



HAL
open science

Utilisation de nouveaux paramètres à base de fractale pour la discrimination des fontes arabes

S. Ben Moussa, A. Zahour, M. Kherallah, A. Benabdelhafid, A. Alimi

► To cite this version:

S. Ben Moussa, A. Zahour, M. Kherallah, A. Benabdelhafid, A. Alimi. Utilisation de nouveaux paramètres à base de fractale pour la discrimination des fontes arabes. Sep 2006, pp.283-288. <hal-00112073>

HAL Id: hal-00112073

<https://hal.science/hal-00112073v1>

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.



HAL Authorization

Utilisation de nouveaux paramètres à base de fractale pour la discrimination des fontes arabes

S. Ben Moussa¹, A. Zahour², M. Kherallah¹, A. Benabdelhafid², A. M. Alimi¹.

¹ Université de Sfax,
REGIM, ENIS, BP W, 3038 Sfax, Tunisie

² Université du Havre,
Quai Frissard, BP 1137-76063, Le Havre, France

{sami.benmoussa, monji.kherallah, Adel.Alimi}@ieee.org ; {abderrazak,benabdelhafid}@univ-lehavre.fr

Résumé : La méthode présentée dans cet article permet l'identification des fontes arabes dans des images de textes imprimés. Cette méthode est basée sur l'utilisation de la géométrie fractale. Deux dimensions fractales sont extraites par la technique de comptage par boîte et par celle de la dilatation élémentaire de l'image. Nous avons extrait trois paramètres pour chaque dimension fractale étudiée. Ces paramètres sont expérimentés dans les deux niveaux : ligne et bloc. Nous avons utilisé le KPPV pour classifier dix fontes. La discrimination est meilleure au niveau bloc. Pour évaluer l'aspect discriminatif de nos paramètres, nous avons utilisé une technique neuronale d'auto organisation en utilisant les six paramètres pour chaque prototype. Les résultats obtenus confirment la pertinence de notre méthode.

Mots-clés : AOFR (Reconnaissance Optique de Fontes Arabes), OCR, analyse de texture, dimension fractale, KPPV, réseau de neurone.

1 Introduction

La reconnaissance de l'écriture arabe demeure au niveau de la recherche et de l'expérimentation. Les applications utilisent un vocabulaire limité et/ou une fonte unique [KHO 02]. Elles aboutissent à des taux de bonne reconnaissance assez élevés. Cependant le taux d'identification des caractères change d'une fonte à l'autre. Les systèmes de reconnaissance optique des caractères arabes les plus efficaces ne permettent pas de résoudre le cas où plusieurs fontes co-existent dans le même texte. Le taux d'erreur reste élevé en absence du contexte [KAH 87] [AMI 00].

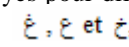
Malgré l'importance de la Reconnaissance Optique des Fontes dans un système de ROC, la ROF reste un problème souvent négligé, et les études dans ce domaine sont rares, en particulier dans le cas des écritures arabes.

Dans ce qui suit, nous allons présenter les problèmes posés par l'identification de fontes arabes. Nous proposerons ensuite notre contribution pour l'identification des fontes arabes, par l'utilisation de

nouveaux paramètres extraits à partir des dimensions fractales. Nous terminerons cet article par une présentation des résultats préliminaires obtenus; suivie de la conclusion et les perspectives.

2 Les problèmes posés dans l'identification des fontes arabes

Nous avons recensé deux types de problèmes pour l'OFR arabe. La première difficulté concerne les écritures calligraphiques et leurs topologies. La deuxième difficulté réside dans la multiplicité des fontes arabes ainsi que la variabilité des caractères dans une même fonte. En effet les caractères de la même fonte peuvent changer de morphologie selon leur position dans le mot. Par conséquent, beaucoup de travaux concernant la reconnaissance des caractères arabes utilisent le contexte mono fonte et s'attaquent rarement au contexte multi-fontes [ABD 98] et [KAN 02].

L'écriture arabe est cursive. Dans un mot, les caractères peuvent être reliés au caractère suivant et/ou précédant, mais pour d'autres, ils ne sont jamais reliés. Certaines formes de caractères sont communes à plusieurs fontes [AMI 00] et [AZM 01]. Chaque caractère a de deux à quatre formes différentes qui dépendent de leur position dans le mot. L'écriture arabe est caractérisée par la présence de formes ambiguës. 16 alphabets arabes ont de un à trois points qui se positionnent au dessus ou au dessous du caractère. Ainsi ces points sont employés pour différencier les caractères semblables comme :  Ce qui pourra poser des problèmes, particulièrement pour l'identification des caractères et des fontes.

La plupart des caractères arabes ont des graphiques et des tailles variables. La taille d'un caractère change selon sa position dans le mot ; c'est un concept d'irrégularité des caractères arabes [ABD 98] [CHE 01].

La complexité morphologique et topologique de l'écriture arabe rend impossible la conception d'un système d'OCR omni fontes. Selon [KAH 87] un système

omni-fontes est capable d'identifier toute fonte, généralement sans apprentissage, basée sur les règles morphologiques et morphologiques de l'écriture.

L'utilisation d'un seul modèle de fonte pour la reconnaissance de caractères est une tâche relativement accessible. Cependant, dans un contexte multi fontes, il est nécessaire d'identifier initialement la fonte ou les fontes utilisées afin de concevoir un système d'OCR robuste. Notre objectif est d'analyser le texte à reconnaître puis d'identifier la fonte utilisée afin de la soumettre au système d'OCR le plus adéquat.

3 La méthode proposée

Notre méthode s'appuie sur l'analyse fractale d'un texte. La géométrie fractale est un ensemble mathématique de haut niveau qui permet de modéliser plusieurs phénomènes naturels que la géométrie Euclidienne classique ne le permet pas. Le concept d'auto similitude est au centre de la géométrie fractale [MAN 97]. Il rend possible de caractériser des objets ayant une forme très irrégulière par un changement d'échelle. La dimension fractale (DF) est un rapport de logarithme du nombre des objets similaires et de logarithme de la taille de l'objet primitif (logarithme du facteur de rapport optique). DF constitue un dispositif scalaire intéressant pour décrire une forme.

La valeur de DF dépend du nombre d'intervalle $N(r)$ de longueur 'r' permettant de couvrir un objet 'E', quelle que soit la manière dont ces intervalles sont choisis. On l'estime théoriquement par [JIN 95] :

$$N(r) \cdot r^{FD(E)} = 1 \quad (1)$$

$$\text{Alors } FD(E) = \frac{\log N(r)}{\log(\frac{1}{r})} \quad (2)$$

Nous avons proposé une méthode qui permet l'extraction, par deux techniques, des dimensions fractales, telles que la dimension de Hausdorff et la dimension de Minkowski-Bouligand (dimension d'auto similitude) [MEH 90] ; dans [CHE 91], ces deux dimensions ne sont pas forcément égales. Plusieurs travaux ont utilisé seulement une technique parmi les deux précités [FRE 00], [BOU 00] et [ZHU 01]. Dans ce travail, nous avons choisi d'étudier les deux techniques pour obtenir deux dimensions fractales. La première technique est celle de comptage par boîte ("box counting method") en faisant varier à chaque fois la taille de la boîte prise comme unité de mesure (c'est l'unité de mesure qui est affectée), afin d'extraire la dimension fractale notée B. Elle permet de décrire l'aspect de la répartition de la texture dans une image 2-D. La deuxième concerne la technique de la dilatation élémentaire appliquée sur la texture de l'image (c'est l'image qui est affectée) en changeant à chaque fois l'ordre de la dilatation, et ceci pour extraire la dimension fractale notée D. Elle permet de décrire l'aspect de vision à différente échelle d'observation. Ces deux

techniques sont expliquées en détailles dans [MOU 05].

La dimension B est appliquée pour n'importe quelle structure dans une surface. Elle est donnée par [ZEN 01] :

$$B(E) = \lim_{r \rightarrow 0} \frac{\log N(r)}{\log(\frac{1}{r})} \quad (3)$$

qui satisfait :

$$N(r) = A \cdot r^{-B(E)} \quad (4)$$

avec A un entier positif.

$B(E)$: la dimension B d'un ensemble E.

$N(r)$: le nombre minimum de boîtes nécessaires pour couvrir l'ensemble E.

r : la taille de la boîte utilisée.

Pour estimer la dimension B, l'espace Euclidien contenant l'image peut être divisé en grille des boîtes de taille 'r', en comptant le nombre des boîtes qui ne sont pas vides [MEH 90]. Ce processus doit être répété L fois, pour différentes tailles de 'r'. Nous avons fixé la valeur initiale de la taille de la boîte. Le pas d'avancement entre deux boîtes de mesures successives, à la valeur un. Alors L définit la taille de la boîte utilisée comme l'unité de mesure la plus grande. B dépend du choix de L. Dans ce travail, la dimension B est exprimée en fonction de plusieurs valeurs de L dont le but est d'étudier l'influence de ces limites de tailles sur les dimensions fractales B obtenues d'une part, et sur leurs impacts dans la discrimination des écritures d'autre part. Cette dimension fractale est caractérisée finalement par la taille maximum L de la boîte. Les paramètres extraits sont notés $SIZE_B(L)$.

La manière pratique d'obtenir la dimension fractale DF est décrite comme suit : un graphique de $\log N(r)$ par rapport à $\log(1/r)$ pour chaque objet 'r' doit être établi. Puis la corrélation entre $N(r)$ et $1/r$ est ajustée par régression linéaire ; DF est estimée par la pente obtenue de la ligne droite [ANG 04].

Nous avons noté que la limite de la taille maximale de boîte à maintenir dépend du niveau d'analyse (mot, ligne, bloc). Nous avons constaté que la dimension fractale d'un bloc est plus stable que celle d'une ligne. Cette dernière est plus stable que celle d'un mot. Ce concept de stabilité dépend non seulement de la taille limite de la boîte, mais également de la taille d'image.

La dimension D utilisée dans ce travail est obtenue par une mesure du volume de la texture pour chaque niveau de la dilatation morphologique de l'image 2-D. Pour chaque valeur de 'd', nous avons obtenu des fonctions dérivées, notée $SIZE_D$; ces dernières sont caractérisées finalement par un ordre maximum de dilatation, noté $SIZE_D(\text{rayon de dilatation})$. Ces fonctions dépendent de deux paramètres suivants : le rayon maximum de dilatation, et la valeur du pas d'avancement entre deux dilatations successives. $SIZE_D(d)$ est égale à [CET 91], [FRE 00] :

$$SIZE_D(d) = \lim_{r \rightarrow 0} \left(n - \frac{\log V(d)}{\log(d)} \right) \quad (5)$$

avec :

FD(E) Mink : la dimension de Minkowski pour un ensemble E,

n : la dimension de l'espace d'étude,

r : le rayon maximum de dilatation,

V : le corps de dilatation d'un ensemble E qui est défini comme suit :

$$V = \{ x \in \mathbb{R}^n : \exists a \in E / d(a, x) \leq r \} \quad (6)$$

d(a,x) : la distance entre le point a et x

V(r) : le volume n dimensionnel de l'ensemble V.

La dimension fractale SIZE_D est obtenue par la ligne linéaire approximative de l'uplet (le rayon de dilatation, volume correspondant de l'ensemble après dilatation).

4 Expérimentation : Classification des fontes Arabes

Dans [MOU 05], les auteurs ont montré la robustesse de leur méthode face aux facteurs, à savoir : la taille de l'écriture et l'inclinaison, pour la discrimination des fontes au niveau ligne. Dans ce travail, la base de test est composée d'images de texte (paragraphe), résultant d'une étape de segmentation. C'est pourquoi ces blocs peuvent avoir différentes tailles; ainsi la largeur (respectivement la hauteur) maximale d'un bloc est celle de la largeur (respectivement la hauteur) d'une page de format A4.

La base des fontes arabes est composée de 450 images textes, numérisées à 300 ppp, dans dix fontes : AH : Ahsa, AT : Arabic_transparent, BU : Buryidah, KA : El_Kharj, KO : Koufi, Na : Naskh, TO : Tholoth, DA : Dammam, Ha : Hada, et HI : Hijez, de diverses tailles. Un exemple est présenté dans la figure 1. L'originalité du choix de la base de données est exprimée par la reconstitution du même texte pour les différentes fontes choisies. Nous avons calculé les deux dimension B et D respectivement par les fonctions SIZE_B(L), SIZE_D(d), avec plusieurs alternatives de 'L' et 'd', comme suit : SIZE_B(15), SIZE_B(30), SIZE_B(100), SIZE_D(15), SIZE_D(20) et SIZE_D(30).

La figure 2 montre un exemple de localisation des zones de fontes, en utilisant SIZE_B(30) et SIZE_D(20). Mais quelques fontes telles que AH, AT et KA sont quasiment superposables. Cet inter localisation graphique peut être résolue par l'utilisation d'un simple classificateur.

L'utilisation d'un classificateur pour la discrimination des fontes permet la résolution du problème de chevauchement entre les zones de fontes. Nous avons utilisé le K-Plus Proche Voisin pour classifier les 10 fontes. Chaque prototype est exprimé par un vecteur de six éléments. Plusieurs valeurs de K sont expérimentées. Nous avons retenu pour K une valeur égale à un. La distance utilisée pour mesurer l'écart entre deux prototypes est celle de la distance Ecludienne.

Le meilleur taux de discrimination des fontes telles que : Ahsa, Arabic_transparent, Buryidah, El_Kharj, Koufi, Naskh, Tholoth, Dammam, Hada, et Hijez, est de l'ordre de 98%. La base d'apprentissage et de test sont composées respectivement de 200 et de 250 images. La matrice de confusion est présentée dans le tableau 1.



FIG. 1- Exemple des fontes de taille 14 points.

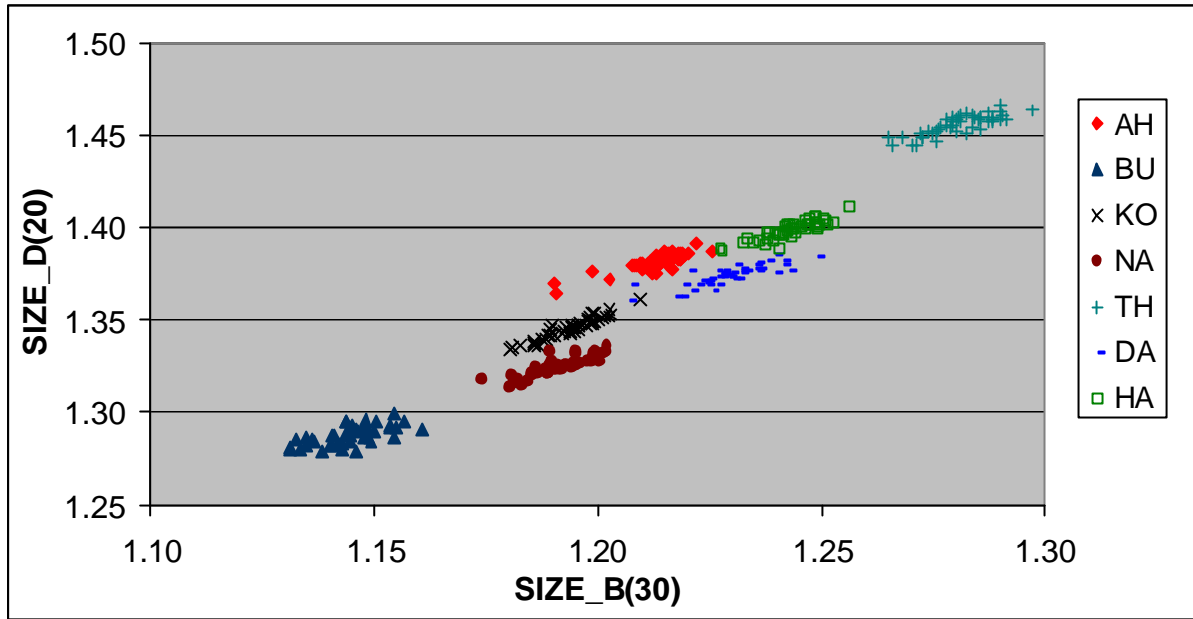


FIG. 2- Localisation graphique de quelques fontes en fonction de deux paramètres SIZE_B(30) et SIZE_D(20), AH : Ahsa, BU : Buryidah, KO : Koufi, NA : Naskh, TH : Tholoth, DA : Dammam, HA : Hada.

Fonte	AH	AT	BU	KA	KO	NA	TO	DA	HA	HI
AH	23	0	0	2	0	0	0	0	0	0
AT	0	25	0	0	0	0	0	0	0	0
BU	0	0	25	0	0	0	0	0	0	0
KA	1	0	0	22	0	0	0	0	0	2
KO	0	0	0	0	25	0	0	0	0	0
NA	0	0	0	0	0	25	0	0	0	0
TO	0	0	0	0	0	0	25	0	0	0
DA	0	0	0	0	0	0	0	25	0	0
HA	0	0	0	0	0	0	0	0	25	0
HI	0	0	0	0	0	0	0	0	0	25

TAB. 1- La matrice de confusion pour les dix fontes

Nous avons noté une bonne discrimination pour certaines fontes telles que : Arabic_transparent, Buryidah, Koufi, Naskh. Les erreurs de classification sont produites par une fusion entre les deux fontes telles que : le Ahsa et le Kharj. Une augmentation du nombre de prototypes dans la base d'apprentissage ne peut pas résoudre ce problème. Pour une base d'apprentissage et de test constituée respectivement par 350 et 100 images de texte, le taux de discrimination obtenu est de l'ordre de 95.14 %. Le taux de classification obtenu par une base d'apprentissage et de test, constitués respectivement 100, 350 images, est de l'ordre de 96 %. Ces résultats obtenus par augmentation/réduction du nombre de prototypes dans la base d'apprentissage, montre que notre méthode nécessite seulement quelques prototypes significatifs pour l'apprentissage. Un meilleur taux de classification n'est pas lié au nombre de prototypes dans la base d'apprentissage, mais il dépend

du choix du constituant d'apprentissage.

Pour évaluer l'aspect discriminatif de nos paramètres nous avons utilisé une technique neuronale d'auto organisation [VER 99]. Le réseau de Kohonen est composé de deux couches ; la première constitue les entrées (en utilisant les six paramètres pour chaque prototype) et la seconde constitue la sortie (le nombre de neurones dans la couche de sortie est égal à 30*30). L'activation des neurones dans la couche de sortie est assurée par un simple calcul de la distance Ecludienne [MEZ 02]. Cette technique nous a permis une observation de répartition des classes en deux dimensions. La carte de Kohonen obtenue est représentée par la figure 3 ; elle montre la puissance de nos paramètres pour la classification des fontes d'une part, et d'autre part l'origine de chevauchement entre les classes pour la détection des fontes similaires. Le nombre de classes sémantiques dans la figure 3 est supérieur au nombre des

classes réelles. L'erreur de quantization finale est égale à 0.005, et l'erreur de topographique finale est égale à 0.040. L'utilisation d'une autre technique pour la classification des fontes peut aboutir à l'amélioration de la performance de notre système.

5 Conclusions et perspectives

Une nouvelle méthode d'identification automatique des fontes est proposée et expérimentée. Cette méthode est basée sur l'analyse globale de textures. Elle ne dépend pas de la langue étudiée. Elle peut être utilisée pour d'autres langues. Les meilleurs taux de classification pour la majorité des fontes arabes sont de l'ordre de 100 %. Le taux moyen de discrimination des fontes au niveau bloc est de l'ordre de 98 %.

Les taux d'identification obtenus sont satisfaisants puisqu'ils sont obtenus avec des blocs ni normalisés ni traités, mais dans des conditions normales d'un document original, contrairement aux autres méthodes, qui ont utilisé plusieurs étapes de prétraitement, tels que les travaux de Avilés et al. [AVI 05], ou les travaux de Zhu et al. [ZHU 01]. Ainsi il n'existe pas des travaux de discrimination des fontes au niveau bloc ou paragraphe dans le cas de l'écriture arabe ; bien que plusieurs auteurs

aient accentué l'influence de l'ROF dans l'ROC arabe.

Néanmoins, il serait intéressant, d'une part de tester notre système sur d'autres langues ; et d'autre part d'affiner les paramètres utilisés pour améliorer leurs aspects distinctifs afin de réduire le nombre de dimensions de B et D, en étudiant d'avantage les alternatives issues de 'd', 'r' et 'L'.

Remerciements

Ce travail est soutenu par l'Agence Universitaire de la Francophonie, dans le cadre du projet "SAuverture du Patrimoine Culturel de Civilisation Ancienne Arabe" (SAPCCAA) du réseau CEMUR. Les auteurs tiennent à remercier la Direction Générale de la Recherche Scientifique et de la Renovation Technologique (DGRSRT) en Tunisie, pour son aide financière fournie par le programme ARUB/01/UR/11/02.

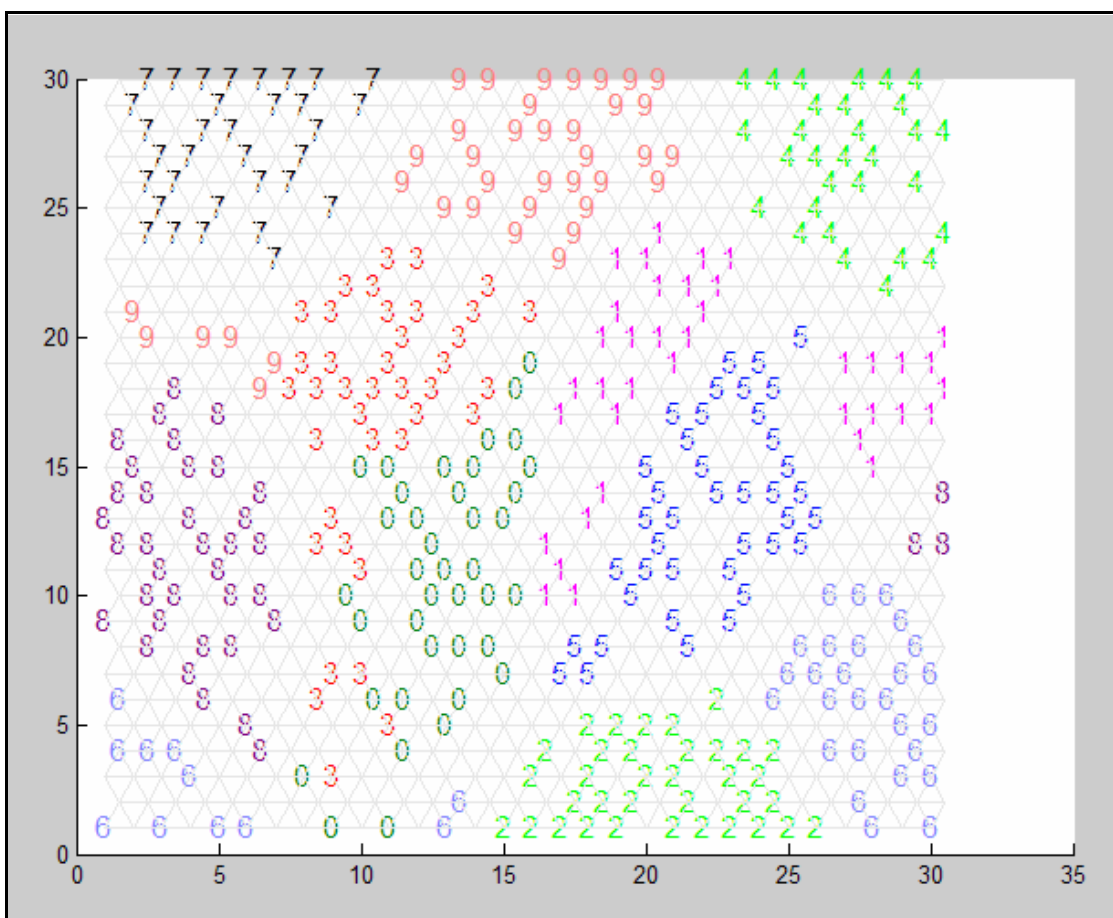


FIG. 3- Localisation graphique de dix fontes (0 : Ahsa, 1 : arabic_transparent, 2 : Buryidah, 3 : El_Kharj, 4 : Koufi, 5 : Naskh, 6 : Tholoth, 7 : Dammam, 8 : Hada, et 9 : Hijez), en fonction des six paramètres extraits.

Références

- [ABD 89] ABDELAZIM H.Y., HASHISH M.A., Interactive font learning for Arabic OCR, *Proc. 1st Kuwait computer conference*, Kuwait, 1989, pp. 463-486.
- [AMI 00] AMIN A., Recognition of printed arabic text based on global features and decision tree learning techniques, *Pergamon, Pattern Recognition Vol. 33*, 2000, pp 1309-1323.
- [ANG 04] ANGELES G. R., PERILO GERARDO M. E., Fractal analysis of tidal channels in the Bahia Blance estuary, *Geomorphology N° 57*, 2004, pp 263-274.
- [AVI 05] AVILES-CRUZ C., RANGEL-KUOPPA R., REYES-AYALA M., ANDRADE-GONZALEZ A., ESCARELA-PEREZ R., High-order statistical texture analysis font recognition applied, *Pattern Recognition Letters Vol. 26*, 2005, pp 135-145.
- [AZM 01] AZMI R., KABIR E., A new segmentation technique for omnifont Farsi text, *Pattern Recognition Letters, Vol. 22*, 2001.
- [BOU 00] BOULÉTREAU V., VINCENT N., EMPTOZ H. AND SABOURIN R., How to use fractal dimension to qualify writings and writers, *Fractals, complex geometry, Paterns and scaling in nature and society, Vol. 8, N° 1*, March 2000, pp. 85-98.
- [CET 01] CETERA A., The Minkowski dimension and critical effects in fractal evolution of defects, *Chaos, Solitons and Fractals Vol. 12, Pergamon*, 2001, pp 475-482.
- [CHE 01] CHEUNG A., BENNAMOUN M., An Arabic optical character recognition system using recognition-based segmentation, *Pattern Recognition, vol. N° 34, Bergmann*, 2001, pp 215-233.
- [CHE 91] CHERBIT G., Fractals : Dimensions non entières et applications, Masson, France, Paris, 2^{ème} édition, 1991.
- [FRE 00] FRECHET., VINCENT N., Mesure du contraste par analyse fractale : du texte en niveau de gris, à la ligne, 2^{ème} Colloque International Francophone sur l'écrit et le document, CIFED' 00, Lyon, Paris, 03-05 Juillet 2000, pp 91-100.
- [JIN 95] JIN X.-C., A pratical method for estimating fractal dimension, *Pattern Recognition Letters, Vol. 16*, May 1995, pp. 457-464.
- [KAH 87] KAHAN S., PAVLIDIS T., On the Recognition of Printed Characters of any Font or Size, *IEEE Trans. PAMI*, Vol. 9, No. 2, March 1987.
- [KAN 02] KANOUN S., ALIMI A., ENNAJI A., LECOURTIER Y., Reconnaissance de mots arabes par approche affixale, 3^{ème} Colloque International Francophone sur l'écrit et le document, CIFED' 02, Hammamet, Tunisie, 21-23 Octobre 2002, pp 21-30.
- [KHO 02] KHORSHEED M. S., Off-Line Arabic Character Recognition – A Review, *Pattern Analysis & Applications Springer-Verlag, Vol. 5*, pp 31-45, 2002.
- [MAN 97] MANDELBROT B., Fractales, Hasard Et Finance, Flammarion, France, Paris, 1997.
- [MEH 90] LE MEHAUTE A., Les géométries fractals, Edition HERMES, N° imp. 002023, Paris, France, 1990.
- [MEZ 02] MEZGHANI N., MITICHE A., AND CHERIET M., Reconnaissance en ligne de caractères arabes manuscrits par un réseau de Kohonen, *15th International Conférence on Vision Interface, Calgary*, May 2002, pp.186-190.
- [MOU 05] BEN MOUSSA S., ZAHOUR A., ALIMI M. A., BENABDELHAFID A., Can Fractal Dimension Be Used On Font Classification, *8th International Conference in Document Analysis and Recognition, Seoul, Korea, Vol. 1*, August 2005, pp 146-150.
- [VER 99] VERSANTO J., HIMBERG J., ALHONIEMI E., PARHANKANGAS J., Self organizing map in matlab: the somtoolbox, *In Proceeding of the Matlab DSP Conference 1999, Finland*, November 1999, pp 35-40.
- [ZEN 01] ZENGL X., KOEHL L., VASSEUR C., Design and implementation of an estimator of fractal dimensions using fuzzy technique, *Pattern Recognition Vol. 34, Pergamon*, 2001, pp 151-169.
- [ZHU 01] ZHU Y., TAN T., YUNHONG W., Font recognition based on global texture analysis, *IEEE, Trans on pattern recognition and machine intelligent, Vol. 23, N° 10*, October 2001, pp 1192-1200.