行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※
※　　　　　　　　　　　　　　　　　　　　　　※
※**以顏色、形狀及空間關係為基礎之影像擷取系統**　※
※　　　　　　　　　　　　　　　　　　　　　　※
※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：¦個別型計畫　　□整合型計畫

計畫編號：NSC 89-2218-E-032-013

執行期間：八十九年八月一日至九十年七月卅一日

計畫主持人：施國琛

共同主持人：鍾興臺

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**

執行單位：淡江大學資訊工程學系

中　華　民　國　九十 年　九 月　二十日

行政院國家科學委員會專題研究計畫成果報告

以顏色、形狀及空間關係為基礎之影像擷取系統

A Flexible Content-based Image Retrieval System Integrating with Color, Shape and Spatial Relations

計畫編號：NSC 89-2218-E-032-013
執行期限：八十九年八月一日至九十年七月卅一日
主持人：施國琛 淡江大學資訊工程學系
共同主持人：鍾興臺 淡江大學資訊工程學系

中文摘要

本計畫發展了一套使用顏色、形狀及空間關係來擷取影像的系統，本系統可讓使用者以調整影像屬性權重的方式，有彈性的從近 5000 張的圖片中，查詢所需之影像。另外，根據自訂的顏色叢集方法及人類對顏色的知覺，我們提出一套自動產生影像資料庫索引的機制，此索引機制可使影像的過濾更有效率。而且，我們提出了一個類似播種(Seed-Filling)的演算法，可以順利的求得影像中物件的形狀及空間的關係。由於一般而言，使用者很難確定物件間的距離，所以我們使用定性的(qualitative)空間關係來分析空間的相似性。另外，本系統所提供的圖形式使用者介面及繪圖工具，可以讓使用者以載入或描繪影像的方式，輕鬆的查詢影像。除此之外，本系統的回饋學習機制更可增強搜尋的準確度。我們的經驗證明本系統所提之方法，可讓使用者有效地找出所求之影像。

關鍵字: 內涵式影像擷取系統，影像資料庫，顏色，形狀，空間關係。

Abstract

In this paper, we provide color, shape and spatial relations approach to retrieve images. User can flexible use this three methods by select their weight to retrieve images from nearly 5000 pictures. We propose a revised automatic indexing scheme of image database according to our color clustering method, which could filter the image efficiently. As a technical contribution, a Seed-Filling like algorithm that could extract the shape and spatial relationship feature of image is proposed. Qualitative approach is applied to the similarity comparison of spatial differences. Besides, a friendly visual input interface and a feedback learning mechanism are provided. And we already have on-line internet system. Our experience shows that the system is able to retrieve image information of a very high satisfaction.

Key Words: content-based image retrieval, image database, color, shape, spatial relation

1. INTRODUCTION

Color, shape and spatial relation are the main features for human beings as well as computers to recognize the image. For all color-based retrieval methods, there are some common issues: the selection of a proper color space and the use of a proper color quantization scheme to reduce the color resolution. Wang *et al* [1] reduce the color resolution by hierarchical clustering, CNS (Color Name System) merging, and an equalize quantization method. Swain and Ballard [2] using histogram intersection as color indexing. Wan and Kuo [3] use hierarchical color clustering method based on the pruned octree data structure. In this paper, we select the HIS color space to cluster the color resolution according to the sensation of human beings.

In general, shape representations can be divided into to categories, boundary-based, and region-based. The most successful representations for these two categories are Fourier Descriptor and Moment Invariant. The main idea of Fourier Descriptor is to use the Fourier transformed boundary as the shape feature [4]. The main idea of Moment Invariant is to use region-based moments, which are shape feature invariant to transformation. In this paper, we proposed a series of methods to solve the image retrieval problems especially in shape representation and matching. Such as defining the signature of the object, statistical recognition and species classify.

Many researchers propose temporal modeling of multimedia objects. The mechanism handles 1-D and 2-D objects, as well as 3-D objects, which treat a time line as the third axis based on the R-tree spatial data structure (proposed in Michael *et al.*, 1996). The 2D Projection Interval Relations (i.e., PIR) [5] is based on both directional and topological temporal relations. Image retrieval algorithms were discussed based on PIR. The use of spatio-temporal relations serves as a reasonable semantic tool for the underlying representation of objects in many multimedia applications. In this paper, we extend temporal interval relation by means of a complete analysis for spatial computation.

The rest of our paper is organized as follows. In Section 2, the color clustering and indexing scheme of image database are discussed. The algorithm for extracting features will describe in section 3. The similarity functions of color, shape and spatial relation features are given in Section 4. Section 5 describes the procedure of image retrieval in our system. Finally, a short conclusion is given in Section 6

2. COLOR CLUSTERING AND INDEXING SCHEME OF IMAGE DATABASE

In this section, we introduce the selection of color space, procedure of color clustering and the normalization of image. Besides, the index scheme and filer mechanism according to clustering scheme for speeding up the retrieval process will be described.

There are several color spaces existing for a variety of reasons. We choice the HIS color space because of its similarity and perceptibility. Similarity means that two perceptually similar colors will be in the same or neighbor quantized color bin and two non-similar colors will not be in the same quantized color bin. Thus, the similarity of two colors can be determine according to the distance in the HIS color space. Besides, the HIS color space is defined based on the human color perception. User could choice the color he/she wants easily by indicate hue, saturation, intensity value independently.

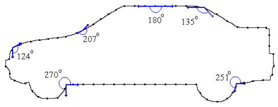
In the procedure of color clustering, firstly, we equally quantized the RGB color space to change color levels from 256 to 16 levels in each axis. Secondly, we linearly convert the 16-level RGB color bins to the HSI coordinates. And, we cluster the hue to 12 levels, since hue is represented as circle and primary hues are located on the equal space at 60 degrees (Red, Yellow, Green, Cyan, Blue and Magenta) in the HSI color space. And, because the human visual system is more sensitive to hues as compared to saturation and intensity, the H axis should be quantized finer than the S axis and the I axis. According to our experiment, we quantized the HSI color space into 12 bins for hue, 4 bins for saturation, and 4 bins for intensity. Finally, We normalize the resolution of all images to 400*300 pixels.

After the quantization and normalization, our system will index the images according to the dominant colors of those images. First, system will calculate the histogram and dominant colors of the image. The color histogram is an array that is computed by differentiating the colors within the images and counting the number of pixels of each color. From the color histogram, we could choose the dominant colors whose numbers of pixels exceeds the threshold. After getting the dominant colors, system will save the unique image ID to each corresponding color bin. According to this indexing scheme, system could load the candidate images that have same dominated colors and eliminate irrelevant images immediately before the more complex and expensive similarity measure.

For a small image database, sequentially searching the image during the retrieval process will be fast and provide acceptable response time. However, it is not feasible for large image database. Therefore, we propose a filtering mechanism to eliminate irrelevant images before the more complex and expensive similarity measure. First, our system will load the image ID arrays according to the dominant colors of query image. Next, system will conjunct and rank the image ID arrays according to the number of appearance. After this step, system could filter out the irrelevant images effectively.

3. THE EXTRACTION OF FEATURES

In this section, we present our algorithms for extracting the features of image. Our feature extraction algorithm is based on the bitmap format. Since its pixel stream could represent the spatial information. For others format image, we will



convert it to the bitmap format only in the feature extraction phase. After the feature extraction, the system will save its feature information and original format image to the database.

The extracting of color feature includes two main procedures. First, system reads the RGB value of pixels from image file sequentially then converts them to the HSI value. Secondly, system constructs the color histogram of the image according to the number of pixels of each color within the image. Besides, system records the color sensation of image including warm - cold, gray - vivid and bright – dark according to the average of hue value, saturation value and intensity value of image for filtering.

There are three main steps of shape extraction: First, we extract the shapes of image by the enhanced Seed Filling algorithm, Then, we normalize the shape and convert those shape from region to contour by edge detection before the shape similarity measurement. In the first step, system will record the color, location, height, width and area information of the objects for the similarity measurement and identify the color and spatial relation. The normalized contour that generated by the second and the third steps is used in the measurement of shape similarity only.

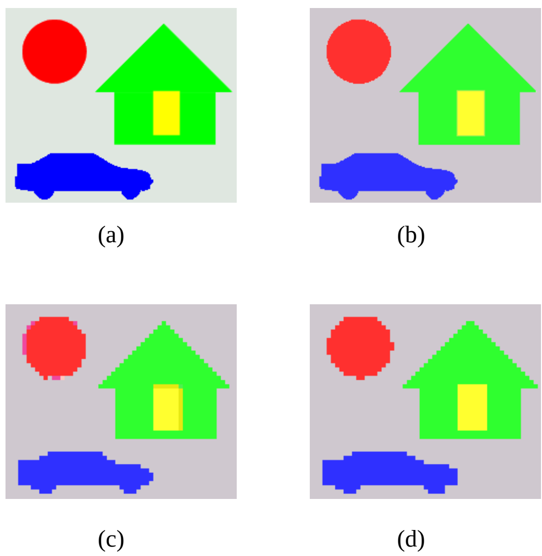


Fig. 1 Example of processed image (a) Original image .(b) Color clustering image. (c) Shape extraction image (d) Refined image

Shape Representation and Normalization: Technically, edge detection is the process of locating the edge pixels, and edge tracing is the process of following the edges, usually collecting the edge pixels into a list. This is done in a consistent direction, either clockwise or counterclockwise around the objects. Chain coding [7] is one of the methods of edge tracing. The result is a non-raster representation of the objects, which can be used to compute shape measurement or otherwise identity or classify the object. But, this method is influenced by the rotation and scale of the object.

Fig. 2. The turning angle representation of the object.

In this paper, we use the turning angle representation to describe the object. Shape representation is a set of turning angles $\theta=\{\theta(1), \theta(2), \theta(3), \ldots \theta(N)\}$ as showed in Fig. 2. This method is invariant for translation and scale of the object. In addition, it is invariant for rotation after the normalization. The turning angle of the object can express the edge's subtle difference including the curvature and distance. According the turning angle variation, we classify feature tokens based on their curvature properties.

In our system, the edge is segmented to be a fixed number of parts at first. And, the turning points are selected from the edge by computing the local maximum curvature points in each segment. So, the number of turning points in the image is fixed. This method is better than selecting the point at the segment point that may lose the important turning angle. After the extraction of shape, our system extracts the spatial relation by projecting the shape from a 2-D space to two 1-D spaces. Our system records the starting and ending points of two objects in the X and the Y coordinate by splitting the shapes orthogonally to the coordinate axis. And, those points will estimate the spatial relation of two objects in the X and Y coordinates according to the proposed mechanism discussed in [8].

4. THE SIMILARITY MEASURE OF FEATURES

In our system, the similarity score of the images is integrated by three components: color similarity score named *CS*, shape similarity score named *SS* and spatial relation similarity score named *SRS*. And, the similarity score *S* is defined as following:

$$S = Wc * CS + Ws * SS + Wsr * SRS$$

where *Wc*, *Ws* and *Wsr* are the weights of color, shape and spatial relation, which could be indicated by user or via training by our system.

4.1 The Similarity Measure of Color

We present the similarity formula of two colors first. Then, we define the similarity of two images according to the similarity of their colors.

The Similarity of two Colors: Let the hue, saturation and intensity value of color C are H, S and I. The similarity between two colors, i and j, is given by:

$$C(i,j) = W_h H(i,j) + S(i,j) + I(i,j)$$

where

$$H(i,j) = \min\left(\left|H_i - H_j\right|, 12 - \left|H_i - H_j\right|\right)$$

$$S(i,j) = \left|S_i - S_j\right|$$

$$I(i,j) = \left|I_i - I_j\right|$$

And, we set *Wh* =2. Because In our experience, the change of hue value affect the conceptual similarity is sensitive than the change of saturation and intensity value. Finally, the degree of similarity between two colors, *i* and *j,* is given by:

$$CS(i,j) = \begin{cases} 0 & \text{if } H(i,j) > H\max \\ 1 - \dfrac{C(i,j)}{C\max} & \text{otherwise} \end{cases}$$

where *Cmax=WhHmax+Smax+Imax.* In which, *Cmax, Hmax, Smax* and *Imax* are the maximum tolerant distance of the color, hue, saturation and intensity in the similar colors, And, *Hmax* is 2, because two colors are not perceive to be similar when the distance of their hue value exceeds 60 degree (60/360=2/12) in the HIS color space. *Smax* and *Imax* are equal to 4-1=3.

The Color Similarity of two Images: The color similarity between two images shall be discussed in two fields: the difference of pixel numbers in the completely same color and the perceptual similar colors. In the first field, the similarity measure between query image Q and database image D for a color i can be determined as:

$$CSc(Q,D,i) = \min(Qi,Di)$$

And, the similarity measure between query image *Q* and database image *D* can be determined as following formula for all colors:

$$CSc(Q,D) = \sum_{i=1}^{198} CSc(Q,D,i) / 400 * 300$$

where 198 is the total number of colors in our clustering HIS color space (12*4*4 + 6 gray colors).

In the second field, the similarity measure between query image *Q* and database image *D* for a color *i* can be determined as:

$$CSp(Q,D,i) = Max(\min(Qi,Dj) \times CS(i,j)) \quad \forall j \in Cp$$

where Cp is the set of colors that are perceptually similar to color i. Then, the similarity formula between query image Q and database image D for all colors is:

$$CSp(Q,D) = \sum_{i \in C_D} CSp(Q,D,i) / Qi$$

Finally, we define the similarity formula of the images according to the color feature:

$$CS(Q,D) = Max(CSc(Q,D), CSp(Q,D))$$

4.2 The Similarity Measure of Shape

Given two shapes represented by their turning angle vectors $\theta 1$ and $\theta 2$. For the best match and partial similarity, we record the minimum distance of the two shapes by rotating one to match another. Then, the distance of similarity between these two shapes is calculated:

$$SS_D(\theta_1,\theta_2)=\min\left(\sum_{i=1}^{N}\left|\theta_1(i)-\theta_{2r}(i)\right|\right)$$

where N is the fixed number of turning points in the image and $\theta 2r$*'s* are in the shape set of rotate angle $\theta 2$. In this formula, if the segment is a straight line, we will set the turning angle to 180 degree for the segment.

Let the normalized similarity measurement degree from 1(completed matching) to 0(most dissimilar matching). The measurement between the requested image and archive images will be:

$$SS(\theta_1,\theta_2) = 1 - \frac{SS_D(\theta_1,\theta_2)}{N*360}$$

where N *360 is the maximum distance measure of the requested image and archives images, the SS_D is the minimum distance measure of the requested image and archives images. If the archives images include the requested image, SS_D is equal to 0, and SS is equal to 1. In addition, we set a threshold for disregarding dissimilar images. If SS_D is greater than the threshold, it means that the requested image is dissimilar to archives image. Then, we set the SS to 0.

4.3 The Similarity Measure of Spatial relation

Before the similarity measure of spatial relation, we identify the object according to the weights of color, shape and area of objects that adjusted by user. In this section, we use the 2-D spatial relational representation that proposed by authors in [8] to demonstrate the usefulness in the spatial domain. It is the simplest kind of spatial information of practical relevance. In 1-D spatial relation case, while different from the temporal relations, can be handled by a simple modification of temporal interval relations.

Let $EPIRD18(r_i, r_j)$ be a EPIRD index function takes as input two relations, $r_i, r_j \in 18REL$, and returns a similarity index from 0 to 8. Note that, the lower the index, the closer the relation to the relations. Then, we have:

$$\text{similarity distance} = EPIRD18(r1, r2)$$

Assume that A and B are spatial relation strings of two object strings in different images, And, $SRS_X(A, B)$and $SRS_j(A, B)$ are the similarity functions of A and B in the X and the Y coordinates. We define a normalized spatial relation similarity function of A and B as:

$$SRS(A,B)=(SRSx(A,B)+SRSy(A,B))/2$$

and

$$SRSx(A,B)=1-\frac{\sum_{i=1}^{\min(m,n)-1}EPIRD(A(i)_x,B(i)_x)+8\times|m-n|}{8\times(Max(m,n)-1)}$$

$$SRSy(A,B)=1-\frac{\sum_{i=1}^{\min(m,n)-1}EPIRD(A(i)_y,B(i)_y)+8\times|m-n|}{8\times(Max(m,n)-1)}$$

where $A(i)_x$, $A(i)_y$, $B(i)_x$ and $B(i)_y$ denotes the ith spatial relation of spatial relation string A and B in the X and the Y coordinates, m and n are of the number of spatial relations in spatial relation string A and B.

Qualitative representation of object positions has an advantage. It is difficult for a user to specify how far two objects are separated. But, use the before or the after relation, approximate reasoning can be used.

5. CONCLUSIONS

We use color, shape and spatial method to retrieve images and promote these approaches more efficiency to search the images. We bring up a color clustering method and use it index images feature to the database and filter the images when the searching. Besides, improve the database hierarchy faster to retrieval the images. And extract the objects to find the shape and spatial information of the image. We use shape to analyze object similarities and qualitative spatial relations to analyze all the image objects distribution. In the system, we proposed a friendly interface that user can query by sketch or example and a feedback learning mechanism. Final, user can select the weight to retrieve the image and can search image through Internet. The system is implemented on Windows 98. We tested nearly 5000 pictures. The performance and accuracy are reasonable as expected.

REFERENCES

[1] Wang, J., Wen-jann yang and Raj Acharya (1997) Color Clustering Techniques for Color-Content-Based Image Retrieval from Image Databases. Proceeding of IEEE International Conference on Multimedia Computing and Systems (ICMCS), pp.442~449, Jun 3-6

[2] Swain, M.J. and .H. Ballard (1991) Color index. Internal Journal of Computer Vision, 7(1), 11~32.

[3] Wan, Xia and C.-C Jay Kuo (1996) Color Distribution Analysis and Quantization for Image Retrieval,.in SPIE storage an Retrieval for Image and Video Database IV, Vol. SPIE 2670, pp.9-16, Feb.

[4] Person, E. and K. S. Fu (1977) Shape discrimination using Fourier Descriptors. IEEE Trans. Sys. Man.

[5] Nabil, M., Anne H. H. Ngu, and John Shepherd (1996) Picture Similarity Retrieval Using 2D Projection Interval Representation. IEEE Transactions on Knowledge and Data Engineering, 8(4), 533-539.

[6] Canny, J. (1986) A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6), 679-698.

[7] Freeman, H. (1974) Computer processing of line drawing images. Computer Survey, 6, 57-97.

[8] Shih, Timothy K. and Anthony Y. Chang (1998) The Algebra of Spatio-Temporal Intervals. 12th International Conference on Information Networking, Japan, January 21-23.