Applied Mathematical Sciences, Vol. 8, 2014, no. 163, 8099 - 8111
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2014.410790

A New Approach for Similar Images Using Game Theory

O. Bencharef [a], B. Jarmouni [b], N. Moussaid [a] and A. Souissi [b]

[a] SAED[2], High School of Technology-Essaouira
Cadi Ayyad University, Marrakech, Morocco

[b] 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

Abstract

A Content Based Image Retrieval (CBIR) system is the result of the combination of computer vision techniques such as segmentation, extraction and classification in order to detect similar images. In our work we propose a CBIR system based on global descriptors, multiple clustering and on multi-criteria optimization problems using game theory. We propose to determine the similar images as a Nash Equilibrium. We tackle the problem by splitting the optimization variables into three players. The first player minimizes its objective function using the first strategy (Color description), the second player by using the second strategy (Gist descriptor) and the third one by using the third strategy (Zernike descriptor). After using the K-means algorithm to divide each level of description on K clusters, we use the Nash equilibrium to detect the classes of membership of the searched image.

Keywords: Multi-criteria Optimization, Game theory,Concurrent optimization, K-means, image retrieval, Color, Gist and Zernike descriptors

1 Introduction

The problem of automatic indexing pictures by their content presents a challenge for researchers in computer vision. The fact that every image is represented by one or more vectors, the research stage becomes very hard. To raise these difficulties, many techniques were proposed. Among these approaches we find the techniques of research for images based on text description known under the name Text-based Image Retrieval (TBIR) [18, 19, 17, 7, 10] which is the oldest approach used until our days. The problem with the techniques is that each image should be annotated by a set of keywords describing their contents. Since the automatic generation of text descriptors for a set of images is not feasible, researchers were oriented to image annotation [18, 3] and content description [20]. And finally CBIR system [6, 5].
In this paper, we propose a search engine for similar images, based on image multiple representations and clustering, reinforced by game theory and K-nearest neighbors (KNN) algorithm. On the representation phase, each image is symbolized by three vectors that represent the visual characteristics of low level (Figue1):

- The color vector using compressed color histogram.
- Gist descriptors.
- Zernike moments.

On the learning phase, the fact that each image is represented by a dependent vector and not by isolated values, without forgetting the enormous number of image existing on the web, makes the research of the similar images with the classical research methods a painful mission. To do this, after having calculated the three vectors of representations for each image of database (DB), we separately applied to each level of representation the clustering algorithm, which automatically divides the images into several classes in each level of representation.

On the classification phase, we solved a multidisciplinary optimization problem using a non-cooperative game (Nash Game) where the strategy of the players is naturally defined by three strategy(color, Zernike, Gist), the optimal solution is used to detect the classes of membership for each level of representation. Finely the results are obtained using KNN on the intersection of membership classes.

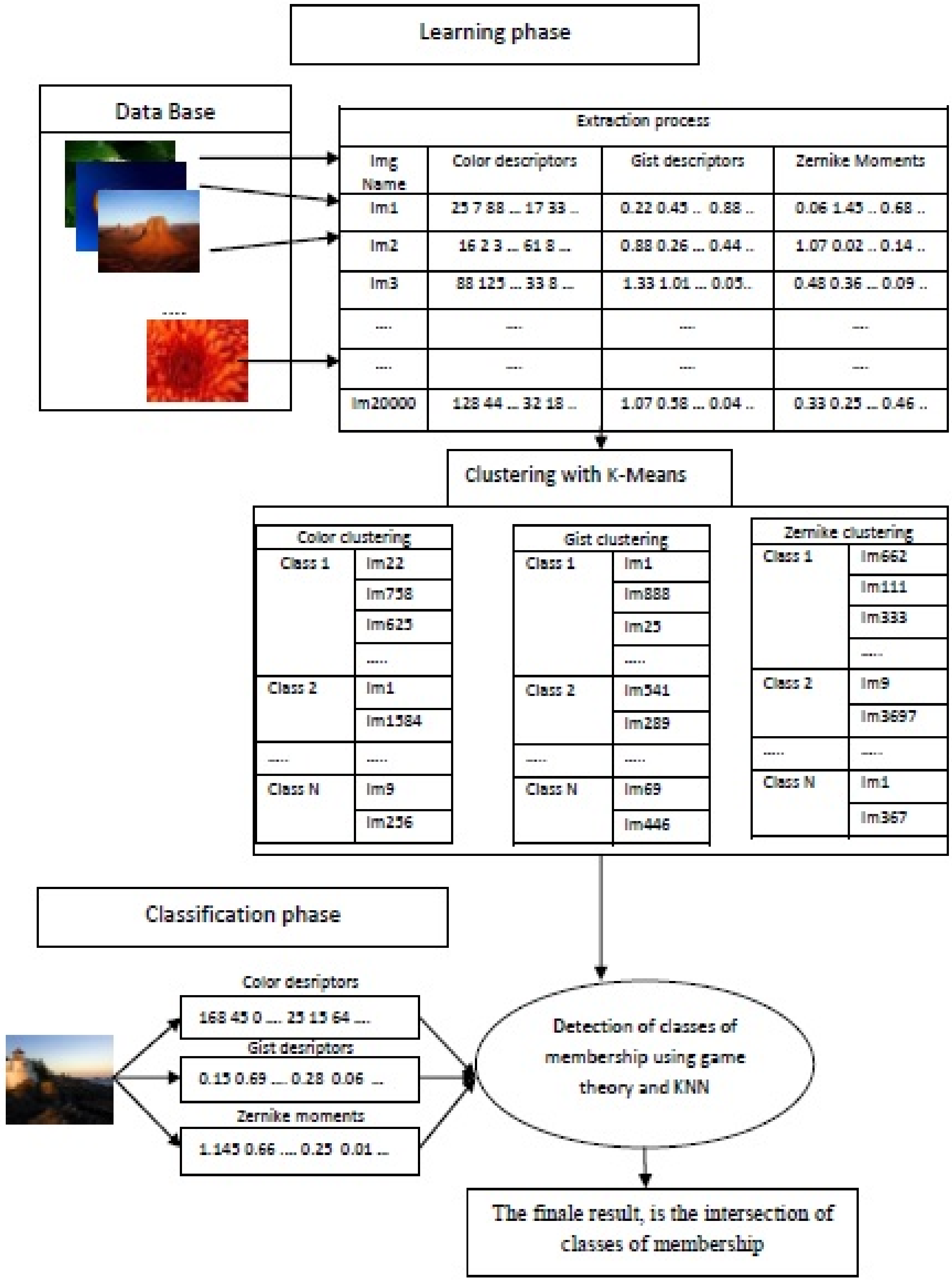

Figure 1: Proposed system.

2 Pretreatment

There is an option in our system that can handle the noisy images. The algorithm that we have integrated is based on Tikhonov regularization:
Let us denote by I an original image, non-corrupted image, which is a function defined on some domain Ω open and bounded, and let F be noisy image defined by:

$$F = I + \mu \tag{2.1}$$

where μ is a white additive Gaussian noise. Reconstructing the original image I from F, by simply minimizing the quadratic mist [9]

$$J(I) = \min_{I} \frac{1}{2}||I - F||^2 + \frac{\epsilon}{2}||\nabla I||^2 \qquad \forall I \in H_0^1(\Omega), \qquad F \in L^2(\Omega) \tag{2.2}$$

where ϵ is some parameter to be adjusted, see Figure 2

Figure 2: left: original image, middle: noisy image, right: restored image using the Tikhonov regularization.

We assume that the image depends on design strategies of the color (C), Gist (G) and shape (S). Hence, we obtain a multidisciplinary optimization problem. To determine the optimal image as Nash equilibrium, we will use in both cases three criteria J_c, J_g and J_s associated with the three players respectively. We use, in this case, a concurrent optimization realized by an algorithm simulating a Nash game [15] between the three players. The three players act following different objectives; in particular, player 1 has to choose his strategies in order to minimize his function J_c, while player 2 has to minimize the function J_g and player 3 has to minimize the function J_s (Paragraph 3).

3 Preparation of the database

First, we will calculate the shape descriptors, Zernike, Gist and color histogram of each image in the database. Secondly, we will apply a clustering algorithm on each level of description. Therefore, each image of the database will belong to three classes (color class, gist class and global shape class).

3.1 Extraction

Extraction is the process of reduction of the size of information that represents the image. This work is based on three following descriptors [6]:

- The color histograme
- Zernike moments
- Gist descriptors

3.1.1 Color histogram

Many studies have shown that color is the most effective descriptor [21]. There exists several descriptors of colors, in our case we work with the reduced color histogram, based on the regular color histogram [6].

3.1.2 Gist descriptors

Gist descriptors are based on the analyze of the spatial frequencies and orientations. The global descriptor is built by combining the amplitudes obtained in the output of K Gabor filters at different scales and orientations. For reducing the size, each image in filter output is resized to a size $N \times N$ [16].

3.1.3 Zernike moments

Zernike moments provide very good properties of theoretical invariance by translation, scaling and rotation. They constitute a vector space in which the image of the form is projected [13, 12]. With a compact and easy coding to compute.

3.2 Preparation of the training database

Each image will be represented by three vectors: a color vector $VC(1, 96)$, a Gist vector $VG(1, 36)$ and a shape vector $VF(1, 24)$:

- VC, the color histogram of each color, is divided on 32 parts of 8 values. We calculate sum of components of each part.

- VG, Gist applied directly on color images. By using small parameters we find 64 components.
- VS, Zernike cannot be performed on color image. So, we calculate the first 8 components of each color level. The data set is composed of the three vectors (VC, VG, VS) of each image of the iaprtc benchmark [11]. It contains 20000 images.
- The color descriptors: all images are collected in a matrix denoted $MC(20000, 96)$. (20000 lines per every picture, 48 the number of color components that we have chosen).
- The Gist descriptors: all images are collected in a matrix denoted $MG(20000, 64)$.
- The shape descriptors: all images are collected in a matrix denoted $MS(20000, 24)$.

3.3 Clustering of the database

In this work, we use the K-means algorithm defined by McQueen [14, 8]. With a simple implementation, the K-means algorithm is the most used clustering algorithm. It partitions data in K clusters. The algorithm divides the inputs data set in K clusters, in which the objects within each cluster are as close as possible to each other as far as possible to other clusters. Each cluster in the partition is defined by its objects and its centroid.
K-means is an iterative algorithm that minimizes the sum of distances between each object and its cluster centroid. The initial position of centroids determines the final result. So, the centroids should initially be placed as far as possible from each other in order to optimize the algorithm. K-means out exchanges the objects of cluster until the sum can not decrease any more. The result is a set of compacts and clearly separated clusters, provided that they have chosen the correct value K, the number of clusters.
The main steps of K-means algorithm are:

Require: Set of N data, denoted by X. Number of desired group, denoted by k
Ensure: A partition of k groups $\{\mathcal{C}_1, \mathcal{C}_2, ...\mathcal{C}_k\}$
Begin
1) Random initialization of the $\mathcal{C}_k$ centers ;
repeat
2) Assignment: create a new partition by assigning each object to the group whose center is the closed;

$$x_i \in \mathcal{C}_k \quad \text{if} \quad \|x_i - \mu_k\| = \min_j \|x_i - \mu_j\| \quad \forall j$$

With μ_k the center of the k classes;

```
3) Representation: Compute the associated centers to the
new partition;
```

$$\mu_k = \frac{1}{N} \sum_{x_i \in \mathcal{C}_k} x_i$$

until `the algorithm converges to a stable partition;`

End

We apply the algorithm K-means to the three matrix MC, MG and MS with different values of k.

4 Proposed methods

Color (C), Gist (G) and shape (S) information have been the primitive image descriptors in our content based image (I) retrieval systems. We propose to determine the equilibrium of similar images as a Nash Equilibrium. Therefore, we split I, the original optimization variable, into three strategies Color, Gist and shape [1, 2] and we denote $I = (C, G, S)$ and $F = (F_c, F_g, F_s)$.
We consider a three-players static game of complete information. The first player is the Color descriptor that is used to control color in image, denoted by C. The second player is the Gist descriptor, denoted by G. Then, the third is the Zernike descriptor who controls the shape, denoted by S.

4.1 Existence of a Nash equilibrium

We consider the functionals $J_C(C, G, S)$, $J_G(C, G, S)$ and $J_S(C, G, S)$ defined by:

$$J_C(C, G, S) = \frac{1}{2}||C - F_c||^2 + \frac{\epsilon}{2}(||\nabla C||^2 + ||\nabla G||^2 + ||\nabla S||^2) \tag{4.1}$$

$$J_G(C, G, S) = \frac{1}{2}||G - F_g||^2 + \frac{\epsilon}{2}(||\nabla C||^2 + ||\nabla G||^2 + ||\nabla S||^2) \tag{4.2}$$

$$J_S(C, G, S) = \frac{1}{2}||S - F_s||^2 + \frac{\epsilon}{2}(||\nabla C||^2 + ||\nabla G||^2 + ||\nabla S||^2) \tag{4.3}$$

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.

$$\begin{cases} \text{Find } (C^\star, G^\star, S^\star) \text{ such that:} \\ \min\limits_{C} J_C(C, G^\star, S^\star) = J_c(C^\star, G^\star, S^\star) \\ \min\limits_{G} J_G(C^\star, G, S^\star) = J_g(C^\star, G^\star, S^\star) \\ \min\limits_{S} J_S(C^\star, G^\star, S) = J_s(C^\star, G^\star, S^\star) \end{cases} \tag{4.4}$$

Theorem 1. *There exists a Nash equilibrium* $(C^\star, G^\star, S^\star)$ *solution of the problem (4.4).*

Proof. The functionals J_C, J_G and J_S are convex and lower semicontinuous. Then we have at least the existence of one Nash equilibrium $(C^\star, G^\star, S^\star)$, for more details see [4]. □

4.2 Algorithm of Nash Equilibrium

Finding the Nash equilibrium requires solving the last problem (4.4). The Nash equilibrium is computed by the following decomposition algorithm.
Set $n = 0$. Starting from an initial design pair $I^{(0)} = (C^{(0)}, G^{(0)}, S^{(0)})$.
Step 1:
Phase 1: solve the problem

$$\min_{C} J_C(C, G^{(n)}, S^{(n)}) \to C^{(n+1)}$$

Phase 2: solve the problem

$$\min_{G} J_G(C^{(n)}, G, S^{(n)}) \to G^{(n+1)}$$

Phase 3: solve the problem

$$\min_{S} J_S(C^{(n)}, G^{(n)}, S) \to S^{(n+1)}$$

Step 2:
set $I^{(n+1)} = (C^{(n+1)}, G^{(n+1)}, S^{(n+1)})$ until convergence, redo the parallel phases 1, 2 and 3.

4.3 Phase of recognition

The researched images must undergo the same pretreatment as the images of the base of training. 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, the method of the k nearest neighbors is to consider the k training samples whose entrance is closed to the new input x. In our case we work with the Euclidean distance to determine the membership of the wanted image to the different classes that represent the three levels of representation. The searched result is the intersection of the three classes of membership.

5 Numerical Results

To test our method, we have used the database iaprtc12 [11] which is composed of a variety of 20000 images, see Figure 3.

Figure 3: Example of images of the database ipartc 12.

The primary objective of this paper is to present a recognition system for similar images, able to reduce research time.
The system allows reducing the number of images to check, theoretically, to a maximum of 1% percent of the training set. Experimentally, the number of images to be checked varied between 0.03% to 0.8% of the training set, according to the nature (the clusters of membership) of the image.
To minimize the number of images to check we work with a number of clusters $k = 50$ for Gist and Zernike and $k = 30$ for color histogram.
To verify the effectiveness of our approach, we have perform a comparative study between our algorithm and the direct use of KNN with the presented descriptors and finally with the matrix distance between images.
Table 1 presents the detection rates and the maximum time needed to recognize an image.

Table 1: Comparative study between the proposed method and classical method.

	Matrix distance$^{(1)}$	KNN Zernike$^{(1)}$	KNN GIST$^{(1)}$	KNN Color$^{(1)}$	K-means+KNN$^{(2)}$	Proposed Method$^{(2)}$
Detection rate	100%	78%	77%	45%	91%	99%
Max time to find an image	34 min	4 min	5 min	7 min	30 sec	35 sec

(1) The outputs of the algorithms are the 5 closest pictures and we verify the existence of the searched image.
(2) The algorithm return between 1 to 3 images and we verify the existence of the searched image.
The clustering allowed us to reduce the number of pictures to be checked to a maximum of 88 pictures, so the minimal time needed in the proposed methods (30s and 35s). The classic methods need to check the entire databse(20000 pictures). The integration of game theory allowed us to increase the classification rate obtained by K-means and KNN from 91% to 99%.

By using a smaller value of the number of clusters k, the proposed system can be used to detect similar images, the following figures present some request results:

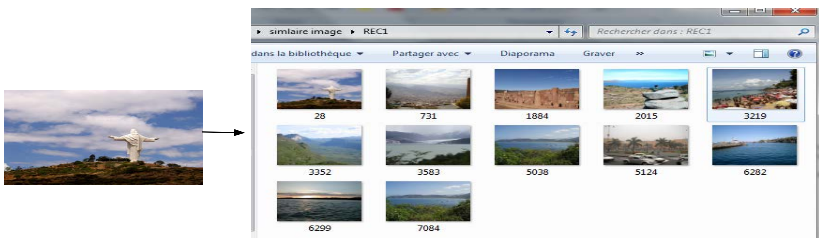

Figure 4: Example of a similar image search, for a picture with a natural background level

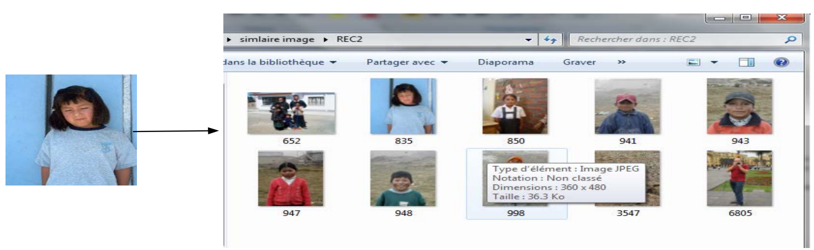

Figure 5: Example of a similar image search, For an image of one person.

6 Conclusion

In this article, we have proposed a search method for similar images based on K-means algorithm and the integration of game theory. The description of images is performed using three descriptors Gist, Zernike and reduced color

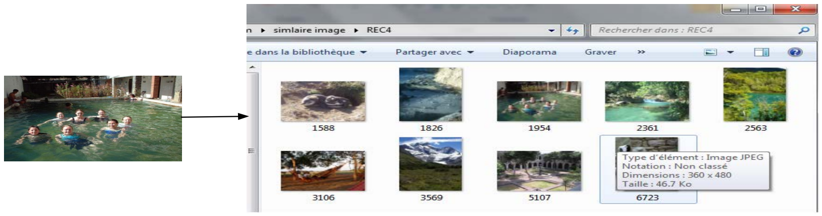

Figure 6: Example of a similar image search, For a complex content.

histogram. Experimentally the proposed method can effectively reduce the search time compared to conventional methods based on the KNN or distance matrix. By using a smaller number of clusters, the proposed system extract a large number of similar images. Thus, it might be used as search engine. The proposed system present quite similar results to the existing search engine.

Acknowledgements. The authors would like to thank the editor and the anonymous referees.
This work has been partially supported by LIRIMA.
The authors would like to express their gratitude to Professor JADIR Ahmed (High School of Technology-Essaouira, Cadi Ayyad University, Marrakech, Morocco) for his valuable help to improve the writing.

References

[1] R. Aboulaich, A. Habbal, N. Moussaid, *Split of an optimization variable in game theory, Math. Model. Nat.* Phenom(MMNP), 5(7)(2010), 106–111. http://dx.doi.org/10.1051/mmnp/20105720

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

[3] P. Agouris, J. Carswell , A. Stefanidis, *An environment for content-based image retrieval from large spatial databases*, Journal of photogrammetry and remote sensing, 54(4)(1999), 263–272. http://dx.doi.org/10.1016/s0924-2716(99)00025-8

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

[5] T.Baranidharan, D.K. Ghosh, *Medical Image Classification Using Genetic Optimized Elman Network*, American Journal of Applied Sciences 9 (1)(2012), 123–126. http://dx.doi.org/10.3844/ajassp.2012.123.126

[6] O. Bencharef, B. Jarmouni, A. Souissi, *Research of Similar Images Based on Global Descriptors and Multiple Clustering*, International Journal of Engineering and Technology, 5(3)(2013), 3142–3151.

[7] M. Braschler, C. Peters, *Cross-Language Evaluation Forum: Objectives, Results, Achievements*, Information Retrieval, 2004. http://dx.doi.org/10.1023/b:inrt.0000009438.69013.fa

[8] G. Celeux, E. Diday, G. Govaert, Y. Lechevallier, H. Ralam-Bondrainy, *Classification Automatique des Données*. Bordas, Paris, 1989.

[9] P. Destuynder and M. Jaoua and H. Sellami. *A dual algorithme for denoising and preserving edges in image processing*. J. Inv. ILL-Posed Problem, 15(2005), 19–35.

[10] J. Gobeill, H. Mller, P. Ruch, *Translation by Text Categorization: Medical Image Retrieval in ImageCLEFmed 2006*, Springer Lecture Notes in Computer Science, LNCS 4730 (2007), 706–710. http://dx.doi.org/10.1007/978-3-540-74999-8_88

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

[12] A. Khotanzad, Y. H. Hong, *Invariant image recognition by Zernike moments*, IEEE Trans. Pattern Anal. Mach. intell, 12(5)(1990), 489–497. http://dx.doi.org/10.1109/34.55109

[13] H. K. Kim, J.-D. Kim, D.-G. Sim, D.-I. Oh, *A modified Zernike moment shape descriptor invariant to translation, rotation and scale for similarity-based image retrieval*, in Proc. IEEE Int. Conf. Multimed, 1(2000), 307–310. http://dx.doi.org/10.1109/icme.2000.869602

[14] B. Mirkin, *Clustering for Data Mining: A Data Recovery Approach*, Chapman and Hall/CRC, 2005. http://dx.doi.org/10.1201/9781420034912

[15] J.F. Nash, *Non-cooperative Games, Annals of Mathematics*, 1951. http://dx.doi.org/10.2307/1969529

[16] A. Oliva, A. Torralba, *Building the gist of a scene: the role of global image features in recognition*, 155(2006).

[17] J. J. Rocchio, *Relevance feedback in information retrieval, The SMART Retrieval System*, Experiments in Automatic Document Processing, NJ(1971), 313–323.

[18] G. Salton, C. Buckley, *Term weighting approaches in automatic text retrieval*, Information Processing and Management, 24(5)(1988), 513–523. http://dx.doi.org/10.1016/0306-4573(88)90021-0

[19] K. Sparck Jones, C. J. Van Rijsbergen, *Progress in documentation*, Journal of Documentation, 32(1976), 59–75.

[20] M. Swain, D. Ballard, *Color indexing* , International Journal of Computer Vision, 7(1991), 11-32. http://dx.doi.org/10.1007/bf00130487

[21] M. H. Yang, N. Ahuja, *Detecting Human Faces in Color Images*. In Proc. of ICIP98, 1(1998), 127–130. http://dx.doi.org/10.1109/icip.1998.723442

Received: October 8, 2014; Published: November 20, 2014