ISSN: 2393-8528

Contents lists available at www.ijicse.in

International Journal of Innovative Computer Science & Engineering

Journal Index In ICI World of Journals – ICV 2016 – 69.57

Volume 4 Issue 6; November-December 2017; Page No. 12-17

Active Learning Methods for Interactive Image Retrieval

B. Jhansi Rani[1], B.Prashanth[2], B.Krishna[3]

[1]Lecturer in TSRWDC (W), Suryapet, Telangana State, India

b.jhansi007@gmail.com

[2]Asst.Professor, Dept of CSE, Vaagdevi College of Engineering, Warangal, Telangana State, India

prashanth19bolukonda@gmail.com

[3]Research Scholar, Dept of CSE, UCE, Osmania University, Hyderabad, Telangana State, India

bandikrishna007@gmail.com

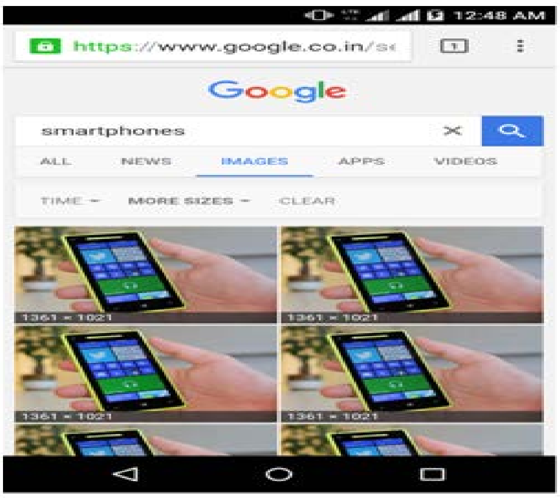
ABSTRACT

Human interactive systems have attracted a lot of research interest in recent years, especially for content-based image retrieval systems. Contrary to the early systems, which focused on fully automatic strategies, recent approaches have introduced human-computer interaction. In this paper, we focus on the retrieval of concepts within a large image collection. We assume that a user is looking for a set of images, the query concept, within a database. The aim is to build a fast and efficient strategy to retrieve the query concept. In content-based image retrieval (CBIR), the search may be initiated using a query as an example. The top rank similar images are then presented to the user. Then, the interactive process allows the user to refine his request as much as necessary in a relevance feedback loop. Many kinds of interaction between the user and the system have been proposed, but most of the time, user information consists of binary labels indicating whether or not the image belongs to the desired concept.

Keywords: Multimedia information retrieval, Content based image retrieval, Image search, Interactive search, Relevance feedback.

1. INTRODUCTION

Terabytes of imagery are being accumulated daily from a wide variety of sources such as the Internet, medical centers (MRI, X-ray, CT scans) or digital libraries. It is not uncommon for one's personal computer to contain thousands of photos stored in digital photo albums. At present, billions of images can even be found on the World Wide Web. But with that many images within our reach, how do you go about finding the ones you want to see at a particular moment in time? Interactive search methods are meant to address the problem of finding the right imagery based on an interactive dialog with the search system. The areas of interactive search with the greatest societal impact have been in WWW image search engines and recommendation systems. Google, Yahoo! and Microsoft have added interactive visual content-based search methods into their worldwide search engines, which allows search by similar shape and/or color and are used by millions of people each day. The recommendation systems have been implemented by companies such as Amazon, NetFlix in wide and diverse contexts, from books to clothing, from movies to music. They give recommendations of what the user would be interested in next based on feedback from prior ratings.



*Corresponding author: B. Jhansi Rani

Text search relies on an notations that are frequently missing in both personal and public image collections. When annotations are either missing or incomplete, the only alternative is to use methods that analyses the pictorial content of the imagery in order to find the images of interest. This field of research is also known as content based image retrieval. This survey is aimed at content-based image retrieval researchers and intends to provide insight into the trends and diversity of interactive search techniques in image retrieval from the perspectives of the users and the systems.

1.1 INTERACTIVE SEARCH FROM THE USER'SPOINT OF VIEW

A rough overview of the interactive search process. Note that real systems typically have significantly greater complexity. In the first step, the user issues a query using the interface of the retrieval system and shortly thereafter is presented with the initial results. The user can then interact with the system in order to obtain improved results. Conceivably, the ideal interaction would be through questions and answers (Q&A), similar to the interaction at a library helpdesk.

Through a series of questions and answers the librarian helps the user find what he is interested in, often with the question "Is this what you are looking for?". This type of interaction would eventually uncover the images that are relevant to the user and which ones are not. In principle, feedback can be given as many times as the user wants, although generally he will stop giving feedback after a few iterations, either because he is satisfied with the retrieval results or because the results no longer improve.

2.1 QUERY SPECIFICATION

The most common way for a retrieval session to start is similar to the Q&A interaction one would have with a librarian. One might provide some descriptive text, provide an example image or in some situations use the favorites based on the history of the user . The query step can also be shipped directly when the system shows a random selection of images from the database for the user to give feedback on. When image segmentation is involved there area variety of ways to query the retrieval system, such as selecting one or more pre-segmented regions of interest or drawing outlines of objects of interest. A novel way to compose the initial query is to let the user first choose keywords from a thesaurus, after which per keyword one of its associated visual regions is selected.

2.2 RETRIEVAL RESULTS

The standard way in which the results are displayed is a ranked list with the images most similar to the query shown at the top of the list. Because giving feedback on the best matching images does not provide the retrieval system with much additional information other than what it already knows about the user's interest, a second list is also often shown, which contains the images most informative to the system. These are usually the images that the system is most uncertain about, for instance those that are on or near a hyper plane when using SVM-based retrieval. This principle, called active learning.

2.3 USER INTERACTION

Many of the systems have interaction which is designed to be used by a machine learning algorithm which gives rise naturally to labeling results as either positive and/or negative examples. These examples are given as feedback to the systems to improve the next iteration of results. Researchers have explored Using positive feedback only, positive and negative feedback, positive, neutral and negative feedback, and multiple relevance levels: four relevance levels, five levels or even seven levels. An alternative approach is to let the user indicate by what percentage a sample image meets what he has in mind. While positive/negative examples are important to learning, in many cases it can be advantageous to allow the user to give other kinds of input which may be in other modalities (text, audio, images, etc.), other categories, or personal preferences. Thus, some systems allow the user to input multiple kinds of information in addition to labelled example. In addition, sketch interfaces allow the user to give a fundamentally different kind of input to the system, which can potentially give a finer degree of control over the results. In the Q&A paradigm, results may be dynamically selected to best fit the question, based on deeper analysis of the user query. For example, by detecting verbs in the user query or results, the system can determine that a videos how the actions will provide a better answer than an image or only text. When the system uses segmented images it is possible to implement more elaborate feedback schemes, for instance allowing the splitting or merging of image regions, or supporting drawing a rectangle inside a positive example to select a region of interest. An interesting discussion on the role and impact of negative images and how to interpret their meaning can be found in. Besides giving explicit feedback, it is also possible to consider the user's actions as a form of implicit feedback, which may be used to refine the results that are shown to the user in the next result screen. An example of implicit feedback is a click-through action, where the user clicks on an image with the intention to see it in more detail. In contrast with the traditional query-based retrieval model, the ostensive relevance feedback model accommodates for changes in the user's information needs as they evolve over time through exposure to new information over the course of a single search session.

2.4 THE INTERFACE

The role of the interface in the search process is often limited to displaying a small set of search results that are arranged in a grid, where the user can refine the query by indicating the relevance of each individual image. In recent literature, several interfaces break with this convention, aiming to offer an improved search experience. These interfaces mainly focus on one, or a combination, of the following aspects: Support for easy browsing of the image collection, for instance through an ontological representation of the image collection where the user can zoom in on different concepts of interest, by easily shifting the focus of attention from image to image allowing the user to visually explore the local relevant Neighborhood surrounding an image or by letting users easily navigate to other promising areas in feature space, which is particularly useful when the search no longer improves with the current set of relevant images. Better presentation of these arch results, with for instance giving more screen space to images that are likely to be more relevant to the query than to less relevant images, dynamically reorganizing the displayed pages into visual islands that enable the user to explore deeper into a particular dimension he is interested in, or visualizing the results where similar images are placed closer together. Multiple query modalities, result modalities and ways of giving feedback, for instance by allowing the user to query by grouping and/or moving images, 'scribbling' on images to make it clear to the retrieval system which parts of an image should be considered foreground and which parts background, or providing the user with the best mixture of media for expressing a query or understanding the results.

3. INTERACTIVE SEARCH FROM THE SYSTEM'S POINT OF VIEW

A global overview of a retrieval system is shown in Fig the images in the database are converted into a particular image representation, which can optionally be stored in an indexing structure to speed up the search. Once a query is received, the system applies an algorithm to learn what kind of images the user is interested in, after which the database images are ranked and shown to the user with the best matches first. Any feedback the user gives can optionally be stored in a log for the purpose of discovering search patterns, so learning will improve in the long run. This section covers the recent advances on each of these parts of a retrieval system.

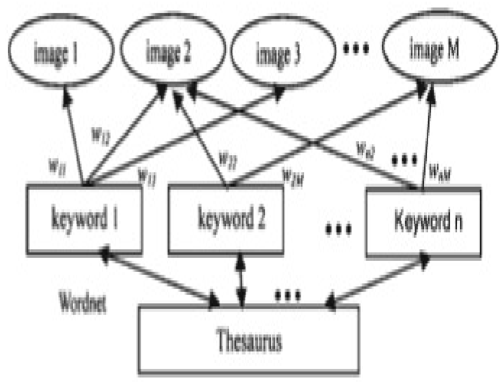


Fig.1: A thesaurus is used to link keywords to images

3.1 IMAGE REPRESENTATION

By itself an image is simply a rectangular grid of colored pixels. In the brain of a human observer these pixels' form meanings based on the person's memories and experiences, expressing itself in a near-instantaneous recognition of objects, events and locations. However, to a computer an image does not mean anything, unless it is told how to interpret it. Often images are converted into low-level features, which ideally capture the image characteristics in such a way that it is easy for the retrieval system to determine how similar two images are as perceived by the user. In current research, the attention is shifting to mid-level and high-level image representations. Mid-level representations focus on particular parts of the image that are important, such as sub-images, regions and salient details. After these image elements have been determined, they are often seen as standalone entities during the search. However, some approaches represent them in a hierarchical or graph-based structure and exploit this structure when searching for improved retrieval results. The multiple instance learning and bagging approach lends itself very well to image retrieval, because an image can be seen as a bag of visual words where these visual words can, for instance, be interest points, regions, patches or objects. By incorporating feedback, the idea is that the user can only give feedback on the entire bag (i.e. the image), although he might only be interested in one or more specific instances (i.e. visual words) in that bag. The goal is then for the system to obtain a hypothesis from the feedback images that predicts which visual words the user is looking for. where the multiple instance learning technique does not assume that a bag is positive when one or more of its instances are positive. High-level representations are designed with semantics in mind. A thesaurus, such as Word Net, is often used to link annotations to image concepts, for instance by linking them through synonymy, hyponymy, hyponymy, etc. (See fig 2). Since manually annotating large collections of images is a tedious task, much research is directed at automatic annotation, mostly offline, but also driven by relevance feedback. Finding the best balance between using keywords for searching and using visual features for searching is one of the newer topics in image retrieval. For instance, in the image ranking presented to the user is composed first using a textual query vector to rank all database images and then using a visual query vector to re-rank them.

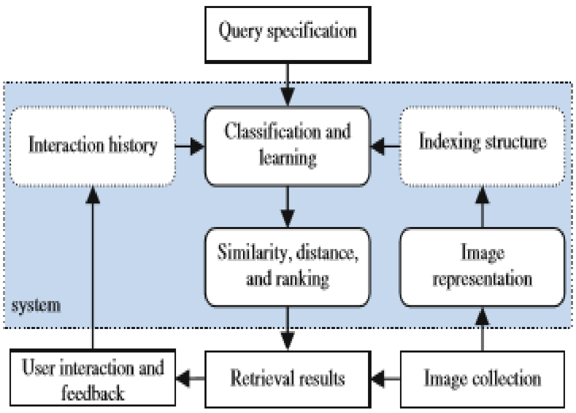


Fig.2: The interactive search process from the system's point of view

3.2 INDEXING AND FILTERING

Finding images that have high similarity with a query image often requires the entire database to be traversed for one-on one comparison. When dealing with large image collections this becomes prohibitive due to the amount of time the traversal takes. In the last few decades various indexing and filtering schemes have been proposed to reduce the number of database images to look at, thus improving the responsiveness of the system as perceived by the user. A good theoretical overview of indexing structures that can be used to index high-dimensional spaces. The majority of recent research in this direction focuses on the clustering of images, so that a reduction of the number of images to consider is then a matter offending out which cluster(s) the query image belongs to. Often the image clusters are stored in a hierarchical indexing structure to allow for a step-wise refinement of the number of images to consider. Alternatively, the set of images that are likely relevant to the query can be quickly established by approximating their feature vectors. A third way to reduce the number of images to inspect is by partitioning the feature space and only looking at that area of space which the query image belongs to. Hashing is a form of space partitioning and is considered to be an efficient approach for indexing.

3.3 ACTIVE LEARNING AND CLASSIFICATION

The core of the retrieval system is the algorithm that learns which images in the database the user is interested in by analyzing the query image and any implicit or explicit feedback. Typical interactive system shave two categories of images to show the user:(1) clarification images, which are images that may not be wanted by the user but that will help the learning algorithm improve its accuracy, and (2) relevant images, which are the images wanted by the user.

How to decide which imagery to select for the first category is addressed by an area called "active learning". Active learning arguably, the most important challenge in interactive search systems is how to reduce the interaction effort from the user while maximizing the accuracy of the results. From a theoretical perspective, how one can measure the information associated with an unlabelled example, so a learner can select the optimal set of unlabelled examples to show to the user that maximizes its information gain and thus minimizes the expected future classification error? This category as pertaining to image search is usually called active learning in the research community and is closely related to relevance feedback, which many consider to be a special case of active learning.

3.4 LONG-TERM LEARNING

In contrast with short-term learning, where the state of the retrieval system is reset after every user session, long term learning is designed to use the information gathered during previous retrieval sessions to improve the retrieval results in future sessions. Long-term learning is also frequently referred to as collaborative filtering. The most popular approach for long-term learning is to infer relationships between images by analyzing the feedback log, which contains all feedback given by users over time. From the accumulated feedback logs a semantic space can be learned containing the relationships between the images and one or more classes, typically obtained by applying matrix factorization or clustering techniques. Whereas the early long-term learning methods mostly built static relevance models, the recent trend is to continuously update the model after receiving new feedback.

4. DISCUSSION AND CONCLUSIONS

Over the years, the performance of interactive search systems steadily improved. Nonetheless, much research remains to be done. This section provides the most promising research directions.

4.1 PROMISING RESEARCH DIRECTIONS

Some top research directions that are based on this article are outlined below.

• Interaction in the question and answer paradigm

The Q&A paradigm has the strength that it is probably the most natural and intuitive for the user. Recent Q&A research has focused significantly more on multimodal (as opposed to mono modal) approaches for both posing the questions and displaying the answers. These systems can also dynamically select the best types of media for clarifying the answer to a specific question.

Interaction on the learned models

Beyond giving direct feedback on the results, preliminary work was started involving mid-level and high-level representations. Multi-scale approaches using segmented image components are certainly novel and promising.

• Interaction by explanation: providing reasons along with results

In the classic relevance feedback model, results are typically given but it is not clear to the user why the results were selected. In future interactive search systems, we expect to see systems which explain to the user why the results were chosen and allow the user to give feedback on the criteria used in the explanations, as opposed to only simply giving feedback on the image results.

•Social interaction: recommendation systems and collaborative filtering

The small training set problem is of particular concern because humans do not want to label thousands of images. An interesting approach is to examine potential benefits from using algorithms from the area collaborative filtering and recommendation systems. These systems have remarkably high performance in deciding which media items (often video) will be of interest to the user based on a social database of ranked items.

REFERENCES

1. Andre P, Cutrell E, Tan D, Smith G (2009) Designing novel image search interfaces by understanding unique characteristics andusage. In: Proceedings of international conference on human– computer interaction
2. Aggarwal G, Ashwin TV, Ghosal S (2002) An image retrieval system with automatic query modification. IEEE Trans. on Multimedia 4(2):201–214
3. Bian W, Tao D (2010) Biased discriminant Euclidean embedding for content-based image retrieval. IEEE Trans Image Process 19(2):545–554
4. Datta R, Joshi D, Li J, Wang JZ (2008) Image retrieval:ideas, influences, and trends of the new age. ACM ComputSurv 40(2): 1–60
5. Datta R, Li J, Wang JZ (2005) Content-based image retrieval: approaches and trends of the new age. In: Proceedings of ACM international workshop on multimedia, information retrieval, pp 253–262
6. Lew MS, Sebe N, Djeraba C, Jain R(2006) Content-based multimedia information retrieval: state of the art and challenges. ACM Trans Multimedia ComputCommunAppl 2(1):1–19
7. Ren K, Sarvas R, Calic J (2010) Interactive search and browsing interface for large-scale visual repositories. Multimedia Tools Appl 49:513–528
8. Smeulders AWM, Worring M, Santini S, Gupta A, Jain R (2000) Content-based image retrieval at the end of the early years. IEEE Trans Pattern Anal Machine Intell

Authors Profile

	B.Jhansi Rani received the MCA and M.Tech(CSE) form JNTUH affiliated colleges in 2012 and 2014 respectively. Now working as Lecturer in TSWRDC(W), Suryapet, Telangana State, India.
	B.Prashanth received the MCA from Kakatiya University affiliated college(KITS Warangal) and M.Tech(CSE) form JNTUH affiliated college in 2008 and 2012 respectively. Now working as an Asst.Professor in Dept of CSE, Vaagdevi College of Engineering, Telangana State, India.
	B.Krishna received the MCA from Kakatiya University affiliated college(KITS Warangal) and M.Tech(CSE) form JNTUH affiliated college in 2007 and 2013 respectively. Now pursuing PhD from Osmania University, Telangana State, India.