Reconnaissance des fontes arabes par l'utilisation des dimensions fractales et des ondelettes

Nizar Zaghden, Sami Ben Moussa, Adel M Alimi

▶ **To cite this version:**

Nizar Zaghden, Sami Ben Moussa, Adel M Alimi. Reconnaissance des fontes arabes par l'utilisation des dimensions fractales et des ondelettes. Sep 2006, pp.277-282. hal-00112058

HAL Id: hal-00112058
https://hal.science/hal-00112058

Submitted on 7 Nov 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reconnaissance des fontes arabes par l'utilisation des dimensions fractales et des ondelettes

Nizar ZAGHDEN – Sami BEN MOUSSA – Adel M ALIMI

Ecole Nationale d'ingénieurs de Sfax
Route de Soukra km 3.5- B.P. W 3038 Sfax, Tunisie

nizar.zaghden@gmail.com, sami.benmoussa@ieee.org, adel.alimi@ieee.org

Résumé : *Nous étudions dans cet article la reconnaissance des fontes arabes imprimées par l'utilisation des dimensions fractales calculées sur les images des différents blocs issues de différentes fontes. Nous présentons aussi une étude comparative avec les autres méthodes élaborées auparavant pour faire la reconnaissance des fontes et nous faisons aussi une comparaison entre les taux de reconnaissance obtenus par les dimensions fractales et les ondelettes. Un test de robustesse est aussi effectué afin de conclure quelle est la meilleure technique qui permet de reconnaître les fontes sur des images bruitées.*

Mots-clés *:* fonte arabe, dimension fractale, ondelette, test de robustesse.

1 Problématique

L'identification de la fonte est sensible à plusieurs facteurs tel que la diversité des formes, l'irrégularité des alignements, l'irrégularité des dimensions et la présence de formes ambiguës.

La reconnaissance optique de la fonte ou OFR (Optical Font Recognition) a pour objectif d'identifier la fonte utilisée dans la région d'édition originale d'un texte donné. Cependant, malgré l'importance de cette procédure aussi bien au niveau de l'identification de la structure d'un document qu'au niveau de la reconnaissance optique des caractères, l'OFR reste souvent négligé [ZHU 01, CRU 05] et notamment dans le cas de l'arabe [BEN 02].

L'identification de la fonte peut être opérée en amont ou en aval de la reconnaissance optique des caractères (OCR : Optical Character Recognition) [ZRA 95]. Dans les deux stratégies, l'OFR permet d'améliorer les performances de l'OCR et de réduire la complexité de la tâche en ramenant une problématique multifonte à des problématiques monofontes.

Il existe principalement trois stratégies de combinaison de l'OFR avec l'OCR : l'approche a posteriori et l'approche a priori et l'approche hybride.

L'identification a posteriori, reconnaît la fonte après avoir reconnu les caractères [ANI 92]. Elle aide à corriger les erreurs de reconnaissance et elle permet de retrouver la description d'origine du document.

L'identification a priori reconnaît la fonte sans pour autant reconnaître les caractères. Elle simplifie considérablement les tâches ultérieures dans un système OCR en optimisant les recherches par la réduction du nombre de graphismes différents à considérer. De plus, elle améliore les performances du système en bénéficiant des différentes informations relatives à la fonte, préalablement disponibles. L'identification hybride peut être adaptée en combinant l'identification a priori et l'identification a posteriori et c'est notamment l'approche utilisée par Shi [SHI 97].

Vu qu'il n'y a pas de base d'images standard de textes arabes, nous avons construit notre propre base de textes arabes imprimés multifontes. Cette base contient 2500 blocs de textes arabes dans cinq tailles différentes : 16, 18, 20, 22 et 24. Nous avons sélectionné dix fontes arabes parmi celles les plus utilisées : Ahsa (FIG. 1), Arabic Transparent, Buryidah, Diwani, Dammam, Koufi, Kharj, Khobar, Naskh et Tholoth. Ces images ont été enregistrées avec le format libre "TIFF" et avec une résolution de 300 ppp (pixels par pouce). Pour chacune des fontes étudiées 2/3 des images sont utilisées pour l'apprentissage et le reste pour le test. Les images prises en considération sont en niveau de gris. Cependant pour pouvoir appliquer la méthode de dilatation et l'approche CDB, une binarisation a été appliquée sur toutes les images. Nous avons ainsi appliqué deux différentes techniques pour caractériser les différentes fontes :

1. Les dimensions fractales (Box Counting, CDB, dilatation).
2. Les ondelettes (db3, sym5, coif2, bior4.4, dmey).

Nous avons aussi comparé la robustesse de ces deux techniques par rapport à l'ajout du bruit Gaussian et à la variation de résolution.

2 Etude des fontes arabes

Plusieurs études ont été faites sur la reconnaissance des caractères arabes, mais la plupart de ces études négligent l'identification de la fonte. L'identification de la fonte peut être opérée en amont ou en aval de la phase de reconnaissance des caractères. En ce qui nous concerne nous avons pris pour problème le phénomène de l'identification de la fonte. C'est pourquoi le type de reconnaissance que nous avons adopté est le système de reconnaissance de fontes à priori. La figure FIG2

présente une phrase de texte arabe écrite dans les différentes fontes étudiées.

لا يشكل الخط العربي فقط
أداة تجسيد اللغة الحاملة
للخصائص الحضارية
والتاريخية والنفسية والمخيال
الجماعي، ، بل يحمل هذا
الخط أقدس رسالة خص بها

FIG. 1 - Exemple d'image de texte illustré dans la fonte Ahsa

Ahsa	لا يشكل الخط العربي فقط أداة
Buryidah	لا يشكل الخط العربي فقط أداة
Diwani	لا يشكل الخط العربي فقط أداة
Dammam	لا يشكل الخط العربي فقط أداة
Kharj	لا يشكل الخط العربي فقط أداة
Khobar	لا يشكل الخط العربي فقط أداة
Tholoth	لا يشكل الخط العربي فقط أداة
Arabic transparent	لا يشكل الخط العربي فقط أداة
Naskh	لا يشكل الخط العربي فقط أداة
Koufi	لا يشكل الخط العربي فقط أداة

FIG. 2 – Représentation des 10 fontes étudiées

Plusieurs techniques d'extraction de caractéristiques ont été appliquées sur les images de l'écrit afin de caractériser les fontes. Parmi ces méthodes nous citons les caractéristiques globales (texture, densité de l'image, fréquence spatiale, orientation du contenu de texte) [ZRA 95, ZHU 01], ou aussi les caractéristiques issues des ondelettes [GAZ 04, BEN 04].

L'approche utilisée dans [ZRA 95] est basée sur les structures typographiques qui dépendent du script et du langage, par contre l'approche utilisée dans [ZHU 01] qui peut être appliquée pour tous les scripts et les langages, nécessite plusieurs étapes de prétraitements.

Pour ceci, nous avons choisi de travailler avec une nouvelle méthode d'extraction de caractéristiques sans faire de prétraitements. Cette approche est aussi indépendante des scripts et langages : la dimension fractale.

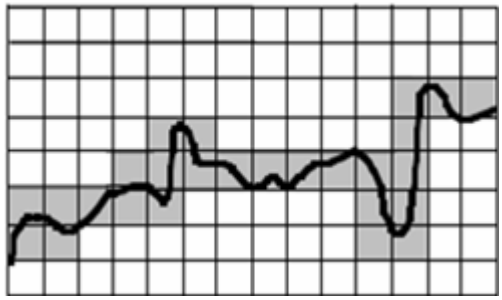
2.1 La dimension fractale

La dimension fractale est l'une des caractéristiques principales de la géométrie fractale. Elle a été utilisée entre autres dans le domaine de reconnaissance de l'écrit pour faire la reconnaissance des scripts, des auteurs et de l'écriture manuscrite. Nous avons choisi d'utiliser cette technique pour faire la reconnaissance des fontes arabes imprimées. Pour ceci nous avons développé une nouvelle méthode de calcul de la dimension fractale que nous avons appelé la méthode de comptage de densité par boîte : CDB.

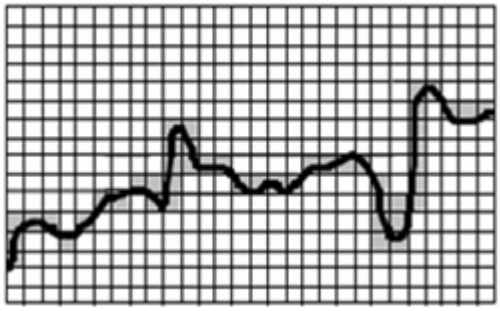
2.1.1 Méthodes de calcul de la dimension fractale

Plusieurs méthodes ont été développées pour calculer la dimension fractale comme la méthode de comptage par boîte (Box Counting) et la méthode de dilatation. La dimension de Box-Counting (FIG. 3) est considérée comme la plus simple méthode pour le calcul de la dimension fractale. Elle est appliquée sur des images non vides [FAL 90].

A partir des propriétés de l'auto-similarité, la dimension fractale d'un ensemble A est définie comme suit : D=log (N)/log (1/r), telle que N est le nombre total des copies distinctes similaires à A et 1/r correspond au facteur d'échelle avec lequel A est divisé [BEN 05].



Recouvrement de l'objet par des carreaux de côté l



Recouvrement de l'objet par des carreaux de côté l/2

FIG. 3 - Méthode de comptage par boîtes (Box-Counting)

La nouvelle méthode que nous avons développée pour le calcul de la dimension fractale (CDB) est inspirée de la méthode de comptage par Boîte. Parmi les autres dérivées de cette méthode, il y a la méthode de comptage de cellules réticulaires [GAN 86, BIS 01] et aussi la méthode de Box Counting différentielle [SAR 92, SAR 94, CHO 99].

Parmi les auteurs qui se sont intéressés au calcul de la dimension fractale par la méthode dilatation, Peleg et all ont adopté l'idée de Mandelbrot et l'ont développé pour calculer l'aire d'une surface [PEL 84].

2.1.2 Méthode proposée : Méthode de comptage de densité par boîte

Nous avons proposé une nouvelle méthode de box counting (CDB: comptage de densité par boîte) que nous pouvons l'appliquer sur des images binaires. En effet cette méthode consiste à décomposer l'image en des grilles, ensuite de calculer pour chaque grille la densité en pixels noirs. Pour chaque image de taille $M \times M$, nous la subdivisons en des grilles de coordonnées (i, j) et de taille $s \times s$, ensuite nous calculons la densité en pixels noirs dans chaque grille alors : n_r (i, j)=densité (i, j).

Nous appliquons cette formule sur toutes les grilles de l'image ce qui nous permet d'obtenir la contribution totale de l'image : $Nr = \sum_{i,j} n_r(i, j)$. Le nombre de boîtes significatives Nr est compté plusieurs fois dans l'image pour des différentes valeurs du facteur d'échelle r (r = s/M) et la dimension fractale peut être estimée de la droite des moindres carrés de log(Nr) en fonction de log(r).

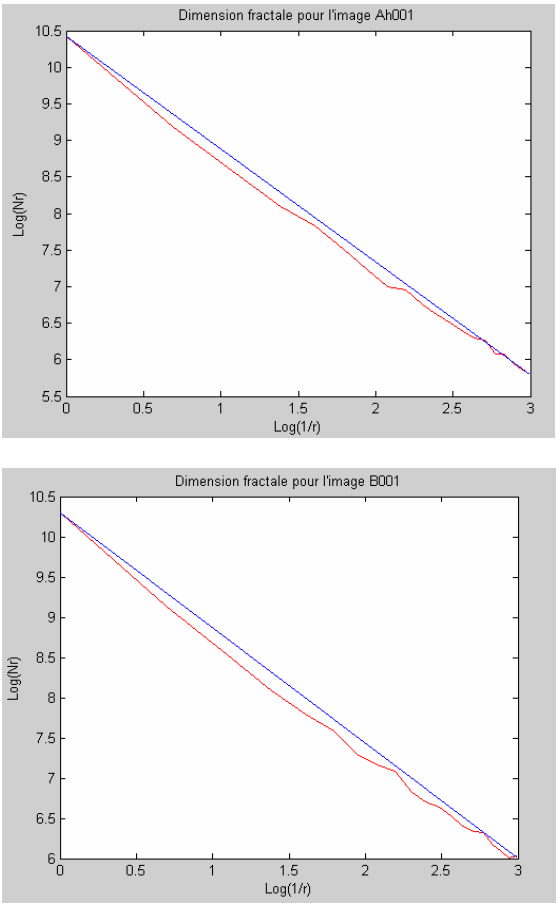


FIG. 4 - Calcul de la dimension fractale par la méthode de comptage de densité par boîte

Ah001 : fonte Ahsa
B001 : fonte Buryidah

Ainsi, les courbes obtenues (FIG. 4) correspondent à une représentation de Log (Nr) en fonction de Log (1/r), tel que r est le facteur d'échelle avec lequel nous subdivisons l'image par des boîtes de taille $s \times s$. Ensuite nous traçons une droite d'approximation de chaque courbe. La dimension fractale est dérivée de la pente de la droite d'approximation et elle correspond alors à: dimension fractale=1-pente. Nous avons comparé cette méthode avec les autres méthodes sur les mêmes images. Plusieurs auteurs ont proposé des méthodes de calcul de cette dimension et les ont appliqués sur les images naturelles de texture (FIG. 5) issues de Brodatz [BRO 66], c'est pourquoi nous avons choisi ces mêmes images références afin de pouvoir comparer notre méthode avec les méthodes proposés dans [PEL 84, GAN 86, PEN 84, KEL 89, SAR 92, SAR 94].

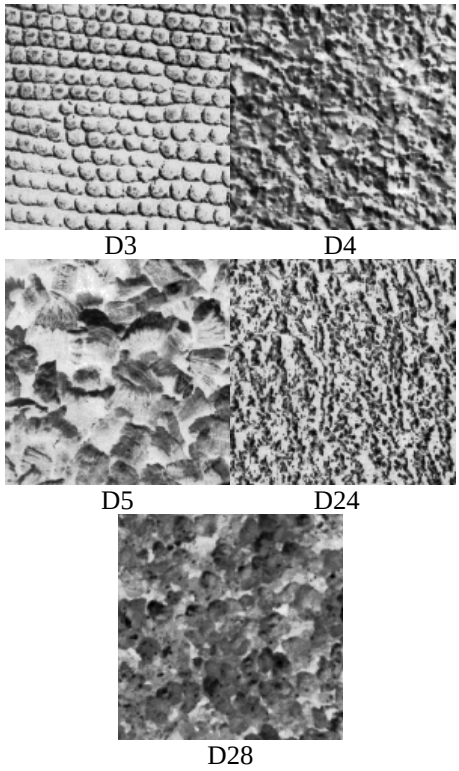


FIG. 5 - Textures naturelles

Nous avons ainsi calculé la dimension fractale de cinq images de textures naturelles et nous les avons comparé avec les autres méthodes appliquées sur les mêmes images.

	D3	D4	D5	D24	D28
CDB	2.77	2.89	2.66	2.83	2.72
Keller	2.63	2.68	2.57	2.57	2.62
Gangepain	2.40	2.45	2.38	2.39	2.41
Peleg	2.69	2.72	2.52	2.59	2.61
Pentland	2.54	2.55	2.38	2.36	2.48
Sarkar	2.60	2.66	2.45	2.45	2.55

TAB.1 - Calcul de la dimension fractale des textures naturelles selon différentes méthodes

D3, D4, D5, D24, D28: Numéros des images dans le livre de Brodatz [BRO 66].

Le tableau TAB1 illustre les différentes valeurs de dimensions fractales obtenues en comparaison avec les méthodes proposées par les autres auteurs.

Les résultats obtenus par la méthode de comptage de densité par boîte (CDB) sont très encourageants puisqu'elles donnent des valeurs discriminantes de dimensions fractales pour ces images de texture. En effet pour les images D5 et D24, la méthode de Sarkar donne le même résultat (2.45), aussi celle de Keller permet d'obtenir la même valeur de dimension fractale (2.57), cependant nous avons obtenu des dimensions distinctes pour les mêmes images: par la nouvelle approche (CDB) que nous avons proposée, nous avons obtenu des dimensions fractales de 2.65 et de 2.83.

Nous avons appliqué cette nouvelle approche (CDB: comptage de densité par boîtes) pour le calcul de la dimension fractale sur les différentes images de texte de notre base d'image textes de taille 16. Afin d'atteindre le meilleur taux de reconnaissance, nous avons opté pour l'utilisation du classificateur kppv (k plus proches voisins) en utilisant trois types de distance: la distance de Hamming, la distance Euclidienne et la distance du maximum.

Outre l'approche CDB, nous avons calculé la dimension fractale avec deux autres méthodes : la méthode de Box Counting et la méthode de dilatation [ZAG 05].

Pour la méthode de Box Counting, le meilleur taux de reconnaissance que nous avons obtenu est de 58.5%, ce qui présente un taux moyen. Pour la méthode de dilatation, nous avons obtenu un taux meilleur que celui obtenu avec la première méthode, et qui correspond à 88%. La nouvelle approche CDB est estimée comme la meilleure méthode puisqu'elle a donné des résultats très satisfaisants, car grâce à elle, nous avons obtenu un taux de reconnaissance fiable correspondant à ***96.5%*** (pour l'ensemble des fontes et pour des images de l'écrit de taille 16). Pour les images de texte de taille 18, 20, 22 et 24, les résultats obtenus sont moins performants et ils varient de 80% à 89,5%.

2.2 Les ondelettes

Nous avons appliqué la transformation en ondelettes sur les différentes images de l'écrit afin d'extraire les caractéristiques des différentes fontes arabes imprimées. Pour chaque niveau de décomposition, nous faisons l'extraction de la moyenne et l'écart type de l'image d'approximation, image diagonale, image verticale et image horizontale. Ainsi pour chaque niveau de décomposition, le nombre d'indices caractéristiques de l'image correspond à 8. Nous avons appliqué plusieurs types d'ondelettes pour faire l'extraction des caractéristiques des différentes fontes arabes imprimées. Les ondelettes utilisées sont : Daubechies (db3), Symlets (sym5), Coiflet (coif2), Bior (bior 4.4), Discrete Meyer (dmey). Pour cette dernière ondelette nous n'avons pas pu dépasser le premier niveau de décomposition. Mais pour les quatre premières ondelettes nous avons fait la décomposition jusqu'au troisième niveau.

En plus de ces coefficients statistiques pour chaque image, nous avons calculé la densité des pixels noirs de l'image (paramètres structurels). En effet nous avons décomposé les images de taille 512* 512 en des blocs de taille 8*8, ensuite nous avons calculé la densité en pixels noirs pour chaque bloc.

Nous avons remarqué une nette supériorité des résultats obtenus avec l'ondelette symétrique coif2, puisque nous avons pu atteindre un taux de reconnaissance de 100% [ZAG 05] pour les textes de taille 16. Cependant les meilleurs taux pour les textes de taille 18, 20, 22 et 24 ont été obtenus aussi avec l'ondelette coif2 et ils varient de 97.5% à 99%.

3 Test de robustesse

Les images sur lesquelles nous avons appliqué la reconnaissance sont des images claires. Cependant il est important de considérer le phénomène du bruit car les images sont en général contaminées par ce processus comme dans le cas du scannage de basse qualité. En premier lieu nous avons ajouté un bruit Gaussian sur les images de l'écrit. En second lieu nous avons effectué la discrimination des fontes sur des images de différentes résolutions. Dans les deux méthodes nous faisons l'apprentissage pour des images de texte de taille 16 ayant une résolution de 300 ppp et ne contenant aucun bruit, ensuite le test est effectué pour des images contenant du bruit Gaussian (première méthode) ou pour des images de résolution inférieure à 300 ppp (deuxième méthode). Nous avons aussi utilisé le classificateur kppv (distance de Hamming, distance Euclidienne et la distance du maximum) pour faire la reconnaissance des images bruitées. De même 2/3 de la base d'images a servi pour l'apprentissage et le reste pour le test.

L'ondelette Coiflet a été testée pour faire la discrimination des images bruitées, cependant cette approche s'est trop dégradée et les taux de reconnaissance obtenus sont de 15% environ. Ainsi, il s'avère que l'ondelette Coiflet, qui auparavant a permis d'atteindre un taux de reconnaissance de 100% pour les images claires, ne permet pas de faire la discrimination des fontes pour les images bruitées.

L'affectation d'un bruit sur des images n'est pas toujours la même, c'est pourquoi nous avons mesuré le rapport du signal bruité (SNR: Signal Noise Ratio) pour les différentes images de test. Le bruit des images est mesuré par la formule suivante:

$$\mathrm{SNR}=10\log\frac{\sum I_{m,n}^2}{\sum (I_{m,n}-J_{m,n})^2}$$

Où $I_{m,\,n}$ et $J_{m,\,n}$ représentent respectivement l'image originale et l'image bruitée.

Nous avons effectué la reconnaissance en utilisant la méthode CDB pour des images bruitées selon différentes valeurs du bruit SNR (FIG. 6). Les résultats sont exprimés dans le tableau TAB2.

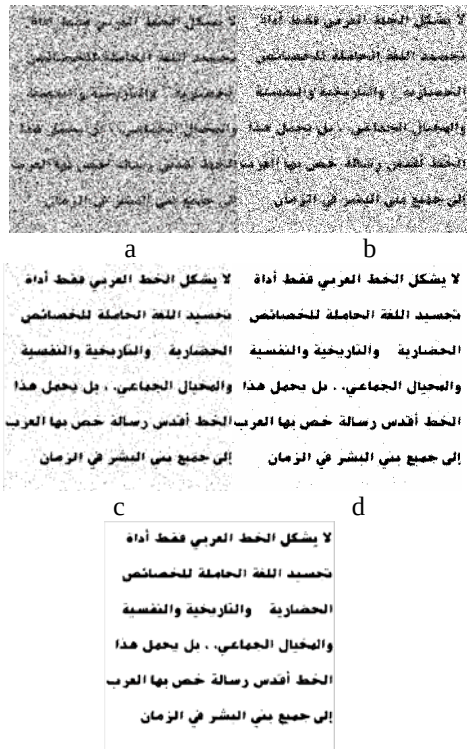


FIG. 6 - Représentation des images pour différentes valeurs du bruit

a: SNR=10; b: SNR=20; c: SNR=30; d: SNR=40; e: SNR=50

SNR	Pas de bruit	50	40	30	20	10
Taux:%	96.5	96.5	96.5	96	90.5	30

TAB. 2 - Taux de reconnaissance des fontes pour différentes valeurs du bruit

Ainsi nous pouvons remarquer que les primitives fractales sont plus discriminantes que les caractéristiques issues des ondelettes et que la méthode CDB est presque invariante pour les images bruitées possédant une valeur du SNR supérieure à 20. Nous avons aussi comparé ces deux méthodes pour des images de résolution inférieure à 300 ppp. Comme dans le cas pour les images bruités, l'ondelette Coiflet s'est avérée insuffisante pour la discrimination des fontes pour des images ayant des résolutions différentes aux résolutions des images de l'apprentissage. En effet le meilleur taux de reconnaissance obtenu avec les caractéristiques issus des ondelettes est de 30%.

Cependant les dimensions fractales issues de la méthode CDB pour des résolutions supérieures à 60 ppp ont permis d'obtenir des taux de reconnaissance presque similaires aux résultats obtenus avec la même approche pour des images claires. Les résultats obtenus pour la variation de résolution sont illustrés dans le tableau suivant (TAB3).

Ainsi nous pouvons conclure que la méthode CDB est robuste dans les deux cas de variation de résolution et des images bruitées et nous pouvons même prévoir que cette méthode peut donner de bons résultats pour la reconnaissance des fontes dans des documents anciens.

Résolution (ppp)	300	200	100	75	60	50	40
Taux (%)	96.5	96.5	96.5	96	95	85.5	80

TAB. 3 - Taux de reconnaissance des fontes pour différentes résolutions

4 Conclusion

Le calcul des caractéristiques des images par la dimension fractale est toujours plus simple puisqu'il permet d'associer à chaque image un seul indice, cependant l'application des ondelettes nécessite un temps de calcul plus long puisqu'elle est utilisée avec plusieurs niveaux de décompositions et elle calcule les coefficients caractéristiques de tous les pixels de l'image. En plus nous avons pu remarquer que la méthode CDB est robuste pour les images bruitées ayant une valeur de SNR supérieure à 20 et ceux de résolution supérieure à 60 ppp. Cependant les ondelettes se sont avérées insuffisantes pour faire la discrimination sur des images non claires.

5 Références

[ANI 92] J.ANIGBOGU : " Reconnaissance des textes imprimés multifontes à l'aide des modèles stochastiques et métriques ", *Thèse de doctorat, Université de Nancy I*, 1992.

[BEN 02] N.E.BEN AMARA : " Sur la problématique et les orientations en reconnaissance de l'écriture arabe" *CIFED'02, Hammamet, Tunisie*, 2002, pp 1-10.

[BEN 04] N.E.BEN AMARA, S.GAZZAH : "Une approche d'identification des fontes arabes", *huitième Colloque International Francophone sur l'écrit et le document, CIFED'04*, 21-25 Juin 2004, La Rochelle, France, 2004.

[BEN 05] S. Ben Moussa, A. Zahour, M.A Alimi, A. Benabdelhafid, "Can Fractal Dimension Be Used In Font Classification," icdar, Eighth International Conference on Document Analysis and Recognition (ICDAR'05), 2005, pp. 146-150.

[BIS 01] A.K.BISOI, J.MISHRA, "On calculation of fractal dimension of images", *Pattern Recognition Letters, vol 22, N°6-7*, 2001, pp 631-637.

[BRO 66] P.BRODATZ, texture: A photographic album for artists and designers. Dover, New York, 1966.

[CHO 99] J.H.CHO, C.W.PARK, E.Y.CHA, "Learning data composition and recognition using fractal parameters", *International Joint Conference on Neural Networks (IJCNN), vol 4*, 1999, pp 2872-2875.

[CRU 05] C.A.CRUZ, R.R.KUOPA, M.R.AYALA, A.A.GONZALEZ, R.E.PEREZ, "High order statistical texture analysis-font recognition applied", *Pattern Recognition Letters, vol 26, N° 2*, 2005, pp 135-145.

[FAL 90] K.FALCONER, "Fractal geometry: Mathematical foundation and applications", Wiley, New York, USA, 1990.

[GAN 86] J.J.GANGEPAIN AND C.R.CARMES, "Fractal approach to two dimensional and three dimensional surface roughness", *wear 109*, 1986, pp 119-126.

[GAZ 04] S.GAZZAH, N. ESSOUKRI BEN AMARA : "Utilisation des ondelettes et des réseaux de neurones en identification des fontes arabes", *Première conférence internationale Signaux, Circuits et Systèmes, SCS 2004, 18-21 mars, Monastir, Tunisie*, 2004.

[KEL 89] J.M.KELLER, R.CROWNOVER AND S.CHEN, "Texture description and segmentation through fractal geometry", *Computer Vision Graphics and Image Processing, vol 45*, 1989, pp 150-166.

[PEL 84] S.PELEG, J.NAOR, R.HARTLEY, AND D.AVNIR, " Multiple resolution texture analysis and classification ", *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), vol6, N° 4*, July 1984, pp 518-523.

[PEN 84] A.P.PENTLAND, "Fractal-based description for natural scenes", *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), vol. PAMI-6, N° 6*, November 1984, pp 661-674.

[SAR 92] N.SARKAR AND B.B.CHAUDHURI, "An efficient approach to estimate fractal dimension of textural images", *Pattern Recocognition, vol 25, N° 9*, 1992, pp 1035-1041.

[SAR 94] N.SARKAR AND B.B.CHAUDHURI, " An efficient differential box-counting approach to compute fractal dimension of image", *IEEE Trans.Systems, man, and cybernitics, vol 24, N° 1*, January 1994, pp 115-120.

[SHI 97] H.SHI, T.PAVLIDIS, "Font recognition and contextual processing for more accurate text recognition", *Proceedings of the International Conference on Document Analysis and Recognition(ICDAR '97)* ,1997, pp 39-44.

[ZAG 05] N.ZAGHDEN, M.CHARFI, A.M.ALIMI , "Optical Font Recognition Based on Global Texture Analysis", *International Conference on Machine Intelligence(ICMI-ACIDCA), Tozeur – Tunisia, November 5-7*, 2005, pp-712-717.

[ZHU 01] Y.ZHU, T.TAN AND Y.WANG: "Font recognition based on global texture analysis*".IEEE trans on Pattern Analysis and Machine Intelligence, vol 23, n°10*, October 2001, pp 1192-1200.

[ZRA 95] A.ZRAMDINI AND R.INGOLD " A study of document image degradation effects on font recognition *",International Conference on Document Analysis and Recognition, vol 2*,1995, pp 740-743.