Image Database Navigation: A Globe-Al Approach

Gerald Schaefer and Simon Ruszala

School of Computing and Informatics,
Nottingham Trent University, Nottingham, United Kingdom
gerald.schaefer@ntu.ac.uk

Abstract. Image database visualisation and navigation tools become increasingly important as image collections keep ever growing. Demanded are easily navigable and intuitive ways of displaying and browsing image databases allowing the user to view images from a collection that facilitates finding images of interest. In this paper we introduce a way of viewing a complete collection of images by projecting them onto a spherical globe for colour-based image database navigation. Taking median hue and brightness of images, features that are useful also for image retrieval purposes, and using these as a set of co-ordinates which then determine the location on the surface of the globe where the image is projected. Navigation is performed by rotation (e.g. choosing a different hue range) and zooming into areas of interest.

Keywords: image database navigation, image database visualisation, content-based image retrieval.

1 Introduction

Due to the large increase in use of digital capturing equipment for both personal and professional use, there is currently a large demand for ways of storing and exploring these image databases. With the size of collections ranging from the average home user owning around 1,000 images to companies with databases in excess of 1,000,000 images, efficient and effective ways of locating and searching for desired images are in high demand. Presently most tools display images in a 1-dimensional linear format where only a limited number of thumbnail images are visible on screen at any one time, thus requiring the user to search back and forth through pages of thumbnails to view all images. Obviously, this is a very time consuming, impractical and exhaustive way of searching images, especially in larger catalogues. Furthermore, the order in which the pictures are displayed does not reflect the actual image contents and hence cannot be used to speed up the search.

Some approaches which provide a more intuitive interface for image database navigation were recently introduced. Often the images are projected onto a 2-dimensional plane represented typically by the screen. Using multidimensional scaling, images can be arranged on screen in such a way that images that are

G. Bebis et al. (Eds.): ISVC 2005, LNCS 3804, pp. 279–286, 2005.

visually close to each other are also located close to each other on the display [4] hence providing an easy way for the user to zoom into an area of interest. Other methods incorporate the application of virtual reality ideas and equipment to provide the user with an interactive browsing experience [3].

In this paper we present a simple and fast approach to image database navigation. All images are projected onto a spherical globe; navigation through the image collection is performed by rotation of the sphere and zooming in and out. The use of a spherical object is not a coincidence, rather it stems directly from the type of features that are used for navigation. We utilise the median hue and median brightness (in HSV colour space) to calculate a pair of co-ordinates for each image in the database. As hue describes a circular quantity ($0^\circ = 360^\circ$) whereas brightness is not, a sphere seems a natural choice of geometrical body to encapsulate the combination of these two.

The rest of the paper is organised as follows: Section 2 describes some related work in the area of image database visualisation and navigation. Our proposed method is introduced in Section 3 whereas Section 4 describes how this system is used for database navigation. Section 5 concludes the paper.

2 Related Work

In here we briefly describe two previous methods for image database navigation: multi-dimensional scaling and MARS 3D. A more detailed review which also includes other approaches is given in [5].

2.1 Multidimensional Scaling

Rubner *et al.* were one of the first to suggest more intuitive interfaces for image database navigation [4]. They suggested the application of multidimensional scaling (MDS) [2] to calculate the locations of images and displaying them in a global 2-dimensional view on a single screen. Using this method all images in a database are (initially) shown simultaneously; their locations are dependent on their visual similarity (based on features such as colour, texture or shape descriptors) compared to all other images features in the database. If two images are very similar in content they will also be located close to each other on the screen and vice versa. The user can browse the database easily from a top-down hierarchical point of view in an intuitive way.

The main disadvantage of the MDS approach is its computational complexity. First a full distance matrix for the complete database, i.e. all pairwise distances between any two images in the collection, need to be calculated. MDS itself is then an iterative process which constantly rearranges the locations of each image minimising the (Euclidean) distances between images on screen and their actual (feature-based) database distances. Interactive visualisation of a large number of images is hence difficult if not impossible to achieve. Furthermore, adding images to the database requires re-computation of (part of) the distance matrix and rerunning MDS.

2.2 3D MARS and ImageGrouper

3D MARS [3] represents a virtual reality (VR) approach to image database navigation. Images are displayed on four walls in an immersive CAVE where browsing the database is done using a special wand to select images and make queries. While being an interesting novel idea it requires very specialist equipment not available for the average user wanting to explore their personal collections.

Although there exists also a desktop version called ImageGrouper which can be used with ordinary CRT displays, using this program is fairly difficult as there is no option for global browsing; only searching for an image first and then browsing from that location in the database is possible.

3 Visualisation on a Hue-Based Globe

Our aim is to provide a simple yet effective interface for image database visualisation and navigation. It should allow fast and intuitive browsing of large image collections and not rely on any specialised equipment not available to the average user.

We separate our approach into two tasks: selection of suitable features and visualisation of the image collection. Among typical features used for image retrieval those describing the colour content are certainly the most popular ones [8]. We follow this and describe each image by its median colour. However rather than employing the standard RGB colour space we use the HSV space [6]. Of this we take only the hue and value attributes as the saturation descriptor is deemed less important for describing image content. Value, which describes the brightness of a colour is defined in HSV as [6]

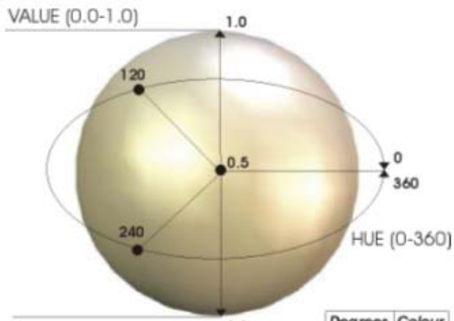
$$V = \frac{R + G + B}{3} \tag{1}$$

where R, G, and B are red, green and blue pixel values. V ranges between 0 and 1 where 0 corresponds to pure black and 1 to pure white. Hue "*is the attribute of a visual sensation according to which an area appears to be similar to one of the perceived colours, red, yellow, green and blue, or a combination of two of them*" [1] and is the attribute that is usually associated as 'colour'. Hue in HSV is defined as [6]

$$H = \cos^{-1} \frac{0.5[(R-G) + (R-B)]}{\sqrt{(R-G)(R-G) + (R-B)(G-B)}} \tag{2}$$

It is apparent that hue constitutes an angular attribute; H goes from red to yellow to green to blue back to red and is also often referred to as hue circle.

We now want to find a geometrical body that can be used to derive a coordinate system for placing thumbnails of the images contained in a database as well as to serve as the actual surface onto which those thumbnails are projected. Looking at the two attributes we have selected, H and V, we almost naturally end up with the body of a sphere, or a spherical globe. The hue circle describes



Degrees	Colour
0 =	Red
120 =	Blue
240 =	Green
360 =	Red

Fig. 1. Hue-Value co-ordinate system used

one dimension of the sphere. As all colours with high V values are similar i.e. close to white and the same holds true for those colours with low V which become similarly close to black, and as black and white by definition don't have a hue quality, the two points $V = 0$ and $V = 1$ describe the poles of the globe.

The use of a globe not only comes naturally with the choice of features, it also has other clear advantages. The concept of a globe will almost certainly be familiar to the average user as it is a direct analogy of the earth globe. It therefore provides a very intuitive interface to the user who will have experience on how to navigate and find something on its surface. Furthermore it allows us to employ a familiar co-ordinate system based on latitude and longitude. Longitude describes the circumference of the globe (i.e. the east-west co-ordinate) and lies in the interval $[0°, 360°]$. Latitude describes the north-south direction from pole to pole and ranges from $-90°$ at the south pole to $+90°$ at the north pole. Clearly each point on the surface of the globe can be uniquely described by a pair of longitude/latitude co-ordinates.

Our approach to visualising image databases is simple and straightforward. It also provides a very fast method for accessing the image collection. Images are transformed to an HSV representation and the median hue and value attributes are calculated. We make use of the median rather than the mean in order to provide some robustness with regards to the image background. Also, we actually calculate the median R, G, and B values and transform those to HSV which is computationally more efficient and results only in a slight deviation in terms of accuracy. From these the co-ordinates on the globe are then determined (see also Figure 1), where H directly translates to longitude and H is rescaled to match the latitude range ($V = 0.5$ corresponds to $0°$ latitude, i.e. a position on the equator). A thumbnail of the image is then projected onto the surface of sphere at the calculated co-ordinates[1]. Once the locations for all images have been

[1] Since those thumbnails are distorted more towards the poles we actually remap all images to be located between the great circles with $\pm 80°$ latitude.

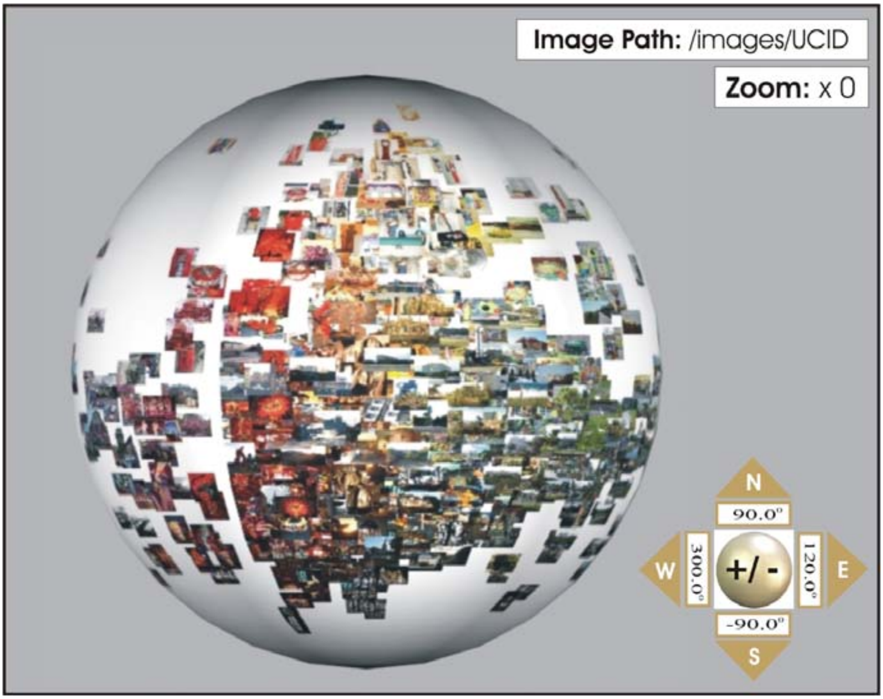


Fig. 2. Initial globe view of the UCID dataset

calculated and their thumbnails projected the system shows an initial view of the globe and is ready for navigation as detailed in Section 4. Figure 2 shows the initial view of our image database globe based on the UCID [7] image database which contains some 1300 images.

Since the co-ordinates are extracted directly from the images our approach is very fast and hence much more efficient than methods such as MDS which rely on a comparably slow and computationally expensive iterative convergence procedure. Also does the addition of images to the database require only the computation of its median hue and brightness with no further calculations necessary as is the case for MDS and similar approaches. In addition are the features that we employ intrinsically suitable for query-based image retrieval [9]. Furthermore, the axes of the co-ordinate system are well defined and meaningful which is in contrast to those obtained by MDS or other similar techniques (such as principal components analysis) where axes are not associated with attributes. Finally, as the database globe is displayed on an ordinary CRT monitor no specialist equipment is necessary.

4 Image Database Navigation

As mentioned above the interface starts with an initial view such as the one shown in Figure 2 (which has a central point of 30° longitude, i.e. an average

Fig. 3. Globe view with central hue of 120° (left), 210° (middle), and 300° (right)

reddish/yellowish hue, and 0° latitude, i.e. medium brightness). From Figure 2 we can also see the controls the user has at his/her disposal for navigation purposes. As these are kept fairly simple and again due to the average user's familiarity of localising places on an earth globe, navigation is straightforward and intuitive.

The controls allow rotation of the globe around both the vertical and the horizontal axis and zooming in and out of a region of interest. The user can perform these operations using the keyboard where the cursor keys are translated to rotation of the sphere and the + and - keys are available for zooming. Similarly the mouse can be used for navigation: moving the mouse prompts a rotation whereas the left and right mouse buttons are responsible for zoom operations. Especially attractive is the application of a trackball: here the ball itself corresponds directly to the image globe, hence rotating the track ball rotates the sphere on screen whereas the two buttons are again used for zooming in and out.

From Section 3 it becomes clear that a rotation around the vertical axis will focus on a different average hue of the displayed images. In Figure 3 we show the view after rotating it clock-wise in 90° steps. While Figure 2 was centred around a red/yellow hue, the spheres in Figure 3 show the display centred around (from left to right) greenish, bluish, and magenta hues. In contrast, rotation around the horizontal axis will either shift the display to images that are darker (rotation towards south pole) or brighter pictures (rotation towards north pole).

While the global view of the sphere allows for easy selection of a general hue/brightness area, it is clear that in this view single images are hard to make out. Therefore a zoom function is provides which allows to user to restrict their attention to a smaller, more localised area of interest. In Figure 4 we show an example of a zoomed-in area where the user was interested in images with a beach/ocean view. It is clear that futher zoom operation(s) can be applied to localise single images of interest.

We note, that for none of the operations used for browsing the images, i.e. neither for rotation nor for changing the zoom factor, any additional calculations need to be performed (in contrast to e.g. MDS which will recalculate all co-ordinates for a zoomed-in area [4]) as the co-ordinates at which the pictures are placed do not change. We are therefore able to provide an image browsing environment that can operate in real time. Furthermore, while we provide a dedicated application for image database navigation we can also export our

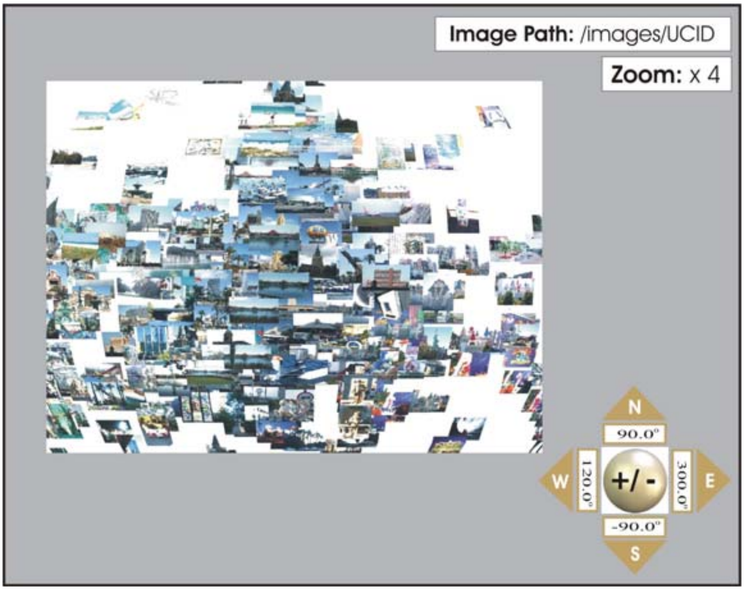


Fig. 4. Zoomed-in area displaying beach/ocean images

image globe to a VRML model [10] which can them be used and browsed with a suitable viewer.

5 Conclusions

We have introduced an efficient and effective approach to visualise and browse large image collections. Thumbnails of images are projected onto a spherical globe which acts as the medium users interact with. Navigation is simple, intuitive, and fast. While we have used colour features to determine the position of the thumbnails on the sphere's surface, the method is generic and other types of features can be equally employed.

References

1. CIE. *International Lighting Vocabulary*. CIE Publications 17.4, Commission International de L'Eclairage, 4th edition, 1989.
2. J.B. Kruskal and M. Wish. *Multidimensional scaling*. Sage Publications, 1978.
3. M Nakazato and T.S. Huang. 3D MARS: Immersive virtual reality for content-based image retrieval. In *IEEE Int. Conference on Multimedia and Expo*, 2001.
4. Y. Rubner, L. Guibas, and C. Tomasi. The earth mover's distance, multi-dimensional scaling, and color-based image retrieval. In *Image Understanding Workshop*, pages 661–668, 1997.

5. S.D. Ruszala and G. Schaefer. Visualisation models for image databases: A comparison of six approaches. In *Irish Machine Vision and Image Processing Conference*, pages 186–191, 2004.
6. J. Sangwine and R.E.N. Horne. *The Colour Image Processing Handbook*. Chapman & Hall, 1998.
7. G. Schaefer and M. Stich. UCID - An Uncompressed Colour Image Database. In *Storage and Retrieval Methods and Applications for Multimedia 2004*, volume 5307 of *Proceedings of SPIE*, pages 472–480, 2004.
8. A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12):1249–1380, 2000.
9. M. Stricker and M. Orengo. Similarity of color images. In *Conf. on Storage and Retrieval for Image and Video Databases III*, volume 2420 of *Proceedings of SPIE*, pages 381–392, 1995.
10. VRML Consortium. The Virtual Reality Modeling Language. *ISO/IEC IS 14772-1*, 1997.