Learning Object Correspondences with the Observed Transport Shape Measure

Abstract. We propose a learning method which introduces explicit knowledge to the object correspondence problem. Our approach makes use of an a priori learning set to compute a dense correspondence field between two objects, where the characteristics of the field bear close resemblance to those in the learning set. We introduce a new local shape measure we dub the “observed transport measure”, whose properties make it particularly amenable to the matching problem. From the values of our measure obtained at every point of the objects to be matched, we compute a distance matrix which embeds the correspondence problem in a highly expressive and redundant construct and facilitates its manipulation. We present two learning strategies that rely on the distance matrix and discuss their applications to the matching of a variety of 1-D, 2-D and 3-D objects, including the corpus callosum and ventricular surfaces.

1 Introduction

From signal processing to pattern recognition, the issue of object matching permeates a broad range of image related fields. In computer vision for instance, the search for target patterns often requires matching a given template to pictorial elements in an input image [1]. In medical imaging, the objects may be instances of a given anatomical structure, for which a statistical model, a shape average, or a segmentation is desired [2]. In computer graphics, matched objects may be used to derive a series of intermediate shapes to “morph” one object into the other [3], etc. In this paper, we approach the issue of object matching as a process of computing a dense correspondence field between two objects.

At a glance, defining a correspondence between two objects entails finding in them pairs of corresponding elements that share particular similarities, in terms of shape, position, or both. More formally, given two objects $O_1$ and $O_2$ with any a priori parameterizations represented by two functions $O_1$ and $O_2$:

$$O_1 : I_1 \subset \mathbb{R}^m \rightarrow \mathbb{R}^n \quad O_2 : I_2 \subset \mathbb{R}^m \rightarrow \mathbb{R}^n (m \leq n)$$

we are looking for a reparameterization of $O_1$ and $O_2$, that is, for two diffeomorphisms $f_1 : I_1 \rightarrow I_1$ and $f_2 : I_2 \rightarrow I_2$, such that $O_1^* = O_1 \circ f_1$ and $O_2^* = O_2 \circ f_2$ and

$$\forall x_1 \in I_1, \forall x_2 \in I_2, x_1 \text{ “close to” } x_2 \Rightarrow O_1^*(x_1) \text{ “very similar to” } O_2^*(x_2)$$

(1)

where “very similar to” is defined with respect to a given similarity metric.
Fig. 1. Illustration of the proposed matching framework in the case of 2-D parametric curves \((m = 1, n = 2)\)

To allow for multiple points in \(I_1\) to be matched to a single point in \(I_2\) and conversely, we restate our problem (see Figure 1) as that of finding a monotonically increasing and continuous function \(\varphi : I \subset \mathbb{R}^m \rightarrow I_1 \times I_2\)

\[
\begin{align*}
\forall x \in I, \quad & \, O_1(\varphi_1(x)) \text{ “very similar to” } O_2(\varphi_2(x)) \\
\text{such that:} & \\
& \varphi_1(x) \in \varphi_1^{-1}(x) \quad \text{and} \quad \varphi_2(x) \in \varphi_2^{-1}(x) \\
& \varphi_1^{-1}(x) \cap \varphi_2^{-1}(x) \neq \emptyset.
\end{align*}
\]

(2)

A number of automated methods for object matching have been presented in the literature, that tackle some or all of the above issues.

Trouvé and Younes detailed in [4] an axiomatic formulation for the 1-D matching problem; they introduced, among others, the concepts of symmetry (\(\varphi_{O_1} \circ O_2 \) should be the inverse of \(\varphi_{O_2} \circ O_1\)) and consistent self-matching (\(\forall\)Object \(O, \varphi_{O \rightarrow O} = (I_d, I_d)\); in the general case, \(\varphi\) should not be not too dissimilar from the identity) and proposed a matching framework for 2-D piecewise lines that satisfies their axioms. In [5], Cohen et al. compared the bending and stretching energies of one curve \((O_1)\) and a reparameterization of the other \((O_2)\) to find the best match. Fleuté et al. [6] minimized the Euclidean distance between an input shape and a registered template which assumed smooth transition paths in between them. Wang et al. [7] used geodesic interpolation to compute the dense correspondence field between two surfaces once an initial sparse set of corresponding points had been obtained with an automated shape-based matching algorithm. In [8], the first elliptical harmonics of the expansion of the input objects (which must have spheroidal shapes) served to establish a correspondence. In [9], Sebastian et al. used a dynamic programming approach similar to [4] to find the best match between two 2-D curves, using a similarity measure based on “alignment” between segments of the curves. Elastic registration and warping approaches have also been investigated. In [10], Thompson et al.
mapped the input surfaces to two spheres whose coordinates were then warped under anatomical feature curve constraints. Davatzikos et al. [11] also identified corresponding points on object boundaries in two images before aligning them using elastic warping. Along different lines, Davies et al. [12] presented a curve matching method, in the context of the search for the most compact statistical shape model. An information theoretic criterion was designed and controlled the correspondence between objects.

The common drawback of those approaches, despite their diversity, lies in their lack of control over the similarity measure introduced in equation 1, which is often defined a priori, once and for all, and uses only limited domain-based information (or information learned from examples). Typically, these matching processes can be reduced to the optimization of a functional whose minimum corresponds to a “good” correspondence field. The difficulty of designing an adequate functional comes from the difficulty of characterizing an adequate correspondence field. In [5] for instance, the authors assume that points with similar curvature should be matched. This may suit some applications, but is not always desirable. Figure 2 illustrates such a situation: here two corpus callosum have been delineated and we wish to compute their average shape:

- Suppose that part of the fornix (a2) has been improperly delineated together with the corpus callosum, then we would like segments \{a1, (b1,b2)\}, \{(a2,a3), b3\} and \{a4, b4\} to be matched, in spite of the fact that the curvature signature of segment a2 more closely resembles that of b2 than that of b3.
- On the other hand, we may decide to trust the delineation and assume that a lesion is the cause of the odd looking bulge (a1) in the corpus callosum in 2.a. Then, we would like a match: \{a1, b1\}, \{a2, b2\}, \{a3, b3\} and \{a4, b4\}.

Clearly, choosing between these 2 scenarios requires the introduction of explicit knowledge into the matching algorithm.

To overcome this issue, we propose a learning approach where a learning set helps the matching algorithm compute a correspondence field whose characteristics bear close resemblance to those of the a priori given learning correspondence fields. Our method relies on the use of a distance matrix derived from the values of a local shape measure which is computed on every pair of points of the objects to be matched. We argue that this shape distance matrix embeds the matching problem in a highly expressive and redundant construct which is more easily
manipulated. This matrix is both visually interesting (as it allows for visual inspection of the specific reparameterization problem at hand) and enables us to recast the matching problem as the search of a geodesic in another metrizable space: the space of reparameterizations (which is a group).

We introduce in section 2 the so-called “observed transport” shape measure and discuss the properties that make it particularly amenable to the matching problem. We then present the various learning techniques that we have developed in section 3 and discuss their applicability to 1-D, 2-D and 3-D objects along with some examples from medical imaging.

2 Observed Transport Local Shape Measure

We first define our local shape measure in a variety of cases before presenting some of its properties.

1-D case. Let \( C : I \subset \mathbb{R} \to \mathbb{R}^2 \) be a 2-D curve (open or closed), parameterized with respect to a scalar \( u \). We define the observed transport measure \( \rho_C \) as follows:

\[
\forall t \in I, \quad \rho_C (t) = \int_{V_C (t)} \left\| C (t) - C (u) \right\| \cdot \left| C' (u) \right| . du
\]

(3)

where \( V_C (t) \) is the arc of \( C \) “visible” within range \( r \in \mathbb{R}^+ \) from \( C (t) \):

\[
V_C (t) = \{ C (u) \ s.t. \ [C (t) C (u)] \cap C = \{ C (t) ; C (u) \} \text{ and } \left\| C (t) C (u) \right\| \leq r \}
\]

with \( [C (t) C (u)] \) the line segment between points \( C (t) \) and \( C (u) \).

\( \rho_C (t) \) can be regarded as the minimal total amount of work it takes to transport the elementary elements \( du \) with mass \( \left| C' (u) \right| . du \) that are visible within range \( r \) from point \( C (t) \), from their location \( C (u) \), to \( C (t) \) (in the fashion of a Monge-Kantorovitch transport problem [13]). Figure 3 displays (thick red lines) the arcs of \( C \) that are visible from point \( P \) at range \( r \), for a given vertebra outline.

2-D case. Let \( S : I^2 \subset \mathbb{R}^2 \to \mathbb{R}^3 \) be a 2-D surface, parameterized with respect to scalars \( u \) and \( v \). \( \rho_S \) becomes:

\[
\rho_S (\mathbf{u}', \mathbf{v}') = \int_{V_S (\mathbf{u}', \mathbf{v}')} \left\| S (\mathbf{u}, \mathbf{v}) - S (\mathbf{u}', \mathbf{v}') \right\| \cdot \left| \frac{\partial (x, y, z)}{\partial (u, v)} \right| . dudv
\]

(4)

where \( \left| \frac{\partial (x, y, z)}{\partial (u, v)} \right| \) is the Jacobian of \( S \), and \( V_S \) is defined analogously to the 2-D case.
Fig. 3. Observed Transport Measure principle

Discrete approximation. We define a discrete version of object $O$ as an unsorted collection of n-D points: $O = \{O_i \in \mathbb{R}^n\}_{i=1}^N$ (that is, we do not assume any a priori parameterization). We then use a centered space finite difference approximation to derive a discrete version of $\rho_O$ in n-D:

$$\forall i \in 1 \ldots N, \quad \frac{\rho_O(i)}{dg} = \sum_{j=1, O_j \in V_O(O_i)}^N \|O_i - O_j\|$$  \quad (5)

with $dg$ the grid step size in the n directions.

Example. Figure 4 shows how our local shape measure $\rho$ behaves on a few 2-D and 3-D objects. Curve (a) demonstrates the ability of our measure to model shape characteristics: even though $\rho$ evidently depends on the curvature of the curve at the point at which it is computed, $\rho(A) \neq \rho(B)$ and $\rho(C) \neq \rho(D)$, which correctly reflects the differences in the shape landscape surrounding those points. As such, the observed transport measure is both a measure of local shape and an indicator of context, that is, of where we are in the object (with large ranges $r$): for instance, it adequately discriminates between the belly and the back of the corpus callosum (Figure 4.d). Note that a curvature measure would not necessarily exhibit such behavior since for instance, in Figure 4.a, the curvatures at A and B, and at C and D, are the same.

Fig. 4. Observed transport measure (red is lowest, blue is highest) computed over: (a) a $u$-parameterized 2-D curve, (b) a set of 2-D points, (c) a $u,v$-parameterized 2-D surface (ventricle) and (d) a set of 3-D points (corpus callosum).
Properties.

- The observed transport measure is independent of (i.e., invariant to) reparameterization.
- Given an observed transport signature (a series of consecutive observed transport values), there is only one curve that can be reconstructed from it, modulo rotation and translation.
- It is invariant with respect to rigid transformation. However, it is not scale invariant as we believe the scale of an object is an important shape characteristic when trying to match anatomical structures. We could easily introduce scale invariance by normalizing it by the largest observed transport value across the entire object, or by using a scale parameter in subsequent optimizations.

3 Learning the Correspondence Field

We present in this section the learning algorithms we have developed to bias the search for a correspondence field between two objects towards instances that are admissible with respect to an a priori given learning set. We first briefly describe a non-learning algorithm before detailing how we can derive learning strategies from this first approach. We have tackled 3 distinct cases, to which all or only some of these methods can be applied:

1-D case: 2-D and 3-D u-parameterized curves: we consider objects defined on an interval of \( \mathbb{R} \), taking values in \( \mathbb{R}^2 \) or \( \mathbb{R}^3 \) respectively; \( m = 1, n \in \mathbb{N}_+ \) with notations of the first section.

2-D case: discrete 2-D point-set (unsorted collections of points of \( \mathbb{R}^2 \)) and u,v-parameterized 2-D surfaces; \( m = 2, n = 2 \) or 3.

3-D case: discrete 3-D point-set (unsorted collections of points of \( \mathbb{R}^3 \)); \( m = 3, n = 3 \).

3.1 Optimal Path in the Shape Distance Matrix [m=1, n\in\mathbb{N}_+]

Following the Trouvé approach [4], we define the best reparameterization \( \varphi_{C_1 \rightarrow C_2}^* \) between curves \( C_1 \) and \( C_2 \) to be that which minimizes the overall cumulative distance between measures computed for all pairs of matched points:

\[
\varphi_{C_1 \rightarrow C_2}^* = \arg \min_{\varphi} \left\{ \int \| \rho_{C_1} (\varphi_1 (u)) - \rho_{C_2} (\varphi_2 (u)) \| \, du \right\}
\]  

In the discrete case (and for piecewise linear curves, see [4] for details), a dynamic programming approach can be used to find the optimal reparameterization. Let \( D \) be the shape distance matrix associated with the curves \( C_1 = \{ C_1^i \}_{i=1}^{N_1} \) and \( C_2 = \{ C_2^j \}_{j=1}^{N_2} \):

\[
D = [d_{ij}]_{i=1 \ldots N_1}, \forall (i, j), d_{ij} = \| \rho_{C_1} (C_1^i) - \rho_{C_2} (C_2^j) \| \]  

\[
j = 1 \ldots N_2
\]
Fig. 5. Non-learning reparameterization: (a & b) reparameterized curves, (c) shape distance matrices and optimal paths (in red), (d) point by point average curves.

Finding the best reparameterization then boils down to finding in \( D \) (see Figure 1) the minimal cost path between points \( S \) (start) and \( E \) (end), which requires that a single matching pair \( (M_1 \in C_1, M_2 \in C_2) \) be given (for open curves, one could choose the extremities; this condition can also be relaxed if circular shifts are included in the optimization as well). A dynamic programming approach then yields an \( O(N_1.N_2) \) complexity.

Note that when a number of consecutive points have the same shape measure (in a circle for instance), there is not a unique best path with respect to the above criterion. To bias the search towards “natural” reparameterizations (the “consistent self-matching” axiom), we introduce in equation 6 a constraint to prevent the path from deviating too much from the diagonal of \( D \), i.e. for some \( \alpha \in [0, 1] \):

\[
\varphi_{C_1 \rightarrow C_2} = \arg \min_\varphi \alpha \cdot \int \left| \rho_{C_1}(\varphi_1(u)) - \rho_{C_2}(\varphi_2(u)) \right| \, du + (1 - \alpha) \cdot \int \left| \varphi_1(u).C_2'(u) - \varphi_2(u).C_1'(u) \right| \, du
\]  

Figure 5 displays some reparameterized curves along with the point by point averages derived from them. Corresponding pairs of points are rendered with the same color (we progress from dark blue to dark red in a “rainbow” fashion). Note in particular how the discriminating power of our shape measure enabled the triangular indentations to be correctly matched together in the first row, and against the corresponding points in the rectangle in the second row.
3.2 A Pattern Matching Approach to the 2-D Reparameterization Problem \( [m = 1, n = 2 \text{ or } 3] \)

An interesting feature of the shape distance matrix is that it embeds, in a highly redundant way, information about all possible reparameterizations between the two input objects. In Figure 5 for instance, we can notice clear patterns corresponding to the triangles on the first line. A local “matching scenario” (e.g., “discarding the fornix” in Figure 2, or “matching the triangles together” in Figure 5) then corresponds to a path in a sub-matrix extracted from the shape distance matrix of the objects. Note that even though our shape measure is independent of reparameterization, pairs of objects with different initial parameterizations will produce different looking shape distance matrices. Care should thus be taken to use the same (or similar) parameterization for the objects to be matched and the ones in the learning set.

We derive the corresponding algorithm (see Figure 6):

**Step 1 (a priori).** Given a number of desired local matching scenarios, a human operator first forms a learning set by selecting instances for each scenario (a careful process as the operator must ensure that the learning set adequately represents the desired matching characteristics). An instance consists of a 2-D sub-matrix \( M_{i,j} \in \mathcal{M}_{m_i \times n_i, n_j} \), and its associated connected path \( P_{i,j} = \{(x_{i,j}^k, y_{i,j}^k)\}_{k=1}^{m_i + n_j} \). The sub-matrices are extracted from shape distance matrices computed from objects which should be “similar” to the
ones the operator wants the algorithm to reparameterize. $P_{i,j}$ is the paths in $M_{i,j}$ which represents a local matching scenario, in the same fashion that the optimal cost path in Section 3.1 represents the “optimal” global matching scenario. For each instance, we also compute the distance map of its path.
Let $S_1 = \{S_{1,1}, \ldots, S_{1,N_1}\}, \ldots, S_K = \{S_{K,1}, \ldots, S_{K,N_K}\}$ be the $K$ scenarios, with their instances $S_{i,j} = (M_{i,j}, D_{i,j})$ where $M_{i,j}$ is the shape distance sub-matrix, and $D_{i,j}$ the associated distance map.

**Step 2.** Once we have computed the shape distance matrix $M \in M_{m,n}$ from the two input objects $O_1$ and $O_2$, a pattern matching algorithm is used to find in $M$ sub-matrices that bear close resemblance to those of the learning set. We have developed a straightforward multi-scale framework where each sub-matrix $M_{i,j}$ in the learning set is matched against sub-matrices, extracted from $M$, at a number of positions and scales. For each $M_{i,j}$, we record the translation $t^*_i$ and scale $s^*_j$ for which the maximal similarity is achieved:

$$
(t^*_i, s^*_j) = \arg \max_{t, s} \text{similarity} (M_{i,j}, M \cdot [t, t_x + s, m_{i,j}] \times [t_y, t_y + s, n_{i,j}])
$$

where $M \cdot [t, t_x + s, m_{i,j}] \times [t_y, t_y + s, n_{i,j}]$ is the sub-matrix of $M$ of size $s_n_{i,j} \times s_m_{i,j}$ which starts at index $t_x, t_y$ (with $t = [t_x, t_y]^T$). We also discard instances for which the associated similarity measure is too low.

**Step 3.** For each scenario in the learning set, we then average the distance maps of the paths associated with their instances (once we have applied the proper translation and scale from step #2). The averaging process is done pixel by pixel. In Figure 6, we average the maps of the two instances of scenario #2; no averaging is required for scenario #1 since it only has 1 instance.

**Step 4.** We then combine the average maps $D_{i,j}$ with the underlying shape distance matrix $M$ to bias the dynamic programming search towards the sub-paths from the learning set:

$$
M_{x,y}^* = M_{x,y} + \sum_{i=1}^{K} \sum_{j=1}^{N_i} \lambda_{i,j-1} [t^*_{i,j,x}, t^*_{i,j,x} + s^*_{i,j,m_{i,j}}] \times [t^*_{i,j,y}, t^*_{i,j,y} + s^*_{i,j,n_{i,j}}] (x, y).
$$

$$
D_{i,j}^* (x, y), \text{ with } t^*_{i,j} = [t^*_{i,j,x}, t^*_{i,j,y}]^T
$$

(9)

Note that the relative weight of the average distance maps with respect to the shape distance matrix $\lambda_{i,j}$ could be controlled by the quality of the match between the sub-matrices from the learning set and the matrix $M$. That quality could also be used to compute a weighted average distance map instead of an equal-weight one.

Figure 7 illustrates this approach on two geometric examples. In the first case (first line), we make sure to match triangles together, whereas in the second case (second line), we discard them as noise, and match them against the directly corresponding rectangle pieces. The learning set sub-matrices were taken from the matrices of Figure 5.

Incidentally, the same method can be used to rule out certain sub-matches. When a pattern in the learning set has no associated sub-path, its distance map is infinite everywhere and thus the dynamic programming algorithm will avoid the corresponding area in the shape distance matrix.
Fig. 7. Pattern matching examples: (a) learning set, (b & c) reparameterized curves (matching pairs of points are rendered with the same color), (d) the resulting point by point average curve

3.3 A Registration Approach to the n-D Reparameterization Problem \([m \in \mathbb{N}, n \in \mathbb{N}, m \leq n]\)

Even though noticeable patterns are still present in higher dimensional distance matrices, the lack of a single-scalar parameterization for n-D objects prevents us from using our pattern matching approach. However, we can still capitalize on the advantageous aspects of the shape distance matrix by considering the problem of reparameterization between two objects to be that of deforming and adapting a hyper-surface given a priori (associated to an a priori shape distance matrix) to the shape distance matrix of the input objects. In doing so, we avoid the issue of the parameterization of the input objects (and can thus consider collections of points).

The resulting algorithm is very similar to that of section 3.2:

- Given a number of 2m-D shape distance matrices computed from pairs of already matched objects (and their associated matching hyper-surfaces), we non-linearly register them to the shape distance matrix of the two input objects.
- The non-linear transforms are then applied to the distance maps of the hyper-surfaces associated to the learning items.
- Those transformed distance maps are then averaged and the zero-level set of the average map becomes the new reparameterization. Note that a matching criterion (the integral of the deformation field for instance) could be used to compute a weighted average.

With this approach, we transform an m-D matching issue into a 2m-D registration problem. Despite the curse of dimensionality, we are left with a simpler problem given the high expressivity of the distance matrices (see [14] for a similar dimension increase for surface matching). Consequently, the performance of our method depends on the robustness and accuracy of the non-linear registration algorithm. In the 1-D case (2-D shape distance matrix) we use the PASHA method [15] where the amount of regularization depends on the estimated discrepancy between the instances in the learning set and the objects to be reparameterized. We have adapted it in 4-D to treat the 2-D case (4-D shape
distance matrix). Even though extending it to 6-D (3-D case) is not theoretically impossible, the size of the search space makes the registration intractable. We are currently experimenting with sparse space techniques to tackle this.

Figure 8 shows how our registration method behaved on a series of 20 caudate nucleus (u,v-parameterized surface). One caudate was selected as a target and the remaining 19 others were resampled together with it, using a 2-item learning set built by an expert neuroanatomist. We show 3 sample caudates (out of the 20) and the resulting mean caudate (right most column), obtained by averaging homologous points across the resampled test set. Visual inspection confirmed the adequacy between the parameterization of the structures in the learning set and those in the test set.

### 3.4 Building the Learning Set

Our approaches require that the correspondence between the objects of the learning set be established \textit{a priori}. This may not be a trivial task for 3-D objects with complex shapes. However, it only has to be done once and for a small number of instances. Also a sparse subset of the correspondence field could be specified by the user to generate a learning set. Most of the fully automated techniques presented in the introduction could produce a meaningful set that could then be manually corrected if need be.

Note that using a learning set implies that the objects we want to reparameterize should not be too different from those in the learning set. In fact, similarities between objects do not matter so much as similarities between the pair of objects to be reparameterized and the pairs of objects in the learning set. Of course, the former is a sufficient condition for the latter. However, a unique advantage of our approach lies in its ability to learn a matching strategy for even very dissimilar objects, provided that we apply it to the same dissimilar matching situations.

### 4 Conclusion

We have presented a learning approach to the object correspondence problem. Our method makes adequate use of known correspondences from an \textit{a priori} learning set to compute between two given objects a dense correspondence field whose characteristics are similar to those of the learning set. This enables us to exert explicit control over the reparameterization. As such, this technique proves
useful to put in correspondence the “outliers” of an object set whose “ordinary” instances may be treated with direct non-learning algorithms.

We have also introduced a new local shape measure, the observed transport measure, and illustrated the highly discriminating properties that make it particularly amenable in this context.

Finally, technical difficulties (curse of dimensionality) prevented us from implementing our method for full 3-D objects. We are currently exploring alternative approaches to alleviate this problem.

References