Shape Decomposition Scheme by Combining Mathematical Morphology and Convex Partitioning

Duck Hoon Kim
School of Electrical Eng.
Seoul Nat'l Univ.
Seoul, 151-742, Korea
ducks@diehard.snu.ac.kr

Il Dong Yun
School of Electronics and Inform. Eng.
Hankuk Univ. of F.S.
Yongin, 449-791, Korea
yun@computer.org

Sang Uk Lee
School of Electrical Eng.
Seoul Nat'l Univ.
Seoul, 151-742, Korea
sanguk@sting.snu.ac.kr

Abstract

Recently, the interests not only in 2-D image but also in 3-D image have significantly increased in terms of quality, as well as quantity. The shape analysis of these images would be required for further applications: matching, retrieval and so on. In this paper, a novel shape decomposition scheme would be proposed for both 2-D and 3-D images. The proposed scheme is composed of the iterative partitioning based on mathematical morphology using the ball-shaped SE(structuring element). In each step of the iterative partitioning, meaningful parts are selected by the adjacency criterion. Finally the best decomposition result of a given shape is selected by the convexity criterion. Experimental results for 2-D and 3-D images show that decomposition results would be in good agreement with human insight also robust to various perturbations.

1. Introduction

Shape analysis is a fundamental problem in image processing and computer vision. The shape decomposition scheme is the basis of shape analysis, which would be developed for various applications. Above all, it is important and useful for shape matching and retrieval. This importance is justified by the fact that the MPEG-7 group will incorporate shape decomposition into shape description[1]. Shape decomposition implies high-level representation of a given shape, which is described in terms of simpler shape parts and the relationships among the shape parts. Thus the shape decomposition should partition the object into a canonical meaningful parts naturally. Unfortunately, these tasks are difficult, not always well-defined for arbitrary shapes, and sensitive to small perturbations: scaling, rotation, and noise.

In this paper, we propose a novel shape decomposition scheme combining mathematical morphology and convex partitioning. The proposed scheme is composed of the iterative partitioning based on mathematical morphology using the ball-shaped SE. In each step of the iterative partitioning, meaningful parts are selected by the adjacency criterion. Finally the best decomposition result is selected by the convexity criterion among decomposition results from each step of the iterative partitioning. The proposed scheme is simple, and does not require many stages of processing and parameters. Moreover, experimental results for 2-D images will show that decomposition results would be in good agreement with human insight and quite robust to scaling, rotation, also noise.

With the advent of computer graphics technology that provides the 3-D visualization using display devices, the demand of 3-D image has increased significantly. Therefore the shape decomposition scheme for 3-D image should be required for further shape-related applications. Unfortunately, most 2-D schemes do not extend directly for 3-D image, of which the boundary representation is different from that of 2-D image. Although 1-D boundary contours of 2-D image have a natural arc length parameterization[2, 3], 2-D surfaces of 3-D image do not. In this paper, the methodology of shape decomposition for 3-D image will be shown.

This paper is organized as follows: In Section 2 we propose the novel shape decomposition scheme. Section 3 shows various experimental results on 2-D and 3-D images. Finally, Section 4 concludes this paper.

2. The Proposed Shape Decomposition

In this section, the novel shape decomposition scheme combining mathematical morphology and convex partitioning would be described. Generally, mathematical morphology is not robust to small perturbations and convex partitioning has the lack of mathematical structure. The proposed scheme improves such shortcomings based on mathematical morphology using the ball-shaped SE. The whole process is summarized in Figure 1.

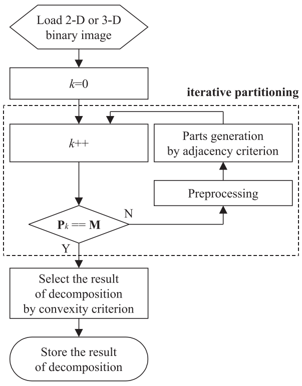

Figure 1. The block diagram of the proposed decomposition scheme.

2.1 Basic definitions

Mathematical morphology is an approach to the shape-oriented image analysis. The language of mathematical morphology is that of set theory. However, the effects of basic morphological operations can be given simple and intuitive interpretations using geometric terms of shape, size and location[4]. In mathematical morphology, a 2-D image, composed of pixels, is defined as a subset of $\{0,1\} \times \{0,1\}$ and a 3-D image, composed of voxels, is defined as a subset of $\{0,1\} \times \{0,1\} \times \{0,1\}$: “0” represents the exterior, while “1” represents the boundary and interior. For an image $\mathbf{M}$ and a point $\mathbf{u}$, “translation” of $\mathbf{M}$ by $\mathbf{u}$ is defined as

$$(\mathbf{M})_{\mathbf{u}} = \{\mathbf{m} + \mathbf{u} \mid \mathbf{m} \in \mathbf{M}\}. \tag{1}$$

There are two fundamental morphological operations: “dilation”,$\oplus$, and “erosion”,$\ominus$, which are defined as follows, respectively:

$$\mathbf{M} \oplus \mathbf{B} = \bigcup_{\mathbf{b} \in \mathbf{M}} (\mathbf{M})_{\mathbf{b}}, \tag{2}$$

$$\mathbf{M} \ominus \mathbf{B} = \bigcap_{\mathbf{b} \in \mathbf{M}} (\mathbf{M})_{-\mathbf{b}}. \tag{3}$$

There are two important morphological operations: “opening”,$\circ$, and “closing”,$\bullet$, which are defined as follows, respectively:

$$\mathbf{M} \circ \mathbf{B} = (\mathbf{M} \ominus \mathbf{B}) \oplus \mathbf{B}, \tag{4}$$

$$\mathbf{M} \bullet \mathbf{B} = (\mathbf{M} \oplus \mathbf{B}) \ominus \mathbf{B}. \tag{5}$$

2.2 Iterative partitioning

In this paper, the ball-shaped SE would be used for the iterative partitioning based on mathematical morphology. If k is a natural number, the ball-shaped SE in 2-D, $\mathbf{B}_k(x,y)$, is defined as $(2k+1) \times (2k+1)$ pixel block with

$$\mathbf{B}_k(x,y) = \begin{cases} 1, & \text{if } |\mathbf{u} - \mathbf{o}| \le k, \\ 0, & \text{if } |\mathbf{u} - \mathbf{o}| > k. \end{cases} \tag{6}$$

where $\mathbf{u}$ and $\mathbf{o}$ imply (x,y) and (k,k) respectively and $|\mathbf{u} - \mathbf{o}|$ implies the Euclidean distance between $\mathbf{u}$ and $\mathbf{o}$. The ball-shaped SE in 3-D, $\mathbf{B}_k(x,y,z)$, is defined as $(2k+1) \times (2k+1) \times (2k+1)$ voxel block with

$$\mathbf{B}_k(x,y,z) = \begin{cases} 1, & \text{if } |\mathbf{u} - \mathbf{o}| \le k, \\ 0, & \text{if } |\mathbf{u} - \mathbf{o}| > k. \end{cases} \tag{7}$$

where $\mathbf{u}$ and $\mathbf{o}$ imply (x,y,z) and (k,k,k) respectively. The ball-shaped SE has point symmetry, thus the proposed scheme can guarantee rotation invariance.

The proposed iterative partitioning would be performed according to increasing k, the number of iterations and size of the ball-shaped SE. The kth partitioning is composed of $\mathbf{M} \circ \mathbf{B}_k$ and $\mathbf{P}_k$, which is defined as

$$\mathbf{P}_k = \mathbf{M} - \mathbf{M} \circ \mathbf{B}_k. \tag{8}$$

Parts, the result of the kth partitioning, are divided into two classes: body class (parts in $\mathbf{M} \circ \mathbf{B}_k$) and branch class (parts in $\mathbf{P}_k$). Then, preprocessing is applied to remove the boundary ambiguity between body and branch classes. The preprocessing is composed of dilations of the body class and branch class sequentially. The preprocessing is similar to “opening”, which breaks narrow isthmuses and eliminates thin protrusions[5]. After preprocessing, merging would be applied to remove redundant parts in branch class. Generally there exist two problems in merging. One is how to select redundant parts and the other is which parts absorb the redundant parts. In this paper, the redundant parts in branch class are selected by the adjacency criterion and the part, which absorb redundant parts, is already determined in body class because the iterative partitioning is based on “opening”. After merging, labeling is applied, which is composed of the extraction of the connected parts using mathematical morphology[5]. This iterative partitioning is completed when $\mathbf{P}_{K+1} = \mathbf{M}$, where K is the last iterative step before body class becomes an empty set.

2.2.1 Adjacency criterion

To explain how to determine the redundant parts, an simple example is presented in Figure 2, which shows the one large part has two adjacent parts. In this paper, the unit of parts would be defined as the entity, which implies the pixel in 2-D and the voxel in 3-D. The number of entities of the 2nd part is equal to that of the 3rd part, but the number of adjacent entities between the 1st part and 2nd part is less than that of adjacent entities between the 1st part and the 3rd part. Generally, the 2nd part represents the more important part than the 3rd part, thus the specific criterion is required to decide the 3rd part should be merged.

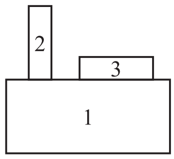

Figure 2. The example for the adjacency criterion.

In this paper, the adjacency criterion is established as follows: The ratio of the ith part in branch class, $R_k(i)$, is defined as

$$R_k(i) = V_k(i)/\{N_k(i)^{\frac{d}{d-1}}\}, \tag{9}$$

where $V_k(i)$ is the number of entities of the ith part, $N_k(i)$ is the number of entities, adjacent to the part which may absorb the ith part, and d is the dimension of entities. $V_k(i)$ and $N_k(i)$ imply the area and length in 2-D image and the volume and area in 3-D image, respectively. Thus $\frac{d}{d-1}$ would be required for $R_k(i)$ to be dimensionless. Then, the adjacency criterion is defined as

$$R_k(i) \quad \leq \quad t_R, \tag{10}$$

where t_R is the threshold for the adjacency criterion. If (10) is satisfied, the ith part should be merged.

2.3 Convexity criterion

All decomposition results, $\mathbf{P}_k$ ranging from 1 to K, from the iterative partitioning would be candidates for the best decomposition result. Thus a good measure to select the best result from candidates should be established. In this paper, the convexity C_k of $\mathbf{P}_k$, which has n parts as the result of the kth partitioning, is defined as

$$\begin{aligned} C_k &= \frac{1}{V}\sum_{i=1}^{n} V_k(i) \cdot C_k(i) \\ &= \sum_{i=1}^{n} \frac{V_k(i)}{V} \cdot \frac{V_k(i)}{H_k(i)}, \end{aligned} \tag{11}$$

where V is the number of entities of $\mathbf{M}$, $C_k(i)$ and $H_k(i)$ is the convexity and number of entities of the convex hull of the ith part in $\mathbf{P}_k$, respectively. This definition of convexity is fairly natural and has been widely used in the field of shape-related analysis[6].

Then the convexity criterion is as follows: $\mathbf{P}_{max}$ which has the highest convexity, C_{max}, among candidates, is the best decomposition result. If there exist decomposition results which share the maximum convexity, the decomposition result with less step would be adopted.

3. Experimental Results

A good shape decomposition should provide an description of a given shape similar to human insight and should be scaling, rotation invariant and insensitive to noise. In this section, we present various experimental results of 100×100 2-D binary images which are consistent with human insight and robust to affine transformations also noise. Moreover, the methodology of shape decomposition for 3-D image would be shown and decomposition results of 3-D image are presented. The threshold for the adjacency criterion, t_R, would be 0.5 for both 2-D and 3-D images.

3.1 For 2-D Binary Image

3.1.1 Several 2-D binary images

Figure 3 shows the iterative partitioning step by step with the number of parts and convexity. In case of "beetle" image, the decomposition result would be Figure 3(c) by the convexity criterion. Figure 4 shows decomposition results of six binary images. We provide the detailed information for 2-D images in Table 1. One can feel these results of decomposition are very similar to human insight, thus the proposed decomposition scheme seems to be adequate to describe a shape.

Table 1. The detailed information for 2-D images

2-D image	# of parts	C_{max}
beetle	9	0.871
bird	4	0.953
device1-1	6	0.968
device5-1	5	0.907
dog	6	0.918
kangaroo	5	0.861
teapot	4	0.954

3.1.2 Robustness to affine transformations

Figure 5 shows decomposition results for scaled images. Figure 5(a)-(c) are scaled images with ratio 1.5 and Figure 5(d)-(f) are scaled images with ratio 0.5. All scaled images have same decomposition results with original images in Figure 3(c), Figure 4(d), and Figure 4(f). As a result of experiments, it could be concluded that the proposed scheme has the property of scaling invariance.

Figure 6 shows decomposition results for rotated images. Figure 6(a)-(c) are rotated images with 9 degree and Figure 6(d)-(f) are rotated images with 45 degree. All rotated images have same decomposition results with original images in Figure 3(c), Figure 4(d), and Figure 4(f). As a result of experiments, it could be concluded that the proposed scheme has the property of rotation invariance.

3.1.3 Robustness to noise

In order to investigate the robustness of the proposed scheme to noise, noise for 2-D binary image is defined as follows: Each boundary pixel and background pixel neighboring with a boundary pixel have some probability of the change of pixel value. According to the random number generation with specific probability, pixel value would be changed stochastically.

Figure 7 shows decomposition results for noise-corrupted images. Figure 7(a)-(c) and Figure 7(d)-(f) show results of three binary images corrupted by 5% and 20% noises, respectively. All noise-corrupted images have same decomposition results with original images in Figure 3(c), Figure 4(d), and Figure 4(f). As a result of experiments, it could be asserted that the proposed scheme is insensitive to noise.

3.2 For 3-D Binary Image

Generally 3-D image is represented by the mesh data. The mesh data is specified by a collection of surfaces, which are represented by triangular meshes, and is obtained by the laser scanner and space encoding range finders. However, the mesh data is described using real coordinates, yielding has the irregular grid structure. Due to such properties, the shape decomposition of the mesh data is very difficult. In this paper, the voxelization[7], which is independent of the extent of the mesh data and can be performed in any desired resolution, would be adopted to make the structure of 3-D image regular. As the result of the voxelization, the 3-D image of triangular meshes is represented as a binary image, of which the resolution is defined as $X \times Y \times Z$, where X, Y, and Z are how the bounding box is divided along x, y, and z axis, respectively. Then the binary image would be composed of the voxel which approximates a cube. The more the resolution of the binary image grows, the more the number of voxels increases. Thus the resolution of the binary image is closely related to the performance of the 3-D image decomposition. Thus the optimal resolution selection scheme is important, but that is beyond the scope of this paper. Thus the voxelization is empirically performed until voxels of of the boundary and interior of the binary image, “1” voxels, are more than 5000.

Figure 8,9 shows the decomposition of six 3-D images. Figure 8(a)-(c) and Figure 9(a)-(c) show that images are represented by wireframe. Figure 8(d)-(f) and Figure 9(d)-(f) show that parts are represented by voxels. As shown in Figure 8(d), “phone” image has 3 parts: a grip, a transmitter, and a receiver. As shown in Figure 8(e), “cactus” image has 3 parts: a body and two branches. As shown in Figure 8(f), “cow” image has 7 parts: a body, a head, a tail, and four legs. As shown in Figure 9(d), “toydog” image has 7 parts: a body, a head, a tail, and four legs. As shown in Figure 9(e), “eagle” image has 7 parts: a body, a head, a tail, two wings and two legs. As shown in Figure 9(f), “santa” image has 6 parts: a body, a head, two arms and two legs.

We provide the detailed information for 3-D images in Table 2. One can feel these decomposition results are very similar to human insight, thus the proposed decomposition scheme seems to be adequate to describe 3-D image as well as 2-D image.

Table 2. The detailed information for 3-D images

3-D image	# of vertices	# of meshes	Resolution	C_{max}
phone	338	672	$52 \times 17 \times 15$	1.000
cactus	620	1236	$58 \times 35 \times 9$	0.994
cow	2904	5804	$46 \times 28 \times 15$	0.952
toydog	1944	3808	$42 \times 34 \times 20$	0.943
eagle	16778	33324	$93 \times 46 \times 25$	0.833
santa	18946	37888	$52 \times 51 \times 22$	0.875

4. Conclusion

In this paper, the novel shape decomposition scheme for 2-D and 3-D images was presented. The proposed scheme is composed of the iterative partitioning based on mathematical morphology and adopts the adjacency criterion to select meaningful parts. Finally the best decomposition result is selected by the convexity criterion. From the experimental results, it was confirmed that the proposed scheme yielded the decomposition which was similar to human insight and was found to be robust to various perturbations: scaling, rotation and noise. Moreover, it was verified that the proposed scheme could cover 3-D image as well as 2-D image using voxelization. This decomposition result can be used in matching, retrieval, and so on. While our results are promising, it is necessary to develop the proposed scheme for various applications. For example, the hierarchical structure should be adopted for more accurate decomposition. The use of shape decomposition for shape description will be the subject of further research.

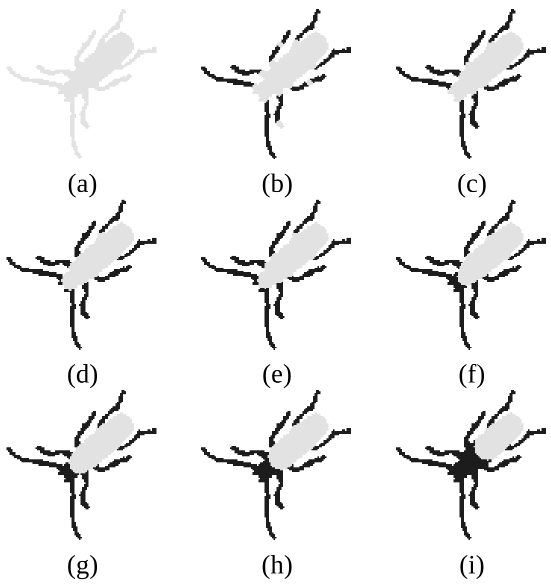

Figure 3. Steps of the iterative partitioning for “beetle" image: (a) 1st partitioning with 1 part and $C_1 = 0.446$, (b) 2nd partitioning with 14 parts and $C_2 = 0.862$, (c) 3rd partitioning with 9 parts and $C_3 = 0.871$, (d) 4th partitioning with 9 parts and $C_4 = 0.870$, (e) 5th partitioning with 9 parts and $C_5 = 0.861$, (f) 6th partitioning with 8 parts and $C_6 = 0.800$, (g) 7th partitioning with 8 parts and $C_7 = 0.789$, (h) 8th partitioning with 6 parts and $C_8 = 0.732$, and (i) 9th partitioning with 4 parts and $C_9 = 0.567$.

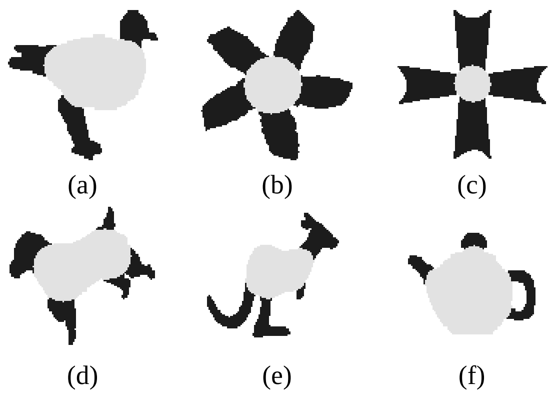

Figure 4. Results of decomposition for six 2-D binary images: (a) “bird" image with 4 parts, (b) “device1-1" image with 6 parts, (c) “device5-1" image with 5 parts, (d) “dog" image with 6 parts, (e) “kangaroo" image with 5 parts, and (f) “teapot" image with 4 parts.

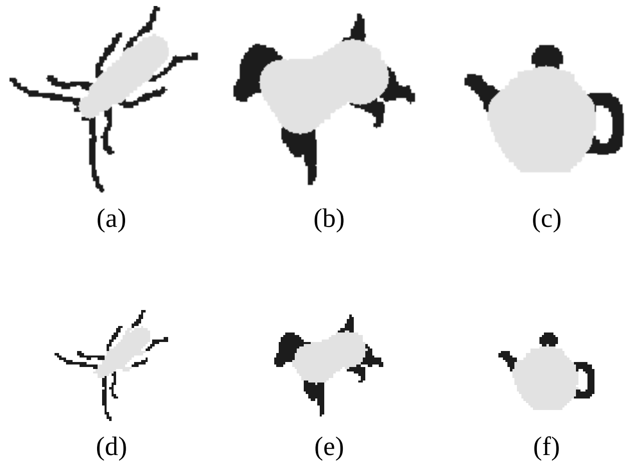

Figure 5. Results of decomposition for scaled images: (a) “beetle" image scaled with ratio 1.5, (b) “dog" image scaled with ratio 1.5, (c) “teapot" image scaled with ratio 1.5, (d) “beetle" image scaled with ratio 0.5, (e) “dog" image scaled with ratio 0.5, and (f) “teapot" image scaled with ratio 0.5.

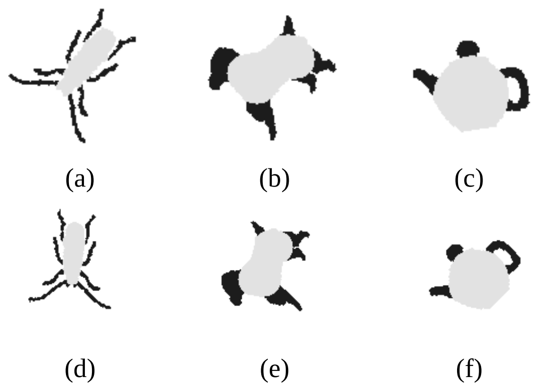

Figure 6. Results of decomposition for rotated images: (a) “beetle" image rotated with 9 degree, (b) “dog" image rotated with 9 degree, (c) “teapot" image rotated with 9 degree, (d) “beetle" image rotated with 45 degree, (e) “dog" image rotated with 45 degree, and (f) “teapot" image rotated with 45 degree.

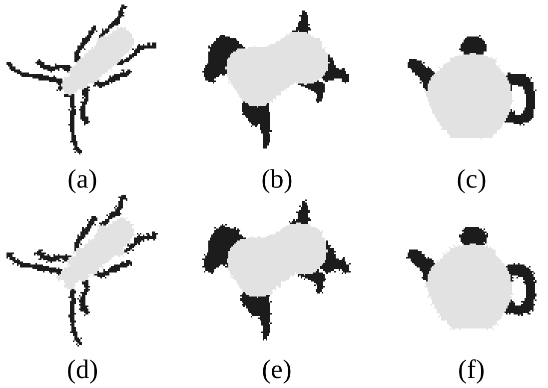

Figure 7. Results of decomposition for noise-corrupted images: (a) “beetle" image with 5% noise, (b) “dog" image with 5% noise, (c) “teapot" image 5% noise, (d) “beetle" image with 20% noise, (e) “dog" image with 20% noise, and (f) “teapot" image with 20% noise.

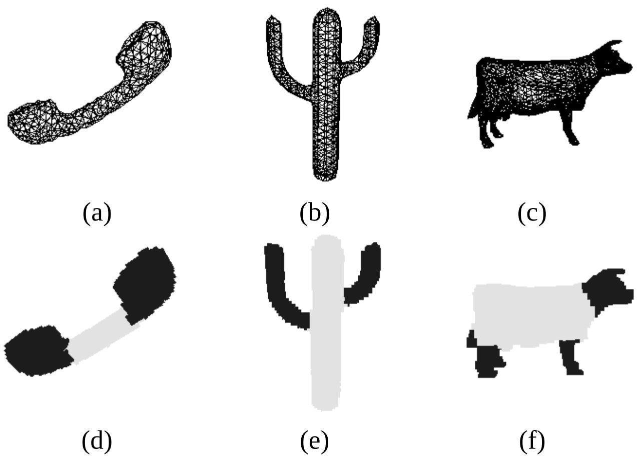

Figure 8. Experimental results for 3-D images; (a) “phone" image, (b) “cactus" image, (c) “cow" image, (d) “phone" image with 3 parts, (e) “cactus" image with 3 parts, and (f) “cow" image with 7 parts.

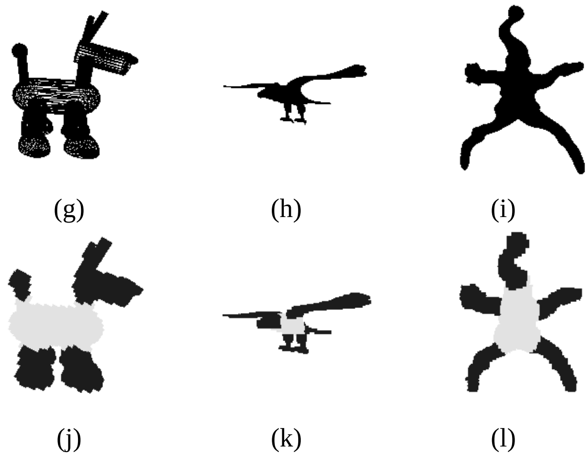

Figure 9. Experimental results for 3-D images; (a) “toydog" image, (b) “eagle" image, (c) “santa" image, (d) “toydog" image with 7 parts, (e) “eagle" image with 7 parts, and (f) “santa" image with 6 parts.

References

[1] M. Bober. Mpeg-7 visual shape descriptors. *IEEE Trans. Circuits and Systems for Video Technology*, 11(6):716–719, June 2001.

[2] F. Mokhtarian and A. Mackworth. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(8):789–805, August 1992.

[3] L. J. Latecki and R. Lakämper. Convexity rule for shape decomposition based on discrete contour evolution. *Computer Vision and Image Understanding*, 73(3):789–805, March 1999.

[4] P. A. Maragos and R. W. Schafer. Morphological skeleton representation and coding of binary images. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 34(5):1228–1244, October 1986.

[5] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison Wesley, 1992.

[6] P. L. Rosin. Shape partitioning by convexity. *IEEE Trans. Systems Man and Cybernatics-Part A: Sytems and Humans*, 30(2):202–210, March 2000.

[7] D. H. Kim, I. D. Yun, and S. U. Lee. Graph representation by medial axis transform for 3d image retrieval. In *Proceedings of SPIE on Three-Dimensional Image Capture and Applications IV*, pages 223–230, 2001.