

Système MMSINFO destiné pour la recherche d'information par des requêtes image dans le réseau mobile 2G et 3G

Abdelmoghit.Souissi, Hassan.Tabout et Abderhmane. Sbihi

Laboratoire des Systèmes de Télécommunications.

et Ingénierie de la Décision Université Ibn Tofail,

Faculté des Sciences, BP. 133, 1400, Kénitra, MAROC

ab.souissi@menara.ma, h.tabout@ieee.org, sbihi@univ-ibntofail.ac.ma

RÉSUMÉ.

Nous présentons dans cet article, un nouveau service mobile MMSINFO qui exploite la technologie RVA du réseau mobile 2G et 3G, et les techniques de reconnaissance des images par le contenu, pour la recherche d'information sur les sites touristique du MAROC, à partir de photos prises par les téléphones mobile. Les images requêtes sont envoyées par MMS au serveur de l'application qui mesure leurs similarités aux images de la base, et retourne une information détaillée sur l'image similaire à l'image requête. Cette information est envoyée au client par SMS. Pour détailler et fiabiliser davantage l'information retournée aux clients, nous avons proposé l'utilisation d'un portail WAP, pour l'affichage des résultats des requêtes, avec cette solution les clients peuvent naviguer et s'assurer de la pertinence des images retournées. Nous présentons aussi les résultats de la méthode de reconnaissance des images et aussi un scénario d'utilisation du service dans l'annexe.

MOTS-CLÉS: Recherche d'information, système CBIR, segmentation image, Réseaux RVA Mobile, localisation géographique, Approche Bayésienne.

1. Introduction sur le projet MMSINFO

Maintenant avec la prolifération des technologies RVA (Réseau à valeurs ajoutés) des réseaux mobiles (SMS, MMS, WAP, GPRS, UMTS). La plupart des opérateurs ont mis à la disposition des clients mobiles des services supplémentaire de recherche d'information et de localisation géographique. Ces services sont basés essentiellement sur les technologies SMS et WAP pour envoyer ou recevoir les requêtes et les réponses des clients. Parmi ces applications, SMSINFO le service d'information sur les horaires du train, de prières, les indicateurs de la bourse, la météo ... etc. Toutes ces informations sont à la disposition des clients mobiles, sur envoi d'SMS contenant des microcodes spécifiés pour la recherche désirée, ou par navigation sur un portail WAP. Maintenant avec les téléphones mobiles munis d'appareils photos et d'OS qui supporte des applications. Beaucoup de nouvelles applications et services vont émerger. Dans cette optique, nous proposons de développer dans le projet MMSINFO un nouveau service de recherche d'information, dans un contexte de mobilité. Il s'agit de l'utilisation des techniques de reconnaissances d'images par le contenu, ainsi que les techniques de géo-localisation dans un service, basé sur les réseaux RVA mobile, destinée à la recherche d'informations sur des sites touristiques du Maroc.

Le présent article est organisé comme suit : dans la seconde partie nous présentons l'architecture matériel et logiciel de l'application, dans la troisième partie nous proposons une méthode d'indexation et de reconnaissance d'images par le contenu visuel, et dans la quatrième partie, nous montrons les résultats de l'expérience d'une telle approche..

2- Architecture matériel et logiciel du service MMSINFO

Le service MMSINFO est une application orienté client/serveur (figure1). Le client est un téléphone portable, muni d'un appareil photo, d'une interface Java et éventuellement d'un détecteur GPS. Le serveur est un ordinateur multiprocesseur, muni d'une application qui reçoit, analyse et traite les requêtes images provenant du client, et retourne la réponse d'information en SMS. Le client a le choix de consulter le portail WAP pour plus du détail et de fiabilisation des résultats.

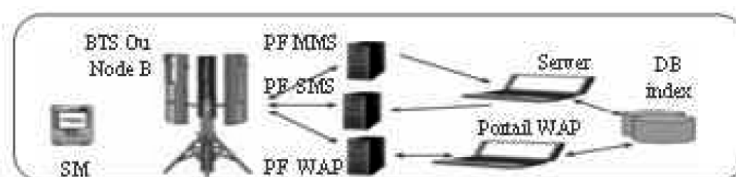


Figure 1: Architecture Matériel du service MMSINFO

Le système est composé de 4 processus (figure2 : Communication MMS, Communication SMS, Géo-localisation, et Mesure Similarité), et fonctionne selon le mode multithread. En effet Le système est constamment en écoute des requêtes provenant de la plate-forme MMS, et stocke les images dans une mémoire tampon avant la phase du traitement et du retour d'information au client. Le portail WAP accède à la table 'Résultat', et permet une affiche des résultats en détail.

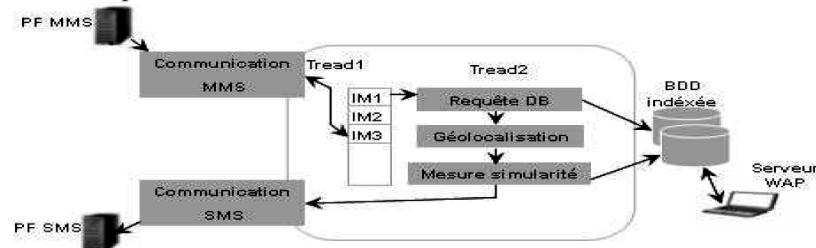


Figure 2 : Fonctionnement de l'application serveur du service MMSINFO

3- Reconnaissance des images par le contenu

L'information de géo-localisation aide à affiner le contexte de la requête, mais n'en capture pas l'intention des clients. Du point de vue technologique, l'obtention d'une information de localisation est d'ores et déjà possible avec un système GPS. Elle peut également être obtenue par l'infrastructure cellulaire du réseau mobile, grâce à l'information CID ou la triangulation. Cependant, connaître la position d'un téléphone, et donc de son utilisateur, n'est pas suffisante pour déterminer ce qui l'intéresse.

Le processus de reconnaissance d'image par le contenu constitue dans le système MMSINFO l'étape de finalisation de la recherche, pour retourner les informations pertinentes. Ce processus est basé sur un système RIPC (Recherche Image par le contenu), caractérisé par 2 aspects indissociables, l'indexation et la recherche. Le premier concerne le mode de représentation informatique ou la description des images et le second concerne l'utilisation de cette représentation dans un but de recherche. La description des images, soit calculée d'une manière globale, à partir de toute l'image, soit locale, le calcul est réduit sur des régions ou des points d'intérêt sur l'image. Les descripteurs globaux sont peu discriminants, vu que l'information portée par les objets de l'image peut être perdue. Au contraire, la description locale de chaque région de l'image est plus précise.

4.1 Segmentation des images

La segmentation est une étape essentielle pour pouvoir s'attaquer à l'extraction d'information locale de l'image. Pour cette étape, nous avons utilisé la méthode **SOM** (Self Organizing MAPs) [1], en utilisant des paramètres de la couleur de l'espace HSV.

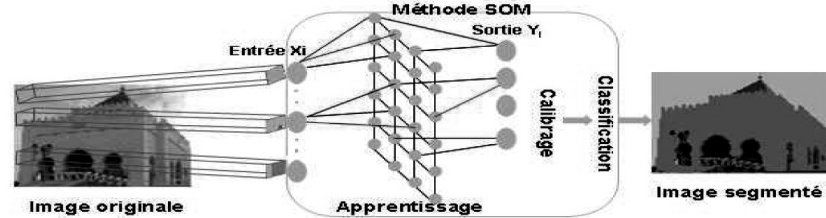


Figure 3 : méthode de segmentation d'une image par la méthode SOM

Après l'étape de segmentation, chaque image $I = \{R_1, R_2, \dots, R_j\}$ de la base B est décomposée en régions. Chaque région est représentée par un descripteur $D(R_j) = \{D_c(R_j), D_t(R_j), D_f(R_j)\}$ décrivant la couleur, la texture et la forme.

4.2 Descripteur de la couleur

Pour la description de la couleur des images de la base, nous avons utilisé le descripteur de la couleur dominante. Dans [2] l'auteur a prouvé que 4 à 6 couleurs suffisent pour modéliser l'histogramme d'une image. Ce descripteur est composé pour chaque classe de couleur i par 3 paramètres : La couleur dominante CD_i , le pourcentage P_i que représente la classe par rapport à l'image et la variance σ_i de la couleur dans la classe. Pour combler l'absence d'information spatiale dans le descripteur de la couleur dominante, on a utilisé le descripteur de la cohérence spatiale [3] SCR des couleurs dominantes. Pour remédier au problème de luminosité, nous avons utilisé seulement les composante H et S , et nous avons évité le composante V qui représente la luminosité, ainsi la distribution du descripteur de la couleur $D_c(R_j)$ d'une région R_j est donné par :

$$D_c(R_j) = \{(CD_1, P_1, \sigma_1, SCR_1), \dots, (CD_n, P_n, \sigma_n, SCR_n)\} \text{ Avec } n = 4$$

4.3 Descripteur de la texture

Il existe plusieurs descripteurs de la texture, les plus utilisés sont les fonctions de corrélation à travers la matrice de cooccurrence. Pour la description de la texture, nous avons utilisé les 2 paramètres de texture l'entropie ENT et le moment inverse IDM . D'après l'étude de [4] les deux paramètres précités admettent le plus fort pouvoir discriminant parmi les 14 paramètres de Haralick. La texture de chaque région sera décrite par le vecteur $D_t(R_j) = \{ENT, IDM\}$

4.2 Descripteur de la forme

Pour la description de la forme, on a utilisé les invariants de Hu [5] basé sur le calcul du moment géométrique centré des régions d'image, donnée par la fonction :

$$\mu_{pq} = \sum \sum I(x, y)(x - x_0)^p (y - y_0)^q$$

Avec x_0 et y_0 le barycentre de l'image. Les cinq premiers paramètres de Hu sont invariants en translation, rotation et changement de l'échelle, ils sont définis dans les équations suivantes, ainsi $D_F(R_j) = \{\phi_1, \dots, \phi_5\}$.

$$\begin{aligned} \phi_1 &= \mu_{20} + \mu_{02} ; \phi_2 = (\mu_{20} + \mu_{02})^2 + 4\mu_{11}^2 ; \phi_3 = (\mu_{20}\mu_{02})^2 - \mu_{11}^2 \\ \phi_4 &= (\mu_{30} - 3\mu_{12})^2 - (3\mu_{21} - \mu_{30})^2 ; \phi_5 = (\mu_{30} - \mu_{12})^2 - (\mu_{21} + \mu_{03})^2 \end{aligned}$$

4.3 Mesure de la similarité

Pour estimer la ressemblance entre deux régions R et R' , il faut mesurer les similarités entre leurs descripteurs correspondant aux mêmes attributs. Les poids sont évalué automatiquement en utilisant une procédure de teste itérative.

$$D_G = w_C D_C + w_T D_T + w_F D_F$$

Où D_i et w_i représentent respectivement la distance normalisée et le poids associé à l'attribut i (c pour la couleur, t pour la texture et f pour la forme).

La distance de Kullback est utilisée pour la couleur tandis qu'il fait appel à la distance Euclidien L2 pour la texture et la distance L1 pour la forme.

$$d_K(R, R') = \sum_{n=1}^N \sum_{m=1}^M (D_{nm} - D'_{nm}) \times \log \left(\frac{D_{nm} + 1}{D'_{nm} + 1} \right) \quad d_L(R, R') = \sqrt{\sum_{n=1}^N (D_n - D'_n)^2}$$

Où M est le nombre d'éléments de chaque composante de couleur, N est le nombre de composante de couleur ($H = 16$, 4 paramètres des 4 couleurs dominantes et $N = 2$ les composantes H et S). D_C et D'_C sont les descripteurs de la couleur de 2 régions R et R' .

4.4 Modélisation du système CBIR

Dans la première phase, nous avons organisé la base d'une manière arborescente, chaque site $S_i(C, CID_k^p, I_j, INF)$ est caractérisé par leur cordonnée GPS, les CID des BTS serveuses avec leurs priorités, des images I_j de site S_i dupliquées sur plusieurs angles, et les informations INF qui décrit le site en question. Chaque image est décomposée en régions, décrites par des descripteurs locaux.

Pour l'étape de l'indexation automatique des images au niveau de la base, nous avons procédé à la réduction des régions R_j de chaque sites S_i à un ensemble M_i qui représente que les régions d'intérêts les plus importantes, autrement dit l'ensemble des régions d'intérêts RI_j différent à l'ensemble M_i' des régions non significatifs RF_k , notamment les régions qui représentent le ciel, les passagers, les arrières plans,....etc.

On dit qu'une région RI_j est une Région d'intérêt de l'ensemble M_i pour une distance ϵ , si la distance entre cette région et toutes les régions RF_k de M_i' est supérieure à ϵ . Ce qui peut s'écrire également :

$$R_j \in I, R_j \in M_i \Leftrightarrow R_j = RI_j; \forall RF_k \in M'; D_g(R_j; RF_k) > \varepsilon$$

Soit $B = \{M_1, M_2, \dots, M_n\}$ une collection de n ensemble des sites S_i . Chaque site est représenté par un ensemble des régions d'intérêt $RI = \{RI_1, RI_2, \dots, RI_j\}$, avec j le nombre de régions d'intérêts caractéristiques. Soit une image requête Q . Nous souhaitons retrouver l'ensemble $M_i \in B$ dont l'une des régions caractéristiques est la plus proche de l'une des régions de l'image requête. Cet ensemble est celui qui a la plus grande probabilité $P(M_i, RI_j / Q)$. C'est la probabilité qu'une région R_Q de l'image Q corresponde à une région d'intérêt RI_j de l'ensemble M_i .

Soit $H = \{h_1 \cup h_2 \cup \dots \cup h_n\}$ l'ensemble des hypothèses pour une image $Q = \{R_{Q1}, R_{Q2}, \dots, R_{Qn}\}$. Une hypothèse h_k correspond à ce que une région d'intérêt RF_j d'un ensemble M_i corresponde à une région R_{Qn} de l'image requête Q . Le signe $\cup$ représente le ou logique. Notons également que si une hypothèse h_k est vraie, toutes les autres hypothèses sont fausses, car une région de l'image requête ne peut correspondre qu'à un seul ensemble M_i . Si une image requête est représentée par l'ensemble de toutes les hypothèses possibles, la probabilité $P(M_i, RI_j / Q)$ peut être exprimée par $P(M_i, RI_j / H)$. Le modèle le plus proche est celui qui contient la région ayant la plus grande probabilité. En utilisant le théorème de Bayes [6], nous avons :

$$P(M_i, RI_j / H) = \frac{P(H / M_i, RI_j)P(M_i)}{P(H)} \text{ Avec } P(H) = \sum_{i=1}^n P(H / M_i, RI_j)P(M_i)$$

Avec $P(M_i, RI_j / H)$ est la probabilité a posteriori, $P(H / M_i, RI_j)$ et $P(M_i)$ sont respectivement la probabilité d'appartenance à une classe M_i et la probabilité a priori. D'après la décision de Bayes, on affecte l'image Q à la classe qui maximise $P(M_i, RI_j / H)$. En effet il faut estimer les probabilités $P(H / M_i, RI_j)$ et $P(M_i)$. Les ensembles M_i contient des RI dissemblables, pour cela, nous avons procédé à la classification automatique de chaque ensemble M_i en classes homogènes C_i . Ainsi nous avons assimilé les classes C_i à des classes gaussienne, $P(H / M_i, RI_j)$ s'est ramené à $P(H / C_i, RI_j)$, les n classes C_i sont équiprobable soit $P(C_i) = \frac{1}{n}$, et pour le calcul de cette probabilité, nous avons utilisé l'approche de Bayes quadratique :

$$P(H / C_i, RI_j) = \frac{1}{(2\pi)^d (\det \Sigma)^{1/2}} \exp\left(-\frac{1}{2}(X - \mu)^T \Sigma^{-1}(X - \mu)\right)$$

Avec μ est la moyenne d'une classe C_i , X est le descripteur local d'une région de l'image requête, d est la dimension du descripteur et Σ représente la matrice de covariance. Par conséquent le problème de l'apprentissage se ramène à l'estimation des moyennes μ_i et les matrices de covariances Σ_i des classes C_i . Notre approche consiste à

appliquer l'algorithme Online-median et déterminer les centres, i.e. Les moyennes, puis estimer la variance à posteriori pour chaque classe C_i des RI_j homogènes. La méthode online median [7] est une solution online du problème de k-median. Au lieu d'optimiser globalement le placement de k centres (k fixé), ceux-ci sont placés un par un, selon le principe des Facility Location. L'algorithme online-median place successivement, et de façon définitive, les centres du clustering. Le processus s'arrête lorsqu'un critère d'arrêt est vérifié, ou lorsque toutes les données ont été choisies comme centre. L'algorithme dispose de paramètres fixés α , β , γ et δ . la valeur d'une boule A de centre x et de rayon r est définie comme suit : $val(A) = \sum_{y \in A} (r - d(x, y))$

Le fils d'une boule A de centre x et de rayon r est un point y qui vérifie $d(x, y) < r\beta$.

Rappelons que $d(x, y)$ représente $\min_{x \in X} d(x, y)$, avec $d(x, y)$ est la distance D_g .

Après l'étape de classification, chaque ensemble M_i de RI_j va être regroupé selon des classes uniformes $M_i = \{C_1, C_2, \dots, C_k\}$, chaque classe C_k est caractérisée par une moyenne μ_k et une matrice de covariance Σ_k .

L'étape de la recherche consiste à retourner l'information du site qui présente au moins une région d'intérêt similaires à une région de l'image requête, en utilisant la règle de Bayes pour estimer les probabilités a posteriori, et avec les paramètres calculés dans la phase d'apprentissage, ainsi l'algorithme de l'opération de la recherche d'information.

$G_i = \text{Géo_localisation} \{S_1, S_2, \dots, S_k\}$; $R_{Q_j} = \text{Segmentation (Image Requête)}$

$D(R_{Q_j}) = \text{Description_Local} \{R_{Q_1}, R_{Q_2}, \dots, R_{Q_j}\}$

Tant que ($j \leq \text{Card} \{R_{Q_j}\}$)

Tant que ($i \leq \text{Card} \{G_i\}$)

$\text{Resultat}_j = \text{Argmax} \{P(C_i, RI_i / R_{Q_j})\}$

$i++ ; j++$

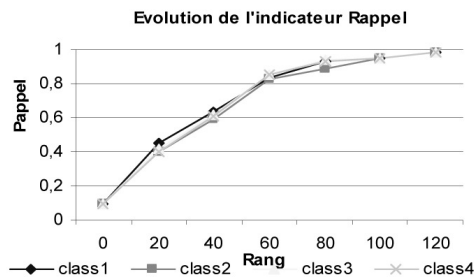
Retourner_et_envoyer_SMS_Information ($S_x / \max(\text{resultat}_j)$)

Pour l'affichage des résultats avec le WAP, le serveur accède à tous les informations de la table 'Résultat' enregistrée et indexée sur la base par le N° de client et la date d'envoi de la requête, et affiche les informations et les images similaires avec l'ordre de similarité.

5 Expérience et résultat

Premièrement, nous avons évalué la méthode CBIR proposée par le paramètre 'Rappel', appliquée sur un corpus de 1000 images de 15 monuments et sites touristiques de MAROC, divisés selon les villes en 4 classes {Rabat, Fès, Marrakech et

Meknès}. Pour tester le service MMSINFO, nous avons simulé les plates formes SMS et MMS par des PC qui communiquent en temps réel avec le client (Émulateur sur PC) et avec le serveur via un intranet, et nous avons évalué le pourcentage de l'information pertinente avec et sans géo-localisation, Aussi en utilisant le service WAP.



class	% d'information pertinente sans Géo-localisation		% d'information pertinente avec Géo-localisation	
	SMS	WAP	SMS	WAP
Class1	62%	86%	82%	100%
Class2	60%	85%	81%	100%
Class3	64%	82%	79%	100%
Class4	59%	82%	80%	100%

Les résultats de reconnaissance obtenus montre que la technique de localisation a permis une très bonne identification des sites par rapport à une recherche global. Aussi le portail WAP a donné plus de fiabilité d'information pour le service, ce qui constitue une complémentarité entre la géo-localisation, technique de reconnaissance par le contenu et le WAP pour une meilleure utilisation du service MMSINFO.



Figure 3 : Déroulement de la recherche d'information par le service MMSINFO

5 Bibliographie

- [1] T. Kohonen, "Self-Organizing Maps", 3rd ed. New York: Springer-Verlag, 2001, vol. 30, Springer Series in Information Sciences
- [2] K. Idressi Segmentation et description de la couleur pour l'indexation, la navigation visuelle et la recherche dans les bases images thèse de doctorat, université claud de Lyon, Mars 2003
- [3] B.S Manjunath, P. Salembier, T.Sikora: Introduction To MPEG-7 Wiley, 2002.
- [4] R.W.Conn,C.A.Harlow,"Theorical comparaison of texture algorithms"IEEE Transaction on PAMI
- [5] R Mukundan S H Ong P A Lee. Image analysis by tchebichef moments. IEEE Transaction on Image Processing, 10 :1357–1364, 2001.
- [6] R. Mettu and C. G. Plaxton. The online median problem. In FOCS '00: Proceedings of the 41st Annual Symposium on Foundations of Computer Science, page 339, 2000.
- [7] Chang, E. Y., K., Sychay, G., WU : Content base soft annotation for multimodal image retrieval using Bayes point machines. IEEE, 13(1), 26-38