Query by Visual Example  
--- Content based Image Retrieval ---  

Kyoji HIRATA  
*NEC Corporation  
Kawasaki, Japan*  

Toshikazu KATO  
*Electrotechnical Laboratory  
Tsukuba Science City, Japan*

1. Introduction

"A picture is worth a thousand words." Visual information is a good man-machine communication medium. Users expect a multimedia database system to manage visual data as well as alphanumeric data. Users also expect it to provide a visual interface to accomplish flexible man-machine communication in a user-friendly manner.

In this paper, we will present the idea and the general framework for visual interaction with multimedia database systems. The paper will focus on the visual interface for content based image retrieval.

How can such intelligent retrieval be accomplished? Our basic idea is QVE (query by visual example). A user has only to draw a rough sketch to retrieve the original image and the similar images. The system evaluates the similarity between the rough sketch, i.e. visual example, and each part of the image data in the database automatically. This paper will also give the fundamental algorithms for QVE.

We have been developing an experimental database system, called ART MUSEUM, as an electronic art gallery [1, 2, 3, 4]. The ART MUSEUM database is a collection of artistic full color paintings. The algorithms for QVE are implemented and examined on this system. This paper also includes some experimental results and our current evaluation. The algorithms are quite effective for content based image retrieval.

Section 2 gives a brief survey on visual interfaces and summarizes the idea and the general framework for visual interaction. Section 3 describes the primitive algorithms for QVE. These algorithm enable the system to build a pictorial index, based on the composition of paintings, and to find the original image and similar images automatically. Section 4 shows some experimental results obtained on QVE and our current evaluation.

2. Visual Interface for Content Retrieval

Several experimental image database systems have been proposed to provide visual interfaces. Old style image database systems only accept a combination of key words and

---

1 ART MUSEUM: multimedia database with sense of color and composition upon the matter of art
index terms in a user's query. Such key words and index terms were prepared by the database manager. Of course, it is quite difficult to describe the image content with such key words. The personnel expense of assigning key words is also an important factor. The QPE system provides a schema of pictorial data in a graphic form as well as in a tabular form [5]. While this system shows the data in a graphic form, its query style is only a substitute for the query languages on alphanumeric data. In the icon-based system, icons and their two dimensional strings are referred to as pictorial keys to image data [6]. A user can specify the target images by placing icons on the graphic display as a visual query. The system evaluates only two dimensional strings of icons in its retrieval process. Therefore, it is difficult to accomplish similarity retrieval, according to the subjective measure presented by the user. The hypermedia system provides an indexing mechanism for multimedia data in a uniformed style [7, 8]. Although this system enables a kind of subjective indexing, its process owes much to the user's effort on defining many links.

We originally want to process visual data and their contents in visual interaction. On this point, although the above systems use graphic devices to show schema, icons and guidelines, their facilities are not sufficient to allow accomplishing such interaction in a user-friendly manner.

Then, what factors are needed in visual interaction? Essential needs are visual queries for content based retrieval of visual data. The requirements for visual interaction can be summarized as follows.

(a) It is necessary to show visual data to the database in a user-friendly manner. For instance, a user would like to retrieve image data based on database content. Users may provide image data themselves, such as hand-written, rough and partial sketches, as pictorial keys to retrieve some image data.

(b) Of course, it is necessary to view visual data from the database in a simple manner. For instance, a user would like to browse the database guided by some intelligent navigation facility.

(c) Users also expect the system to automatically build the pictorial index for content retrieval. Image analysis facilities map individual parts of image data into their abstracted representations, e.g., abstract images on general composition, according to a user's visual perception process.

These needs can be answered by a visual interface, called QVE (query by visual example). The QVE interface proposed in this paper has the following characteristics;

(i) Simply showing a rough sketch is sufficient to enable retrieving some image data from the system. The powerful pattern recognition algorithms search for the best match candidates on the pictorial index. Currently, the system can accept a hand-drawn rough sketch, a monochrome photo, or a xerographic copy, as well as a full color fair copy as a visual example.