Integrated Search Based on Image Contents

Ramon Mas Sansó, Xisco Tous Llull, Joan Oliver Mesquida, Felip Salas
Ramon Mas Sansó:
Universitat de les Illes Balears
Unidad de Gráficos y Visión por Ordenador
Universitat de les Illes Balears
ramon.mas@uib.es.
Xisco Tous Llull, Joan Oliver Mesquida and Felip Salas:
Fundació Illes Balears per la Innovació Tecnològica (IBIT)
Parc BIT Ed.17(3,D-2), 07121, Palma de Mallorca
{xisco,fsalas,joliver}@ibit.org

Abstract. This works describes a new metaphor to make a contents search in an image database and to recover the results obtained using an automatically selected bandwidth. The increasing use of electronic commerce and the consequent publication of image catalogues through internet make evident the necessity to incorporate alternative mechanisms to the traditional search by key-words. Our work explores the scope of image-content access of databases. We use textual description, shapes, patterns and color similarities. At a later stage and due to the cost of the transmission of high quality images, we analyze the available bandwidth and we adapt the size of the resulting images to cope with the user requirements.

1 Introduction

The increasing use of electronic commerce and the consequent publication of image catalogues through internet make evident the necessity to incorporate alternative mechanisms to the traditional search by key-words. This is especially true in cases where the search is based on criteria subjective and hard to include as text fields. In this article we explore the scope of content based access to databases of images by means of alternative searches to the textual description such as shapes, patterns or color percentages. The methods and metaphors presented try to provide any customer who deals with huge image databases with a portable and easy to integrate methodology that allows the use of search and recovery of images by content descriptions. The methodology we propose is based on two phases. First we use off-line image processing tools to automatically recognize features. Second, the information obtained is used to index a database of images. In a later stage, the user

performs a search using any of the available criteria and the resulting series of images is returned. The quality of the resulting images is automatically adapted to the available bandwidth using the constraints imposed by the user (maximum download time).

The results obtained can be easily added to existing on-line product catalogues of industries where appearance of the goods plays a significant role, so as clothes, ceramic tiles, wallpaper and other. This methodology has been applied to the interactive search in a real clothes catalogue. In such applications a customer often is not sure of what his looking for, but has some clues concerning the final destination of the product (the color which best suits in the environment, the shapes it should represent or even a similar pattern he has previously seen or thought about).

2   Image search in internet

Content-based search techniques have been applied to medical images [1], art work[2], satellite imagery indexing [3][4] and videoclips [5]. However, the search of images in internet servers traditionally uses textual descriptions. Nowadays, there is ongoing research in the description of images using external annotations [6], author annotations [7] or anthologies. Except in a number of cases where alternative mechanisms can be used to search images, as for instance color contents and distribution [2] or direct selection [8], there is no tool using multimodal integration. In [2], a IBM's experimental query by image content is used. The user can arrange areas of color to approximate the visual organization to look for work of art in a museum database. The user-drawn canvas is used as a grid of colored areas that is matched to other images stored in a database. This tool is classified as a color distribution search.

In this paper we present a contribution to the semantic web by specifying a system that allows a simultaneous search by either textual or contents description integrating author described metadata labels following the Dublin Core standard [9], color tolerance, color percentages, spatial similitude or object shapes.

Those applications, often linked to internet, work with three different kinds of actors: users, image owners and server owners. Users are identified as clients browsing the application in internet; they connect to a server to perform an image search. Image owners and server owners may or may not be the same. In our research we also have in mind three corresponding levels of application. First, the access module or front-office attends the users requests, implements the search process and sends the resulting images in the most convenient size and resolution (depending on available bandwidth). Second the analysis module or owners back office that helps the actor to maintain the information system and automatically classify the images when they are inserted into the system. Third, the administration back office (in case of a hosting service) which should establish the system configuration parameters and the management of the owners.

To make the description process easier we consider that the owner of the images is also in charge of the maintenance of the server. With this assumption in mind, we do not lose any generality as the user can still add images to the server through

XML-SOAP messages (see figure 1). One of the most important constraints is the preservation of managed images with the same quality as they were entered in the system. We use compressed tiff format to store the image, a low quality image for image previewing and jpeg compression for image downloading, although other available image compression formats can be easily integrated.
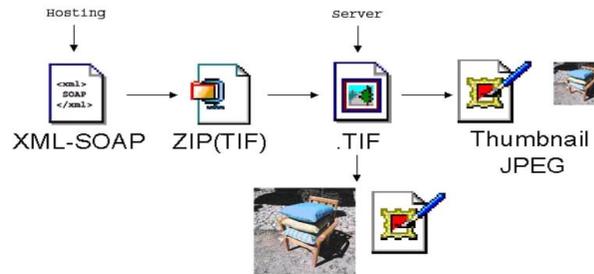

Fig. 1. Image storage and retrieval

3 Image analysis

First, during the database creation or addition of information, we automatically execute image processing algorithms to classify the images. When images enter the system we first apply a histogram analysis for color properties classification. A gradient operator detects the edges in the image and they are considered as characteristic points during the Hough transform. This transformation allows us to create a shape taxonomy. Similarities comparison is done through correlogram computation [10]. This preprocess is very important in the search process, as it drastically reduces the time search at the only cost of a non-significant amount of additional space on the database.

Although the histogram ignores the spatial organization of the pixels in an image, it is easy and fast to compute and it provides a graphical representation of the color frequencies [11]. Chromaticity moments have also been used [12] but they have a lower effectiveness in this area of application as they use a chromaticity diagram that characterizes too globally the shape and distribution of the image.

Correlogram describes the correlation of the image colors as a function of their spatial distance. Due to the huge amount of data to be processed, we use the autocorrelogram. The autocorrelogram is a subset of the correlogram that computes the probability of finding identical colors at a given distance. We keep a table describing the variation of spatial correlation between pixels of the same color.

The detection of object shapes in the image is performed using the Hough transform. We have successfully tried Hugh [13] to detect the existence and position of related lines in the image although it can be easily extended to other geometric entities [14].

While entering an image into the database, the user can also insert annotations using the metadata of the standard Dublin Core ontology. Metadata is coded in RDF (Resource Description Framework http://www.w3c.org/RDF) and is added to the image file.

The decision of including metadata in the image file has been taken to facilitate the interchange of this kind of data between different applications. So that the image creator can use his own application to label images directly.

4   The search process

The search process describes how the user can request images from the server database. Mainly, three criteria can be used:

- Color search: using as input a color tolerance, a percentage distribution of the colors and its tolerance we search in the image database comparing the frequencies of colors in each image. The color information is stored in normalized histograms.

- Similarity search: this criterion is based on the use of correlogram. The input parameter is an image, either from the system or a newly supplied one. Prior to the search, we compute the correlogram of the input image. Once we have the correlogram we compare it to those of the stored images. As a result we get a vector of pointers to images with a decreasing similarity level.

- Shape search: we use the results obtained when the Hough transform was applied to search the existence and/or the position of geometric entities in the image. As stated previously, the initial tests are done using only line detection although we currently extending to other simple shapes. When we detect a geometric attribute, we compute both its orientation and its location in the image. The image is divided in nine sectors in order to make the search request easier to the user. We identify the sectors where the shape belongs and later we use them to establish the matching criteria for the search.

- Textual search: as in any database, we can search using a text description of the image. In this case the affected fields are metadata included in the image.

5   Preliminary results

The resulting methodology of our research, based on the integration of the image analysis, metadata and geometric description has been applied to the interactive search in a real clothes catalogue. In such application a customer often is not sure of what his looking for, but has some clues concerning the final destination of the product (the color which best suits in the environment, the shapes it should represent or even a similar pattern he has previously seen or thought about).

To verify the efficiency and response time of the searches we have used a database containing more than 200 images of regional typical clothes of different colors and patterns. Following we show the results we have obtained with the searching criteria.

Let ask for the images 25% ping (RGB 266,141,146). The results obtained are three images that we show in figure 2. The response time is less than a second (including transmission time). In fact, on a test of 25 color searches we computed a mean of 0.5 seconds.
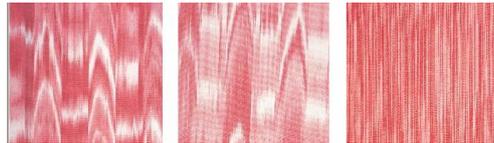
Fig. 2. a search using a 25% of RGB(226,141,146)

As can be easily foreseen, a decreasing importance in the percentage of appearance of a certain color remarkably increases the number of recovered images (figure 3).

Fig 3. a 6% RGB(10,79,136) search

Concerning the similarity searches, we have noticed an important increase in the response time which is due to the much more time consuming process needed in the server.

Figure 5 shows the result of a similarity search using the pattern in figure 4. The response time (from the request to the visualization of the images) in this search is of 18 seconds. We have to take into account that most of this time is spent comparing the pattern against the 200 correlograms in the data base.

The fastest searches are those using the text information as they can profit the indexing system of the database management system. The response in these cases is instantaneous and the only time delay is the transmission of the thumb images that need to be displayed.

**Fig. 4.** Pattern used in the similarities search
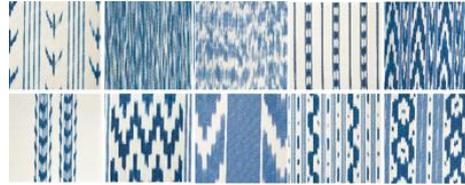
Fig. 5. It took 18 seconds to obtain those 10 images.

Finally we will discuss the results obtained using the geometric feature detection in the images. Figure 6 shows the results for a request of all the images with three horizontal lines distributed in the upper centre and lower sectors of the images. Let observe that the request does not exclude the images with more than three lines, as we are asking for images with at least three horizontal lines. The images were obtained in less than a second.
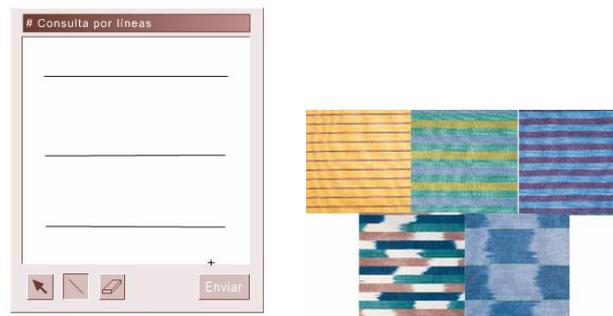

Fig .6. Three parallel horizontal lines (in the upper, middle and lower sectors)

The most important drawback in this search is the detection of characteristic points in the image to be used in the Hough transform. Although they are automatically detected using image filtering and binarization, they are sensitive to the nature of the image and a change in the application nature may require some minor changes in the parameters.

6   Image download

Downloading images includes the computation of the available bandwidth and the automatic selection of the image size and resolution to the effectiveness of the real time connection.

We determine the bandwidth available by sending information packets from the server to the client and tracking the elapsed time. The algorithm uses the following steps:

- First a bandwidth approximation is computed sending a small data package. The size of the package is system configured.

- Using the values obtained in the previous step we set the size of the next package to be sent. The main goal being to refine the computation of the real bandwidth.
- Then we send the image and use its size to verify the bandwidth.

The elapsed time since the request is computed in the client side. Once we effectively know the bandwidth of the connection, we compute the size of the image depending on the user configurable maximum transmission time. Our compression algorithm uses a precomputed table of values to adjust the size of the image.

7   Conclusion

In this paper we have shown that image taxonomy in terms of characteristics extracted automatically is a difficult process due to the heterogeneity of the images. However, if we work in controlled domains the process is highly simplified and the results are more than acceptable. Although performance comparison with other described methods is never easy due mainly to the difference of size in image databases and image quality, the mean response times we have obtained are of the same magnitude than other techniques despite the additional search criteria we incorporate. Qualitative results in our field of application, as shown in section 5, are good. However, to obtain quantitative comparisons we should all work using the same image database and this is not always possible as property rights are involved.

We have presented a new method for multimodal searches that allows the user to search images by their contents using textual descriptions, color percentages, geometric features or similarities with a defined pattern.

Although an annotation based description of images is easier and more reliable as it is a person who describes the images using his own criteria, the process is expensive and very subjective. The annotations have to be reusable so we must use standard formats and interoperativity techniques to ensure the compatibility with most of the applications. In our research we use the Dublin Core ontology.

We also automatically adapt the quality of the final images so as to ensure that the available bandwidth allows a fast transmission of the results.

In our research we have tried to unify both the automatic and the annotated classifications in a unique tool following the semantic web specifications and offering additional functionalities oriented to the search and transmission of images in internet.

References

1. H. Müller, N. Michoux, D. Bandon, A. Geissbuhler, A Review of Content Based Image Retrieval Systems in Medical Applications- Clinical Benefits and Future Directions, International Journal of Medical Informatics,vol.73, pp 1-23,2004.

2. The state Hermitage Museum: digital collection, last visited may 2006, http://www.hermitagemuseum.org.

3. G.G. Wilkinson, "Results and Implications of a Study of Fifteen Years of Satellite Image Classification Experiments", IEEE Transactions in Geoscience and Remote Sensors, **43**(3), pp. 433-440, 2005.

4. A. Parulekar, R. Datta, J. Li, J. Z.Wang, Large-scale Satellite Image Browsing using Automatic Semantic Categorization and Content-based Retrieval, International Workshop on Semantic Knowledge in Computer Vision, Beijing, 2005.

5. S.W. Smoliar, H. Zhang, Content Based Video Indexing and Retrieval, IEEE Multimedia, **1**(2), pp 62-72, 1994

6. Marja-Riitta Koivunen y Ralph R. Swick. Collaboration through Annotations in the Semantic Web. Annotation for the Semantic Web. IOS Press. 2003.

7.Siegfried Handschuh and Steffen Staab. Annotating of the Shallow and the Deep Web. Annotation for the Semantic Web. IOS Press. 2003.

8. http://www.classicroses.co.uk, last visited may 2006

9.Dublin   Core   Metadata   Element   Set   Reference   Description, http://purl.org/dc/documents/rec-dces-19990702.htm, last visited april 2006.

10. J. Huang, S. R. Kumar, M. Mitra, W.-J. Zhu, R. Zabih, Image indexing Using Color Correlograms, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico (1997).

11. J. Hafner, et al., Efficient Color Histogram Indexing for Quadratic Form Distance Functions, IEEE Transactions on Pattern Analysis and Machine Intelligence, **19**(7), 1995, pp. 729-736.

12. G. Paschos, I. Radev, N. Prabakar, Image Content-Based Retrieval Using Chromaticity Moments, IEEE Transactions on Knowledge and Data Engineering, **15**(5), 2003, pp 1069-1072.

13. R.K.K. Yip, Line patterns Hough transform for line segment detection, *IEEE Transactions on Image Processing*, 1995, pp. 319-323.

14. V. Leavers, Shape Detection in Computer Vision Using the Hough Transform, Springer-Verlag,   New   York,   1992.