MIR system for mobile information retrieval by image querying

A.SOUISSI , H.TABOUT and A. SBIHI

Department of Telecommunication Systems and Decision Engineering, Ibntofail University, Kenitra, Morocco

Summary
This paper presents a new mobile service called MIR (Multimedia Information Request). The proposed approach relies on VAS (Add value service**)** technology of the mobile network (MMS, SMS and WAP) and content based image retrieval techniques. The aim of this service is to facilitate the search of historical monuments information and architectural heritage of Morocco. Photos taken by user mobile phone are sent by MMS to a server. After the geographic location stage the server analysis and measures their similarities and compares them with prestored images. In order to visualise results information, we have proposed two options: the first consists of a detailed message SMS about the monument described in the request image. The second is concerned with a WAP portal in order to display detailed results to each similar image. The last technique allows users to browse and extend their research. An implemented prototype system has demonstrated a promising retrieval performance for a test database containing more than 3000 color images.

Key words:
Information search, Content based image retrieval, Median Online, ***VAS Mobile Network****, Geographic Localization.*

1. Introduction

With the rapid evolution of mobile networks and technologies used in mobile phones, a lot of applications and services of information were made available to mobile clients. These applications are either integrated directly on mobile phones, or exploit SMS and WAP technologies to send queries to a server of information. These applications include train schedules; pray time, stock market indicators, weather and many other applications….

Currently all theses information are available on users mobile phones, any they can either send an SMS containing a microcode specified for the desired search or by navigating on a WAP portal. Now with the exponential growth of cell phones equipped with cameras, many new applications and services have been emerged. In this context, we have proposed a prototype that provides information using images taken by a mobile phone. In fact, the number of mobile phones equipped with cameras exceeds the number of digital cameras which creates new applications with cameras integrated in mobile. Imagine yourself as a tourist in Rabat contemplating the mausoleum Mohamed V (Fig 1-1), and you would like more information about this typical monument, you can learn about it by visiting your books and travel guides, but you can do better: take a picture with your cell phone; send this picture to a service provider by a multimedia message (MMS). After a moment, you receive an SMS text message which describes the monument and gives you more information.

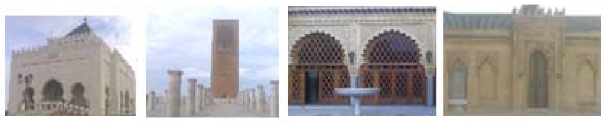

Fig. 1-1 Mausoleum Mohamed V images

That is a kind of service we propose in the MIR system: access to information, wherever you are by a simple and visual query on a mobile terminal. The user will then be localised with its physical context that will offer the best conditions of service. Another option is to specify the city in the user application. From a technological point of view, getting to location information is possible with a GPS system. It may also be available through the cellular telephone infrastructure in the mobile network and through information CID or triangulation. However, knowing the position of a phone, and consequently its user, is not sufficient to determine what interests him. Futhermore, he may want to get information about a monument which is far by hundreds of meters. The location information surely helps to refine the context of the request, but does not capture the intention.

A more explicit query is needed, as a photo taken from the cell phone. That's exactly what we proposed to look for the relevant information.
To find a picture stored in a large database of images from a basic photo is a difficult operation. We need to represent the image by visual descriptors (color and texture).
The aim of the MIR project is to develop new search services of multimedia information in a context of mobility. We make use of image recognition techniques by exploiting their contents and techniques for location-based services VAS mobile technology, to searching information on tourist sites and architectural heritage of Morocco.

Manuscript received April 5, 2008
Manuscript revised April 20, 2008

The remainder of the paper is organized as follows. The next section summarizes related works. Section 3 provides a brief review of the hardware and software architecture. In section 4, the content based image retrieval method is described. In section 5, simulation results and evaluation are provided, and finally, concluding remarks are offered in section 6.

2. Related work

There are many similar approaches to MIR; Snaptotell [1] is a system for tourist. It intends to retrieve information using mobile phones. Authors have used color histograms as descriptor to retrieve images. In [2], the used client device is a PDA (Personal Digital Assistant) connected to the Internet via wireless communication (WLAN).
This architecture requires a wireless access point installed in the neighborhood. The system uses an iPAQ 3870, a camera NexiCam, a probe of orientation, and a GPS receiver. The position detection is provided by a GPS connected to the PDA. However, the direction and tilt sensor are connected to the PDA through a laptop. Our approach is similar to the Snaptotell solution. We have further developed the method of indexing and searching by content on the basis of a Local Descriptor calculated from the interest images' regions, instead of a global descriptor based only on color.
We have also proposed an alternate method of visualization of search results using WAP technology that offers many convenient displays compared to the text message limited by the small size and which can contain only characters.

3. The functioning and system architecture.

3.1 Hardware and software architecture

The service MIR (Fig 3-1) is a client / server system. The client is a mobile phone equipped with a camera, a Java interface and possibly a GPS detector. The server is a powerful and multiprocessor computer with an application that receives, analyzes and processes the images requests from the client and returns the response information via SMS. Also, it allows users to visit the WAP portal, which deals with the same resource of data for more detail and reliability of results.

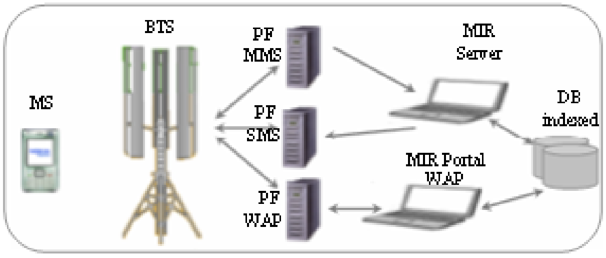

Fig. 3-1: Hardware architecture of MIR service

The system is made of 2 principal processes, the first is MMS communication. This process manages the images' reception, coming from the MMS platform, and stores them on a buffer memory; it is often listening to the requests and functions in a multithread manner with the other process. The second, Response Info process, is divided into 3 subprocesses which function in a sequential manner: Localization, recognition and manging results. The third Indexing process functions independently of the two other processes to index and manage data on the database.

When the system receives an image query that launches the sub-process of localization starting from a parameter of géo-localization, this stage is not conclusive, but narrows the field of research in the database, which minimizes the response time. After the stage of recognition, the system returns the information of the most similar image to the original one with a threshold of confidence determined by the experimentation.

The sub process of image recognition is based on content based image retrieval system. Images are characterized by local descriptors; calculated by making use of low levels features such as color and texture. Theses features are extracted from interest regions. (Fig 3-2) exhibits an overview of the system.
The MIR service proposes two methods to enhance retrieval performances. The first consists of exploiting the geo-localization techniques to locate the mobile phone, and to determine the zone of research, this method is employed, if the sent requests' site is in the neighborhood of the place of the user. In the second method, the customer program founded in the mobile phone invites the user to specify the city where the monument is, thus the research is limited on the city. We limited our study to the first technique.

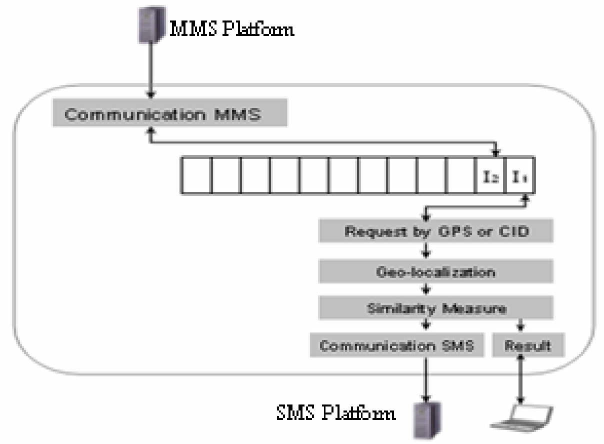

Fig. 3-2: Operation of the MIR server software

3.3 Geographic-location technique

Mobile technology provides several methods for the mobile phones' localization, The CID (Cell Identification), is the cheapest technology because it does not need any material set up. When the mobile is in a zone covered by the network, it is connected to a BTS antenna. By identifying this antenna we can easily locate the mobile phone. This localization is very fast (less than 5 seconds) but it is not very precise because it depends on the antennas number and the distance between each one (the more the antenna is isolated, more the zone of cover is vast, and the localization is less precise). In an urban zone, the precision varies between 100 and 700 meters. On the rural area, that can reach 25 kilometers. The triangulation is a technique which requires preliminary the set up of a program on the SIM cards. The localization is done in about five seconds and is more precise than the CID (150 meters in urban zone and 5 kilometers in rural area).

The use of the worldwide system GPS (Global Positioning System), requires the set up of a GPS module in the mobile, which is more expensive than the other means used for the geo-localization, but the localization is very precise. In this case another problem arises: Buildings or high agglomerations may cause many problems in capturing the satellite signal. Moreover, the time of localization is longer.

MIR service uses two techniques of geo-localization which are GPS and CID. For the GPS, the system calculates the distance between GPS co-ordinates of the request site Sr, and those of the base Si by the distance given by:

$$d(M_1, M_\varepsilon) = R\left(\frac{\pi}{2} - ASin\left(Sinx_1 \times Sinx_2 + Cos\left(y_1 - y_2\right) \times Cosx_1 \times Cosx_2\right)\right)$$

After the Geo-localization stage, the system returns over the images, whose distances with the site request, vary between 15 and 50m (precision of the standard systems), then compares them with the query image by the visual contents. The most similar image presents the threshold of confidence and the minimal distance at the request site.

The use of the localization by CID remains the best option to use, because the majority of mobile customers do not have a mobile phone with a GPS detector. This technique requires the classification of the base of the sites by CID, which is very complicated for the urban areas, considering the density of the BTS. As a matter of fact, it is impossible to plan an optimal cellular network (Fig 3-3) with standards of proximity, and the operations of Handower between BTS and the breakdowns. A mobile phone receives the signals of several cells of the BTS, and dialogues with the cell which has the best signal and quality. The distance is not necessarily the nearest one.

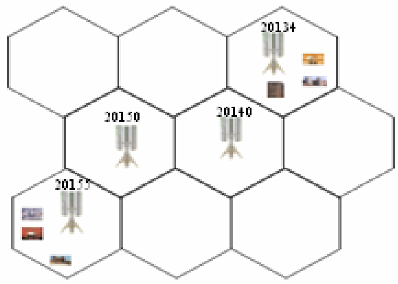

Fig. 3-3: Scheduling an optimal mobile cellular network

To take into account all the above mentioned constraints, it is necessary to add the priority of classification of a site to a CID, by using a trace station to recover the CID of the BTS with their priorities in each site. This is very fastidious work in the case of thousands of sites to be indexed, or by using an information processing system of planning the mobile cellular network which simulates the cover on a platform SIG.

3.4 WAP portals

The WAP portal is used in parallel with MIR system (Fig 3-4), to provide more options of results models, so that, the client can navigate easily in the page, containing the results, and be ensured of the relevance of images. In return, SMS message gives information about only one relevant image, showing limited research capabilities.

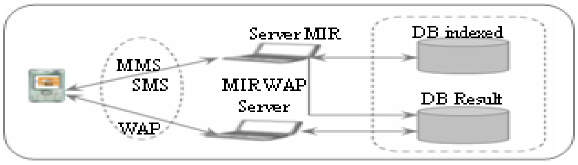

Fig. 3-4: operation of WAP portal

In each research stage, the system records the references of the relevant images in a results table, structured by the number of the customer, the date and the hour of sending of the request. In addition the WAP server reaches the table of results of the database via a dynamic script on request of the customer. With this solution, MIR service behaves as a CBIR system.

4. Image recognition by visual content

The process of the content image retrieval constitutes the decision making in the MIR system. Many applications use this technique, in particular, biometrics, medical images analysis, etc.

The description of the image by its visual contents is less subjective and richer than a textual description, and on the other hand, the image is independent of the language of research, consequently, the result of research by the images contents is often effective compared to research by keywords.

4.1 CBIR system

A CBIR system is characterized by two aspects, the indexation and the research. The first one relates to the data-processing of image representation and the second one relates to the use of this representation to retrieve the most similar images to the query image. Traditional architecture of CBIR system, presented in (Fig 4-1)

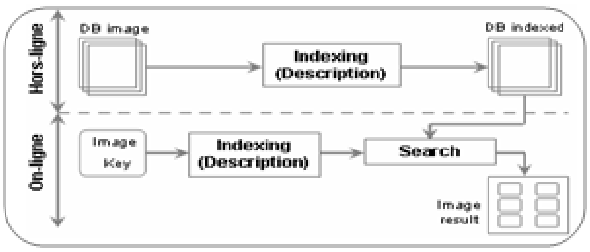

Fig. 4-1: Classic architecture e of a CBIR system

breaks up into two data processing runs: a “Out-line” phase for indexing or the calculation of the descriptors of the images, and an "On-line" phase of research.

Two principal constraints exist in CBIR systems. The first one relates to the types of the used descriptors, and their discriminating capabilities for images identification, with the various variations: scaling, rotation, translation and lighting change. The second constraint is related to the speed in the research stage in the real time context, and especially in a mobile environment which depends on other network constraints. So the descriptor must have a reduced size, invariant with the changes mentioned above and adapted to the entirety of the images in the database.

Two principal approaches exist to calculate the descriptors of images. Local descriptors, in which, features are extracted from interest regions or by using interest points on the image. Global descriptors, calculated from the whole image; the information carried by the objects of the image can be lost. The figure 4-2a represents the ‘Hassan tower’, it is noticed that the color of the sky is dominant in the image, knowing that the color of the sky is variable; consequently the global description of this image by the color gives bad results. On the opposite, the local description of each region of the image (Figure 4-2b) is more precise and gives better results.

(a) (b)

Fig. 4-2: Hassan Tower Image: (a) original image, (b) segmented image

The local descriptors constitute another vision of content based image indexing. Here, we are not interested to recognize the elements constituting the image, but, rather with the similarity between the images, or more precisely, between the objects and the regions which are in the images. For that it is necessary to pass by a segmentation step to divide the image into uniform regions and to thereafter extract the local descriptors from each region.

4.1 Image segmentation

Segmenting image into interests regions *R*,i, is an essential stage which allows to tackle the extraction of local information of the image. It consists of finding structures of joint dependences between pixels or block of pixels and to gather them on uniform visual classes by making use of automatic classification algorithms.

For segmentation stage, we have used the median online method [5], by using hybrid entry parameters of color and texture. The parameters of the color are composed by the averages of the components H and S of space HSV of the square zones Znm of dimension n * m (Figure 3), the parameter of texture is the entropy of the spectrum of the Z_{nm} zones.

$Z_{nm}\ (P_H\ , P_S\ , P_T\ )$ Avec

$n = k \times height(image)$ and $m = k \times width(image)$

Fig. 4-3: Image segmentation example

$$P_H = Moy_H(Z) = \tfrac{1}{n.m}\sum_{l}^{l+n}\sum_{c}^{c+m} H(i,j)$$

$$P_S = Moy_S(Z) = \tfrac{1}{n.m}\sum_{l}^{l+n}\sum_{c}^{c+m} S(i,j)$$

$$P_T = Ent_V(Z) = \tfrac{1}{n.m}\sum_{1}^{l+n}\sum_{c}^{c+m} \widetilde{V}(i,j) \times Log\left(\widetilde{V}(i,j)\right)$$

$$\widetilde{V}(i,j) = \tfrac{1}{n.m}\sum_{l}^{l+n}\sum_{c}^{c+m} V(i,j) \times Exp\left(-2j\pi\left(\tfrac{Un}{n} + \tfrac{Vm}{m}\right)\right)$$

With $\widetilde{V}(i,j)$ is the Discret Fourier Transform of $V(i,j)$

The algorithm online median [4], proposed by R. Mettu and C Plaxton is an online solution for the k-means problem [3]. Instead of optimizing overall the placement of *K* centers (*K* fixed), those are placed one by one, according to the principle of Facility Location. The online median algorithm calculates successively, and in a definite way, the centers of clusters. The process stops when a stop criterion is checked, or when all the data were selected like center. The algorithm has fixed parameters α, β, γ and δ clean with the algorithm. First of all, define the value of a ball *A*, center *x* and radius *r*

$$val(A) = \sum_{y \in A} \left(r - d(x,y)\right)$$

The child of a ball *A* of *x* center and *r* radius there is a point that checks $d(x,y) \prec r\beta$. Recall that $d(x,\ y)$ represents $\min_{x \in X} d(x,y)$, with $d(x,y)$ is the L1 distance

Each step of adding a center goes as follows:

- Calculating the value of all the centered balls at a *x* point *D* data and $\frac{d(x,Z)}{\gamma}$ radius
  Where *Z* is all centers have already been placed. If $Z = \varnothing$; (Case of the first center), the radius is chosen $\max_{x \in X} d(x,y)$
- Selection of the A_0 ball of maximum value
- While A_i contains several sons
  Consider the centered balls by checking $d(x,y) \prec \beta.r_i$, of radius $r_{i+1} = \frac{r_i}{\alpha}$.
  Either A_{i+1} the ball maximum value, and x_{i+1} the correspondent center. Its radius is the r_{i+1} precedent.
- When the ball A has only one son, we chose as a new center of the cluster. This is tantamount to making
  update a ball with a radius decreases with each iteration by $r_{i+1} = \frac{r_i}{\alpha}$, and that moves at each step towards the region of higher density, estimated in value.

After the segmentation stage, we assume the following: for each image *I* of the base *B*, we lay out many regions. Each region R_i with image *I* is represented by a descriptor $D(R_j)$ which describes the color and texture. Thereafter, we associate each image to several visual descriptors modeled:

For all $I \in B, R_j \in I$, the descriptor of image *I* is given, by $D(R_J) = \{D_C(R_j), D_T(R_j)\}$ with D_C and DT are respectively the color and the texture descriptors of region Rj. The database is organized in an arborescent way (Fig 4-5), a monument contains several sites $S_i(C, CID_k^p, I_j, INF)$, If those are characterized by their GPS characters, the CID of the BTS servers with their priorities, the images of the sites duplicated on several I_j angles, and information *INF* that concerns them. Finally each image is described by local descriptors extracted from its interests regions. Thus the structure of indexing the base appears:

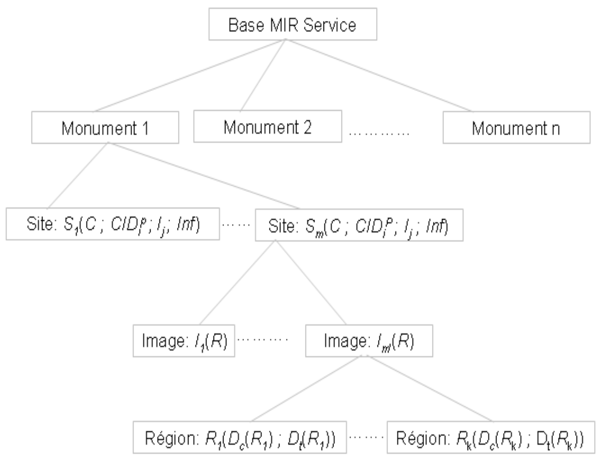

Fig. 4-5: MIR Indexation structure

4.2 Color descriptor

Color is a very popular descriptor used for indexing and retrieval, several studies show that the color descriptor are effective [5]. The intersection of histograms is a very used technique for similarity measurement and comparison between images [6]. The histogram is easy and fast to calculate, and robust with rotation and translation, but presents some drawbacks, since it has a big size, which penalizes the retrieval time. Also it does not have spatial information on the positions of colors, and is very sensitive to small changes of luminosity.

There are other approaches to characterize the color, the statistical of color moments [7], the correlogram and the auto-correlogram [8]. But generally, the color histogram remains the most used feature in CBIR systems, since it keeps a strong capacity of image description. We have used here the dominant color descriptor. In [9] the author proved that 4 to 6 colors are enough to model the histogram of an image, the dominant colors are defined like the maxima of the histogram of color. This descriptor is composed for each class of color i by 3 parameters: The dominant color CDi, the percentage pi which represents the class compared to the image and the variance σ_i of the color in the class. This descriptor then composed of N = 4 to 6 triplets (CD_i, P_i, σ_i), N is the number of dominant colors detected in the image. This descriptor eliminated the problem from the size of the histogram for indexing and retrieval. To fill the absence of spatiocolorimetric information in the descriptor of the dominant color, we have used space coherence descriptor (SCR) [10] of the dominant colors on the level of the image region. In this descriptor, a low value of SCR indicates that the color is dispersed in the image, while for a homogeneous dominant color, SCR will be close to 1.
As for the sensitivity of light, to solve this problem, we used the space of HSV color which is less sensitive to the luminosity and close to the human vision compared to other spaces RVB and L*a*b. For the calculation of the color descriptor, we limited on the two components H and S and we avoided the V component which represents the luminosity, thus the distribution of the descriptor of $D_C(R_j)$ color of a R_j region is given by:

$$I \in B, R_j \in I, D_C(R_j) \subset D(R_j) \text{ Avec}$$

$$D_C(R_j) = \{(CD_1, P_1, \sigma_1, SCR_1); \ldots\ldots; (CD_n, P_n, \sigma_n, SCR_n)\}$$

4.3 Texture descriptor

The concept of texture is used to translate a homogeneous aspect of an object surface on an image. Thus texture appears by visual information which makes it possible to describe it qualitatively thanks to following adjectives: coarse, fine, smooth, mottled, granulous, marbled, regular or irregular. There are several descriptors of texture, the most used are related to correlation through the matrix of co-occurence, Tamura [11] proposed 6 parameters having a visual significance, and characterizing it directly. Certain authors combine several approaches in order to benefit from the various types of analysis, co-occurence matrices, and decomposition in subbands [12, 13], wavelets and neural networks [14].

For the description of texture, we carried out a spectral analysis of the texture of the interests regions by parameters of entropies extracted the spectral field. These parameters measure specific characteristics of the distribution of the components of the spectrum and quantify properties such as the regularity, the direction, the linearity and the smoothness. Also let us note that the entropy measured is invariant by rotation and translation. In fact, that makes it possible to see whether a texture contains repetitive motifs. Also announcing that disturbed textures which have many frequential components give a greater value of the entropy. On the contrary, highly structured textures have low entropy. Thus each R_j region identified after the segmentation step is described directly by the high and low values of the entropy of the spectrum of the Z_{nm} zones.

$$I \in B, R_j \in I, D_T(R_j) \subset D(R_j) \text{ Avec}$$

$$D_T(R_j) = \left\{ \begin{array}{c} Arg\max_{i \le n}\left(Entp_L(Z)\right); Arg\min_{i \le n}\left(Entp_L(Z)\right) \\ n = 4 \end{array} \right\}$$

4.3 similarity measurement

To estimate the similarity between two images, it is necessary to measure the similarities between their local descriptors corresponding to the same attributes. There are two ways to combine the two similarities. The first one:

$$d_g = w_c d_c + w_t d_t$$

Where d_i and w_i respectively represent the standardized distance and the weight associated with attribute (c for the color and t for texture). The distance from Kullback [15] is used for the color while it calls upon the Euclidian distance of texture.

$$d_c(R, R') = \sum_{n=1}^{N}\sum_{m=1}^{M}(Dc_{nm} - Dc'_{nm}) \times Log\left(\frac{Dc_{nm}+1}{Dc'_{nm}+1}\right)$$

Where M is the number of elements of each component of color, N is the component count of color (H = 16, 4 parameters of the 4 dominant colors and N = the 2 two components H and S). Dc and Dc' are respectively the color descriptor of a query image region and database image regions.

$$d_t(R,R') = \sqrt{\sum_{n=1}^{N} (Dt_n - Dt'_n)^2}$$

Where N is the number of the texture descriptors component (N=8), D_t and $D_{t'}$ are respectively the texture descriptor of the regions R and R' of the image request and an image from the database. The second approach is hierarchical, the attributes are then considered continuously by a descending order of importance, by evaluating distances d_c and d_t. The weights can be fixed directly by the experiment or evaluated automatically by using a procedure of iterative tests. We carried out an automatic classification not supervised with ascending hierarchical clustering method (AHC) [16]. On a sample of the images, by using the distance dg, and the aggregation criterion (or average distances): The distance between two classes is the average between the elements of the two classes. In continuation we measured the clustering relevance of similar images, by changing the weight w_i color and texture (Fig 4-1) the best score is estimated by the combination (w_c = 70% et w_t = 30%).

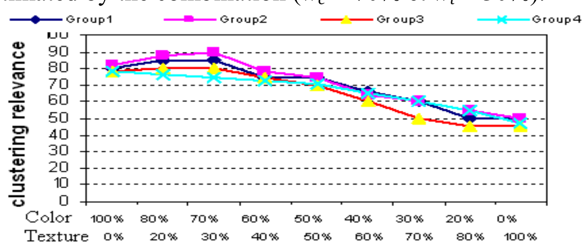

Fig. 4-1: Clustering relevance

4.4 CBIR Algorithm

For the automatic images indexing, we have to carry out initially the elimination of nonsignificant regions *FR*, in particular the sky, the greenery, the passengers etc,… We have used a K-NN classifier [17] who tests if an region R_i belongs to a class *FR*, the distance used is dg, then the algorithm of this operation is given as follows:

$For \quad All\,(I_k \in B)$
$R_i = clustering(I_k)$
$D(R_i) = Extractsignature(R_i)$
$IF\,(KNN[D(R_i); D(FR_j)] = False) Then$
$Save[R_i(D)] inDatabase$
$else$
$Save[reference(R_i)] inLogFileFalse$
$Endif$
$incr?ent \quad k$

After classification, a log file is created to evaluate this operation; it contains the detail of the refused regions. The research phase consists in returning the information of the site which presents the similar regions at the query image interest regions with the distance dg, thus the algorithm of the operation.

$S_i = Geo - localizati\,on(CID \quad or \quad GPS)$
$R_k = clustering\,(imagereq)$
$D_{req}(R_k) = Extractsig\,nature\,(R_k)$
$For \quad All\,(R_j \in S_i)$
$result_k = dg\left(Dreq\,(R_k); D(R_j)\right) \le threshold$
$Next \quad k$
$sort\,(result_k)$
$Rrelev_k = \arg\min(result_k)$
$IF\,(Rrelev_k \in unique \quad S_i) then$
$return\,(\inf(S_i))$
$else$
$return\,(\inf(S_i))$
with takin g into account CID priorrity or ,
minimal distance between tw o GPS points
END IF

5. Experiment and result

For the evaluation of MIR system, we evaluated CBIR method, applied to a corpus of 3000 images of the monuments and touristy places, divided into four groups {Hassan, Oudaya, Chalah, Rabat Medina} (Fig 5-1), by using Recall / precision metrics [18], and we compared this method with that based on the colors histogram (Fig 5-2).

Fig 5-1: Example of Rabat monuments images

The results obtained show that the returned images relevance is improved using our CBIR method compared to the histogram method. For a recall of 50%, we obtained 70% of precision against 48%.

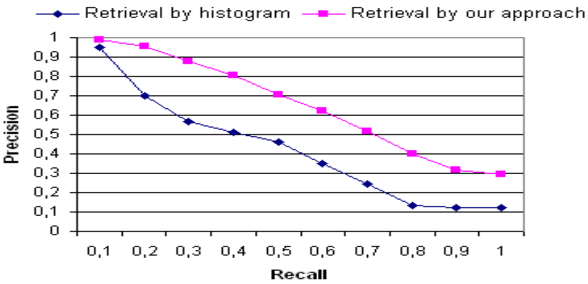

Fig. 5-2: Recall/Precision diagram

For MIR system simulation, we developed a client emulator that communicates with the MIR system software via LAN network, an overview of the proposed emulator is shown in (Fig 5-3). In table1, we present the simulation results with different options (SMS and WAP), with and without geo-localization. The use of geo-localization gives better results, especially when we use the WAP interface which provides more reliable results.

Fig. 5-3a: user interface and image selection

Fig. 5-3b: example of a query image and returned information

Fig. 5-3c: WAP portal interface and retrieved images

Table1: Simulation result

Monument Image	information relevance without Geo-localization		information relevance with Geo-localization	
	SMS	WAP	SMS	WAP
Group1	52%	76%	80%	100%
Group2	50%	75%	71%	100%
Group3	54%	72%	79%	100%
Group4	47%	70%	80%	100%

6 Conclusions

MIR Service will attract the client's attention towards the culture of the visited cities, considering its ease of use. For the telecommunication operators, to apply such a service, will increase the use of the **VAS** networks, and will create another partnership dimension with information services providers, travel agencies, museums... etc. Also this technique can be projected for another application, in particular, information retrieval on products by shape, text and bars' codes.

References

[1] CHEVALLET J.-P., LIM J.-H., VASUDHA R., « SnapToTell: a Singapore Image Test Bed for Ubiquitous Information Access from Camera », Poster on European Conference on Information Retrieval ECIR'05, 2005.

[2] MAI W., DODDS G., TWEED C., « A PDA-Based System for Recognizing Buildings from User-Supplied Images », *Mobile and Ubiquitous Information Access: Mobile HCI 2003 International Workshop*, vol. 2954 / 2004, Lecture Notes in Computer Science, Springer-Verlag Heidelberg, Sept 2003, p. 143–157.

[3] DUDA R. O., HART P. E.: Pattern classification and scene analysis. *John Wiley and Sons* (1973).

[4] R. Mettu and C. G. Plaxton. The online median problem. In FOCS '00: Proceedings of the 41st AnnualSymposium on Foundations of Computer Science, page 339, 2000.

[5] Smeulders, M. Worring, S.Santini, A.Gupta and R. Jain. Content-Based Image Retrieval at the End of the Early Years. *IEEE Trans. on Patt. Anal. and Machine Intell.*, vol. 22, no 12, pp. 1349-1380, Dec.2000.

[6] Swain and D.H. Ballard. Color indexing. *Int. J. of Computer Vision*, vol. 7, no. 1, pp. 11-22, 1991.

[7] Stricker and A. Dimai. Color indexing with weak spatial constraints. *SPIE Proc.*, vol. 2670, pp. 29 - 40, 1996.

[8] Huang, S.Kumar, M. Mitra, W.J. Zhu and R. Zabih. Image Indexing Using Color Correlograms. *Proc. of Conf. on Comp. Vision and Patt. Recog.(CVPR)*, San Juan (Puerto Rico), pp. 762-768, 1997.

[9] K.Idrissi Segmentation et description de la couleur pour l'indexation, la navigation visuelle et la recherche dans les bases images thèse de doctorat, université claude bernard de lyon, Mars 2003

[10] B.S Manjunath, P. Salembier, T.Sikora: Introduction To MPEG-7 Wielly, 2002, ISBN: 0-471-48678-7

[11] R. Haralik: *Statistical and structural approach to textures,* Proceedings of the IEEE, vol 67 NO. 5, pp 786-804, 1979

[12] K. S THYAGARAJAN, T. NGAYEN, C. PERSONS: *Maximun like lihood approach to texture classification using Wavelet transform*, In Proc, IEEE International conference on image Processing 1994

[13] A. KUNDU, J.L. CHEN: *Texture classification using qnf bank-based Subband decomposition* CVGIP

[14] C. BUSH, M. H. GROSS: *Interactive neural network texture analysis and visualization for surface reconstruction in medical imagin*, In R. J. Hubbold and R. Juan, editors,

Enrographics 93 pp 49-60, Oxford, UK Eurographics, Backwell Publishers, 1993
[15] S. Kullback: *Information theory and statistics*. New York: John Winley & Sons, 1959.
[16] WILLIAMS, WT. 1967 *A general theory of classificatory sorting strategies: I. Hierarchical systems*. Computer Journal. 9, 378 – 380
[17] T. HASTIE, R. TIBSHIRANI, AND J. FRIDMAN: *The Elements of Statistical Learning: Data Mining, Inference,and Prediction*. Springer-Verlag, 2001.
[18] J.R. SMITH: *Image Retrieval Evaluation*, CBAIVL" 98: Proceedings of the IEEE Workshop on Content - Based Access of Image and Video Libraries, pp. 112, 1998.

Abdelmoghit SOUISSI is an engineer in a telecom company since 2001. He received the master and the Diploma of higher Studies in Computer Engineering from the University , Hassan Premier, Settat, Morocco in 1999 and 2001 respectively. Currently, he is a PhD student in the Telecommunication Systems and Decision Engineering laboratory, Ibntofail University, Kenitra, Morocco. His research interests have been oriented toward the pattern classification and content based image retrieval.

Hassan TABOUT is an engineer in a telecom company since 2002. He received the master and the Diploma of the higher Studies in Computer Engineering from the University , Hassan Premier, Settat, Morocco in 1999 and 2001 respectively. Currently, he is a PhD student in the Telecommunication Systems and Decision Engineering laboratory, Ibntofail University, Kenitra, Morocco. His research interests includes pattern recognition, content based image retrieval, interactive learning in multimedia systems.

Abderrahmane SBIHI received the "License" of Science degree in physics from the University Mohammed V of Rabat, Morocco, in 1983 and the "Docteur ès Sciences" degree from the university Ibn Tofaïl of Kénitra, Morocco, in 1995. In 1985, he joined this last university where he worked as an Assistant Professor. From 1992 to 1995, he joined the University of Science and Technology of Lille, France, as a Guest Assistant Professor at the Institute of Technology and as a Research Assistant in the Automatic Control Center where he worked on the adaptation of numerical mathematical morphology to pattern classification problems. In 1995, he returned to the university Ibn Tofaïl where he is currently Professor and Head of the Computer Vision and Pattern Recognition Laboratory (LIRF). His current research interests include computer vision, pattern recognition and multidimensional data analysis.