Paper

A Modified Exoskeleton for 3D Shape Description and Recognition

Rajalida Lipikorn* Member
Akinobu Shimizu* Non-member
Yoshihiro Hagihara* Non-member
Hidefumi Kobatake* Non-member

Three-dimensional(3D) shape representation is a powerful tool in object recognition that is an essential process in an image processing and analysis system. Skeleton is one of the most widely used representations for object recognition, nevertheless most of the skeletons obtained from conventional methods are susceptible to rotation and noise disturbances. In this paper, we present a new 3D object representation called a modified exoskeleton (mES) which preserves skeleton properties including significant characteristics about an object that are meaningful for object recognition, and is more stable and less susceptible to rotation and noise than the skeletons. Then a 3D shape recognition methodology which determines the similarity between an observed object and other known objects in a database is introduced. Through a number of experiments on 3D artificial objects and real volumetric lung tumors extracted from CT images, it can be verified that our proposed methodology based on the mES is a simple yet efficient method that is less sensitive to rotation, noise, and independent of orientation and size of the objects.

Keywords: shape representation, shape recognition, skeleton, modified exoskeleton, distance transformation.

1. Introduction

Due to the progress of modern technology, 3D volumetric images with high resolution, such as multi-slice CT images, are made available for use in image processing, and the structure of a 3D object is more complicated than that of the 2D which makes an image analysis by a human observer even more difficult. Therefore, we need a computer system for 3D shape description and recognition.

This paper presents a new 3D object representation and a shape recognition system for 3D binary images. The system is comprised of two main processes: shape description and matching processes. Given a database of known objects which hereafter are called models, the objective is to identify an unknown observed object with one of the models in the database regardless of its size, position, and orientation. In shape description process, an object is generally reduced to its compact representation, such as boundary or skeleton, where important information for use in the recognition process can be retrieved. Besides, it frequently occurs that an object may be distorted by noise, movement and other errors during the process, therefore, a good selection of object representation that is less susceptible to such errors is also essential. Moreover, our concentration is on shape representations which transform an object into some meaningful numerical values used to calculate a similarity measure between two objects. We define efficient representations as those which give small similarity measure for similar shapes whereas large value for different shapes. In the recognition system, once the object is represented by its compact form, the observed object is compared with each model by their representations, and the similar models ranked by similarity measures are returned as results.

During the past few decades skeleton and its applications have been attractive research areas in image processing and analysis. Skeleton is a structural representation that provides an effective object representation for shape recognition [1] and visualization [2] [3] [4] in both two dimensions and three dimensions. A skeleton of a volumetric object provides a compact description, and can be extracted by various methods such as those introduced in [3] [4] [5] [6] [7] [8]. It is, however, sensitive to subtle shape changes caused by rotation or noise. Existence of a small hole or cavity can cause major change in skeleton structure. This sensitivity is an open problem that motivates us to derive an object representation which is more tolerant of noise.

We have proposed a new object representation called a modified exoskeleton (mES) for two dimensional binary images [9]. The mES is a special type of skeleton that is less sensitive to distortion caused by rotation and noise. It contains significant information for use in object recognition and reconstruction as the skeleton does. In our previous work, exoskeletonization is based on the 4-neighbor distance transformation and the effectiveness of the proposed method has been demonstrated only for 2D objects [9]. It is well known that the skeleton obtained by the 4-neighbor distance transformation

*Tokyo University of Agriculture & Technology
Koganei, Tokyo 184-8588, Japan

is sensitive to rotation, and the utilization of Euclidean distance transformation is effective in alleviating the rotation sensitivity. This paper shows that the mES is effective in recognizing 3D objects whose skeletonization is based on the Euclidean distance transformation. The proposed 3D object recognition system, which uses mES for shape description and has a matching process based on the mES, is simple, fast, and independent of size, position, and orientation of the 3D observed object. In order to evaluate the effectiveness, we demonstrate how well it can recognize 3D artificial and real objects, and compare the results against those of the skeleton. A few methods for shape recognition based on 3D skeletons have been introduced in the literature, e.g., the method presented in [4] uses the mix application of the medial axis transformation and the symmetry transformation to extract a skeleton that makes the skeleton rotation invariant, however, it is still sensitive to noise.

The remaining of this paper is organized into 4 sections. Basic notions and the exoskeletonization algorithm for 3D objects are presented in section 2. The matching algorithm is described in section 3. Several experiments are discussed in section 4 followed by conclusions in section 5.

2. Basic Notions

In this section we review a fundamental definition of the original exoskeleton and present a definition for 3D mES. Our proposed mES has been developed to alleviate the distortion caused by image noise and segmentation error by combining the skeleton with the original exoskeleton and the invariant property of a symmetrical object. Then an exoskeletonization algorithm based on Euclidean distance transformation is described along with illustrations.

2.1 Original Exoskeleton In a binary image, if we define an object as a set of feature pixels, a skeleton can be extracted from either a set or the complement of a set. A skeleton extracted from a set of feature pixels is a typical skeleton, whereas a skeleton of the complement of the set is called an *exoskeleton*. Even though the exoskeleton is not used directly as an object representation, its concept is partially included in many other object representations, such as Voronoi diagram, Delaunay triangulation [10] [11], and a well-known tool in segmentation, namely $SKIZ$.

J. Prewitt used the skeleton of the background of isolated objects(called by her an "exoskeleton") in order to characterize their zones of influence [12]. Let F be a binary image or a set and $X_1, X_2, ..., X_n$ be isolated objects of F:

$$F = \bigcup_{i=1}^{n} X_i, i \neq j \rightarrow X_i \neq X_j \cdots\cdots\cdots\cdots (1)$$

If $S(F)$ defines the skeleton of set F then the skeleton of the complement of the set F called the exoskeleton of F can be expressed as

$$ES(F) = S(\bar{F}) \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (2)$$

2.2 Modified Exoskeleton in Three Dimensions The 3D mES is generalized from the 2D mES [9] and is defined as follows:

Definition Let $X = \{(x, y, z)\}$ be an object in the binary image F, and $C = \{(x, y, z)\}$ be the intersection of $\bar{X}$ and a sphere that circumscribes the object X as shown in Fig. 1. The mES and the exoskeleton function $mesf$ can be expressed as

$$\begin{aligned} mES(X) = \{(x, y, z) | (x, y, z) \in C, \\ D(x, y, z) \geq \max_{(p,q,r) \in N(x,y,z)} [D(p, q, r)], \\ D(p, q, r) = \min_{(l,m,n) \in \bar{C}} [d((p, q, r), (l, m, n))]\}, \end{aligned}$$

$$mesf(x, y, z) = \{D(x, y, z) | (x, y, z) \in mES(X)\}$$

where $N(x, y, z)$ is a set of neighbors of (x, y, z), and $d((x, y, z), (l, m, n))$ is the Euclidean distance from a voxel (x, y, z) in the intersection C to a voxel (l, m, n) in $\bar{C}$. In general an exoskeleton can be represented by a local parametrization as $mES = (x, y, z, mesf(x, y, z))$.

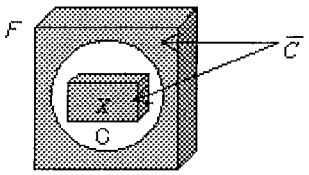


Fig. 1. Illustration of an object in spherical background (Cross section of 3D image).

The general concept of the 3D mES is analogous to the 2D mES except that a sphere is used as a symmetrical background and it is, in particular, a skeleton of a sphere with the original object embedded in the center. Beside the properties previously mentioned, the mES also allows the original object to be partially or fully reconstructed.

A typical example of the skeleton and the mES of an ideal sphere and a sphere with a small hole and a tiny crack are shown in Fig. 2. For better visualization, only the central parts of the original object and the deformed object are extracted and displayed. It can easily be notified that the skeleton structure is very sensitive to noise.

2.3 Exoskeletonization Algorithm The exoskeleton extraction in this paper is based on the algorithms for exact Euclidean distance transformation (EDT), and the skeleton extraction introduced in [13] [14]. The EDT converts a binary image into an image whose each voxel has a value corresponding to the distance to its nearest background voxel. In the past, Euclidean distance transformation was seldomly used because its computation time was quite costly. Fortunately, many improved algorithms have been introduced in the literature [13] [15] [16] which reduce the computation time significantly and make it possible to compute Euclidean distance for large binary images.

The exoskeleton extraction process is analogous to the skeleton extraction, except that it requires a preprocess to generate a spherical background. This process is applicable to any binary images, and can be performed in the following steps:

Step 1. An object is embedded in the center of a generated sphere whose radius R is equal to the longest distance from the centroid to the boundary of the object plus a constant value ϵ to ensure that the sphere covers the original object thoroughly. The radius R can be obtained from

$$R = \max_{(l,m,n)\in\bar{X}} \{d((l_g, m_g, n_g), (l, m, n))\} + \epsilon \cdots (3)$$

where (l_g, m_g, n_g) is the centroid, (l, m, n) is any voxel in the object, and $d(.,.)$ is the Euclidean distance between two voxels (see Fig. 3(b)).

Step 2. Next, the distance value $D(x, y, z)$ for each voxel in the intersection C of the sphere and $\bar{X}$ is computed by using Euclidean distance transformation [14].

Step 3. Finally, the intersection C is reduced to the mES in two scans [14] and the distance corresponding to each voxel is stored in the $mesf$ (see Fig. 3(c)).

The advantage of the proposed mES is that it is robust against rotation and noise disturbances, and the reasons are as follows: (1) The main structure of any conventional skeleton concentrates, in general, around the center of gravity of an object where the skeleton function values are large which makes the skeleton very sensitive once the object is distorted by noise. The mES, on the contrary, disperses around the object (see Fig. 2(c)) and so every voxel has its own significant. (2) The effects of noise and/or rotation spread throughout the entire skeleton and the values of the skeleton function are also influenced in a wide range. Distortion can easily disrupt the main structure of the skeleton as well as the skeleton function. The variations of these two components cause the similarity measure to diverge. Not like the mES, the extent of noise disturbance and/or rotation merely affects the local parts of the mES nearby the noise and the variation of the $mesf$ is relatively small (see Fig. 2(f)).

3. Matching Algorithm

Our proposed matching algorithm has been developed for use with either the skeleton or the mES, it estimates the similarity between two objects based on the distance measure and the (exo)skeleton function. The fundamental scheme has been implemented on 2D binary images [17] and is extended to 3D version.

In the case of the mES, the algorithm searches for the local similitude between the mES of the observed object and the mES of a model in two steps. First, it compares the size of the two objects, the mES of the larger object is assigned as a mES_1 and that of the smaller one as a mES_2. To handle translation, the origin of the axial system is moved to the center of gravity of the objects so that both objects have center of gravity at the origin and one object is superimposed on the other. It, then, matches each voxel in the mES_1 with its nearest one in the mES_2. In finding the best match for each voxel, the $mesf$ is also taken into account. When matching voxels of the two mESs that represent similar components, their $mesf$s should be similar. On the other hand, these values should be different if they represent different components. The $mesf$ is used as a weight function to ensure that the matching of voxels representing similar components has less effect on similarity measure whereas the matching of different components has greater effect. The distance between the voxels and the difference between their $mesf$s are accumulated and used to compute the similarity measure E_1 with respect to the first object. Second, each voxel in the mES_2 is matched with its nearest one in the mES_1, and the similarity measure E_2 with respect to the second object is computed. The similarity measure E, which provides a mean for determining the overall similarity, can be calculated from the following equations:

$$E = \min_a \{E_1(a) + E_2(a)\} \cdots\cdots\cdots\cdots (4)$$

$$E_1(a) = \{\sum_u \sum_v \sum_w sd_1(u, v, w)\}/K_1 \cdots\cdots (5)$$

$$E_2(a) = \{\sum_x \sum_y \sum_z sd_2(x, y, z)\}/K_2 \cdots\cdots (6)$$

$$sd_1 = \sqrt{(x' - au)^2 + (y' - av)^2 + (z' - aw)^2 + [mesf_2(x', y', z') - a(mesf_1(u, v, w))]^2}$$

$$sd_2 = \sqrt{(x - au')^2 + (y - av')^2 + (z - aw')^2 + [mesf_2(x, y, z) - a(mesf_1(u', v', w'))]^2}$$

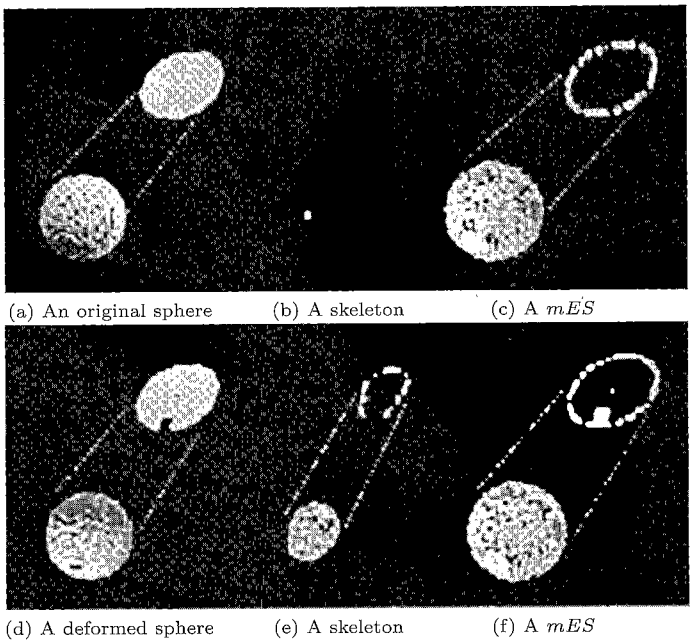
$$(x', y', z')^t = Rot(\alpha, \beta, \gamma)(x, y, z)^t \cdots\cdots\cdots (7)$$

$$(u', v', w')^t = Rot(\alpha, \beta, \gamma)(u, v, w)^t \cdots\cdots\cdots (8)$$

where (x, y, z), (x', y', z'), (u, v, w), (u', v', w') are coordinates of voxels in the two mESs, a is a scaling factor, and K_1, K_2 are the numbers of voxels in mES_1 and mES_2, respectively. The $Rot(\alpha, \beta, \gamma)$ represents a rotation matrix around the three major axes. These angles are iteratively varied within the range of $-90^\circ \leq \alpha, \beta, \gamma \leq 90^\circ$ to find the values that best match with the orientation of the object. Meanwhile, the scaling factor is exhaustively searched within a range. This search range can be set from ρ which is obtained from the ratio of the longest distances from the boundary voxels to the centroids of the objects. This longest distance is actually equal to the radius of the spherical background when ε is equal to zero, and thus, the ρ is computed from:

$$\rho = \left.\frac{R_2}{R_1}\right|_{\epsilon=0}, \cdots\cdots\cdots\cdots (9)$$

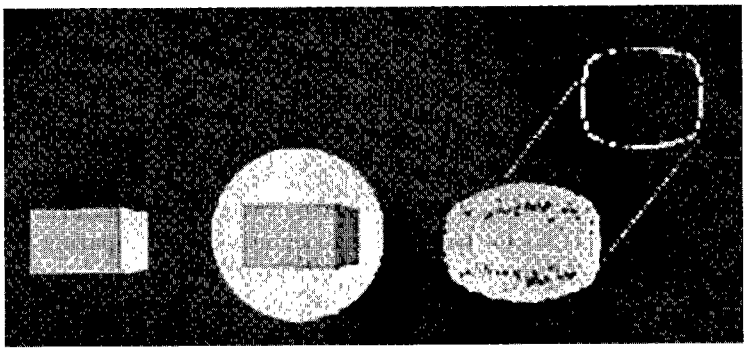
where the subscripts of R denote the name of the mESs. This matching process is performed on every pair of objects and the algorithm finally recognizes the observed object as a model in the database whose similarity measure E is minimal.



(a) An original sphere (b) A skeleton (c) A mES

(d) A deformed sphere (e) A skeleton (f) A mES

Fig. 2. Examples of the skeletons and the mESs of 3D objects with their central slices.



(a) An original object (b) After step 1 (c) A mES and its central slice

Fig. 3. Exoskeletonization process for 3D object.

4. Experimental Results

To evaluate the robustness of the mES and the performance of the recognition system, we compared the results obtained from the mES-based method with those obtained from the skeleton-based method. Both the skeleton and the mES extractions were based on the Euclidean distance transformation. The experiments were conducted on a database of nine 3D artificial binary objects in images of size 80 x 80 x 60 voxels, and a database of seventeen volumetric lung tumors segmented from CT images of size 152 x 152 x 152 voxels, where the scaling range used in all the experiments was $\rho * 1.01^{-15} \leq a \leq \rho * 1.01^{15}$ interval of 1.01.

4.1 Artificial Data The experiments for artificial objects were divided into three different cases. The first experiment was the matching between different shapes; i.e., between the observed object and each model in the database. The second experiment consisted of two tests, one was the matching between each object and its rotated versions, and the other one was the matching between each rotated version and models in the database. The degree of rotation $\theta_x, \theta_y, \theta_z$ around the three coordinate axes varied within the range of $-90° \leq \theta_x, \theta_y, \theta_z \leq 90°$. Finally, the third experiment was the matching between noisy versions of each object and models in the database where two kinds of noise were added to the original object, namely internal noise and boundary noise.

In the first experiment, the matching between different shapes was performed on the object database consists of nine different shapes. Examples of the original objects and their mESs are as shown in Fig. 4. The confusion matrix of the similarity measure E based on the skeleton and the mES is shown in Table 1, where the upper-triangular matrix shows the matching results of the skeleton-based method, and the lower-triangular matrix shows the results of the mES-based method. Smaller value of E indicates similarity whereas larger value of E indicates dissimilarity. It can be verified that there is an anology between these two approaches, i.e., most of the matching results give the same order of similarity. However, some differences can be seen from the

confusion matrix. For example, the order of similarity for a 'Cross' shape based on the skeleton in ascending order is an 'E', a 'Cylinder', and a 'Box' while the order of similarity for the same shape based on the mES in ascending order is a 'Box', a 'Sphere', and a 'Cylinder'. The reason that the 'Cross' shape is best matched with the letter 'E' for the skeleton-based method is because the letter 'E' is composed of both vertical and horizontal components similar to the 'Cross'.On the other hand, the position of these vertical and horizontal components with respect to the spherical background can be determined from the $mesf$ since the values stored in the $mesf$ can tell how far the boundary of an object is from the spherical background, thus the 'Cross' is more like a 'Box' for the mES-based method. It is difficult to explicitly declare which method gives better matching results, it depends on each observer's perception. Most of the matching results obtained from both methods are reasonable and coincide with human perception. The quantitative evaluation is given in the second experiment. Another interesting result is the matching of a 'Sphere' shape, the order of similarity based on skeleton seems to be reasonable, but the similarity measures are too extreme to be acceptable. The main reason for such extreme similarity measures between the 'Sphere' and other objects is because the skeleton of the sphere contains only one voxel.

The matching based on the mES was repeated several times with different sizes of spherical background. Various sizes of the sphere were used in order to examine the effect of the radius over the similarity measures. The constant value ϵ was varied from 1 to 50 in 10 steps. From a number of experiments, the results reveal that the size of the sphere imposes no effect on the matching results because the orders of similarity are always the same no matter how large or how small the sphere is. The only difference is that the similarity measure gradually increases as the sphere gets larger, this is because the similarity measure relies on the distance between two voxels and the $mesf$ value, and these values increase as the sphere gets larger. However, this increment is propotional to others and not significant enough to affect the order of similarity. Thus, smaller size is more preferable in terms of memory requirement, and computation time as the size of the mES depends on the size of the sphere. In our experiments, $\epsilon = 2$ was used.

In the second experiment, each object was rotated around the three coordinate axes in the range $-90^\circ \leq \theta_x, \theta_y, \theta_z \leq 90^\circ$, and 36 rotated versions were produced. The experiment was divided into two cases. The first case was to compare each object with its rotated versions to observe the variation of its representation. The results obtained from the skeleton-based and the mES-based methods were evaluated by their relative dispersions or coefficients of variation (CV). It measures variability in relation to the mean, and lower value of CV is more preferable since it indicates that less variation occurs. To calculate the coefficient of variation, we find the mean μ and the standard deviation σ of the similarity measures E between each object and its rotated versions. The coefficient of variation expresses the variation of the throughput as a percentage of the mean and can be calculated as follows:

$$CV = \left(\frac{SD}{mean}\right) * 100 = \left(\frac{\sigma}{\mu}\right) * 100. \quad \cdots\cdots (10)$$

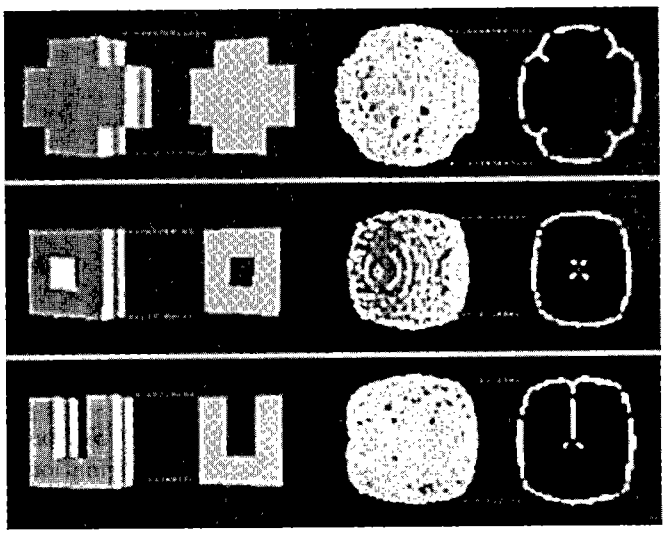
The results shown in Table 2 indicate that all of the coefficients of variation of the mES-based method are lower than those of the skeleton-based method which means that the mES is less sensitive to distortion caused by rotation even though the skeleton obtained from Euclidean distance transformation is known to be insensitive to rotation.

In the latter case of the second experiment, each rotated version was compared with models in the database and the following relative distance with respect to the two distributions was calculated.

$$Relative\ \ distance(RD) = |\mu_1 - \mu_2|/\sigma, \quad \cdots\cdots (11)$$

where $\sigma = (\sigma_1 + \sigma_2)/2$. The mean μ_1 and the standard deviation σ_1 were calculated from the similarity measures between rotated versions of the object and the models except the object itself, whereas the mean μ_2 and the standard deviation σ_2 were calculated from those between the original object and its rotated versions. The relative distance for each object is shown in Table 3. It can be notified that most of the relative distances derived from the mES-based method are larger than those derived from the skeleton-based method. We also evaluated the performance of the mES-based shape recognition by comparing the error rates based on the mES with those based on the skeleton as shown in Table 4. The average error rate for the mES is 0.068 which is smaller than 0.133 for the skeleton. By observing the experimental results in detail, we found that most of the errors occurred when the degrees of rotation were in the vicinity of 45° for both methods.

The third experiment was performed regarding recognition of noisy objects. Most of the conventional skeletons are known to be very sensitive to noise. If we consider a hole or a cavity as noise, only one additional voxel of noise can absolutely transform the skeleton structure. In the last experiment, each object was distorted by internal noise and boundary noise. Noise was added to the object by converting object voxels to background voxels with a probability δ. The amount of added noise can be varied by changing the value of δ within the range of $0 \leq \delta \leq 1$. For internal noise, noise was added to the object in arbitrary position, and fifteen noisy versions were generated for each object. For boundary noise, noise was added to the object by randomly selecting the boundary voxels and their neighbor background voxels and reversing these selected voxels, and fifteen deformed versions were generated for each object. Each noisy object was then compared with models in the database. The skeletons and the mESs of the 'Cross' shape and its noisy version are shown in Fig. 5. The error rates of both methods are shown in Table 4. It can be seen that a small cavity caused the main structure of the skeleton to vary in a wide range, whereas, the main structure of



(a) An original object (b) A mES of (a)

Fig. 4. Examples of 3D objects and their mESs with their cross-section slices.

Table 1. Confusion matrix of similarity E based on the skeleton and the mES

Object	Box	Ring	d	Cylinder	Sphere	Cross	E	U	Ribbon
Box	0	20.10	19.00	9.17	101.43	11.38	11.75	21.73	13.22
Ring	11.20	0	7.50	19.04	129.16	13.36	6.17	3.89	7.44
d	37.76	17.64	0	22.38	135.92	15.09	5.59	7.26	10.63
Cylinder	9.21	15.70	49.74	0	88.09	10.06	16.47	18.02	14.79
Sphere	11.84	18.49	56.68	6.28	0	109.93	125.31	130.24	132.306
Cross	9.28	15.99	37.33	11.32	9.31	0	8.92	16.13	11.64
E	17.29	9.78	13.43	25.85	28.85	18.28	0	6.81	10.35
U	10.37	5.99	27.36	13.94	18.06	15.73	12.04	0	6.308
Ribbon	14.57	11.44	23.87	17.26	22.25	16.81	11.55	9.53	0

(Upper-triangular matrix: skeleton-based method, lower-triangular matrix: mES-based method)

Table 2. Variation of matching results between each object and its rotated versions based on the skeleton and the mES

	skeleton				*mES* ($\epsilon = 2$)			
Object	maximum	mean (μ)	SD (σ)	CV	maximum	mean (μ)	SD (σ)	CV
Box	4.25	2.67	1.76	65.92	3.00	2.22	1.00	45.14
Ring	8.28	2.99	2.65	88.42	10.37	4.02	3.13	77.97
d	8.94	3.21	2.88	89.91	8.86	3.64	2.78	76.35
Cylinder	7.10	2.87	2.33	81.46	4.94	2.1146	1.55	73.29
Sphere	80.05	57.08	39.01	68.33	3.68	2.21	1.21	54.89
Cross	9.81	5.03	3.81	75.76	8.96	4.48	3.38	75.28
E	10.33	4.70	3.69	78.59	8.22	4.02	2.97	74.00
U	6.65	3.11	2.46	79.17	5.19	2.78	1.82	65.44
Ribbon	5.12	1.88	1.56	83.12	3.92	1.57	1.28	81.63

Table 3. Relative distances of the skeleton-based and the mES-based methods on rotation.

	skeleton					*mES* ($\epsilon = 2$)				
Object	μ_1	μ_2	σ_1	σ_2	relative distance	μ_1	μ_2	σ_1	σ_2	relative distance
Box	23.09	2.67	30.11	1.76	1.28	13.50	2.22	10.25	1.00	2.01
Ring	22.96	3.00	40.38	2.65	0.93	11.80	4.02	6.01	3.13	1.70
d	24.82	3.21	42.24	2.88	0.96	29.32	3.64	17.97	2.78	2.47
Cylinder	22.00	2.87	25.66	2.33	1.37	16.59	2.11	14.39	1.55	1.82
Sphere	105.82	57.08	42.83	39.00	3.98	19.08	2.21	16.57	1.21	1.89
Cross	21.83	5.03	33.37	3.81	0.90	14.89	4.48	10.11	3.78	1.49
E	21.26	4.70	39.28	3.69	1.21	15.23	4.02	8.66	2.97	1.93
U	23.37	3.11	40.71	2.46	0.94	12.56	2.78	7.72	1.82	2.05
Ribbon	22.96	1.88	41.23	1.56	0.99	14.15	1.57	7.18	1.28	2.97
average					*1.40*	*average*				*2.04*

the mES was almost unchanged and only a few voxels nearby the cavity were added to the mES. The majority of errors occur when the noise density increases, however, improvement in error rates for noisy objects based on the mES is exemplified in Table 4, the error rates of the mESs are lower than those of the skeletons for both boundary and internal noise. The results yield the average error rate of 0.081 for the mES and 0.174 for the skeleton.

Table 4. Error rates of the skeleton-based and the mES-based ($\epsilon = 2$) methods regarding rotation and noise disturbances.

	Rotated version		*Boundary noise*		*Internal noise*		*Noisy version*		*Total*	
Object	Skeleton	mES	Skeleton	mES	Skeleton	mES	Skeleton	mES	*Skeleton*	*mES*
Box	0	0	0.200	0.133	0.333	0.267	0.267	0.200	*0.121*	*0.091*
Ring	0.361	0.306	0.067	0.067	0.067	0.133	0.067	0.100	*0.227*	*0.212*
d	0.278	0.028	0	0.133	0.067	0.333	0.033	0.233	*0.167*	*0.121*
Cylinder	0	0	0	0	0	0	0	0	*0*	*0*
Sphere	0	0	0.933	0	0.933	0	0.933	0	*0.424*	*0*
Cross	0.083	0.111	0	0	0.133	0.067	0.067	0.033	*0.076*	*0.076*
E	0.056	0.028	0	0	0	0	0	0	*0.030*	*0.015*
U	0.250	0	0	0	0.067	0	0.033	0	*0.152*	*0*
Ribbon	0.167	0.139	0.133	0.133	0.200	0.200	0.167	0.167	*0.182*	*0.152*
average	*0.133*	*0.068*	*0.148*	*0.052*	*0.200*	*0.111*	*0.174*	*0.081*	*0.153*	*0.074*

4.2 CT Lung Data We conducted the experiment on a database of seventeen real volumetric images of lung tumors segmented from CT images by a 3D region growing method. The CT data have a resolution of 512 x 512 where the pixel size is 0.38 $\sim$ 0.42 mm and slice interval and thickness are 1.0 $\sim$ 2.0 mm which were altered to be equal to the pixel size by using cubic interpolation. After segmentation of each volumetric lung tumor, the image size was cropped to 152 x 152 x 152 voxels. Each image has one tumor and there are nine malignant and eight benign tumors as shown in Fig. 6. In general, benign tumor has smooth and round shape, whereas malignant tumor tends to have irregular shape with bulges.

In this experiment, we first calculated the similarity measure between every pair of lung tumors, then the k-clustering algorithm was used to classify lung tumors as either benign or malignant. The initial prototypes(seeds) of the two classes are selected from a pair of the tumors whose similarity measure is maximal, then each tumor is classified into one of these classes using k-clustering algorithm ($k = 2$) with the sum of the similarity measure as the cost function. To evaluate the classification results based on the mES, we compared them with those of the skeleton, and the results indicate four misclassifications for the mES-based method and five misclassifications for the skeleton-based method. In Fig. 6, misclassified tumors are surrounded by rectangles. As can be seen from this figure, the errors of the mES-based classification mostly occur in the case of benign tumors with protrusions which may result from parts of the tumor that touch the blood vessels, and the malignant tumors with smooth surface.

5. Conclusions

In this paper, a 3D binary object representation called a modified exoskeleton (mES) and a methodology for 3D shape recognition are presented. The mES integrates the original exoskeleton with rotation invariant property of symmetrical object and the skeleton function into an efficient object representation. Its main purpose is to alleviate distortion caused by rotation and noise that most conventional skeletons have encountered. A recognition system is then used to evaluate the robustness of the mES against that of the skeleton in how well they can represent the object. The system extracts the skeleton/mESs from the objects using the Euclidean distance transformation, and matches the observed object with the most similar model in the database regardless of its size, position, and orientation. The matching algorithm presents the similarity between two objects in term of distance measure where the object is recognized as a shape model whose distance is minimal.

The results obtained from the experiments on artificial objects are quite promising. The statistical results in Tables 2, and 3 indicate that the mES is more stable to distortion caused by rotation, and the error rates shown in Table 4 ensure that the mES is also less sensitive to noise disturbance than most conventional skeletons. From the experiments on artificial objects, the overall recogintion rate of the mES-based method is 92.6% whereas that of the skeleton-based method is 84.7% which indicates an improvement of 7.9%. For real volumetric lung tumor classification, we obtained the recognition rate of 70.6% for the mES-based method and 64.7% for the skeleton-based method which accounts for 5.9% improvement.

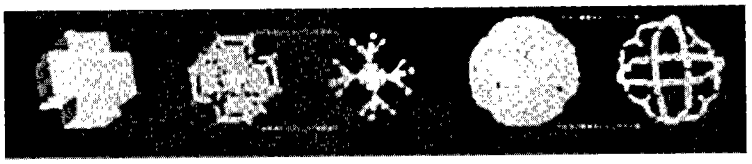
Furthermore, the experiments on lung tumor indicate that shape feature is one of the features that can be used to classify tumor, however, the experimental results reveal that we may need to combine shape with other features to make our classification more accurate since the problem arises when both classes of tumors have similar shapes.

This research work is our first step in 3D shape recognition using the mES. The development of better lung tumor classification is in the progress. Further extensions concerning recognition and classification system based on the mES for 3D gray scale images are our future research projects.

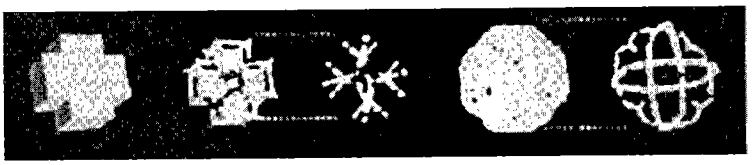
Acknowledgments

The authors would like to express our gratitude to Dr. K. Miyakawa at the National Cancer Center, Japan for providing the volumetric lung data sets, and H. Okada in Shimizu lab for segmenting and constructing 3D lung tumors. This work was supported in part by the Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan and the Grant-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare, Japan

(Manuscript received January 23, 2002)

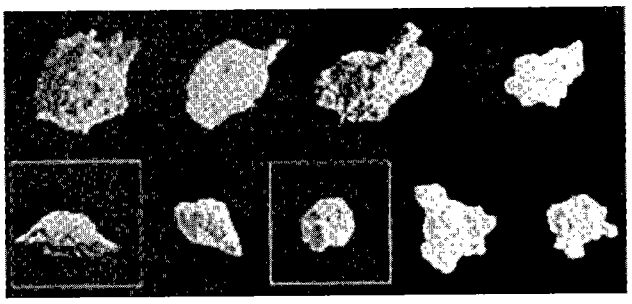


(a) A cross (b) A skeleton (c) A mES

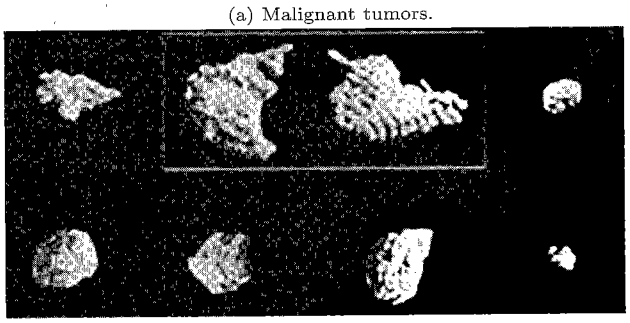


(d) A cross with a cavity (e) A skeleton (f) A mES

Fig. 5. A skeleton and a mES of a cross and a cross with a cavity.



(a) Malignant tumors.



(b) Benign tumors.

Fig. 6. Two clusters of volumetric lung tumors based on their mESs.

References

(1) P.E. Trahanias: "Binary Shape Recognition Using the Morphological Skeleton Transform", *Pattern Recognition*, Vol. 25, pp.1277-1288 (1992)

(2) T. Kaneko, L. Gu, and H. Fujimoto: "Abdominal Organ Recognition using 3D Mathematical Morphology", *ICPR Proceedings*, Vol. 2, pp.263-266 (2000)

(3) C. Min Ma and M. Sonka: "Fully Parallel 3D Thinning Algorithm and its Applications", *Computer Vision and Image Understanding*, Vol. 64, pp. 420-433 (1996)

(4) M. Naf, G. Szekely, R. Kikinis, M.E. Shenton, and O. Kubler: "3D Voronoi Skeletons and Their Usage for the Characterization and Recognition of 3D Organ Shape", *Computer Vision and Image Understanding*, Vol. 66, pp.147-161 (1997)

(5) G. Borgefors, I. Nystrom, and G. Sanniti di Baja: "Computing Skeletons in Three Dimensions", *Pattern Recognition*, Vol.32, pp.1225-1236 (1999)

(6) T.C. Lee, R.L. Kashyap, and C.N. Chu: "Building Skeleton Models via 3-D Medial Surface/Axis Thinning Algorithms", *CVGIP: Graphical Models and Image Processing*, Vol. 56, pp.462-478 (1994)

(7) G. Sanniti di Baja and S. Svensson: "Surface Skeletons Detected on the D6 Distance Transform", *Excerpt from Proceedings of SSPR* (2000)

(8) Y. Zhou and A.W. Toga: "Efficient Skeletonization of Volumetric Objects", *IEEE Trans. Visualization and Computer Graphics*, Vol. 5, pp.196-209 (1999)

(9) R. Lipikorn, A. Shimizu, Y. Hagihara, and H. Kobatake: "A Modified Exoskeleton and Its Applications to Object Representation and Recognition", *IEICE Trans. INF. & SYST.*, Vol. E85-D, No. 5, pp.884-896 (2002)

(10) F.P. Preparata and M.I. Shamos: Computational Geometry, Springer-Verlag, New York (1985)

(11) P. Soille: Morphological Imgae Analysis, Springer (1999)

(12) J. Serra: Image Analysis and Mathematical Morphology, Academic Press, London (1982)

(13) T. Saito and J.I. Toriwaki: "New Algorithm for Euclidean Distance Transformation of an n-Dimensional Digitized Picture with Applications", *Pattern Recognition*, Vol. 27, pp.1551-1565 (1994)

(14) T. Saito, and J. Toriwaki: "Reverse Distance Transformation and Skeletons Based upon the Euclidean Metric for n-Dimensional Digital Binary Pictures", *IEICE Trans. INF. & SYST.* Vol. E77-D, pp.1005-1016 (1994)

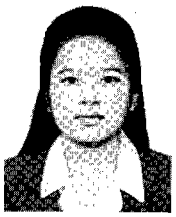
(15) I. Ragnemalm: "The Euclidean Distance Transform in Arbitrary Dimensions", *Pattern Recognition Lett.*, Vol. 14, pp.883-888 (1993)

(16) H. Yamada: "Complete Euclidean Distance Transformation by Parallel Operation", *ICPR Proceedings*, pp.69-71 (1984)

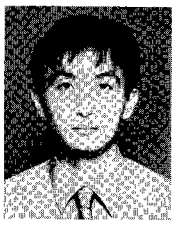
(17) S. Doida, Y. Hagihara, A. Shimizu, R. Lipikorn, and H. Kobatake: "A Fundamental Study on Shape Analysis Based on

Skeleton Transform", *Technical Report of IEICE, PRMU* (2000)
(18) R.O. Duda and P.E. Hart: Pattern Classification and Scene Analysis, John Wiley, New York (1973)
(19) P.M. Griffin and B.L. Deuermeyer: "A Methodology for Pattern Matching of Complex Objects", *Pattern Recognition*, Vol. 23, pp.245-254 (1990)
(20) C.W. Niblack, P.B. Gibbons, and D.W. Capson: "Generating Skeletons and Centerlines from the Distance Transform", *CVGIP: Graphical Models and Image Processing*, Vol. 54, pp.420-437 (1992)
(21) T. Pavlidis: "A Thinning Algorithm for Discrete Binary Images", *Computer Graphics and Image Processing*, Vol. 13, pp.142-157 (1980)

Rajalida Lipikorn (Member) was born in Bangkok, Thailand in December, 1962. She received the B.S. degree in Mathematics and the M.S. degree in Computer Science from California State University, Northridge in 1987, and 1992 respectively. Since 1993, she has been a lecturer at Chulalongkorn University, Thailand. She had been the PhD. student at Graduate School of Bio-Applications and Systems Engineering, Tokyo University of Agriculture and Technology, Japan from April, 1999 till Sept. 2002. She is a member of IEEE.



Akinobu Shimizu (Non-member) was born in October, 1965. He received his B.E. and Ph.D. degrees from Graduate School of Engineering, Nagoya University in 1989 and 1994. He became a research associate at Nagoya University in 1994, and has been an associate professor in the Graduate Shcool of Bio-Applications and Systems Engineering, Tokyo University of Agriculture and Technology since 1998. His research interests include image processing. He is a member of the Japanese Society of Medical Imaging Technology, the Japanese Society for Medical and Biological Engineering, the Japan Society of Computer Aided Diagnosis of Medical Images, and IEEE.



Yoshihiro Hagihara (Non-member) was born in Kanagawa Prefecture, Japan in May, 1964. He received the B.E. and the Ph.D. degrees from Tokyo University of Agriculture and Technology in 1990 and 1996. From 1993 to 1997 he worked as researcher with Systems Development Laboratory, Hitachi, Ltd. In 1997 he joined Tokyo University of Agriculture and Technology. In 2002 He has joined Iwate University as a lecturer. His research interests range from pattern recognition to image processing with industrial applications.

Hidefumi Kobatake (Non-member) was born in November, 1943. He received the B.E., M.E., and Ph.D. degrees from The University of Tokyo, Tokyo, Japan, in 1967, 1969 and 1972, respectively. He is now a Professor at Graduate School of Bio- Applications and Systems Engineering, Tokyo University of Agriculture and Technology, Tokyo, Japan. His research activities are in the areas of speech processing, image processing, and the applications of digital signal processing. He has received several awards, including a 1987 Society of Instrument and Control Engineers' Best Monograph Award and a 1998 Three Dimensional Image Conference's Best Paper Award. He is the member of IEEE, Society of Instrument and Control Engineers of Japan, The Acoustical Society of Japan, etc.