OBJECT RECOGNITION AND RECOVERY BY SKELETON GRAPH MATCHING

Lei He[1], Chia Y. Han[2], William G. Wee[2]

[1]Information Technology Department, Armstrong Atlantic State University
[2]ECECS Department, University of Cincinnati

ABSTRACT

This paper presents a robust and efficient skeleton-based graph matching method for object recognition and recovery applications. The novel feature is to unify both object recognition and recovery components into an image understanding system architecture, in which a complementary feedback structure can be incorporated to alleviate processing difficulties of each component alone. The idea is firstly to recognize the preliminary extracted object from a set of models using the new skeleton graph matching method, then to apply the a priori shape information of the identified model for accurate object recovery. The output of the system is the recognized and segmented object. The skeleton graph matching method is illustrated by recognizing a set of tool and animal silhouette examples with the presence of geometric transformations (translation, rotation, scaling, reflection), shape deformations and noise. Experiments of object recovery using MR knee images, have shown satisfactory results.

1. INTRODUCTION

Image understanding generally includes two key interrelated components: image segmentation and object recognition. Image segmentation approach such as deformable contour method (DCM) [1-3] yields contours, either exact or approximate, of objects of interest in images for recognition. Given the segmented object shape, object recognition performs shape matching to identify the object from a set of models. The recognition results can be fed back into the image segmentation to enhance the accuracy and robustness of the segmentation results, which is referred to as object recovery. Our work focuses on object recognition component, which is implemented through a novel object skeleton matching approach, i.e., matching the skeleton graph of an input DCM [3] contour with those of models. Skeleton is selected here from a group of well-known shape descriptors (e.g., chain code, B-spline, Fourier and wavelet descriptors) due to its significant features on the desired representational properties, such as invariance to object geometric transformations (translation, rotation and scaling) and reversibility to the original shape. Moreover, skeleton is the only descriptor providing object structural information, such as location of convex-parts, width and length of each part, which is important for the object recognition applications. The "best" matching skeleton pair determines the correct model and constructs the contour feature point (landmark) correspondences. The correspondences of the contours' segments follow automatically. For the DCM contour segments with a large error when compared with the corresponding model segments, a fine-tuning process, which is formulated as a maximization of a posteriori probability [4], is performed for final object recovery.

The objective of this paper is to present a new skeleton-based graph matching approach for object recognition, finally yielding a robust and efficient object recovery, which is different from most of previous related works [5-11]. Using object morphology skeleton, Ruberto [5] derived the attributed skeletal graph with the selected attributes like distance function, curvature variation and size. With the unordered tree representation, it does not necessarily preserve the order of skeleton edges/branches at nodes in the final match. Luo and Hancock [6] proposed a purely structural approach for inexact graph matching between two point sets, which is formulated as a maximum-likelihood estimation. It can only handle small rotation and non-rigid shape deformations. Zhu and Yuille [7] constructed object skeleton model in terms of the principal deformation modes for object recognition. The algorithm is sensitive to noise on primitive segmentation and computationally demanding with multiple parameters to be tuned. Shock graph [8-11] consider object skeleton as a set of singularities (shocks), which can be further represented as a shock tree/graph for shape representation and matching. The matching algorithm is to find the shock graph node correspondences based on both the graph topological and geometrical similarities, which are constructed by the shock category and attributes, such as location, orientation and time of formation. The matching algorithms are usually implemented by assignment algorithm [9], finding subgraph isomorphism [8,10] or edit-distance algorithm [11]. The shock segmentation and matching algorithms are complex and sensitive to noise. Moreover, these algorithms are for object recognition purpose, and not well suited for object shape recovery since correspondences on contour landmarks cannot be derived from the skeleton edge correspondences.

The rest of the paper is organized as follows. The notation for skeleton entities is presented in Section 2. Section 3 gives the algorithm description. Experiments on matching and recovering shapes are provided in Section 4. Section 5 draws the conclusions.

2. SKELETON STRUCTURE NOTATION

The object skeleton consists of the locus of centers of maximal disks (CMD) that can be inscribed within the object and a maximal disk is not completely contained in any other disks in the object. The notation used for the major *skeleton graph* entities is illustrated in Figure 1(a) with a skeleton example of a side view of a quadruped animal. The curve segments of actual skeleton edges are simplified as line segments. The attributes of a skeletal point on the derived skeleton graph are its distance value and if it is a CMD or not. In Figure 1(a), a skeletal point having only one skeletal point in its eight neighbors is defined as an *ending node* (E-node) (e.g. A, B) and a skeletal point having at least three points in its

eight neighbors is a *bifurcation node* (B-node) (e.g. C, D). All others are normal nodes. After the skeleton computation, skeletal points can be linked to form skeleton edges (SE). A SE includes the skeletal points between a B-node (or an E-node) and another B-node. A SE with two B-nodes is a *primary* SE (e.g. CD); otherwise it's a *normal* SE (e.g. AC).

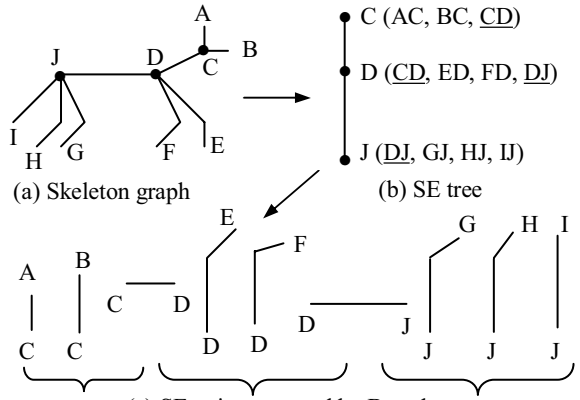


Figure 1. Skeleton graph notation

3. PROPOSED APPROACH

The skeleton-based object recognition and recovery approach can be described in four major steps: (I) skeleton processing, (II) skeleton model construction, (III) skeleton matching and model detection, and (IV) contour segment correction, as shown in Figure 2. Two types of object models: skeleton model and contour segment models are involved in our approach. The skeleton model is constructed for shape matching and object recognition in Step (III), and the contour segment models are for fine-tuning the final object shape recovery result in Step (IV).

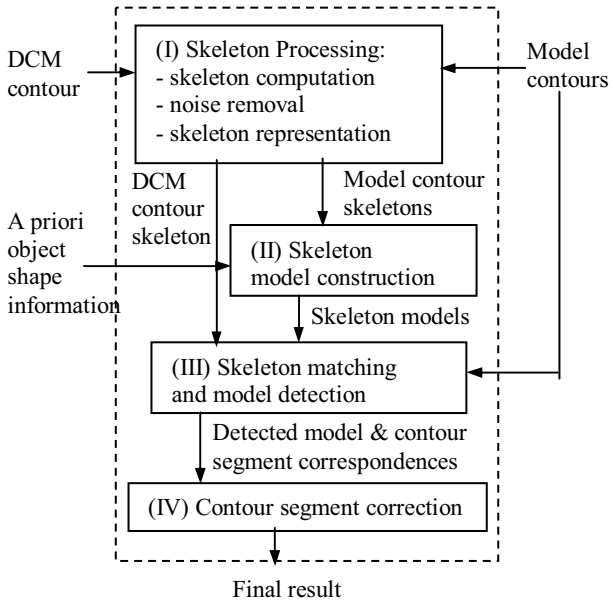


Figure 2. Object recognition and recovery flowchart

(I) Skeleton Processing

The first step consists of skeleton computation [12], removing SEs produced by shape noise, and representing the skeleton. After deriving the skeletons for input and model contours, the skeletons can then be pruned based on the importance of the SEs. For each normal SE, we can determine its importance by comparing the original shape with the shape reconstructed from the skeleton without that SE. A large difference in the shapes indicates that the SE is important; otherwise, it is due to shape noise and can be cut out from the skeleton graph. For effective processing, the skeleton graph is further represented by a tree and a string of SEs, as shown in Figure 1(b) and 1(c), respectively. A B-node tracing algorithm to generate a skeleton tree is applied to record SEs sequentially, starting from the B-node connected with the shortest normal SE. The tree nodes are the B-nodes and the tree edges are the primary SEs as underlined in Figure 1(b). Finally, the skeleton tree can be represented as a SE string (Figure 1(c)) grouped by the B-nodes, and each group consists of a set of SEs connected by the common B-node.

(II) Skeleton model construction

Two major types of information: rigid transformations and nonrigid deformations, are collected to specify the model and used in matching. The first includes the admissible connectivity relationship among the skeleton edges, parameterized by the bifurcation angles between adjacent SEs and the variations of skeleton edges caused by the bifurcation delay/splitting phenomena (BDP/BSP) [7] (e.g. the B-node D will split to two B-nodes as FD shifts left along DJ in Figure 1(b)). The second includes the object shape variations, described as the Gaussian distribution on the distance values of skeletal points.

(III) Skeleton matching and model detection

In our application, a similarity function is proposed to measure the resemblance between the input and model skeleton graphs. Given an input shape skeleton string $D=(d_1, d_2, \ldots, d_R)$ and a model skeleton string $M=(m_1, m_2, \ldots, m_S)$, the matching is to find the SE correspondences in the two strings to maximize the similarity function: $S(D,M) = e^{P}/(E_1E_2E_3)$, where P is the thickness similarity of the input shape D and a model M, weighted by the shape error weight E_1, the thickness error weight E_2, and the length error weight E_3.

Assume the skeleton edge d_i matches with m_j (i=1, 2, …R, j=1, 2, …S) and only T skeleton edge pairs (e.g. T=min(R,S)) are matched, we have $P = \frac{1}{T}\sum_{i,j=1}^{T} p(d_i, m_j)$, with $p(d_i, m_j) = \prod_{k=1}^{N} \frac{1}{\sigma_k\sqrt{2\pi}} e^{-\frac{(rd_{ik}-rm_{jk})^2}{2\sigma_k^2}}$, rd_{ik} and rm_{jk} are the normalized distance values of the kth corresponding points on the matched skeleton edges d_i and m_j, respectively, after normalizing the skeleton edges d_i and m_j to be the same length N. rd_{ik} and rm_{jk} are normalized with respect to the largest distance value on d_i and m_j, respectively. σ_k is the variance of the normalized rm_{jk} to represent the nonrigid deformations. In the experiments, we use $\ln(p(d_i, m_j))$ instead of $p(d_i, m_j)$ to avoid a very small P. The E_1 is to measure the shape difference between D and M. After transforming the input SE d_i to the coordinate system of its corresponding model

SE m_j as d_i' and normalizing it to be the same length N as m_j, E_1 can be computed as: $E_1 = \frac{1}{\mathrm{T}} \sum_{i,j=1}^{\mathrm{T}} e_1(d_i', m_j)$, with

$$e_1(d_i', m_j) = \frac{1}{\mathrm{N}} \sum_{k=1}^{\mathrm{N}} \sqrt{(xd_{ik}' - xm_{jk})^2 + (yd_{ik}' - ym_{jk})^2}$$, (xd_{ik}', yd_{ik}') and (xm_{jk}, ym_{jk}) are the coordinates of the kth corresponding points on d_i' and m_j, respectively. The E_2 is to measure the difference of the average thickness ratio between D and M, and it is formulated as: $E_2 = \frac{1}{\mathrm{T}} \sum_{i=1}^{\mathrm{T}} e_2(d_i, m_j)$, and $e_2(d_i, m_j) = \frac{|\overline{rd_i} / \overline{rd} - \overline{rm_j} / \overline{rm}|}{\overline{rm_j} / \overline{rm}}$, $\overline{rd_i}$ and $\overline{rm_j}$ are the average distance values of d_i and m_j, $\overline{rd}$ and $\overline{rm}$ are the average distance values of all the skeletal points on D and M. Likewise, the length error weight E_3 is to measure the difference of the length ratio between D and M.

The optimization process is divided into two sequential subtasks: structural and statistical matching. In the structural matching, a branch-and-bound string matching algorithm is used to search over all possible matches between the two strings for the same number and connectivity relationship of B-nodes, and the same numbers of primary and normal SEs in each matched B-node group pair. For each match found, several validity checks based on the rules and information collected in Step (II) can be used to take out the invalid matches, thus reduce the searching space for following steps. The statistical matching is to detect the correct model from a set of models passed above validity checking and determine the correct SE correspondences based on the similarity function between the input and model skeletons. The model producing the largest similarity is selected and the SE correspondences in that match are considered correct.

(IV) Contour segment correction

The contour landmarks are determined by the E-nodes of skeleton edges. These landmarks are then used to determine the contour segment correspondences. Contour segments are compared for errors. The error computation is the same as the computation of the shape error weight E_1, in which the skeleton edges are replaced as the contour segments. A fine-tuning process, which is formulated as a maximization of a posteriori probability [4], given the contour segments model and image features, is performed on the segments with large errors for final result.

4. EXPERIMENTS

In this section, two experiments are used to illustrate our algorithm: the first experiment uses a set of animal silhouette shapes from Brown University Stimuli and [7], as well as a set of biological and tool shapes from Rutgers's database and [8] to demonstrate the skeleton graph matching algorithm for object recognition; the second uses biomedical image samples, MRI knee images to show the object recovery process.

For the first shape set in the object recognition experiment, it is to select the most similar shape for the input test shapes, shown in Figure 3 (a)-(h), from a set of animal model shapes shown in Figure 3 (1)-(9). After skeleton matching is used to locate the corresponding SEs between the input and all the possible models, the similarity functions are computed for all valid matches. Table 1 shows the linearly normalized similarity values (0-10), with the highest values being shown in boldface. The quadruped mammals don't match with birds due to their different skeleton structures.

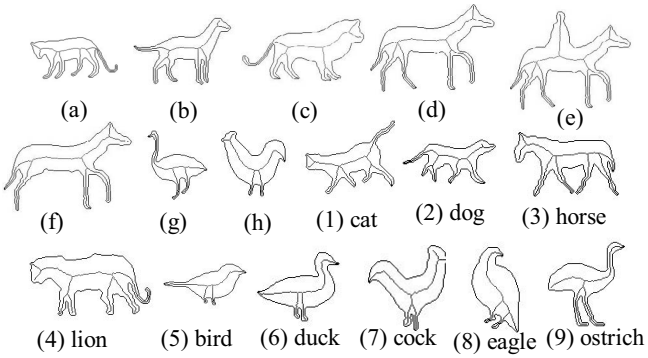


Figure 3. The first shape set for object recognition

Table 1. Similarity values between input and model shapes

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
(1)	**8.90**	2.79	0.18	0.65	0.64	**3.81**	-	-
(2)	0.01	**4.66**	0.01	2.22	1.98	2.05	-	-
(3)	4.45	3.27	0.55	**3.70**	**3.61**	0.97	-	-
(4)	5.12	0.28	**5.12**	0.08	0.06	0.88	-	-
(5)	-	-	-	-	-	-	0.44	2.49
(6)	-	-	-	-	-	-	1.74	0.86
(7)	-	-	-	-	-	-	3.94	**3.53**
(8)	-	-	-	-	-	-	3.17	3.30
(9)	-	-	-	-	-	-	**4.08**	1.70

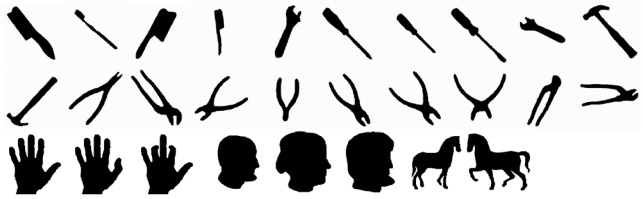


Figure 4. Second set: tool and biological shapes

Table 2. Top five matches for each input shape

Input shape	Top five similar shapes				
	1	2	3	4	5
	3.39	1.76	0.97	0.54	0.05
	2.97				
	5.28	3.88	1.52	0.59	
	7.16				
	6.53	4.51	4.47	3.44	3.14
	4.14	3.94	3.24	3.02	2.79
	9.84	7.53			
	5.87	3.84			
	3.76				

The second shape set is shown in Figure 4, with the twenty-eight objects being classified as nine classes (models). Similar to

[5,8], an object was selected from each class as the input shape. When single SE error is allowed, the top five matches with the linearly normalized similarity values (0-10) are shown in Table 2.

For object recovery, two midline sagittal MRI knee images of size 256 by 256 (Figure 5) are used in the second experiment to extract the femoral condyle (top portion of the knee). The challenge is that there is a blurry edge segment along the middle top boundary, while the left and right portions of the femoral condyle are rather darker than the middle region. This prevents the deformable contour to reach the real boundary on the two sides before it flows out from the top. The two input knee contours (Figure 6(b), (c)) obtained by a DCM [3] are used to match with the knee model (Figure 6(a)) from a radiologist. The skeletons of the knees after noise removal are shown in Figure 7. The contour landmarks are determined from the skeleton E-nodes and the skeleton matching algorithm is applied to construct the correspondences, e.g., the E-nodes of *a*, *b*, and *c* on the model and input skeletons (Figure 7) correspond to landmarks of A, B and C on the model and input contours (Figure 6). Two large error segments in both input contours (BC & CA in Figure 6(b), AB & CA in Figure 6(c)) are corrected and the final recovery results are shown in Figure 8. It can be seen the final results have less shape error than those before correction.

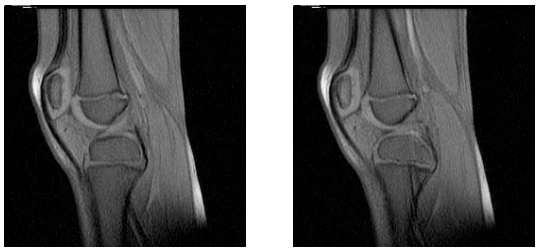
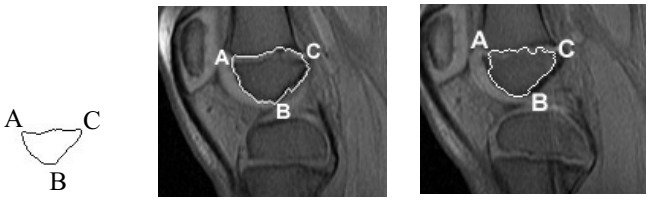
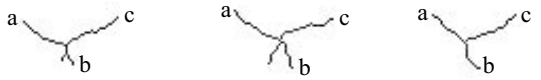


Figure 5. Input MRI knee 1 and 2 images



(a) Model skeleton (b) Knee 1 skeleton (c) Knee 2 skeleton

Figure 6. The knee model and input shapes obtained by a DCM [3]



(a) Knee model skeleton (b) Knee 1 skeleton (c) Knee 2 skeleton

Figure 7. The knee skeletons of shapes in Figure 6 after noise removing

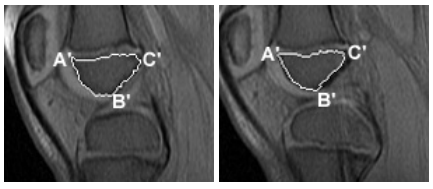


Figure 8. The knee recovery results 1 and 2 after segments correction

5. CONCLUSIONS

In this paper, a robust and efficient skeleton-based shape matching method is presented to solve the object recognition and recovery problems for image understanding. The presented method uses a combination of both structural and statistical approaches that are applied in a sequential manner. The connectivity relationship among skeleton edges and the geometrical features of skeleton edges are used for skeleton graph matching. The object recovery using the skeleton-based shape matching approach is invariant to object geometric transformations. Thus the initial condition requirements and the searching space for object recovery are reduced significantly compared with many other model-based DCMs. The experiments with the animal and tool shape matching and the MRI knee shape recovery demonstrate the capability and potential of this new approach.

6. REFERENCES

[1] M. Kass, A. Witkin and D. Terzopoulos, "Snake: active contour models", *Int. J. Comput. Vision*, vol.1, no.4, pp.321-331, 1988.

[2] R. Malladi, J.A. Sethian, and B.C. Vemuri, "Shape modeling with front propagation: a level set approach", *IEEE PAMI*, vol.17, no.2, pp.158-175, 1995.

[3] X. Wang, L. He, and W.G. Wee, "Deformable contour method: a constrained optimization approach", *Int. J. Comput. Vision*, vol.59, no.1, pp.87-108, 2004.

[4] L.H. Staib and J.S. Duncan, "Boundary finding with parametrically deformable models", *IEEE PAMI*, vol.14, no.11, pp.1061-1075, 1992.

[5] C.D. Ruberto, "Recognition of shapes by attributed skeletal graphs", *Pattern Recognition*, vol.37, no.1, pp.21-31, 2004.

[6] B. Luo and E.R. Hancock, "Structural graph matching using the EM algorithm and singular value decomposition", *IEEE PAMI,* vol.23, no.10, pp.1120-1136, 2001.

[7] S. Zhu and A.L. Yuille, "FORMS: a flexible object recognition and modeling system", *Int. J. Comput. Vision,* vol.20, no.3, pp.187-212, 1996.

[8] K. Siddiqi, A. Shokoufandeh, S.J. Dickinson, and S.W. Zucker, "Shock graphs and shape matching," *Int. J. Comput. Vision*, vol.35, no.1, pp.13-32, 1999.

[9] D. Sharvit, J. Chan, H. Tek, and B.B. Kimia, "Symmetry-based indexing of image databases," *J. Visual Comm. and Image Representation*, vol.9, pp.366-380, 1998.

[10] M. Pelillo, K. Siddiqi, and S. Zucker, "Matching hierarchical structures using association graphs," *IEEE PAMI*, vol.21, no.11, pp.1105-1120, 1999.

[11] T.B. Sebastian, P.N. Klein, and B.B. Kimia, "Recognition of shapes by editing their shock graphs," *IEEE PAMI*, vol.26, no.5, pp.550-571, 2004.

[12] S. Svensson and G. Borgefors, "On reversible skeletonization using anchor-points from distance transforms", *J. of Visual Comm. and Image Representation*, vol.10, pp.379-397, 1999.