Image and Its Semantic Role in Search Problem

Nasir Touheed[1], Muhammad Saeed[2], M. Atif Qureshi[2], and Arjumand Younus[2]

[1] Institute of Business Administration, Karachi, Pakistan
[2] Department of Computer Science, University of Karachi, Pakistan
ntouheed@iba.edu.pk, saeed@uok.edu.pk, matifq@unhp.com.pk,
arjumand_younus@unhp.com.pk

Abstract. The world has now shrunk and information today exists in many forms ranging from text to videos. An overloaded World Wide Web, full of data makes it difficult to extract information from that data and this has given birth to a new phenomenon in the computer industry which is the search engine technology. Image that contains dense information has not yet found its real interpretation over search engines. In this paper we practice application of Semantic Web concepts and propose a standard dimension in image structures in order to improve searching ability in image search engines. An image annotation tool, called "SemImage", was developed which allows users to annotate an image with various ontologies and JPEG was taken as a case-study. This work describes our initial research efforts in semantics-based searching driven by ontological standards for images which we refer to as Image SemSearch.

Keywords: Search problem redefinition, semantics, ontological standard, search seed.

1 Introduction

The World Wide Web is a huge repository of information and this information is really varied and diverse in nature. Multimedia content has now become an integral part of the Web and it adds to the appeal of web sites; among these varied forms of media images are the most primitive and widely-used form. They have a unique appeal of their own as it is said

"A picture is worth a thousand words"

In primitive ages when Man did not know any language he learnt through the language of images. The same phenomenon still holds and an evidence of this is the huge amount of visual information on the Internet. In fact it would not be wrong to say that without visual information the Web would not have attained the popularity it enjoys today. We use our visual ability to see and understand visual information more than any other medium to communicate and collect information. Visual information can be found in Web documents or even as stand-alone objects. These include images, graphics, bitmaps, animations and videos [1]. With such a large volume of unstructured digital media there is a need for effective and additional techniques for image

D.-S. Huang et al. (Eds.): ICIC 2008, CCIS 15, pp. 388–397, 2008.

retrieval [2]. Further people quite often need such visual information as it finds a variety of applications like education and training, criminal tracking, law enforcement etc. Even though image search engines do exist like Google Image Search, Lycos and AltaVista photo finder but their search relies on text to look for images which yields a great percentage of irrelevant results [3].

Both types of visual information (still or moving) could not be left unorganized or unexplored. This led us to explore and dive deep into the world of images. Realizing the significance of the task there is extensive research being conducted on the topic yet there is a need of effective tools that can search for images and videos [4]. In this paper we propose a technique for modifying existing image formats that can give new directions to the field of image searching; furthermore we have devised a new technique for image search engines on basis of underlying image information and its semantics.

Image Retrieval systems are now an active area of research especially after the explosive growth of the World Wide Web. However although much work has been done in Content-Based Image Retrieval-CBIR; techniques to retrieve images on basis of semantic content are yet in a stage of infancy [5]. As pointed out by Eakins [5] in order to achieve the goal of intelligent image retrieval one has to go for a paradigm shift. In this paper we follow this approach by borrowing ideas from Semantic Web and Ontology.

This paper describes the first phase of our research and lays the groundwork for image searching on basis of ontology and semantic annotations. It explores the role that images themselves can play in getting themselves known to search engines: the mechanics which we refer to as ***"Semantic Image."*** SemSearch is a type of search engine which is based on ontology-driven principles leading to semantic-based searching mechanism. It has at its foundations a modification of image formats to include in them information about image taxonomy which will assist in the concept of the "Semantic Image" and will eventually lead to a redefinition of the image search problem. The idea presented in this paper is to standardize the image structure so as to introduce a strong association between the content to be searched and the search engine itself which is lacking in current implementations.

1.1 Motivation

On analysis of image retrieval systems we find that there are commercial engines in the form of Google [6], Alta-Vista [7] etc and the research prototypes are mostly CBIR-based systems. Even though there have been major improvements in text search especially after the famous PageRank algorithm [8], yet image search has not found its ground. We explored this problem with a paradigm shift by borrowing a majority of ideas from the Semantic Web concept [9] and considered the enhancements that could be achieved by giving ontological descriptions of images thereby laying the foundation for ontology-driven semantic search [10].

The rest of this paper is organized as follows. In section 2 we explore the existing image searching techniques and their underlying architecture so that our research can refine existing ideas. We also discuss drawbacks of the previous approaches. In section 3 we describe the proposed architecture and standardizations for image formats in detail. In section 4 we present a summary of our case-study which at this point is the JPEG image format. Section 5 concludes the paper.

2 Existing Technologies for Image Searching

Traditional techniques for image searching have been keyword indexing or browsing. These are employed in the popular commercial engines like Google Image Search or Yahoo! Photo Finder. According to Kherfi et al [3] services users would require from an image retrieval system can be of three types: query-based retrieval, browsing or summarization. Each of these is explored briefly.

In query-based retrieval systems users specify queries and the system retrieves images corresponding to their queries. The queries can be text-based or image-based, or a combination of the two. A text-based query may contain names of objects within the image to be sought, name of the image, or a description of the image or a citation that can be associated with these images. In an image-based query user may be given set of sample images or may be asked to provide his/her own image. After query specification, the system goes through its index to put the image in its proper class that most closely relates to the query. Certain problems may then arise in the retrieved images but Kherfi [3] proposes that these problems can be removed with relevance feedback. The browsing service is provided by text search engines like Lycos, AltaVista and Yahoo! These systems surf the web on a regular basis, record textual information on each page and through automated or semi-automated analysis create catalogs of the information on the Web. These catalogs are hierarchical in nature. We propose a somewhat similar mechanism but extend the ideas to make images self-descriptive as you will see in Section 3. Another service that may be required of a Web image search engine is provision of summaries; the user may require the system to give a title to a set of images collected from the Web in order to perform the images on categorized images under a sub-title. Summaries can complement the catalog if each category of images in the catalog is represented by some representative words or images provided by the summarization service.

2.1 Existing Systems' Review

Many prototypes have been proposed for image retrieval from the Web; these include ImageRover [11], Atlas WISE [12], WebSeek [1], ImageScape [13] etc. We will now explore and throw light on these systems in comparison with Image SemSearch mechanism.

Image retrieval has been a very active research area since the 1970's with thrust from mainly two major research communities: computer vision and database management [14]. Text-based image retrieval borrows ideas from database management

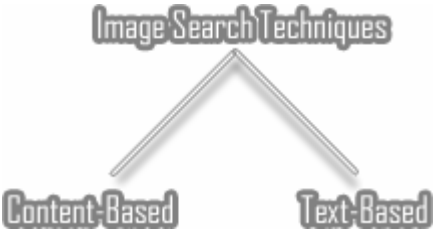

Fig. 1. Current Image Search Techniques

field in which images are first annotated by text and then text-based DBMS's are used for image retrieval. Then in the 1990's content-based image retrieval was proposed in which images are indexed by their own visual content such as color and texture.

With the size and diversity of digital image collections growing at exponential rate efficient image retrieval is becoming increasingly important. In a general sense current image search systems/engines can be categorized into two categories: text-based and image content-based as shown in figure 1.

2.2 Limitations of Existing Approaches

Although these approaches yield effective results yet they suffer from some limitations and this work attempts to remove these limitations by giving a redefinition of the "image search problem." In text-based retrieval systems there is the problem that accompanying the relevant search results, there could be a large number of irrelevant search results i.e. their precision could be low [15]. Content-based image retrieval systems have been on the research scene for quite sometime now but there are still some impediments in their widespread user acceptance as also pointed out by Eakins [5]. This is not because the need for such systems is lacking- there is ample evidence of user demand for better image data management in fields as diverse as crime prevention, photo-journalism, fashion design, trademark registration, and medical diagnosis. It is because there is a mismatch between the capabilities of the technology and the needs of users. The vast majority of users do not want to retrieve images simply on the basis of similarity of appearance. They need to be able to locate pictures of a particular type (or individual instance) of object, phenomenon, or event. A successful solution lies in semantic image retrieval which intelligently retrieves images on basis of users' specific, focused needs and it requires a significant paradigm shift as suggested by Eakins [5]. The information that these systems derive about the images is in a loosely connected form, it can be given more worth by embedding it into the image structure itself in the form of tags and metadata. In short an important limitation in current research directions is lack of a standard foundation upon which to build the searching framework, be it manual annotation of an image or automatic through content-based image processing techniques. What we propose is a modification to the image formats in order to standardize the image retrieval process through ***"Semantic Encapsulation and Annotation of Images."***

3 Proposed Strategy

We now move on to explain our strategy for image searching. We have sought to collect the afore-mentioned techniques on a single platform (refer figure 2) and treat images as being important entities in themselves. By this we mean to make an image more and more self-sustainable and not heavily reliant on the text around it and/or on the HTML tagged information. The image search problem needs a proper redefinition and in such an attempt modifications to image formats are suggested from our end which can prove to be helpful in standardizing the image search problem.

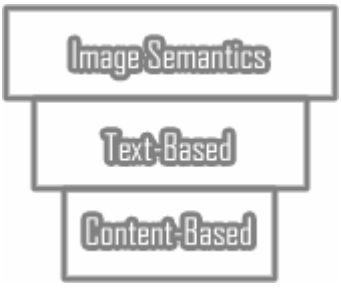

Fig. 2. Proposed Framework for a Layered Approach towards Semantic Searching of Images

3.1 The Search Problem Redefined

We feel that one fundamental aspect has to be an important part of all image retrieval systems. This fundamental aspect lies in the famous phenomenon: "Picture is worth a thousand words." Unfortunately current image retrieval systems are not realizing this phenomenon in essence and when it comes to a computer a picture is worth almost zero words. What has made this task immensely difficult is the computer's reluctance in understanding images. This limitation could be overcome by redefining an image as a self-descriptive entity through ***"semantics"*** and in it the 1000 words should be contained so that it can tell by itself what it is. They already have a visual appeal and tell a lot but this information is conveyed to humans when humans look at the pictures. Image search engines cannot grasp the 1000 words that the picture wants to reveal about itself. This led us to a new idea; what if the image formats that are used on the Web are enhanced and revised with an ad-on in standard. This can be embedded into the image formats in such a way that image itself becomes self teller of the compressed knowledge of 1000 words and that knowledge is compacted into limited words which creates actual sense.

What we are proposing is a modification in the structure for images such that their properties like class (taxonomy), type and description is hidden in the structure. So alongside display information (e.g. RGB) of all pixels there can be an image header which can tell about what exactly a particular image represents. This information about the image can also be added somewhere else in the image structure for instance towards the end of file marker or in the bytes reserved for future aspects. This would naturally lead to ***"a redefinition of the image searching problem."*** Hence unlike the previous approaches where images were searched as integral part of web pages and their categorization was carried out on basis of the text around the images and their tags but this structure will enable the image to have a separate entity and existence of its own.

3.2 Nth Dimensional Image Structure

Data is a not just a simple term today; it takes multi-varied forms and can be considered to possess a multi-dimensional nature. This is what has given rise to new data processing and mining techniques like Online analytical processing [16]. The similar concepts can be applied to images by extending images to nth dimensions where n is an arbitrary number. Greater the amount of detail to be added to the image, greater the value of n.

Now we discuss the concept of image dimensionality in detail. An image is physically two-dimensional view port having certain width and height; the existing search engines do give details of some parameters of images like width, height and pixel depth are available but for image retrieval systems to get more intelligent the image formats should include additional information. It is this information which will add new dimensions to the image and search results can then be refined by virtue of this information.

We suggest embedding image description into the standards for image formats thereby taking the "Semantic Web" idea ahead. The Semantic Web is a philosophy, a set of design principles, a scheme to make data on the Web defined and linked in a way so as to enable efficient discovery, automation and integration of information. The proponents of the Semantic Web philosophy put forward the argument that the vision of making the Web function as an intelligent agent can only flourish when standards are well established. The need for ontological standards has led to development of an independent data representation scheme like XML [9] [17]. Likewise what we are putting forward is an ***"ontological standard for images"*** so as to make image search on the World Wide Web easy and efficient.

We believe that the searching seed should be within the image rather than chiefly within the context like in today's engines. Hence the idea is to combine content-based image retrieval with context-based image retrieval. Following is a representation of the modified image format hierarchy:

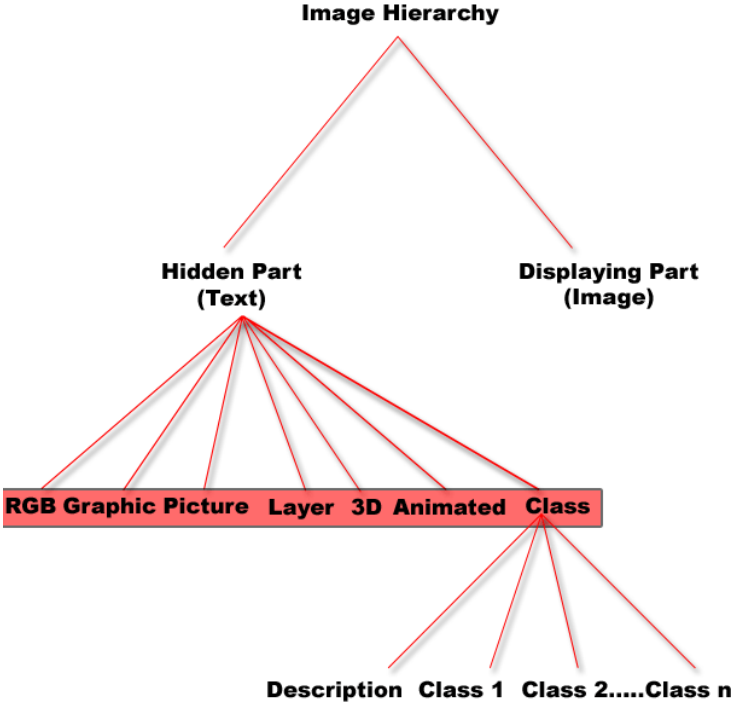

Fig. 3. Image Format Hierarchy

Moreover the proposed additions to image formats would contain seven bits for TYPE in its structure:

Table 1. Description of TYPE bit of semantic image

Bit 1	Bit 2	Bit 3	Bit 4	Bit 5	Bit 6	Bit 7
RGB	G	P	L	3D	A	C

Here RGB represents whether it is palette-based or RGB.

G and P tell if it's a graphic or a picture.
L represents whether image is layered or not.
3D tells whether it a three-dimensional view of an object or simple 2D-image.
A represents whether the image is animated or still.
C tells whether class definition exists or not.

Amongst these seven parameters class is of the greatest significance because on basis of class we can classify the image and it is this attribute of the image which will enhance the search process. Class is further categorized into description or sub-classes. For example with family photographs etc. names of all people in the picture can be given as description of the image and class can be personal family.

Class could be categorized as landscape, sports, personal family, etc. Even these classes can be further classified like sports into cricket, football, hockey etc.

3.3 Semantic Role of Image in Search Problem

Some researchers might argue as to what breakthroughs the proposed scheme can achieve in image search engines: modification of image structures is a mere programming activity and hence would not be capable of achieving much. However what is of paramount significance is the fact that the proposed scheme can introduce standardization into the entire image search process. The main focus of our research is to create a standard for description of an image. Since images have always been a visual entity rather than textual information that is why researches on image retrieval have expressed much weight towards computer vision techniques. However today the processed data from images is only available to a limited scope and becomes unknown or useless to the world due to the absence of a standard. This is what motivated us to take an initiative for supporting available computer vision algorithms in order to enable them to share their processed information by utilizing the capabilities and features of ***"Image SemSearch."*** Hence the trigger point of our research starts with an Image Labeler "***SemImage"*** which can then help realize the semantic role of an image in search engines.

3.4 Towards Intelligent Image Retrieval

As pointed out earlier in this paper though automated retrieval systems have been around for quite some time now yet their widespread user acceptance still has to be achieved. We feel that this has been due to the fact that automation (through image analysis techniques) should be the next step in the process of image searching; first

there is a need of standardization in placing image labels so that image semantics can be defined. One fundamental problem with existing systems is the lack of a standardized approach for image retrieval; once this is addressed we can move towards intelligent image retrieval. Techniques like Google Image Labeler [18] and Facebook's image tags utilize the manual element in a better sense but the lack arises due to absence of a standardized approach.

The image searching process can then be standardized by making the proposed image format technology open for standard image search engines already available and then through their content-based retrieval procedures like edge detection, convolution, power spectrum analysis etc. [5], the image annotation process can be automated. Then to insert that automatically deduced information into the images already available on the World Wide Web the owners of Websites that have those images can be asked to go through some authentication mechanism to verify that they own those retrieved and annotated images. After the authentication process is over, the information about the image that is deduced by the engine can be given to the owner of that image and he can then be asked to insert that information into the image. Of course website owners would want traffic to be redirected to their site and they would surely want to make the digital content of their sites easily searchable through insertion of semantic content into them.

4 Experiments with JPEG File Format

In this section we discuss our image annotation tool "SemImage" that allows users to create the Semantic Image by annotating it with various ontologies in the form of image type, classes and description.

At this point we have experimented with the JPEG File Format hence the type bit will already have some bits filled out; that is the picture and RGB bit will be set to 1 for the reason that JPEG contains both these features. Further the bits for layered, animated and 3D field would be set to 0 since the JPEG file format does not contain support for these features.

Following shows a screenshot of the image viewer of our prototype model in which the class description and type attribute of a particular image can be viewed. The image is that of a brick wall, its name is "khi.jpg" Four classes have been defined for this JPEG image namely Location: Karachi, Pakistan, Building: Karachi Port Trust, Bridge: Kemari over hear Bridge and Link: http://www.pakistanpage.net/gallery/main/cities/karachi.html. All this information is now a part of the image structure.

The approach presented allows a maximum of 255 classes for the image to be specified; further one can set the seven bits for the TYPE parameter.

We have employed a technique similar to steganography in which some information is hidden in an image. As we mentioned in earlier sections the image has a hidden part (semantic part) and display part. Steganography utilizes the concept of hiding information so it served the purpose. Data about the image has been inserted at the end of the EOI (End of Image) marker. This enables the image to be viewed in any JPEG compliant Internet browser or Picture Viewer. Work of a similar nature has been performed in the EXIF file format [18]. However EXIF inserts camera-specific information into the JPEG image unlike content that can give semantic meaning to the image which is the approach that we have followed.

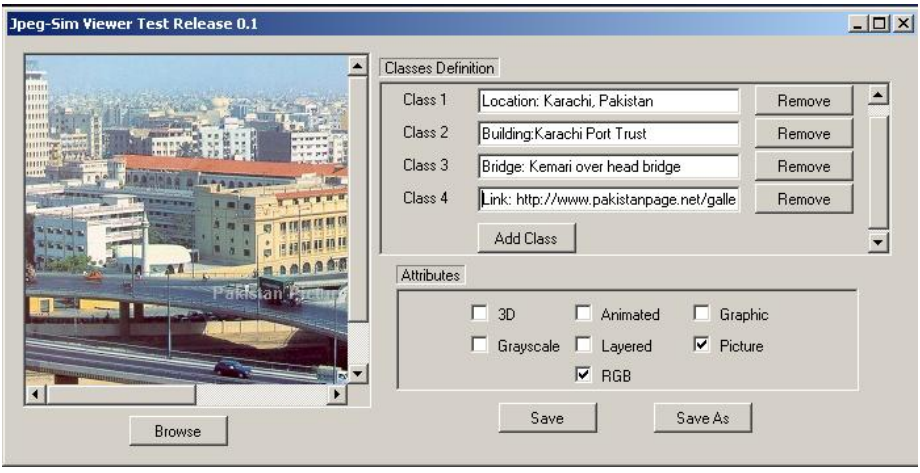

Fig. 4. Sample JPEG for Application of Semantics and Semantic Information about it

One can now easily see that semantics have given a new self-descriptive nature to this image: the search seed is now located within the image itself. This approach can give a major boost to existing techniques by giving them a platform through which semantics can be associated with images. Moreover the idea of custom and personalized search can be realized through the proposed technique.

5 Conclusion

The answer to the "Image Search" problem lies in building a strong association between image and search engine. The shortcoming is not in the database management schemes used for the search engine but rather in the image self-description. This is where image semantics shares the load and distributes it in a better way to lead to a semantic and effective solution to the image search problem.

In this paper we have laid the foundations for Semantic Search in the field of image searching and have attempted to implement Semantic Web technologies within images. These concepts cannot just be limitized to the Internet in fact searching for images on local PC's would also be much easier and the "Semantic Web" can be taken a step further and things such as "Semantic Multimedia" can be introduced.

6 Future Directions

As mentioned earlier this work describes the initial stages of our approach which are geared towards standardization of image formats so that to make their semantic information and their visual content tightly integrated with each other. The future phases of ***"Image SemSearch"*** include introduction of an ontological standard language ***"Image Semantic Language"*** which will assist the process of querying the engine thereby making image retrieval efficient and effective through Semantic Web technologies.

References

1. Smith, J.R., Chang, S.-F.: An Image and Video Search Engine for the World-Wide Web. In: Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases V (IS&T/SPIE, San Jose, CA), pp. 84–95 (1997)
2. Halaschek-Wiener, C., Schain, A., Golbeck, J., Grove, M., Parsia, B., Hendler, J.: A Flexible Approach for Managing Digital Images on the Semantic Web. In: Proceedins of the Fifth International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot) (2005)
3. Kherfi, M.L., Ziou, D., Bernadi, A.: Image Retrieval from the World Wide Web: Issues, Techniques and Systems. ACM Comput. Surv. 36(1), 35–67 (2004)
4. Scarloff, S.: World Wide Web Image Search Engines. In: NSF Workshop on Visual Information Management, Cambridge, MA (June 1995)
5. Eakins, J.P.: Towards Intelligent Image Retrieval. Pattern Recogn. 35(1), 3–14 (2002)
6. Parbury, B.: Google Empire – Technology Behind the Giants. In: 7th Annual Multimedia Conference (2006)
7. Silverstein, C., Henzinger, M., Marais, H., Moricz, M.: Analysis of a Very Large Alta Vista Query Log. Systems Research Center Technical Note, pp.1998–2014 (1998)
8. Brin, S., Page, L.: The Anatomy of a Large-scale Hypertextual Web Search Engine. In: Proceedings of the Seventh International World Wide Web Conference, pp. 107–117, Brisbane, Australia (April 1998)
9. Shadbolt, N., Hall, W., Berners-Lee, T.: The Semantic Web Revisited. IEEE Intell. Syst. 21(3), 96–101 (2006)
10. Bonino, D., Corno, F., Farinetti, L., Bosca, A.: Ontology Driven Semantic Search. WSEAS Trans. Inform. Sci. Appl. 1(6), 1597–1605 (2004)
11. Taycher, S., La Cascia, M.: ImageRover: A Content-Based Image Browser for the World Wide Web. In: Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries, San Juan, Peurto Rico, pp. 2–9 (1997)
12. Kherfi, M.L., Ziou, D., Andbernardi, A.: Atlas WISE: A Web-based image retrieval engine.In: Proceedings of the International Conference on Image and Signal Processing(ICISP), Agadir, Morocco, pp. 69–77 (2003)
13. Lew, M.S.: Next Generation Web Searches for Visual Content. IEEE Comput. 33(11), 46–53 (2000)
14. Rui, Y., Huang, T.S., Chang, S.F.: Image Retrieval: Current Techniques, Promising Directions and Open Issues. J. Vis. Commun. Image R. 10, 39–62 (1999)
15. Baeza-Yates, R., Rebeiro-Neto, B.: Modern Information Retrieval. ACM Press, New York (1993)
16. Somani, A., Choy, D., Kleewein, J.C.: Bringing Together Contenet and Data Management Systems: Challenges and Opportunities. IBM Syst. J. 41(4), 686–696 (2002)
17. `http://www.w3.org/XML/`
18. `http://images.google.com/imagelabeler/`
19. Exchangeable Image File Format for Digital Still Cameras, Exif Version 2.2, Standard of Japan Electronics and Information Technology Industries Association (2002)