Proposal for ICADL 2012 Tutorial on
"Querying and browsing image libraries"

Gerald Schaefer
Department of Computer Science
Loughborough University, U.K.
gerald.schaefer@ieee.org

Tutorial length

Half-day.

Level of audience

Anybody interested in finding information in image libraries. As we will start with the basics no prior knowledge is necessary.

Outline

This tutorial will give an introduction to managing, querying and browsing image libraries for the case when no textual annotations for the images are available. It will cover the basics of content-based image retrieval (CBIR) which extracts visual features directly from the images than can then be employed for retrieving images, and image database navigation which provides a visual overview of (potentially large) image libraries coupled with tools for effective intuitive browsing through the image collection.

Topics covered:

- Image databases and problems of manual annotation
- Content-based image retrieval
  - by colour, texture, and shape
- Image classification
- Image annotation
- Retrieval vs. visualisation and browsing
- Image database visualisation
  - by dimensionality reduction, clustering, graph analysis
- Image database navigation
  - horizontal browsing, vertical browsing, time-based browsing
- Current trends
  - immersive browsing, mobile browsing
- Image database retrieval and browsing – challenges and future directions

Contact information

Gerald Schaefer
Department of Computer Science
Loughborough University
Loughborough, LE11 3TU
U.K.
Phone: +44-1509-635707
Fax: +44-1509-635722
E-mail: gerald.schaefer@ieee.org
http://www-staff.lboro.ac.uk/~cogs/

Proposal for ICADL 2012 Tutorial on "Querying and browsing image libraries"

Gerald Schaefer
Department of Computer Science
Loughborough University, U.K.
gerald.schaefer@ieee.org

Extended outline

Speaker biography

Gerald Schaefer gained his PhD in Computer Vision from the University of East Anglia. He worked at the Colour & Imaging Institute, University of Derby (1997-1999), in the School of Information Systems, University of East Anglia (2000-2001), in the School of Computing and Informatics at Nottingham Trent University (2001-2006), and in the School of Engineering and Applied Science at Aston University (2006-2009) before joining the Department of Computer Science at Loughborough University where he leads the Vision, Imaging and Autonomous Systems Research Division.

His research interests are mainly in the areas of colour image analysis, image retrieval, physics-based vision, medical imaging, and computational intelligence. He has published extensively in these areas with a total publication count exceeding 250. He is a member of the editorial board of more than 10 international journals, reviews for over 70 journals and served on the programme committee of more than 200 conferences. He has been invited as keynote or tutorial speaker to more than 30 conferences, is the organiser of various international workshops and special sessions at conferences, and the editor of several books, conference proceedings and special journal issues.

Similar/related tutorials were given/invited at a number of conferences, including SoCPaR 2009, NaBIC 2009, ISDA 2010, SITIS 2010, PAKDD 2011, BNCOD 2011, ICMMI 2011, VCIP 2011, with further tutorials scheduled at SMC 2012, ICPR 2012 and VCIP 2012.
See http://www-staff.lboro.ac.uk/~cogs/talks/talks.html for more details on previous tutorials/talks.

Extended outline

Nowadays, the majority of people possess some form of digital camera to use in their everyday lives. Devices range from relatively low quality web cameras, to medium range cameras integrated into mobile devices, to higher quality cameras aimed at the average user, on to high-end cameras used by professional photographers. Affordability of devices and storage media coupled with increased capabilities and the 'to hand' availability of camera equipment has led to a dramatic increase in the number of digital images the average end user creates and stores.

With the reduction in digital photography costs, a shift in the attitude towards photo taking can be observed. Users tend to take more images now than before, particularly of the same objects or

scene (e.g. from different perspectives. This is certainly a change from the past, where one would generally be concerned about the number of exposures left on the current film roll or the cost of developing photographs, whereas a digital camera user not happy with a photo can simply delete it from the camera's memory and images can be printed on home printers.

Personal image collections nowadays are typically in the range of hundreds to thousands of images. The rapid increase in the number of digital images taken by individuals has also caused an exponential growth in the number of images available online. Social networking sites allow users to instantly share images with friends, family or a wider community of users that also have the ability to comment and even 'tag' who or what may be in an image. Commercially, professional photography companies may store millions of digital images in their databases. These are generally manually annotated image collections used by journalists from a variety of publications to search for images suited to their particular needs. As one can imagine, the search for any particular image in collections of either personal or commercial magnitude can be tiresome and exhaustive. Generally, images are arranged in a one-dimensional linear arrangement, whereby an image has no correlation to any of its neighbours. Images are usually grouped together in a manually named folder or on the basis that they were uploaded to the computer at the same time.

This organisation of images is not ideal for a variety of reasons. Firstly, the cost of storage media has dramatically decreased whilst storage capacity has increased. Therefore an average end user may take many photos of many different events (such as birthdays, holidays etc.) on a camera before uploading them to their computer. If not sorted manually, multiple events may get grouped together, potentially making it difficult for the user to locate specific images in the future.

In the proposed tutorial we will therefore look at a variety of techniques and methods for effective and intuitive image database querying, visualisation and browsing. A lot of research in the last two decades has focussed on techniques to extract useful data directly from images to facilitate searching large image repositories. In this tutorial we will explain the underlying techniques, highlight some of the challenges to be overcome, and introduce some recent approaches that provide interesting and useful methods of working with image datasets.

Image browsing systems attempt to provide the user with a more intuitive interface, displaying more images at once in order to harness the cognitive power of the human mind in order to recognise and comprehend an image in seconds. Interaction with a traditional QBE system can often lead to a confusing and frustrating user experience. Formulating queries from images can prove difficult for the user, and the 'black-box' state of such approaches means that users typically cannot derive how the system is retrieving these results, and are thus unable to modify the query in order to improve the results returned by the system.

Browsing systems give a useful alternative to QBE. Providing an overview of the database to the user allows for intuitive navigation throughout the system. This is particularly the case when images are arranged according to mutual similarity which has been shown to speed up target search (i.e. looking for a particular image). QBE systems cannot be used when the user does not have a specific image in mind, as no query image can be provided. Image browsing systems overcome this problem by showing an overview of the image database. An overview of the collection will give the user a good indication whether or not an image or image class they have in mind might actually be present in the database. In some cases, the entire database will be displayed to the user on a single display. The user can then focus on regions of the visualisations that they are attracted to or believe will harbor a particular concept they have in mind. Browsing such visualisations when arranged according to image similarity, can increase the rate of retrieval. These visualisations are usually achieved through dimensionality reduction, whereby the relationships between images in a high-dimensional feature space are maintained as best possible in a reduced 2D (or 3D) space which is more comprehensible to the user.

In case image collections are too large to fit to a single display, images can be grouped according to similarity through the application of a clustering procedure. The user is then able to navigate through these clustered groups of images in order to browse the collection. An overview of the database is provided by initially presenting the user with a representative image for each cluster. Clustering can also be performed in a hierarchical manner which in turn allows for visualisation of very large datasets.

Another way in which image databases can be displayed is through graph-based visualisations. In these approaches, links are formed between images that are deemed similar or that share a common concept, while the images themselves form the nodes of the graph. The whole connected graph, or part thereof, is then displayed to the user for visualisation and navigation.

Similarity-based visualisation is not the only useful form of arranging image databases. In particular for personal collections, grouping according to the time images were created has shown to be useful. This approach can be adopted to automatically cluster event images. In cases where time information is not always available or not necessarily reliable, this approach can be combined with similarity-based systems.

The fundamental issue with the development of a browsing system is how to present the user with the images in a database. With image collections ranging in the size of millions, any browsing system needs to utilise the limited screen space provided by a typical computer monitor in a manner which is intuitive and easily navigable by the common user. Immersive environments and virtual reality allow for a completely new way of visualising information with a unique user experience. It is only natural that this approach has also been adopted for visualising image databases. The user is immersed into the actual database, while the addition of a third dimension coupled with the larger visualisation space can lead to a more effective approach of navigation.

While a visualisation of an image collection is useful for providing an overview of the contained images, it provides only part of a useable image database navigation system. Once a collection is visualised, users should have the ability to interact with it in order to arrive at the image(s) they are looking for. Typical operations here include panning and zooming which allow the user to focus on images in a different part of the visualisation space, respectively on images which were previously hidden.

Finally, it is evident that the usefulness of image browsing systems needs to be evaluated. While there is relatively little work on this, we will look at various measures to attempt to measure how well image browsers are suited for various tasks.

Tutorial outline

- Image databases and problems of manual annotation
- Content-based image retrieval
    - By colour
    - By texture
    - By shape
- Challenges
    - Image compression
    - Colour variations
- Image classification
- Image annotation
- Retrieval vs. visualisation and browsing
- Image database visualisation
    - Through dimensionality reduction
    - Through clustering
    - On graphs and networks
    - Time-based image database visualisation and hybrid visualisation approaches
- Image database navigation
    - Horizontal image database browsing
    - Vertical image database browsing
    - Immersive image database browsing
- Evaluating image database browsing systems
- Image database retrieval and browsing – challenges and future directions