Implementation and Evaluation of a Content-Based Image Retrieval System

賴鼎陞(Ting-Sheng Lai)
國立故宮博物院 資訊中心
台北市士林區 111 至善路二段 221 號
(02) 2881-2021 Ext. 741
sam@npm.gov.tw

Abstract

CHROMA is a fully functional prototype content-based image retrieval system designed as part of the support mechanism for a digital photographic library. The paper outlines CHROMA's colour based indexing scheme and the access system which combines a Thumbnail Viewer, index structure Navigation Tool, a Sketch Tool and Query-by-Image-Example tool. Some recent results of evaluations of the effectiveness and usability of CHROMA and the methodology behind them are then reviewed. Although the evaluations are limited in some ways they show that whether users are attempting to find a specific known image or an image like another one the combination of tools represented by CHROMA is effective and useful.

Keywords: Content-Based Image Retrieval, User-Centered Evaluation

1. Introduction

CHROMA is a novel image retrieval system which provides easy access to collections of general photographic images such as might be found in a digital photographic library or archive. It uses a simple and comprehensible indexing method which facilitates the building of an easy to use user interface. Underlying this is a model of the retrieval process which allows the user to move between analytic query, browsing and similarity retrieval (in effect relevance feedback of a sort) in a natural and flexible manner.

CHROMA uses global colour as its primary indexing key. Images are organised in virtual hierarchical clusters using a set of 10 colour descriptors derived from the work of Berlin and Kay [2]. Users may navigate this hierarchy, use a sketch tool to outline the required image, or select a particular image and retrieve further images similar to the selected one. They may also reorganise the virtual hierarchy to exclude, for example, images which are very dark or to nullify the effect on indexing of very bright areas or reflections. The system is a fully implemented prototype capable of allowing rapid retrieval from collections of thousands of images.

2. The CHROMA Image Retrieval System

CHROMA (Colour Hierarchical Representation Oriented Management Architecture) [9], a full-scale prototype image retrieval system, was developed to run on MS Windows-based personal computers. The system was developed

for use by a large and varied group of users who may need to manage large numbers of photographic images that have not been indexed using manual text-based methods. For example, such images may come from CD-ROM collections, digital cameras or scanned images. Therefore, the system was not intended for use by a single group of users.

The overall architecture of the CHROMA system can be divided into two sub-systems according to their different purposes, these sub-systems are the Indexing and the Access System. The Indexing performs the operations of automatic image indexing. The Access System provides a user interface to accept users' instructions. The Access System then accesses the image database operating the indexing database to output the retrieval results.

2.1 Automatic Indexing Using the Perceptual Colour Model

A unique element of the CHROMA system is the use of a defined "Perceptual Colour Model", used in both the indexing scheme and the design of visual retrieval tools. The definition of this model is illustrated in Table1. Note that the group "0" represents uncertain colours such as dark shadows and bright reflections. A central assumption of the model is that people find it easier to use particular base colours only, such as red, blue, green, etc, rather than variant shades of that base colour. For example, users may simply use "blue" to query for a clear sky instead of using "light blue" or "dark blue".

Evidence from the psychological literature seems to support this view. For example Arnheim [1] suggests that although we can distinguish between subtly different shades, when relying upon our memory for a colour our powers of discrimination are severely limited. In addition, the studies of the linguists Berlin and Kay [2] introduced the eleven basic colour terms, which are denoted by "*" in Table 1.

Table 1. The Perceptual Colour Model

Colour Descriptor	Colour Terms Mapped
0	Uncertain Colours: "very dark" or "very bright"
1	White*
2	Grey*
3	Black*
4	Red*, Pink*
5	Brown*, Dark Yellow, Olive
6	Yellow*, Orange*, Light Yellow
7	Green*, Lime
8	Blue*, Cyan, Aqua, Turquoise
9	Purple*, Violet, Magenta

The indexing scheme generates a 1-D "Class Key" for each image according to the predominance of the major perceptual colours, which is expected to be assignable by human users at least in principle for manually indexing or using as a cue for retrieval. Also, a virtual hierarchical classification exists, which can support structured browsing and navigation. The method was inspired by the concept of the well-known Dewey Decimal Classification [10], and on a method of representation of image quantisation inspired by Harmon's work [6]. A 2-D "Index Vector" is generated to represent spatial information (or local features) in an image using the

perceptual colours, which can be used for similarity measures of images. More details about the indexing scheme can be found in [9].

2.2 The Access System

The Access System was implemented to perform different interactive retrieval strategies based on the use of image indices. The layout of the user interface of the Access System is divided into three windows, illustrated in Figure 1.

The top-left window, called the Navigation Tool, allows users to explore the hierarchy of classification for the image database. The top-right window, called the Thumbnail Viewer, outputs the retrieved images by displaying sets of thumbnail images, which are selected by using the Navigation Tool. The bottom window, called the Information Viewer, displays a list of related information about the retrieved images.

These three tools can be operated interactively, by selecting a specific class with the Navigation Tool, all of the images that are grouped into this class would be listed in the Information Viewer. In addition, the browsing operation of the Thumbnail Viewer can be performed to browse sets of thumbnails by clicking the "scroll buttons" above once a time (i.e. to display previous and next sets).

The design of the Navigation Tool is based on a tree-based directory (or folder) structure, which represents the hierarchical classification. It aims to allow users to navigate specific "image classes" on the basis of their dominant colours. Since the Navigation Tool and the overall layout of the Access System adopted the common design of current file management tools, it is hoped that the adoption of this interface will help users learn to use the tool rapidly. In addition, to assist the user in understanding the visual correspondence between specific groups and photos, the coloured icons which are embedded in the tree structure can assist the user to intuitively identify different classes.

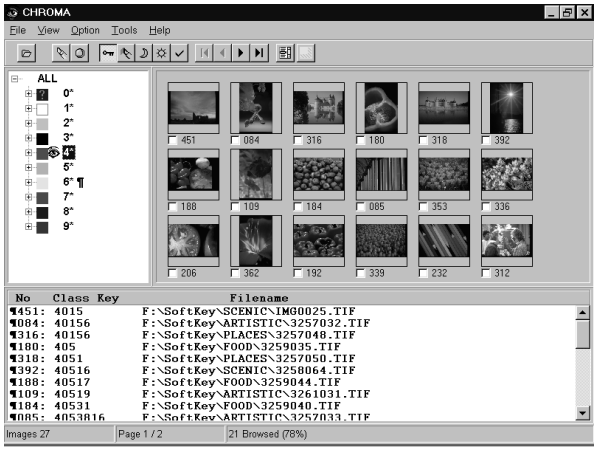

Figure 1. Navigation and Browsing by Hierarchical Classification.

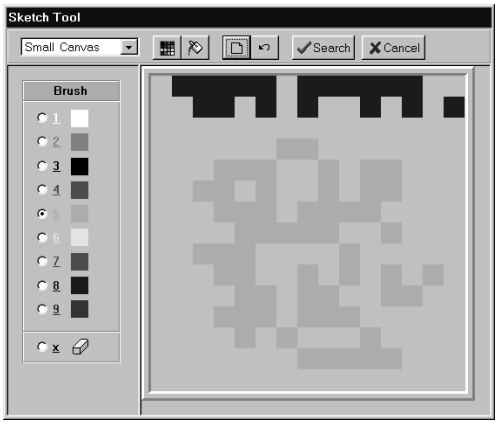

Figure 2. The Sketch Tool and an Example.

2.3 The Navigation Tool

The CHROMA system provides a method for navigating and browsing the hierarchical classification using the Navigation Tool. The initial state of the Access System is to

display all of the images in order of the hierarchical classification. More specifically, the "cursor" of the Navigation Tool is located on "ALL", which means displaying all images. Some current systems, such as QBIC [6], merely display a set of thumbnails in "random" orders in the initial state.

The Navigation Tool performs "Filtering" operations for limiting the scope of browsing and querying images to a specific group, for example mainly blue (sky) with some green (grass, tree). By selecting a specific navigable group, images are filtered and displayed by the Information Viewer and the Thumbnail Viewer. The effects of browsing levels 1 and 2 of the class hierarchy are shown in Figure 1 and 2.

In order to minimize the possibility that users might become lost in the navigable hierarchy we limited the tree structure to two levels of colour groupings. I believe that this would not adversely affect users during navigation and search since it seems likely that they would identify only the most dominant colours in a given image. I also sought to provide navigational support by helping users distinguish between sets of thumbnails they had previously viewed and those that remain to be seen. My inspiration for this feature came from existing methods used to indicate revisitation patterns in hypertext systems. Most current World Wide Web browsers support different colour appearances for distinguishing between links that have been followed and those that have not been examined. The CHROMA system provides the Visited Tag, "¶", to perform a similar purpose to label the viewed images and the viewed classes in both the Navigation Tool and the Thumbnail Viewer. By using the Visited Tags, users may skip tagged classes, this facility is expected to assist users in gradually decreasing the "searching domain" among the whole image database.

2.4 The Sketch Tool

The CHROMA system also supports similarity-based methods for querying images. The Sketch Tool provides a user construction tool to outline the approximate contents in terms of colours and spatial locations of objects in the required image using a paintbrush (or a mouse). This tool aims to allow users to retrieve an existing image which they have seen or used before. The tool also allows users to construct approximate image contents from their imaginations or experience of how a particular scene might look. For example, landscapes could be viewed as having very similar composition.

The layout of the Sketch Tool is illustrated in Figure 2. Note that users are proposed to use only the perceptual colour groups, except the group "0". An example is given to search for images depicting desert scenes by sketching "blue" sky ("8") and "brown" sands ("5").

2.5 Query-by-Image-Example

CHROMA also provides a tool for "*Query-by-Image-Example*", which is a well-known method to allow users to point out a candidate image and the system then retrieves some more images similar to it.

The operation of the method is

performed by the action of “click-to-query”. Users may choose a candidate image among the set of thumbnails on the browser. Since the left button of a mouse is used for the feature of “click-to-enlarge”, the action of selecting the candidate image is conducted by click the right button. By clicking the right button of the mouse, a pop-up menu would then appear and a selectable item, “Query-by-Image-Example”, is available. This design is intend to allow additional features to be added in the future, such as the choice of different algorithms of similarity measures. It also keeps the user interface simple.

3. Evaluation

3.1 Current Evaluation Approaches

In text information retrieval it is conventional to assess the effectiveness of a system using banks of standard queries with previously (perhaps manually assessed) performance for standard test document collections. However the limitations of this approach for interactive retrieval are well-known. In image retrieval the position is worse, since at present there are no standard test collections. Therefore assessing the retrieval of effectiveness of CHROMA (an interactive image retrieval system) presented special challenges.

Traditionally, the effectiveness of information retrieval (IR) systems is evaluated by the standard measures of recall and precision [3]. However, the approach is not without its critics and a number of objections have been raised. These objections are primarily concerned with the fact that precision and recall metrics fail to recognise the importance of the end user, and the impact their behaviour might have on the effectiveness of any interactive system.

A number of investigators have highlighted the advantages offered by the use of user-centred evaluation techniques in image and information retrieval [4, 5, 7, 8]. For example, Dunlop [5] points out that standard precision and recall metrics merely show the retrieval effectiveness of the underlying system and do not take account of user interface and speed issues. Therefore, if such measures were used to compare two IR systems, they would not be able to predict which system would perform best under certain task conditions and with different groups of end users. User-centred evaluation however, allows us to compare system performance with different users and for a large variety of tasks [5]. Jose, Furner and Harper [8] also support this view and stress the need for evaluating real world systems in real world settings. Draper [4] suggests that one of the main advantages of the user-centred approach is that it allows us to observe how people actually use our systems as opposed to how we expect them to be used. Therefore, this type of evaluation may lead to positive system refinements that precision and recall would not have identified.

The use of user-centred measures also allows us to test claims regarding the effectiveness of our systems under a variety of different task conditions and with different user groups. However, there are a number of disadvantages with user-centred evaluation. First, ill conceived or poorly

designed studies can lead to spurious or misleading results. Therefore, great care and attention must be taken during experimental design. Second, conducting a user-centred evaluation is time consuming and labour intensive. However, the wealth of data that this approach provides justifies the additional effort required.

3.2 The Experimental Studies and the Results

The evaluation of the CHROMA system focused primarily on the system's visual search tools. It is important to note the evaluation was not attempting to support the claim that visual search tools are superior to text-based queries for image retrieval. The CHROMA system was intended for use in situations in which manual indexing is neither possible nor practical. Therefore, the evaluation did not compare the CHROMA visual search tools to the text-based query facilities of other systems.

The CHROMA system has a number of visual search tools that can be used both on their own and interactively. Therefore, it was decided to evaluate small units of the system individually rather than attempting to construct one large-scale evaluative study of the system as a whole. This approach mirrors that described by Harper and Hendry [7] who suggest the use of small controlled tasks that are directed at evaluating a given claim or hypothesis.

The evaluation was undertook by means of three studies: a comparison of the Navigation Tool with the Sketch tool; a comparison of the CHROMA Sketch Tool with the Colour Layout Editor of IBM's QBIC [17]; and a comparison between the three primary retrieval tools of CHROMA (Navigation, Sketch and Query-by-image-example) plus thumbnail browsing through the whole collection. (N.B. The pre-April 1999 version of QBIC was used. IBM suggests the current version of QBIC provides a better user interface which might or might not affect the experimental results.)

A special test collection of 1000 images extracted from a commercial royalty free CD-ROM was prepared. Two types of image query sets were used: "*target specific*", in which the users were shown in advance an image and then asked to find it again, and "*target non-specific*" in which they had only an example of a similar image was provided. In each study the usability of the system was assessed as well as the retrieval effectiveness. Usability and retrieval effectiveness are obviously separate issues but they interact in a complex way.

The form and results of each study will be briefly outlined there: a more detailed analysis of the results may be found in [9].

The first experiment aimed to asses whether structured browsing and navigation (as supported by the hierarchical indexing and the Navigation Tool) was a useful alternative to the more common similarity based query method exemplified by the Sketch Tool. There were 24 subjects and the experiment used a within-subjects (repeated measures) design.

The experiment found that the results for two tools were broadly similar, but the Navigation Tool was significantly faster for target non-specific tasks and there was suggestive evidence that the users preferred

to use it.

The second experiment compared the Sketch Tool and that used by the Colour Layout Editor of IBM's QBIC. The purpose of this experiment was to try to eliminate the possibility that the Sketch Tool performed poorly in experiment one because it is a poor example of its kind. 10 subjects worked with each tool in turn. The experiment used a within subject design which also attempted to eliminate possible practice and order effects. Overall there was no significant difference in the user satisfaction with the images retrieved by the two tools, but the users did find the Sketch Tool easier to use.

The final experiment compared the four main retrieval tools, which are provided by CHROMA:

- The Navigation Tool
- The Sketch Tool
- Thumbnail Browsing
- Query-by-Image-Example

for both target specific and target non-specific images. We used a count of the number of times each tool was used by each user in the process of carrying out each assigned task both during initial query formulation and refinement of the query. Again a within-subjects design was used.

12 subjects were used in these tests with five images which existed in the database (target specific) and five more which were not but were similar to those in the database (target non-specific). In all cases a (or the) satisfactory relevant image was found. There was no time limit for these tests. Analysis of the results revealed that there was a statistically significant preference for the use of the Navigation Tool in initial query formulation, and a statistically significant preference for query-by-image-example during query refinement. Furthermore these conclusions hold for both target specific and target non-specific tasks, although the differences between methods is less striking in case of target non-specific tests. Figure 3 illustrates the results of the tests. As can be seen there was almost no use of thumbnail browsing or query-by-image-example in initial query formulation, but these tools were heavily used when in the final stages of searching for the required images.

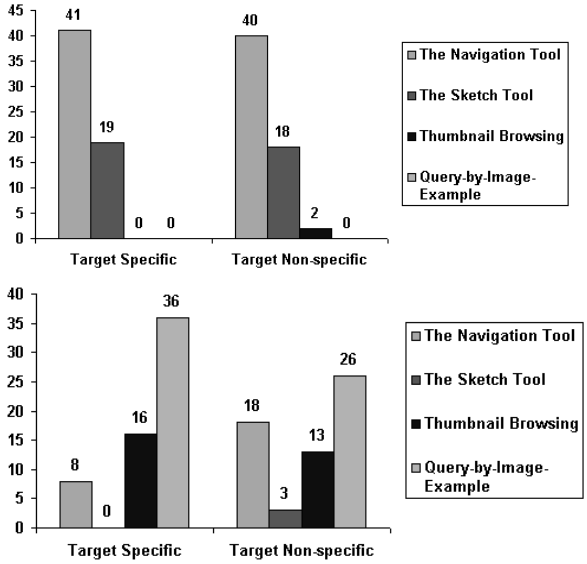

Figure 3. Total Number of Occassions Each Tool Was Used (a) As the First Point of Access to the Image Collection, (b) During Query Refinement Stage.

4. Conclusions

The work with CHROMA has demonstrated the feasibility and effectiveness of constructing a Content Based Image Retrieval tool using fairly simple and easy to understand indexing methods combined with a variety of retrieval tools.

Although the tests are inevitably

limited in scale at the present time, they have yielded statistically significant results. Although it would be useful to conduct more experiments to investigate further such issues as gender differences in interface use, age differences, differences between image collections and more refined notions of target non-specific tasks types, the results obtained so far are sufficiently clear cut that there seems little point in merely conducting the same experiments with more subjects.

Digital photographic libraries are an increasingly important part of the digital library scene, and, indeed, digital images are increasingly important within more general digital libraries. It is often the case that it is not feasible to undertake any more than rudimentary cataloguing of images especially on first acquisition. However undertaking automatic image indexing in ways which effectively facilitates end uses image retrieval has to date proved very challenging. The evaluation work reported here has indicated that the combined use of an easy to understand colour indexing scheme, support for browsing index structures, query-by-image-example and a sketch tool is an extremely promising combination.

Acknowledgements

I would like to thank Prof. John Tait and Dr. Sharon McDonald, University of Sunderland, U.K., for their contributions to the development and the evaluation of the CHROMA system. Thanks also go Dr. Dragutin Petkovic and IBM for offering the QBIC system to be the comparison system for the evaluation.

References

[1] Arnheim, R. Art and Visual Perception: A Psychology of the Creative Eye. The New Version (The expanded and revised edition of the original publication of 1954). University of California Press, Berkeley and Los Angeles, California, 1974.

[2] Berlin, B., Kay, P. Basic Colour Terms: Their Universality and Evolution. The Paperback Edition (The first publication: 1969). University of California Press, California, 1991.

[3] Cleverdon, C. W., Mills, J., Keen, E. M. Factors Determining the Performance of Indexing Systems. Aslib Cranfield Research Project, College of Aeronautics, Cranfield, 1966.

[4] Draper, S. Evaluation in HCI and IR. In: Dunlop, M. (ed.) Proceedings of the Second Mira Workshop. Glasgow University Research Report, 1996. Available at http://www.dcs.gla.ac.uk/mira/workshops/padua_procs

[5] Dunlop, M. The Problems with Precision and Recall. In: Dunlop, M. (ed.) Proceedings of the Second Mira Workshop. Glasgow University Research Report, 1996.

[6] Harmon, L. D. The Recognition of Faces. Scientific American. 229(5), November 1973. pp. 71-82.

[7] Harper, D., Hendry, D. Evaluation light. In: Dunlop, M. (ed.) Proceedings of the Second Mira Workshop. Glasgow University Research Report, 1996.

[8] Jose, J., Furner, J., Harper, D. Spatial Querying For Images Retrieval: A User-Oriented Evaluation. In: ACM SIGIR'98, August 24-28, 1998. pp. 232-240.

[9] Lai, T.-S. CHROMA: A Photographic Image Retrieval System. Unpublished Ph.D. Thesis, University of Sunderland, UK, January 2000. Available at http://osiris.sund.ac.uk/~cs0sla/thesis/

[10] Winslade, B. Introduction to the Dewey Decimal Classification for British Schools. Third Edition. School Library Association. Oxford, 1977.