3D MODIFIED EXOSKELETON AND ITS APPLICATIONS FOR SHAPE RECOGNITION

Rajalida LIPIKORN, Akinobu SHIMIZU, Yoshihiro HAGIHARA, and Hidefumi KOBATAKE

The Graduate School of Bio-Applications and Systems Engineering
Tokyo University of Agriculture & Technology
2-24-16 Naka-cho Koganei-shi, Tokyo 184-8588, Japan

ABSTRACT

This paper presents a new 3D object representation called a *modified exoskeleton* (mES) which is another type of skeleton that is more stable and less sensitive to rotation and noise than the conventional skeletons. The mES preserves significant characteristics about an object, that are meaningful for object recognition and reconstruction, as the skeleton does. Then a fast and reliable matching algorithm for recognition and classification system which can identify and classify an observed object is introduced. The main advantage of our matching algorithm is that it can recognize the observed 3D object regardless of its orientation and noise. The first experiment was conducted on a set of artificial 3D objects and the average recognition rate of 92.60% based on the mES is achieved. The second experiment was tested on the classification of volumetric lung tumor data sets and the obtained results are promising.

1. INTRODUCTION

The task of object recognition is essential for biological visual systems and robotics. Due to modern technology, large 3D volumetric data sets become commonly used in these fields. In object recognition, comparing an observed 3D object with known objects (here called *models*) in a database based on measurable characteristics known as *geometry (or shape)* is one of the most widely used techniques. Skeleton is an effective object representation for shape recognition, visualization, and reconstruction [2, 3]. It is, however, very sensitive to subtle shape changes caused by noise. Existence of a small hole or cavity causes extreme difference in skeleton structure. Our proposed mES combines the basic concept of the original exoskeleton introduced by J. Prewitt [5] with the rotation invariant property of a symmetrical object to makes it more stable and less sensitive to rotation and noise disturbances. Once, the observed object and the models are reduced to their compact shape representations, the tasks of the recognition system is to identify a model which corresponds to the observed object in the matching process.

The major contributions of this paper are to propose the 3D mES that alleviates deformations caused by rotation and noise and to introduced an effective recognition system for 3D objects whose matching results are independent of the position and orientation of the observed object. The definition of the mES for 3D objects is described in Section 2. The recognition system is presented in Section 3. Several experiments and their results are discussed in Section 4. The summary and conclusions are presented in Section 5.

2. MODIFIED EXOSKELETON

The mES is analogous to the skeleton, except that we do not extract the skeleton directly from a 3D object. Instead we generate a symmetrical background such as a sphere around the object and then extract the skeleton from this spherical background, and the radius of the sphere can be calculated from

$$rs = \max_{(l,m,n)} \{d((l_g, m_g, n_g), (l, m, n))\} + \varepsilon, \qquad (1)$$

where (l_g, m_g, n_g) is the centroid, (l, m, n) is any voxel on the boundary of the object, $d(.,.)$ is the Euclidean distance between two voxels, and ε is a constant value added to ensure that the sphere covers the original object thoroughly.

Definition Let $X = \{(x, y, z)\}$ be a 3D object in a 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 Figure 1. The mES is defined as a locus of voxels of C, or a skeleton of C, whose distance to any voxel of $\bar{C}$ is the shortest, whereas none of its neighbors has greater distance value. The mES and its exoskeleton function $mesf$ can be expressed as

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

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

where $d((x,y,z),(l,m,n))$ represents the Euclidean distance between voxels (x,y,z) and (l,m,n), and $N(x,y,z)$ is a set of neighbors of (x,y,z). The actual algorithm to extract the skeleton of C is presented in [4].

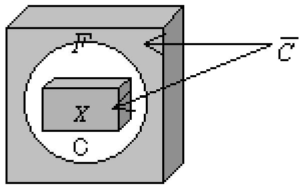
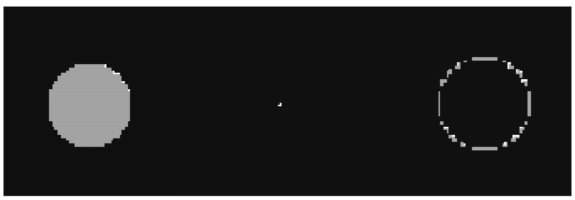
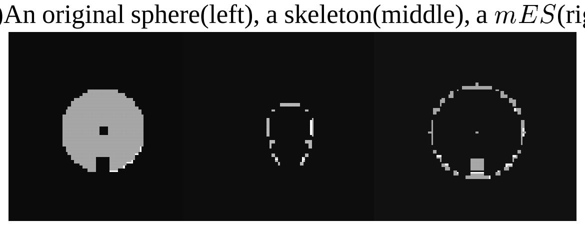


Fig. 1. Illustration of an object embedded in a spherical background (Cross section of 3D image).

The typical example in 3D case in which the skeleton and the mES of an ideal sphere and a sphere with a small hole and a tiny crack are shown in Figure 2. For better visualization, only the central parts of the original object and the deformed object are segmented and displayed.



(a)An original sphere(left), a skeleton(middle), a mES(right).



(b) A deformed sphere(left), a skeleton(middle), a mES(right).

Fig. 2. Central slides of an original sphere and its deformed version (segmented from the 3D objects).

The advantage of the proposed mES is that it is robust against noise disturbances and rotation, and the reasons are as follows: (1) The main body of conventional skeletons concentrate , in general, around the center of gravity of an object where the skeleton function values are large. The mES, on the contrary, disperses around the object (see Figure 2(a)). (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. Not like in the case of the mES, the extent of noise disturbance and/or rotation have merely local effect to parts of the mES nearby the deformed parts and the variation of the $mesf$ is relatively small (see Figure 2(b)).

3. MATCHING ALGORITHM

To ensure scale invariance, a scaling factor ρ is obtained from the ratio of specified lengths which in our case are the longest distances from the boundary voxels to the centroids of the objects. This longest distance is actually equal to the radius of the sphere when ε is equal to zero, and thus, the scaling factor is computed by:

$$\rho = \frac{rs_1}{rs_{2_i}}, \tag{2}$$

where rs_1 denotes the radius of the spherical background of the object and rs_{2_i} denotes that of the model i.

In the case of the mES, the matching algorithm, extended from 2D version in [1], searches for the local similitudes between the two mESs by matching each voxel in the object mES with its nearest one in the model mES. In this process, the $mesf$ is also taken into consideration as a weight function. When matching voxels between 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 distance between voxels and the difference between their $mesf$s are accumulated and used to computed the similarity measure E_1 with respect to the object. Then, each voxel in the model mES is matched with its nearest one in the object mES, and the similarity measure E_2 with respect to the model i is computed. The similarity measure E, which provides a mean for determining the overall similarity, can be calculated from the following equations:

$$E = \min_a \{E_1(a) + E_{2_i}(a)\} \tag{3}$$

$$E_1(a) = \{\sum_u \sum_v \sum_w sd_1(u,v,w)\}/M_1 \tag{4}$$

$$E_{2_i}(a) = \{\sum_x \sum_y \sum_z sd_2(x,y,z)\}/M_{2_i} \tag{5}$$

$$sd_1 = \sqrt{\begin{array}{l}(x'-au)^2+(y'-av)^2+(z'-aw)^2+\\ \{mesf_{2_i}(x',y',z')-a(mesf_1(u,v,w))\}^2\end{array}}$$

$$sd_2 = \sqrt{\begin{array}{l}(x-au')^2+(y-av')^2+(z-aw')^2+\\ \{mesf_{2_i}(x,y,z)-a(mesf_1(u',v',w'))\}^2\end{array}}$$

where (x,y,z), (x',y',z'), (u,v,w), (u',v',w') are voxels in the two mESs, a is a scaling range within $\rho * 1.1^{-15} \leq a \leq \rho * 1.1^{15}$, and M_1, M_{2_i} are the numbers of voxels in mES_1 and mES_{2_i}, respectively. This matching algorithm then recognizes an object as a model whose similarity measure E is minimal.

4. EXPERIMENTAL RESULTS

Our experiments were conducted on a set of several 3D artificial binary objects in the image of size 80 x 80 x 60 and a set of nine volumetric lung tumor data sets segmented from CT images: two benign and seven malignant. The CT data have a resolution of 512 x 512, where the interval between voxels is 0.3-0.4 mm. Example of artificial objects and tumor data are shown in Figure 3, and Figure 4.

The robustness of the mES was compared against that of the skeleton using our matching algorithm. The first experiment on matching between different shapes was performed on a set of artificial objects. From the experimental results, it can be verified that most of the matching results give almost the same order of similarity. In the next session, we evaluate each performance quantitatively.

Our main concern is to show how efficient the mES is when the observed object is distorted by noise or rotation. Thus, in the second experiment the object was rotated in every 5^{o} steps around the three coordinate axes in the range $-90^{o} \leq \theta_x, \theta_y, \theta_z \leq 90^{o}$, thus 36 rotated versions were produced. The effect of rotation on the mES and the similarity E was evaluated using the relative dispersions or coefficient of variations (CV).

The coefficient of variation expresses the variation of the throughput as a percentage of the mean and can be calculated as follows:

$$CV = \left(\frac{STDEV}{mean}\right) * 100 = \left(\frac{\sigma}{\mu}\right) * 100. \qquad (6)$$

where μ and σ are the mean and the standard deviation of the similarity values E between each object and its rotated versions, respectively. The experimental results in Table 1 indicate that most of the relative dispersions of the exoskeleton-based method is lower than those of the skeleton-based method which means that the mES is more stable with respect to rotation.

Besides, the relative distance is also used to evaluate the effectiveness of the mES. The relative distance can be calculated as follows:

$$Relative \;\; distance(RD) = |\mu_1 - \mu_2|/\sigma, \qquad (7)$$

$$\sigma = (\sigma_1 + \sigma_2)/2$$

where μ_1, σ_1 are the mean and the standard deviation of the similarity measures between rotated versions and other models in the database, μ_2, σ_2 are the mean and the standard deviation of the similarity measures between rotated versions and their original objects. The relative distance of good representation should be large. The relative distance of each objects is shown in Table 1. It shows that most of the relative distances derived from the exoskeleton-based method are larger than those derived from the skeleton-based method which indicates that the mES of each rotated version is more similar to its original than other models, and these statistical results are also confirmed by the recognition rates shown in Table 2.

Another problem with 3D skeletons arises when the observed object contains cavities or tunnels. The skeleton of an object with internal noise, such as cavities or tunnels, could differ completely from the original skeleton. In the last experiment on the effect of noise, each object was distorted by two kinds of noise, namely internal noise and boundary noise. For internal noise, noise was added to an object in arbitrary positions, and 15 noisy versions were generated for each object. Next boundary noise was added to an object by randomly converting the boundary voxels of the object to background voxels and 15 deformed versions were generated for each object. Each noisy object was then compared with the models. The experimental results in Table 2 indicate that the mES gives better recognition rate when dealing with noise disturbance.

In the preliminary stage, we also apply our method to classifiy real volumetric images of lung tumors segmented from CT images. In our experiment, a tumor is classified as either benign or malignant by the simialrity measure on their mESs. The initial prototypes of these two classes are selected from a pair of the tumors whose similarity measure is maximum, then each tumor is classified by using the k-clustering algorithm with the sum-of-squares of the similarity measure as the cost function and $k = 1$. To evaluate the classification results obtained from using mES, we compared them with those obtained from the skeletons, and the results indicate one misclassification for the mES-based method and two missclassifications for the skeleton-based method.

5. CONCLUSION

A binary object representation called the *modified exoskeleton* mES and a matching algorithms for 3D shape recognition are proposed in this paper. The main features of the mES is the integration of the original exoskeleton concept with the rotation invariant property of symmetrical object and the skeleton function to alleviate distortation problem due to rotation and noise. The average recognition rate of the mES-based method for artificial objects is 92.6% whereas that of the skeleton-based method is 84.7% which indicates an improvement of 7.90%. For real images, the similarity measure based on the mES is used to classify the lung tumors. We obtained the recognition rate of 89% for the mES-based method compared to 78% for the skeleton-based method, and the advantage of using the mES in classifying the tumor is that it is less susceptible to noise and the orientation of the tumor than the skeleton.

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.

6. REFERENCES

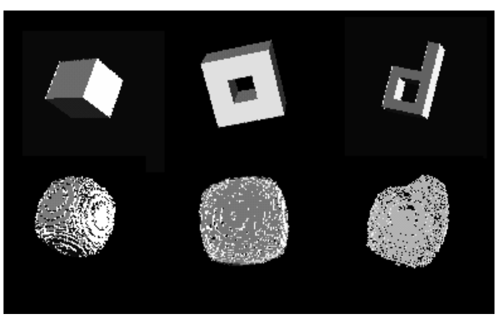
[1] S. Doida, Y. Hagihara, A. Shimizu, R. Lipikorn, and H. Kobatake, "A Fundamental Study on Shape Analysis Based on Skeleton Transform," *Technical Report of IEICE, PRMU99-245,2000-2* (in Japanese).

[2] T. Kaneko, L. Gu, and H. Fujimoto, "Abdominal Organ Recognition using 3D Mathematical Morphology," *Proc. of ICPR*, Vol. 2, pp. 263-266, 2000.

[3] 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, No. 2, pp. 147-161, 1997.

[4] 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, No. 9, pp. 1005 - 1016, 1994.

[5] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982.



(a) A cube (b) A letter 'o' (c) A letter 'd'

Fig. 3. Example 3D binary objects in a database

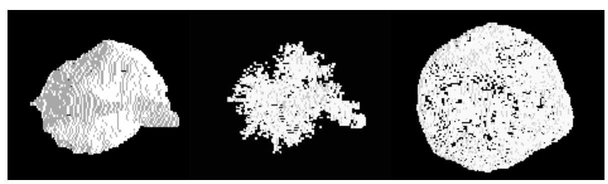
Table 1. Variation of matching results and relative distance between each object and its rotated versions based on the skeleton and the mES.

	skeleton		mES (ϵ = 2)	
Object	CV	RD	CV	RD
Box	65.92	1.28	45.14	2.01
Ring	88.42	0.93	77.97	1.70
d	89.91	0.96	76.35	2.47
Cylinder	81.46	1.37	73.29	1.82
Sphere	68.33	3.98	54.89	1.89
Cross	75.76	0.90	75.28	1.49
E	78.59	1.21	74.00	1.93
U	79.17	0.94	65.44	2.05
Ribbon	83.12	0.99	81.63	2.97
average	*78.96*	*1.40*	*69.33*	*2.04*

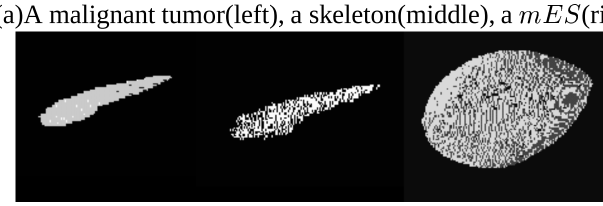
Table 2. Recognition rates(%) between the skeleton-based and the exoskeleton-based (ϵ = 2) methods on rotation and noise disturbances.

	Rotated version		*Noisy version*		*Total*	
Object	S	mES	S	mES	S	mES
Box	0	0	73.3	80.0	87.9	90.9
Ring	63.9	69.4	93.3	90.0	77.3	78.8
d	72.2	97.2	96.7	76.7	83.3	87.9
Cylinder	0	0	0	0	0	0
Sphere	0	0	6.7	0	57.6	0
Cross	91.7	88.9	93.3	96.7	92.4	92.4
E	94.4	97.2	0	0	97.0	98.5
U	75.0	0	96.7	0	84.8	0
Ribbon	83.3	86.1	83.3	83.3	81.8	84.8
average	86.7	93.2	82.6	91.9	84.7	92.6

(Note: S = Skeleton)



(a)A malignant tumor(left), a skeleton(middle), a mES(right).



(b) A benign tumor(left), a skeleton(middle), a mES(right).

Fig. 4. The original lung tumors with their skeletons and mESs.