Automatic Web Image Annotation for Image Retrieval Systems

Hsien-Tsung Chang
Computer Science and Information Engineering
Chang Gung University
No.259, Wenhua 1st Rd., Guishan Shiang, Taoyuan County 333
Taiwan (R.O.C.)

Abstract: - The number of images is growing in a rapid speed with the popularity of digital cameras and for enriching the visual aesthetics. WWW is becoming a large image library for browsing, and it is an important issue that how to retrieve the images accurately on the web. In this paper we sketch the architecture of an image retrieval system with image annotation. And we propose four methods to generate the annotation automatically for every image from its hosted web page, by analyzing the structural blocks, collecting anchor text of link structures, and gathering shared annotation with other images with the same visual signature.

Key-Words: - Image Annotation, Image Retrieval

1 Introduction

From the invention of World Wide Web by Tim Berners-Lee on November 12, 1990, the population on the Internet is growing in a rapid and uncontrolled speed. The population brings commercial potential onto World Wide Web. In order to attract users staying and returning to the web site, the looks of webs pages are more and more elegant. A large number of images are used to enrich the visual aesthetics. On the other hand, more and more digital images are posted onto the web space with the popularity of digital cameras. Lots of free web album services are announced for users put their images for sharing. For the above two reasons, the number of images on the WWW is increasing in a fast way.

Such a large number of images are posted on the WWW, the text content of web pages is a large free text library as we know, the images are also formed another library. It is becoming an important issue to retrieve images with a quick and accurate method on the web. Unlike the web page search, a lots of text retrieval technique have already developed for a long time. There are many difficult problems for a web image retrieval system.

In the traditional images retrieval systems, they add annotations to each image manually. Although it is a good methodology to retrieve images through text retrieval technologies, it is becoming impossible to annotate images manually due the huge and rapid growing number of web images. Some researches [1] [2] [3] [4] create image retrieval systems by analyzing the image content. They find out the size, shapes, patterns, colors, and distribution of hues [5] [6] [7] of each image, and transfer the properties into a signature for the image. User can specify the properties as the query signature, or submit a similar image and generate the query signature with the same method by the image retrieval system. However, how to select the properties for images is a difficult task. In the past research paper, we can notice that the accuracy is not good enough by this method. The properties of the image with “Tokyo Tower” or one with “Eiffel Tower”, as an example, may be similar, but the semantic meanings of the two images are poles apart. Therefore, even we can develop an accuracy method to select the best properties for an image retrieval system, it is still hard to recognize the semantic difference only by signatures generating. On the other hand, it is also hard for a user to submit a query by specifying the properties or submitting a similar image as a query. Therefore, in some other researches [8] [9], they allow users to submit a text query, and then use the retrieved images by text based retrieval method to generate the signatures. The generated signatures will use as the input for the second phase image retrieval. However, it is easy to understand that the signatures generating is a time consuming process, it will need a lot of hardware and time to obtain the results. It is becoming hard to accomplish through the same reason with annotating image manually.

Large image retrieval systems [10] [11] [12] nowadays on the web allow users submitting keyword queries similar to web page search. They use automatic image annotation and text inverted index to match the users’ queries. In this paper, we will sketch the architecture of the image retrieval system and how it works with the automatic image annotation techniques.

2 Architecture of the Image Retrieval System

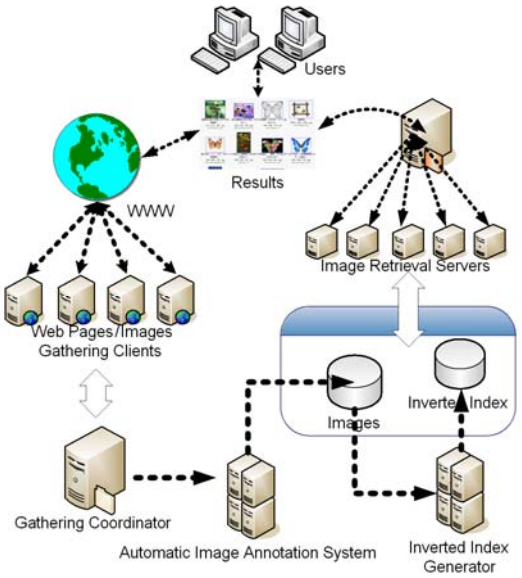

Figure 1 The architecture of the image retrieval system on the WWW

We can model the WWW as a large database. There is a main difference to the standard information retrieval system, that is, we should generate the results to the users in a short time (usually in one second). It is impossible for us to access the remote web pages and images through the network, because it is too slow to bear. We need to store a special copy of the web images, and those would be thumbnails from the original ones.

There are many tasks a web image retrieval system has to accomplish. A whole web image retrieval system can be generally separated into four sub-systems. They are data gathering sub-system, automatic image annotation sub-system, index generator sub-system, and image retrieval sub-system. Figure 1 shows the architecture of the web image retrieval system.

The purpose of data gathering sub-system is to gather the data for the rest of the sub-systems. A simple data gathering sub-system starts with a set of URLs through the protocol HTTP. It then extracts the URLs from the starting pages, following the new URLs recursively by the method of breadth-first or depth-first tracing. After gathering the web pages on the web, we need to extract the image URLs from the pages. And then dispatches the image URLs to the gathering clients for gathering the images. Generally speaking, we design the data gathering sub-system in a distributed model to increase the concurrency and gathering speed. One problem is how to coordinate the gathering robots to avoid traversing a URL many times. Traversing one same URL many times wastes the bandwidth to fetch and time to process. We can design the data gathering sub-system in a centralized model, which is putting one centralized gathering coordinator to arrange the URLs for the gathering clients. Figure 2 shows the flow chart of a data gathering sub-system.

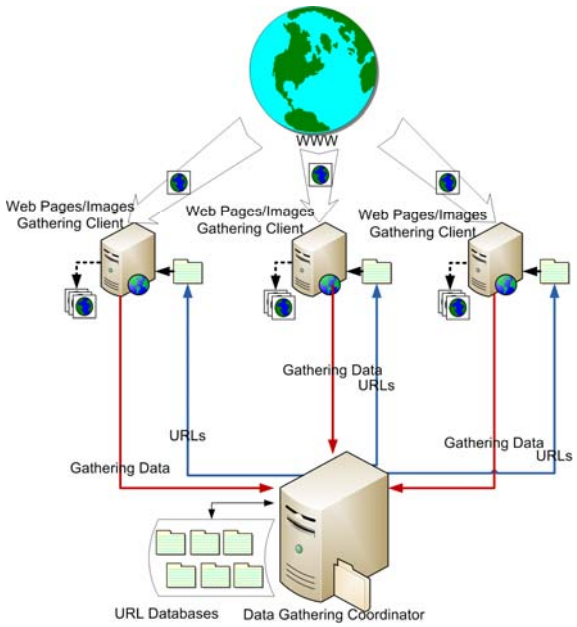

Figure 2 The flow chart of a data gathering sub-system

After the gathering phase, we now have the images and pages in our database for our next phase automatic image annotation. We will make the thumbnails and collect the basic information, like image width, height, format… for all the images and transfer them into our internal format. And then generates the annotation automatically from the database, this will be described in detail next section.

The annotation for all gathered images can be used for building the inverted index through the traditional information retrieval techniques, and we call it as index generator sub-system. The generated inverted index, the thumbnails, and the basic information from the images can be utilized for the image retrieval sub-system. When a user submits a query term contains some keywords to the image retrieval sub-system, it will look up the inverted index to matching the images with the automatic generated annotation which is relative to the keywords. And then show the image thumbnails to the user according to some pre-defined ranking methods.

3 Automatic Web Image Annotation

Web pages are composed by HTML. There are a lot of clues can be used for automatically generating image annotation, especially the links characteristics

and structures. We will introduce the methods how we generate the image annotation automatically by following sub-sections.

3.1 Annotation from the Hosted Web Page

We use the ***<img>*** tag to show an image in a HTML page, the value of the property "***alt***" is a statement composed by some words defined by the author. It is generally to tell page consumers what the image represents when the browser fails to load it. Therefore, it is the best annotation candidate for an image. Generally speaking, the text around the ***<img>*** is describing the image. We can also say that the image is used for emphasizing the text around it. It is impossible for a web page author putting some irrelevant images in the web page. The text around the image can also be treated as an important annotation for an image.

Figure 3 An example of Clown Anemonefish from URL: http://animals.nationalgeographic.com/animals/fish/clown-anemonefish.html

The author of the web page may specify a meaningful title to the page. The specified title can be treated as the name of the page. Generally speaking, the name contains the essential meaning of the page. The specified title can also be the annotation to images within the page. By observation, we can find out that the image with the largest area within the page can share the meaning of the page title. Figure 3 shows that, the web page with title ***"Clown Anemonefish, Clown Anemonefish Profile, Facts, Information, Photos, Pictures, Sounds, Habitats, Reports, News - National Geographic"*** from National Geography for describing clown anemonefish. Although there are many images within the page, we can easily notice that only the image with largest area just share the meaning of the specified title.

3.2 Annotation from Structural Blocks

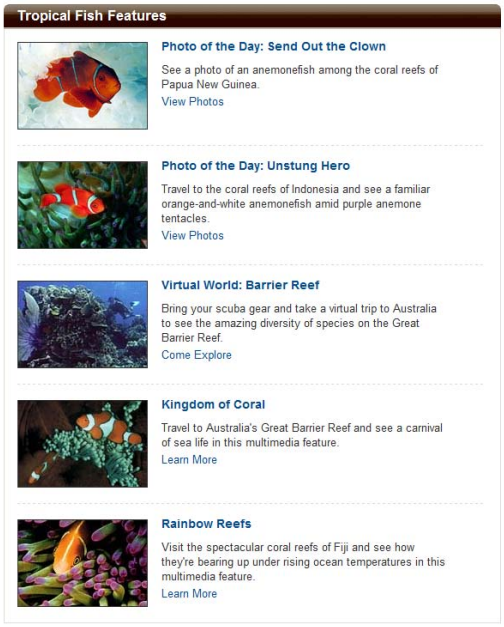

Figure 4 An example of structural blocks annotation in URL: http://animals.nationalgeographic.com/animals/fish/clown-anemonefish.html

The structural blocks are very common in a web page. To analyze and discover the structural blocks can gather the annotation for the image accurately. Figure 4 is an example of the annotation from structural blocks. In fact, we can build the DOM tree for a web page, the DOM tree can help us to find out the structural blocks easily.

3.3 Anchor Text

A link is represented as *** Anchor Text ***. The anchor text of a link assigned by users usually is a good annotation for the linked URL. In other words, the anchor text is the name that others like to call. We can gather all kinds of annotation from different authors to a single image by collecting the anchor text that link to that image. Even we can collect the annotation for the image in different languages, because the anchor text of the links could be assigned by the people from any countries.

The anchor text is widely used for web page search engine. Besides the anchor text for the image, the anchor text for the hosted web page is also acting a good role for the image annotation. As we mention in section 3.1, since the image with the largest area can share the meaning of the page title. It can also share the anchor text of the page.

Besides the anchor text from a link, the structure of links can also be used for ranking the image like popular ranking algorithms, PageRank, and HITS. However, because of the rampant drowthing spam

techniques, the first positions of results ranking by links structure will be occupied by spam sites if you do not apply any spam filterring technologies.

3.4 Shared Annotation

In the traditional images retrieval system, we treat two images as different ones if they have different URLs. However, web page authors offten duplicate images from others sites, and put the original duplicated image or resized one into their web sites. In fact that if the two image are visually the same, even the size or brightness are different, we still can treat these two images the same in our image retrieval system. It is because that, image retrival system users do not concern about which image is the orignal one, they only concern about if they can get the images that are relevant to their submitted query terms. Therefore, we can share the annotation generated from sections 3.1 to 3.4 from visually the same images.

In order to share the annotation, we will generate a visual signature for each image. How to generate the visual signatures perfectly is not the main issue in this paper. We can just use a simple method for visual signatures by resizing every image to a pre-defined dimentions, and then use the MD5 digest method to generate the signature for each image. The visul signatures of duplicated images and original one will be the same, even the duplicated images are resized. Images with the same visual signature could share the annotation to each others.

By sharing annotation from images with the same visual signature can substantially increase the source of annotation for every image. Although the share annotation can not increase the result number of retrieved images with different visual signatures, it can help the ranking algorithm higher the score if one image can gather more shared annotation from others. It is because one image might be good if many authors duplicate it and put it in their own sites.

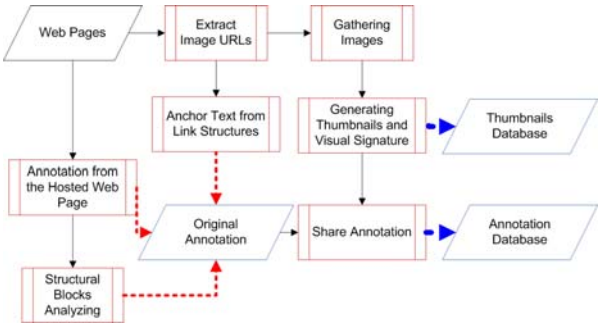

Figure 5 The data flow of automatic image annotation generation.

Figure 5 displays the data flow of automatic image annotation generation. We will generate the annotation for each image through the methods we described in the section 3.1 to 3.3. And then generate the thumbnails and visual signature for each image. The shared annotation can be generated the by sorting the internal format according to their visual signatures. The images with the same visual signature will be put in the continuous location. It will be very easy to collect all the shared annotation in linear time.

4 Experimentation

We make an experimentation to gather some free web albums and e-commerce sites in Taiwan. We first mark some of the images as functional ones by following rules.

A. *Width/Height or Height/Width > 4*
B. *Width * Height < 2500*
C. *Image in GIF format with more than one layer*
D. *The image appears in the single page more than once*

We will take out the functional images, because users are not interesting in these functional ones, for examples: buttons, logos, separators and icons.

# of Web Pages	20,219,113
# of functional images	303,190,414
# of images	49,104,883

Table 1 The information of gathering data

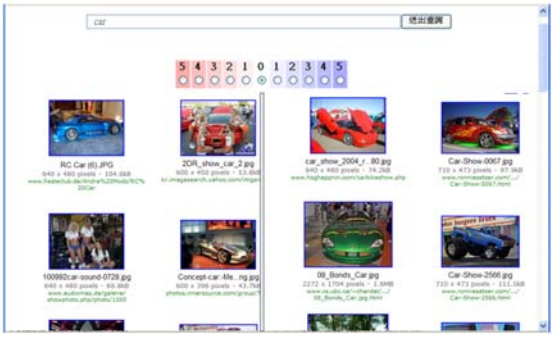

Figure 6 Blind test for the experimentation

Table 1 shows the number of web pages, functional images, and normal images we gathered for this experimentation. We will create two set of annotation for the experiment, the first set is the annotation generated by the methods in section 3.1 to 3.3, and the second set is adding the shared annotation. We select the top 10 query terms from the query logs, and use blind test as shown in Figure 6 to

20 users. The retrieved result from two set will randomly put in the two sides of the result page. Testing users will assign one value from 1 to 5 for the better side. The average value to the first set is 0.8, and the value to the second set is 1.65. We can say that shared annotation can improve the results in an image retrieval system.

5 Conclusion

In this paper, we describe the architecture of the image retrieval system with image annotation. It can be generally separated into four sub-systems, they are data gathering sub-system, automatic image annotation sub-system, index generator sub-system, and image retrieval sub-system. In order to generate the annotation automatically for the image retrieval system, we proposed four methods to gather the annotation for every image. We can first gather the annotation from hosted web page, analyze the structural blocks for accuracy annotation, collect the anchor text through the link structure, and last share the annotation from images with the same visual signature.

How to generate the annotation for every image automatically is an important issue in the context-based image retrieval system. The amount of annotation can be widely increased by our proposed methods. Especially the concept of shared annotation can substantially increase the source of annotation for every image. This will be helpful to an image retrieval system.

References:

[1] V.N. Gudivada and V.V. Raghavan, "Content based image retrieval systems", Journal of Computer, Volume 28, Issue 9, Sep 1995

[2] 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 Transactions on Pattern Analysis and Machine Intelligence, Volume 22, Issue 12, Dec 2000

[3] Y. Rui, T.S. Huang, M. Ortega and S. Mehrotra, “Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval,” IEEE Trans. Circuits and Systems for Video Technology, 1998

[4] J. R. Smith and S.-F. Chang, "VisualSEEk: a fully automated content-based image query system", Proceedings of the fourth ACM international conference on Multimedia, Nov. 1997, Massachusetts, United States

[5] T.A.S. Coelho, P.P. Calado, L.V. Souza, B. Ribeiro-Neto, and R. Muntz, “Image Retrieval Using Multiple Evidence Ranking,” IEEE Trans. Knowledge and Data Eng., vol. 16, no. 4, pp. 406-417 Apr. 2004.

[6] A.K. Jain and A. Vailaya, "Image Retrieval Using Color and Shape", Pattern Recognition, vol. 29, no. 8, pp. 1,233-1,244, 1996

[7] A.K. Jain and A. Vailaya, "Shape-Based Retrieval: A Case Study with Trademark Image Databases", Pattern Recognition, vol. 31, no. 9, pp. 1,369-1,390, 1998

[8] S.-F. Chang, J.R. Smith, M. Beigi and A. Benitez, "Visual Information Retrieval from Large Distributed Online Repositories", Comm. ACM, vol. 40, no. 12, pp. 63-71, 1997

[9] C. Frankel, M.J. Swain and V. Athitsos, "WebSeer: An Image Search Engine for the World Wide Web," Technical Report TR-96-14, Univ. of Chicago, Aug. 1996

[10] Z. Chen, W. Liu, F. Zhang, M. Li, H. J. Zhang, "Web Mining for Web Image Retrieval", Journal of the American Society for Information Science andTechnology 52 (10) (2001) pp. 831-839

[11] H. J. Zhang, W. Liu, C. Hu, "iFind – A System for Semantics and Feature BasedImage Retrieval over Internet" , Proc. 8thACM Multimedia Conf., Los Angeles,USA, 2000, pp. 477-478. 24. Z. Chen, W. Liu, F. Zhang, M. Li, H. J. Zhang, Web Mining for Web ImageRetrieval, Journal of the American Society for Information Science andTechnology 52 (10) (2001) pp. 831-839

[12] S. Bloehdorn, K. Petridis, C. Saathoff, N. Simou, V. Tzouaras, Y. Avrithis, S. Handschuh, Y. Kompatsiaris, S. Staab and M.G. Strintzis, “Semantic annotation of images and videos for multimedia analysis”, Proceedings of the 2nd European Semantic Web Conference (ESWC 2005) 29 May–1 June 2005, Heraklion, Greece

[13] Google Image Search, http://images.google.com/

[14] Yahoo Image Search, http://images.search.yahoo.com/images

[15] MSN Image Search, http://search.msn.com/images/