M + 1)^{\text{th}}$ *outlier cluster* with uniform distribution. The likelihood of an observation to be an outlier is uniformly distributed over the volume V which contains the embedded shapes. This yields:

$$P(x_n|z_n = m) = \mathcal{N}(x_n|\mathbf{R}y_m, \Sigma), 1 \leq m \leq M,$$

$$P(x_n|z_n = M + 1) = \mathcal{U}(x_n|V, 0).$$

One can write the likelihood of an observation as a mixture of these $M + 1$ distributions weighted by their priors:

$$P(x_n) = \sum_{m=1}^{M+1} \pi_m P(x_n|z_n = m).$$

We introduce the latent variables $\mathcal{Z} = \{z_1, \dots, z_N\}$, that assign each observation to a cluster. By assuming independent and identically distributed observations, we obtain the log-likelihood of the point-set $\bar{\mathcal{X}}$:

$$\ln P(x_1, \dots, x_N) = \sum_{n=1}^N \ln \sum_{m=1}^{M+1} \pi_m P(x_n|z_n = m). \quad (7)$$

Hence, the problem of point-to-point assignment can be formulated as the maximization of the log-likelihood in eq. (7). Due to the presence of the latent variables this will be carried out using the EM algorithm. With our notations, EM should solve:

$$\mathbf{R}^* = \arg \max_{\mathbf{R}} E_{\mathcal{Z}}[\ln P_{\mathbf{R}}(\bar{\mathcal{X}}, \mathcal{Z})|\bar{\mathcal{X}}]. \quad (8)$$

That is, maximize the *conditional expectation* taken over \mathcal{Z} of the joint log-likelihood (of observations and assignments) *given the observations*. The joint likelihood is parameterized by \mathbf{R} . In contrast to standard EM, we estimate a *global* orthogonal transformation \mathbf{R} (instead of M independent K -dimensional means), as well as a *global* $K \times K$ covariance matrix Σ , constrained to be common to all the clusters (in the interest of avoiding convergence problems when a point x_n is infinitely close to a cluster center $\mathbf{R}y_m$).

In order to formally derive an analytic expression for the conditional expectation of eq. (8), we need to compute the posterior class probabilities, namely $P(z_n = m|x_n) = P(z_n = m, x_n)/P(x_n)$. This can be interpreted as the posterior probability of a point x_n to be registered with point y_m and we denote it by α_{nm} . We assume that all points m have equal prior probabilities and hence $\pi_1 = \dots = \pi_M = \pi_{in}$. Therefore we have $\pi_{M+1} = 1 - M\pi_{in} = \pi_{out}$. Denoting d_{Σ} the Mahalanobis distance³, we obtain:

$$\alpha_{nm} = \frac{\exp(-d_{\Sigma}(x_n, \mathbf{R}y_m))}{\sum_{i=1}^M \exp(-d_{\Sigma}(x_n, \mathbf{R}y_i) + \kappa)}. \quad (9)$$

²For simplicity we drop the superscript of \mathbf{R}^K .

³ $d_{\Sigma}(\mathbf{a}, \mathbf{b}) = (\mathbf{a} - \mathbf{b})^{\top} \Sigma^{-1} (\mathbf{a} - \mathbf{b})$

The parameter κ represents the contribution of the outlier cluster; it is possible to verify that: $\kappa = (2\pi)^{K/2} (\det \Sigma)^{1/2} \pi_{out} / V \pi_{in}$. Next, we write $P(\bar{\mathcal{X}}, \mathcal{Z})$ as:

$$P(\bar{\mathcal{X}}, \mathcal{Z}) = \prod_{n=1}^N \prod_{m=1}^{M+1} (P(x_n|z_n = m)P(z_n = m))^{\delta_m(z_n)},$$

where the function $\delta_m(z_n)$ is equal to 1 if $z_n = m$ and to 0 otherwise. When taking the conditional expectation of the log of this likelihood, we need to estimate the conditional expectation of the functions δ_m . One may notice that:

$$E[\delta_m(z_n)|\mathcal{X}] = \sum_{i=1}^{M+1} \delta_m(z_n = i)P(z_n = i|x_n) = \alpha_{nm}.$$

Finally, denoting the current parameter estimate with superscript (q) and dropping out the constant terms, the negative expectation in eq. (8) can be written as:

$$E(\mathbf{R}|\mathbf{R}^{(q)}) = \sum_{m=1}^M \xi_m \left(d_{\Sigma}(\bar{x}_m, \mathbf{R}y_m) + \ln \det \Sigma \right), \quad (10)$$

where,

$$\bar{x}_m = \sum_{i=1}^N \alpha_{im} x_i / \xi_m \quad \text{and} \quad \xi_m = \sum_i \alpha_{im}. \quad (11)$$

Our algorithm finds one-to-one assignments $\bar{x}_m \Leftrightarrow y_m$. \bar{x}_m is a mean over all $x_i \in \bar{\mathcal{X}}$, built by weighting each x_i with its posterior probability of being assigned to $y_m \in \bar{\mathcal{Y}}$, and normalizing each such assignment by ξ_m . In practice, these posteriors (α_{nm}) and these weights (ξ_m) allow the *classification* of the observations into inliers and outliers. For this reason, our algorithm performs *unsupervised robust point registration*.

To summarize, the formal derivation outlined above guarantees that the maximization of eq. (7) is equivalent to the minimization of eq. (10) [13]. Therefore, one can apply the EM algorithm to eq. (10) as detailed in the next section.

5. Experimental results

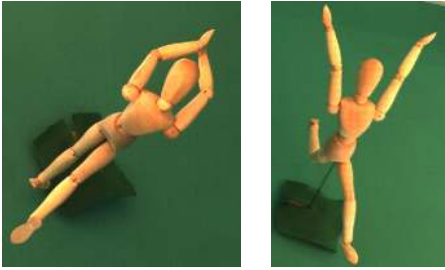
We applied the method described above to estimate the matching of 3-D articulated shapes described as dense sets of voxels. In this section, we show the results of the algorithm on three different objects: a wooden mannequin, a person, and a hand (see Table 5). We gathered several data sets of poses of these objects using six calibrated cameras and a voxel reconstruction software package. The results show how most of the points are matched in few iterations of the algorithm.

Figure 3 shows two widely separated poses of the same articulated object, a wooden mannequin. The algorithm

Example	EM iterations	Number of voxels	Outliers+ Unmatched
Mannequin	6	12667 - 13848	472
Manneq-person	8	11636 - 12267	388
Hand	4	13594 - 13317	39

Table 1. Three challenging experiments

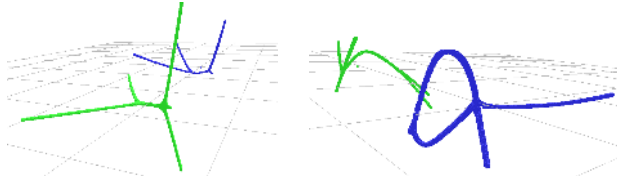
starts by building the Laplacian matrices of these shapes in order to map them onto a lower dimensional space (Figure 4). Then, the signatures of the associated eigenfunctions are computed and matched as described in Section 3 in order to find a permutation matrix between the eigenvectors associated with the two embeddings and solve for the sign ambiguity (see Figures 2 and 6-a). In this experiment, the dimension of the embedding space turned out to be $K = 5$. Notice how the contact between the two hands alters the topology of the embedded shape (Figure 4); without the eigenfunction matching stage, i.e. using only eigenvalue to order the eigenfunctions, the point matching becomes a difficult task and the EM algorithm would easily fall into a local minimum.



(a) pose 1

(b) pose 2

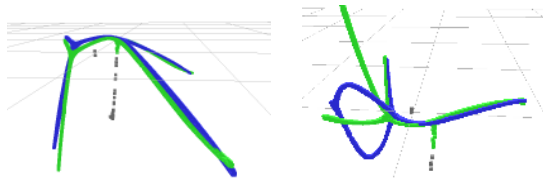
Figure 3. Two poses of a wooden mannequin. In (a) the hands touch each other introducing important topological differences between the two associated graphs.



(a) eigenfunctions 1-2-3

(b) eigenfunction 1-4-5

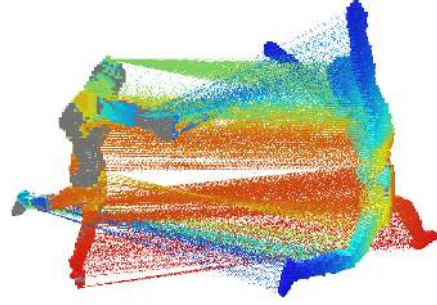
Figure 4. Embeddings before alignment: Embeddings of the two poses of the wooden mannequin in Figure 3 (blue: pose 1, green: pose 2). The embeddings are represented as their three dimensional projections on two different subspaces. (a) subspace spanned by eigenfunctions 1-2-3, (b) subspace 1-4-5.



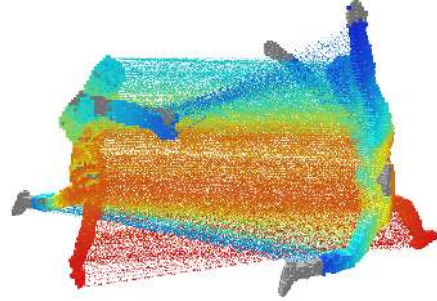
(a) eigenfunctions 1-2-3

(b) eigenfunction 1-4-5

Figure 5. Aligned embeddings: result after eigenfunction alignment and EM. Only the K retained eigenfunctions are shown.



(a) After eigenfunction alignment (EM initialization)



(b) After EM convergence

Figure 6. Dense match between the two poses of the wooden mannequin in Fig. 3. (a) Assignment found after aligning the eigenspaces with the eigenfunction signature histograms. (b) Final result after convergence of EM. Points in grey illustrate outliers.

Then, the algorithm performs point registration under a $K \times K$ orthogonal transformation, using the EM algorithm described in Section 4. This procedure treats the first point-set as observations and the second point-set as cluster centers. EM needs to be initialized, by provision of an initial orthogonal transformation \mathbf{R} and a covariance matrix Σ , as well as values for the cluster priors $\pi_1 \dots \pi_M$ (see below). We use the eigenfunction alignment scheme of Section 3 to provide an initialization for \mathbf{R} . We also use an isotropic covariance $\sigma^2 \mathbf{I}$ for the initialization, where σ is chosen sufficiently large to allow evolution of the algorithm and \mathbf{I} is the $K \times K$ identity matrix. In practice the initial covariance is of the order of 15 voxels. The cluster priors are chosen such that $\pi_1 = \dots = \pi_M = \pi_{in}$ and $\pi_{M+1} = 1 - M\pi_{in} = \pi_{out}$. Moreover π_{in} is chosen as a small proportion of the working volume V .

The *E-step* computes the posterior probabilities associated with each observation using eq. (9); in addition to the current estimates of \mathbf{R} and Σ , one needs to specify the uniform-component parameter κ . With the explained choice for π_{in} and π_{out} we have $\kappa \propto (\det \Sigma)^{1/2}$. The *M-step* estimates the transformation \mathbf{R} and the covariance Σ . The estimation of \mathbf{R} is done in closed form using an extension of [2] to deal with orthogonal matrices of arbitrary dimension rather than with 3-D rotations. The covariance matrix can be easily estimated from eq. (10) as is classically done [4]. At convergence, EM provides an optimal

value for the maximum likelihood, as well as final estimates for the posterior probabilities. This posteriors assign each “mean” observation \bar{x}_m (eq. 11) to a “cluster center” y_m .

Figure 5 shows the final alignment of the two embedded shapes of Figure 4. Even though the hands self-contact induces significant differences in the graph topology and thus in the embedded shapes, there is still a significant set of voxels which preserves the same structure in both poses. Our sub-graph matching algorithm is capable to recover this maximal subset and successfully match the point-sets, despite the presence of a large number of unmatched points. This performance is due the Laplacian embedding, but mainly to the outlier rejection mechanism introduced in Section 4, which rejects this unmatched points to a class of their own and prevents them to influence the correctness of the orthogonal alignment.

Figure 6 shows the registered sets of voxels after the eigenfunction alignment and EM. It is interesting to see how the initial registration, based on aligning the eigenfunctions of the Laplacian, is already capable to provide a good assignment for the voxels in the limbs. This has a simple justification in terms of spectral clustering, since each eigenfunction corresponds to a well-identified group of voxels.

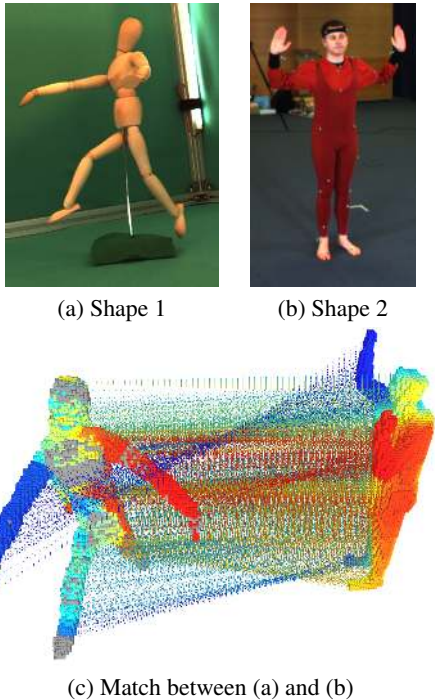


Figure 7. Matching the wooden mannequin to a person is possible in our framework, as the maximal sub-graph isomorphism is sought. As long as two shapes have a common sub-structure, the algorithm is normally able to find it.

Figure 7 illustrates the capability of the proposed algorithm to handle shapes of *different datasets*. It shows the matching between the wooden mannequin and a person.

Even though they have roughly the same “shape” (described as a graph whose nodes are voxels), the relative dimensions of the limbs are not the same, the shape of the rigid links forming them is also altered, and finally they are sampled with a different number of voxels. Nevertheless our methodology delivers remarkable results.

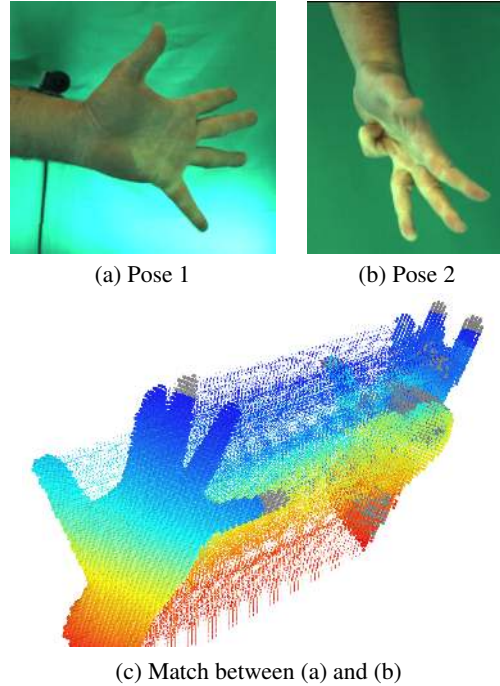


Figure 8. Matching two different poses of a hand showing bending and a different type of self-contact. Points in the extended finger are matched to points in the bent finger. The unmatched region in the palm reflects the local change of the connectivity.

In a final test, Figure 8 shows an alignment between two different poses of a hand. The bending of the small finger creates a different type of self-contact. The algorithm solves for most of the matches and in particular finds appropriate matchings for the small finger. The unmatched region in the palm reflects the change in the connectivity between the palm and the finger which alters the embedded shapes.

5.1. Discussion

The good performance of the algorithm is the result of several choices. First, the use of voxel-sets as input for our experiments instead of other shape representations such as meshes; indeed, due to the statistical properties of our large datasets, the voxel representations leads to highly stable embeddings, they are easier to generate and to maintain than surfaces, and allow to exploit the full volumetric information of the shape. Furthermore, the fact that they are regularly sampled contributes to the convergence of the Laplacian embedding towards a “geometric-aware” eigenbasis [3]. Second, the definition of a *local* graph-connectivity allows the treatment of the otherwise difficult

cases related to self-contacts and topology changes. These cannot be handled when completely connected graphs are defined since a self-contact would imply changes in all the pairwise distances. Furthermore, locality gives rise to very sparse Laplacian matrices and thus to an efficient calculation of the eigendecomposition (we use ARPACK). Additionally, the initialization is usually good for EM to converge in a few iterations. As a consequence, the procedure is time-efficient, and that whole matching procedure of large voxel-sets takes only a few seconds. Finally, it is important to mention that our method can be applied to other type of data since it relies only on geometric cues and does not use the photometric information as suggested in [18].

6. Conclusions

This paper describes a new method for establishing dense correspondences between the points associated with two articulated objects. We address the problem using both spectral matching and unsupervised point registration. We formally introduce graph isomorphism using the Laplacian matrix, and we provide an analysis of the matching problem whenever the number of nodes in the graph (or, equivalently the number of voxels in the shape) is very large, i.e. of the order of 10^4 . We show that there is a simple equivalence between graph isomorphism and point registration under the group of orthogonal transformations, when the dimension of the embedding space is much smaller than the cardinality of the point-sets.

The eigenvalues of a large sparse Laplacian cannot be reliably ordered. We propose an elegant alternative to eigenvalue ordering, using eigenfunction histograms and alignment based on comparing these histograms. The point registration that results from eigenfunction alignment yields an excellent initialization for the EM algorithm, subsequently used only to refine the registration. Indeed, each eigenfunction corresponds to a “mode” within the framework of spectral clustering.

Moreover, the proposed EM algorithm differs significantly from previously proposed methods: it leads to a general purpose unsupervised robust point registration method. The algorithm can deal with large discrepancies between the two sets of points, by incorporating a uniform component in the mixture model. Eventually, it assigns an optimal “mean” observation to each cluster center, which amounts to a dense point-to-point matching.

In the future we plan to study more thoroughly the link between spectral matching and spectral clustering. We believe that this may solve the problem of dimension selection for the embedding space and that this is equivalent to choosing the number of modes in spectral clustering algorithms. The combination of spectral matching with probabilistic methods, along the lines described in this paper,

can be applied to many other representations, such as 3-D meshes, 2-D silhouettes, bags of features, and so forth.

References

- [1] D. Anguelov, P. Srinivasan, H.-C. Pang, D. Koller, S. Thrun, and J. Davis. The correlated correspondence algorithm for unsupervised registration of nonrigid surfaces. In *NIPS*, 2004.
- [2] K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3-d point sets. *PAMI*, 9(5):698–700, 1987.
- [3] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. In *Neural Computation*, 2003.
- [4] C. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [5] M. Carcassoni and E. R. Hancock. Correspondence matching with modal clusters. *PAMI*, 25(12):1609–1615, 2003.
- [6] M. Carcassoni and E. R. Hancock. Spectral correspondence for point pattern matching. *Pattern Recognition*, 36:193–204, 2003.
- [7] F. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.
- [8] G. Dewaele, F. Devernay, R. Horaud, and F. Forbes. The alignment between 3-D data and articulated shapes with bending surfaces. In *ECCV*. Springer, 2006.
- [9] A. Elad and R. Kimmel. On bending invariant signatures for surfaces. *PAMI*, 25(10):1285–1295, 2003.
- [10] V. Jain and H. Zhang. Robust 3d shape correspondence in the spectral domain. In *IEEE International Conference on Shape Modeling and Applications*, 2006.
- [11] C. W. Jenkins, O. C. Jenkins, and M. J. Mataric. Markerless kinematic model and motion capture from volume sequences. In *CVPR*, 2003.
- [12] B. Luo and E. Hancock. A unified framework for alignment and correspondence. *CVIU*, 92(1):26–55, October 2003.
- [13] G. McLachlan and T. Krishnan. *The EM Algorithm and Extensions*. Wiley, New York, 1997.
- [14] A. Ng, M. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *NIPS*, 2002.
- [15] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.
- [16] G. Scott and C. L. Higgins. An algorithm for associating the features of two images. *Biological Sciences*, 244:21–26, 1991.
- [17] L. Shapiro and J.M.Brady. Feature-based correspondence an eigenvector approach. In *Image and Vision Computing*, volume 10, pages 283–288, 1992.
- [18] J. Starck and A. Hilton. Correspondence labelling for wide-tiemframe free-form surface matching. In *ICCV*, 2007.
- [19] S. Umeyama. An eigen decomposition approach to weighted graph matching problems. *PAMI*, 10:695–703, 1988.
- [20] W. Wells III. Statistical approaches to feature-based object recognition. *IJCV*, 28(1/2):63–98, 1997.
- [21] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *IJCV*, 13:119–152, 1994.