Shape matching through particle dynamics warping

Gady Agam and Suneel Suresh
Department of Computer Science, Illinois Institute of Technology, Chicago, IL 60616
{agam,suresun}@iit.edu

Abstract
Shape matching is fundamental to numerous computer vision algorithms and may be used for similarity determination and registration. Establishing correspondence and measuring similarity between shapes is of great importance. Shape matching often involves simultaneous estimation of both a correspondence and an alignment transformation. Such an estimate is particularly difficult when the alignment transformation is non-linear and so contains a large number of degrees of freedom. We describe a novel approach for shape matching that is based on shape contexts and uses particle dynamics warping to maximize the similarity of shapes while satisfying structural constraints. The approach is based on an iterative solution of a system of first order ordinary differential equations. The main advantage of the proposed approach is its ability to incorporate shape constraints into the matching process. Furthermore, the proposed approach does not require a solution of an optimal assignment problem which is sensitive to outliers, and does not require thin-plate spline warping which is computationally expensive. To illustrate the applicability of our approach we address the problem of offline signature recognition. An example of the use of the proposed approach for shape matching is the application of online signature recognition. Online signature recognition differs from offline signature recognition in that it provides a simple 1D parametrization of the signature curves. This approach of warping the parameter values to establish correspondence between two curves is known to be useful for measuring curve similarity. However, it cannot be extended directly to offline signature recognition in which a simple 1D parametrization of the signature curves does not exist and in which the curve deformation cannot be described by a simple deformation of the parameter values. An approach to map the offline problem into an online problem is presented in [3] where instead of using the actual complex signature curves, simple exterior curves approximating the top and bottom parts of the signature are used. These curves are easily characterized by parametric curves with 1D parametrization and so are suitable for DTW. This, however, comes at a cost of simplifying the actual signature data.

1. Introduction
Shape matching is fundamental to numerous computer vision algorithms and commonly used for establishing correspondence and measuring similarity between shapes. Generally speaking, shape matching techniques can be classified [13, 6] into feature-based and intensity-based techniques. Intensity-based techniques often employ some global transformation to obtain a translation/rotationScale invariant representation [13, 6]. By nature, such techniques are limited to linear deformations of shapes and cannot cope with non-rigid deformations. To perform shape matching in an image, preprocessing operations may be used to convert the image into a more compact representation. Examples include the boundaries of segmented regions [9, 14], and the medial axis [12]. Simple representations such as the silhouettes of shapes often lead to 1D parameterizations of curves. Such parameterizations have the advantage of being suitable for matching through dynamic programming. However, the use of silhouettes may unnecessarily eliminate some shape information.

1-4244-1180-7/07/$25.00 ©2007 IEEE
points along curves. A particularly successful shape matching technique is the shape context approach of Belongie et al. [2]. This approach is based on the characterization of shape points using a 2D histogram of pixel distributions as a function of distance and angle. Given two shapes, the matching of points between them is treated as a linear assignment problem in which the points in one shape are matched to the points in the second shape based on the shape context distance between points. As the linear assignment will most likely produce errors, an iterative solution is employed in which the shape context and the assignment between points is refined. This refinement is done by using the initial correspondence solution to warp one of the shapes using thin-plate splines, recomputing the shape context descriptors of points in the warped shape, and repeating the assignment and warping steps. An extension of this approach that is capable of better preserving locality is described by Zheng and Doermann [15]. An approach for point matching with some robustness to noise has been proposed by Chui and Rangarajan [4]. Shape contexts are rich localized descriptors that are useful in various applications. It is interesting to note that the rough equivalent of 2D shape contexts in 3D is commonly referred to as spin images [7]. Spin images again are rich localized descriptors that are based on 2D histograms and have proved very useful in various applications.

In this paper, we describe a novel approach for shape matching that is based on shape contexts and uses particle dynamics warping to maximize the similarity of shapes while satisfying structural constraints. The proposed particle dynamics warping approach obtains an optimal solution to the constrained maximization problem using an iterative solution of a system of first order ordinary differential equations. The main advantages of the proposed approach are its ability to incorporate shape constraints into the matching process, and the fact that it does not require a solution of an optimal assignment problem which is sensitive to outliers, and that it does not require multiple thin-plate spline warping steps which are computationally expensive. To illustrate the applicability of our approach we address the problem of offline signature recognition which in contrast to online signature recognition does not provide for a simple parametrization of the signature curves. The proposed approach is evaluated by measuring the precision and recall rates of documents based on signature similarity. To facilitate a realistic evaluation, the signature data we use was collected from real world documents spanning a period of several decades.

The following sections describe our approach in greater detail. Section 2 describes the proposed particle dynamics warping approach. Section 3 presents quantitative evaluation results obtained by the proposed approach. Section 4 concludes the paper.

2. The proposed approach

Given two shapes the similarity between them is measured in the proposed approach by warping one of them onto the other, and measuring an energy term. The energy term is composed of both intrinsic and extrinsic components. The intrinsic component measures the amount of deformation that was applied to the warped signature to make it similar to the target signature, whereas the extrinsic component measures the degree of similarity between the warped and the target signature. This similarity measure is based on both the shape context similarity as well as spatial similarity. The warping of one of the signatures onto the other is obtained using a particle dynamics warping (PDW) approach [1]. This PDW approach employs an iterative solution to a system of ordinary differential equations, which results in a minimum energy term. The obtained minimum energy term then represents the degree of similarity between the shapes. Note that in contrast to the original shape context algorithm, the proposed approach does not require optimal assignment, and includes structural constraints. This results in improved ability for handling similarity between shapes involving noise and/or large deformations.

2.1. Shape model construction

The first step in the proposed approach for shape matching is to obtain one or more sets of connected points on the shapes that need to be compared. The connection between points need not follow a simple 1D parametrization. Obtaining the sets of connected points depends on the application domain. However, a fundamental requirement of this step is that the points be obtained in different instances of the shape in a consistent way. To facilitate this requirement, multiple points are selected in each shape, without attempting to select a small subset of points of interest with good localization. Thus the selected points are not constrained to be corner points. The connection between points may be either real with support from shape elements, or imaginary based on vicinity constraints. Again, the connectivity between points may be obtained in different ways depending on the application domain so long as the connectivity is obtained in a consistent way in different instances of the shape. A reliable measure for consistent connectivity relations in binary shapes is connectivity in the image domain (e.g. 4- or 8-connectivity).

The application domain that is used to illustrate the proposed approach for shape matching in this paper is offline signature recognition. Signatures form complex shapes with high levels of non-rigid deformations. Note, however, that the proposed approach is general in nature and may be applied to general shape matching problems. Specific to the application domain of offline signature recognition, we begin the shape model construction my segmenting the sig-
nature image, thinning it, and obtaining a polyline approximation of strokes. The polylines are obtained so that each line segment in the polyline does not deviate by more than a preset threshold \( \tau_1 \) from the original stroke, and so that each vertex in the polyline is not separated from its neighboring vertices by more than a preset threshold \( \tau_2 \). The resulting shape model is then composed of 2D vertices and links (polygon edges) between them. Note that the connectivity between vertices produced in this approach is not constrained to be simple and that it is likely that multiple disjoint segments be produced.

2.2. Shape normalization

Given two shape models that need to be compared, an alignment step is performed to take into account rotation, scale, and translation differences. While, the proposed approach is likely to converge to a correct solution even without performing this step, the convergence is faster and more reliable when performing this step. The alignment of the shape models is achieved by estimating a global affine transformation between the shapes and applying it to rectify them. The affine mapping estimated for each shape is estimated so that the center of mass of its vertices is translated to the origin and so that the principal directions of the covariance matrix of the vertices be aligned with the horizontal and vertical coordinate axes. Let \( \{v_i\}_{i=1}^{n} \equiv \{(x_i, y_i)\}_{i=1}^{n} \) be the vertices of the shape represented in homogeneous coordinates. Their center of mass is given by \( \bar{v} = \frac{1}{n} \sum_{i=1}^{n} v_i \) whereas their \( 3 \times 3 \) covariance matrix is given by: \( C = \frac{1}{n} \sum_{i=1}^{n} (v_i - \bar{v})(v_i - \bar{v})^T \). Let \( \lambda_1 > \lambda_2 > \lambda_3 \) be the eigenvalues of \( C \). Let \( e_1, e_2, \) and \( e_3 \) be the corresponding eigenvectors which are assumed to be normalized. The desired affine mapping \( M \) is given by a translation transformation which is followed by a rotation and scale transformations. Combining the transformation matrices in homogeneous coordinates we get:

\[
M = \begin{pmatrix}
\frac{\lambda_1^{1/2}}{\bar{v}} & -\lambda_1^{-1/2} \bar{v} \\
0 & 1
\end{pmatrix}
\]

where the first and second rows of the \( 2 \times 2 \) matrix \( R \) are set to the first two elements in the eigenvectors \( e_1 \) and \( e_2 \), respectively, and \( 0^T \) is a zero vector. To avoid arbitrary rotations in the affine mapping when the principal directions are not uniquely defined, the rotation matrix \( R \) is set to identity when \( \lambda_2/\lambda_1 > \tau \), where \( \tau \in [0, 1] \) is a threshold value (set to 0.7 in our study). In order to prevent reflection in the affine mapping due to orientation ambiguity, it is necessary to ensure that the diagonal elements of \( R \) are positive. Thus, denoting the elements of \( R \) by \( r_{ij} \), such ambiguity may be prevented by setting \( r_{ij} = \text{sgn}(r_{ii}) \cdot r_{ij} \), and using this modified \( R \) in Equation (1).

An example of the results obtained after normalizing the signatures is presented in Figure 1. The top part shows ten signatures overlaid on top of each other by aligning their bounding boxes without normalization. The middle figure shows the signatures overlaid after normalization. The bottom figure shows the signatures after the PDW alignment as described later. As can be observed, the normalization results in better alignment of the signatures. In all these figures, the signatures were thinned to a single pixel width to reduce the clutter in them.

2.3. Particle dynamics warping

Given a shape model, obtained as described before, a particle model is created. The particle model is composed of identical mass particles which are placed at vertices of the shape model, and damped springs which are placed between particles that are connected by edges in the shape model. The proposed approach for measuring the similarity between shapes is based on two types of competing forces that are used to warp one of the shapes onto the other. The first type are intrinsic forces which define structural constraints and are formed by damped springs in the particle model. The second type are extrinsic forces which are formed by attraction between particles in the two particle models. The attraction is defined using a shape context similarity measure, where a high shape context similarity results in high attraction. For efficiency reasons, the shape context similar-
ity attraction force is applied only between particles with a distance smaller than $R$ between each other.

Given the the particle systems that represent the shapes, a system of ordinary differential equations (ODE) which describe the particle system dynamics [5] is formed. This system is then solved iteratively by propagating the system in time toward a solution of minimum energy. We refer to this approach as particle dynamics warping (PDW). The proposed approach is generic and can easily support additional forces. Examples of such possible forces include distance-based attraction between the particles of different shapes and repulsion between the particles of the same shape. Note, however, that distance based forces require spatial sorting and so substantially reduce time efficiency.

This system of equations is formed in a phase space in which the position and velocity of all the particles in the system are concatenated to form a $4n$-dimensional vector:

$$X = (x_1^1, x_1^2, v_1^1, v_1^2, \ldots, x_n^1, x_n^2, v_n^1, v_n^2)^T \quad (2)$$

where $x_i \equiv (x_i^1, x_i^2)$ is the position of the $i$-th particle, and $v_i \equiv (v_i^1, v_i^2)$ is the velocity of the $i$-th particle. Assuming Newtonian particles we have $\dot{x}_i = v_i$ and $\dot{v}_i = \frac{1}{m} f_i$, and so the ODE system is given by:

$$\dot{X} = (v_1^1, v_1^2, \frac{1}{m} f_1^1, \frac{1}{m} f_1^2, \ldots, v_n^1, v_n^2, \frac{1}{m} f_n^1, \frac{1}{m} f_n^2)^T \quad (3)$$

where $m$ is the mass of each particle and $f_i \equiv (f_i^1, f_i^2)$ is the total force which is exerted on the $i$-th particle. We denote Equation (3) by $\dot{X} = F(X, t)$. The system (3) can be solved iteratively using the fourth order Runge-Kutta method [11].

PDW should not be confused with active contours [8]. While in both PDW and active contours a curve is deformed, the means and purpose of the deformation are substantially different. Active contours start with a simple curve in an image and attempt to deform it so as to fit the boundary of an object of interest. The deformation in active contours relies on intrinsic energy terms measuring the regularity and smoothness of the produced curve and an extrinsic energy term derived from intensity gradients. The curve deformation in active contours is based on a solution of a partial differential equation. In contrast to active contours, in PDW the deformed curve is not simple, the propose of the deformation is the alignment of the curve with another curve, image gradients are not used, the intrinsic energy measures the amount of deformation introduced into the original curve, and the curve is deformed based on the solution of a system of ordinary differential equations.

### 2.4. The force equations

The forces involved in the proposed PDW system are designed to guide the warping of one shape onto the other while preserving its original form as much as possible, thus minimizing an energy term that is composed of both intrinsic and extrinsic components. The forces affecting the intrinsic energy term are composed of spring-damper pairs connected particles of the warped shape. The forces affecting the extrinsic energy term are attraction forces between the particles of the warped shape and the particles of the target shape, whereas these forces are computed based on shape context similarity.

Given two particles $i$ and $j$ that are connected with a damped spring, the force exerted on particle $i$ by the spring is given by:

$$f_i^{(s)} = - \left( k_s(|\Delta| - r) + k_d \frac{\Delta}{|\Delta|} \right) \frac{\Delta}{|\Delta|} \quad (4)$$

where $\Delta = x_i - x_j$, $\dot{\Delta} = v_i - v_j$, $r$ is the rest length of the spring (set to the initial value of $|x_i - x_j|$), and $k_s$ and $k_d$ are the spring and damping constants respectively. The force exerted on particle $j$ by the spring is given by: $f_j^{(s)} = -f_i^{(s)}$. The damper is added to the spring to prevent oscillations and increase the stability of the system. The potential energy due to the spring force is given by:

$$E_i^{(s)} = \frac{1}{2} k_s (|\Delta| - r)^2 \quad (5)$$

The spring potential energy is referred to as spring energy in the remainder of this paper.

The attraction forces between particles in the two shapes are computed based on the complement of the shape context similarity measure of Belongie et al. [2]. Let $h_i(k)$ be the shape context of the $i$-th particle in the deformable shape, and $h_j(k)$ be the shape context of the $j$-th particle in the target shape. The shape context similarity between these particles is given by:

$$C_{ij} = 1 - \sum_{k=1}^{n} \frac{|h_i(k) - h_j(k)|}{h_i(k) + h_j(k)} \quad (6)$$

The force excerpted on particle $i$ by particle $j$ is given by:

$$f_i^{(sc)} = -C_{ij} \Delta \quad (7)$$

The potential energy due to this attraction is given by:

$$E_i^{(sc)} = \frac{1}{2} (C_{ij})^2 \quad (8)$$

This energy term depends only on the shape context of the particles. It does not depend on the distance between them. The shape context attraction potential energy is referred to as shape context energy in the remainder of this paper. Theoretically, a shape context attraction force exists between each particle in the warped shape and each particle in the
target shape. However, to improve the computational efficiency of the approach we only consider shape context attraction between particles with a distance of R or smaller. Finally, to enhance numerical stability a viscous drag force is added to each particle. The viscous drag force applied to the i-th particle is given by:

\[ f_i^{vd} = -k_{vd}v_i \]  

(9)

2.5. Measuring the similarity between shapes

Using the PDW process as described above, the similarity between two shapes may be obtained by warping one of the shapes onto the other and measuring the intrinsic and extrinsic energy in the system after the warping. The intrinsic energy is the spring energy \( E_i^{(s)} \) as defined in Equation (5), whereas the extrinsic energy is the shape context energy as defined in Equation (8). An additional extrinsic energy term may be computed based on the distance between each particle in the warped shape and its closest particle in the target shape. We refer to this term as displacement energy \( E_i^{(d)} \). Using these energy terms, the similarity between two shapes \( C_a \) and \( C_b \) is given by:

\[ S(C_a, C_b) = \omega_s \sum_{i=1}^{n-1} E_i^{(s)} + \omega_{sc} \sum_{i=1}^{n} \hat{E}_i^{(sc)} + \omega_d \sum_{i=1}^{n} \hat{E}_i^{(d)} \]  

(10)

where \( n \) is the number of vertices in the warped shape model, \( \hat{E}_i^{(s)} \), \( \hat{E}_i^{(sc)} \), and \( \hat{E}_i^{(d)} \), are normalized energy terms in the range [0, 1], and the weights \( \omega^{(s)} \), \( \omega^{(sc)} \), \( \omega^{(d)} \) are set to be convex combination coefficients: \( \omega^{(s)} + \omega^{(sc)} + \omega^{(d)} = 1 \) and \( \omega^{(s)}, \omega^{(sc)}, \omega^{(d)} \geq 0 \). These weights are set empirically as described later. Note that there is a trade-off between the values of the spring and displacement energy terms. Obtaining lower displacement energy is possible when ignoring the structural constraints imposed by the springs of the system. Similarly, obtaining lower spring energy is possible when ignoring the displacement between the shapes. The simultaneous minimization of these terms through the proposed PDW scheme results in a minimization that satisfies both the displacement and structural constraints.

When several instances of the same shape are available for training, the similarity between the training set and a test shape is determined by computing the combined energy term using the test shape and each of the shapes in the training set, and taking the minimum value of these energy terms.

3. Results and Discussion

The application domain that was used to evaluate the proposed approach for shape similarity is that of offline signature recognition. We have created a dataset of signatures which were collected by searching a large collection of scanned documents (containing several millions of documents) for documents containing signatures. The obtained documents were then labeled for identity and location of the signatures in them. A subset of 76 subjects was extracted such that each subject has more than five sample signatures in the test collection. The overall number of signatures in the test collection is 627. The obtained test collection is highly realistic in that the signatures were obtained over a period of time spanning a few decades, the signatures were signed in a normal business context, and the scanning of the documents was conducted by different scanners at different times and so produced varying levels of image quality and noise.

Qualitative results of warping multiple signatures onto a single signature using the proposed approach is presented in figure 1. To quantitatively evaluate the performance of the proposed approach we selected the application of content-based retrieval. Given a single signature we measure its similarity to the remaining signatures in the dataset. Using a decision threshold, we then measure the true positive (TP), false negative (FN), and false positive (FP) rates. These rates are then converted to precision (TP/(TP+FP)) and recall (TP/(TP+FN)) rates. Given multiple signatures of the same person, we measure similarity based on each of the available instances, and select the minimum distance result. This normally, results in improved precision and recall rates. The obtained precision-recall graphs are presented in Figure 2. The data used in this evaluation contained 76 subjects with roughly 10 signatures per subject. The query signature contained 1-5 signatures. Note that the query signatures were not used in the evaluation. The weights used in this evaluation were taken as \( \omega^{(d)} = 0.9 \) for the displacement energy, \( \omega^{(sc)} = 0.1 \) for the shape context energy, and \( \omega^{(s)} = 0.0 \) for the spring energy. The selection of these weights is explained later in the paper. As can be observed, when increasing the number of signatures in the query set the performance improves. The performance gain between a four signature query and a five signature query is small. With five instances of a signature in the query set we obtained, for example, 90% precision with 80% recall or alternatively 80% precision with 85% recall. Note that the use of real world data in the experiment imposes four difficult conditions: only a small number of signatures (five) is available for the query; the signatures are scanned by different scanners at different and normally law resolutions; the signatures were obtained over a long period of time thus substantially increasing their variability; and various levels of noise such as segments of underlying printed text are present in the data.

The combined similarity measure for distance between curves in Equation (10) uses a weighted sum of displacement, shape context, and spring energy terms. These weights can be estimated based on a retrieval quality cri-
Figure 2. Precision recall curves for a dataset of 76 subjects. Each of the subjects has roughly 10 signatures of which one to five were used in the query. As can be observed, as the number of signatures in the query increases, the performance increases.

Figure 3. The area under the 76-subjects precision recall curve as a function of the spring and displacement energy weights. The weight of the shape context energy term is the one-complement of the spring and displacement energy weights.

Figure 4. Sensitivity to shape deformation evaluation. A deformed shape is aligned with a non-deformed one using different techniques. As can be observed, the proposed PDW approach produced lower error compared with the LAP and LAP-TPS approaches.

4. Conclusion

We propose a novel approach for shape matching through a particle dynamics warping scheme which uses shape contexts as a shape similarity measure. In this
scheme, a particle system is formed by inserting particles at various shape points, and damped springs between connected points. The springs form intrinsic constraints that preserve the original shape. Attraction forces between selected particles in the warped and target shape are set with a magnitude which is proportional to the shape context similarity between them. A system of first order ordinary differential equations is formed and solved iteratively. The obtained solution minimizes simultaneously the spring energy and the shape context energy. A weighted sum of these energy terms as well as the obtained displacement energy is used to measure the similarity between shapes. The evaluation scheme employs real world signatures signed over a long period of time and scanned by different scanners. This is in contrast to common evaluation scheme in which the signatures are collected at the same time and are scanned in high resolution using the same scanner. Precision-recall graphs based on queries of five signatures are presented for a set of 76 subjects. The area under the precision-recall curve is used to select the weights of the energy terms in the similarity measure. The noise sensitivity of the proposed approach is evaluated and compared to that of known techniques.

Acknowledgments

This work was supported in part by a challenge grant from ARDA.

References


