

# Image Retrieval of Digital Crime Scene Images

Che-Yen Wen,<sup>1,\*</sup> Ph.D. ; Chiu-Chung Yu,<sup>1</sup> M.S.

<sup>1</sup> Department of Forensic Science, Central Police University, 56 Shu Jen Road, Taoyuan 333, Taiwan ROC.

Received: February 25, 2005 / Accepted: April 06, 2005

## ABSTRACT

The digital image has a predominant position among multimedia data types, and it plays a valuable role in numerous human activities, such as law enforcement, forestry management, environment management, weather forecasting and entertainment. In forensic sciences, digital images have been widely used, such as fingerprints, scene photos, firearm tool marks, and so on. Therefore, the arrangement of tremendous image data becomes a big issue, such as "how to get an interested image fast?" In this paper, we provide a retrieval method for digital image databases of crime scene photos. We use experimental results to show the capability of the proposed method. We also show the potential of applying the image databases to forensic sciences, such as case data management.

**Keywords:** Forensic science, Image retrieval, Image databases

## Introduction

Photographic images are increasingly being stored and transmitted in the digital format. Their applications range from personal use to media, advertising, art, medicine, education and even research in forensic sciences. In the consumer market, digital cameras have been quickly replacing traditional film based cameras. The characteristics of devices for acquiring, displaying, storing and printing digital images have been decreasing their cost. In recently years, the color accuracy and resolution of digital cameras have been improving. The advancement of the storage technology make it possible to store a large number of pictures in digital cameras before uploading them to a personal computer. Digital images are also easy to be shared and disseminated. They can be posted on personal web sites or sent via e-mail to everyone all over the world.

The digital image databases maintain large collection of images. They can be used to store, archive, browse and search interested data. Nowadays, fingerprint images have been used in law enforcement and become increasingly popular for access control to secure information and identity. The relevance of the digital image databases for crime scene applications is to

organize collections of scene pictures, evidence pictures, trace evidence images, vehicle number images, identification documents and investigation documents, and so on. In this paper, we provide a retrieval method for digital image databases of crime scene photos.

## Visible Image Retrieval

Recently, the availability of large digital archives has attracted research efforts in providing tools for effective retrieval of image data based on their content (content-based image retrieval, CBIR). An automatic content annotation is required for large digital archives. In CBIR, "how to reduce the semantic gap between user expectation and system supports" is a big challenge. For instance, a discrepancy between the query a user ideally would and the one that he actually could submit to an information retrieval system. We need a retrieval method based on raw images properties, and all retrieved images will look like the example image with respect to their color, texture, text, shape, etc. In some case, textually annotating visual content is a hard job. Unlike text, images cannot be regarded as a syntactically structured collection of words. The semantic gap is wider for images than for text.

In the last few years, a number of CBIR systems

\* Corresponding author, e-mail: cwen@mail.cpu.edu.tw

using image-recognition technologies, that is applied in industrial automation, social security, biomedicine, etc. In proof of biometric and security, face recognition systems are now widely used [1]. Similarly, automatic image-based detection of tumor cells in tissues is being used to support medical diagnosis and prevention [2].

However, there is much more to image retrieval than simple recognition. Image retrieval by similarity is the true distinguishing feature of a CBIR system (see Table 1). In this paper, we focus on general-purpose systems for retrieval of photographic imagery,

**Table 1** Typical features of recognition and similarity retrieval systems [3].

	Recognition	Similarity Retrieval
Target performance	High precision	High recall, any precision
System output	Database partition	Database reordering
Interactivity	Low	High
User modeling	Not important	Important
Built-in intelligence	High	Low
Application domain	Narrow	Wide
Semantic level	High	Application-dependent
Annotation	Manual	Automatic
Semantic range	Narrow	Wide
View invariance	Yes	Application-dependent

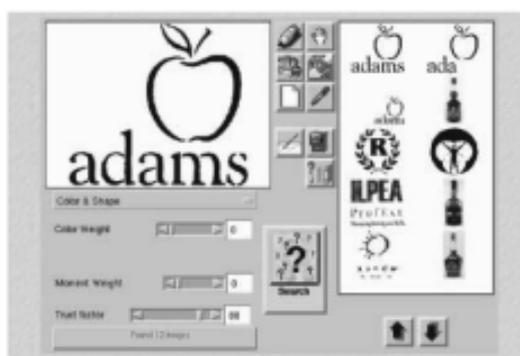
Every application of retrieval system is characterized by a typical set of possible queries reflecting a specific semantic content. These are partitioned into three main levels [3]: (1) Low Level: the basic perceptual features of visual content, as dominant colors, color distributions, texture patterns, relevant edges, 2D shapes, uniform image regions and their spatial arrangement; (2) Intermediate Level: this level is characterized by a deeper involvement of users with the visual content. This involvement is peculiarly emotional and is difficult to express in rational and textual terms. (3) High Level: there are the queries that reflect data classification according to some rational criterion.

### Visible Image Retrieval Examples

In this section, we show several examples of image retrieval according three main levels.

Because of the characteristics of trademark images is very simple. Image retrieval can be based on color and shape features. Figs. 1 to 3 show three different retrieval tasks. Fig. 1 shows the result in the case of retrieval by shape. This trademark is composed of the shape of apple and "adams". Fig. 2 shows this trademark contains two dominant colors. Fig. 3 performs retrieval based on both

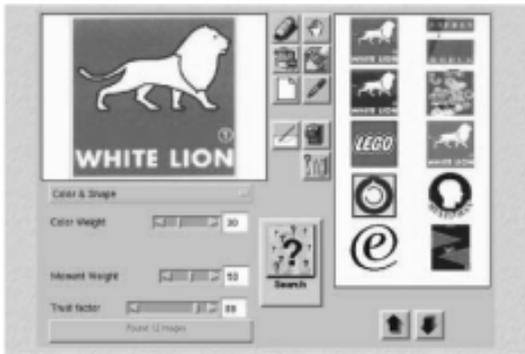
color and shape.



**Fig.1** Retrieval of trademarks by shape.  
[ftp://wiley.com/public/sci\\_tech\\_med/image\\_databases](ftp://wiley.com/public/sci_tech_med/image_databases).



**Fig.2** Retrieval of trademarks by color.  
[ftp://wiley.com/public/sci\\_tech\\_med/image\\_databases](ftp://wiley.com/public/sci_tech_med/image_databases).



**Fig.3** Retrieval of trademarks by shape and color.  
[ftp://wiley.com/public/sci\\_tech\\_med/image\\_databases](ftp://wiley.com/public/sci_tech_med/image_databases).



**Fig.4** Retrieval of art paintings by color similarity.  
[ftp://wiley.com/public/sci\\_tech\\_med/image\\_databases](ftp://wiley.com/public/sci_tech_med/image_databases).

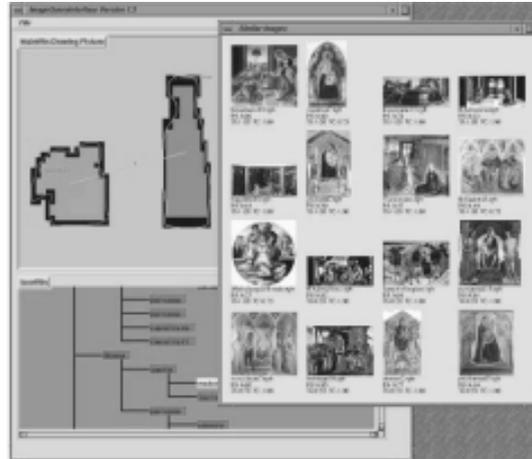


**Fig.5** Retrieval of art paintings by shape similarity.  
[ftp://wiley.com/public/sci\\_tech\\_med/image\\_databases](ftp://wiley.com/public/sci_tech_med/image_databases).

In Fig. 4 and Fig. 5, it is apparent that color and shape are the most important image characteristics for feature-based retrieval of paintings. As this demonstrates retrieval from modern art paintings, both low level and intermediate level queries are supported.

Bimbo and Pala provide an interesting example of concurrent exploitation of low level and high level

descriptors [4]. In this retrieval example, spatial relationships, color, shape and texture are combined with textual annotations of visual entities.



**Fig.6** Retrieval of iconographic pictures. .  
[ftp://wiley.com/public/sci\\_tech\\_med/image\\_databases](ftp://wiley.com/public/sci_tech_med/image_databases).

## Methods

A typical content-based image retrieval system is described in Fig 7 [5]. The image collection database contains raw images for the purpose of visual display. The visual features database is used to store visual features, which are extracted from the image collection database. The text annotation repository contains keywords and free-text descriptions of images. Multidimensional indexing is used to achieve fast retrieval and to make the system scalable to large image collections. The retrieval engine includes a query interface and a query-processing unit. The query interface, typically employing graphical displays and direct manipulation techniques, collects information from the user and displays retrieval results. The query-processing unit is used to translate user queries into an internal form. In order to bridge the gap between visual features and semantic means, the query-processing unit is usually used to communicate with the search engine in an interactive way.

In this paper, we will address "feature extraction", such as color, texture, shape, object spatial relationship, and Arabic numeral recognition.

A feature is defined as a distinctive characteristic of an image and a descriptor is a representation of a feature. In the image processing and computer-vision literature, the terms feature and descriptor are often used synonymously.

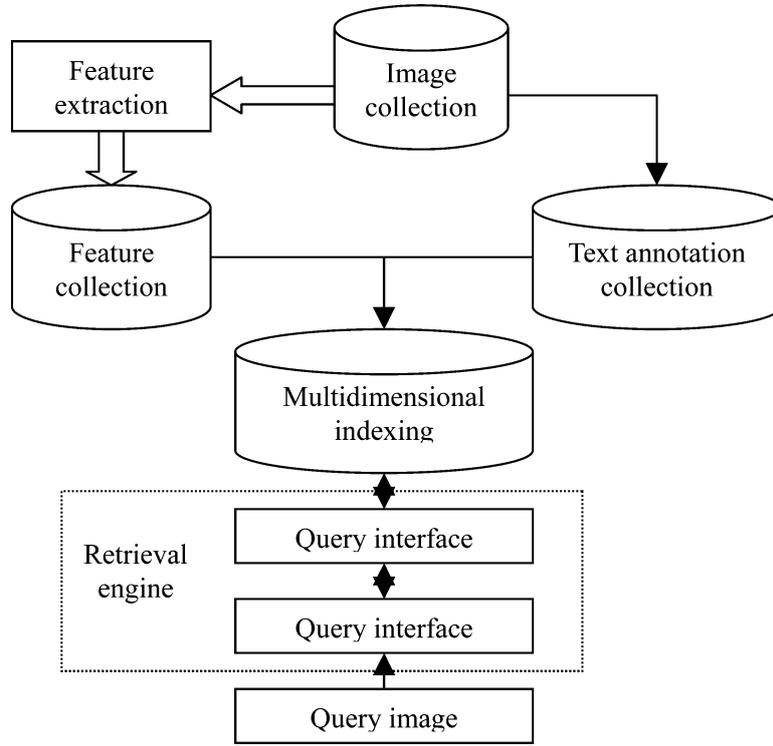


Fig. 7 A flow chart of image retrieval system.

### Color Descriptor Extraction

This section introduces several methods for image retrieval based on color features of images. Color is an important dimension of human visual perception that allows discrimination and recognition of visual information. Accordingly, color features have been found to be effective for indexing and searching of color images in image databases. We focus on the design and extraction of color descriptors and the matching methods. A color descriptor is a numeric quantity that describes a color feature of an image.

### Color Histograms

The generality of color descriptors are color histograms. If  $I$  denotes an image of size  $W \times H$ ,  $I(i, j)$  is the gray value of the pixel at position  $i, j$ , and  $Y_m$  is the  $m$ th code of the vector quantity, the color histogram  $h_c$  has defined by

$$h_c[m] = \sum_{i=0}^{W-1} \sum_{j=0}^{H-1} \delta(I(i, j), Y_m), (m = 1, \dots, M) \quad (1)$$

$$\delta(i, j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where the Kronecker delta function,  $\delta(\cdot, \cdot)$  is equal to 1 if its two arguments are equal and zero otherwise.

### Region Color

One of the drawbacks of global color histograms is that it cannot present the spatial distribution of color across different areas of the image. A number of methods have been developed for integrating color and spatial information for the content-based query. Sticker and Dimai illustrated an example of extracting localized color descriptors in [6]. As Fig 8, an image can be partitioned into sixteen uniform regions:  $r_0 - r_{15}$ . The dissimilarity,  $D_{q,t}$  of images based on the color spatial descriptors can be measured by computing the weighted sum of individual region dissimilarities as follows:

$$D_{q,t} = \sum_{m=0}^{M-1} w_m d_{q,t}(r_m^q, r_m^t) \quad (3)$$

where  $r_m^q$  is the color descriptor of the  $m$ th region of the query image,  $r_m^t$  is the color descriptor of the  $m$ th region

of the target image,  $w_m$  is the weight of the  $m$ th region and satisfies  $\sum w_m = 1$ ,  $d_{q,t}(\cdot, \cdot)$  is the distance between query and target images.

Smith and Chang presented a system for querying for images by the spatial and feature attributes of regions [7]. This system enables users to find the images that contain an arrangement of regions similar to that diagrammed in a query image (see Fig. 9).

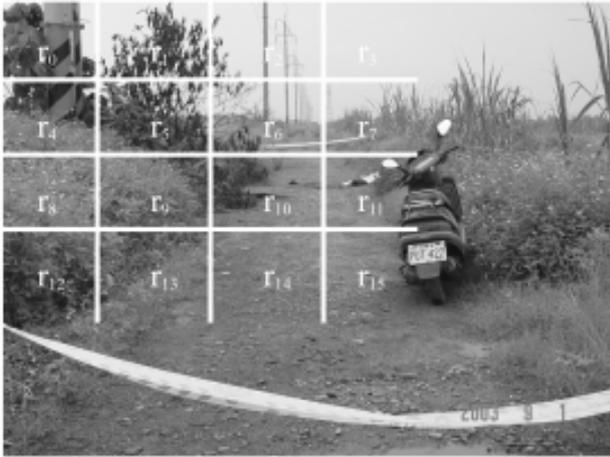


Fig.8 Sixteen uniform regions.

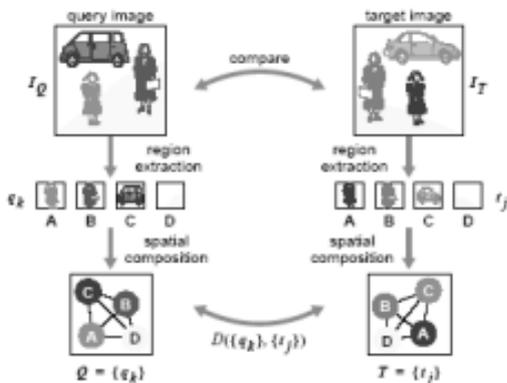


Fig.9 The integrated spatial and color feature query approach matches the images by comparing the spatial arrangements of regions [7].

### Texture Descriptor Extraction

The "texture" feature can distinguish many natural and man-made objects. In Fig. 10, pictures of cloud, stones, leafages and brick wall contain strong examples of image texture. Texture is a property of image regions. However, it is not easy for us to describe them directly in qualitative terms. In this paper, we show some analysis techniques to retrieval of textures.

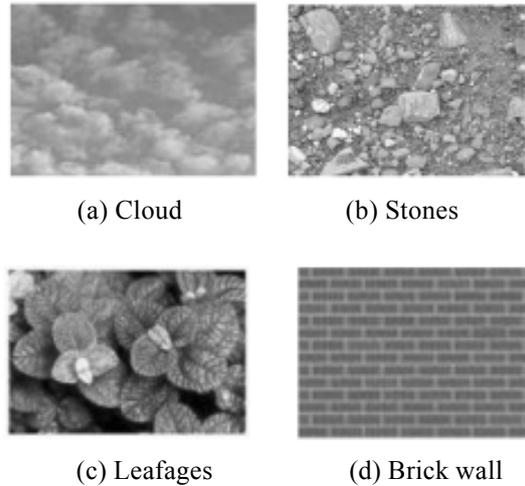


Fig.10 Examples of some textured images.

### Co-occurrence Matrices

Texture manifests itself as regular variations of the image intensity within a given region. The co-occurrence matrix is a popular descriptor for texture. Co-occurrence matrices are based on the second-order statistics of pairs of intensity values of pixels in an image. Let  $I(x, y)$  be the intensity value of an image,  $I$ , at location  $(x, y)$ . For a given displacement vector  $\Delta = (\Delta_x, \Delta_y)$ , the joint probability of a pixel at location  $(x, y)$  has a gray level  $i$ , and the pixel at location  $(x + \Delta_x, y + \Delta_y)$  has a gray level  $j$  in the texture image. The co-occurrence matrices are defined as follows:

$$P(i, j|\Delta) = \frac{1}{N} \sum g[f(x, y) = i, f(x + \Delta_x, y + \Delta_y) = j] \quad (4)$$

where  $N$  is the number of pairs of pixels separated by the displacement  $\Delta$ . If  $f(x, y) = i$  and  $f(x + \Delta_x, y + \Delta_y) = j$ , we have  $g[\cdot] = 1$ ; otherwise, we have  $g[\cdot] = 0$ .

### Coarseness

Some approaches try to measure the level of the spatial change rate in the image intensity, and therefore indicate the level of coarseness of the texture. The particular procedure can be briefly described as [8]:

**Step 1:** Build  $k$  images in which each element is the average of intensities of neighboring  $2^k + 2^k$  elements, where  $k=0, 1, \dots, 5$ .

$$A_k(x, y) = \frac{1}{2^{2k}} \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} T(i, j) \quad (5)$$

**Step 2:** Take differences between pairs of average that correspond to non-overlapping neighborhood both in

vertical and horizontal orientation.

$$\begin{aligned} E_{k,horizontal}(x,y) &= |A_k(x+2^{k-1},y) - A_k(x-2^{k-1},y)| \\ E_{k,vertical}(x,y) &= |A_k(x,y+2^{k-1}) - A_k(x,y-2^{k-1})| \end{aligned} \quad (6)$$

**Step 3:** For each point, compute the size of the neighbor mask that is the structural element as the maximum difference among adjacent regions.  $S_{best}(x,y) = 2^k$  where  $k$  maximizes the differences:  $E_k = \max\{E_{k,horizontal}, E_{k,vertical}\}$ .

**Step 4:** Compute a global coarseness value as the average of the  $S_{best}$ :

$$coarseness = \frac{1}{n \times m} \sum_{i=1}^n \sum_{j=1}^m S_{best}(i,j) \quad (7)$$

### Contrast

The local contrast is commonly defined for each pixel as an estimate of the local variation in a neighborhood. It measures the amount of local intensity variation in an image. Contrast also refers to the overall picture quality. Given a pixel  $p=(i,j)$  and neighbor mask  $W \times W$  of the pixel, the local contrast is computed by

$$local\_contrast(i,j) = \frac{\max_{p \in W \times W}(P) - \min_{p \in W \times W}(P)}{\max_{p \in W \times W}(P) + \min_{p \in W \times W}(P)} \quad (8)$$

### Gabor Features

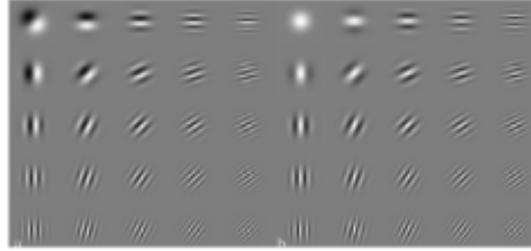
The use of Gabor filters to extract image texture features has been extensively applied in image processing, such as pattern segmentation, texture classification and image recognition. It can be considered as orientations, scale-tunable edges, and line detectors [9].

A 2D Gabor function is defined as

$$g(x,y) = \left( \frac{1}{2\pi\sigma_x\sigma_y} \right) \exp\left[ -\frac{1}{2} \left( \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right] \cdot \exp[2\pi j Ux] \cdot (9)$$

To compute the Gabor texture feature vector, an image  $I(x,y)$  is filtered with a set of scale- and orientation-tuned Gabor filters, as Fig 11. Let  $m$  and  $n$  index the scale and orientation of Gabor filters. Let  $\alpha_{mn}$  denote the mean of the energy distribution and  $\beta_{mn}$  denote the standard deviation of the transform coefficients. If  $M$  is the number of scales and  $N$  is the number of orientations, a texture feature vector can be constructed as

$$\vec{f} = [\sigma_{00}\beta_{00}\sigma_{01}\beta_{01}\dots\sigma_{(M-1)(N-1)}\beta_{(M-1)(N-1)}] \quad (10)$$



**Fig.11** A set of scale- and orientation-tuned Gabor filters.

## Roman Numerals Recognition in scene images

For an optical character recognition (OCR) system, the component-based character locating method plays an important role. However, it is very difficult to locate and extract characters from scene images, these reasons are [10]:

1. The characters are often mixed with other objects.
2. The characters may be of any color, and the background color may differ only slightly from the characters.
3. The font style and size of the characters may vary.
4. The lighting conditions in the images may vary.

It is fortunate that we can find number plates in crime scene images (Fig.12). We can use them to locate the numeral character by shape and segment for numeral recognition.

For Roman numerals recognition, many literatures had been presented [11-16]. Roman numerals recognition systems typically involve two steps: feature extraction and classification. (1) Feature extraction: the patterns are represented by a set of features. (2) Feature classification: decision rules for separating pattern classes are defined. Features can be classified into two different categories: structural features (as strokes and bays in various  $\sigma$  directions, end points, fork points, intersections of line segments, loops and stroke relations, ..., etc.) and statistical features (derived from the statistical distribution of points like zoning, moments, n-tuples, characteristic loci,...). For printed Roman numerals recognition, it has been recognized accurately.

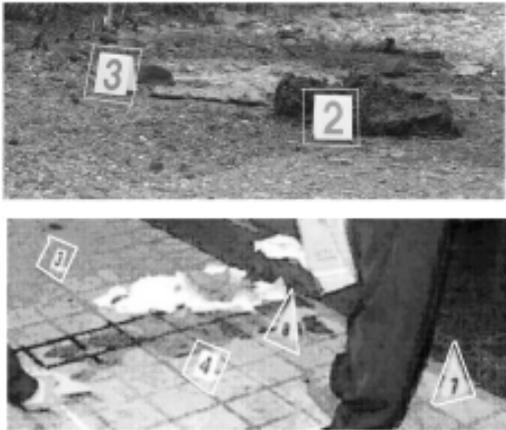


Fig.12 Some number plates in crime scene images.

### Experimental Results and Conclusion

Fig. 13 shows the ten similarity-ranked images with the query image that is retrieved by our method. Color similarity is estimated for the relationship of these crime scene images. By this way, we can check if there are same or similar images in the database.

In Fig 14, the motorcycle is the only red object in the crime scene. The ratio of red color to the whole image can

be used to retrieve images that contain red motorcycle. We can take all motorcycle pictures from various angles. With the ratio of red color to the whole image, we can find all images related to the red motorcycle.

Texture features associated with salient geographic regions can be used to index the image data. Fig. 15 shows the retrieval of crime scene images by the stone texture. The stone texture is a conspicuous feature in the crime scene. We can apply this feature to collect all images with physical evidences over the stones.

In Fig 16, we can label or index these scene images automatically based on the Roman numerals. According to the numerals, the arrangement of the evidence images is easily completed in a sequence.

In this paper, we provide a retrieval method for digital image databases of crime scene photos. We use experimental results to show the capability of the proposed method. We also show the potential of applying the image databases to forensic sciences, such as case data management. The integration of multiple features for better characterizing images is our future work. Besides, we will work on how to figure out the relationship between different cases from the crime scene images.

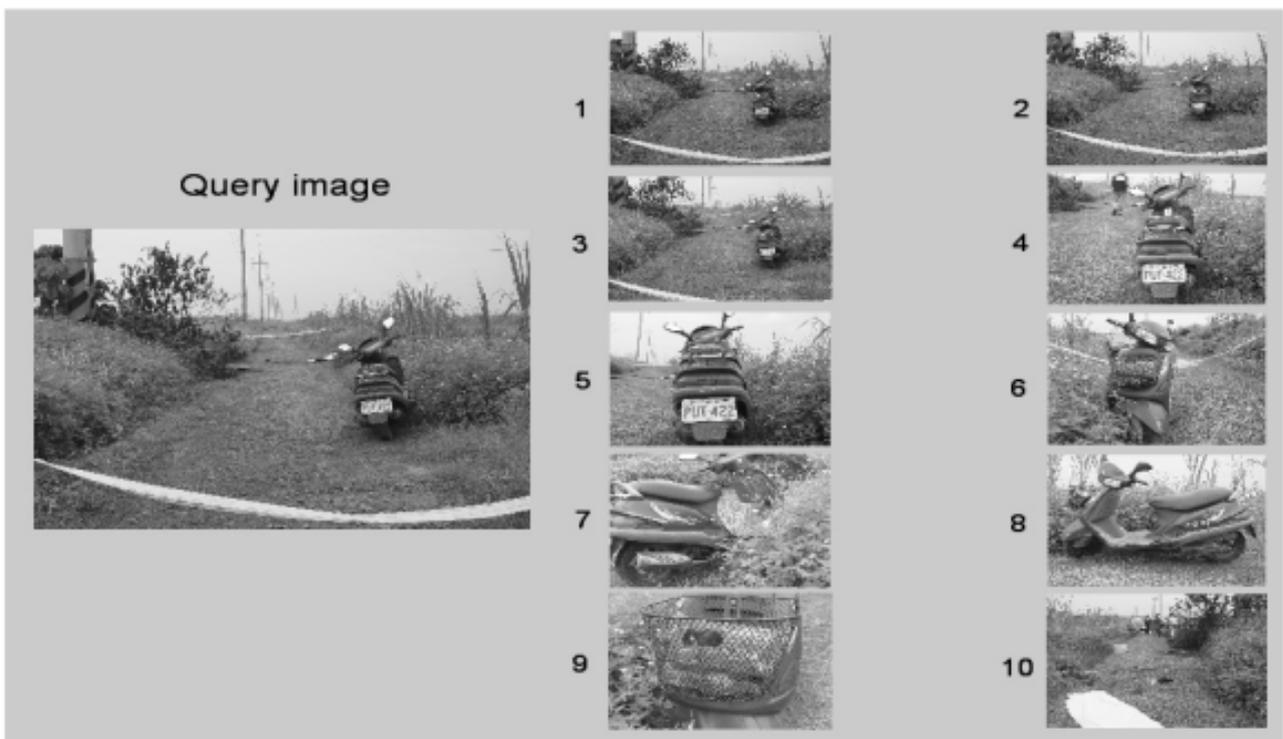


Fig.13 Retrieval of crime scene images by color similarity.



Fig.14 Retrieval of crime scene images by the spatially localized color.

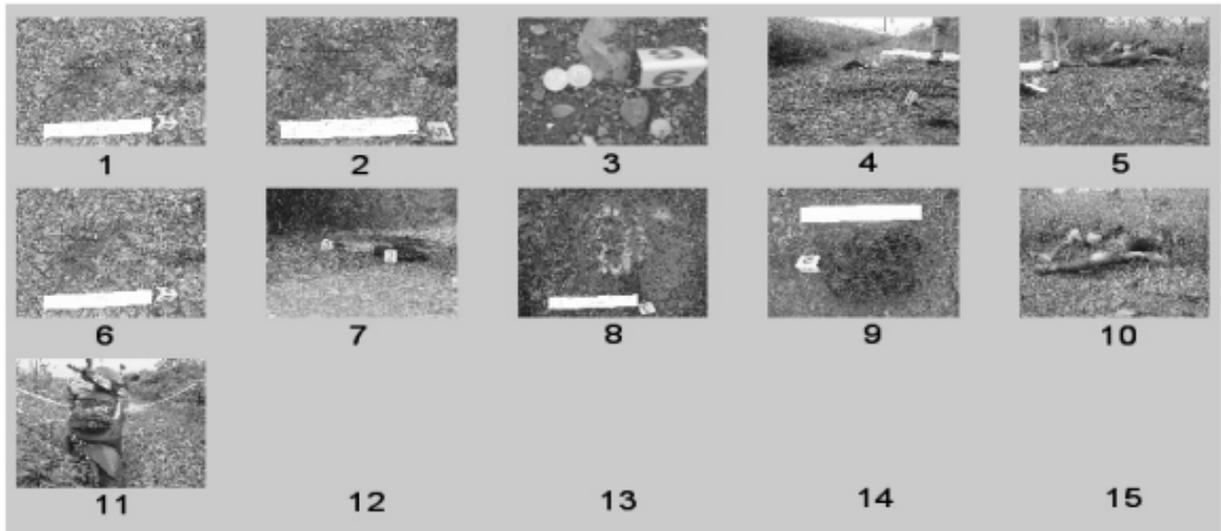


Fig.15 Retrieval of crime scene images by the stone texture.

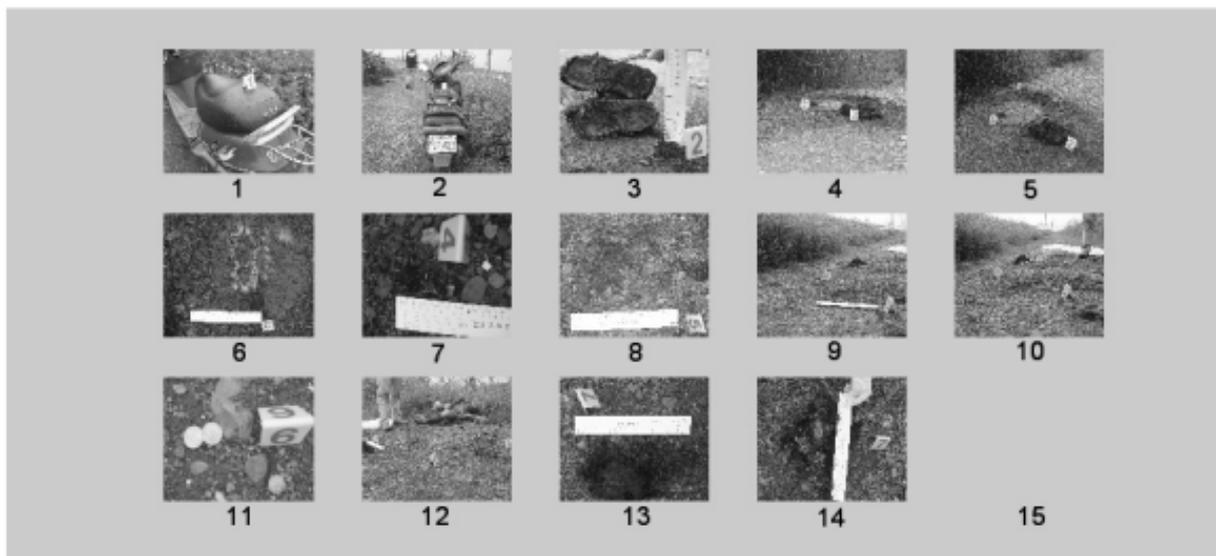


Fig.16 Retrieval of crime scene images by the Roman numeral sequence.

## Reference

1. R.C.L. Wilson and S. Sirohey, Human and Machine Recognition of Faces: A Survey, *Proc. IEEE*, 83(5), 705-740, 1995.
2. C.R. Shyu et al., A Physician-in-the-loop content-based retrieval system for HRCT Image Databases, *Computer Visual Image Understand*, 75, 175-195, 1999.
3. V. Castelli and L.D. Bergman, *Image databases*, WILEY, 2002.
4. A. Del Bimbo and P. Pala, Retrieval by Elastic Matching of User Sketches, *IEEE Trans. Pattern Anal. Machine Intell*, 19(2), 121-132, 1997.
5. Y. Rui, T.S. Hang and S. F. Chang, Image Retrieval: Current Technique, Promising Directions and Open issues, *Commun. Image Represent*, 10 39-62, 1999.
6. M. Stricker and A. Dimai, Color Indexing with Weak Spatial Constraints, in *Symposium on Electronic Imaging: Science and Technology – Storage & Retrieval for Image and Video Databases IV*, Proc. SPIE 2670 IS&SPIE, 1996.
7. J.R. Smith and S.F. Chang, Integrated spatial and feature image query, *Multimedia System*, 7(2), 129-140, 1999.
8. H. Tamura, S. Mori, and T. Yamawaki, Texture features corresponding to visual perception, *IEEE Trans. Sys. Man, and Cybernetics* 8(6), 460-473, 1978.
9. Che-Yen Wen and Chiu-Chung Yu, 2003, "Fingerprint Pattern Restoration by Digital Image Processing Techniques," *Journal of Forensic Sciences*, 48 (5), p. 973-984
10. Chung-Mong Lee, et al., Automatic Extraction of Characters in Complex Scene Images, *Pattern Recognition*, 9(1), 67-82, 1995.
11. L. Heuute, T. Paquet, J. V. Moreau, Y. Lecourtier and C. Olivier, A Structural/statistical feature based vector for handwritten character recognition, *Pattern Recognition Letters* 19, 629-641, 1998.
12. Z. Chi, J. Wu and H. Yan, Handwritten Numeral Recognition Using Self-Organizing Maps and Fuzzy Rules, *Pattern Recognition*, 28(1), 59-66, 1995.
13. Y. Donggang and Y. Hong, Reconstruction of broken handwritten digits based on structural morphological features, *Pattern Recognition*, 34, 235-354, 2001.
14. H. Hishida, Shape Recognition by Integrating Structural Descriptions and Geometrical/Statistical Transforms, *Computer Vision and Image Understanding*, 64(2), 248-262, 1996.
15. H. Jianming and Y. Hong, Structural Primitive Extraction and Coding For Handwritten numeral recognition, 31(5), 493-509, 1998.
16. V. S. Chakravarthy and B. Kompella, The Shape of Handwritten Characters, *Pattern Recognition Letters*, 24, 1901-1913, 2003.