A study on object recognition using morphological shape decomposition

Chang-Sun Ahn*, Kyoung-Bae Eum**
* Janghang Technical High School
** Dept. of Computer Science, Kunsan Univ.

ABSTRACT

Mathematical morphology based on set theory has been applied to various areas in image processing. Pitas proposed a object recognition algorithm using Morphological Shape Decomposition(MSD), and a new representation scheme called Morphological Shape Representation(MSR). The Pitas's algorithm is a simple and adequate approach to recognize objects that are rotated 45 degree-units with respect to the model object. However, this recognition scheme fails in case of random rotation. This disadvantage may be compensated by defining small angle increments. However, this solution may greatly increase computational complexity because the smaller the step makes more number of rotations to be necessary.

In this paper, we propose a new method for object recognition based on MSD. The first step of our method decomposes a binary shape into a union of simple binary shapes, and then a new tree structure is constructed which can represent the relations of binary shapes in an object. Finally, we obtain the feature informations invariant to the rotation, translation, and scaling from the tree and calculate matching scores using efficient matching measure.

Because our method does not need to rotate the object to be tested, it could be more efficient than Pitas's one. MSR has an intricate structure so that it might be difficult to calculate matching scores even for a little complex object. But our tree has simpler structure than MSR, and easier to calculated the matchng score. We experimented 20 test images scaled, rotated, and translated versions of five kinds of automobile images. The simulation result using octagonal structure elements shows 95% correct recognition rate. The experimental results using approximated circular structure elements are examined. Also, the effect of noise on MSR scheme is considered.

1. INTRODUCTION

An ultimate goal of object recognition is to provides a computer the recognition ability that is similar to human being. There are various methods in object recognition. Model based method is the technique that analyzes, and identifies objects by searching the data base that keeps the information of objects. The model based object recognition algorithm is realized by two steps, training and recognition steps. In the training step, a data base is made and in the recognition step unknown object is identified by searching the data base. There are several tasks to realize an efficient model based vision system. Those are (1) what kinds of information to be captured from the image and how to construct a model, and (2) how to seach the data base or match a unknown object to the stored model.

There are two kinds of information that are used in the vision system, global and local information.[13] The shape description method using global information is divided into the external shape description based on the shape contour and the internal shape description based on shape area or shape structure. The former ones include the description methods such as Fourier transform and B-spline. The latter ones are the description methods including quad-trees, skeleton and shape decomposition algorithms.[4-5] It is impossible to recognize partially occluded and overlapped objects in the methods using global information, and is sensitive to noise. In order to handle these problems managable, the objects are devided into small partitions and the information of small partitions is used to recognize objects in the local information based vision.[14] In these method using local information, the problem of inexact boundary segmentation may occur.

Mathematical morphology was proposed by Matheron and Serra, and had developed by Maragos and Doughterty. Mathematical morphology was made use of shape analysis, feature extraction, recognition, image coding, and nonlinear filtering efficiently.[8-11] Mathematical morphology in image processing is performed on the image feature in visible spatial domain rather than frequency domain. So, the transfer of physical meaning is direct. Also the morphological basic operations based on set theory is easy to be implemeted by the logic circuits, and it is possible to process in parallel.[1-3] In particular, morphological operations can simplify image data preserving their essential shape characteristics and eliminate irrelevancies, and can easily extract geometrical information of image. So it might be useful in the description and representation of image.

MSD has been proposed by Pitas. There are three decomposition methods depending on the construction of primitive elements.[7] These are MSD with maximal inscribable elements, minimum enclosing elements and minimum error elements. In the method using maximal inscribable element the largest homothetic element is formed that can be completely included witin the object. In the method using minimum enclosing element, the object is made up of an union of the smallest homothetics that can completely enclose the remainder of the object. The method using minimum error element is designed to minimize the total representation error in which elements may be overlapped. The set of centers for the minimum error element is selected so that the number of points in the set is as small as possible. Pitas proposed object recognition algorithm based on MSD with maximal inscibable elements. This scheme constructs database according to the decomposition of model objects and compare the first decomposed elements of the unknown object with the first decomposed elements of model objects. When the size of structuring element is not equal, it is adjusted so that it is equal. The mass center of each cluster is found and is translated so that the mass center of model image is equal to that of unknown object. After that, it is determined by area differences whether the model object coincide with the unknown object which is rotated by 45-degree units. Repeted calls of this procedure with various model objects are used in order to identify a given object.[6]

In this paper, we propose a new method for object recognition based on MSD. In the proposed method a new tree is constructed which can represent the relations of simple binary shapes in an object. Also, we propose a new shape matching algorithm based on the feature which is obtained from proposed tree. The matching in Pitas's algorithm is determined by area differences between the model object and a rotated target object by 45-degree units. Pitas's algorithm is simple and adequate to the problem of recognizing objects which may be rotated by multiples of 45 degrees with respect to the model objects. However, this recognition scheme fails in case of random degrees of rotations. This disadvantage can be compensated by defining a small angle increments. However this solution may increase computational complexity because the smaller the step prodeuces the more comparision.

Our method does not need to rotate the object to be tested so that it could take less time than Pitas's one. Our method calculates the matching scores based on the information invariant to translation and scaling and rotation. But, we did not get perfect results in the rotated objects in the experiment. The reason is caused by the nature of the discrete grid, which does not support rotation at arbitrary angles and by the fact that there is no a complete disk in discrete grid. MSR proposed by Pitas has an intricate structure so that it might be difficult to calculate matching scores even for a little complex object. But our tree has simpler than MSR to calculate the matching scores.

The organization of this paper is as follow. In section II, the MSD, new tree structure and the corresponding matching algorithm are proposed. The simulation results of 20 test images are shown and examined in section III. In section IV, we made the conclusion.

2. PROPOSED OBJECT RECOGNITION METHOD

2.1 The oprations of mathematical morphology

In mathematical morphology, shape transfor- mation can be represented by set operations. The primary morphological operations are dilation, erosion, opening, and closing.[1-3] Those morphological operations are defined as follows.

(1) Dilation

$A \oplus B = \bigcup_{b \in B} A_b = \{ x : x = a + b \ \ where \ \ a \in A \ and \ b \in B \}$

(2) Erosion

$A \ominus B = \bigcap_{b \in B} A_b = \{ x : -B + x \subset A \}$

(3) Opening

$A \cdot B = (A \ominus B) \oplus B$

(4) Closing

$A \cdot B = (A \oplus B) \ominus B$

In eq. (1)-(4), A is the image and B is the structuring element..

In this paper we assme that the structuring elements are symmetric about the origin. Geometrically, the dilation has the effect of expanding an image, the erosion has the effect of shrinking an image, an opening smooths sharp positive edges and eliminates small isolated points, and the closing operation smooths sharp negative edges and fills the gaps of an image.

2.2 MSD(Morphological Shape Decomposition)

MSD, the one of internal description methods of an object is proposed by Pitas. There are three decomposition methods depending on the construction of primitive elements. These are MSD with maximal inscribable elements, minimum enclosing elements and minimum error elements. In the method using maximal inscribable element the largest homothetic element is formed that can be completely included witin the object.[7]

MSD using maximal inscribable primitive element is assumed in this paper. In this case, a MSD can be treated as a process finding the skeleton. The problem arises in the shape decomposition because one assumes that objects are defined on R^n. In most practical case we work objects in digitized images. Therefore, we have to define the MSD over the Euclidean grid Z^n. Let a subset X of Z^n represent a discrete binary image. We assume that X is nonempty and bounded. One can assume structuring elements as shown in Fig. 1. However, MSD using square and rhombus structuring element is not invariant to rotation. Of course, even though the one using an octagonal and approximated circle is not invariant to rotation, the former one causes worse effect than the latter one in the decomposition of rotated images.

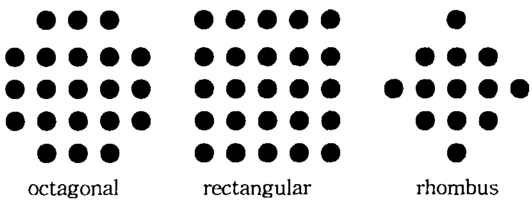


Fig. 1. Structure elements in discrete space Z^2

MSD decompose an object X into an union of simple subsets $X_1, X_2, \ldots, X_{n-1}, X_n$. The number n and sum of these subsets are finite for a bounded set X. The discrete notion of a simple object is given

(5) $X_i = n_iB = B \oplus B \oplus \ldots \oplus B$

where B is a discrete structuring element of size one and n_i is an integer. n_iB is a simple element of size n_i. A more general definition of a simple discrete shape is given

(6) $X_i = L_i \oplus n_iB$

where L_i is a subset of Z^n of zero thickness and it consists of lines or isolated points.

Morphological decomposition is performed in the following way. The first set of maximal inscribable disks in the object, that has the maximum radius, is found. The first set of the decomposition is subtracted from the object. Then the second set of the maximal inscribable disks in the remaining of the object, that have the maximum radius, is found. The procedure is repeated until the remainder of object become an empty set. The procedure can be performed recurrsively as follows.

(7) $X_i = ((X - X'_{i-1}) \ominus n_i B^s) \oplus n_i B$,

where $X'_i = \bigcup_{0<j\leq i} X_j$ and $X'_0 = \emptyset$,

until $(X - X'_k) \ominus B^s = \emptyset$

In eq. (7), n_i denote the maximal size of the inscribable object $n_i B$ in the set $X - X'_{i-1}$.

Another formulation of the discrete morphological decomposition can be possible as follows.

(8) $L_i = (X - \bigcup_{0<j\leq i-1} (L_j \oplus n_j B)) \ominus n_i B^s$

where $L_0 = \emptyset$ and $L'_i = \bigcup_{0<j\leq i} L_j$

until $L_{k+1} = \emptyset$

In eq. (8), L_i is the locus of centers of the maximal inscribable disks $n_i B$ in object $X - X'_{i-1}$. An example is given in Fig. 2. In the figure the ratangle is decomposed by using 5*5 circular structuring element.

The discrete morphological decomposition has the following properties.[4]

(1) X'_k is bounded by X. In cetain cases $X - X'_k \neq \phi$.

(2) The object X_i are simple.

(3) It is unique, scale and translation invariant, antiextensive, and idempoetent.

MSD is invariant to rotation in continuous space. But, the rotation invariance is no longer valid in discrete grid. There are two reasons for this. The first one is due to the very nature of the discrete grid Z^n, which does not support rotation at arbitrary angles θ. The second problems is that there is no exact equivalent of a disk in Z^n.

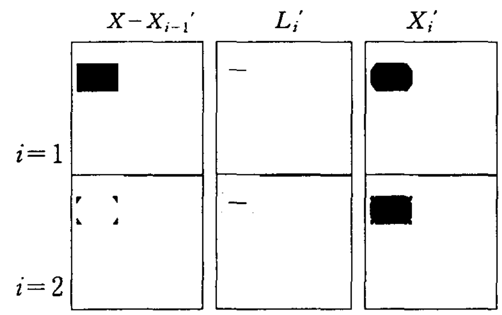


Fig. 2. Shape decomposition of rectangle image

2.3 The representation of tree structure

MSR proposed by Pitas is a scheme to combine the CSG(Constructive Solid Geometry) and MSD. MSR is not unique. However, if certain restrictions are posed, special cases of the MSR can become unique.[5,12] Its nonterminal nodes represent set operations or morphological operations or rigid solid motions or scaling. The terminal nodes represent either points or lines or structuring elements or rigid motion arguments or scaling factors.

MSR has an intricate structure so that it might be difficult to represent an object even for a little complex object. But, our tree has more simple structure than MSR and is unique under certain restrictions. The relations of simple binary shapes obtained after decomposing an object shows how the body of object is related to the remaining decomposed components.

The generation procedure of our tree which represents the relations of primitive decomposed objects is as follows. In the first step, the line set L_1(or mass center) is generated by MSD and the body (X_1) of object is produced by dilating L_1. In the second step, it is examined whether the line sets (or mass center) generated in next step is included the subimage produced by MSD in previous step. If the line sets is included in the subimage, these are inserted into the child of previous node. Otherwise, it is inserted under the finding node after it is compared to the previous steps until searching for the subimage to be involved. In this step, the seperating subimage including maximal inscribable disk is inserted into the comparing list. This comparing list is referened after this step. The second step is repeated until the remainder of the object is an empty set.

The tree structure generated by this method has the shape that the decomposed elements are arranged around the node for a body of the object. But the decomposition levels(the order of decomposition in the procedure) are not always equal to the tree levels. The decomposition levels can be used as a weighting parameter for calculating matching score. A node of our tree is shown in Fig. 3.

Tree level	Decompo-sition level	Line set (or mass center)	Shape of structuring element	Size of structuring element	Angle

Fig.3. A node of tree

In Fig. 3, the relative angle is calculated by using the mass center of child nodes assuming the mass center of parent node is at the origin. The set of approximated circles or the set of octagons can be used as structuring elements. The type and size of structuring element and the relative angle are possible to be used as the feature of the object recognition. The tree of Bongo(a brand name of Korean car) is described by the tree in Fig 4. In this figure, the only tree level and decomposition levle are represented on account of the limited space.

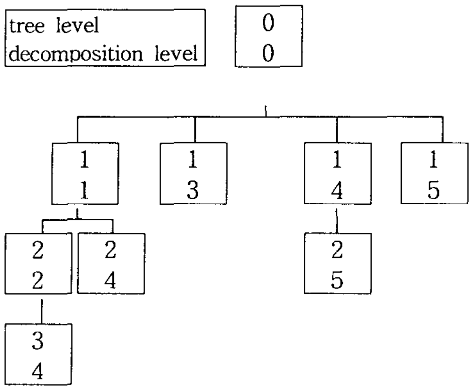


Fig.4. Tree representation of Bongo image

2.4 Matching algorithm

In order to perform matching, we construct the database using feature information of model objects. One can descriminate a unknown object by matching to the stored model objects. The proposed algorithm is as follows.

Step 1. Decompose the input object and obtain the corresponding tree in which feature information is included at each node.

Step 2. Compare the tree of the input object with that of the model objects from the root node. If the matching score of a model is sufficiently small at a node, then remove the model in the next comparision. Other- wise repeat this step. In this step, the matching score is calculated by

$$MS_j = \left\{ \frac{1}{1+k|N_t - N_r|} + \frac{1}{1+d({}_n\theta_t, {}_n\theta_r)/k} \right\} \frac{1}{1.5^i} \quad (10)$$

where i is the shape decomposition level, j is the tree level, and k is a constant. In the simulation we set k = 5. In eq. (10), N_r and N_t represent the number of children nodes of a model and input object at the current nodes to be compared, respectively. In eq. (10),

$$d({}_n\theta_t, {}_n\theta_r) = \min \{ |\theta_{1t} - \theta_{1r}| + |\theta_{2t} - \theta_{2r}| + \ldots. + |\theta_{nt} - \theta_{nr}| \} \quad (11)$$

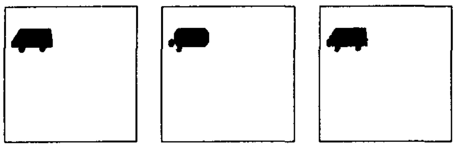
for all combined pairs of children nodes one from the model object and the other from input object, where θ is the relative angle to the parent node.

Step 3. Find the model that has the maximum matching score among the survived model in Step 2.

The rationale behind Step 2. is that the model that provides the smaller matching score at the higher level of tree is removed in the subsquent matching to save the unnecessay matching time. Note that the matching score becomes large when the difference of the number of children nodes between input and a model object. Also, note that the relative angles to the parent node in a model is matched to those in input object and the corresponding matching score gets large value when the angle difference becomes small. In the angle comparision, the number of children in a model does not coinside with that of input object. In that case, we permute the angle of children in such a way to get the minimun distance in eq. (11). Also as we see in eq. (10), the maching score is weighted in such a way that the primitive components that larger area(the higher decomposition level) havily contributes the matching score.

3. RESULTS OF SIMULATION

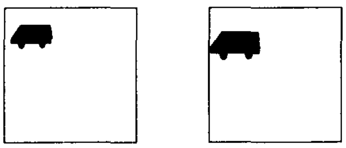
The SDT-200 work staton was used in the experiment. The images that are used in the simulation are the 256*256 binary images of 5 kinds of Korean-brand automobiles. ; Bongo, Stellar, Korando, Pride-beta, Truck. Each automobile image is rotated, translated, and scaled to make 20 images. The original image of Bonge is shown in Fig. 5 (a). Fig. 5 (b) shows the shape decomposed image up to decomposition level 3 using octagonal structuring elements. Fig. 5(c) shows the same image decomposed up to level 6.



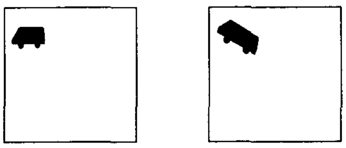
(a) original image (b) decomposed image up to level 3 (c) decomposed image up to level 6

Fig.5. Original image and shape decomposed image of Bongo

After 5 original images are decomposed by using octagonal structuring elements (or the approximated disks), we generate the trees by using the proposed method in Section 2.3. Each node of the generated trees includes the feature informations. The matchings between the generated tree of the original object and that of each scaled, translated, and rotated objects are performed according to the maching algorithm in Section 2.4. In Fig. 6, the rotated, scaled, and translated images of Bongo are represented.



(a) translated image (b) large scaled image



(c) small scaled image (d) rotated image

Fig.6. Transformed images of Bongo

Model image \ Test image	Translated Bongo	Large scaled Bongo	Small scaled Bongo	Rotated Bongo	Translated Stellar
Bongo	2.7709	2.4636	2.0631	0.8492	0.3823
Stellar	0.4099	0.4281	0.3108	0.2731	2.8148
Korando	0.9983	1.1309	0.9641	0.3706	0.8829
Pride	0.7468	0.9830	0.6161	0.3157	1.4700
Truck	0.8982	0.9647	0.8273	0.1971	0.6256

Model image \ Test image	Large scaled Stellar	Small scaled Stellar	Rotated Stellar	Translated Korando	Large scaled Korando
Bongo	0.3967	0.3955	0.2321	0.9003	0.2330
Stellar	1.9259	2.2469	1.1107	0.8829	0.4795
Korando	0.5099	0.8846	0.5258	3.4074	1.8653
Pride	1.8177	1.4938	1.4785	0.5454	0.7832
Truck	0.6254	0.6099	0.5546	1.3109	0.1845

Model image \ Test image	Small scaled Korando	Rotated Korando	Translated Pride	Large scaled Pride	Small scaled Pride
Bongo	0.5450	0.9362	0.7153	0.7164	0.6427
Stellar	0.8437	0.2521	1.4700	0.9046	1.4601
Korando	0.9450	1.2230	0.5454	0.5569	0.5503
Pride	0.8218	0.6396	2.8148	1.4740	2.4444
Truck	0.2367	0.8644	0.6258	0.8015	0.6258

Model image \ Test image	Rotated Pride	Translated Truck	Large scaled Truck	Small scaled Truck	Rotated Truck
Bongo	0.3927	0.8036	0.6917	1.0498	0.6287
Stellar	0.1867	0.4743	0.5913	0.5158	0.5388
Korando	0.1508	1.2258	0.5288	1.1879	0.4903
Pride	0.5515	0.4744	0.6545	0.5545	0.5728
Truck	0.3439	2.9465	1.6666	1.7777	0.7187

Fig.7. Matching scores of input object and model object

In Fig. 7, the matching scores are tabulated in case that the octagons are used as structuring elements. We calculated the matching score up to tree level 3 because a good recognition rate is obtained even though it is not calculated up to the bottom level. We obtained 95 percentages of correct recognition ratio. The perfect recognition was obtained in the translated and scaled objects. However, the rotated Stellar image was not correctly recognized. There are several reasons for the incorrect recogniton. The first one is the very nature of the discrete grid Z^n, which does not support rotation at arbitrary angles θ. The second one is that threre is no exact equivalent of a disk in Z^n.

Two scaled images were incorrectly recognized when we use a set of approximated circles as structuring elements. In rotation, we could get better result than the one using an octagonal structuring element, because an approximated circle becomes more close to an circle than an octagonal. For scaled objects the result was worse than the one using octagonal structuring elements. This reason is that it is hard to obtain the set of discrete approximated circles, in which a circle is just the scaled scaled version of the other differently from the set of octagons.

The recognition rate was not good for the noisy image without preprocessing. In this experiment, each original images was corrupted in such a way that 20 percentages of the boundary was dislocaed in a small distance. After preprocessing the noisy image by opening-closing filter, three rotated images and two scaled images were incorrectly recognized. One can realize that MSD using maximal inscribable element has a prominent flaw when the noise make the boundary intruding an object.

5. REFERENCES

1. J. Serra, "Image Analysis and Mathematical Morphology," Academic Press, CA, 1982.
2. Charles R.Giardina and Edward R. Dougherty, "Morphological Methods in Image and Signal Processing," Prentice Hall, 1987.
3. R. M. Haralick, S. R. Sternberg and X. Zhuang, "Image analysis using mathematical morphololgy," IEEE Trans. Pattern Anal. Mach. Intelligence PAMI-9, 4. pp. 532-551, July 1987.
4. I. Pitas and A. N. Venetsanopoulos, "Morphological shape decomposition," IEEE Trans. Pattern Anal. Mach. Intelligence PAMI-12, 1, pp. 38-46, Jan. 1990.
5. I. Pitas and A. N. Venetsanopoulos, "Morphological shape representation," Pattern Recognition, Vol.25, No.6, pp.555-565, 1992.
6. I. Pitas and N. D. Sidiropoulos, "Pattern recognition of binary image objects using morphological shape decomposition," CVIP, pp. 279-305, 1992
7. J. M. Reinhardt, W. E. Higgins, "Toward efficent morphological shape representation," ICASSP-93, Vol. 5, pp. 125-128, 1993.
8. P. Maragos and R. W. Schafer, "Morphological skeleton representation and coding of binary images," IEEE Trans. Acoust. Speech Signal Proc. ASSP-34. 5, pp. 1228-1244, Oct. 1986.
9. X. Zhuang and R. M. Haralick, "Morphological structuring element decomposition", Comput. Vision Graphics Image Process. 35, pp. 370-382, 1986.
10. I. Pitas and A. N. Venetsanopoulos, "Nonlinear Digital Filters: Principles and Applications," Kluwer Academic, Dordrecht, 1990.
11. Y. Zhao and R. M. Haralick, "Binary shape recognition based on automatic morphological shape decomposition," in Proceedings, IEEE International Conference on Acoustics Speech and Signal Processing, Glasgow, pp. 1691-1694, 1989.
12. D. Hearn and M. P. Baker, "Computer Graphics," Prentice-Hall, 1986.
13. G. J. Ettinger, " Large hierachical object recognition using libraries of parameterized model sub-parts," in Proc. IEEE Comput. Vision Patt. Recogn. (AnnArbor,MI), June 1988.
14. W. E. L. Grimson," Object recognition by computer - the role of geometric constraints", Cambridge, MA : MIT Press, 1990