

User-oriented evaluation of color descriptors for Web image retrieval

Otávio A. B. Penatti and Ricardo da S. Torres

Institute of Computing, University of Campinas (Unicamp),
Campinas, Brazil

`{penatti,rtorres}@ic.unicamp.br`

Abstract. This paper proposes a methodology for effectiveness evaluation in content-based image retrieval systems. The methodology is based on the opinion of real users. This paper also presents the results of using this methodology to evaluate color descriptors for Web image retrieval. The experiments were performed using a database containing more than 230 thousand heterogeneous images that represents the existing content on the Web.

Key words: user evaluation, color descriptors, content-based image retrieval, web

1 Introduction

The growth in the size of image collections and the availability of these collections worldwide is an evident trend. The Web is, nowadays, one of the biggest and most heterogeneous image databases existent. The existence of this huge amount of visual information has increased the demand for image retrieval systems. A promising approach to address this demand is to retrieve images based on image content (Content-Based Image Retrieval - CBIR). This approach considers image visual properties for indexing and retrieval. One of the main visual properties considered by human vision in images is color. In CBIR systems, color is the most common property analyzed and it is the most studied in the literature [1].

A CBIR system is based on *image descriptors*. The image descriptor is composed by [2]: (i) an algorithm for extracting feature vectors and (ii) a distance function. The feature vector encodes information about image properties. Given two feature vectors, the descriptors' distance function computes a distance value. Given a query image, the distance value is used to rank database images.

The query processing time and the results of a search in a CBIR system depends on the descriptors used. Therefore, different descriptors can be used to achieve the system goals. This illustrates the importance of evaluating image descriptors considering different criteria. One important criterium is the descriptor effectiveness, which refers to the quality of the retrieved images.

In the literature, descriptors' effectiveness evaluation is usually made automatically. This evaluation is easier to be made, however it is possible to be

performed only in classified image databases. In a Web environment it is impossible to know the exact number of relevant images given a query image. Another problem with automatic effectiveness evaluation is that the previous database classification assumes that all users agree with the classification made. However, different users tend to have different interpretations about the same image and a single classification cannot be considered as true for all users.

An alternative to solve these problems is to evaluate descriptors considering the opinion of real users. One of the main advantages of a user-oriented evaluation is that the descriptors' effectiveness is measured by the opinion of potential users, what reflects a real environment of use for a CBIR system. Despite the large number of color descriptors in the literature, they are rarely tested in really heterogeneous environments like the Web and using potential users.

This paper performs experiments using a heterogeneous image database containing more than 230 thousand images collected from the Web. The main contributions are: (i) a methodology for effectiveness evaluation for CBIR systems by real users and (ii) the effectiveness evaluation of color descriptors for Web image retrieval.

2 A methodology for user-oriented effectiveness evaluation

The evaluation of image descriptors for CBIR tasks is very important and can be made based on different criteria. Important criteria include the complexity of feature extraction algorithms and distance functions, the storage requirements and the effectiveness of the descriptors. A theoretical comparative study of color descriptors for Web image retrieval is presented in [1].

In this paper we focus on the effectiveness evaluation. Effectiveness measures the descriptor's ability to retrieve relevant images. A descriptor with good effectiveness retrieves the most similar images at the first positions of the results (ranked list), for a given query image. The success of a CBIR systems is closely related to the quality of their results. A user can tolerate a not so fast response, but he or she will not tolerate non-relevant results.

In [3] a user-oriented evaluation is used for digital libraries. The effectiveness of structured and non-structured queries in digital libraries is evaluated by real users. A similar evaluation is proposed here to assess the effectiveness of image descriptors. A Web interface shows to the user the query image and a set of retrieved images. Users are asked to indicate which of that images they consider similar to the query image.

The set of images showed to the user is created as follows. For each descriptor, a list containing the 30 most relevant images is obtained for a given query image. The lists of each descriptor are combined, eliminating duplicates. Every image in the list has a reference to the descriptor(s) that retrieved it and to the rank in its original list. The final list is shuffled and showed to the user.

The query image is showed highlighted at the top of the page and the shuffled list is showed below it. The user has no information about which descriptor

retrieved each image. After the user has indicated the images they considered similar to the query image, they click on a button to finish the evaluation of that query image. The process repeats until all query images are evaluated.

When the user finishes the evaluation, effectiveness measures are computed. The measures used here are: P_{10} , P_{20} , and P_{30} . These measures stand for precision values and they indicate the percentage of images marked as similar among the top 10, 20, and 30 results, respectively. The measures are computed for each descriptor, for each query image and for each user. Therefore, it is possible to compute the descriptors' effectiveness for each query image independently.

3 Experiments and results

The objective of the experiments was to evaluate color descriptors for Web image retrieval. The image database used was collected by researchers from Federal University of Amazonas (UFAM), Brazil. The database collection was made recursively from addresses found in the Yahoo directory. The final collection contains 234.828 images and more than 1 million of HTML documents. The image database has no classification.

Five color descriptors were evaluated in this paper: *global color histogram* (GCH) [4], *color autocorrelogram* (ACC) [5], *color structure* (CSD) [6], *border/interior pixel classification* (BIC) [7], and *color bitmap* [8]. They were chosen because of their simple algorithms for feature extraction and distance computation. Also, GCH, ACC, and CSD are important descriptors from the literature.

The effectiveness evaluation used 16 query images. These images have a well defined semantics and represent different image categories. 69 subjects were invited to take part in our experiments. Users that have not evaluated all the 16 queries were discarded. Therefore, the complete evaluation considered 15 users.

Table 1 presents the average values of P_{10} , P_{20} , and P_{30} among all query images for each descriptor. The results indicate that BIC descriptor has the best average precision for 10, 20, and 30 retrieved images. ACC and GCH achieved the second and third best average precision, respectively. CSD and Color Bitmap achieved similar average precision values. Although BIC descriptor has presented the best average precision, it was not the best for all query images ¹.

Table 1. Average values of P_{10} , P_{20} , and P_{30} among all query images.

Descriptor	P_{10}	Descriptor	P_{20}	Descriptor	P_{30}
BIC	0.31	BIC	0.21	BIC	0.17
ACC	0.27	ACC	0.18	ACC	0.15
GCH	0.25	GCH	0.17	GCH	0.13
CSD	0.18	Color Bitmap	0.12	Color Bitmap	0.10
Color Bitmap	0.17	CSD	0.12	CSD	0.09

¹ The precision tables for each query image are available at: http://www.lis.ic.unicamp.br/~otavio/cbir_eval/tables.pdf

Considering a Web scenario, it is very important that the relevant images appear at the top positions of the results. This indicates that it is more important to have higher values of P_{10} than P_{30} .

It is possible to note that the precision values decreased as the number of images increased. This indicates that when more images are retrieved, the descriptors retrieve more non-relevant images than relevant images.

The analysis of the precision values for each of the query images led us to some conclusions. Descriptors have presented higher precision values for queries with homogeneous background. Another aspect was the influence of the image semantics during the evaluation. Users tend to consider similar images with similar semantics and not necessarily with similar visual properties.

4 Conclusions

This paper presented a methodology for user-oriented effectiveness evaluation and the results of the evaluation of color descriptors for Web image retrieval. The experiments were performed in a heterogeneous image database with more than 230 thousand images collected from the Web. The results indicate that good color descriptors are BIC and ACC. Future work includes the evaluation of more descriptors and the inclusion of more users in the evaluation process.

Acknowledgments. Thanks to Fapesp (process numbers 2006/59525-1 and 2009/10554-8), Capes, CNPq, and Microsoft Research for financial support.

References

1. Penatti, O.A.B., Torres, R.d.S.: Color descriptors for web image retrieval: A comparative study. *XXI Sibgrapi* (2008) 163–170
2. da S. Torres, R., Falcão, A.X.: Content-Based Image Retrieval: Theory and Applications. *Revista de Informática Teórica e Aplicada* **13**(2) (2006) 161–185
3. Gonçalves, M.A., Fox, E.A., Krowne, A., Calado, P., Laender, A.H.F., da Silva, A.S., Ribeiro-Neto, B.A.: The effectiveness of automatically structured queries in digital libraries. *JCDL* (2004) 98–107
4. Swain, M.J., Ballard, D.H.: Color indexing. *International Journal of Computer Vision* **7**(1) (1991) 11–32
5. Huang, J., Kumar, S.R., Mitra, M., Zhu, W., Zabih, R.: Image indexing using color correlograms. In: *CVPR*. (1997) 762
6. Manjunath, B.S., Ohm, J.R., Vasudevan, V.V., Yamada, A.: Color and texture descriptors. *IEEE Trans. on Circuits and Systems for Video Tech.* **11**(6) (2001) 703–715
7. de O. Stehling, R., Nascimento, M.A., Falcão, A.X.: A compact and efficient image retrieval approach based on border/interior pixel classification. In: *CIKM*. (2002) 102–109
8. Lu, T., Chang, C.: Color image retrieval technique based on color features and image bitmap. *Information Processing and Management* **43**(2) (2007) 461–472