

A CBIR System for the Recognition of Agricultural Machinery

Rodolfo Rojas Ruiz¹, Lisbeth Rodríguez-Mazahua¹, Asdrúbal López-Chau²,
Silvestre Gustavo Peláez-Camarena¹, María Antonieta Abud-Figueroa¹,
Isaac Machorro-Cano¹

¹ División de Estudios de Posgrado e Investigación, Instituto Tecnológico de Orizaba,
Orizaba, Veracruz,
México

² Universidad Autónoma del Estado de México, Centro Universitario UAEM Zumpango,
Zumpango, Estado de México,
México

Rudy_199191@hotmail.com, lrodriguez@itorizaba.edu.mx, imachorro@gmail.com,
{gpelaez, mabud}@ito-depi.edu.mx,
alchau@uaemex.mx

Abstract. Content-Based Image Retrieval (CBIR) is a technique to obtain images based on visual characteristics like color, texture and shape. In this article, an analysis of different CBIR systems is presented; also an inventory system that allows content-based queries is proposed to help the employees of a company that sells agricultural machinery and spare parts, called AGROMAQ. This system identifies different products, all with the aim of reducing customer service time, since an expert is currently needed to identify them, which increases attention time not only in hours but also in days.

Keywords: CBIR system, BoofCV, SURF descriptor.

1 Introduction

Multimedia data provide a direct representation of the physical world in the digital format and include photographs, X-ray images, sound, and video [1]. A multimedia database is a collection of multimedia data that provides: 1) Content-based access; 2) Generation of knowledge; 3) Handling of large volumes of data, and 4) Good response times [2].

This article presents an inventory system for the company AGROMAQ. The system uses CBIR to perform the search for spare parts and machinery based on an image, everything with the aim of improving customer service times and reducing the dependence on an expert.

The descriptor SURF (Speeded Up Robust Features) was used as a method for the extraction of characteristics of the pieces and their subsequent automatic recognition. SURF [3] is one of the most widely used algorithms in the literature for extraction of

invariant points of interest. The extraction of the points is done by first detecting the possible points of interest and their location within the image. Subsequently, the neighborhood of the point of interest is represented as a feature vector (by default with a size of 64, although it is possible to increase it). The SURF algorithm uses a basic approximation of the Hessian matrix to reduce computation time. The Hessian matrix is used because of its good relationship between accuracy and time cost. The main feature of SURF points of interest is repeatability, if the point is considered reliable, the detector will find the same point under different perspectives (different scale, orientation, among others).

This article is made up of five sections. Section 2 in this paper is an analysis of different articles related to CBIR. Section 3 describes the different phases of the CBIR system. The evaluation of the system is discussed in Section 4. Finally, the conclusions and future work are presented in Section 5.

2 State of the Art

Navel et al. [4] proposed an enhanced SURF for CBIR systems, the enhanced was made by extracting Hu moments and eigen values from the image thus forming a 88 dimension enhanced SURF descriptor. The experiments with this enhanced version of SURF proved that this version is better than the original SURF. In contrast, Kommineni and Chava [5], developed a CBIR system that uses the SIFT (Scale-Invariable Feature Transform) descriptor in conjunction with SPM (Linear Spatial Pyramid Matching), which improves object recognition and results in Sparse + SIFT.

Table 1. Comparison of the related works.

Article	Evaluation	Descriptor name
[3]	Enhanced SURF descriptor is better than SURF in precision (increased in 9.2%) and recall (increased in 3%).	Enhanced SURF
[4]	Sparse+SIFT descriptor proved to be more accurate than the SIFT descriptor.	Sparse+SIFT
[5]	The CBIR system has a precision in the range between 69.51% and 91.19%.	SURF
[6]	The CBIR system has a precision of the 94.2% with the use of geometric constraints.	SIFT
[7]	SURF with standard deviation works better to find images in a database with sketches.	SURF
[8]	The CBIR system has a precision of the 73.17% using SIFT as descriptor.	SIFT
[9]	The precision of the system was 80% for MRI, 75% for Iris and 90% for Bones.	SURF
[10]	Dynamic delimitation is more effective and robust than static delimitation.	TOP-SURF JCD (Joint Composite Descriptor).
[11]	LSH (<i>Locality-sensitive hashing</i>) was the fastest technique.	LSH
[12]	The proposed method has an average precision of the 78.97%	SURF

Ronald et al. [6] constructed a CBIR system, Plantcyclopedia, that can identify plants given by the user using SURF, is divided into three main steps: 1) Retrieve the image information; 2) Define the orientation of the image, and 3) Obtain feature points from both images using K-Means as a classification technique. Lee et al. [7] presented a CBIR system for a database of tattoo images based on SIFT. The goal of the system is to obtain visually similar images to the query image as well as determining some extra feature to the shape and design of the tattoo, such as the area of the body where it is located and class tags of tattoos.

Pradnya and Pravin [8] developed a SBIR (Sketch Based Image Retrieval) system, this system uses SURF with a standard deviation to recover the images from the database. Whereas SIFT was used in a system for detection and pairing of drug pills presented in [9], the system used as the basis the detection of the pills depending on their shape, stamp / figure, color, color histogram and Hu moments. Kaur and Jindal [10] proposed a feature extraction on medical images using SURF under the Open CV platform. The steps of the system are five: 1) Load test images; 2) Extract key points, contours and textured data; 3) Store the features in a vector; 4) Display images highlighting key points, and 5) Compute performance.

Avi et al. [11] presented a two-step method for a CBIR system in large multimedia databases, in order to improve both the effectiveness and efficiency of the traditional CBIR by exploring secondary media, such as pre-filtered text or images. The two-step recovery method at the beginning classifies by a secondary means and then performs CBIR only on the best ranked K-objects. The proposal of an architecture for efficient image recovery, as well as Web analysis from the point of view of images, using Big Data technologies, is the main theme on [12]. The phases of the proposed architecture are: 1) Signature of the input image; 2) Search with binary trees; 3) Distributed calculation of distances, and 4) Ordination of similar images.

Alexandra et al. [13], proposed a method for CBIR and classification, the method used SURF combined with Bag-of-Visual-Words (BoVW). This combination has a good retrieval and classification results compared to other methods. SURF has four major steps: 1) Integral image, 2) Fast Hessian, 3) SURF descriptor, and 4) Salient features. BoVW consists of three main steps: 1) Automatically extract the interest points and descriptor from the images, 2) Quantize the keypoints and descriptors to form the visual dictionary, and 3) Find the occurrences of each visual words in the image in order to build the BoW histogram.

In conclusion, CBIR systems are important because they facilitate the task of identifying objects to people who are not experts in the field, and are not limited to a single area such as medicine, botany, forensic medicine, among others. Finally, it is possible to say that SIFT and SURF have good results (precision) in the CBIR systems as shown in Table 1. Nevertheless, Section 4 will show that the time and recall are better in SURF than SIFT, thus, SURF will be used in the CBIR system of the company AGROMAQ.

3 CBIR System

Figure 1, shows the process followed by the system for the extraction and storage of the descriptors each time an image is recorded and thus saving time in the search.

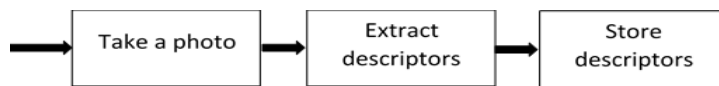


Fig. 1. Process during the recording of images.

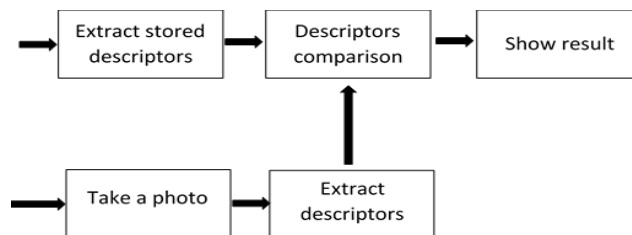


Fig. 2. CBIR process.



Fig. 3. Interface of the content based query.

During the registration of a product the user takes one or several photographs of the product to be registered and uploads them to the system. Once the product is registered, the system extracts the points of interest of each of the images and the descriptors using SURF, finally, these descriptors are stored in the form of objects for their future use.

In Figure 2, on the other hand, it is represented the process that the system follows for the image-query and the images in the database during the search process. The user chooses to write a search field to help reduce the sample of products to be compared during the process. If this is done, the system will search for images with those fields in their description and will return those that fit. After doing this, it extracts the descriptors stored in objects of each of these images, in case no search field has been written, the system will take the complete sample of all the registered images.

Likewise, the user must enter a photograph of the product to be searched to which the points of interest will be extracted and the descriptors will be calculated again with SURF. Once the descriptors of both the images in the database and the query image are obtained, the system proceeds to compare the query image with each of the images in the database to obtain the points where the pairing is correct, once this process is finished, the images are sorted with respect to the number of correct pairings and the results are displayed.

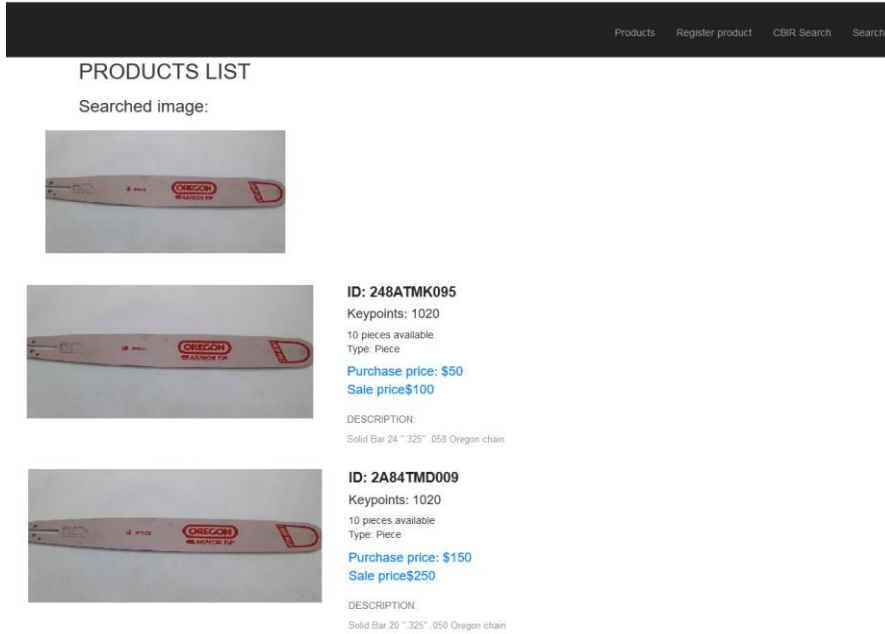


Fig. 4. Interface of the result of the content based query.

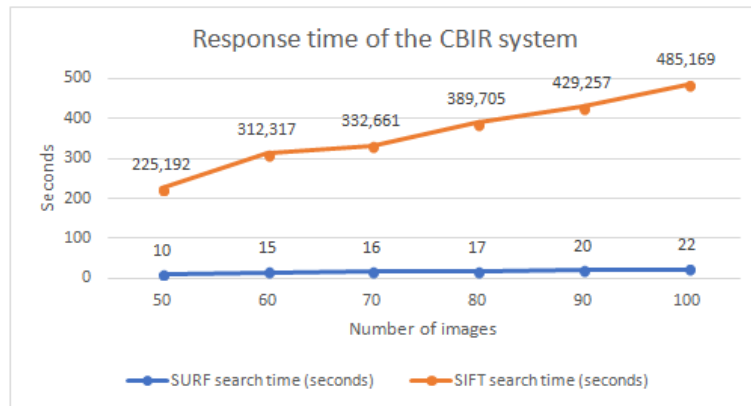


Fig. 5. Comparative diagram of efficiency using SIFT and SURF.

The CBIR system was implemented using Java with the JavaServerFaces framework. The multimedia database was implemented using PostgreSQL. The search interface in Figure 3 shows two fields to proceed, which are the number of images most similar to the given image, and the search image (query). Once the search is done, the system displays the results in list form, detailing the product or products that are similar to the given image as shown in Figure 4.

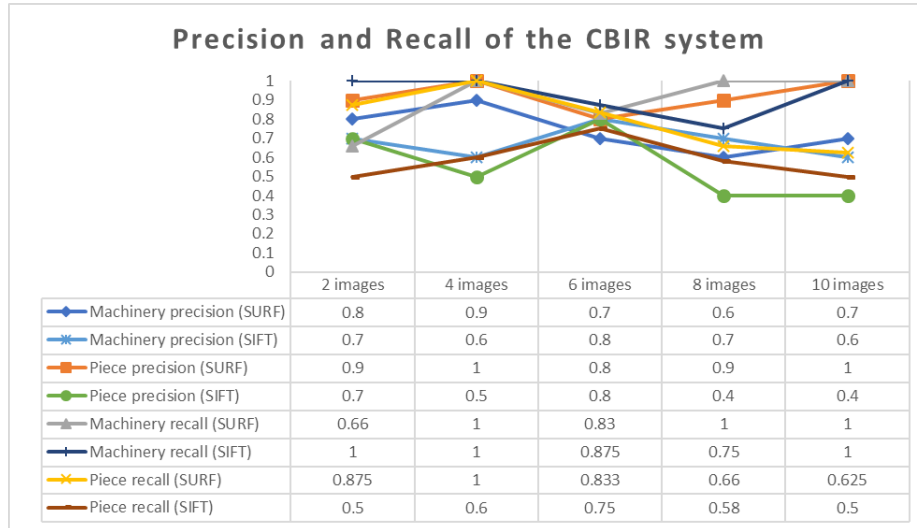


Fig. 6. Comparative diagram of effectiveness between SIFT and SURF.

4 Experiments

There are two main parameters used to measure the performance of a CBIR system: efficiency and effectiveness. Efficiency is mainly concerned with the response time of the system, while effectiveness is related to the system's ability to retrieve relevant items and discard irrelevant items [13]. In order to evaluate the performance of the CBIR system presented in this paper, the follow experiments were made with the descriptor SURF from BoofCV v0.26 and the descriptor SIFT from OpenCV 2.4.13.

4.1 Efficiency of the CBIR System

A performance evaluation was realized on the response time of the system, the content-based query of the figures 3 and 4 was executed in different cases which varied from each other by the number of images in the database, ranging from 50 images to 100 images increased 10 images in each case both for the SIFT and SURF descriptors. In all cases the number of images retrieved was 5.

Figure 5 shows clearly that the CBIR system applying the SURF descriptor has a better efficiency than when it implements SIFT.

4.2 Effectiveness of the CBIR System

Based on the state of the art presented in Section 2, recall and precision are the most used measures to evaluate the effectiveness of CBIR systems. Recall measures the ability to retrieve relevant items, and precision measures the ability to reject irrelevant items. A good system should have both high recall and precision values [3, 14]. Therefore, during the tests performed in the system where the evaluation was realized

to two different groups of images (Pieces and Machinery) in 10 iterations, applied to the 100 images in the database, the results that were obtained by equations (1) and (2) are shown in Figure 6. Five cases were considered; 2, 4, 6, 8 and 10 queries executed and in all cases the total number of images retrieved was 5.

$$\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in database}}, \quad (1)$$

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}. \quad (2)$$

Figure 6, shows a comparison between SIFT and SURF and as it can be seen the effectiveness of the CBIR system using SURF is better most of the times taking into account the precision and the recall.

5 Conclusion and Future Work

CBIR systems are important at the commercial level because they make easier to identify objects, spare parts and machinery in this case, to people who are not experts in the field, as well as the fact that they no longer depend so much on experts in the area, with the aim to make customer service times shorter.

In this article, a CBIR system for the recognition of agricultural machinery and spare parts was showed, also the efficiency and effectiveness of the system was evaluated through the results of the experiments realized with SIFT and SURF descriptors as well as a brief comparison between them, with those results it was concluded that in the case of this article SURF is faster than SIFT and has better performance in precision and recall.

The performance of the CBIR process will be tried to improve by implementing a system that allows an environment in which it is possible to have a control of the brightness during the shooting of the products.

Acknowledgments. The authors are very grateful to the National Council of Science and Technology (CONACYT) for supporting this work. Also, this research paper was sponsored by National Technological of Mexico, as well as by the Public Education Secretary (SEP) through PRODEP.

References

1. Phillip, K.C.: *Multimedia Information Storage and Retrieval. Techniques and Technologies*, IGI Publishing, Hershey (2008)
2. Assent, I.: *Efficient adaptive retrieval and mining in large multimedia databases* in Dissertation. Fakultät für Mathematik, Informatik und Naturwissenschaften, RWTH Aachen (2008)
3. Romero, M. et al.: *Comparativa de detectores de características visuales y su aplicación al SLAM* in *Comparativa de detectores de características visuales y su aplicación al SLAM*. Cáceres, pp. 55–62 (2009)

4. Navel, G. et al.: Content Based Image Retrieval using Enhanced SURF in Computer Vision, Pattern Recognition, Image Processing and Graphics (2015)
5. Kommineni, D., Chava, S.: CBIR Based On Linear SPM. Using SIFT Sparse Codes in IEEE International Conference on Computational Intelligence and Computing Research (2016)
6. Ronald, A. et al.: Plant Encyclopedia with Image Matching on a Web-Based Application. In International Conference on Information Management and Technology, Indonesia, pp. 89–92 (2016)
7. Lee, J. et al.: Image Retrieval in Forensics: Tattoo Image Database Application in Multimedia in Forensics. Security and Intelligence, 19(1), pp. 2–11 (2012)
8. Content Based Image Retrieval by Using Sketches: Information Communication and Embedded Systems (2014)
9. Lee, Y. et al.: Pill-ID: Matching and retrieval of drug pill images. In Pattern Recognition Letters, 33, pp. 904–910 (2012)
10. Kaur, B., Jindal, S.: An implementation of Feature Extraction over medical images on OPEN CV Environment. In Devices, Circuits and Communications (2014)
11. Avi, A. et al.: Dynamic two-stage image retrieval from large multimedia databases. In Information Processing and Management, 49, pp. 274–285 (2013)
12. Sergui, R. et al.: Tecnologías big data para análisis y recuperación de imágenes web en el profesional de la información, 23(6), pp. 567–574 (2014)
13. Alexandra, A. et al.: Content Based Image Retrieval and Classification Using Speeded-Up Robust Features (SURF) and Grouped Bag-of-Visual-Words (GBoVW). In International Conference on Technology, Informatics, Management, Engineering & Environment, Indonesia, pp. 77–82 (2013)
14. Lu, G.: Multimedia Database Management Systems. Artech House Computing Library (1999)