附件一

行政院國家科學委員會補助專題研究計畫 ■ 成 果 報 告 □期中進度報告

（計畫名稱）

計畫類別：■ 個別型計畫　　□ 整合型計畫
計畫編號：NSC 91-2213-E-032-015
執行期間：　91 年　8 月　1 日至　92 年　7 月 31 日

計畫主持人：林慧珍
共同主持人：顏淑惠
計畫參與人員：　高楊達

成果報告類型(依經費核定清單規定繳交)：■精簡報告　□完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：

中　華　民　國　92 年　10 月　1 日

二、中英文摘要及關鍵詞：

關鍵詞：形狀、空間位置、特徵抽取、影像分析、相似度測量、影像分割

中文摘要

隨著網際網路迅速成長，以往在網頁上呈現資料的方式已無法滿足使用者的需求下，多媒體資料的呈現已是必然的趨勢，多媒體所能呈現的包括影像、聲音、Video 以及虛擬實境等等。就影像來說，如何讓使用者在龐大的影像資料庫中快速地有效查詢出自己想要的影像，已成為當務之急。因此，本計劃將研究並提出一個能根據影像內容特徵做搜尋之影像查詢系統，並希望此系統能將查詢影像減少在一可容許範圍，如此不但能減少網路負載量，亦能快速的讓使用者得到想要的影像。

Keywords： shape、spatial、structure、pattern recognition、centroid、features

ABSTRACT

Because of recent advances in computer technology and the revolution in the way information is processed, increasing interest has developed in automatic information retrieval from huge databases. In particular, content-based retrieval has become a hot research topic and, consequently, improving the technology for content-based querying has become an important challenge.

There are four basic components for content-based retrieval from images: color, texture, shape, and spatial relationship. In this research project, we will focus our attention on both shape and spatial-based retrieval from images as described in the following: (1). Shape features: A variety of shape features have been exploited, such as area, perimeter, boundary segments, points of maximal curvature change,…, etc.. All of these features have proved to have advantages and disadvantages. In this project, we will make an attempt to collect and overview some presented shape features, exploit some new features, and then assess their effectiveness. As a result, we will present an approach for shape retrieval based on a combination of the most effective and suitable image features. (2). Spatial features: Graphs are used as models in a variety of areas. When used to represent spatial structures, graph representation allows the system to efficiently perform spatial similarity measurements and indexing. In this research, the spatial relationship between objects in an image is modeled using a simple (directed) graph. This graph consists of vertices which represent object centroids and a specific vertex, called an origin, which represents the centroid of the object centroids or, simply, just the centroid of the entire image, and (directed) edges from the origin to the object centroids. Each edge may be assigned a distance (between the origin and the centroid of the indicated object). Each vertex may be associated with important characteristics (such as location, orientation, area, shape, or color) of the corresponding object.

After the shape features are determined, graph models for the spatial representation are constructed, and the best similarity measurements are determined, different types of features

(shape, spatial, color, and texture) associated with the weight values, which represent the levels of relative importance of different types of features received from the user query are then combined to retrieve the desired images.

Our approach attempts to achieve the following objectives:

1. It is able to handle images distorted by different viewpoint translation.
2. The employed features are easy to compute.
3. The features require little storage.
4. The feature representation and measure enables to support efficient indexing.

三、內容：

Introduction and Background

Because of recent advances in computer technology and the revolution in the way information is processed, increasing interest has developed in automatic information retrieval from huge databases. In particular, content-based retrieval has become a hot research topic and, consequently, improving the technology for content-based querying systems is of more challenge. [11]

There are four basic feature components for content-based image retrieval from images: color, texture, shape, and spatial relationship. Among these features, shape contains most attractive visual information for human perception. An important step before shape extraction is edge point detection. Many edge detection methods have been proposed [8-10][15]. Z. Hou [8] uses a one-way design detector to determine an edge point existence. J. Weickert [10] uses a partial differential and morphology method to formulate an optimization problem. M. Kass [18] has proposed an active contour model to surround the contour points. They have succeeded to overcome the multiple objects in an image problems. Most of these research results show that they are robust but ineffective, because of their long processing time.

Another important issue next to edge detection is shape representation. A good shape representation requires invariance to translation, rotation and scale. F. Mahmoudi [11] uses an edge orientation autocorrelogram to represent an object. This approach offers more effectiveness if the edge point detection is not perfect. In T. Bernier [2] a distance and angle of the contour points relative to center is invariant to transformation as well as translation, rotation and scale. The representation method presented here is boundary based, produces a serial numerical sequence to preserve the shape information; however, an incompletely edge detection gets a bad result. J. Zhang [7] uses a shape spaces method to deal with the noise and occlusion problems. H. Nishida [3] has proposed a new representation to tolerate the deformed contour points. Both methods in [3] and [7] provide effective representations in practice.

參考文獻

1. M. S. Chor, W. Y. Kim, A novel two stage template matching method for rotation and illumination invariance. Pattern Recognition 35(2002) 119-129.
2. T. Bernier, J.A. Landry, A new method for representing and matching shapes of natural objects. Pattern Recognition 36(2003) 1711-1723.
3. H. Nishida, Structural feature indexing for retrieval of partially visible shapes. Pattern Recognition 35(2002) 55-67.
4. K. K. Yu, N. C. Hu, Two-Dimensional Gray-Level object recognition using shape-specific points. Journal of the Chinese Institute of Engineers, Vol. 24, No. 2 (2001) 245-252.
5. J. H. Chang, K. C. Fan, Y. L. Chang, Multi-modal gray-level histogram modeling and decomposition. Image and Vision Computing 20 (2002) 203-216.
6. H. Wei, D. Y. Y. Yun, Illumination-Invariant Image indexing using Directional gradient angular histogram. Proceedings of the IASTED International Conference Computer Graphics and Imaging, August, 2001 13-16, Honolulu, Hawaii, USA.
7. J. Zhang, X. Zhang, H. Krim, G. G. Walter, Object representation and recognition in shape spaces. Pattern Recognition 36 (2003) 1143-1154.
8. Z. Hou, T. S. Koh, Robust edge detection. Pattern Recognition 36 (2003) 2083-2091.W. P. Choi, K. M. Lam, W. C. Siu, An adaptive contour model for highly irregular boundaries. Pattern Recognition 34 (2001) 323-331.

9. J. Weickert, Efficient image segmentation using partial differential equations and morphology. Pattern Recognition 34 (2001) 1813-1824.
10. F. Mahoudi, J. Shanbehzadeh, A. M. Eftekhari-Moghadam, Image retrieval based on shape similarity by edge orientation autocorrelogram. Pattern Recognition 36 (2003) 1725-1736.
11. H. Araujo, J. M. Dias, An Introduction to the Log-Polar Mapping. Proceedings of Cybernetic Vision, 1996, Second Workshop, 9-11 Dec. 1996, 139-144.
12. L. Xiping, T. Jie, Modified Active Contour Based Segmentation of Medical Image Series. Proceeding of the IASTED International Conference, Computer Graphics and Imaging, August, 2001 13-16, Honolulu, Hawaii, USA.
13. R.R. Ramon. G. L. Juan Frnasco. A measure of quality for evaluating methods of segmentation and edge detection. Pattern Recognition 34(2001) 969-980.
14. G. P. Daniel, C. Gu, Extensive Partition Operators, Gray-Level Connected Operators, and Region Merging/Classification Segmentation Algorithms: IEEE Transactions on Image processing, VOL. 10, NO. 9 Sept. 2001, 1332-1345.
15. N. S. Roussopoulos, F. Vincent, Nearest neighbor queries. Proceeding of ACM SIGMOD International Conference On the Management of Data, June, 1995, 71-79.
16. J. Shanbehzadeh, A. –M. Eftekhari-Moghadam, F. Mahmoudi, Image indexing and retrieval techniques: past, present and next, Proceedings of the SPIE: Storage and Retrieval for Multimedia Database 2000, Vol. 3972, San Jose, California, USA January 2000, 461-470.
17. M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, Int. J. Comput. Vision 9 (1988) 321-331.
18. H. J. Lin, Y. T. Kao, A prompt contour detection method, International Conference On the Distributed Multimedia Symtems, 2001.
19. G. Rafel, E. G. Richard Woods, Digital Image Processing.
20. G. Rafel, E. G. Richard Woods, Digital Image Processing.

Proposed Method

- Edge Detection

Edge detection is a front-end preprocessing step in most of the computer vision and image understanding systems. The accuracy and reliability of edge detection is critical to the overall performance of these systems. Earlier researchers have paid a lot of efforts on edge detection, but up to now, edge detection is still a high challengage. In this section, we will briefly illustrate two common methods of edge detection, and point out their drawbacks. In addition, we will proposed a simple and efficient method the detect edges.

i. Active Contour Model

The first common used method is Active contour model or called snake [9]. A Snake model is an energy minimizing spline that can be operated under the influence of internal contour forces, image forces, and external constraint forces. An energy function of the Snake model can be defined using the following equation:[9]

$$E^*_{snake} = \int_0^1 \left(E_{internal}[v(s)] + E_{image}[v(s)] + E_{constrant}[v(s)] \right) ds \qquad (1)$$

The location of the final template for the Snake model corresponds to a minimum of the energy function. However, the original Snake model meets some problems. The first problem is that without initially surrounding the object, the template will fail in detection. Second, the original Snake cannot detect multiple objects.

Wai-Pak Choi, and Kin-Man Lam [9] have proposed a modified version of the Snake model to solve these disadvantages. They used some critical points for splitting and connecting neighboring contour points. The idea is

that if the distance between two adjacent contour points is greater than a given threshold, a new contour point is inserted between these two points. In contrast to the smaller distance between two adjacent contour points, a contour point between these two points is deleted. This method overcomes the problem of detecting multiple objects, but the need to initially surround the objects still remains, and it needs several iterations to fit the boundaries of objects and thus requires a lot of run time.

ii. Sobel Operation

Another common edge detection method is Sobel method [20]. The Sobel operation uses four masks to detect edge points, as illustrated in Fig. 1.

Fig. 1. Sobel Masks

Sobel operation computes the derivation of a local pixel, called the gradient. For the gray-level function f(x,y), the gradient of f at pixel (x,y) is defined as the vector:

$$\nabla f = \left[\partial f / \partial x \quad \partial f / \partial y \right]^T \qquad (2)$$

The magnitude of ∇f can be approximated as the following equation:

$$\begin{aligned} \|\nabla f\| \approx & \left|(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)\right| \\ & + \left|(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)\right| \\ & + \left|(z_6 + 2z_8 + z_9) - (z_1 + 2z_2 + z_4)\right| \\ & + \left|(z_4 + 2z_7 + z_8) - (z_2 + 2z_3 + z_6)\right| \end{aligned} \qquad (3)$$

Fig. 1(a) approximates the derivative in the x direction, (b) get the derivative in the y direction, and (c) and (d) compute the derivatives the diagonal directions. Equation (3) computes the magnitude level of the current interesting pixel, if $\|\nabla f\|$ is greater than a threshold, this pixel is marked as an edge point.

Although, the Sobel's edge detector generally works well, it must repeatedly the whole image many times. As a result, this method requires Width*Height*3^2 time to operate on an image of size Width*Height. We compare this result with that of the pro posed method in the following section.

iii. A Prompt Edge Detection Method

n this subsection, an easy and effective scanning edge point algorithm is introduced, including threshold value determination and edge point detection algorithm.

The Prompt edge detector detects edge points by checking the differences of the gray value of each point from those of its neighbors. For a pixel (x,y) with gray value g(x,y), let g0(x,y), g1(x,y), …, and g7(x,y) denote the gray values of its neighbors in 8 directions as shown in Fig. 2

g_3	g_2	g_1
g_4	g	g_0
g_5	g_6	g_7

Fig. 2

Let $h_d(x,y) = |g(x,y)\, g_d(x,y)|$ be the difference of the gray value of pixel (x,y) from that of its neighbor in direction d. Let $B_d(x,y)$ be the number of those differences that exceed a threshold T, where T = a + c, c is a constant, and a is the average difference all pairs of adjacent pixels' gray values. In this work, we take the value 2 for c and, instead of taking a single value for T, we take the local average differences for T by dividing the whole image into $M \times M$ blocks and computing the average difference for each block.

The pixel (x,y) is indicated as an edge point if the following inequalities hold :

$$3 \le B_d(x,y) \le 6 \qquad (4)$$

These inequalities can also avoid indicating noisy points as edge points.

- Shape Representation

A good representation can manifest object's characteristics. Furthermore, it would achieve a high recognition rate when designing a content-based image retrieval system. Some researchers [2-4] have shown that objects are very distinguishable on the basis of their visible features, among these features, an object's shape is frequently an important key to its recognition. In this section, we will describe a shape representation, which is invariant to both rotation and translation but not invariant to scale.

i. Central Point Determination

The first step is to locate the central point of one object. In order to permit representational n variance to position, rotation and scale, the geometric center of the shape is selected as a reference point.

Vertical/Horizontal Medial Cruves

Let x_t, and x_b be the upper and lower bound lines of a contour image, and we equally divided them into m partitions. The line equation of each partition is $x = x_t + k * d_x$, where k = 0 ~ m-1, and $d_x = (x_b - x_t)/m$. Assume that there are two intersection points between contour and each line equation, which are a_k and b_k respectively. We compute the central point p_k of the a_k and b_k. The final process is to link $p_0 \rightarrow p_1 \rightarrow \ldots \rightarrow p_{m-1}$. With the similar process, we can get the horizontal curvature line. The visualization is depicted in Fig. 3.

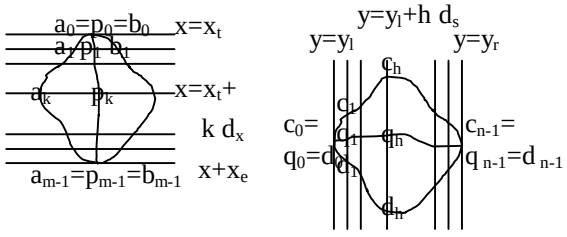


Fig. 3.The process of vertical/horizontal

The intersection point of vertical and horizontal medial curves is the central point of the contour object. Fig. 4 depicted the central point location.

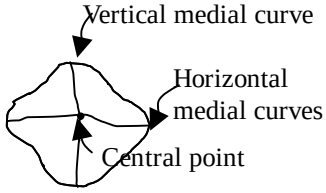


Fig. 4.Central Point Detection

ii. Polar Representation and Distance Sequences

A shape representation method [3] is outlined in this section. We characterize the contour by a sequence of contour points described in polar from. Here, we take the pole at the centroid (x_c,y_c), then the contour graph can be represented by a polar equation $d = f(\theta)$, and each contour point (x,y) has the polar description (d,θ), where x, y d and θ are related by equation (5) and (6). A sketch map was shown in Fig. 5:

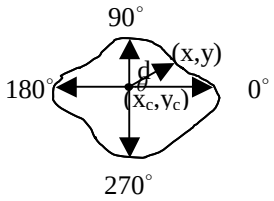
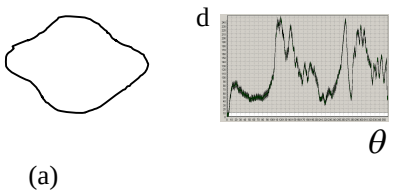


Fig. 5.Distance and angle of the contour points relative to center

$$d = \sqrt{(x-x_c)^2 + (y-y_c)^2} \quad (5)$$

$$\theta = \tan^{-1}(y-y_c / x-x_c) \quad (6)$$

With polar description, we may represent contour points as a sequence $(d_0, d_1, \cdots, d_{n-1})$, where $d_i = f(\theta_i)$ and $\theta_i = i * 2\pi/n$. By equation (5), d_i is known to be the distance between the corresponding contour point and the centroid. Fig. 6 illustrates the graph of the polar equation $d = f(\theta)$ for a given contour.



(a)

Fig. 6. An example of polar representation. (a) A contour of a object. (b) Graph of polar representation for the contour in (a)

We can obtain the distance sequence $(d_0, d_1, \cdots, d_{n-1})$ by successively rotating a ray emanating from the pole counterclockwise through an fixed angle $\Delta\theta$, where $\Delta\theta = 2\pi/n$ for a positive integer n, and at each step of rotation, recording the distance of intersection point of the current ray L_i with the contour. This method of representing and describing the shape of an object is adequate for convex objects. However, for an object containing concavities, a ray may intersect the contour at more than one points. If we only record the distance of one point (say, smallest distance) at each step of rotation, the contour points of some protrusions may be missed, see Fig. 7. We can eliminate this problem by recording the distance of the farthest intersection point at each step of rotating the scanning ray, or for more detailed description, recording the distances, or furthermore associating the number of all the intersection points, the result is shown as Fig. 8. We got points A and D instead of B and C points, otherwise we will got the similar features of two different objects. Besides, in order to provide scale invariance, the maximum distance is computed and all distances are normalized to it. Thus, all values fall between 0 and 100 regardless of the scale image of object.

Fig. 7. Features selection of two different concave objects.

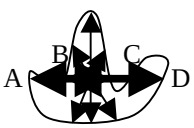
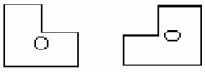


Fig. 8.Features selection of concave object

iii. Minimum Low-to-High Sequence (LHS)

Since we take the pole at the centroid of the object, the representation of a distance sequence is

translation-invariant. To achive rotation-invariance, Bernier [2] rotates the polor coordinate system, such that the maximum distance is associtated with the angle of zero. But there is no any guarantee of this method, if there are more than one maximum distances, the representation is not unique. As the example shown in Figure 9, both objects have three maximal distances and so each has three possible representaions: ($\sqrt{2}$, 1, 1, $\sqrt{2}$, 1, $\sqrt{2}$,1), ($\sqrt{2}$, 1, $\sqrt{2}$,1, $\sqrt{2}$, 1,1), and($\sqrt{2}$,1$\sqrt{2}$, 1,1, $\sqrt{2}$,1). To reduce this problem, we propose another representation method that deals with all distances rather than only the individual maximal distances in the sequence. First, we evaluate the ordering-consistency functions c_i's, as defined in Equation (7), at the seguence D = (d_0, d_1, …, d_{n-1}) as described in sub-section 3.1, and find out the index, s, of the function having smallest value, as defined in Equation (8). Then shift forward the distances in the sequence D for s positions to yield a new sequence D_s= (d_s,d_{s+1}, …,d_{s+n-1}), where $d_k = d_{k \bmod n}$ for each k.



(a) (b)

Fig. 9. An example of rotated object with three maximum distances

$$c_i = \sum_{k=i}^{i+(n-1)} \left(n - \left(k - i\right)\right) * d_k \qquad (7)$$

$$s = \arg\min_{0 \le i < n} \left\{c_i\right\} \qquad (8)$$

Without any surprsing, the new representations for Fig. 9(a) and 9(b) are identical and come out as (1, 1, $\sqrt{2}$, 1, $\sqrt{2}$, 1, $\sqrt{2}$), and as a result, they are considered having the same shape.

iv. Shape Matching

The final step of content-based image retrieval is matching stage. As discussed in the preceding subsection, we use the LHS to represent the shape of objects. And then we may proceed to the matching stage to measure the similarity between two objects by simply evaluating the Euclidean distance of the LHSs of them. The Euclidean distance between two sequences is given in Equation (9).

$$\begin{aligned} d\left(U,V\right) &= \left\|V - U\right\| \\ &= \sqrt{\left(v_0 - u_0\right)^2 + \cdots + \left(v_{n-1} - u_{n-1}\right)^2} \\ &= \sqrt{\sum_{i=0}^{n-1} \left(v_i - u_i\right)^2} \end{aligned} \qquad (9)$$

where U,V is the query and dababase images, v_i, u_i is their i_{th} feature, and n is the number of features. Any arbitrary two-dimensional shape can then be compared to any other by the outline method.

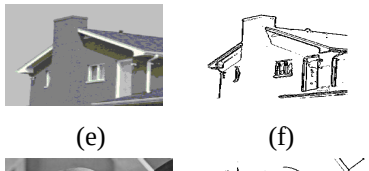
- Experiments and Results

This section consists of two subsections as follows: The first subsection shows the proposed edge detection method compared with Sobel operation. The second subsection presents the contour retrieval results.

i. Edge Detection Results

In this subsection we demostrated the results of our edge detection method. We also compared the processing time with Sobel operation. Fig. 10 shows some results of our porposed method. The test images include binary, gray and true color images.

(a) (b) (c) (d)



(e) (f)

(g) (h)

We transferred RGB color model to HSI color model, then we got the I vector to process edge detection, respectively as Fig. 10 (d). Fig. 11 shows the running time of Sobel and our proposed methods. We find that with the growing of imgae size, the time efficiency of the proposed method is more evident compared with that of Sobel's method.

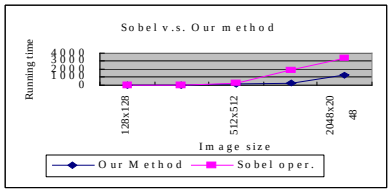


Fig. 11. The comparison on running time of the proposed method with Sobel's method.

ii. Contour Retrieval Results

The shape feature database was constructed for the image data set publicly available through the web site: http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html from the VSSP Center of the University of Surrey, UK. The data set is composed of 1100 boudnary contours of marine creature images, originally scanned from some printed books. Some experimental results for shape retrieval are presented in Fig. 12. In each figure, the query image is presented at the top, and the query results are arranged from top to left in the ascending order of matching distance.

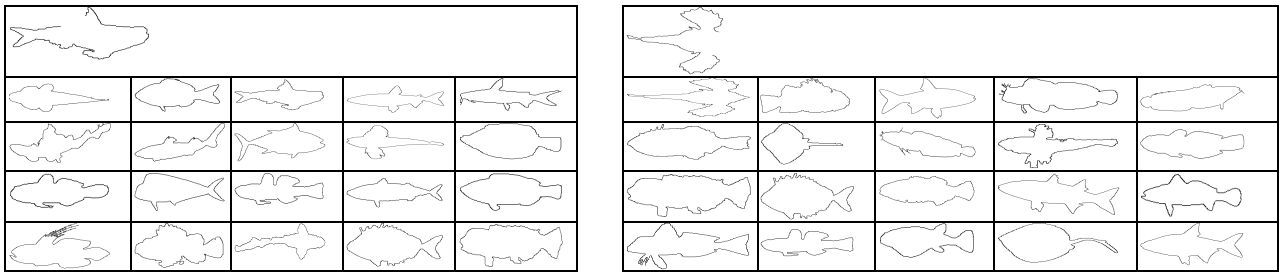


Fig. 12. Example of shape retrieval from the shape database of marine creature images. The query shape is presented at the top, and retrieved results are arranged in the ascending order of distance from top-left.

The results are also compared with H. Nishida [3] obtained by the authors, as shown in Table 1. Table 1 presents the average query rates for top 1%, 2%, 3%, 4%, 5%, and top 10% choices when the query images with a fixed degree of rotation. The choices rate meant that a curve segmentation is composed of how many percentage portions of a complete contour. We found that our proposed method almost has the same outstanding results as Nishida, and has much excellent than Stein.

- Conclusions

This paper presents a new edge detection method [19], which can cut down the processing time compared with Sobel operation.

We also proposed a new method of shape representation, the low-to-high sequence, that is invariant to translation, rotation, and scale. The experimental results of our method show superiority over the following schemes: edge detection, polar representation, a minimum one dimension sequence re-order, and similar matching. We have also compared our proposed method with Nishida and Steins' methods [3] to prove its robustness and effectiveness.

Table 1
Average query rates (%) of deformed patterns of Proposed, Nishida and Stein methods.

β	Choices (%)	Method	r=1%	r=2%	r=3%	r=4%	r=5%	r=10%
0^o	100	Proposed	100	100	100	100	100	100
		Nishida	100	100	100	100	100	100
		Stein	92.8	96.8	98.3	98.6	98.8	99.7
	90	Proposed	100	100	100	100	100	100
		Nishida	100	100	100	100	100	100
		Stein	92.1	96.1	97.6	98.2	98.7	99.6
	80	Proposed	98.1	99.2	99.9	100	100	100
		Nishida	99.6	99.7	99.7	99.7	99.8	99.8
		Stein	86.3	92.3	94.8	96.1	97.3	98.9
	70	Proposed	97.2	99.1	99.9	99.9	99.9	100
		Nishida	97.8	98.6	99.1	99.3	99.4	99.6
		Stein	75.8	86.0	89.2	91.3	93.0	96.2
90^o	100	Proposed	96.3	98.7	99.9	99.9	99.9	99.9
		Nishida	99.8	99.8	99.8	99.9	99.9	99.9
		Stein	80.6	87.4	89.9	92.1	92.8	96.5
	90	Proposed	97.2	98.6	99.9	99.99	99.9	100
		Nishida	99.2	99.6	99.7	99.7	99.7	99.8
		Stein	76.8	84.6	88.7	90.5	91.6	95.8
	80	Proposed	96.1	98.2	99.1	99.5	99.9	99.9
		Nishida	96	97.8	98.2	98.6	99.1	99.5
		Stein	69.5	79.8	83.9	86.4	88.1	93.1
	70	Proposed	92.3	94.5	95.4	98.3	99.1	99.5
		Nishida	87.2	91.7	93.7	94.5	95.4	97.4
		Stein	52.2	64.8	70.8	75.1	77.9	86.5

- 計畫成果自評

本計畫執行成果所提之形狀查詢方法能有效且快速的解決物件表示之唯一性問題，執行成果顯示與當初所提之原計畫內容大致相符，也達到預期目標。本計畫內容已整合成一篇論文，submitted 於 The 6th International Workshop on Multimedia Network Systems and Applications (MNSA'2004)。