Image Browsing and Navigation Using Hierarchical Classification

Ting-Sheng Lai

School of Computing, Engineering and Technology, University of Sunderland
Sunderland, UK

John Tait

School of Computing, Engineering and Technology, University of Sunderland
Sunderland, UK

Sharon McDonald

School of Computing, Engineering and Technology, University of Sunderland
Sunderland, UK

Abstract

In this paper, we attempt to address the inadequacies of query-based image retrieval, and argue the importance of structured browsing and navigation methods. We propose to build a hierarchical classification model, which is the basic structure for image browsing and navigation. Currently a prototype system has been implemented, called CHROMA. CHROMA provides both a visual tool to browse or navigate through the image collection, which can be used in combination with existing query-based methods. In addition, by measuring the required iterations of thumbnail browsing we evaluate the improvements in retrieval effectiveness of the browsing and navigation approach.

1 Introduction

In recent years there has been a great deal of research on the development of content-based image retrieval. Current research mainly focuses on querying images by their visual features, such as colour, shape, texture, etc. [6, 16] However, image retrieval is a complicated task which may require the user to implement a variety of search strategies, in addition to the query-based method. It is the aim of this paper to describe and evaluate the use of navigation and browsing tools as an alternative method for image retrieval.

Before going on to discuss the use of browsing and navigational tools, we must first explore and identify the limitations of the query-based strategy. To accomplish this, we will use a scenario based on the retrieval of photographic images from a large CD-ROM collection. Suppose a user had a number of CD-ROMs containing thousands of photographic images. Unfortunately, because text-based indices are too time-consuming and therefore costly to produce, the collection has no text search facility. Instead, the image retrieval system provides visual feature extraction which can be used by some search tools such as “query-by-example” or “query-by-sketch” [6, 16]. However, for the user to be able to retrieve a target image using these facilities they must either have an example of the target image they would like to retrieve, or they must be able to produce a sketch which matches suitable candidates in the collection. Clearly, if the user does not have such information at their disposal, the image retrieval system is rendered useless. One way to overcome these problems would be to provide the user with a structured browsing and navigation system that would allow them to identify candidate images.

1.1 The Need for Browsing and Navigation

The query-based approach is dependant upon the user being able to explicitly and coherently express a precise analytical query. However, there is evidence to suggest that users may not always be able to formulate queries [12] especially in circumstances in which they do not know the likely contents of the collection, or have only a vague idea of what it is they wish to retrieve.

In these circumstances, structured navigation and browsing would be useful to assist users in the construction of a query at both the pre-search stage and at the post-search or refinement stage of query formulation. At the pre-search stage, users may have a vague idea about what the target image looks like. However, they may find it difficult to construct a precise query. Browsing and navigating part of the image collection using hierarchical classification would assist them to refine their queries.

After this initial search period the user may be dissatisfied with the search results, perhaps because a large number of images may have been returned. At this post-search stage, structured browsing and navigation may help the user to refine, narrow and reformulate their query to produce an improved set of relevant images.

Therefore, structured browsing and navigation can be used to iteratively refine queries at all stages of the search process. [11]

Given the problems with query-based retrieval, we argue that structured browsing and navigation should play an important role in image retrieval. For example, Abdel-Mottaleb et al, (1996) point out that browsing and navigation are important, especially when the user may not have a precise query in mind. They suggest that users should be allowed to browse part of the contents. However, to help users refine their queries, it is important to give immediate feedback during browsing and navigation [1].

Batley (1988) suggests that browsing is an essential strategy in retrieving photographs [2]. She confirmed this through two experiments utilising search time and quality of images retrieved. In addition, Markkula and Sormunen [15] point out that, in their investigations “Browsing was used a lot since single-key queries often retrieved large sets of thumbnail images,” and “heavy browsing was necessary when the retrieved sets were large.”

Furthermore, a critical issue for browsing and navigation is how to present information to the user [1]. Del Bimbo [5] points out, “If data structures are employed to store image data, browsing and navigation tools provide facilities for focussing on some part of the information and non-sequential inspection of information content. Tools for defining perspective filters on visual data allow restriction of the interest in an image subset by focussing on a few attributes of images, and placing constraints on their values; tools for browsing and navigation permit the user to inspect the database and visit instances that are relationally linked.”

Although the methods of content-based image retrieval were not involved with the above experiments [2, 15], we propose to address the similar problems associated with the difficulties of constructing a visual query, and extensive browsing for a large amount of visually similar images. Therefore, we suggest that there is a need for the development of browsing and navigation methods to enhance the inadequacy of the use of visual query tools alone. However, little work has been done on the combination of visual query tools and support for navigation and browsing.

1.2 Proposed Methods

We propose a method for image browsing and navigation using a hierarchical classification model, that can improve the retrieval effectiveness in combination with other query processes.

The terms browsing and navigation have various definitions because they are used in a number of different research communities. Therefore, for reasons of clarity we must define what we mean by browsing and navigation more closely. For the purposes of this paper, our definition of the terms hinges on the degree of support provided by the system. By browsing we mean that the system provides no support in terms of the ordering in which the images are encountered. So for example, a picture of a white mouse could follow a picture of a red apple. Therefore, the user must inspect an arbitrarily ordered set of images, which they may choose to either select or reject. By contrast, navigation involves the process of inspecting an image collection that has been ordered according to the image features (e.g. colour, etc). In other words, the user may choose to examine different or distinct parts of the collection rather than viewing the collections as an unsorted or disorganised whole.

We propose a navigational tool to facilitate image retrieval via navigation and browsing by supporting the automatic generation of a hierarchical ordering of the images. Such a system would support a variety of user tasks and search goals that are not currently supported by query-based systems. For example, structured navigation and browsing would be particularly useful to those users who are unclear about the likely contents of the collection, or who do not have a clear search query in mind, or do not have a sample image to work from.

In this paper, we introduce an indexing scheme for building a hierarchical classification model, which is the basic structure for image browsing and navigation. Currently a prototype system has been implemented, called CHROMA. CHROMA provides both a visual tool to browse or navigate through the image collection, and two visual query tools: “query-by-example”, and “query-by-sketch”. In addition, based on measuring the maximum iterations of thumbnail browsing, we evaluate the improvement of the retrieval effectiveness using this approach.

2 The Hierarchical Classification Model

To develop an effective method for image browsing and navigation, we propose to build a hierarchical classification model, which can represent the correlation between different images in the entire collection. Given a suitable interactive user interface, users can navigate through a hierarchical representation of the image collection.

Our recent work [13, 14] introduced a two-stage automatic indexing scheme which models human colour perception.

2.1 Filtering

The first stage is to transform natural colours on photographs into several major perceptual colour groups. We attempted to pick up a few common colours which may be identified intuitively by most people. It aims to define a common colour model to reduce the gap between the colour distribution described by human eyes and the actual colour appearances on photographs.

In 1969, the linguists Berlin and Kay presented the eleven basic colour terms [3]. They argued that there are commonalities of basic colour term use that extend across languages and cultures, and probably express universal features of perception and cognition. In recent years, visual scientists and psychologists met with linguists and anthropologists for the first time to examine how these claims have fared in the light of current knowledge [10].

To design an experimental colour model, we attempt to list a set of ten perceptual colour groups which are defined as shown below in Table 1. We represent each group by using a descriptor (a number). Those basic colour categories which Berlin and Kay present are marked by a star sign “*”.

Note that we introduce a unique colour group “0”, which represents “uncertain” colour. It aims to represent specific colour appearances that humans generally fail to identify correctly. It contains “very dark” colours that seem “black”, and “very bright” colours that seem “white” for human eyes. For example, a photo of night scene might contain a large area of “black” sky, and yellow or orange colours of road lamps. However, the sky may not be “pure black”; it may be an “unsaturated” red colour because of the influence of the road lamps. To take an other example, the sky in daylight in a photo may seem white; however, it might really be “very light blue”.

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

Table 1: The Perceptual Colour Groups Defined

To group various real-world colours into the perceptual colour groups for implementing the above theory, we attempt to subdivide the Smith’s (1978) HSV colour space [8] according to the definition of the boundary between different groups. Recent research present various definition of ranges [4, 9]. We emphasis that the differences would not have a great influence on our experiments.

Though the definition of the perceptual colour groups do not completely adopt the current research on linguistics and image retrieval; however, in the future we propose to allow users to adjust the definition of the colour groups according to their own colour perception if a suitable user interface is provided.

Furthermore, we found that using the perceptual colour groups can also overcome the problem of different lighting conditions [13, 14]. For instance, searching for a clear sky could simply query for the colour blue, no matter whether it is light blue or dark blue.

2.2 Indexing

The second stage is to generate an expressive class key for images for global colour distributions of image contents. The indexing key should be able to identify a specific class of images, and should be assignable both automatically and manually at least in principle (to ensure it is comprehensible to users).

The primary role of the indexing method is to add colour descriptors in the order of the predominance of colours. The predominance of a colour depends on the amount of a colour covering in an image, which are decided by calculating the colour histograms [9] of the colour groups.

Initially, the descriptor of the most dominant colour is added, and then the less dominant colour follows. As each digit is added, the class key becomes a more and more expressive representation of the colour distributions of an image. In other words, the colour covering the most area of an image is regarded as the main colour feature. Its colour group number is represented by the first digit of the class number. The second digit of the class number is added by the colour group number of the colour which covers the second largest area of an image. Then the third digit is added in the same way. Then the rest of the digits follow.

For instance, we attempted to index a picture of a sunset scene. The content consists of an orange sky, black clouds and a bright sun. The class number would start with “4”, which represent the large area of red sky (cf. Table 1). Following the first descriptor, “3” is added to represent the shape of black landscape. The third digit would be “1” representing the near white sun. Thus, the class number would become “431”.

2.3 Hierarchical Representation

Strictly speaking, this indexing scheme does not group similar images into specific clusters, but merely generates a 1-D vector from each image. However, this vector (the index key) contains the intrinsic value of a hierarchical content representation using the colour descriptors. In addition, we point out that an invisible hierarchical classification exists, though every index key in the collection is neither connected nor are clusters successively linked.

Given a suitable human-computer interface, users are able to navigate through the image collection with the “dynamic” hierarchies which is generated on-line. It allows users to adjust the classification structure according to their preferences. For instance, users may like to omit specific colours to adapt different search tasks (See section 3.2), or change the definition of perceptual colour model by their own colour perception.

3 CHROMA: A Prototype System

A prototype system, called CHROMA, has been implemented. It provides a navigation tool using the hierarchical classification model, and also two visual query tools, “query-by-example” and “query-by-sketch”. The user may employ either the navigation tool, or the query tools alone. They may also use a combination of navigation and querying for retrieving images.

3.1 The Interactive User Interface

The principal concept behind CHROMA is to provide the user with an effective interface for navigating through the image collection. To represent the perceptual colour classification by a tree structure, the retrieval interface of CHROMA adopts the common design of current file management tools, which allow users to navigate the tree-based directories (or folder) structure. It is hoped that the adoption of this interface will help users learn to use the tool rapidly.

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 groups. Given time the user can easily navigate through the entire image collection using the visual navigation tool.

Figure 1 illustrates the layout of the user interface which is divided into three windows or sections. The top-left window is the navigation tool as described above. The top-right window is the content area which displays the thumbnail images which are clustered into a specific group. The user can browse through the retrieved set of thumbnail images using the scroll buttons page by page. In addition, the bottom window displays a list of information for the retrieved set of images which are clustered into this specific group.

In addition, we found that most people only identify the two or three most predominant colours in an image. To design an easy-to-use interface, at present the navigation tool provides the top two levels of colour groups. The navigation tool may be regarded as an “information filter” for limiting the scope of browsing and querying images to a specific cluster.

3.2 Refinement of Hierarchical Classification

Our research suggests that there is a need to improve the processing of shadows and reflections in photographs. For instance, different lighting conditions result in the variations of colour saturation. These result in errors in the similarity measure of colour matching. In addition, the occurrence of shadows and reflections have a great influence on measuring sizes and contours of images [13, 14]. These would result in the incorrect effectiveness of classification.

In order to improve the above problem, we introduce a unique method to refine the original clusters. This retrieval task attempts to simulate the human eye’s ability to ignore unexpected colours. Since the unexpected colours are excluded temporarily the hierarchical classification could be refined.

Figure 1: The Layout of the User Interface

Figure 2: Refinement of Hierarchical Classification

For example, we attempted to find a group of images which are mainly constructed by blue sky and green grass. However, our initial results showed that the group “87*”, which is combined blue “8” and green “7”, contained only one image. Other “similar” images are grouped into inappropriate clusters because they contain dominant unsaturated shadows in sunlight. Navigating through the group “0”, we found that there are 193 images containing dominant “unsaturated” colours. This confirms our expectation.

Therefore, we may temporarily exclude all images within the colour group “0” out of the hierarchical classification, and then the classification was dynamically reconstructed. The experimental results revealed that the method provides a significant improvement of effectiveness.

Figure 2 illustrates the clustering results after refinement. There are 16 images which were re-grouped into the group “87*”, and the retrieval effectiveness seem reasonable. Some image containing dominate appearance of shadows which the original class key were “087*” and “807*” are clustered into the group “87*.”

3.3 In Combination with Existing Query Methods

To improve the retrieval effectiveness of the combination of the browsing and navigation tool and query-based method, CHROMA provides two image search strategies, “query-by-sketch” and “query-by-example”, which exist in current systems, such as QBIC[7], VisualSEEK[17]. Our recent work [13] presented the methods of operation of these query methods.

4 The Results

In this section, we propose to assess the usability of the system utilising a large image collection from a CD-ROM, which contains 5777 images in 37 categories [18]. It is subjective judgement whether the scale of the test image collection is sufficient. However, if the user browses a set of 8 or 18 thumbnails on each iteration, the user may need to make at most 723 or 321 iterations to browse through the entire collection, which is an extensive task. It is likely therefore that the user needs the assistance of effective navigation for structured browsing.

We attempt to validate whether the hierarchical classification model can assist the users in browsing and navigating a image collection. The first part of validation is to assess the effectiveness of the hierarchical classification. The second part is to assess the differences in browsing efficiency with and without using the navigation tool.

4.1 Assessment of Classification Effectiveness

Table 2 shows the results of classification using CHROMA, which presents the first two levels of the hierarchical classification model. Level 1 represents the image collection classified by the predominant colour of each image. Level 2 represents the collection classified by the two most predominant colours. Further levels could be considered by the system, but two will suffice to illustrate our point here. The values in the table represent the number of the retrieved images for each cluster. We assume initially the user will move through the clusters using the navigation tools to order to obtain an overview of the contents the complete image collection. We wish to determine whether our classification could provide users with an effective means of doing this.

Level 1	Level 2											Total
	N/A	0	1	2	3	4	5	6	7	8	9	
0*	0	N/A	162	43	93	174	521	54	192	256	26	1521
1*	0	104	N/A	49	31	18	530	10	56	61	1	860
2*	0	22	13	N/A	28	7	19	0	12	26	2	129
3*	0	71	28	16	N/A	7	22	0	14	30	1	189
4*	0	126	36	5	8	N/A	65	10	12	12	14	288
5*	0	410	64	25	46	61	N/A	45	77	69	1	798
6*	1	30	9	0	2	10	45	N/A	7	2	0	106
7*	1	226	40	11	44	8	78	24	N/A	98	2	532
8*	22	526	211	101	120	29	110	7	150	N/A	27	1303
9*	0	20	2	2	3	14	1	0	2	7	N/A	51

Table 2: Results of Classification

Level 1	Level 2										Total
	N/A	1	2	3	4	5	6	7	8	9	
1*	0	N/A	70	77	50	601	20	73	119	12	1022
2*	0	20	N/A	46	13	28	0	17	45	3	172
3*	0	50	36	N/A	23	61	5	51	55	1	282
4*	0	129	11	22	N/A	165	36	28	30	41	462
5*	1	307	70	170	236	N/A	190	187	154	4	1319
6*	1	28	0	4	23	85	N/A	16	3	0	160
7*	8	95	26	138	27	182	84	N/A	162	2	724
8*	32	448	176	271	68	236	14	258	N/A	56	1559
9*	1	9	4	5	36	2	0	6	14	N/A	77

Table 3: Results of Classification After Refinement (Group 0 Omitted)

Subject	Classification
Fireworks	“34*”, “35*”, “36*”
Sunsets	“46*”, “64*”
Landscape	“78*”, “87*”
Red flowers	“47*”, “74*”
Yellow flowers and leaves	“67*”, “76*”
Purple flowers	“97*”
Underwater	“8”
Air show	“81*”

Table 4: Results of Classification

However, we found that the group “0*” is the most dominant group among this image collection. It confirms our observations of the importance of intrinsic colour appearance for general photographic images [14]. It is likely that few people have such excellent colour perception that they can easily identify “very dark” or “very bright” colours. Therefore, we recommend users omit the group “0” while navigating images by the hierarchical classification. After omitting the group “0”, the hierarchical classification was refined. The more natural effects of classification after refinement are shown in Table 3.

In Table 3, note that there are 8 clusters containing “no retrieved images”; these clusters will not appear in the user interface. Thus, there are 71 navigable clusters are accommodated in the navigation tool. To simulate the actual search task, we attempt to browse and navigate through all of the clusters inspecting one page of thumbnails for each cluster. It takes approximately 15 seconds to browse a set of 18 thumbnails, and takes approximately 20 minutes to “skip” through the entire collection. We found that this collection may at least contain the following possible groups of photos, which are as illustrated in Table 4.

In addition, photo in the “5*” groups contain some objects about skin colours, animals, rocks, buildings, sands, land, etc., which are “brown-like” colours. Those photos whose backgrounds are “blue” sky or “blue” sea mainly contain the blue group “8”. Some photos which have white background are located in “1*”. In the group “2”, however, it seems that many photos are shot in dim lighting conditions; thus the colours of objects are dull. We also found “purple” is unusual among the photos. Some sunset photos have purple-like sky, which have visual impact.

Informal testing indicates the retrieval effectiveness is reasonable.

Figure 3: Numbers of Trials for Level 1 Groups

Figure 4: Numbers of Iterations for Level 1 and Level 2 Groups

4.2 Assessment of Browsing Efficiency

The baseline for measuring the performance of browsing is to predict the maximum iterations for browsing through the search space. We assume that it is acceptable for users to have to inspect up to 10 sets of thumbnails while browsing, and that it is progressively more unacceptable to have view more than 10 sets.

Without using the navigation tools, the user can only browse through the entire image collection linearly. Assuming the user browses through the collection by sets of 18 thumbnails, the numbers of the maximum iterations the collection is 321. On average the user will only hit the specific image by chance in 161 iterations, which is too extensive to be acceptable.

Next, we compared the performance by using three different numbers of simultaneously visible thumbnails, 8, 18, 50. Different size of thumbnail set would have an influence on the efficiency of browsing.

Providing the navigation tool, the average number of iterations would be reduced provided the user can make correct navigation decisions. Figure 3 shows the reduced numbers of browsing iterations for different sets of thumbnails. Filtering by the level 1 group, the numbers of the maximum iterations are reduced greatly. However, the group “1”, “5”, “8” still contain a large number of retrieved images. Furthermore, to narrow down the information search space more, the first two groups 1 and 2 are used. The numbers of the maximum browsing iterations are shown in Figure 4. Overall, the numbers of browsing iterations are under 10.

4.3 System Evaluation and Usability Testing

The CHROMA system is currently undergoing user centred evaluation using an image set of 1,000 images. The evaluation will proceed in three stages. First, the CHROMA sketch tool will be evaluated in comparison with the IBM QBIC sketch tool. This study will use a repeated measures (within-subjects) design. The independent variable is sketch tool (CHROMA vs. QBIC), the dependent variables are the number of correct searches and usability satisfaction ratings obtained through the administration of a post-test questionnaire. The experimental task consists of typical user scenarios and target search sets. The second stage of the evaluation process will focus on the evaluation of the CHROMA navigation tool. The navigation tool will be evaluated in isolation since there is no other currently available navigation tool based on automatic clustering. This second study will also use scenarios, target search sets and the post-test questionnaire. The final stage of the evaluation will involve participants performing the same type of experimental tasks as the first two studies, however for this study users will have access to both the CHROMA sketch tool and the CHROMA navigation tool. This will hopefully allow us to determine how the tools are used together and will provide further information about user search preferences.

5 Conclusion and Future Work

To summarise, this paper suggests that there is a need to develop the methods of browsing and navigation to overcome the inadequacy of querying images alone.

The initial usability validation reveals that the effectiveness of the hierarchical classification model is reasonable. The navigation tool of CHROMA which is the visualised implementation of the model may provide assistance for navigation.

In the future, we propose to build a customisable user interface to allow user's preferences for adjusting the perceptual colour model. In addition, we also suggest that the query methods of text-based image retrieval can be integrated into this system to provide other forms of query support.

References

1. Abedl-Mottaleb M, Wu H-L, Dimitrova N. Aspects of multimedia retrieval. *Philips Journal of Research* 1996;50: 227-251
2. Batley S. Visual information retrieval browsing strategies in pictorial data. PhD thesis, University of Aberdeen, Aberdeen, UK, 1988
3. Berlin B, Kay P. Basic color terms: their universality and evolution. The Paperback Edition. University of California Press, California, 1991
4. Cox I, Miller M, Omohundro S, Yianilos P. Target testing and the PicHunter Bayesian multimedia retrieval system. *Proceeding of 3rd International Forum on Research and Technology Advances in Digital Libraries*, IEEE, 1996, pp 66-75
5. Del Bimbo A, Guest Editor. Image and video databases: indexing, browsing and retrieval, *Journal of Visual Languages and Computing* 1996;7: 353-359
6. De Marsicoi M, Cinque L, Levialdi S. Indexing pictorial documents by their content: a survey of current techniques 1997;15: 119-141
7. Flickner M, Sawhney H, Niblack W, et al. Query by image and video content: the QBIC system. In: Maybury M (ed) *Intelligent multimedia information retrieval*. AAAI/MIT Press, 1997, pp 1-22
8. Foley J, van Dam A, Feiner S, Hughes J. *Computer graphics: principles and practice*, 2nd Edition in C. Addison-Wesley, 1996
9. Gong Y, Chua H, Guo X. Image index and retrieval based on color histograms. *Multimedia Tools and Applications* 1996;2: 133-156
10. Hardin C, Maffi L (eds). *Color categories in thought and language*. University of Cambridge, Cambridge, 1997
11. Jorna G, Wouters M, Gardien P, et al. The multimedia library: the center of an information rich community. In *ACM CHI 97 Electronic Proceedings: Design Briefings*. Available at <http://www.acm.org/sigchi/chi97/proceedings/briefing/gj.htm>
12. Jul S, Furnas G. ACM CHI 97 Workshop Report: Navigation in electronic worlds. in *SIGCHI Bulletin*. 1997;29: 44-49

13. Lai T-S, Tait J. General photographic image retrieval simulating human visual perception. Paper presented at the 21st Annual International ACM SIGIR Conference Post-Conference Workshop on Multimedia Indexing and Retrieval. Melbourne, Australia, 28 August, 1998. Available at <http://osiris.sund.ac.uk/~cs0sla/mir/mir98.htm>
14. Lai T-S, Tait J. Using global colour features for general photographic image indexing and retrieval. Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Melbourne, Australia, 24-28 August, 1998, pp 349-350
15. Markkula M, Sormunen E. Searching for photos – journalist’s practices in pictorial IR. The Challenge of Image Retrieval: A Workshop and Symposium on Image Retrieval. University of Northumbria at Newcastle, UK, 5-6 February, 1998
16. Rui Y, Huang T, Chang S-F. Image retrieval: past, present, and future. Accepted to Journal of Visual Communication and Image Representation, 1998, Available at <http://www.ifp.uiuc.edu/~yrui/ps/jvcir98.ps.Z>
17. Smith J, Chang S-F. Querying by color regions using the VisualSEEK. In: Maybury M (ed) Intelligent multimedia information retrieval. AAAI/MIT Press, 1997, pp 23-42
18. Software Partners Ltd. Launched at CD-ROM 82, PC Preview, Future, June 1998