SHAPE MATCHING BASED ON GRAPH ALIGNMENT USING HIDDEN MARKOV MODELS

Xiaoning Qian[1] *and Byung-Jun Yoon*[2]

[1]University of South Florida, Department of Computer Science & Engineering, Tampa, FL 33620
[2]Texas A&M University, Department of Electrical & Computer Engineering, College Station, TX 77843

ABSTRACT

We present a novel framework based on hidden Markov models (HMMs) for matching feature point sets, which capture the shapes of object contours of interest. Point matching algorithms provide effective tools for shape analysis, an important problem in computer vision and image processing applications. Typically, it is computationally expensive to find the optimal correspondence between feature points in different sets, hence existing algorithms often resort to various heuristics that find suboptimal solutions. Unlike most of the previous algorithms, the proposed HMM-based framework allows us to find the optimal correspondence using an efficient dynamic programming algorithm, where the computational complexity of the resulting shape matching algorithm grows only linearly with the size of the respective point sets. We demonstrate the promising potential of the proposed algorithm based on several benchmark data sets.

Index Terms— Shape matching, hidden Markov model (HMM), graph alignment.

1. INTRODUCTION

Shape analysis and image registration are two key problems in computer vision and image analysis [1]. They can provide useful information and insights in various disciplines, including biology and medicine [2]. These two problems are often formulated as a "point matching" problem, since feature points are representative of the original image or shape and also amenable to analysis. Finding the correct dense correspondence between such feature points is the crucial first step for image analysis. In this paper, we focus on the problem of shape matching, especially, establishing the optimal correspondence between object contours, represented by sets of sampled contour points.

Shape matching can be categorized into rigid and nonrigid shape matching. The former problem is relatively easy and it has been well studied [3]. Currently, active research is mainly focused on nonrigid shape matching. In recent years, a number of successful algorithms have been developed, where examples include iterated closest point (ICP) algorithm [4], thin-plate spline-robust point matching (TPS-RPM) algorithm [5], shape context algorithm [6], free-form deformation (FFD) algorithm [7], and landmark-sliding algorithm [8]. As noticed in [9, 10, 11, 12], many of these algorithms ignore the local neighborhood structure around each point, although it is important to preserve the local structure so that the objects of interest can deform only within certain physical constraints. More research efforts [10, 11, 12] have been made to incorporate such structural restrictions, including higher-order constraints to find physically reasonable correspondence. However, most of the available algorithms are computationally expensive and the most recent algorithm given in [10] has cubic complexity with respect to the size of the point sets. Due to the high computational cost of finding the optimal solution, many algorithms resort to various heuristics that lead to suboptimal solutions.

In this paper, we address the shape matching problem in the framework of hidden Markov models (HMMs). A similar framework has been recently applied to biological network alignment by the authors [13, 14]. Based on HMMs, we propose a general probabilistic framework for scoring shape similarity and present an efficient search algorithm that can find the optimal correspondence between shape contour points with the highest score. By appropriately parameterizing the HMMs, local neighborhood structure at each contour point can be naturally incorporated into this probabilistic framework. In addition, the HMM-based approach provides a natural mechanism for rejecting outliers, and it prevents cross-overs (neighboring points with swapped correspondence), which may occur in many algorithms [4, 5, 6, 12]. Furthermore, our HMM-based shape matching algorithm has linear complexity with respect to the number of the contour points in each set, which makes matching large point sets computationally feasible.

2. SHAPE MATCHING USING HMM

2.1. Shape matching as graph alignment

We focus on pairwise shape matching in this paper. However, as we show later, our algorithm can be extended for matching multiple shape contours in a straightforward manner. Assume we have two shape contours: $\mathcal{U} = \{u_1, u_2, \ldots, u_{N_1}\}$ with N_1 points; and $\mathcal{V} = \{v_1, v_2, \ldots, v_{N_2}\}$ with N_2 points, which are in fact the sample points along the respective shape con-

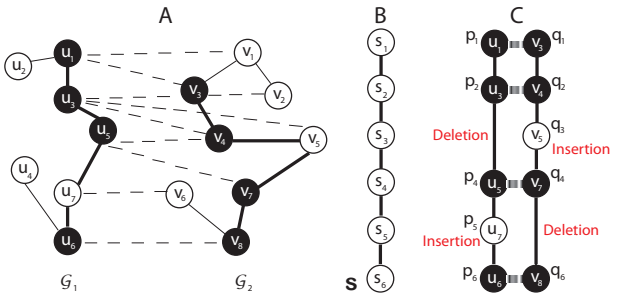

Fig. 1. (A) Example of two graphs $\mathcal{G}_1$ and $\mathcal{G}_2$ representing two shape contours (each circle corresponds to a point). (B) A virtual path $\mathbf{s}$ that corresponds to the alignment of best matching paths. (C) The best-matching pair of paths $\mathbf{p}$ (in $\mathcal{G}_1$) and $\mathbf{q}$ (in $\mathcal{G}_2$).

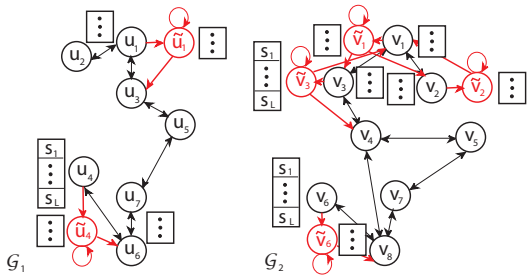

Fig. 2. HMMs that allow insertions and deletions in the best-matching pair of paths. The dots next to the hidden states represent all possible symbols corresponding to virtual nodes in $\mathbf{s}$ that can be emitted. For simplicity, the detailed illustration of the hidden states in the HMMs are shown only for the nodes u_1 and u_4 in $\mathcal{G}_1$; v_1, v_2, v_3, and v_6 in $\mathcal{G}_2$.

tours. Based on the prior information about these points, we can define the edges between neighboring contour points and represent them as graphs $\mathcal{G}_1$ and $\mathcal{G}_2$. In 2D shape matching, these graphs have a linear structure. However, it should be noted that the framework also admits more general structures for 3D shape matching.

To establish the correspondence between contour points, we formulate a graph alignment problem, where we aim to find the most similar pair of linear paths $(\mathbf{p}, \mathbf{q})$, where $\mathbf{p}$ belongs to the graph $\mathcal{G}_1$ and $\mathbf{q}$ belongs to $\mathcal{G}_2$. Figure 1A shows an example with two graphs representing two shape contours. Our task is to find the optimal pair of linear paths in the respective graphs, where the constituent points have the best matching correspondences (see Fig. 1C). As the contour points from different sets can be sampled in different ways, the best matching paths may have flexible number of insertions and deletions, as shown in Fig. 1C, which provides an automatic outlier-rejection mechanism. Note that "insertions" and "deletions" are relative terms. They can also be viewed as surrogates for introducing "many-to-one" or "one-to-many" correspondences between points.

2.2. Hidden Markov models

To find the best matching pair of paths, we adopt the hidden Markov model (HMM) formalism. Let us first focus on the construction of a HMM for $\mathcal{G}_1$. Each node $u_i \in \mathcal{U}$ (a contour point) in $\mathcal{G}_1$ corresponds to a hidden state in the HMM. For convenience, we represent this hidden state using the same notation u_i. For neighboring nodes (points) u_i and u_j, we add an edge between states u_i and u_j in the HMM. At this stage, the resulting HMM has an identical structure as the graph $\mathcal{G}_1$. To accommodate gaps (insertions and deletions), we add an accompanying state $\tilde{u}_i$ for every state u_i. Next, we add an outgoing edge from each state to the corresponding accompanying state and we also add outgoing edges from the accompanying state to all the neighboring states of the original state u_i. We also allow *consecutive* insertions or deletions by enabling self-transitions at $\tilde{u}_i$. The HMM for $\mathcal{G}_2$ can be constructed in a similar way. Figure 2 illustrates the HMMs that correspond to the graphs shown in Fig. 1A. In order to find the best matching pair of paths, we define the concept of a "virtual" path $\mathbf{s} = s_1 s_2 \dots s_L$ that contains L nodes, as shown in Fig. 1B. A node s_t in the virtual path can be viewed as a symbol that is emitted by a pair of hidden states $(p_t, q_t) = (u_j, v_\ell)$ in the respective HMMs. From this point of view, the two HMMs can be regarded as generative models that *jointly* emit the virtual path $\mathbf{s}$, and the underlying state sequence for $\mathbf{s}$ will be a pair of state sequences $\mathbf{p}$ and $\mathbf{q}$ in the respective HMMs. Therefore, the concept of a virtual path can naturally couple a path in $\mathcal{G}_1$ with another in $\mathcal{G}_2$, providing a convenient framework for identifying the correspondence between sample points in the given contours. Note that, the matching paths $\mathbf{p}$ and $\mathbf{q}$ contain L states each, but they may contain one or more accompanying states which represent gaps. The proposed framework does not impose any restriction on the number of gaps and their locations in the matching paths.

2.3. Finding matched paths in HMMs

Based on the described HMM structure, we now define a matching score $S(\mathbf{p}, \mathbf{q})$ for a pair of paths in these HMMs. In order to obtain meaningful results, $S(\mathbf{p}, \mathbf{q})$ should sensibly integrate the similarity $h(p_t, q_t)$ between matched states (corresponding to contour points) $p_t = u_j$ and $q_t = v_\ell$ ($1 \leq t \leq L$), the connectivity $w_1(p_{t-1}, p_t)$ between $p_{t-1} = u_i$ and $p_t = u_j$, the connectivity $w_2(q_{t-1}, q_t)$ between $q_{t-1} = v_k$ and $q_t = v_\ell$ ($2 \leq t \leq L$), and the penalty for possible gaps. The HMM-based representation in Sec. 2.2 provides an effective probabilistic framework that can be used for this purpose.

We first define two mappings $\mathbf{f}_1 : w_1(u_i, u_j) \mapsto t_1(u_j|u_i)$ and $\mathbf{f}_2 : w_2(v_k, v_\ell) \mapsto t_2(v_\ell|v_k)$, which convert $w_1(u_i, u_j)$ and $w_2(v_k, v_\ell)$ to the following transition probabilities

$$P(p_t = u_j | p_{t-1} = u_i) = t_1(u_j|u_i) = \mathbf{f}_1(w_1(u_i, u_j)) \quad (1)$$

$$P(q_t = v_\ell | q_{t-1} = v_k) = t_2(v_\ell|v_k) = \mathbf{f}_2(w_2(v_k, v_\ell)) \quad (2)$$

between the corresponding hidden states in the HMMs. The mappings $\mathbf{f}_1$ and $\mathbf{f}_2$ should follow the constraints in [14] so

that $t_1(u_j|u_i)$ and $t_2(v_\ell|v_k)$ are legitimate probabilities. In addition to this, we define another mapping $\mathbf{g} : h(u_j, v_\ell) \mapsto e(u_j, v_\ell)$ that converts $h(u_j, v_\ell)$ to the following "pairing" probability

$$P(p_t = u_j, q_t = v_\ell) = e(u_j, v_\ell) = \mathbf{g}(h(u_j, v_\ell)), \tag{3}$$

which is the emission probability of a virtual symbol s_t at a pair of underlying hidden states $(p_t, q_t) = (u_j, v_\ell)$ in the respective HMMs. The mapping $\mathbf{g}$ should be defined such that the resulting probability $e(u_j, v_\ell)$ is legitimate [14].

Now the problem of finding the best-matching pair of paths is transformed into the problem of finding the optimal pair of state sequences in two HMMs that jointly maximize the observation probability of $\mathbf{s}$. We can find the optimal solution in polynomial time by using a dynamic programming algorithm, which is conceptually identical to the Viterbi algorithm. We first define $\gamma(t, j, \ell)$ as the log-probability of the most probable pair of state sequences for a subsequence $\widehat{\mathbf{s}} = s_1 \dots s_t$ of length $t(\leq L)$, where the two underlying states for s_t are $p_t = u_j$ and $q_t = v_\ell$. We have the following recursive equation:

$$\begin{aligned}\gamma(t, j, \ell) = \max_{i,k} \big[\gamma(t-1, i, k) \quad &+ \quad \log t_1(u_j|u_i) + \log t_2(v_\ell|v_k) \\ &+ \quad \log e(u_j, v_\ell)\big].\end{aligned} \tag{4}$$

We repeat the above iterations until $t = L$. The maximum log-probability of the virtual path $\mathbf{s}$ is given by:

$$\log P(\mathbf{p}^*, \mathbf{q}^*) = \max_{\mathbf{p},\mathbf{q}} \Big[\log P(\mathbf{p}, \mathbf{q})\Big] = \max_{j,\ell} \gamma(L, j, \ell), \tag{5}$$

where $\{\mathbf{p}^*, \mathbf{q}^*\} = \arg\max_{\mathbf{p},\mathbf{q}}[\log P(\mathbf{p}, \mathbf{q})]$ is the optimal pair of state sequences that correspond to the best matching paths in the given HMMs. Once we have computed $\log P(\mathbf{p}^*, \mathbf{q}^*)$, it is straightforward to find $\{\mathbf{p}^*, \mathbf{q}^*\}$ by tracing the recursive equations that led to the maximum log-probability $\log P(\mathbf{p}^*, \mathbf{q}^*)$. The computational complexity of the above optimization algorithm is $O(LM_1M_2)$, where L is the length of the aligned paths that we want to find, M_1 is the number of edges in $\mathcal{G}_1$, and M_2 is the number of edges in $\mathcal{G}_2$. $S(\mathbf{p}, \mathbf{q}) = \log P(\mathbf{p}, \mathbf{q})$ can serve as a good similarity score between the matching paths. In principle, we can also use non-stochastic pairing scores $s_{em}(u_j, v_\ell)$ and transition scores $s^1_{tr}(u_j|u_i)$ and $s^2_{tr}(v_\ell|v_k)$ in the recursive equation (4), in place of the log-probabilities $\log e(u_j, v_\ell)$, $\log t_1(u_j|u_i)$, and $\log t_2(v_\ell|v_k)$, respectively. This will yield a non-stochastic score instead of an observation probability.

2.4. Extension to point matching for multiple sets

It is straightforward to extend the described pairwise matching algorithm for matching the points in multiple sets. Without loss of generality, we only consider the extension to the alignment of three sets. Given three graphs $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_3$, we construct the corresponding HMMs based on their structures. We again use the concept of a virtual path, and now we assume that a virtual path $\mathbf{s}$ is jointly emitted by these three HMMs. The emission of a virtual symbol s_t is now governed by a pairing probability $e(u_j, v_\ell, x_n)$ of three hidden states u_j, v_ℓ, and x_n in the respective HMMs. We can derive a similar recursive equation as (5) to find the optimal matching.

2.5. Parameterization of HMMs

The quality of the point matching results obtained from our HMM-based algorithm is determined by the transition and pairing probabilities (scores) of the HMMs. In this section, we present an effective scheme for choosing the parameters based on the local structure of the feature points.

To capture the local structural properties of each contour point, we extract its $2K$-nearest neighbors on the shape contour graph. These nearest neighbors form a contour segment. Let us denote the coordinates of these neighbors as

$$\Sigma_j = \begin{pmatrix} u^x_{j-K} & u^x_{j-K+1} & \cdots & u^x_j & \cdots & u^x_{j+K} \\ u^y_{j-K} & u^y_{j-K+1} & \cdots & u^y_j & \cdots & u^y_{j+K} \end{pmatrix}, \tag{6}$$

where u^x_j and u^y_j are u_j's x and y coordinates, respectively. We subtract $\mathbf{u}_j = (u^x_j, u^y_j)^T$ from each column of Σ_j as the matching should be invariant to translation of shape contours. We further take the singular value decomposition (SVD): $\Sigma_j - \mathbf{u}_j \times \mathbf{1} = R_j \Lambda_j U_j$. Λ_j is a diagonal matrix and will have at most two non-zero diagonal entries, which tells the shape of the distribution of these nearest neighbor points. Assuming that λ^1_j and λ^2_j are the two non-zero diagonal entries in descending order, we use the ratio λ^2_j/λ^1_j as our local shape feature, since shape matching should also be invariant to the size of the given contours. R_j in fact gives the overall direction of these points θ_j:

$$R_j = \begin{pmatrix} \cos\theta_j & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j \end{pmatrix}, \tag{7}$$

We now use these two local structural characteristics to define the pairing and transition scores in the HMMs. First, the pairing score between two hidden states that correspond to contour points is defined as follows:

$$s_{em}(u_i, v_k) = G_h(\frac{\lambda^2_j}{\lambda^1_j}, \frac{\delta^2_k}{\delta^1_k}), \tag{8}$$

where $G_h(\cdot, \cdot)$ is a similarity function whose value is determined by the difference of the ratios of these diagonal entries so that we make sure the local shape are similar around the "paired" states; and δ is used to denote the diagonal entries of the second shape contour. Note that the pairing score can also incorporate the similarity between additional image features, for example, local image features or invariants like SURF (Speeded Up Robust Features) [15] and SIFT (Scale Invariant Feature Transform) features [16]. If either u_i or v_k is an accompanying hidden state in the HMMs, we set a constant Δ_{em} as the gap penalty. We disallow matching two accompanying hidden states by setting the corresponding pairing score to $-\infty$.

We can view the process of emitting the virtual path as a Markovian walk on a product graph of the two HMMs. Therefore, instead of individually specifying the transition scores $s^1_{tr}(u_j|u_i)$ and $s^2_{tr}(v_\ell|v_k)$, we can combine these two scores into one function $s^{ijk\ell}_{tr}$:

$$s^{ijk\ell}_{tr} = \begin{cases} G_w(\theta_i - \theta_j, \phi_k - \phi_\ell) & \text{if } d_{ij} > 0 \text{ and } e_{k\ell} > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

where $G_w(\cdot, \cdot)$ is a similarity function determined by the difference of the direction changes between θ_i and θ_j, and between ϕ_k and ϕ_ℓ, respectively. d_{ij} and $e_{k\ell}$ are edges in $\mathcal{G}_1$

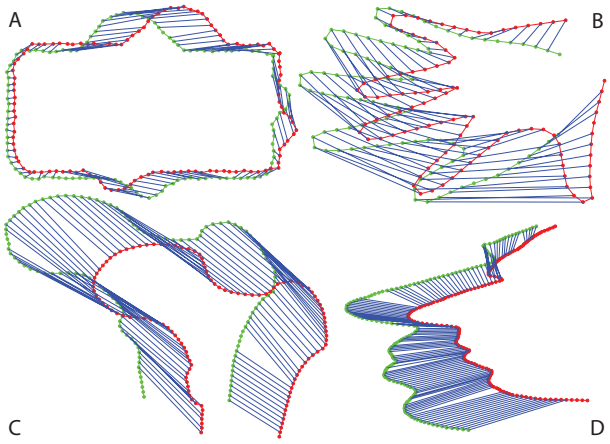

Fig. 3. Shape matching results of the proposed HMM-based algorithm.

and $\mathcal{G}_2$, respectively. By using the relative direction changes, we make shape matching invariant to rotation. If any of the four states is an accompanying hidden state in the HMMs, we set $s_{tr}^{ijk\ell} = \Delta_{tr}$ for gap penalty.

3. EXPERIMENTAL RESULTS

We tested our HMM-based matching algorithm on four benchmark shape data sets, including synthetic bumpboxes (Fig. 3A), hand profiles (Fig. 3B), contours of femurs (Fig. 3C), and silhouette profiles (Fig. 3D), which are obtained from [11, 17, 18]. We used identical settings to get all four results shown in Fig. 3, where $\Delta_{em} = \Delta_{tr} = -6.0$, $K = 5$, and $L = 1.5 \times \min(N_1, N_2)$. It should be noted that although we fix L, we can in fact find all optimal matching paths with length $\leq L$, since all the intermediate scores for shorter paths are stored while running the dynamic programming algorithm. For $G_h(\cdot,\cdot)$ and $G_w(\cdot,\cdot)$, we used Gaussian functions with 0 means and standard deviations σ_h and σ_w, and we set σ_h to be one-tenth of the deviation of the local shape ratios in the given sets and σ_w to $0.5°$. As we can see in Fig. 3, the proposed algorithm achieved good point matching results for all four benchmark sets. Especially, the algorithm was able to establish accurate correspondence also for contour points with high curvatures (in peaks and valleys). Unlike other existing algorithms, the HMM-based approach naturally prevents any cross-over. As mentioned earlier, the contour points that correspond to the accompanying hidden states in the HMMs can either be viewed as outliers or may be used to obtain "many-to-one" correspondence between feature points.

4. DISCUSSIONS

We demonstrated the potential of our HMM-based shape matching algorithm to establish accurate correspondence between feature points in two or more images. This can serve as a useful first step in many image analysis applications. The proposed framework can be extended for shape registration and motion analysis by integrating additional features into the proposed HMM-based graph alignment framework. Our future research will focus on shape matching for more general graph structures and thorough experimental validation of the proposed algorithm based on a quantitative scheme given in [12].

Acknowledgments The authors would like to thank Dr. Yifeng Jiang for sharing the data sets.

5. REFERENCES

[1] D.H. Ballard and C.M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ, 1982.

[2] I.L. Dryden and K.V. Mardia, *Statistical Shape Analysis*, John Wiley and Sons, West Sussex, 1998.

[3] L.G. Brown, "A survey of image registration techniques," *ACM Computing Survey*, vol. 24, no. 4, pp. 325–376, 1992.

[4] P.J. Besl and N.D. McKay, "A method for registration of 3-d shapes," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, February 1992.

[5] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Computer Vision and Image Understanding*, vol. 89, no. 2-3, pp. 114–141, 2003.

[6] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 24, pp. 509–522, April 2002.

[7] X. Huang, N. Paragios, and D.N. Metaxas, "Shape registration in implicit spaces using information theory and free form deformations," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 28, no. 8, pp. 1303–1318, 2006.

[8] S. Wang, T. Kubota, and T. Richardson, "Shape correspondence through landmark sliding," in *CVPR 2004*. IEEE, pp. 1063–1069.

[9] I. Biederman, "Recognition by components: A theory of human image understanding," *Psychological Rev.*, vol. 94, no. 2, pp. 115–147, 1987.

[10] O. Duchenne, F. Francis Bach, I. Kweon, and J. Jean Ponce, "A tensor-based algorithm for high-order graph matching," in *CVPR 2009*. IEEE.

[11] Y. Jiang, J. Xie, D. Sun, and T. Tsui, "Shape registration by simultaneously optimizing representation and transformation," in *MICCAI 2007*. LNCS 4792-II, pp. 809–817.

[12] Y. Zheng and D. Doermann, "Robust point matching for nonrigid shapes by preserving local neighborhood structures," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 28, no. 4, pp. 643, 2006.

[13] X. Qian, S.H. Sze, and B.J. Yoon, "Querying pathways in protein interaction networks based on hidden Markov models," *J Comput Biol*, vol. 16, no. 2, pp. 145–157, 2009.

[14] X. Qian and B.J. Yoon, "Effective identification of conserved pathways in biological networks using hidden Markov models," *PLoS One*, vol. 4, no. 12, pp. 8, 2009.

[15] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[16] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60(4), pp. 91–110, 2004.

[17] H.H. Thodberg, "Minimum description length shape and appearance models," in *IPMI*. LNCS, 2003, pp. 51–62.

[18] R. Davies, C. Twining, T. Cootes, J. Waterton, and C. Taylor, "A minimum description length approach to statistical shape modeling," *IEEE Trans. Med. Imag.*, vol. 21, no. 5, pp. 525–537, 2002.