# A Visual Search of Multimedia Documents in ImageCLEF 2014

Sana FAKHFAKH, Belhassen AKROUT, Mohamed Tmar, and Walid MAHDI

Laboratoire Miracl ISIMS Pole Technologique de Sfax BP 242-3021, Sakiet Ezzit Sfax. TUNISIE. sanafakhfakh@gmail.com, bakrout@gmail.com, mohamed.tmar@isimsf.rnu.tn, walid.mahdi@isimsf.rnu.tn http://www.miracl.rnu.tn

**Abstract.** The Image search by content is an area that is based on a set of low-level features such as histograms, textures, the distribution of colors, shapes an brightness. The structure element is a shortcut of the image may vary depending on the designer of the XML document and which may change according to application needs. The visual appearance of the image is a permanent factor that undergoes no change. Therefore, we present in this paper a new method of search by content based on Harris detector, Haar wavelet and Color histogram for the step of feature extraction to extract a binary signature for all multimedia documents.Finally We estimate the similarities between codes and search relevant images with Hamming distance.

**Keywords:** Future extraction, Plant identification, Hamming distance, Signature extraction

# 1 Introduction

The research of the image content was the subject of numerous studies in recent years. Techniques developed aim to extract and represent effectively the content of images. The image indexing can be done either on the basis of low-level features or those of semantics.

The need for methods of indexing and retrieval directly based on the content of the image is no longer evident. The first prototype system has been proposed in 1970 and this system has attracted the attention of many researchers. Some systems become commercial systems such as QBIC (Query By Image Content) IBM [1] and CIRES [2].

The Frip (Finding Regions in the Pictures) system [3] KoB05 proposes to research by regions of interest to drawn by the user [4]. RETIN system (REcherche and Hunt Interactive) [5] developed at the University of Cergy-Pontoise France, selects a random set of pixels in each image to extract their color values. The texture of these pixels is obtained by applying the method of Gabor filters [6]. These values are grouped and classified using a neural network.

The comparison between the images is done by calculating the similarity between their feature vectors [7]. Some studies have proposed to change their field of research, for example changing the color feature space [8].

Sometimes the user wants to query such as "Find all the images of the database with certain regions similar to those of the query image" [9]. Some work has been done, or the system segments an image or a spatial structure is used. In the first case, having a segmented image, the user selects some regions that we want from the segmented image. In the second case, the spatial structure is used. A structure is often used quadtree [9]. This structure is used to store the visual characteristics of different image regions and filter images by increasing progressively as the level of detail.

Some work has proceeded to minimize the search space using the technique of calculating the nearest neighbors to group similar data in classes [10]. Thus the search for an image is done by searching a class. The disadvantage of these systems is that the user does not always have an image that expresses real need, which makes use of such a system difficult.

One solution to this problem is the technique of vectorization, which allows to find relevant images to a query and which are not returned by the initial system. Its sequence requires the choice of a set of said reference material. The choice of these references is still problematic for the construction of the vector space.

In this context, it is necessary to select a heterogeneous set of documents. The number of references is also problematic. Indeed, the number of documents is the new dimension of vector space. If we choose a small set of documents, we may not be able to properly represent all documents. These references are randomly selected by [11], or the first results of an initial search or through the selection of the centers of gravity of homogeneous grouping images [12].

# 2 Proposed approach

In this section, we propose a description of the proposed visual system. In fact, the search for multimedia documents requires a critical step in the abstract of-fline indexing. This step requires the feature extraction from the image to give a general description of documents. Following this course indexing we can go in search of the visual image based on the descriptors already studied. Figure 1 describes the steps mentioned by an explanatory diagram.

#### 2.1 Feature extraction from the image

The objective of extracting descriptors is the set up the search method contained in the images. One example is the search for a database containing images of plants mentioned. Images are indexed from the local descriptors such as Harris detector [13] to calculate point of interest, the Haar wavelet [14] to analyze the

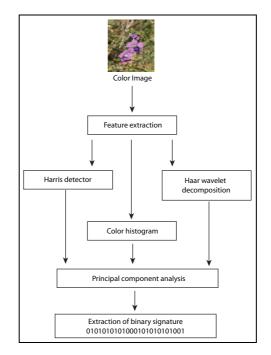


Fig. 1. Diagram of indexing steps

textures and the calculation of the histogram color to represent the distribution of intensities.

Harris detector For twenty years, several detectors points of interest have been developed. Schmid has compared the performance of several of them [13] [15]. The famous Harris detector was evaluated as the best point detector and has a good reputation in the field of automatic extraction points of interest. In fact, Harris and Stephens [13] defined a sensor that recognizes the first derivatives on a window signal representing an image and provide better results in two important criteria [15]. The first criterion is based on their robustness with respect to rotation and scale changes of the camera and points of view. The second

is summed up in his power with respect to variations of illumination. Harris detector was applicable at the beginning only the grayscale image. Montesinos [15] generalized Harris detector for color images. Figure 2 shows an example of detection of interest points in an image.



Fig. 2. Figure showing the detection of points of interest by the Harris detector

The decomposition of the image by Haar wavelet The wavelet transform of the sinusoid replaces the Fourier transform by a family of translations and dilations of a single function. The translation parameters and expansion of the two arguments are the wavelet transform.

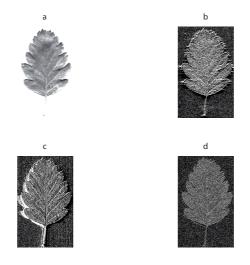
Haar transform was introduced in 1910 and is the oldest wavelet transform. It is based on the demonstration of Haar (equation 1).

$$h(j,k;x) = \frac{2^{j/2}}{2H(2^j x - k)} \tag{1}$$

With H(x) = 1 for x in ]0, 1/2[, H(x) = -1 for x in ]1/2, 1[ and H(x) = 0 ailleurs form a complete orthogonal set for space  $L^2(R)$ . The Haar wavelet is simple to study and implement. Figure 3 shows the result of decomposition of an image texture to analyze and extract the diagonal details, the vertical details, the horizontal details and finally the description.

**Image analysis by color histogram** For a monochrome image, the histogram is defined as a discrete function that maps each intensity value the number of pixels taking this value. The determination of the histogram is made by counting the number of intensity for each pixel of the image. The quantification, which includes several intensity values in one class can better visualize the distribution of image intensities.

Our job is to build a histogram of a color image in the RGB space, the red, green and blue. Each component is quantized into eight intervals. In each interval of each Red, green and blue component; we calculate the number of pixels normal-



**Fig. 3.** Decomposition of the image using the Haar wavelet, (a) image description, (b) horizontal details, (c) vertical details, (d) diagonal details

ized frequency. We can then represent the image by a color histogram as shown in Figure 4 following.

## 2.2 Principal component analysis and signature extraction

The principal component analysis (PCA), developed in France in the 1960s by JP.Benzcri is an exploratory statistical method used to describe a wide array of data type individuals / variables. When individuals are described with 5 large numbers of variables, no simple graphical representation allows to visualize the point cloud formed by the data.

The PCA provides a representation in a space of reduced size, thereby allowing to highlight any structures in the data. For this, we look for subspaces where the projection of the cloud deforms the least possible initial cloud.

Applying the PCA to the descriptors, previously computed, A signal f(x), that includes all the characteristic values, is obtained. This signal is then subjected to an automatic thresholding (equation 2) to the transformed into a binary signature C(x) (equation 3).

$$M_f = \sum_{n=1}^{m} f(x) \tag{2}$$

In this way each image is represented by a binary code which is stored in a database for later use in the research phase.

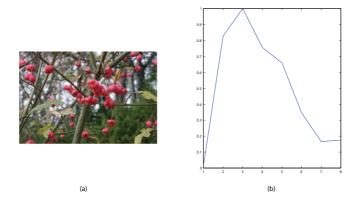


Fig. 4. An example of color histogram applied to an RGB image

$$f(x) = \begin{cases} C(x) = 1 \text{ if } f(x) >= M_f \\ C(x) = 1 \text{ if } f(x) < M_f \end{cases}$$
(3)

#### 2.3 Estimate similarities between codes and search relevant images

Once the signatures are extracted, we must think of a technique that quantifies the distance between them, in order to identify similar images. The bit-wise comparison of each two image codes A and B each detail is given by the normalization of the Hamming distance (HD) [16] defined as following:

$$HD = \frac{1}{N}A(k) \otimes B(k) \tag{4}$$

With  $\otimes$  is the boolean operator (XOR) described by the equation 5 follows

$$a \otimes b = \overline{a}b + a\overline{b} \tag{5}$$

The probability that two bits of arbitrary code images are equal or not, worth P = 0.5. When the majority of codes A(j) and B(j) are equal, is therefore in the case where the two codes compared are equivalent. In the opposite case, the two codes are different. Generally, when HD tends to zero it is called relevant images.

Our group submitted just one run in your rst participation in the ImageCLEF Plant task 2014. In this paper we described a plant species retrieval model based on a extraction of unique signature for each image in collection. This signature is composed by Haar and Harris features.

## **3** Experiments and Results

### 3.1 Evaluation Metric

The metric evaluation S was related to the rank of the correct species in the list of retrieved species as follows [17]:

$$S = \frac{1}{U} \sum_{u=1}^{U} \frac{1}{P_u} \sum_{p=1}^{P_u} \frac{1}{N_{u,p}} \sum_{n=1}^{N_{u,p}} S_{u,p,n}$$
(6)

Where U: is the number of users.

 $P_u$ : number of individual plants observed by the  $u^{th}$  user.

 $N_{u,p}$ : number of pictures taken from the  $p^{th}$  plant observed by the  $u^{th}$  user.  $S_{u,p,n}$ : score between 1 and 0 equals to the inverse of the rank of the correct species (for the  $n^{th}$  picture taken from the  $p^{th}$  plant observed by the  $u^{th}$  user).

#### 3.2 The Results

IV processing team has submitted a single automatic run to the multi-image plant observation queries. In this run, we extracted visual descriptors from images, used texture, color and interest points model for generating the feature vectors. The result is showed in Table 1. At a closer look, the score values of our results are rather low. We cannot fully explain the failure with these plant image as there were no indications for the weak performance during the training stage. One reasonable interpretation of the results is that the conducted learning runs led to an over-fitting.

Table 1. Results for the run in plant collection

Run name	Retrieval type	Score
IV Processing Run 1	Visual	0,043

A total of 10 participating groups submitted 27 runs focusing on plant observations. 6 teams submitted 14 complementary runs on images in ImageCLEF 2014. The following graphic shows the scores obtained on the main task on multi-image plant observation queries.

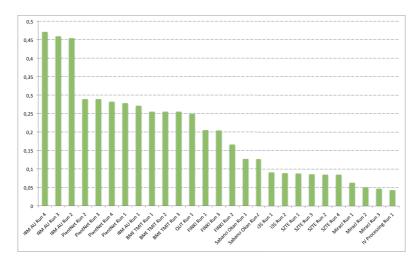


Fig. 5. Scores for Plant task

# 4 Conclusion

We presented in this paper a research method by which the content is based on extracting a signature for a given image. This signature is composed by a set of descriptors, such as the analysis of the histogram of the color image, the extraction of the interest points and Haar wavelet. This signature undergoes a step dimensionality reduction by PCA method. An experimental study on ImageCLEF 2014 showed the effectiveness of our descriptors.

## References

- Myron Flickner, Harpreet Sawhney, Wayne Niblack, Jonathan Ashley, Qian Huang, Byron Dom, Monika Gorkani, Jim Hafner, Denis Lee, Dragutin Petkovic, David Steele, and Peter Yanker. Query by image and video content: The qbic system. volume 28, pages 23–32, Los Alamitos, CA, USA, September 1995. IEEE Computer Society Press.
- 2. Kashif Iqbal, Michael O. Odetayo, and Anne James. Content-based image retrieval approach for biometric security using colour, texture and shape features controlled by fuzzy heuristics. volume 78, pages 1258–1277, Orlando, FL, USA, July 2012. Academic Press, Inc.
- 3. ByoungChul Ko and Hyeran Byun. Frip: a region-based image retrieval tool using automatic image segmentation and stepwise boolean and matching. volume 7, pages 105–113, 2005.
- Yves Caron, Pascal Makris, and Nicole Vincent. Caractérisation d'une région d'intérêt dans les images. In EGC, pages 451–462, 2005.
- J. Fournier, M. Cord, S. Philipp-Foliguet, and F Cergy pontoise Cedex. Retin: A content-based image indexing and retrieval system. 2001.

- 6. Carlos Rivero-Moreno and Stphane Bres. Les filtres de Hermite et de Gabor donnent-ils des modles quivalents du systme visuel humain? In ORASIS 2003, Journes francophones des jeunes chercheurs en vision par ordinateur, pages 423– 432, May 2003. Article apparu en confrence nationale avec actes de congrs et comit de lecture.
- 7. Sabrina Tollari. Indexation et recherche d'images par fusion d'informations textuelles et visuelles. PhD thesis, 2006.
- 8. Jean pierre Braquelaire and Luc Brun. Comparison and optimization of methods of color image quantization. volume 6, pages 1048–1051, 1997.
- Mimoun Malki, Madjid Ayache, and Mustapha Kamal Rahmouni. Rétro-ingénierie des bases de données relationnelles : approche basée sur l'analyse de formulaires. In *INFORSID*, pages 55–74, 1999.
- Sid-Ahmed Berrani, Laurent Amsaleg, and Patrick Gros. Recherche par similarités dans les bases de données multidimensionnelles: panorama des techniques d'indexation. volume 7, pages 9–44, 2002.
- 11. Vincent Claveau, Romain Tavenard, and Laurent Amsaleg. Vectorisation des processus d'appariement document-requête. In *CORIA*, pages 313–324, 2010.
- Hanen Karamti. Vectorisation du modèle d'appariement pour la recherche d'images par le contenu. In CORIA, pages 335–340, 2013.
- Chris Harris and Mike Stephens. A combined corner and edge detector. In Proc. of Fourth Alvey Vision Conference, pages 147–151, 1988.
- F. Javier Díaz, Angel M. Burón, and José M. Solana. Haar wavelet based processor scheme for image coding with low circuit complexity. volume 33, pages 109–126, Tarrytown, NY, USA, March 2007. Pergamon Press, Inc.
- P. Montesinos, V. Gouet, and R. Deriche. Differential invariants for color images. In Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on, volume 1, pages 838–840 vol.1, Aug 1998.
- B. Akrout, I. Khanfir Kallel, C. Benamar, and B. Ben Amor. A new scheme of signature extarction for iris authentication. In Systems, Signals and Devices, 2009. SSD '09. 6th International Multi-Conference on, pages 1–8, March 2009.
- 17. Hervé Goëau, Alexis Joly, Pierre Bonnet, Jean-François Molino, Daniel Barthélémy, and Nozha Boujemaa. Lifeclef plant identification task 2014.