

The Correlated Correspondence Algorithm for Unsupervised Registration of Nonrigid Surfaces

**** Stanford AI Lab Technical Report SAIL-2004-100 *****

Dragomir Anguelov*
Stanford University

Daphne Koller†
Stanford University

Praveen Srinivasan‡
Stanford University

Sebastian Thrun§
Stanford University

Hoi-Cheung Pang¶
Stanford University

James Davis||
Honda Research Labs

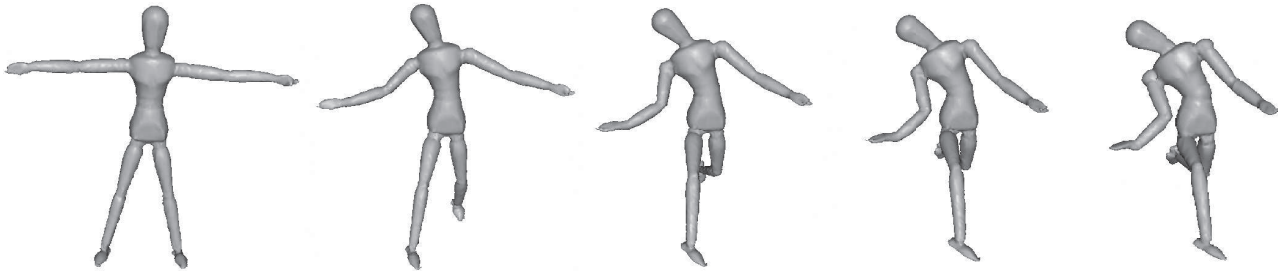


Figure 1: Several frames from a motion animation generated by interpolating two scans of a puppet (far left and far right), which were automatically registered using the Correlated Correspondence algorithm.

Abstract

We present an unsupervised algorithm for registering 3D surface scans of an object undergoing significant deformations. Our algorithm does not use markers, nor does it assume prior knowledge about object shape, the dynamics of its deformation, or scan alignment. The algorithm registers two meshes by optimizing a joint probabilistic model over all point-to-point correspondences between them. This model enforces preservation of local mesh geometry, as well as more global constraints that capture the preservation of geodesic distance between corresponding point pairs. The algorithm applies even when one of the meshes is an incomplete range scan; thus, it can be used to automatically fill in the remaining surfaces for this partial scan, even if those surfaces were previously only seen in a different configuration. We evaluate the algorithm on several real-world datasets, where we demonstrate good results in the presence of significant movement of articulated parts and non-rigid surface deformation. Finally, we show that the output of the algorithm can be used for compelling computer graphics tasks such as interpolation between two scans of a non-rigid object and automatic recovery of articulated object models.

1 Introduction

The construction of 3D object models is a key task for many graphics applications. It is becoming increasingly common to acquire these models from a range scan of a physical object. This paper deals with an important subproblem of this acquisition task —

the problem of registering two deformable surfaces corresponding to different configurations of the same non-rigid object.

The main difficulty in the 3D registration problem is determining the *correspondences* of points on one surface to points on the other. Local regions on the surface are rarely distinctive enough to determine the correct correspondence, whether because of noise in the scans, or because of symmetries in the object shape. Thus, the set of candidate correspondences to a given point is usually large. Determining the correspondence for all object points results in a combinatorially large search problem. The existing algorithms for deformable surface registration make the problem tractable by assuming significant prior knowledge about the objects being registered. Some rely on the presence of markers on the object [Allen et al. 2003], while others assume prior knowledge about the object dynamics [Lin 1999], or about the space of nonrigid deformations [Leventon 2000; Blanz and Vetter 1999]. Algorithms that make neither restriction [Shelton 2000; Hähnel et al. 2003] simplify the problem by decorrelating the choice of correspondences for the different points in the scan. However, this approximation is only good in the case when the object deformation is small; otherwise, it results in poor local maxima as nearby points in one scan are allowed to map to far-away points in the other.

Our algorithm defines a joint probabilistic model over all correspondences, which explicitly model the correlations between them — specifically, that nearby points in one mesh should map to nearby points in the other. Importantly, the notion of “nearby” used in our model is defined in terms of geodesic distance over the mesh, a more appropriate measure in this context than the standard Euclidean distance. We define a probabilistic model over the set of correspondences, that encodes these geodesic distance constraints as well as penalties for link twisting and stretching, and high-level local surface features [Johnson 1997]. We then apply *loopy belief propagation* [Yedidia et al. 2003] to this model, in order to solve for the entire set of correspondences simultaneously. The result is a registration that respects the surface geometry. To the best of our knowledge, the algorithm we present in this paper is the first

*e-mail: drago@cs.stanford.edu

†e-mail: koller@cs.stanford.edu

‡e-mail: praveens@cs.stanford.edu

§e-mail: thrun@stanford.edu

¶e-mail: hcpang@cs.stanford.edu

||e-mail: jedavis@graphics.stanford.edu

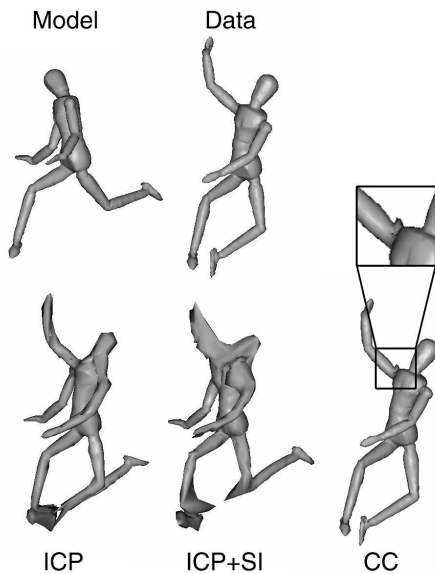


Figure 2: Registration results for meshes (Model) and (Data) using different algorithms. (ICP) Nonrigid ICP gets stuck in a local minimum, due to incorrect initial correspondences. Points on the head are mapped to the right arm, while points on the right shoulder are mapped to the head. (ICP+SI) Incorporating spin-images in the nonrigid ICP distance function does not address the problem of incorrect correspondences. (CC) The Correlated Correspondence algorithm produces a largely correct registration, although with an artefact in the right shoulder (inset).

algorithm which allows the registration of 3D surfaces of an object where the object configurations can vary significantly, there is no prior knowledge about object shape or dynamics of deformation, and nothing whatsoever is known about the object alignment. Moreover, unlike many methods, our algorithm can be used to register a partial scan to a complete model, greatly increasing its applicability.

We apply our approach to three datasets containing models of a wooden puppet, a human arm and entire human bodies in different configurations. The datasets consist of object models acquired with a 3D laser range scanner. We demonstrate very good registration results for scan pairs exhibiting articulated motion, non-rigid deformations, or both. We also describe three applications of our method. In our first application, we show how a partial scan of an object can be registered onto a fully specified model in a different configuration. The resulting registration allows us to use the model to “complete” the partial scan in a way that preserves the local surface geometry. In the second, we use the correspondences found by our algorithm to smoothly interpolate between two different poses of an object. In our final application, we use a set of registered scans of the same object in different positions to recover a decomposition of the object into approximately rigid parts, and recover an articulated skeleton linking the parts. All of these applications are done in an unsupervised way, using only the output of our Correlated Correspondence algorithm applied to pairs of poses with widely varying deformations, and unknown initial alignments. These results demonstrate the value of a high-quality solution to the registration problem to a range of graphics tasks.

2 Previous Work

Surface registration is a fundamental building block in computer graphics. The classical solution for registering rigid surfaces is the Iterative Closest Point algorithm (ICP) [Besl and McKay 1992; Chen and Medioni 1991; Rusinkiewicz and Levoy 2001]. Recently, there has been work extending ICP to non-rigid surfaces [Shelton 2000; Chui and Rangarajan 2000; Hähnel et al. 2003; Allen et al. 2003]. These algorithms treat one of the scans (usually a complete model of the surface) as a deformable template. The links between adjacent points on the surface can be thought of as springs, which are allowed to deform at a cost. Similarly to ICP, these algorithms iterate between two subproblems — estimating the non-rigid transformation Θ and estimating the set of point-to-point correspondences C between the scans. The step estimating the correspondences assumes that a good estimate of the nonrigid transformation Θ is available. Under this assumption, the assignments to the correspondence variables become decorrelated: each point in the second scan is associated with the nearest point (in the Euclidean distance sense) in the deformed template scan.

The nonrigid ICP framework, outlined above, allows the decomposition of the original problem into two subproblems — estimating the correspondences and estimating the transformation — allows efficient solutions for models containing a very high number of points. However, the decomposition also induces the algorithm’s main limitation. By assigning points in the second scan to points on the deformed model independently, nearby points in the scan can get associated to remote points in the model if the estimate of Θ is poor (Fig. 2).

Modeling applications that require the registration of complex 3D surfaces obtain a good initial estimate by placing a sparse set of markers on the scanned objects [Allen et al. 2002; Allen et al. 2003]. The markers allow the registration algorithm to obtain a good initial surface alignment, simplifying the correspondence problem.

In the absence of markers, several techniques have been found to alleviate the problem of incorrect initialization. [Shelton 2000] performs registration in a multi-resolution pyramid, and employs local features, such as color. The TPS-RPM method of Chui and Rangarajan [2000] maintains beliefs over the correspondence estimates, which are annealed to become more deterministic. This makes the algorithm more tolerant of incorrect initialization than the others in its class. However, TPS-RPM algorithm attempts to find a registration that preserves the Euclidean distances between all pairs of points, making it inappropriate for articulated objects where global Euclidean distances can change drastically. In general, although the above solutions can improve convergence, the applicability of nonrigid ICP methods remains largely limited to problems where the deformation is local, or the initial alignment is approximately correct.

Another set of approaches uses prior knowledge about the space of transformations an object can undergo. Given previously registered meshes from the same object class, they create a parametric representation of the surface variability. For this, principal component analysis is applied either to a set of registered meshes [Blanz and Vetter 1999; Allen et al. 2003] or to aligned volumetric representations such as active level sets [Leventon 2000]. A registration of a new surface in the same class to the model can be established by optimizing for the best alignment and for the best set of principal components describing the deformation of the model. Such algorithms are often quite sensitive to the initial alignment. Moreover, the types of deformations that can be well-encoded through linear PCA over points in Euclidean space is quite restricted. Thus, these approaches often work well for largely convex objects, but are unsuccessful at representing the deformation space of surfaces that have branching parts such as arms.

Our algorithm is most closely related to computer vision algorithms for non-rigid template matching. In the 3D case, this framework is used for detection of articulated object models in images [Huttenlocher and Felzenszwalb 2003; Yu et al. 2002; Sigal et al. 2003]. These algorithms assume the decomposition of the object into a relatively small number of parts is known, and that a detector for each object part is available. Like our algorithm, they optimize for a joint embedding of all articulated parts into the scene (usually an image). When the articulated models are tree-structured, efficient dynamic programming algorithms can give the most likely match [Huttenlocher and Felzenszwalb 2003]. When the correlation graph has loops, graph-partitioning algorithms can be applied [Yu et al. 2002]. When the orientation of an articulated 3D human body template is being inferred from image data, reasoning for both correspondence and orientation can be performed simultaneously for a body model consisting of nine parts [Sigal et al. 2003].

Template matching approaches have also been applied to deformable 2D objects. The method of Felzenszwalb [2003], finds a globally-optimal embedding of a morphable 2D template, represented as a set of deformable triangles, in an image. The algorithm solves for the optimal value of the correspondence variables via dynamic programming. However, the model uses a deformation parameterization which is particular to 2D; it also relies on a strong prerequisite that the template triangulation belongs to a constrained set of triangulations, preventing its practicality for 3D meshes. The triangulation constraint is relaxed in the work of Coughlan and Ferreira [2002]. They register a morphable 2D template to image data by defining a probabilistic graphical model, which is optimized by loopy belief propagation. However, both methods above do not extend easily to the case of 3D range data.

3 The Correlated Correspondence Algorithm

The input to the algorithm is a set of two meshes (surfaces tessellated into polygons). The *model mesh* $X = (V^X, E^X)$ is a complete model of the object, in a particular pose. $V^X = (x_1, \dots, x_N)$ denotes the mesh points, while E^X is the set of *links* between adjacent points on the mesh surface. The *data mesh* $Z = (V^Z, E^Z)$ is either a complete model or a partial view of the object in a different configuration. Each data mesh point z_k is associated with a *correspondence variable* c_k , specifying the corresponding model mesh point. The task of registration is one of estimating the set of all correspondences C and a non-rigid transformation Θ which aligns the corresponding points.

3.1 Probabilistic Model

We formulate the registration problem as one of finding an embedding of the data mesh Z into the model mesh X , which is encoded as an assignment to all correspondence variables $C = (c_1, \dots, c_K)$. The main idea behind our approach is to preserve the consistency of the embedding by explicitly correlating the assignments to the correspondence variables. We define a joint distribution over the correspondence variables c_1, \dots, c_K , represented as a Markov network. For each pair of adjacent data mesh points z_k, z_l , we want to define a probabilistic potential $\psi(c_k, c_l)$ that constrains this pair of correspondences to reasonable and consistent. This gives rise to a joint probability distribution of the form $p(C) = \frac{1}{Z} \prod_k \psi(c_k) \prod_{k,l} \psi(c_k, c_l)$ which contains only single and pairwise potentials. Performing probabilistic inference to find the most likely joint assignment to the entire set of correspondence variables C should yield a good and consistent registration.

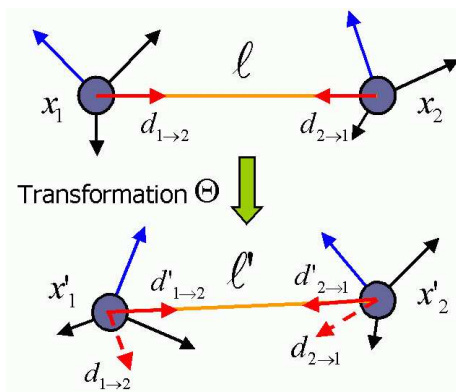


Figure 3: Illustration of the link deformation process. The nonrigid transformation Θ moves the locations and rotates the local coordinate systems of the link endpoints.

3.1.1 Deformation Potentials

We want our model to encode a preference for embeddings of mesh Z into mesh X , which minimize the amount of deformation Θ induced by the embedding. In order to quantify the amount of deformation Θ , applied to the model, we will follow the ideas of Hähnel *et al.* [Hähnel et al. 2003] and treat the links in the set E^X as springs, which resist stretching and twisting at their endpoints. Stretching is easily quantified by looking at changes in the link length induced by the transformation Θ . Link twisting, however, is ill-specified by looking only at the Cartesian coordinates of the points alone. Following [Hähnel et al. 2003], we attach an imaginary *local coordinate system* to each point on the model. This local coordinate system allows us to quantify the “twist” of a point x_j relative to a neighbor x_i . A non-rigid transformation Θ defines, for each point x_i , a translation of its coordinates and a rotation of its local coordinate system.

To evaluate the deformation penalty, we parameterize each link in the model in terms of its length and its direction relative to its endpoints (see Fig. 3). Specifically, we define $l_{i,j}$ to be the distance between x_i and x_j ; $d_{i \rightarrow j}$ is a unit vector denoting the direction of the point x_j in the coordinate system of x_i (and vice versa). We use $e_{i,j}$ to denote the set of edge parameters $(l_{i,j}, d_{i \rightarrow j}, d_{j \rightarrow i})$. It is now straightforward to specify the penalty for model deformations. Let Θ be a transformation, and let $\tilde{e}_{i,j}$ denote the triple of parameters associated with the link between x_i and x_j after applying Θ . Our model penalizes twisting and stretching, using a separate zero-mean Gaussian noise model for each:

$$P(\tilde{e}_{i,j} | e_{i,j}) = P(\tilde{l}_{i,j} | l_{i,j}) P(\tilde{d}_{i \rightarrow j} | d_{i \rightarrow j}) P(\tilde{d}_{j \rightarrow i} | d_{j \rightarrow i}) \quad (1)$$

In the absence of prior information, we assume that all links are equally likely to deform.

In order to quantify the deformation induced by an embedding C , we need to include a potential $\psi_d(c_k, c_l)$ for each link $e_{k,l}^Z \in E^Z$. Every probability $\psi_d(c_k = i, c_l = j)$ corresponds to the deformation penalty incurred by deforming model link $e_{i,j}$ to generate link $e_{k,l}^Z$ and is defined in Eq. (1). We do not restrict ourselves to the set of links in E^X , since the original mesh tessellation is sparse and local. Any two points in X are allowed to implicitly define a link.

Unfortunately, we cannot directly estimate the quantity $P(e_{k,l}^Z | e_{i,j})$, since the link parameters $e_{k,l}^Z$ depend on knowing the nonrigid transformation, which is not given as part of the input. (Indeed, estimating it is part of the goal of the algorithm.) The key issue is estimating the relative rotation of the link endpoints that is induced

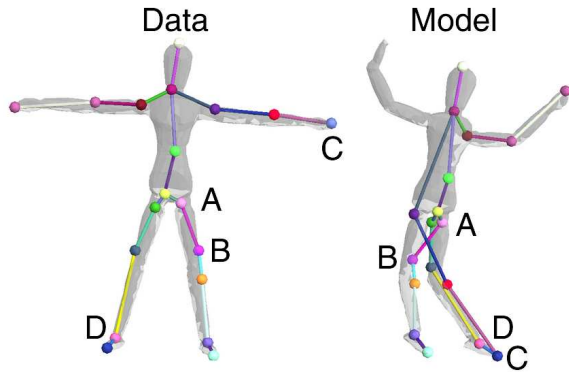


Figure 4: The CC algorithm which uses only deformation potentials can violate mesh geometry. Near regions can map to far ones (segment AB) and far regions can map to near ones (points C,D).

by the (unknown) transformation. In effect, this rotation is an additional latent variable, which must also be part of the probabilistic model. To remain within the realm of discrete Markov networks, allowing the application of standard probabilistic inference algorithms, we discretize the space of the possible rotations, and fold it into the domains of the correspondence variables. For each possible value of the correspondence variable $c_k = i$ we select a small set of candidate rotations, consistent with local geometry. We do this by aligning local patches around the points x_i and z_k using rigid ICP. Specifically, we align the normals at x_i and z_k , and then run ICP on these local regions from a number of different starting points (we have found that two diametrically opposite points suffice). We extend the domain of each correspondence variables c_k , where each value encodes a matching point *and* a particular rotation from the precomputed set for that point. Now the edge parameters $e_{k,l}^z$ are fully determined and so is the probabilistic potential.

3.1.2 Geodesic Distances

Our proposed approach raises the question as to what constitutes the best constraint between neighboring correspondence variables. The literature on scan registration — for rigid and non-rigid models alike — relies on the preserving Euclidean distance. While Euclidean distance is meaningful for rigid objects, it is very sensitive to deformations, especially those induced by moving parts. For example, in Fig. 4, we see that the two legs in one configuration of our puppet are fairly close together, allowing the algorithm to map two adjacent points in the data mesh to the two separate legs, with minimal deformation penalty. In the complementary situation, especially when object symmetries are present, two distant yet similar points in one scan might get mapped to the same region in the other. For example, in the same figure, we see that points in both an arm and a leg in the data mesh get mapped to a single leg in the model mesh.

We therefore want to enforce constraints preserving distance along the mesh surface (geodesic distance). The insight that geodesic distance is the right way of parameterizing mesh surfaces has already been extensively used in graphics, one example is the system of Krishnamurthy and Levoy [2002], which allows the geometric detail manipulation in dense polygon meshes.

Our probabilistic framework easily incorporate such constraints as correlations between pairs of correspondence variables. We encode a *nearness preservation constraint* which prevents adjacent points in mesh Z to be mapped to distant points in X in the geodesic distance sense. For *adjacent* points z_k, z_l in the data mesh, we define

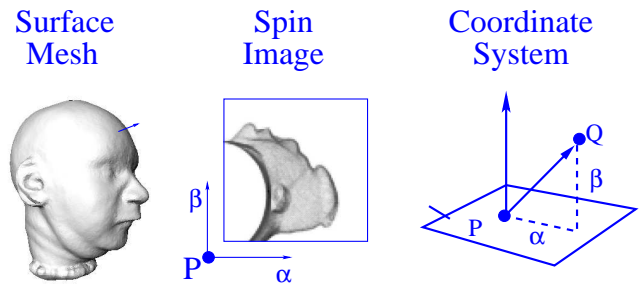


Figure 5: Spin images are two-dimensional histograms computed at an oriented point P on the surface mesh of an object.

the following potential:

$$\psi_n(c_k = i, c_l = j) = \begin{cases} 0 & \text{dist}_{\text{Geodesic}}(x_i, x_j) > \alpha\rho \\ 1 & \text{otherwise} \end{cases} \quad (2)$$

where ρ is the data mesh resolution and α is some constant, chosen to be 3.5.

The *farness preservation* potentials encode the complementary constraint. For every pair of points z_k, z_l whose geodesic distance is more than 5ρ on the data mesh, we have a potential:

$$\psi_f(c_k = i, c_l = j) = \begin{cases} 0 & \text{dist}_{\text{Geodesic}}(x_i, x_j) < \beta\rho \\ 1 & \text{otherwise} \end{cases} \quad (3)$$

where β is also a constant, chosen to be 2 in our implementation. The intuition behind this constraint is fairly clear: if z_k, z_l are far apart on the data mesh, then their corresponding points must be far apart on the model mesh.

3.1.3 Local Surface Signatures

Finally, we encode a set of potentials that correspond to the preservation of local surface properties between the model mesh and data mesh. The use of local surface signatures is important, because it helps to guide the optimization in the exponential space of assignments. We use spin images [Johnson 1997] compressed with principal component analysis to produce a low-dimensional *signature* s_x of the local surface geometry around a point x . When data and model points correspond, we expect their local signatures to be similar. We introduce a potential whose values $\psi_s(c_k) = i$ enforce a zero-mean Gaussian penalty for discrepancies between s_{x_i} and s_{z_k} .

3.2 Optimization

In the previous section, we defined a Markov network, which encodes a joint probability distribution over the correspondence variables as a product of single and pairwise potentials. Our goal is to find a joint assignment to these variables that maximizes this probability. This problem is one of standard probabilistic inference over the Markov network. However, the Markov network is quite large, and contains a large number of loops, so that exact inference is computationally infeasible. We therefore apply an approximate inference method known as *loopy belief propagation (LBP)* (see, for example, [Yedidia et al. 2003]), which has been shown to work well in a wide variety of applications. LBP is a message passing algorithm over the variables in the Markov network. Roughly speaking, it maintains for each variable a probability distribution over its possible values. In each iteration, each variable sends its distribution to its neighbors — those variables to which it is directly connected via a probabilistic potential — and uses the distributions it receives to update its beliefs. Running LBP until convergence results in a set of

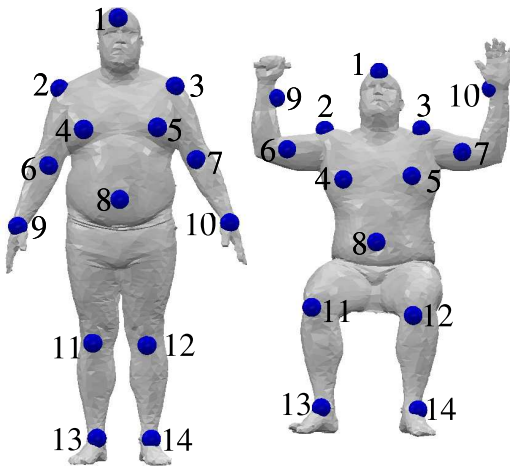


Figure 6: Registration of two poses of the same human. Meshes taken from the CAESAR dataset.

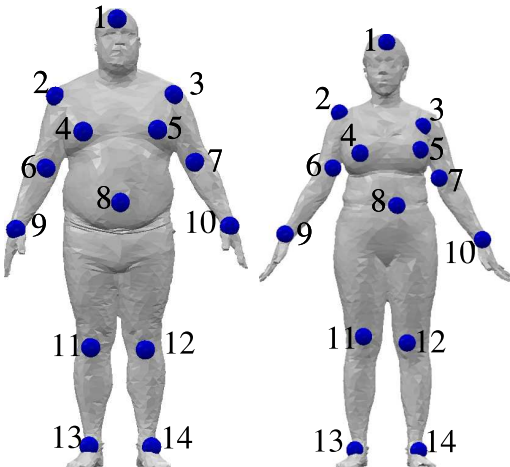


Figure 7: Registration between two different humans in the same pose. Meshes taken from the CAESAR dataset.

probabilistic assignments to the different correspondence variables, which are locally consistent. We then simply extract the most likely assignment for each variable to obtain a correspondence.

One remaining complication arises from the form of our fairness preservation constraints. In general, most pairs of points in the mesh are not close, so that the total number of such potentials grows as $O(M^2)$, where M is the number of points in the data mesh. However, rather than introducing all these potentials into the Markov net from the start, we introduce them as needed. First, we run LBP without any fairness preservation potentials. If the solution violates a set of fairness preservation constraints, we add it and rerun BP. In practice, this approach adds a very small number of such constraints.

We note that the LBP algorithm is an approximate inference algorithm, and may encounter some difficulties. In certain cases, LBP may not converge, and when it does, there are no theoretical guarantees on the quality of the results. In particular, although the algorithm does find correspondences that are locally consistent, it may not produce the optimal alignment. We discuss this issue further in the next section.

4 Experimental Results

In this section, we show some results for the Correlated Correspondence algorithm. We first show that it successfully solves the surface registration problem, even for challenging data sets. We then show that the high-quality correspondences obtained by the algorithm enable us to provide completely unsupervised solutions to several different challenging graphics tasks.

4.1 Basic Registration

We applied our registration algorithm to three different datasets, containing meshes of a human arm, wooden puppet and the CAESAR dataset of whole human bodies [Allen et al. 2003], all acquired by a 3D range scanner. The meshes were not complete surfaces, but several techniques exist for filling the holes (e.g., [Davis et al. 2002; Liepa 2003]).

We ran the Correlated Correspondence algorithm using the same probabilistic model and the same parameters on all data sets. We use a coarse-to-fine strategy, using the result of a coarse sub-sampling of the mesh surface to constrain the correspondences at a finer-grained level. The resulting set of correspondences were used as markers to initialize the non-rigid ICP algorithm of Hähnel *et al.* [Hähnel et al. 2003], which registers the model mesh onto the data mesh. (Note that the CC algorithm works the opposite way, by computing an embedding of the data mesh into the model mesh).

The Correlated Correspondence algorithm successfully aligned all pairs of meshes in the human arm data set. In the puppet data set the algorithm correctly registered four out of six data meshes to the model mesh. In the two remaining cases, the algorithm produced a registration where the torso was rotated, so that the front was mapped to the back. This problem arises from ambiguities induced by the symmetries of the puppet, whose front and back are almost identical. Importantly, however, our probabilistic model assigns a higher score to the correct solution, so that the incorrect registration is a consequence of local minima in the LBP algorithm.

This fact allows us to address this issue in an unsupervised way simply by running the algorithm several times, with different initialization. The initialization conditions were obtained by automatically partitioning the puppet data mesh into parts. Our algorithm looks for extremal points in the data mesh, and then extends the regions around the extremal points to be of a predefined size, which is set to be a fraction of the total object size. Each part thus obtained is then aligned to several different places in the model mesh by using the Correlated Correspondence algorithm. We computed the probability for each part and each of its candidate alignments. We then selected the six non-overlapping part assignments whose total probability was highest. These alignments were used to initialize our algorithm by restricting the set of possible correspondences for the mesh points in the different parts, as dictated by the part-level alignment. We ran the algorithm for each of these six initializations, and selected the one which gave the highest score.

We ran this algorithm to register one puppet mesh to the remaining six meshes in the dataset, obtaining the correct registration in all cases. In particular, as shown in Fig. 2, we successfully deal with the case on which the straightforward nonrigid ICP algorithm failed. Note, however, that the results of the algorithm do contain a small artefact in the puppet’s right shoulder. This artefact is a consequence of the large deformation in the right arm configuration between the two meshes. In this case, the correct registration of the arm cannot be determined from the data, and the algorithm makes an arbitrary decision, leading to the observed effect. We also applied the same algorithm to the CAESAR dataset and produced very good registration for challenging cases exhibiting both articulated motion and deformation (Fig. 6), or exhibiting deformation and a (small) change in object scale (Fig. 7).



Figure 8: Several frames from a motion animation generated by interpolating two scans of an arm (far left and far right), which were automatically registered using the Correlated Correspondence algorithm. The animation was produced by linear interpolation in link transformation space.

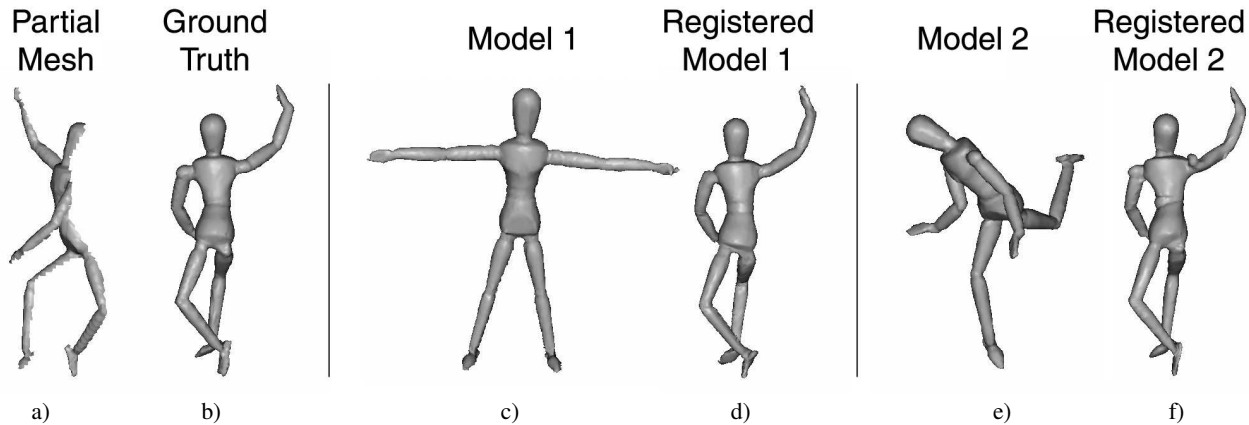


Figure 9: Partial mesh completion. (a) Partial view used as a data mesh. (b) Complete model from which (a) was taken, which serves as the ground truth. The complete model is oriented to display the hidden part of the surface. Registration of the model mesh (c) to (a) produces a completion, displayed in (d), which closely approximates the ground truth. Registration of model mesh (e) to (a) also produces a reasonable reconstruction, shown in (f). The reconstruction in (f) contains an artefact in the right shoulder, resulting from an attempt to preserve the original geometry.

An unoptimized version of the Correlated Correspondence algorithm runs for 1.5 minutes on an Intel Xeon 2.4GHz processor to register a pair of arm meshes. This process includes all the pre-processing steps, including the mesh subsampling phase and the spin-image computation. The algorithm applied to the puppet data, which also involves the computation of the different part embeddings and the execution of the Correlated Correspondence algorithm for the different initialization points, takes a total of 10 minutes per puppet pair.

Overall, the algorithm performs robustly, producing a close-to-optimal registrations even for pairs of meshes that involve large deformations. It deals successfully both with transformations resulting from articulation, where entire parts undergo large motion transformations, and with non-rigid surface transformations. The registration is accomplished in an unsupervised way, without any prior knowledge about object shape, dynamics, or alignment.

4.2 Partial View Completion

The Correlated Correspondence algorithm allows us to register a data mesh containing only a scan of part of an object to a known complete surface model of the object, which serves as a template. We can then transform the template mesh to the partial scan, a process which leaves undisturbed the links that are not involved in the partial mesh. The result is a mesh that matches the data on the observed points, while completing the unknown portion of the surface using the template.

We take a partial mesh, which is missing the entire back part of the puppet in a particular pose. The resulting partial model is displayed in Fig. 9(a); for comparison, the correct complete model in this configuration (which was not available to the algorithm), is shown in Fig. 9(b). We register the partial mesh to models of the object in two different poses (Fig. 9(c) and (e)), and compare

the completions we obtain (Fig. 9(d) and (f)), to the ground truth represented in Fig. 9(b). The results demonstrate a largely correct reconstruction of the complete surface geometry from the partial scan and the deformed template.

The experiment also demonstrates the limitations of this approach. The completion method that we described leaves unchanged links that do not appear in the data mesh. In cases where the template is significantly different from the data configuration on parts that are not visible in the data, this assumption can lead to incorrect completions. For example, the configuration of the left arm in Fig. 9(e) changes significantly, leading to an artefact in the right shoulder of Fig. 9(f). The reason for the artefact is that the shoulder links in the model prefer the original orientation, and no data is available to additionally constrain them.

4.3 Interpolation between Two Meshes

The task of interpolation between different object poses has been extensively studied in graphics and animation. For example, Allen *et al.* [2002] build a model for articulated upper torso deformations from range scan data. They obtain multiple scans of the arms and torso in different positions, and use prior knowledge in terms of a body skeletal structure and markers to build a model of the deformations. In general, there are well-known solutions to the interpolation problem in cases where the object's skeleton is known (e.g., [Chadwick *et al.* 1989; Wang and Phillips 2002]).

As we now show, our Correlated Correspondence algorithm can provide an alternative method for interpolation, which applies directly to meshes. It is therefore applicable even in cases where an object's articulation structure is unknown and in cases where the object is not articulated. Our approach uses the Correlated Correspondence algorithm to register two meshes, which recovers the non-rigid transformation Θ deforming the model mesh. The trans-

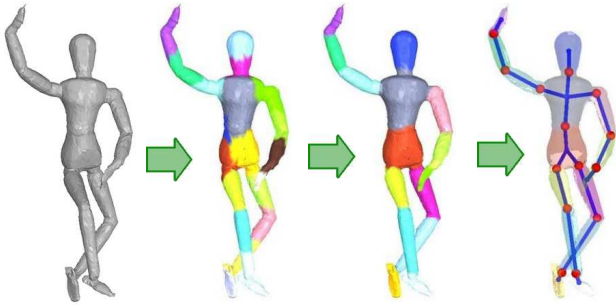


Figure 10: Illustration of the part-finding process: (a) A template mesh is registered to all other meshes by the Correlated Correspondence algorithm. (b) The mesh is randomly divided into small patches of approximately equal areas, different parts are color-coded. (c) Results in (b) are used to initialize an iterative algorithm which estimates the rigid parts and their transformations. (d) The joints linking the rigid parts are estimated.

formation Θ can be expressed in terms of local edge geometry, by using the local transformations of the mesh links, as opposed to movement of mesh points in Cartesian space. We can now interpolate linearly between the two meshes, where the interpolation is done *in the space of link transformations*.

Specifically, for each link $e_{i,j}$ in the model mesh, the Correlated Correspondence algorithm recovers the coordinate system rotations of the link endpoints, and the new link parameters $\tilde{e}_{i,j}$. Any intermediate mesh between the two can be obtained by linearly interpolating the local edge parameters. In particular, we interpolate the rotations of the link endpoints in Euler angle space, and we interpolate the directions $d_{i \rightarrow j}, d_{j \rightarrow i}$ and the lengths $l_{i,j}$. This form of interpolation tries to preserve both the link lengths and their local geometry, to the extent possible. Thus, links whose configuration in both meshes is unchanged will be unchanged throughout the interpolation.

The resulting linear interpolation, executed independently for each link, may not result in a consistent mesh. We therefore solve for a consistent mesh, which is closest (in squared distance) to the linearly interpolated model. Solving for a consistent mesh is equivalent to applying non-rigid ICP (see [Hähnel et al. 2003]) with a deformation prior specified by the linear interpolation defined above. The interpolation process tends to result in natural shapes, generating correct-looking animation sequences, as shown in Fig. 1 and Fig. 8.

4.4 Recovering Articulated Models

Articulated object models have a number of applications in animation and motion capture, and there has been work on recovering them automatically from 3D data [Cheung et al. 2003] and from feature tracking in video [Song et al. 2003].

We show that our unsupervised registration capability can greatly assist articulated model recovery from meshes corresponding to different configurations of an object. First, we register one mesh to all the remaining meshes of the object using the Correlated Correspondence algorithm. Subsequently, we perform *Expectation Maximization* by iterating between finding a decomposition of the object into rigid parts, and finding the location of the parts in the object instances. Finally, we use the recovered rigid parts and their transformations to automatically estimate the joints. The steps of the algorithm are visualized in Fig. 10. The part-finding algorithm [Anguelov et al. 2004] is a separate contribution from the registration work described in this report, and you can refer to



Figure 11: Four different poses from the puppet dataset display the 15 rigid parts and the articulated skeleton, both of which are recovered automatically.

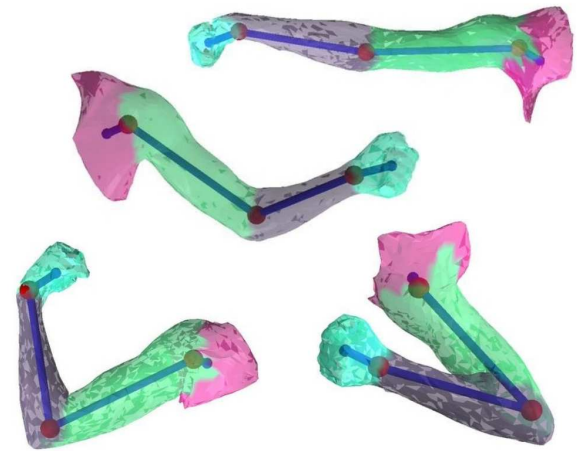


Figure 12: Four different poses from the arm dataset display four (approximately) rigid parts and the articulated skeleton, both of which are recovered automatically.

<http://robotics.stanford.edu/~drago/Papers/uai-parts.pdf> for additional details.

The Correlated Correspondence algorithm successfully registers all seven poses of the puppet, giving us information about the dynamics of the different points in the mesh. As a result of applying the algorithm for clustering the object surface into rigid parts, we automatically recover all 15 rigid parts of the puppet, as well as the joints between them. Several poses of the puppet, along with the recovered skeleton structure and position, are displayed in Fig. 11. To our knowledge, this is the first implementation that estimates such a complex skeleton from real world data with very few poses, in a completely unsupervised way.

In comparison, the algorithm of Cheung *et al.* [2003] is applied to sequences where only 2 parts move at a time, recovering an articulated human model with 9 parts by composing the results of the various sequences. Their approach is essentially a generalization of the ICP algorithm to multiple rigid parts. As we demonstrated in Fig. 2a), ICP is known to be prone to local minima. We hypothesize that the additional degrees of freedom provided by the possible part decompositions make the problem more severe, preventing them from dealing with multiple parts. Solving the registration problem with Correlated Correspondences constrains us enough to allow the

use of global inference technique for rigid part clustering, which is robust even in the presence of multiple parts.

Our algorithm for recovering articulation works well even when the object parts are not purely rigid, as is the case with the human arm. Even in this case, however, we get the intuitive articulated decomposition by using the meshes from our arm data set (see Fig. 12).

5 Conclusion

In this report, we describe an algorithm for unsupervised registration of non-rigid 3D surfaces in significantly different configurations. Our results show that the algorithm can deal with articulated objects subject to large joint movements, as well as with non-rigid surface deformations. The algorithm is not provided with markers or other cues regarding correspondence, and makes no assumptions about object shape, dynamics, or alignment.

We show that a solution to the registration problem can be used as a component in several applications. In our first application, we show that it allows a smooth interpolation between two different meshes of an object, in a way that tends to preserve the local geometry. We note that this interpolation process does not rely on the knowledge of even the existence of an underlying articulated skeleton. In our second application, we show that a partial data mesh (e.g., one arising from a single-view scan) can be registered to a complete model mesh, allowing the missing part of the data mesh to be completed using the model. Finally, we can use a set of scans of an articulate object in different configurations to determine its partition into parts.

The most important limitation of our approach is the fact that it makes the assumption of (approximate) preservation of geodesic distance. Although this assumption is a good heuristic in many cases, it is not always warranted. In some cases, the mesh topology may change, for example, when an arm touches the body. In these cases, our nearness preservation constraints are violated. In other cases, occlusions may eliminate paths in our data mesh, making nearby points appear geodesically distant, and violated our fairness preservation constraints. We can try to extend our approach to handle these cases by trying to detect when they arise, and eliminating the associated constraints. However, even this solution is likely to fail in some cases. A second limitation of our approach is that it assumes that the data mesh is a subset of the model mesh. If the data mesh contains clutter, our algorithm will attempt to embed the clutter into the model. We feel that the general nonrigid registration problem becomes underspecified when significant clutter and occlusion are present simultaneously. In this case, additional assumptions about the surfaces will be needed.

Despite the fact that our algorithm performs quite well, there are limitations to what can be accurately inferred about the object from just two scans. Given more scans of the same object, we can try to learn the deformation penalty associated with different links, and bootstrap the algorithm. Such an extension would be a step toward the goal of learning models of object shape and dynamics from raw data.

References

- ALLEN, B., CURLESS, B., AND POPOVIC, Z. 2002. Articulated body deformation from range scan data. In *Proc. SIGGRAPH*.
- ALLEN, B., CURLESS, B., AND POPOVIC, Z. 2003. The space of human body shapes: reconstruction and parameterization from range scans. In *Proc. SIGGRAPH*.
- ANGUELOV, D., KOLLER, D., PANG, H., SRINIVASAN, P., AND THRUN, S. 2004. Recovering articulated object models from 3d range data. In *In Proc. UAI*.
- B. CURLESS, AND M. LEVOY. 1996. A volumetric method of building complex models from range images. In *Proc. SIGGRAPH*.
- BESL, P., AND MCKAY, N. 1992. A method for registration of 3d shapes. *Transactions on Pattern Analysis and Machine Intelligence* 14, 2, 239–256.
- BLANZ, V., AND VETTER, T. 1999. A morphable model for the synthesis of 3d faces. In *Proc. SIGGRAPH*.
- BLANZ, V., AND VETTER, T. 2002. Face recognition based on 3d shape estimation from single images. In *CG Technical Report No.2, University of Freiburg*.
- CHADWICK, J. E., HAUMANN, D. R., AND PARENT, R. E. 1989. Layered construction for deformable animated characters. In *Proceedings of the 16th annual conference on Computer graphics and interactive techniques*, ACM Press, 243–252.
- CHEN, Y., AND MEDIONI, G. 1991. Object modeling by registration of multiple range images. In *Proc. IEEE Conf. on Robotics and Automation*.
- CHEUNG, K., BAKER, S., AND KANADE, T. 2003. Shape-from-silhouette of articulated objects and its use for human body kinematics estimation and motion capture. In *Proc. IEEE CVPR*.
- CHUI, H., AND RANGARAJAN, A. 2000. A new point matching algorithm for non-rigid registration. In *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR)*.
- COUGHLAN, J., AND FERREIRA, S. 2002. Finding deformable shapes using loopy belief propagation. In *In Proc. ECCV*, vol. 3, 453–468.
- DAVIS, J., MARSCHNER, S., GARR, M., AND LEVOY, M. 2002. Filling holes in complex surfaces using volumetric diffusion. In *Symposium on 3D Data Processing, Visualization, and Transmission*.
- DAVIS, J., RAMAMOORTHY, R., AND RUSINKIEWICZ, S. 2003. Spacetime stereo: A unifying framework for depth from triangulation. In *Proc. CVPR*.
- FELZENSZWALB, P. 2003. Representation and detection of shapes in images. In *PhD Thesis*, Massachusetts Institute of Technology.
- FISCHLER, M. A., AND BOLLES, R. C. 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. In *Comm. of the ACM*, vol. 24, 381–395.
- HÄHNEL, D., THRUN, S., AND BURGARD, W. 2003. An extension of the ICP algorithm for modeling nonrigid objects with mobile robots. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI)*, IJCAI.
- HUTTENLOCHER, D., AND FELZENSZWALB, P. 2003. Efficient matching of pictorial structures. In *Proc. CVPR*.
- JOHNSON, A. 1997. *Spin-Images: A Representation for 3-D Surface Matching*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA.

- KRISHNAMURTHY, V., AND LEVOY, M. 2002. Fitting smooth surfaces to dense polygon meshes. In *Symposium on 3D Data Processing, Visualization, and Transmission*.
- KRISHNAMURTHY, V. 2000. Fitting smooth surfaces to dense polygon meshes. In *PhD Thesis*, Massachusetts Institute of Technology.
- LEVENTON, M. 2000. Statistic models in medical image analysis. In *PhD Thesis*, Massachusetts Institute of Technology.
- LIEPA, P. 2003. Filling holes in meshes. In *Proc. of the Eurographics/ACM SIGGRAPH symposium on Geometry processing*, Eurographics Association, 200–205.
- LIN, M. H. 1999. Tracking articulated objects in real-time range image sequences. In *ICCV (1)*, 648–653.
- PEARL, J. 1988. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann.
- RUIZ-CORREA, S., SHAPIRO, L., AND MEILA, M. 2001. A new signature-based method for efficient 3-d object recognition. In *Proc. IEEE CVPR*, vol. 1.
- RUSINKIEWICZ, S., AND LEVOY, M. 2001. Efficient variants of the ICP algorithm. In *Proc. Third International Conference on 3D Digital Imaging and Modeling (3DIM)*, IEEE Computer Society, Quebec City, Canada.
- SHELTON, C. 2000. Morphable surface models. In *International Journal of Computer Vision*.
- SIGAL, L., ISARD, M., SIGELMAN, B. H., AND BLACK, M. J. 2003. Attractive people: Assembling loose-limbed models using non-parametric belief propagation. In *In Proc. NIPS*.
- SONG, Y., GONCALVES, L., AND PERONA, P. 2003. Unsupervised learning of human motion. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- TAYCHER, L., FISHER, J. W., III, AND DARRELL, T. Recovering articulated model topology from observed motion.
- WANG, X. C., AND PHILLIPS, C. 2002. Multi-weight enveloping: least-squares approximation techniques for skin animation. In *Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*, ACM Press, 129–138.
- YEDIDIA, J., FREEMAN, W., AND WEISS, Y. 2003. Understanding belief propagation and its generalizations. In *Exploring Artificial Intelligence in the New Millennium*, Science & Technology Books.
- YU, S., GROSS, R., AND SHI, J. 2002. Concurrent object recognition and segmentation with graph partitioning. In *Proc. NIPS*.
- ZHANG, D., AND HEBERT, M. 1997. Multi-scale classification of 3-d objects. In *Proc. CVPR*.