Annals of the University of Craiova, Mathematics and Computer Science Series
Volume 42(1), 2015, Pages 202–210
ISSN: 1223-6934

Image retrieval using global descriptors and multiple clustering in Nash game

O. Bencharef, B. Jarmouni, N. Moussaid, and A. Souissi

Abstract. In order to improve the retrieval accuracy of content-based image retrieval systems, we consider the Content-based image retrieval formulated as a Nash game. In the present approach by splitting the optimization variables into three players. The first player minimizes its objective function by using the first strategy (Color descriptor), the second player acts by using the second strategy (Gist descriptor), and the third one by using the third strategy (SIFT descriptor). We will use the Nash equilibrium to detect the classes of membership of the searched image.

Key words and phrases. Nash Equilibrium, concurrent optimization, fuzzy K-means, color, Gist and SIFT descriptors.

1. Introduction

Generally, an image is composed of three matrixes (RGB) with sizes exceeding 200 * 200 px, which provides vast amounts of information to analyze to predict the degree of similarity between images. Until few years ago, Text Based Image Retrieval (TBIR) systems were the best choice of large-scale content-based image retrieval. Based on text annotations (image names, tags) to collect images relevant to a query word, from very large image collections, they shows good results [13]. The problem with this kind of systems is related to the textual description accompanying the images that remains unable to really describe the visual content of the images, and since the automatic generation of text descriptors for a set of Image is not feasible, there are several methods that come to complete the gaps in TBIR systems.

Currently the Content Based Image Retrieval (CBIR) systems based on the visual content of the image, is becoming more popular despite the disadvantage of the data size. They are used to re-rank TBIR results [5] or to retrieve small data set (less than 100000 pictures) [12].

In this context we propose a Content Based Image Retrieval system based on the Fuzzy K-means clustering and a classification system based on game theory[10].

On the extraction phase, each image is symbolized by three vectors that represent the visual characteristics of low level:

- The color vector using Compressed Color Histogram CCH.
- Gist descriptor.
- SIFT descriptor.

After having calculated the three vectors of representations for each image of database (DB), the finale data will contain three Matrices, matrix one for color feature, matrix two for Sift descriptors and matrix three for Gist descriptors.

This paper has been presented at Congrès MOCASIM, Marrakech, 19-22 November 2014.
This work has been partially supported by LIRIMA..

Browse an entire database to find a similar image does not seem a good idea, for Big database the mission is more delicate. For this we applied a clustering algorithm on each level of representation, which automatically divides the images into several classes in each level of representation. In this way each level of representation will be spread on a number of predefined classes.

Our purpose is to introduce an original method to solve the image retrieval problem, based on a game theory approach. We then show that the control formulation naturally leads to a Nash game of static nature with complete information, which involves a the color discreptors, Gist and SIFT. Hence, we obtain a multidisciplinary optimization problem. To determine the optimal image as Nash equilibrium. The existence of the Nash equilibrium is proved, and when the similar image solution exists, it turns out that the Nash equilibrium is exactly the triplet of information have been the primitive image descriptors in content-based image retrieval systems.

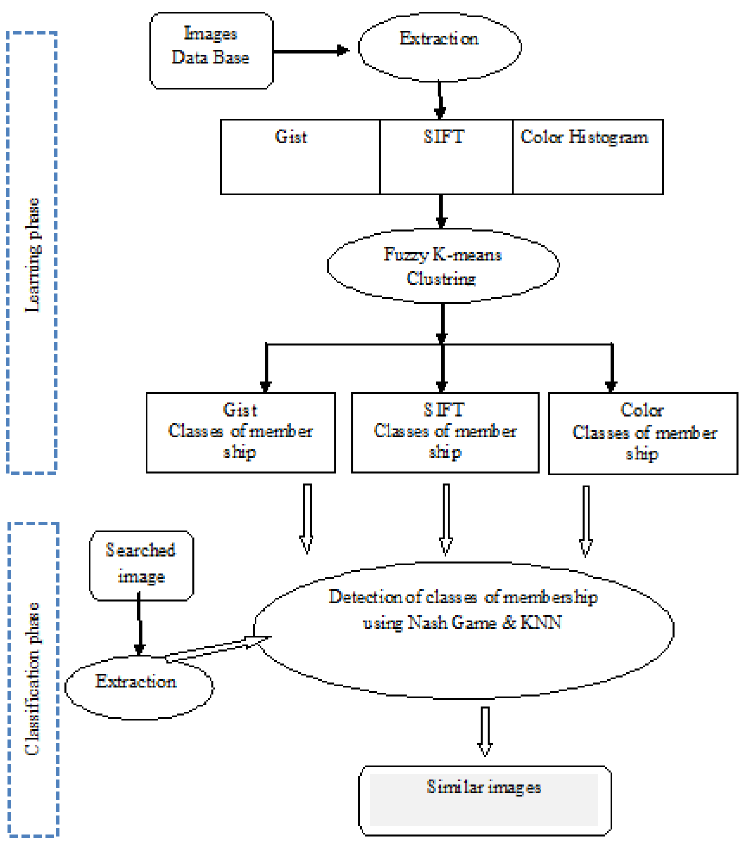

FIGURE 1. Proposed system.

2. Extraction and clustering the database

First, we will calculate the shape descriptors, SIFT, Gist and color histogram of each image in the database. We will have three matrices: a matrix that contain SIFT features for each image in the database, a matrix for Gist features and a matrix for color features. Secondly, we will apply a clustering algorithm separately on every matrix. Therefore, each image of the database will belong to three classes (SIFT class, Gist class and Color class).

2.1. Extraction. Extraction is the process of reduction of the size of information that represents the image. In this work it is based on three shape descriptors [4]: SIFT descriptors, Gist descriptors and Compressed Color Histogram CCH.

2.1.1. *SIFT descriptors.* SIFT (Scale Invariant Feature Transform) is a detector of points of interest (invariant to scale changes). To describes the intensity of the pixels in a neighborhood around each interest point SIFT computes the magnitude and direction of the gradient to a selection of points of the regions. These amplitudes are then weighted by a Gaussian which allows taking into account the distance from the center of the region. These gradients are then accumulated in orientation histograms for each sub-region. In practice, first the image is transformed into an integral picture to accelerate calculation. Secondly, the image with high change of intensity are searched. Finally, each of the extracted points in the previous step is described by a vector composed of $3 * 3 * 4$ values that is 36 dimensions.

2.1.2. *Color histogram.* Color descriptors are the most effective descriptor [11]. There exist several descriptors of colors, in our case we work with the reduced color histogram, based on the regular color histogram [4] (see Figure 2).

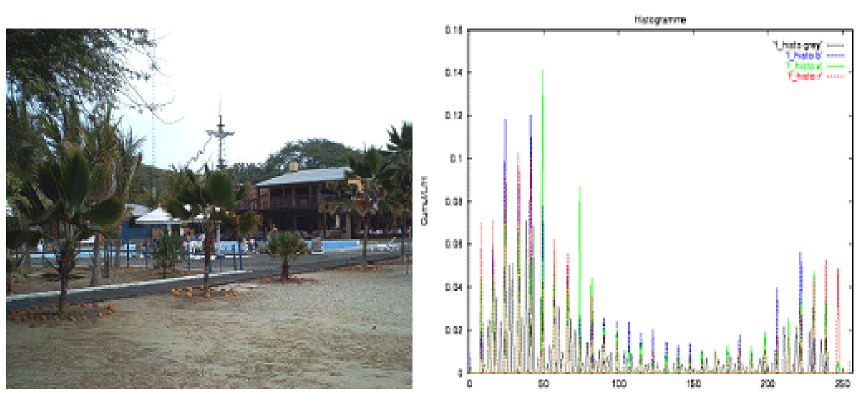

FIGURE 2. Example of color histogram.

2.1.3. *Gist descriptors.* Gist descriptors is a kind of low dimensional representation of the scene they are based on the amplitudes obtained in the output of K Gabor filters at different scales and orientations. To reduce the size, each image in filter output is resized to a size $N \times N$ (N between 2 and 16) [6].

3. Preparation of the training database

Our data base is composed of 20000 image; for each image we are going to calculate three vectors: a color vector VC(1; 96), Gist vector VG(1; 36) and SIFT vector VF(1; 32).
Each image will be represented by three vectors the three vectors VC,VG ans VF(1; 32).
Finely the training data base is composed as follow:

- •: All VC are collected in a matrix denoted MC(20000; 96).
- •: All VF are collected in a matrix denoted MF(20000; 32).
- •: All VG are collected in a matrix denoted MG(20000; 36).

3.1. Clustering of the database. The Fuzzy C Means (FCM) is a fuzzy version of k-Means proposed by Besdek [7], the algorithm allows the separation of data, in to C classes. Each class is represented by its center.
We consider a proximity matrix that describes the degree of membership of each vector to the different C classes, by calculating the distance between each vector and the center of each class using:

$$u_{ik} = \left| \sum_{j=1}^{C} |\frac{d_{ik}}{d_{jk}}|^{\frac{2}{m-1}} \right|^{-1}$$

With:
d_{ik} represents the distance between the vector V_k and the center of class C_i.
u_{ik} represents the degree of membership of the vector V_k and the center of class C_i.
$k = 1..n$ (number of vectors).
$i = 1..C$ (number of classes).
m is any real number greater than 1 (fuzziness coefficient)
Computing the class center is calculated in the same manner as the K-means [14].

4. Proposed methods

Content-Based Image Retrieval (CBIR) allows to automatically extract target images according to objective visual contents of the image itself. In this paper we will discuss the color, SIFT and Gist named from now on as strategies within the scope of CBIR. The three strategies are the primitive image descriptors in content-based image retrieval systems. This paper presents a novel approach for classifying all the three strategies to achieve higher retrieval efficiency. Each image $F = (F_C, F_G, F_S)$ of the database will be extracted according to three classes (SIFT class, Gist class and Color class) [1].
Where F_C is the vector color (Color descriptors);
Where F_G is the vector Gist (Gist descriptors);
Where F_S is the vector SIFT (SIFT descriptors).

We consider a three-players static game of complete information. The first player is the Color descriptor that is used to control color classes in an image, denoted by C. The second player is the Gist classes, denoted by G. Then, the third is the SIFT descriptors which controls the SIFT classes, denoted by S.

4.1. Existence of a Nash equilibrium. The algorithm that we have integrated is based on Tikhonov regularization:

Let us denote by $I(C, G, S)$ an similar images, non-corrupted image, which is a function defined on some domain $\Omega =]0, l[\times]0, L[$ open and bounded, and let F be image of database defined by three vectors (Colors(F_C), GIST(F_G), SIFT(F_S)).

$$F = F(F_C, F_G, F_S) \tag{1}$$

Retrieve the similar images I from F, by simply minimizing the quadratic misfit.
We consider the functionals $J_C(I(C, G, S))$, $J_G(I(C, G, S))$ and $J_S(I(C, G, S))$ defined by:

$$J_C(I(C, G, S)) = \frac{1}{2}||C - F_C||^2 + \frac{\epsilon}{2}(||\nabla C||^2 + ||\nabla G||^2 + ||\nabla S||^2), \quad C \in H_0^1(\Omega) \tag{2}$$

$$J_G(I(C, G, S)) = \frac{1}{2}||G - F_G||^2 + \frac{\epsilon}{2}(||\nabla C||^2 + ||\nabla G||^2 + ||\nabla S||^2), \quad G \in H_0^1(\Omega) \tag{3}$$

$$J_S(I(C, G, S)) = \frac{1}{2}||S - F_S||^2 + \frac{\epsilon}{2}(||\nabla C||^2 + ||\nabla G||^2 + ||\nabla S||^2), \quad S \in H_0^1(\Omega) \tag{4}$$

where ϵ is some parameter to be adjusted [2].
We address the problem of the splitting of the optimization variable between three players. The first player minimizes his objective function $J_C(C, G, S)$ with using the first strategy color, the second player uses the second one Gist $J_G(C, G, S)$, and the third uses global shape $J_S(C, G, S)$.
We consider the following optimization problem.

$$\begin{cases} \text{Find } I(C^\star, G^\star, S^\star) \in H_0^1(\Omega) \text{ such that:} \\ \qquad \min\limits_C J_C(C, G^\star, S^\star) = J_C(C^\star, G^\star, S^\star) \\ \qquad \min\limits_G J_G(C^\star, G, S^\star) = J_G(C^\star, G^\star, S^\star) \\ \qquad \min\limits_S J_S(C^\star, G^\star, S) = J_S(C^\star, G^\star, S^\star) \end{cases} \tag{5}$$

Theorem 4.1. *There exists a Nash equilibrium $(C^\star, G^\star, S^\star)$ solution of the problem (5).*

Proof. Since the functionals J_C, J_G and J_S are the respective compliances of color, GIST and SIFT equations, they could always be seen as supremum envelopes of convex functions with respect to, respectively, C, G and S. Hence, these three objective functionals are also convex. On the other hand, weak star compactness of the strategy spaces, which are convex, as well as weak star lower semicontinuity and weak compactness of the objectives with respect to their strategies are known to hold see e.g. Reference [9]. Notice that the weak star lower semicontinuity of J_C with respect to C is due to the use of compact filters, which preserve the convexity thanks to the linearity of the filters.
The assumptions are fulfilled in order to apply the Nash existence theorem,which yields the existence of a Nash equilibrium see [3]. □

4.2. Algorithm of Nash Equilibrium. Finding the Nash equilibrium requires solving the last problem (5). The Nash equilibrium is computed by the following decomposition algorithm.

(1) Initial guess $I^{(0)} = (C^{(0)}, G^{(0)}, S^{(0)})$, Set $n = 0$.
(2) repeat
(3) $\overline{C}^{(n)} = \arg\min\limits_C J_C(C, G^{(n)}, S^{(n)})$
(4) $\overline{G}^{(n)} = \arg\min\limits_G J_G(C^{(n)}, G, S^{(n)})$
(5) $\overline{S}^{(n)} = \arg\min\limits_S J_S(C^{(n)}, G^{(n)}, S)$

(6) $I^{(n+1)} = (C^{(n+1)}, G^{(n+1)}, S^{(n+1)}) = \lambda(C^{(n)}, G^{(n)}, S^{(n)}) + (1-\lambda)(\overline{C}^{(n)}, \overline{G}^{(n)}, \overline{S}^{(n)})$, $0 < \lambda < 1$

(7) until $I^{(n)}$ converges.

From the game theory viewpoint, this algorithm represents an iterative, partially cooperative game (with partial exchange of information on the partially optimal strategies between the players at the end of each iteration

4.3. Phase of recognition. The researched images must go through the same pretreatment as the training data base images. Firstly, we compute descriptors of the three levels of representations. Secondly, we compute the solution of the Nash equilibrium. Finally, the solution is used to define the classes of membership by using the KNN algorithm. To estimate the output associated with a new input x, we consider the intersection of the three classes of membership.

5. Experimental Results

The image database in our current system contains 20000 images from iaprtc12 data base [8]. There are several semantic categories: images in open sky, in the ocean, or on the city, images with green trees or plants, images with mountain and building under different time (daytime, sunset and nighttime), a big number of image present persons (young, old ...) and objects from our day life (Figure 3). Our work is implemented using Matlab 9.

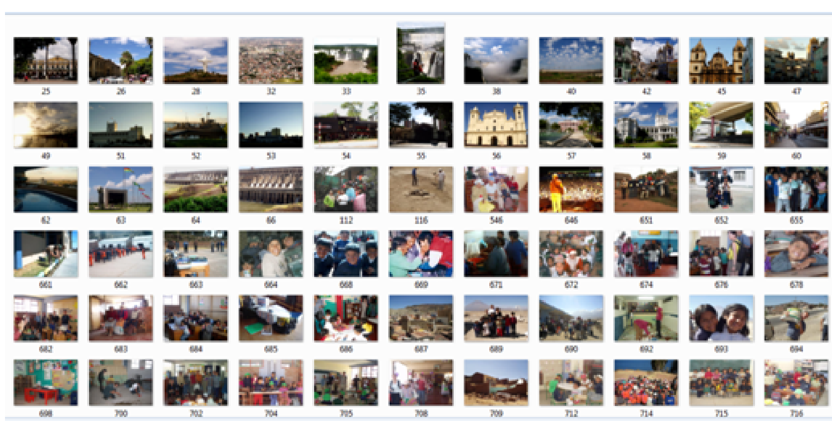

FIGURE 3. Example of images of the database ipartc 12.

Our main objective was to develop a system for big data base so the execution time was our main priority. Unlike classic system based on the calculation of the distance of the requested Image and all database Image, our system reduce the number of Image to be checked in to small number(less than 30). To minimize the number of images to check we work with a number of clusters k = 30 for Gist, k=20 for SIFT, and k = 35 for compressed color histogram. To verify the effectiveness of our approach, we have performed a comparative study between KNN with the presented descriptors and our algorithm. Figure 4, shows the execution time needed for different size of data base.

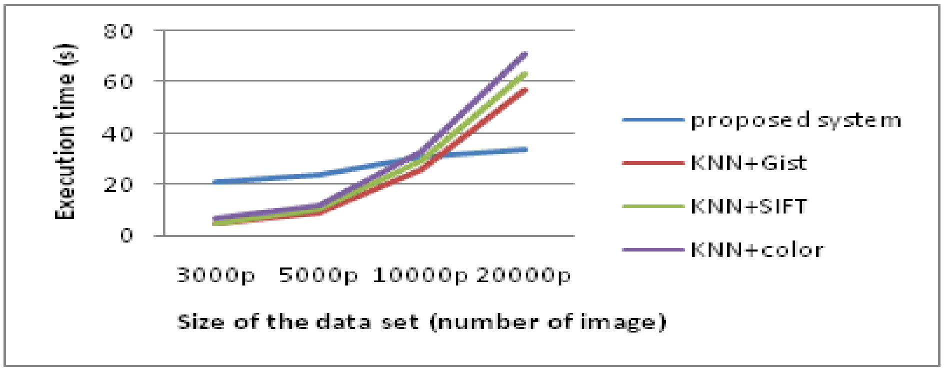

Figure 4. Execution time for different size of the data base; for the proposed system, KNN+Gist, KNN+SIFT and KNN+Color.

The extraction of three components and the integration of Nash game need a considerable time. For a small data base it's clear that a simple classification method (KNN+descriptor) need less time, for big data base (> 12000) the proposed system present better performance. To evaluate classification performance we perform as follow:

- For KNN algorithm we choose a value of k=10.
- The proposed system return a number of Image between 6 an 15.

Then we evaluate manually the outputs of every system, if the number of Image considered similar is more than five the result is considered correct. Table 1 present results for 100 aleatory image:

Table 1. Comparative study between the proposed system and classical method for different data set sizes.

	3000p	5000p	10000p	20000p
proposed system	98%	97%	94%	92%
KNN+Gist	72%	66%	33%	21%
KNN+SIFT	69%	52%	29%	17%
KNN+color	97%	67%	33%	28%

The proposed system allowed us to have results that share the maximum amount of information with the requested Image, at least a double similarity (Gist/SIFT, Gist/Color or SIFT/Color) and sometimes triple similarity(Gist, SIFT and Color). Even for big data set the proposed system presents a recognition rate of 92%. The following figure present examples of request results.

For the five requests there is always a semantic sense; in request 1 (Figure 5) the majority of results share the same color or texture, request 2 the white background, request 3 the river background and boat, request 4 the ocean and the sky are all blue colors.

6. Conclusion

In this paper, we propose an efficient content-based image retrieval system based on the global description of the Image. There are three differences between our system

FIGURE 5. Request number 1.

FIGURE 6. Request number 2.

FIGURE 7. Request number 3.

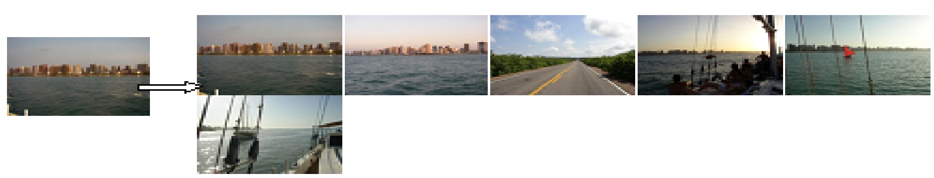

FIGURE 8. Request number 4.

and other systems. First a triple description of the image, second multi fuzzy clustering and finally the integration of Nash game. The proposed system shows effective results with a smaller execution time for big database. In the future, we plan to add more images to our image database and to explore more effective descriptors.

Acknowledgements. The authors would like to thank the editor and the anonymous referees.

This work has been partially supported by LIRIMA.

References

[1] R. Aboulaich, A. Habbal, and N. Moussaid, Split of an optimization variable in game theory, *Math. Model. Nat. Phenom(MMNP)* **5** (2010), no. 7, 106–111.

[2] R. Aboulaich, A. Habbal, and N. Moussaid. Optimisation multicritère : une approche par partage des variables, *ARIMA* **13** (2010), 77–89.

[3] J.P. Aubin, *Mathematical methods of game and economic theory*, North-Holland Publishing Co., Amsterdam, New York, 1979.

[4] O. Bencharef, B. Jarmouni, and A. Souissi, Research of Similar Images Based on Global Descriptors and Multiple Clustering, *International Journal of Engineering and Technology* **5**(2013), no. 3, 3142–3151.

[5] V. Jain and M. Varma, Learning to re-rank: query-dependent image re-ranking using click data, In: *Proceedings of the 20th international conference on World wide web Pages*, ACM, New York (2011), NY, USA, 277–286.

[6] H. Jégou, M. Douze, and C. Schmid, Hamming embedding and weak geometric consistency for large scale image search, In *Computer Vision ECCV 2008*, Lecture Notes in Computer Science **5302** (2008), 304–317,

[7] S.L. Chiu-Fuzzy, Model Identification Based on Cluster Estimation, *Journal of Intelligent and Fuzzy System* **2** (1994), 267–278.

[8] M. Grubinger, *Analysis and Evaluation of Visual Information Systems Performance*, Ph.D. thesis, Victoria University, Melbourne, Australia, 2007.

[9] A. Habbal, J. Petersson, and M. Thellner, Multidisciplinary topology optimization solved as a nash game, *Int. J. Numer. Methods in Eng.* **61** (2004), 949–963.

[10] J.F. Nash, Non-cooperative Games, *Annals of Mathematics* **54** (1951), no. 2, 286–295.

[11] M. Swain and D. Ballard, Color indexing, *International Journal of Computer Vision* **7** (1991), 11–32.

[12] F. Thollard and G. Qunot, Content-Based Re-ranking of Text-Based Image Search Results, *Lecture Notes in Computer Science* **7814** (2013), 618–629.

[13] P.A. Vikhar, Content Based Image Retrieval (CBIR): State-of-the-Art and Future Scope for Research, *The IUP Journal of Information Technology* **6**(2010), no. 2, 64–84.

[14] J.B. MacQueen, Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability 1*, University of California Press (1967), 281–297.

(O. Bencharef) SAED^2, School of Technology-Essaouira, Cadi Ayyad University, Marrakech, Morocco

(B. Jarmouni, A. Souissi) Numerical Analysis Group, Applied Mathematics Laboratory, Department. of Mathematics, Faculty of Sciences, Mohammed V University, 4,Ibn Battouta Avenue, P. O. Box 1014, PC 10090 Rabat, Morocco

(N. Moussaid) SAED^2, School of Technology-Essaouira, Cadi Ayyad University, Marrakech, Morocco; LIRIMA, International Laboratory for Research in Computer Science and Applied Mathematics

E-mail address: `no.moussaid@uca.ma`