

An Outline of Ontology Based Image Retrieval Using Semantic Image Descriptions

Ancy S^{1*}, Sonia Jenifer Rayen², Karthikraj M³

¹Assistant Professor, Department of Information Technology, Jeppiaar Institute of Technology, Kanchipuram, Tamil Nadu, India

²Assistant Professor, Department of Information Technology, Jeppiaar Institute of Technology, Kanchipuram, Tamil Nadu, India

³Assistant Professor, Department of Information Technology, Jeppiaar Institute of Technology, Kanchipuram, Tamil Nadu, India

Review Article

***Corresponding author**

Ancy S

Article History

Received: 07.05.2018

Accepted: 22.05.2018

Published: 30.06.2018

DOI:

10.36347/sjet.2018.v06i06.001



Abstract: This architecture addresses the issues of keyword based image retrieval and content-based image retrieval through the use of qualitative spatial representations over semantic image annotations. Three types of image retrieval Semantic Retrieval based on Global Labels of Images, Semantic Retrieval based on Image concepts, Semantic Retrieval based on Qualitative Relations. This research proposes an ontology based image retrieval framework from a corpus of natural scene images by imparting human cognition in the retrieval process. The proposed architecture addresses the issues of keyword based image retrieval and content-based image retrieval through the use of qualitative spatial representations over semantic image annotations. Domain ontology has been developed to model qualitative semantic image descriptions and retrieval, thereafter can be accomplished either using a natural language description of an image containing semantic concepts and spatial relations, or in a query by example fashion. A psychophysical evaluation has also been carried out to evaluate the effectiveness of our approach and results of different experiments are quite promising in terms of retrieval accuracy and relevance of retrieved images.

Keywords: Content-based image retrieval, content-based visual information retrieval.

INTRODUCTION

Content-based image retrieval

Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases.

"Content-based" means that the search will analyze the actual contents of the image. The term 'content' in this context might refer colors, shapes, textures, or any other information that can be derived from the image itself. Without the ability to examine image content, searches must rely on metadata such as captions or keywords. Such metadata must be generated by a human and stored alongside each image in the database.

Problems with traditional methods of image indexing [Enser,1995] [1] have led to the rise of interest in techniques for retrieving images on the basis of automatically-derived features such as colour, texture and shape – a technology now generally referred to as Content-Based Image Retrieval (CBIR). However, the technology still lacks maturity, and is not yet being used on a significant scale. In the absence of hard evidence

on the effectiveness of CBIR techniques in practice, opinion is still sharply divided about their usefulness in handling real-life queries in large and diverse image collections. The concepts which are presently used for CBIR system are all under research.

Images

Let us start with the word "image". The surrounding world is composed of images. Humans are using their eyes, containing 1.5×10^8 sensors, to obtaining images from the surrounding world in the visible portion of the electromagnetic spectrum (wavelengths between 400 and 700 nanometers).[2] The light changes on the retina are sent to image processor centre in the cortex.

In the image database systems geographical maps, pictures, medical images, pictures in medical

atlases, pictures obtained by cameras, microscopes, telescopes, video cameras, paintings, drawings and architectures plans, drawings of industrial parts, space images are considered as images.

There are different models for colour image representation. In the seventeenth century Sir Isaac Newton showed that a beam of sunlight passing through a glass prism comes into view as a rainbow of colours. Therefore, he first understood that white light is composed of many colours. Typically, the computer screen can display 2^8 or 256 different shades of gray. For colour images this makes $2^{(3 \times 8)} = 16,777,216$ different colours.

Clerk Maxwell showed in the late nineteenth century that every colour image could be created using three images – Red, green and Blue image. A mix of these three images can produce every colour. This model, named RGB model, is primarily used in image representation. The RGB image could be presented as a triple(R, G, B) where usually R, G, and B take values in the range [0, 255]. Another colour model is YIQ model (luminance (Y), phase (I), quadrature phase (Q)). It is the base for the colour television standard. Images are presented in computers as a matrix of pixels. They have finite area. If we decrease the pixel dimension the pixel brightness will become close to the real brightness. The same image with different pixel dimension is shown below.

Image Database systems

Set of images are collected, analyzed and stored in multimedia information systems, office systems, Geographical information systems(GIS), robotics systems, CAD/CAM systems, earth resources systems, medical databases, virtual reality systems,

Information retrieval systems, art gallery and museum catalogues, animal and plant atlases, sky star maps, meteorological maps, catalogues in shops and many other places.

There are sets of international organizations dealing with different aspects of image storage, analysis and retrieval. Some of them are: AIA (Automated Imaging/Machine vision)[3], AIIM (Document imaging), ASPRES (Remote Sensing/Program) etc.

There are also many international centers storing images such as: Advanced imaging, Scientific/Industrial Imaging, Microscopy imaging, Industrial imaging etc. There are also different international work groups working in the field of image compression, TV images, office documents, medical images, industrial images, multimedia images, graphical images, etc.

Logical Image Representation in Database Systems

The logical image representation in image databases systems is based on different image data models. An image object is either an entire image or some other meaningful portion (consisting of a union of one or more disjoint regions) of an image. The logical image description includes: meta, semantic, colour, texture, shape, and spatial attributes.

Colour attributes could be represented as a histogram of intensity of the pixel colours. A histogram refinement technique is also used by partitioning histogram bins based on the spatial coherence of pixels. Statistical methods are also proposed to index an image by colour correlograms, which is actually a table containing colour pairs, where the k -th entry for $\langle i, j \rangle$ specifies the probability of locating a pixel of colour j at a distance k from a pixel of colour i in the image.

Classification and indexing schemes

Many picture libraries use keywords as their main form of retrieval – often using indexing schemes developed in-house, which reflect the special nature of their collections. A good example of this is the system developed by Getty Images to index their collection of contemporary stock photographs. Their thesaurus comprises just over 10 000 keywords, divided into nine semantic groups, including geography, people, activities and concepts. Index terms are assigned to the whole image, the main objects depicted, and their setting. Retrieval software has been developed to allow users to submit and refine queries at a range of levels, from the broad (e.g. “freedom”) to the specific (e.g. “a child pushing a swing”).

Probably the best-known indexing scheme in the public domain is the Art and Architecture Thesaurus (AAT), originating at Rensselaer Polytechnic Institute in the early 1980s, and now used in art libraries across the world. AAT is maintained by the Getty Information Institute and consists of nearly 120,000 terms for describing objects, textural materials, images, architecture and other cultural heritage material. There are seven facets or categories which are further subdivided into 33 sub facets or hierarchies. The facets, which progress from the abstract to the concrete, are: associated concepts, physical attributes, styles and periods, agents, activities, materials, and objects. AAT is available on the Web from the Getty Information Institute at http://www.ahip.getty.edu/aat_browser/. Other tools from Getty include the Union List of Artist Names (ULAN) and the Getty Thesaurus of Geographic Names (TGN). Another popular source for providing subject access to visual material is the Library of Congress Thesaurus for Graphic Materials (LCTGM). Derived from the Library of Congress Subject Headings (LCSH), LCTGM is designed to assist with the indexing of historical image collections in the

automated environment. Greenberg [1993] provides a useful comparison between AAT and LCTGM.[4,5,6]

A number of indexing schemes use classification codes rather than keywords or subject descriptors to describe image content, as these can give a greater degree of language independence and show concept hierarchies more clearly. Examples of this genre include ICONCLASS from the University of Leiden [Gordon, 1990][6], and TELCLASS from the BBC [Evans, 1987][7]. Like AAT, ICONCLASS was designed for the classification of works of art, and to some extent duplicates its function; an example of its use is described by Franklin [1998]. TELCLASS was designed with TV and video programmes in mind, and is hence rather more general in its outlook. The Social History and Industrial Classification, maintained by the Museum Documentation Association, is a subject classification for museum cataloguing. It is designed to make links between a wide variety of material including objects, photographs, archival material, tape recordings and information files.

A number of less widely-known schemes have been devised to classify images and drawings for specialist purposes. Examples include the Vienna classification for trademark images [World Intellectual Property Organization, 1998], used by registries Worldwide to identify potentially conflicting trademark applications, and the Opitz coding system for machined parts [Opitz et al, 1969][8], used to identify families of similar parts which can be manufactured together.

A survey of art librarians conducted for this report suggests that, despite the existence of specialist classification schemes for images, general classification schemes, such as Dewey Decimal Classification (DDC), Library of Congress (LC), BLISS and the Universal Decimal Classification (UDC), are still widely used in photographic, slide and video libraries. The former scheme is the most popular, which is not surprising when one considers the dominance of DDC in UK public and academic library sectors. ICONCLASS, AAT, LCTGM, SHIC are all in use in at least one or more of the institutions in the survey. However, many libraries and archives use in-house schemes for the description of the subject content. For example, nearly a third of all respondents have their own in-house scheme for indexing slides.

When discussing the indexing of images and videos, one needs to distinguish between systems which are geared to the formal description of the image and those concerned with subject indexing and retrieval. The former is comparable to the bibliographical description of a book. However, there is still no one standard in use for image description, although much effort is being expended in this area by a range of organizations such as the Museum

Documentation Association, the Getty Information Institute, the Visual Resources Association the International Federation of Library Association/Art Libraries and the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM).

The descriptive cataloguing of photographs presents a number of special challenges. Photographs, for example, are not self-identifying. Unlike textual works that provide such essential cataloguing aids as title pages, abstracts and table of contents, photographs often contain no indication of author or photographer, names of persons or places depicted dates, or any textual information whatever. Cataloguing of images is more complex than that for text documents, since records should contain information about the standards used for image capture and how the data is stored as well as descriptive information, such as title, photographer (or painter, artist, etc). In addition, copies of certain types of images may involve many layers of intellectual property rights, pertaining to the original work, its copy (e.g. a photograph), a digital image scanned from the photograph, and any subsequent digital image derived from that image.

Published reviews of traditional indexing practices for images and video include many writers discuss the difficulties of indexing images. The problem of managing a large image collection. He notes that, unlike books, images make no attempt to tell us what they are about and that often they may be used for purposes not anticipated by their originators. Images are rich in information and can be used by researchers from a broad range of disciplines. As Baser comments:

“A set of photographs of a busy street scene a century ago might be useful to historians wanting a ‘snapshot’ of the times, to architects looking at buildings, to urban planners looking at traffic patterns or building shadows, to cultural historians looking at changes in fashion, to medical researchers looking at female smoking habits, to sociologists looking at class distinctions, or to students looking at the use of certain photographic processes or techniques.”

Svenonius [1994][9] discusses the question of whether it is possible to use words to express the “abruptness of a work in a wordless medium, like art. To get around the problem of the needs of different users groups, van der Starre [1995] advocates that indexers should “stick to ‘plain and simple’ indexing, using index terms accepted by the users, and using preferably a thesaurus with many lead-ins,” thus placing the burden of further selection on the user. Shatford Layne (1994) suggests that, when indexing images, it may be necessary to determine which attributes provide useful groupings of images; which attributes provide information that is useful once the

images are found; and which attributes may, or even should, be left to the searcher or researcher to identify. She also advocates further research into the ways images are sought and the reasons that they are useful in order to improve the indexing process. Constantopoulos and Doerr (1995)[7] also support a user centred approach to the designing of effective image retrieval systems. They urge that attention needs to be paid to the intentions and goals of the users, since this will help define the desirable descriptive structures and retrieval mechanisms as well as understanding what is ‘out of the scope’ of an indexing system.

When it comes to describing the content of images, respondents in our own survey seem to include a wide range of descriptors including title, period, genre, subject headings, keywords, classification and captions (although there was some variation by format). Virtually all maintain some description of the subject content of their images. The majority of our respondents maintain manual collections of images, so it is not surprising that they also maintain manual indexes. Some 11% of respondents included their photographs and slides in the online catalogues, whilst more than half added their videos to their online catalogues.

Research Trends in the Image Database Systems

Most image database systems are products of research, and therefore emphasize only one aspect of content-based retrieval. Sometimes this is the sketching capability in the user interface; sometimes it is a new indexing data structure, etc. Some systems are created as a research version and a commercial product. The commercial version is usually less advanced, and shows more standard searching capabilities. A number of systems provide user interface that allows more powerful query formulation than is useful in demo system. Most systems use colour and texture features, few systems use shape features, and yet less use spatial features. The retrieval on colour usually yield images that have similar colours. The larger the collection of images, the greater is the chance that it contains an image similar to the query image.

EXISTING SYSTEM

SIRNS (Semantic Image Retrieval of Natural Scenes) TEHCNIQUE An ontology based image retrieval approach using qualitative semantic image

PROPOSED SYSTEM

SIRNS (Semantic Image Retrieval of Natural Scenes) TEHCNIQUE with RDF (Resource Description Framework triples)

Architecture System

The earliest use of the term content-based image retrieval in the literature seems to have been by, to describe his experiments into automatic retrieval of images from a database by colour and shape feature. The term has since been widely used to describe the process of retrieving desired images from a large collection on the basis of features (such as colour, texture and shape) that can be automatically extracted from the images themselves. The features used for retrieval can be either primitive or semantic, but the extraction process must be predominantly automatic. Retrieval of images by manually-assigned keywords is definitely not CBIR as the term is generally understood – even if the keywords describe image content.

CBIR differs from classical information retrieval in that image databases are essentially unstructured, since digitized images consist purely of arrays of pixel intensities, with no inherent meaning. One of the key issues with any kind of image processing is the need to extract useful information from the raw data (such as recognizing the presence of particular shapes or textures) before any kind of reasoning about the image’s contents is possible. Image databases thus differ fundamentally from text databases, where the raw material (words stored as ASCII character strings) has already been logically structured by the author. There is no equivalent of level 1 retrieval in a text database.

CBIR draws many of its methods from the field of image processing and computer vision, and is regarded by some as a subset of that field. It differs from these fields principally through its emphasis on the retrieval of images with desired characteristics from a collection of significant size. Image processing covers a much wider field, including image enhancement, compression, transmission, and interpretation. While there are grey areas (such as object recognition by feature analysis), the distinction between mainstream image analysis and CBIR is usually fairly clear-cut. An example may make this clear. Many police forces now use automatic face recognition systems. Such systems may be used in one of two ways. Firstly, the image in front of the camera may be compared with a single individual’s database record to verify his or her identity. In this case, only two images are matched, a process few observers would call CBIR. Secondly, the entire database may be searched to find the most closely matching images. This is a genuine example of CBIR.

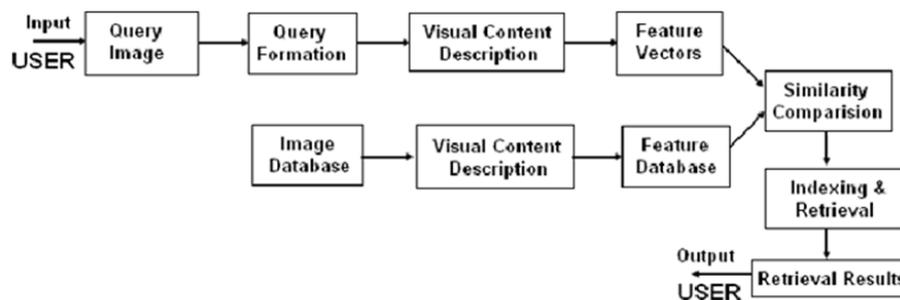


Fig 1 Architecture diagram of cCBIR system

The process of retrieving desired images from a large collection on the basis of features (such as colour, texture and shape) that can be automatically extracted from the images themselves. The features used for retrieval can be either primitive or semantic, but the extraction process must be predominantly automatic

In typical Content-based image retrieval systems (Figure I), the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. The feature vectors of the images in the database form a feature database. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors.

The similarities/distances between the feature vectors of the query example or sketch and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing scheme. The indexing scheme provides an efficient way to search for the image database.

To modify the retrieval process in order to generate perceptually and semantically more meaningful retrieval results. In this chapter, we introduce these fundamental techniques for content-based image retrieval.

CONCLUSION

This work addresses the issues of keyword based image retrieval and content-based image retrieval through the use of qualitative spatial representations over semantic image annotations. A psychophysical evaluation has also been carried out to evaluate the effectiveness of our approach and results of different experiments are quite promising in terms of retrieval accuracy and relevance of retrieved images. Hence this method proves to produce better results.

REFERENCES

1. W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and

G. Taubin, 'The QBIC Project: Querying Images By Content Using Color, Texture, and Shape', In IS&T/SPIE Symposium on Electronic Imaging Science & Technology, San Jose, CA, USA, February 1993.

2. Yong Rui and Thomas S. Huang, Shih-Fu Chang, 'Image Retrieval: Current Techniques, Promising Directions and open issues', Journal of Visual Communication and Image Representation, 1999;10, 39–62,
3. Gerard Salton and Christopher Buckley, 'Term-weighting approaches in automatic text retrieval', Information Processing and Management, 1998.
4. Deb S and Zhang. "An Overview of Content-based Image Retrieval techniques". Proceedings of 18th International Conf on Advanced Information Networking and Application (AINA04), 1, issue 2004:5964.
5. Veltkamp and M Tanase. "Content-Based Image Retrieval Systems: A Survey. Tech. Rep. UU-CS 2000-34, Univ. Utrecht, Utrecht, The Netherlands, 2000.
6. Wang Hui, Dzulkifli Mohamad & N.A. Ismail "Semantic Gap in CBIR: Automatic Objects Spatial Relationships Semantic Extraction and representation International Journal of Image Processing (IJIP), 2010; (4): 3.
7. Qayyum, Z U, Cohn, A G, 'Qualitative Approaches to Semantic Scene Modeling and Retrieval', in Research and Development in Intelligent Systems XXIII, Springer Verlag, 2006.
8. Qayyum, Z U, Cohn, A G and Klippel, A, 'Psychophysical Evaluation for a Qualitative Semantic Image Categorization and Retrieval Approach', Trends in Applied Intelligent Systems, LNCS, pp 321-330, Springer, 2010
9. Giannis Varelas, Epimenidis Voutsakis, Paraskevi Raftopoulou, "Semantic Similarity Methods in WordNet and their Application to Information Retrieval on the Web", In proc. of 7th ACM international workshop on Web information and data management, pp. 10-16, 2005.