

TRAITEMENT MORPHOLOGIQUE DE SIGNAUX UNIDIMENSIONNELS

B. ZOUABI, N. ELLOUZE, ET A. BEN SLIMANE

LSTS - ENIT - TUNISIE

Résumé - Cette communication présente les premiers résultats des travaux relatifs à une nouvelle approche de traitement de signal unidimensionnel basée sur les techniques de la morphologie mathématique qui sont utilisées en traitement d'images. Après avoir défini les opérations morphologiques de base sur les signaux temporels et présenté leur application à la détection du fondamental du signal de parole. Nous présentons une série de transformations morphologiques dites adaptatives et particulièrement la transformation par les extrêmes. Nous introduisons pour terminer, la notion de filtrage morphologique des signaux temporels en donnant des exemples d'application liés au lissage des signaux bruités, à la détection de formes d'ondes particulières pour l'estimation de la puissance instantanée du signal de parole.

I. INTRODUCTION

La morphologie mathématique a pris naissance en 1964 sous l'impulsion de G. Matheron et J. Serra [1] [2] [3]. C'est une théorie tout d'abord ensembliste, car elle s'appuie sur les opérations de base de la théorie des ensembles. Elle emploie ensuite les transformations en tout ou rien par éléments structurants. Elle a été appliquée aux images binaires dans l'analyse des structures et l'extraction d'objets, elle s'est étendue ensuite aux images à niveau de gris particulièrement grâce au développement de la théorie du filtrage morphologique [4] [5] [6].

L'idée fondamentale de notre approche est d'adapter les concepts de la morphologie mathématique aux fonctions définies dans un espace unidimensionnel et particulièrement aux signaux temporels. Cette approche de réduction de l'espace qui semble à priori relativement simple sur le plan de sa mise en œuvre, n'en est pas moins hardue sur le plan de l'interprétation des résultats. En effet le traitement morphologique sur les images constitue une approche naturelle, ce qui l'est moins sur les signaux temporels. Bien évidemment le traitement morphologique des signaux temporels est une étude de la structure du signal et de ses variations en fonction du temps. Des méthodes ont été élaborées tant sur le plan du filtrage morphologique que sur le plan de la caractérisation de certaines formes du signal. Le signal de parole, particulièrement riche en applications, présente des caractéristiques

Summary - This paper presents the initial results which concern a new approach to unidimensional signal processing, based on mathematical morphology techniques used in signal processing. After defining the basic morphological operations on temporal signals and discussing their use in the pitch detection, we will present a series of so-called adaptive morphological transformations, particularly transformation through extrema. In conclusion we will introduce the idea of morphological filtering of temporal signals through the presentation of concrete examples dealing with the smoothing of noisy signals and the detection of particular wave-forms for the estimation of the instant power of speech signal.

propres qui s'adaptent à ce type de traitement. Il est toutefois possible d'utiliser ces méthodes sur d'autres types de signaux tels que les signaux radars, sonars, etc...

Après avoir défini les opérateurs morphologiques de base à savoir l'érosion et la dilatation sur les signaux temporels, nous présentons une application utilisant ces opérateurs et qui est liée à la détermination du fondamental du signal de parole.

De nouvelles transformations morphologiques sont introduites en particulier la transformation par les extrêmes qui trouve son application dans la mise en forme des signaux.

Après avoir présenté les transformations par ouverture et par fermeture, nous aborderons le problème du filtrage morphologique en donnant quelques exemples d'application à savoir: le lissage, la détection et la paramétrisation des signaux.

II. LES TRANSFORMATIONS PAR LE MINIMUM ET PAR LE MAXIMUM OU EROSION ET DILATATION

Soit un signal échantillonné représenté par x_k ; $k=1, \dots, N$. Définissons une fenêtre temporelle de largeur $P=L+M$ ayant son origine placée au point $j=L+1$. On fait déplacer cette fenêtre sur le signal au pas de l'échantillonnage. Dans chaque position on affecte au point j la valeur minimale ou maximale trouvée dans la fenêtre. L'ensemble des valeurs relatives aux points j et auxquelles on aurait rajouté $P-1$ valeurs nulles placées aux



extrémités, forme un nouveau signal y_k qu'on appelle **signal transformé**.

Cette démarche conduit à deux transformations différentes, une transformation par le minimum et une transformation par le maximum, qui ne sont autres que les transformations morphologiques de base à savoir respectivement l'**érosion** et la **dilatation** exprimées dans un espace unidimensionnel. La fenêtre qui a été définie correspond à l'élément structurant et la largeur $P=L+M$ définit la taille de la transformation.

- Le signal érodé issu de la transformation par le minimum est entièrement défini par l'expression :

$$y_k = \min [x_{k-L}, \dots, x_{k+M}] ; \quad k = L, \dots, N-M$$

- Le signal dilaté issu de la transformation par le maximum est entièrement défini par l'expression :

$$y_k = \max [x_{k-L}, \dots, x_{k+M}] ; \quad k = L, \dots, N-M$$

La fenêtre utilisée dans cette définition est de type uniforme. Il est possible d'introduire l'application d'une fonction de voisinage sur le signal dans la fenêtre avant la transformation morphologique. Cette application peut être additive ou multiplicative. Ainsi :

- Dans le cas additif, le signal transformé s'exprime par :

$$y_k = \min [x_{k-L} + V_{-L}, \dots, x_{k+M} + V_M],$$

pour l'érosion et

$$y_k = \max [x_{k-L} + V_{-L}, \dots, x_{k+M} + V_M],$$

pour la dilatation, où V_{-L}, \dots, V_M sont les coefficients additifs. Ceci permet d'envisager des éléments structurants autres que linéaires [8].

- Le signal transformé dans le cas multiplicatif est :

$$y_k = \min [\beta_{-L} x_{k-L}, \dots, \beta_M x_{k+M}],$$

pour l'érosion et

$$y_k = \max [\beta_{-L} x_{k-L}, \dots, \beta_M x_{k+M}],$$

pour la dilatation, où $\beta_{-L}, \dots, \beta_M$ représentent les coefficients de pondération, ils permettent d'envisager un type de convolution dans lequel la somme est remplacée par l'opération minimum ou maximum, donc morphologique.

Pour simplifier les représentations, on adoptera selon le cas les notations suivantes:

$$y_k = E_P[x_k] \quad \text{ou} \quad y_k = E_{L,M}[x_k],$$

pour l'érosion et

$$y_k = D_P[x_k] \quad \text{ou} \quad y_k = D_{L,M}[x_k]$$

pour la dilatation.

- **Remarque :** Selon la position de l'origine dans la fenêtre, le signal x_k peut être érodé ou dilaté d'une manière symétrique lorsque $L = M$ ou d'une manière asymétrique lorsque $L \neq M$. Il est à noter toutefois que quelque soit la manière dont le signal x_k a été transformé, le signal résultant est le même à un retard près, affecté par les échantillons nuls. La figure (1) donne des exemples de transformations symétriques et asymétriques opérées sur un signal échantillonné avec une fenêtre de largeur égale à dix pas d'échantillonage.

D'après les cas de figure présentés, on remarque que l'érosion réduit les maxima du signal et affecte un palier aux valeurs minimales. Dans le cas de la dilatation, il se produit la situation inverse, un palier se crée autour des maxima et une réduction s'opère aux niveaux des minima. La largeur des paliers correspond à la taille de la transformation.

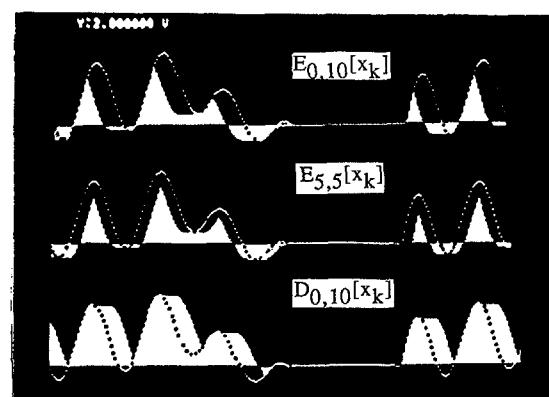


Fig. 1. Transformations par érosion et dilatation.

- Il est possible d'obtenir par érosion, des paliers représentant toutes les valeurs minimales d'un signal si la taille de l'érosion est choisie inférieure ou égale à la plus petite distance séparant deux minima relatifs.

- Il est possible d'obtenir par dilatation, des paliers représentant toutes les valeurs maximales d'un signal si la taille de la dilatation est choisie inférieure ou égale à la plus petite distance séparant deux maxima relatifs.

- **Propriété d'itérativité :** L'érosion et la dilatation sont des transformations itératives. En effet :

$$E_P [x_k] = E_1 [E_1 \dots [x_k]] \quad P \text{ fois}$$

$$D_P [x_k] = D_1 [D_1 \dots [x_k]] \quad P \text{ fois}$$

On peut déduire de ces expressions que la transformation par érosion ou dilatation de taille égale à P peut être obtenue en appliquant successivement P fois la transformation de taille unitaire.

III. APPLICATION DE L'EROSION ET DE LA DILATATION A LA DETERMINATION DU PITCH

Dans certaines applications de traitement du signal de parole, la fréquence du fondamental ou "pitch" joue un rôle important notamment dans les applications de synthèse, de codage et de segmentation. Cependant, la détermination du pitch est une opération délicate à cause de la variabilité des formes d'onde du signal vocal [9]. Il existe plusieurs méthodes, aussi bien temporelles que fréquentielles qui s'intéressent à la détermination du pitch [10],[11]. On peut citer les méthodes directes (passage par zéro, codage des extrema)[12] dont l'élaboration est simple mais qui nécessitent des systèmes de décision complexes et les méthodes indirectes (Cepstre, codage prédictif, filtrage inverse)[13],[14] qui donnent une bonne estimation de la valeur du pitch mais elles sont coûteuses en temps de calcul.

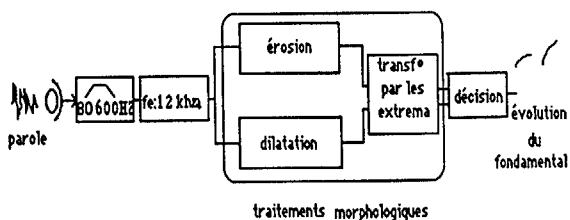


Fig. 2. DéTECTEUR morphologique du pitch

Nous présentons une nouvelle méthode qu'on classe parmi les méthodes temporelles. Sa particularité réside dans l'application des techniques de la morphologie mathématique sur le signal de parole dans le but de déterminer l'évolution du pitch.

La méthode proposée dont le principe est illustré par la figure (2) consiste à effectuer les traitements suivants :

- On filtre le signal de parole dans la bande de fréquence (80...600) hz. Cette bande est choisie pour couvrir aussi bien le pitch relatif à une voix masculine, féminine et même celle de l'enfant.

- Le signal filtré est échantillonné à la

fréquence de 12 khz pour avoir une précision suffisante de la valeur instantanée du pitch.

- On opère de façon parallèle sur le signal échantillonné, les transformations par érosion et dilatation. La taille de ces transformations est choisie en fonction du type du locuteur (homme, femme ou enfant) de façon à éroder ou dilater les oscillations harmoniques se trouvant dans la période du pitch (fig.3).

- On applique sur les signaux érodés et dilatés, une transformation par les extrema qui sera développée au paragraphe suivant et qui consiste à garder les valeurs minimales dans les phases de croissance et les valeurs maximales dans les phases de décroissance. Pour déterminer la valeur instantanée du pitch, il suffit de mesurer l'intervalle de temps séparant deux fronts descendants du signal multiniveaux relatif à l'érosion et l'intervalle séparant deux fronts montants du signal multiniveaux relatif à la dilatation. Ainsi sont détectées deux estimations de la valeur instantanée du pitch.

- Un système de décision permet de comparer les deux estimations issues des deux transformations à un instant donné avec les valeurs les plus récentes. La valeur la plus proche est déclarée comme étant la période du pitch à cet instant.

Cette procédure permet entre autre de déterminer si le son est voisé ou non voisé. En effet dans le cas d'un son non voisé, il apparaît une incohérence évidente entre les estimations.

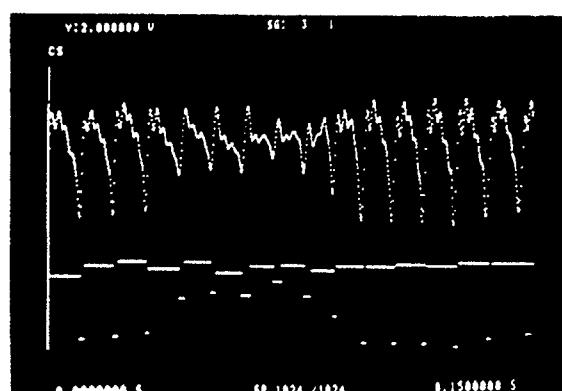


Fig. 3. DéTECTION morphologique du pitch.

IV. TRANSFORMATIONS MORPHOLOGIQUES ADAPTATIVES

Les transformations par érosion et par dilatation développées jusqu'alors sur un signal unidimensionnel s'opèrent d'une manière systématique quelque soit la taille de l'élément structurant et indépendamment de la morphologie du signal. Il en résulte que le signal est affecté dans sa globalité et particulièrement aux niveaux des extrema. Nous introduisons une série de transformations que nous appellons adaptatives, qui consistent à adapter les



opérations d'érosion et de dilatation en fonction du résultat recherché. Nous retrouvons une approche similaire pour la squelettisation des objets par la méthode de la bissectrice conditionnelle [4].

Pour effectuer ces transformations, on se fixe un certain nombre de paramètres liés à la morphologie du signal à savoir : la croissance, la décroissance, les maxima et minima, les passages par zéro ou par un seuil, les points stationnaires, etc... On sélectionne ensuite les paramètres représentant les caractéristiques du signal qu'on désire préserver. On doit choisir enfin comme élément structurant une fenêtre temporelle adéquate dans laquelle on suit l'évolution du signal. La transformation est adaptée selon les paramètres sélectionnés et le résultat recherché. Le paragraphe suivant donne un exemple d'application qui illustre ces transformations.

- Transformation par les extrema

Pour réaliser cette transformation, on choisit une fenêtre symétrique définie par $L = 1$ et $M = 1$ qu'on déplace le long du signal. Pour chaque position on affecte à l'échantillon se trouvant au milieu de la fenêtre la valeur du premier minimum ou maximum rencontré. Cette valeur est gardée jusqu'au prochain extrémum. Le signal résultant de cette transformation est un signal multiniveaux dont les paliers correspondent à tous les extrema du signal (fig.4).

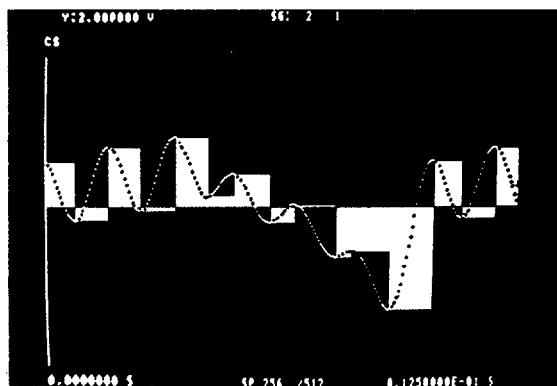


Fig. 4. Transformation par les extrema.

La transformation par les extrema peut être aussi réalisée par une suite d'opérations d'érosion et de dilatation opérées sur le signal jusqu'à l'idempotence. En effet, appelons T_E^1 la transformation par les extrema relative à une seule itération. Le signal résultant de n itérations est donné par l'expression :

$$x_k^n = T_E^n[x]$$

tel que: $x_k^n = T_E^1 \circ T_E^{n-1}[x]$; relation qui traduit

l'idempotence. L'algorithme associé à une itération est le suivant :

- Si x_k est croissant :

$$T_E[x_k] = E_{1,1}[x_k]$$

- Si x_k est décroissant :

$$T_E[x_k] = D_{1,1}[x_k]$$

- Si x_k est un extrema ou un point stationnaire :

$$T_E[x_k] = x_k$$

Les transformations adaptatives peuvent s'étendre, moyennant une fenêtre adéquate, aux points de passage par zéro ou par un seuil, aux points stationnaires etc... Elles trouvent naturellement leurs applications dans la mise en forme, le codage et la détection des signaux.

V. TRANSFORMATIONS PAR OUVERTURE ET FERMETURE.

Il est possible d'opérer sur un signal x_k la transformation qui consiste à effectuer une érosion du signal suivie d'une dilatation, dans ce cas en parlera de transformation par ouverture, ou à lui appliquer une dilatation suivie d'une érosion auquel cas on parlera d'une transformation par fermeture. La figure (5)

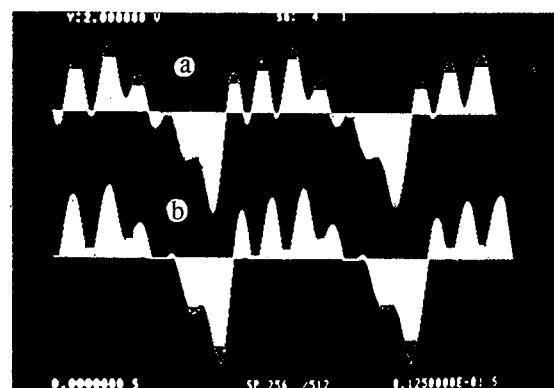


Fig. 5. Transformations par ouverture et fermeture.

donne un exemple de ces transformations. Nous remarquons que l'ouverture et la fermeture stabilisent les portions du signal situées autour des extrema, indépendamment de leur amplitude, et dont la largeur correspond à la taille de ces transformations. Ces opérateurs permettent d'envisager des applications importantes qui seront mentionnées par la suite notamment en matière de filtrage morphologique.

- Propriété d'idempotence : L'ouverture et la fermeture ne sont pas des transformations itératives, elles sont par contre idempotentes ; en

effet :

$$O_p [O_p [x_k]] = O_p [x_k]$$

$$F_p [F_p [x_k]] = F_p [x_k]$$

L'idempotence de l'ouverture et de la fermeture est une propriété fondamentale des filtres morphologiques.

VI. FILTRAGE MORPHOLOGIQUE DES SIGNAUX UNIDIMENSIONNELS.

Nous venons de voir que l'ouverture et la fermeture ne modifient le signal que dans les régions des extrema, ainsi toutes les variations du signal qui correspondent à la taille de ces transformations sont lissées. Il en résulte que le signal transformé, compte tenu de la taille choisie, présente une structure plus régulière que le signal initial. Ces transformations filtrent bien certaines caractéristiques du signal et sont à la base du filtrage morphologique.

Il est possible d'autre part d'utiliser une suite d'opérations d'ouverture et de fermeture de façon à réaliser des filtres medians tels que : les filtres **OFO**, **FOF** etc..., qui ont des propriétés remarquables dans le filtrage de formes d'ondes particulières d'un signal.

D'autres transformations permettent d'extraire les contours lissés par l'ouverture et la fermeture. En effet, la transformation :

$$T [x_k] = x_k - O_p [x_k]$$

permet d'extraire les régions des maxima (fig. 6-a) et la transformation :

$$T [x_k] = F_p [x_k] - x_k$$

permet d'extraire les régions des minima (Fig. 6-b). Toutefois ces transformations ne présentent pas les propriétés des filtres morphologiques [17].

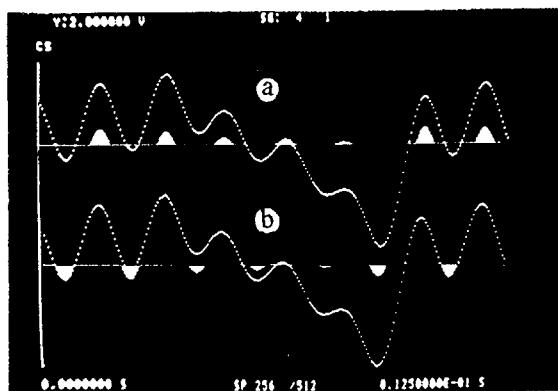


Fig. 6. Détection des régions des extrêmes par ouverture et fermeture.

Le filtrage morphologique a été développé dans le cadre du traitement d'images à niveau de gris [4], [5], [6], [7]. Les premiers résultats que nous avons obtenus montrent que le filtrage morphologique s'adapte bien aux signaux temporels et peut être à la base de diverses applications :

- Soit dans le sens du lissage des signaux bruités avec une amélioration notable du rapport signal sur bruit.

- Soit dans le sens de la détection et de l'extraction de certaines formes d'onde, ce qui permet entre autre de fournir une estimation de la puissance instantanée du signal.

- Soit dans la paramétrisation des signaux, ce qui permet de faire ressortir certaines caractéristiques pertinentes ouvrant des possibilités à l'analyse quantitative.

Les figures suivantes donnent quelques exemples d'illustration. Ainsi :

La figure (7) montre le lissage d'un signal bruité en utilisant un filtre median de type FO. Les tailles de la fermeture et de l'ouverture sont choisies en fonction de l'occupation spectrale du bruit.

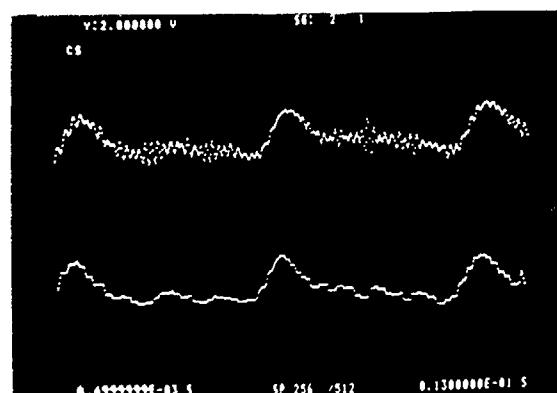


Fig. 7. Lissage morphologique d'un signal bruité.

La figure (8) montre la précision du filtrage morphologique dans la détection d'une onde sinusoïdale cachée dans un train d'ondes de même type, mais dont la période diffère de 1%. L'opération utilisée est une ouverture de taille égale à la période des ondes voisines.

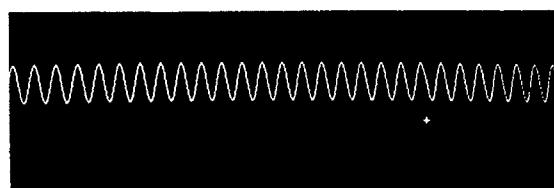


Fig. 8. Détection morphologique d'une forme d'onde.

Enfin la figure (9) donne le résultat d'une estimation de la puissance instantanée d'un



signal de parole (fig. 9-a) obtenue à partir d'un détecteur morphologique multicanal que nous avons réalisé. La courbe obtenue (fig. 9-b), après lissage morphologique, correspond à une sommation toutes les 10 msec des amplitudes relatives aux extréma du signal recueillies dans le premier canal. Il est à remarquer que l'examen des deux cas de figures, correspondantes à un échantillon de parole arabe, doit s'effectuer de droite à gauche. Les autres canaux serviront à la caractérisation de certaines formes d'onde du signal de parole.

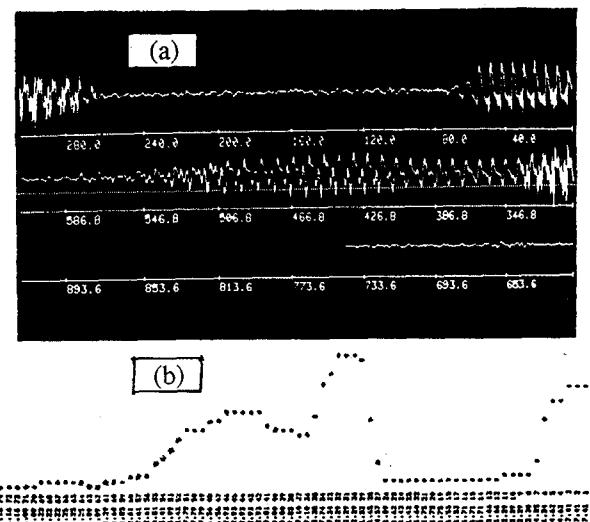


Fig. 9. Estimation de la puissance instantanée par détection morphologique, (a) Signal de parole, (b) puissance instantanée.

VII. CONCLUSION

L'originalité de l'approche morphologique, réside dans le choix de techniques, dont la mise en œuvre est facile, ce qui permet d'élaborer des traitements en temps réel, particulièrement au vue du développement de processeurs spécialisés. La contribution que nous avons apportée dans l'application de la morphologie mathématique aux signaux unidimensionnels, particulièrement au signal de parole, est le fruit d'un travail continu depuis quelques années [15], [16], suite à un premier contact avec le Professeur J. SERRA. Notre apport ne constitue qu'une première approche à ce type de traitement qui mérite de toute évidence un intérêt certain, particulièrement en matière de filtrage.

REFERENCES

- [1] G. Matheron, "Les variables regionalisées et leur estimation", Masson, 1965.
- [2] G. Matheron, "Elements pour une théorie des milieux poreux", Masson, 1967.
- [3] J. Serra, "Introduction à la morphologie mathématique" Cahiers du Centre de Morphologie Mathématique, Ecole des Mines, Fontainebleau, N° 3, 1969.
- [4] F. Meyer, Thèse de Docteur Ingénieur, Ecole des Mines de Paris, 1979.
- [5] J. Serra, "Image analysis and mathematical morphology," Academic Press, 1982.
- [6] G. Matheron, "Les applications idempotentes," Rapport du CGMM, Fontainebleau, 1982.
- [7] S. Beucher, "Extrema of gray tone functions and mathematical morphology," Ecole des Mines, Fontainebleau, 1983.
- [8] M. Coster, J. Chermant, "Précis d'analyse d'images," CNRS, Paris, 1985.
- [9] J.S. Liénard, "Les processus de la communication parlée," Masson, 1977.
- [10] M. J. Cheng, "A Comparative Performance Study of Several Pitch Detection Algorithms," M. S. Thesis, Mass. Inst. Technol., Cambridge, June 1975.
- [11] N. Ellouze, B. Zouabi, "Détecteur Electronique de Pitch", 3^e JTEA, Tunisie, 1982.
- [12] B. Gold, L. R. Rabiner, "Parallel Processing Techniques for Estimating Pitch Periods of speech in the time Domain," J.A.S.A, vol. 46, 1969.
- [13] A. M. Noll, "Capstrum pitch determination", J.A.S.A, Vol 41, 1967.
- [14] J. D. Markel, "The SIFT Algorithm for Fundamental Frequency Estimation," IEEE, Audio Electroacoust, vol. AU-22-, 1972.
- [15] A. Ben Slimane, "Traitement du signal de parole par les méthodes de morphologie mathématique", Rapport ENIT, Tunisie, 1985.
- [16] A. Ben Slimane, "Rapport de D E A". ENIT, Tunisie, 1986.
- [17] J. Serra, S. R. Sternberg, "Summer School on Mathematical Morphology" Fontainebleau, 1981.
- [18] N. Ellouze, F. Vernieres, F. Castanie, "Telegraph Signal Based Uniformly Distributed Analog Random Voltage Generator," IEEE, vol. 69, No.6, 1981.
- [19] N. Ellouze, B. Zouabi, M. Folsched "Modulation Delta Synchrone", 3^e JTEA, Tunisie, 1982.
- [20] B. Zouabi, A. Gacem, "Introduction à la Reconnaissance de Parole Arabe", 4^e JTEA, Tunisie, 1983
- [21] B. Zouabi, N. Rehif, "Système de Reconnaissance Globale de Mots Isolés", Arab School on Sciences and Technology, Maroc, 1983.