

An International Publication of Engineering Science



AJOT

ALGERIAN JOURNAL OF TECHNOLOGY

**E l e c t r o n i c s
C o m p u t e r E n g i n e e r i n g
E l e c t r i c a l E n g i n e e r i n g
C o n t r o l E n g i n e e r i n g**

1999

Series B Volume 14 Number 1

ISSN 1111 - 357 X

ALGERIAN JOURNAL OF TECHNOLOGY

Serie B

Directeur de la Publication	Pr. M.K.Berrah
Directeur de la Rédaction	Pr.A.Zerguerras
Diffusion et Abonnement	C. Larbes
Publicité et Sponsoring	M.Bouabdellah

Comité Scientifique

M.Attari	USTHB , Alger , Algerie.	M.M.Guessoum	Université de Blida ,Algerie
Y.Baghzouz	University of Nevada, Las Vegas , USA	S.Leulmi	Université de skikda ,Algerie
M.Barkat	Université de constantine,Algerie	N.K.M'Sirdi	Université de Versailles, France
H.Baudran	INP , Toulouse, France	L.Refoufi	INELEC ,Université de Boumerdes,Algerie
M.F.Benkhoris	LARGE CRTT , Saint Nazaire, France	M.A Rezzoug	Université de Nancy, France
M.Benmohamed	INELEC ,Université de Boumerdes,Algerie	R.Toumi	USTHB , Alger , Algerie
A.Beroual	ECL , Lyon , France	A.M Sharaf	University of new Branwick , Canada
H.Bourdoucen	Université Sultane Qaboos, Oman	Vander Voste	University Catholique Louvain, Belgique
G.Dubost	Université de Rennes 1 , France	M.E.H.Zaim	LARGE CRTT ,Saint Nazaire, France
I.M.Elamin	KFUPM ,Dahran,Arabie Saoudite		

CORPS DES ENSEIGNANTS DE RANG MAGISTRAL DE L.E.N.P

En Electronique et Génie Electrique

Professeurs

Boubakeur	Ahmed
Boudjema	Farès
Berkani	Daoud
Farah	Ahcène
Ibtiouen	Rachid
Touhami	Omar
Louam	Nadjib
Zerguerras	Ahmed

Maîtres de conférence

Aksas	Rabia
Boucherit	Med seghir
Berkouk	El madjid
Derras	Belkacem
Guerti	M'hania
Haddadi	mourad
Hellal	Abdelhafid
Mehenni	Mohamed
Mahmoudi	Med Oulhadj
Mekhaldi	Abdelouahab

Le Comité de lecture peut consulter tout spécialiste reconnu pour un domaine donné , dont le nom ne figure pas dans cette liste.

Avant – propos

Dans les volumes 12 et 13 de l'A.J.O.T. Le thème de l'avant propos portait sur la recherche scientifique et sa contribution au développement, ainsi que sur les publications qui permettent de diffuser ses résultats. Celui-ci est centré sur l'évaluation de la recherche dans un contexte de mondialisation tout en restant cohérent et ouvert à tout enrichissement ultérieur.

La globalisation du marché progresse aux dépens des prérogatives nationales. Elle se réalisera soit sous notre conduite réfléchie avec nos alternatives et nos choix, soit sous notre participation passive sans en avoir compris les motivations réelles. Un peuple mature, jouissant des libertés d'une démocratie authentique, est mieux préparé pour sélectionner les dispositions de cette mondialisation qui lui conviennent et de rejeter ou négocier les autres. Pour ne pas perdre et rester dans les limites des avantages comparatifs, une formation continue de la population active s'avère nécessaire et en particulier en direction d'une « conscience collective » et d'une « intelligence économique ». Ce type de formation est déjà mis en œuvre dans de nombreux pays occidentaux et aussi tel que l'Inde pour éviter toute déviation de la démocratie acquise. L'objet de cette formation est que chaque individu doit être apte à se gérer en connaissance de cause, de discerner ce qui lui est bon de ce qui ne l'est pas, de fonder ses décisions sur des éléments fiables en temps réel.

Dans le cas des groupes, il y a en plus l'échange d'informations et le bénéfice du débat. Pour intégrer la mondialisation sans mal, la démarche scientifique à suivre couvre plusieurs étapes la première est celle de la connaissance, de l'observation, de l'identification, de la classification, de la mise en ordre, de l'analyse de ces données en concomitance avec les actions nécessaires à leur obtention, leur accumulation, leur classement et leur traitement.

L'étape suivante passe de l'analyse qualitative à l'établissement de relations causales quantifiantes entre les grandeurs des différents facteurs entrant en jeu. Cette formation des lois doit préciser la qualité de l'inférence causale ainsi produite, et prendre en compte tous les facteurs essentiels, nécessaires à une bonne prédiction des effets attendus. Les contributions d'ordre supérieur seront prises en charge dans un modèle plus affiné, permettant une simulation plus proche de la réalité, avec possibilité de synthèse de tout système qu'il soit compliqué, linéaire ou non linéaire, mais néanmoins commandable.

L'application de ces résultats, au de là de l'aide à la prise de décision, aussi bien dans un cadre micro-économique que macro-économique, répond à un objectif final qui peut être global, ou le simple maintien des paramètres présentés par les différents indicateurs, courbes du 'tableau de bord', ... dans la plage d'intérêt ; puisqu'il est désormais question de gouvernement d'Entreprise ou 'Corporate governance'.

Evidemment en sciences humaines la diversité considérable existante entre les individus, ou les groupes sociaux fait que les comportements ne sont pas prévisibles de façon exacte et univoque. Le but ultime reste alors une modélisation statistique à partir d'une connaissance même partielle des paramètres caractéristiques du sujet étudié. Ce type de modélisation doit être suffisant pour prédire correctement le comportement d'individus isolés ou en groupes en réponse à des stimulus choisis. Cet objectif de modélisation ne fonctionnera de façon satisfaisante que si les données, statistiques en particulier, sont assez fiables et disponibles en temps opportun.

L'évaluation en général de la démarche et de la précision des résultats d'une recherche dans la durée, devrait se faire sans erreur de jugement, surtout lorsque l'on sait que des périodiques de

recherche, de grande renommée se sont laissés tromper plus d'une fois, dans la sélection des articles à publier. Même l'Internet qui constitue une facilité de communication n'est pas à l'abri d'informations erronées issues de la propagande, de l'intoxe, de la publicité... Le tri n'est pas aisé.

Alors quel esprit critique faut-il adopter pour évaluer les résultats d'une recherche ?

Un esprit critique est une attitude positive, de libre examen, orientée vers l'objectivité, de façon méthodique, mathématique dans l'observation, l'interprétation et l'évaluation des travaux d'une recherche et de ses résultats, à l'opposé des critiques fondées seulement sur des a priori contre-productives qui n'apportent rien à la démarche scientifique. Dans ce dernier cas de l'attitude négative, subjective, l'évaluation d'une publication de recherche se fait souvent seulement à l'aune de la célébrité de la revue qui en constitue le support, sans tenter d'apprécier l'intérêt du travail et la qualité de ses résultats.

L'invalidité ou la mise en doute d'une théorie peut se présenter quand elle ne peut expliquer l'existence d'une relation relevant de son domaine d'application. Cependant, une théorie n'est pas nécessairement à rejeter par ce qu'elle n'explique pas tous les faits, ou que certains lui soient contradictoires, sauf si une meilleure théorie est proposée. Celle-ci pouvant se fonder sur des méthodes analytiques ou numériques plus appropriées à la nature du problème à traiter et /ou étant les plus efficaces du moment. Toute théorie ne s'avère valable, ou ne donne des précisions acceptables que dans un domaine bien défini de chacune de ses variables.

Une recherche expérimentale soignée, se base sur des protocoles qui ne conduisent à aucune ambiguïté dans l'évaluation des résultats. La validité, interne et externe d'une expérience, est nécessaire pour pouvoir procéder à une généralisation induite par l'authenticité des faits qu'elle démontre.

Vis à vis de l'objet de ses travaux, la conviction du chercheur n'est pas sans importance. Un chercheur non convaincu de la validité des résultats de ses travaux, peut être assez habile pour faire une présentation convaincante, mais heureusement pas tout le temps, quels que soient ses talents dans ce domaine. Il se peut aussi, qu'un chercheur soit parfaitement convaincu de la validité des résultats de son travail, et que maladroit en communication, ne parvienne pas à être convaincant dans son exposé. Cette situation est moins fréquente chez les Enseignants - Chercheurs que chez les Chercheurs plein temps. Le fait du chercheur convaincu et convaincant ne constitue pas une garantie totale pour que sa théorie et les résultats de sa recherche soient dans le vrai. Des recherches complémentaires doivent continuer à confirmer les prévisions de la théorie. Il suffit d'un seul cas qui ne la vérifie pas pour l'infirmer, ou poser matière à problème. Cela peut paraître particulièrement laborieux, mais le chercheur se doit de vaincre l'ignorance au profit de la meilleure connaissance tout comme l'artiste se doit de surmonter les problèmes que lui imposent la nature, les règles de son art.

L'évolution, des moyens techniques disponibles et celle de l'état des connaissances, conduit le chercheur à reformuler la problématique objet de ses travaux dans une approche renouvelée, tenant compte de son environnement du moment et peut être même de celui d'un futur proche. Cet environnement n'est pas toujours identique d'un pays à un autre, malgré la disponibilité d'un réseau de communication multi-services à l'échelle planétaire tel que l'Internet. En conséquence la productivité moyenne du chercheur, dépendant en parti de cet environnement, y sera très inégale, très élevée dans les pays développés où la recherche est devenue un phénomène culturel, très faible ailleurs où elle n'est pas organisée voir même peu considérée au point que chaque chercheur est perçu comme 'un Doctus Cum Libro'. Dans la majorité des pays riches post- industriels, le potentiel

humain voué à la recherche est constamment valorisé par une sorte de formation continue, par une organisation sans cesse performée, par une meilleure intégration des Enseignants - Chercheurs, des Universités et des Grandes Ecoles au redéploiement économique, à leur association dans des projets de développement des entreprises productrices de biens et de services, à leur contribution à la culture de la connaissance et de l'action innovante. Alors que les pays sortis dévastés d'une dure et longue colonisation, croulent de nouveau sous le poids d'un endettement qui perdure quoique le principal ait été remboursé plusieurs fois dans certains cas. Les peuples de ces pays peu instruits, déculturés pendant la colonisation n'ont pas songé à réclamer le dédommagement pour faits de colonisation auquel ils pouvaient prétendre ou même y avoir droit, parce que soucieux seulement de retrouver leur indépendance, en souhaitant garder malgré tout des rapports de coopération avec l'ex-puissance occupante et son peuple. Même après avoir valorisé leur potentiel humain post- indépendance, ces pays n'ont pas réussi à le faire contribuer efficacement à leur décollage économique. Bien plus, les chercheurs sont ressentis comme une charge et non comme une force dynamisante. D'où l'émigration vers les pays riches des chercheurs, i.e on continue à assister à un transfert de richesse et de technologie des pays pauvres vers les plus riches. Les pays pauvres ne devraient pas avoir une vision seulement symptomatique de leurs maux ; ils est plus important d'en cerner aussi l'origine, les causes. Par exemple, si l'on veut maintenir les chercheurs dans leur pays d'origine, il ne suffit pas d'attribuer une enveloppe budgétaire plus conséquente au secteur de la recherche, il est nécessaire d'accompagner cette attribution d'une organisation efficiente dotée d'une dynamique non entravée par de lourdes procédures. L'uniformisation civilisationnelle progressive actuelle, cache des ERM⁽¹⁾ toujours croissants aussi bien entre pays, qu'entre les individus d'un même pays, ce qui constituera à la longue une source de conflits sociaux entre les marginaux et les nantis. Il importe donc d'organiser le pays en maintenant l'ERM à un niveau optimal, i.e un compromis entre les motivations favorables à l'esprit d'entreprise et celui d'une justice sociale qui ne dévalorise pas le travail.

Par ailleurs, des institutions comme les universités américaines de Stanford, du M.I.T, de Harvard, de Yale, de Princeton, ... donnent des exemples d'intégration réussie de leur travail aux préoccupations d'ordre social et économique de leurs pays. Une analyse comparée de ces intégrations réussies dans les pays avancés ou même moins avancés permettra d'identifier les facteurs clés qui contribuent à cette réussite, et à partir de cette analyse il sera alors possible de lancer une première esquisse à adapter à notre pays.

Avant de clore cet exposé, nous tenons à souligner l'intérêt que suscitent les conférences organisées conjointement par l'Ecole Nationale Polytechnique (E.N.P) et l'Association Des Anciens Elèves (A.D.E.P), tant par les thèmes débattus que par la qualité des membres qui y participent. Nous encourageons cette initiative et nous lui souhaitons une bonne continuation. Nous pensons que cette démarche est également un pas en direction de l'intégration citée plus haut.

Le directeur de rédaction
Professeur : A.ZERGUERRAS

⁽¹⁾ ERM est l'Ecart de Revenu Moyen, par exemple l'(ERM)₁₀ est l'écart des revenus moyens entre les 10% des plus riches et les 10% des plus pauvres de la population d'un pays.

Full-Wave Analysis of Microstrip Antennas Employing Coplanar Parasitic Elements Based on Moment Method Solution in Spectral Domain

A. AZRAR and R. AKSAS

Laboratoire de Télécommunication de ENP, d'El-Harrach, 10 Avenue Hassen Badi 16200, Alger.

Abstract

The method of moment in spectral domain is used to perform an analysis of microstrip structures with and without parasitics. The analysis enables the prediction of the current distribution on the surface of the radiating elements; with which the radio-electric properties of the antenna are determined. We have verified the validity of this model at the high frequencies (25 GHz) and shown that the circular polarization is possible with an appropriate single feed.

The increase of the inherent narrow band-width of the microstrip antenna is possible by adding parasitic elements coplanar to the main patch. While highering the directivity this kind of configurations introduces an elevation in the level of secondary lobes and may change the direction of the maximum radiation.

1. Introduction

In the last few decades, extensive effort has been devoted to the utilization of the high frequency-wave portion of electromagnetic spectrum. This is motivated by the willing to use non-saturated channels. When it is noted that the frequency range 10^9 to 10^{12} Hz contains a thousand sections like the frequency spectrum from 0 to 10^9 , the value of developing this band as a means of increasing the available frequency may be readily appreciated.

With the move to exploit more fully the high frequency-wave spectrum and technological development in the high frequency-wave generation, using both solid state devices (very small active elements) and tubes, and in high frequency-wave components and circuitry (resulted from the advances in computer means for accurate characterization) has also come the requirement for new concepts in high frequency-wave antennas. To a large degree, the antennas to be used for many applications can be derived by simply scaling the counterpart microwave design [1]. However, this approach is not always practical due to the reduction in size and tighter monitoring tolerances that arise. In addition to, the trend toward miniaturization and use of integrated circuits, particularly in regard to the utilization of the high frequency-wave systems in missiles and subminitions, does not support the use of scaled conventional antennas, and these would be too large physically and cumbersome to integrate or use comfortably [2]. Thus, the interest in developing low-profile substrate-based planar arrays has grown steadily. Microstrip antennas and printed circuit antennas are obvious candidates for such applications.

Microstrip antennas have evolved from the undesirable radiation effect of microstrip lines; the concept which is proposed by Deschamps [3] in 1953. It was the tendency in the past to avoid rather than investigate the phenomena. It was Lewin [4] who first investigated the radiation from microstrip discontinuities. They have received little attention until the early 1970's [5-6], when there came a new demand for compact and lightweight antennas in many applications. Prompted largely [7] by the effort the Air Rome Development Center, Hanscom AFB, Massachusetts, vigorous research activities have taken place in the recent time, and the development of this family of antennas has finally

evolved into a new era [8-9] with products vastly different from their ancestor.

2. Generalities

Microstrip antennas, as shown in Fig 1, consists of a very thin ($t \ll \lambda_0$ with λ_0 is the free space wavelength) metallic strip (also called patch) printed over a grounded dielectric substrate with thickness commonly very small as compared to the wave-length ($h \ll \lambda_0$; usually $0.003\lambda_0 < h < 0.05\lambda_0$).

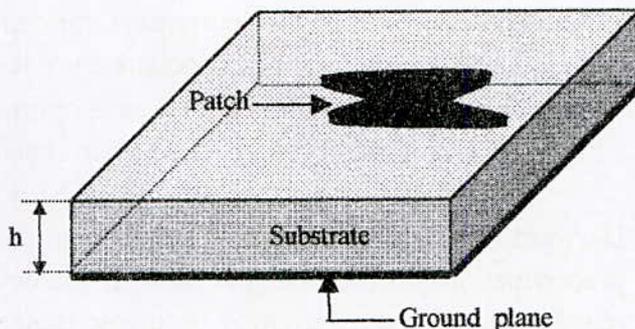


Fig. 1 Microstrip antenna structure

Despite the severe limitation in the bandwidth, microstrip antennas have many unique attractive features: compact in structure, light in weight, low in profile, conformable to planar and non-planar surfaces, simple and inexpensive to manufacture using modern printed circuit technology, mechanically robust when mounted on rigid surfaces, integrable with feeding networks and microwave devices, particularly compatible with the modern MMIC technology, and when the particular patch shape and mode are selected they are very versatile in terms of resonant frequency, polarization, pattern and impedance. Lately, much efforts has been made to search for new designs, such as the so-called multifunctional microstrip antenna [7]. For example, a single patch can be designed to function for any polarization state without power dividing network and phase shifters, even possibly at two different frequencies.

However, these antennas are inherently inefficient and narrow-band radiators, which could have significant

dielectric loss, and surface wave loss. Thus, for these antennas, a low Q-factor, or SWR (standing wave ratio) at the input port, does not imply a good antenna. It is therefore important to emphasize that their bandwidth can be meaningfully defined only by examining over the band the stability of all antenna performances, including gain, pattern, and polarization, in addition to the commonly used criteria of the input impedance.

Most antennas are passive devices that can be used for both transmission and reception. Reciprocity theorem permits the unique characterization of the antenna in either of the two modes. Several factors suggest that, it is easier to study the antenna in the transmit mode. Hence, the antenna should be fed. Different methods exist for energizing the structure. When choosing the most appropriate antenna configuration for a particular application, the means of excitation of the radiating element is an essential and important factor which requires careful consideration since it can considerably change the radiation characteristics of the structure. Matching is usually required between the feed line and the antenna, because antenna input impedances differ from the line impedance, (customary 50 Ω). Matching may be achieved by properly selecting the location of the feed line. However, the location of the feed line may also affect the radiation characteristics by parasitic radiation.

The principle considerations for good design are:

Low insertion loss;

Low space utilization;

Simplicity in realization;

Minimum number of discontinuities in the feed so it does not contribute to main-element radiation, thereby increasing the cross-polarization level.

Several methods are available for the analysis of the microstrip patch antenna. They can generally be divided into three groups according to the models assumed (empirical models, semi-empirical models and Full-wave analysis model). These models have different levels of complexity and require vastly different computation efforts. Although there is no simple clear-cut rule as to which one of the three is best to use, the first guideline would be the antenna thickness. Others are the particular antenna performance to be evaluated, the antenna geometry, the feeding system, and so on.

3. Choice of the analysis method and its steps

The only approach to a complete solution of a microstrip patch antenna employing parasitic elements coplanar with the main patch is the full-wave analysis technique. In this category, one finds formulation that are electromagnetically rigorous as well as computationally extensive. In this class, either the electric field integral equation (EFIE) or mixed potentials integral equation (MPIE) is used in characterizing the microstrip radiating structure. The relevant equation is solved using the moment method in space or spectral domain [10-11] or other techniques [12].

The full-wave analysis based on EFIE is used to deal with the problem at hand. Here, the structure under study is completely characterized by an integral equation where the unknown being the scattered electric field. The

steps involved in this technique are:

1. Obtain an equivalent problem in which the physical structures are replaced by equivalent sources (whose values are to be determined) radiating in an infinite homogeneous region or a region for which the dyadic Green's function is analytically known;
2. Apply boundary conditions appropriate to the problem (e.g. total electric field tangential to a perfectly conducting surface is zero) to obtain an equation which the equivalent sources must satisfy;
3. Discretize the domain over which the equivalent sources exist (e.g. the antenna surface) and expand the equivalent sources in terms of basis functions defined on the discretized domain. The coefficients of the basis functions are the unknown quantities to be determined;
4. Enforce the boundary condition equation from step two in some sense over the various regions of the discretized structure to obtain a system of simultaneous equations. This is often referred to as "testing procedure";
5. Solve the system of equations for the unknown coefficients of the basis set representing the equivalent sources;
6. Compute other quantities of interest which depend on the equivalent source distribution, such as antenna impedance and radiation pattern.

This technique presents some attractive features among which we state: ease of formulation and application to arbitrarily shape of the patch, dielectric as well as conductor losses are easily included in the analysis, and the surface waves are taken into account. However, it presents some limitations in that ground plane is considered infinite but its effect can be added according to Ref. [13], and the computational requirements are very demanding (memory space and execution time).

The principle steps of the analyses of a microstrip radiating structure involved in this these are illustrated in the Fig. 2.

The full-wave analysis technique used in the study of the structures proposed in the current work is based on solving the characterizing electric field integral equation with the unknowns, instead of the scattered electric field, being the surface current distributions on the patches. The passage from electric field as unknown to the currents is accomplished by the use of the exact dyadic Green's function for the Grounded dielectric substrate. For the solution of the integral equation, the moment method is used. The solution is done in the spectral domain, this is because of the less number of the basis functions needed in the decomposition of the surface current. The obtained solution with the help of the Schelkunoff principle is used to obtain the radioelectric properties of the antenna, and to evaluate the influence of different dimensions and electrical parameters.

4. Theory and formulation

4.1 Problem statement

The analysis of a microstrip radiating structure is needed. The treatment involves the use of integral equation model. The current model is able to handle

patches of arbitrarily shapes where no educated guess of the surface-current distribution is possible. The model takes into account mutual coupling between elements. The effect of surface waves is included with dielectric losses too. First, an arbitrarily shaped microstrip radiating structure is characterized by an integral equation where the main unknown is the electric surface current density on the patch. This model then is applied to a rectangular patch as a narrow-band structure. To broaden the bandwidth, a rectangular patch with parasitic elements is used under this analysis. In the aforementioned applications, the integral equation is solved by a method of moments.

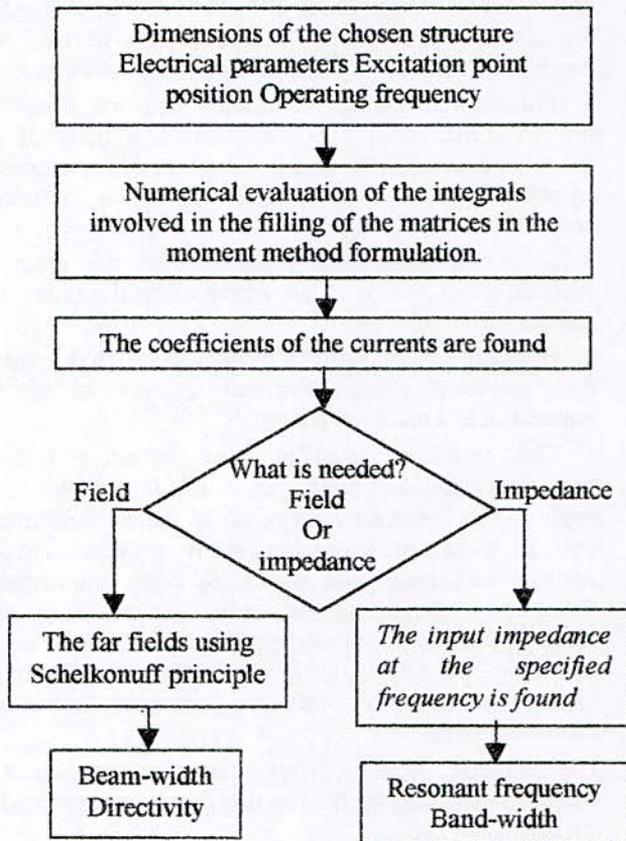


Fig.2 Flow chart showing the analysis steps

4.2 Geometry of the model

For the theory to be clear, the electric surface current model will be developed for the simple microstrip structure shown in Fig. 3. The substrate is assumed to extend to infinity in the transverse directions and is made of nonmagnetic, isotropic, homogeneous material with losses. So its dielectric constant can be written as

$$\epsilon = \epsilon_0 \epsilon_r = \epsilon_0 \epsilon_r' (1 - \tan \delta) \quad 0 < z < h \quad (1)$$

With h is the dielectric substrate thickness. The antenna is excited by a coaxial cable, with which \vec{E}^{inc} and \vec{H}^{inc} are equivalently created.

4.3 Integral equation formulation

For the configuration shown in Fig. 3, an integral

equation can be obtained by making use of the condition that the tangential component of the electric field should be zero on the conducting patch surface. This condition is stated as:

$$\vec{E}^{inc}(x, y, h) + \vec{E}^{scat}(x, y, h) = 0 \quad (2)$$

As stated previously, \vec{E}^{inc} is the electric field created by the feed (probe) current situated at position (x_p, y_p) . So \vec{E}^{inc} can be called excitation field that induces surface currents on the conductor patch. The induced currents in turn create diffracted or scattered electromagnetic fields.

To work out the problem in spectral domain, introduce the following Fourier transform (FT) pair

$$\tilde{A}(\vec{k}_t, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \vec{A}(\vec{\rho}, z) \exp(-j\vec{k}_t \cdot \vec{\rho}) d\vec{\rho} \quad (3)$$

$$\vec{A}(\vec{\rho}, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{A}(\vec{k}_t, z) \exp(j\vec{k}_t \cdot \vec{\rho}) d\vec{k}_t \quad (4)$$

where

$$\vec{k}_t = k_x \hat{x} + k_y \hat{y} = \beta(\cos \alpha \hat{x} + \sin \alpha \hat{y})$$

$$= \beta \hat{k}_t \text{ is the transverse vector wave number}$$

$$\text{and } \beta = |\vec{k}_t|.$$

The symbols $\hat{\cdot}$ and \sim are used to denote respectively the unit vectors and spectral domain quantities.

Application of Equ. (3) in Equ. (2) yields

$$\tilde{E}^{inc}(k_x, k_y, h) + \tilde{E}^{scat}(k_x, k_y, h) = 0 \quad (5)$$

The unknown in Equ. (4) is the scattered field, however, as stated in the introduction of this chapter, the objective of the analysis is to determine the surface current \vec{J}_s on the patch conductor of the antenna, and then to deduce all the antenna's characteristics.

By virtue of the reaction theorem [14], Equ. (5) can be rewritten as:

$$\int_{\Sigma} \vec{J} \cdot \vec{E}^{scat} ds + \int_{V_s} \vec{J}_f \cdot \vec{E}^{scat} dv = 0 \quad (6)$$

Σ denotes the surface of the conductor on which current \vec{J} is located.

V_s denotes the volume containing the feed source.

\vec{J}_f denotes the feed current.

Equ. (6) is the required integral equation that characterizes the radiating structure of Fig. 3. As far as the solution of this equation is concerned, it is more interesting to work with currents instead of fields. This

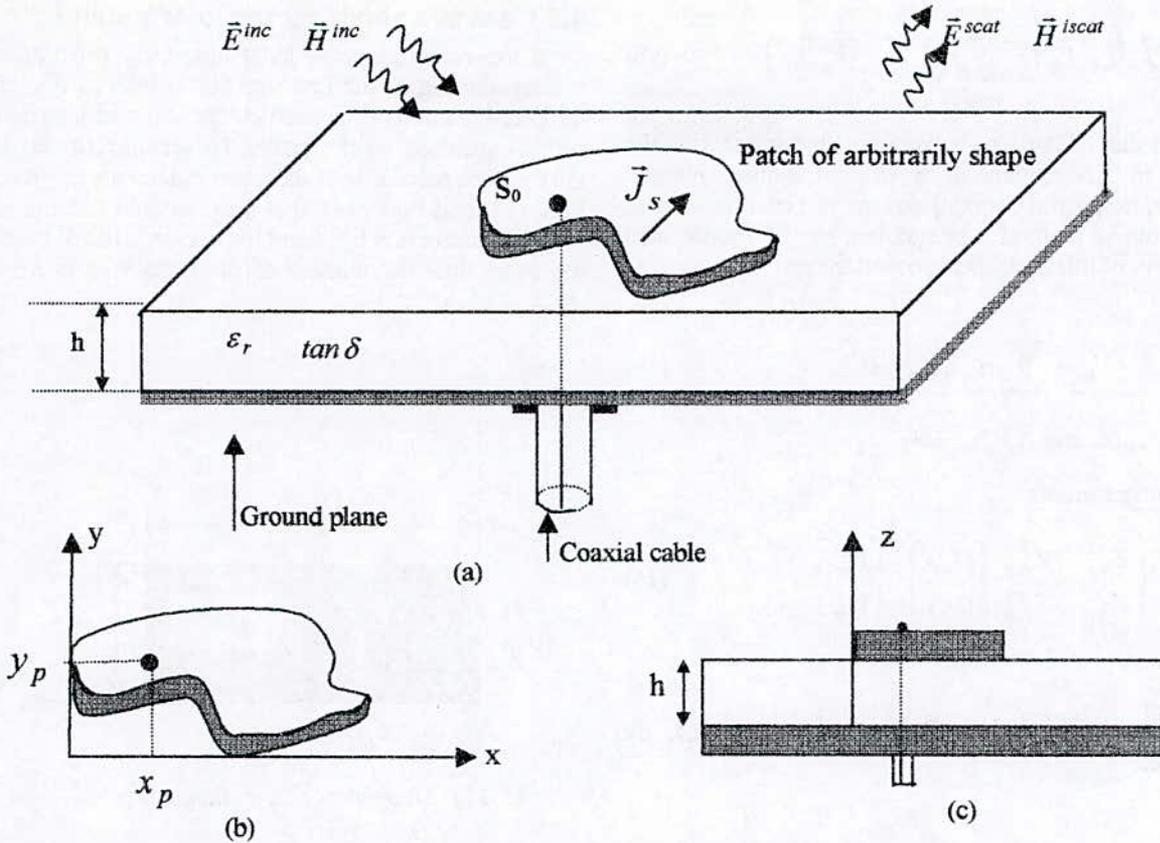


Fig. 3 (a) Geometry of the microstrip structure (b) Top view (c) Side view

suggests another look to Equ. (6) and to the microstrip structure. In view of the linearity of the field equations and most passive structures (including antennas), the scattered-wave amplitudes are linearly related to the incident-wave amplitudes. So in space domain, the electric field is expressed as the convolution integral of the function characterizing the structure (Green's function) and the inducing current. And the theory of linear systems enables the transformation of the convolution integral in space domain into a simple product in spectral domain. This will be applied to Equ. (2-5) in the foregoing sections

4.3.1 Green's function

The best way to visualize the dyadic Green's function is to consider the figure below.

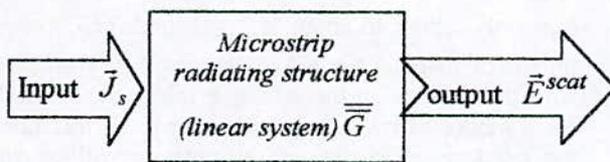


Fig.4 Microstrip structure characterized by \tilde{G} simulated by a linear system

From Fig. 4 we see that \tilde{G} gives a relation of cause to an effect between the surface current \vec{J}_s and the electric

field \vec{E}^{scat} that it creates [16]. From the theory of linear systems this relation can be written as [17]:

$$\vec{E}^{scat} = \int_{s_0} \tilde{G} \vec{J}_s ds \quad (7)$$

In spectral domain, the above integral is transformed to a simple product. Since, the inducing current and the resultant field are vector quantities, Equ. (7) can be written as

$$\begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} = \begin{bmatrix} \tilde{G}_{xx} & \tilde{G}_{xy} & \tilde{G}_{xz} \\ \tilde{G}_{yx} & \tilde{G}_{yy} & \tilde{G}_{yz} \\ \tilde{G}_{zx} & \tilde{G}_{zy} & \tilde{G}_{zz} \end{bmatrix} \begin{bmatrix} \tilde{J}_x \\ \tilde{J}_y \\ \tilde{J}_z \end{bmatrix} \quad (8)$$

4.4 Moment method implementation

Analytical solution of Equ. (6) is precluded even for simple geometries. The moment method [MM] [18] is used to obtain a numerical solution. In MM formulation, the surface current \vec{J}_s is decomposed into a set of orthogonal basis functions \vec{J}_n such that each component of \vec{J}_s can be written (in spectral domain) as:

$$\tilde{J}_x(k_x, k_y) = \sum_{n=1}^{N_x} \alpha_{xn} \tilde{J}_{xn}(k_x, k_y) \quad (9a)$$

$$\tilde{J}_y(k_x, k_y) = \sum_{n=1}^{N_y} \alpha_{yn} \tilde{J}_{xn}(k_x, k_y) \quad (9b)$$

By means of Parseval's theorem (that transforms the integral in space domain in (6) to another integral expressed in spectral domain) and use of Galerkin scheme of the moment method . the problem can be transformed to a system of linear algebraic equations as:

$$V_m + \sum_{n=1}^N \alpha_n Z_{mn} = 0 \quad (10)$$

$m=1,2, \dots, N$ and $N = N_x + N_y$

Or in matrix form as:

$$\begin{bmatrix} Z_{mn}^{xx} & Z_{mn}^{xy} \\ Z_{mn}^{yx} & Z_{mn}^{yy} \end{bmatrix} \begin{bmatrix} \alpha_{xn} \\ \alpha_{yn} \end{bmatrix} = \begin{bmatrix} V_m^x \\ V_m^y \end{bmatrix} \quad (11)$$

where

$$Z_{mn}^{pq} = \frac{1}{4\pi^2} \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} \tilde{J}_{pm}^*(k_x, k_y) \tilde{G}_{pq}(k_x, k_y) \tilde{J}_{qn}(k_x, k_y) dk_x dk_y \quad (12)$$

and

$$V_m^p = -\frac{1}{4\pi^2} \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} \tilde{J}_{pm}^*(k_x, k_y) \tilde{G}_{pz}(k_x, k_y) \tilde{J}_z(k_x, k_y) dk_x dk_y \quad (13)$$

4.4.1 Basis functions

The basis functions employed in the Greens function solutions of the microstrip antenna problem, as indeed in all moment-method solutions, are of two varieties: entire-domain or sub sectional [19]. In the first case the unknowns are expanded in a complete set of basis functions, having the entire structure (e.g. entire patch) as the domain, in the second case, the currents are approximated piecewise by simple functions whose domains are considerably smaller than the structure itself. Examples of the two classes can be found in ref [20]. Previous studies of the microstrip antennas reveal that a few judiciously chosen entire-domain expansion functions usually suffice to represent the patch currents accurately with low order MM matrices. Use of subsectional bases leads to a larger matrices. However, they are more versatile in their applicability to a much wider class of geometries.

The choice of the basis function determines the accuracy and the efficiency of the solution. This choice is guided by previous investigations of similar geometries and dictated by the inherent geometrical features. For the current analysis, rectangular patch with and without parasitic elements, the piecewise sinusoidal basis functions are chosen. (from the study of the resonant cavities, it is found that the sinusoidal functions are attached to rectangular cavities and Bessel functions to circular).

4.5 Case of a single rectangular patch

A top view of a rectangular microstrip patch antenna is shown in Fig. 5. the first step that will be carried out is the determination of the coefficients of the surface current series expansion on the patch. To accomplish this Equ. (11) will be solved, with the matrix elements as given by Equ. (12) and Equ. (13). It is good to note that the order of the Z-matrix is N by N and the vector V has N rows. At the same time the number of the unknowns is N , with $N = N_x + N_y$.

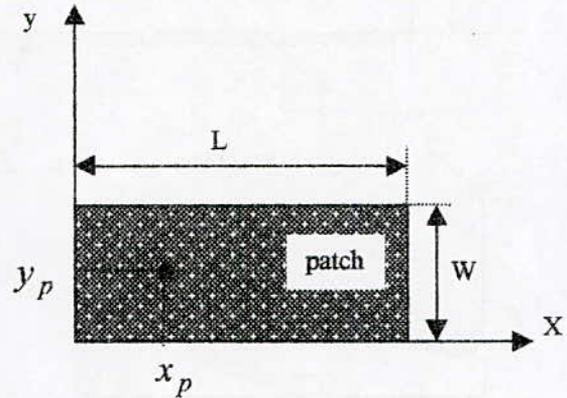


Fig. 5 Top view of a rectangular patch

4.6 Case of Patch with parasitic elements

All the theories available for the analysis of patch antennas including model based techniques and numerical methods for solving the relevant integral or differential equations end to say that microstrip antenna may be seen as a high Q single resonant circuit. Generally, several ways for improving the bandwidth of microstrip antennas can be used as follows.

1. *Impedance matching networks* [21-22]: this method is convenient and effective. microstrip antenna can be seen as a high Q-value RLC resonant circuit, so it is difficult to obtain the broadband matching network. Moreover, because of the matching network, the microstrip antenna causes additional radiation and the size of the structure becomes very large.
2. *Reducing the Q-value:* In fact, the radiation ability of microstrip antenna can be improved by increasing the height of substrate or decreasing the substrate dielectric constant. This method is also relatively simple, but ϵ_r may only reduce to about $\epsilon_r \cong 1$, and the thickness of microstrip antenna can not be increased very much. First with the increase in the substrate thickness, the surface-wave modes of microstrip antenna may be increased, so the radiation efficiency of the antenna will be varied. Second, the weight of the antenna can be increased by the use of the thick substrate, and the cost can also be raised.
3. *Modifying the resonant circuit:* it is well known that microstrip can be seen as a high Q-value single resonant circuit. To improve the bandwidth of microstrip antenna, we can manage to vary a single resonant circuit into two-resonant or multiple resonant circuit. Examples are: i) addition of parasitic patches on top of the original patch [13], [23-25] and ii) addition of gap-coupled patches to

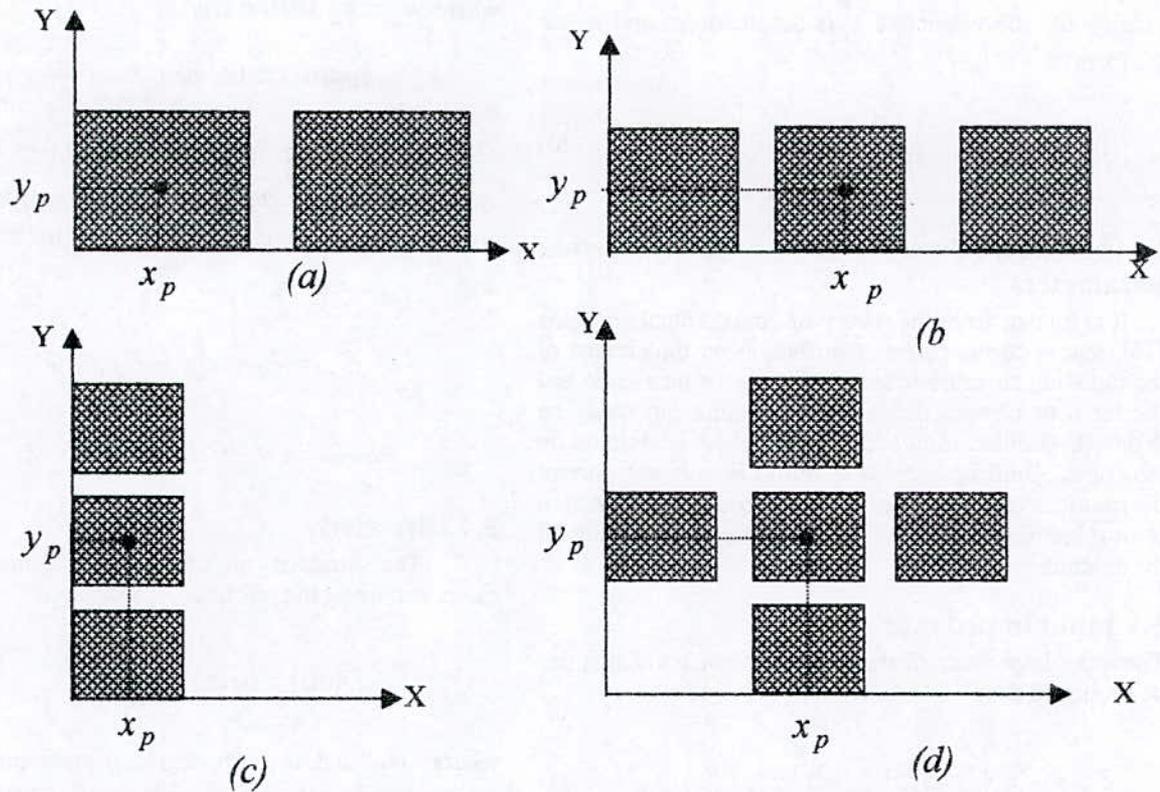


Fig.6 Parasitic elements gap-coupled to coaxial fed rectangular patch antenna for improving the bandwidth

the radiation and non-radiation edges [19], [26-28]. In the current work the latter technique as shown in Fig. 6 is used to improve the bandwidth of microstrip antenna.

In Fig. 6 parasitic elements and the driven patch are in close proximity they interact in a complicated manner. This interaction is called mutual coupling and the effect is to change the current on an antenna from that which it would have if it were isolated in free space. For the case of having more than a patch, multiple resonances of the patches together will constitute a broaden bandwidth. In general, mutual coupling is primarily attributed to the fields that exist along the air-dielectric interface. The fields can be decomposed to space waves, higher order waves, surface waves, and leaky waves.

Furthermore, the patches interact with one another, and produce a resulting radiation pattern that is the vector sum of the individual ones.

In the case of parasitic elements gap-coupled to a coaxial fed rectangular patch, the characterizing integral equation can easily be found as [29]

$$\sum_{k=1}^n \iint_{S_p} \overline{\overline{G}}(r, r') \vec{J}_p^k(r') dr' + \iint_{V_f} \overline{\overline{G}}(r, r') \vec{J}_f(r') dr' = 0 \quad (14)$$

where $\vec{J}_s = \sum_{k=1}^n \vec{J}_p^k$, \vec{J}_f and \vec{J}_p^k are the currents in

the probe and on the k^{th} patch respectively, S_p is the patch surface, V_f is the volume containing the fed current, and n is the number of the patches in the antenna

structure.

As we have done for the simple patch, the mathematical solution of the problem is reduced, from Equ. (14), to a matrix inversion given as:

$$\begin{bmatrix} Z^{11} & Z^{12} & \dots & Z^{1n} \\ Z^{21} & Z^{22} & \dots & Z^{2n} \\ \vdots & \vdots & \vdots & \vdots \\ Z^{n1} & Z^{n2} & \dots & Z^{nn} \end{bmatrix} \begin{bmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{bmatrix} = \begin{bmatrix} V^1 \\ V^2 \\ \vdots \\ V^n \end{bmatrix} \quad (15)$$

The sub matrix Z^{ij} has the form of

$$[Z^{ij}] = \begin{bmatrix} Z_{mn}^{xx} & Z_{mn}^{xy} \\ Z_{mn}^{yx} & Z_{mn}^{yy} \end{bmatrix}^{ij} \quad (16)$$

And its element has a general form of

$$(Z_{mn}^{pq})^{ij} = \frac{1}{4\pi^2} \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} \tilde{J}_{pm}^{i*} \tilde{G}_{pq} \tilde{J}_{qn}^j dk_x dk_y \quad (17)$$

The sub vector $[V^i]$ has the form of

$$[V^i] = \begin{bmatrix} V_x^p \\ V_y^p \end{bmatrix}^i \quad (18)$$

And its element has the form of

$$(V_m^p)^{ij} = -\frac{1}{4\pi^2} \int_{-\infty-\infty}^{\infty} \int_{-\infty-\infty}^{\infty} \tilde{J}_{pm}^{i*} \tilde{G}_{pz} \tilde{J}_z dk_x dk_y \quad (19)$$

Finally, the sub-vector $\left[\alpha^i \right]$ is the unknown and is has the form of

$$\left[\alpha^i \right] = \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix}^i \quad (20)$$

5. Theoretical determination of some antenna parameters

It is known from the theory of conventional antennas [15], that once the current distribution on the surface of the radiating structure is assumed, then the near zone and the far zone characteristics of the antenna can easily be deduced. Similar things can be applied to microstrip structures. Until now, what is found is only the current distributions on the patches. So the purpose of this section is to relate this current to the radioelectric properties of the antenna.

5.1 Input impedance

The input impedance of the microstrip patch antenna can be evaluated by:

$$Z_{in} = - \sum_{k=1}^p \left\{ \sum_{n=1}^{N_x} \alpha_{xn}^k V_{xn}^k + \sum_{n=1}^{N_y} \alpha_{yn}^k V_{yn}^k \right\} \quad (21)$$

where p is the number of patches

5.2 Bandwidth

Frequently, the used measures of an antenna bandwidth are the impedance bandwidth which indicates the frequency band over which the SWR at the circuit port of the antenna remains below 2.

$$B = 200 \frac{f_2 - f_1}{f_2 + f_1} \quad (\%) \quad (22)$$

with f_1 and f_2 are the intersecting frequencies between the input SWR curve with the line SWR = 2.

5.3 Radiation pattern

Once the current distributions on the patches are known, the field equivalent principle of Schelkonuff can be used to determine the radiated fields. the relevant expressions are:

$$\vec{E}(r, \theta, \phi) = E_\theta(r, \theta, \phi) \hat{\theta} + E_\phi(r, \theta, \phi) \hat{\phi} \quad (23)$$

with

$$E_\theta = j \frac{k_0 e^{-jkr}}{2\pi r} \left\{ \tilde{E}_x(k_x, k_y, h) \cos\phi + \tilde{E}_y(k_x, k_y, h) \sin\phi \right\} \quad (24a)$$

$$E_\phi = j \frac{k_0 e^{-jkr}}{2\pi r} \left\{ \cos\phi \left[\tilde{E}_x(k_x, k_y, h) \sin\phi + \tilde{E}_y(k_x, k_y, h) \cos\phi \right] \right\} \quad (24b)$$

where $k_x = k_0 \sin\theta \cos\phi$

$$k_y = k_0 \sin\theta \sin\phi \quad \text{and} \quad k_0 = \frac{2\pi}{\lambda_0}$$

$\hat{\theta}$ and $\hat{\phi}$ are spherical coordinates unit vectors.

According to Ludwing [30], the co-polarized and the cross-polarized field components, noted respectively E_{co} and E_{cross} are given by

$$E_{co} = E_\theta \cos\phi - E_\phi \sin\phi \quad (25)$$

$$E_{cross} = E_\theta \sin\phi + E_\phi \cos\phi \quad (26)$$

5.4 Directivity

The directivity of the microstrip antenna can be calculated using the practical expression of

$$D(\text{dB}) = 10 \log_{10} \left(\frac{26000}{\theta_E \theta_H} \right) \quad (27)$$

where θ_E and θ_H (in degrees), represent the -3 dB beam-width respectively in the E-plane and H-plane.

6. Numerical results and discussion

A software program is implemented and it enables the determination of the current distribution on the patches of the microstrip antenna structure. With the obtained currents the characteristics of the antenna may be determined. The software program can handle all the structures that are shown in Fig. 5 and in Fig. 6.

6.1 single patch

A rectangular patch as shown in Fig. 5, with parameters: L = 11.43 cm, W = 7.62 cm, h = 1.59 mm, $X_p = 6.86$ cm, $Y_p = 3.81$ cm, $\epsilon_r = 2.64$, $\text{tg}\delta = 0.003$ Is studied.

Fig. 7 shows a draw of the resistive and reactive parts of the input impedance of the antenna described in section 3.3, as a function of a frequency. The frequency shows the resonance behavior of the microstrip structure. For a purpose of the validity of the analysis, the measure done by Richards at al [31] are superposed with the theoretical results on the same figure. We notice that the results compare well with.

The resonant frequency denoted, f_r , corresponds to input impedance purely real. For the current case, $f_r = 1.19$ GHz. The bandwidth, can be calculated from Fig. 8, it is found equal to 1%.

Far field pattern of the radiating antenna described previously can be drawn. The model based techniques as well as the Schelkunoff equivalence principle end to say that, in the direction perpendicular to the ground plane the components of the field add in phase and give maximum radiation normal to the patch; thus broadside antenna. so the radiation characteristics of the microstrip antenna structures gradually decrease to zero at $\pm 90^\circ$.

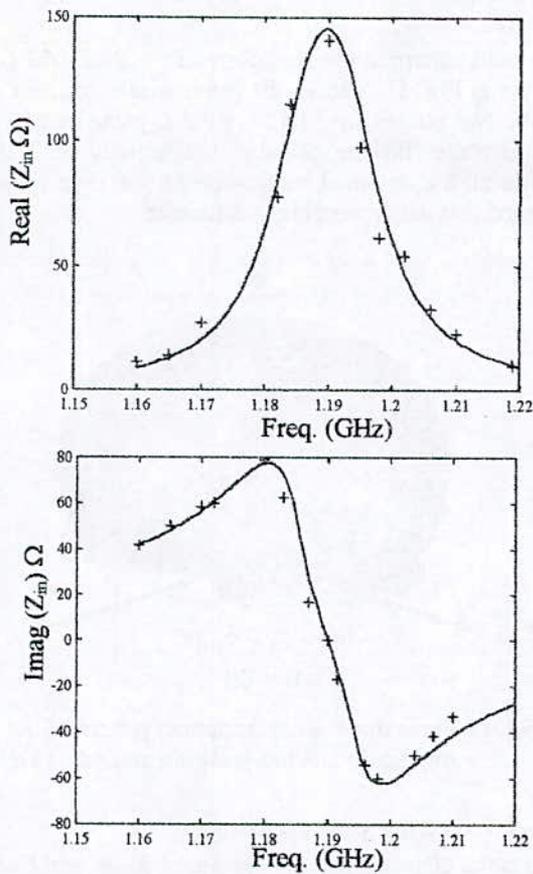


Fig-7 Input impedance of rectangular patch antenna

— Calculated results
+ measures of ref. [31]

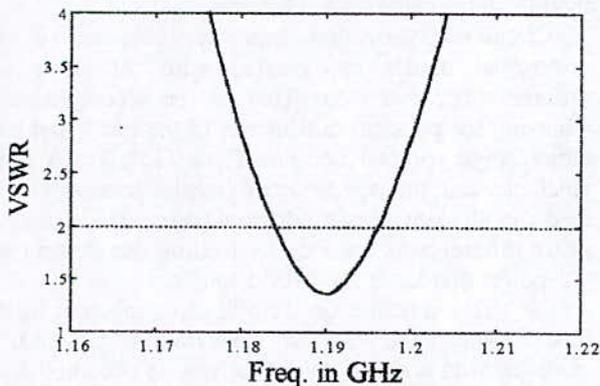


Fig. 8 Input VSWR of the rectangular antenna

To ensure the validity of the analysis and to study the polarization purity of the microstrip antenna, the co-polarization and the cross-polarization components of the radiated electric field are drawn in both E- and H-planes as shown in Fig. 9. In both planes, the co-polarization components compare well with the measures performed by Richards *et al* near the direction of maximum radiation. However, a shift more or less important is noticed, out of the maximum radiation region, between the calculated results and measures of the same reference. The minimization of this shift is tried by considering more basis functions but without significant improvement. Thus, the shift is attributed to the interfering field from the surface waves. In addition, the ground plane is considered infinite. The finite ground plane effect can be considered [13].

The minus 3 dB beam-widths of the structure in the E- and H-planes are respectively, 120° and 89° , so, the calculated directivity is 5.44 dB, and the one of the experiment [31] is 5.50 dB.

With the number of the basis functions considered ($N_x = N_y = 5$), we notice that the cross-polarization components of the electric field in both planes are practically null (less than -40 dB). This enhances the assumption that microstrip antenna is a linear polarization structure.

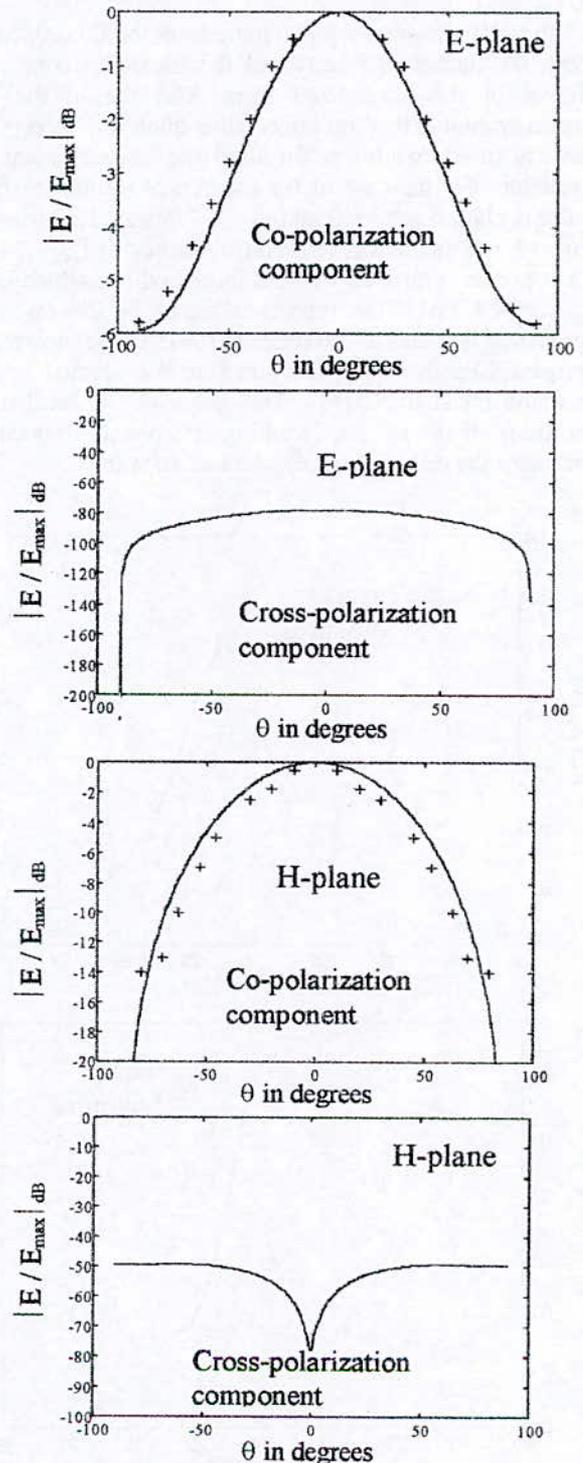


Fig. 9. Far field pattern of the patch antenna
— calculated.
++ measures of [31]

6.1.1 Microstrip antenna working around 24 GHz

The full-wave analysis is also able to handle the analysis of the microstrip structures operating at high frequencies. Although no previous analytical publications are available to compare the obtained results, but some measures from laboratory at microwave U.C.L., Louvain La-Neuve, Belgium are used for this purpose. The physical and electrical parameters of the antenna under study are $L = 3.55$ mm, $W = 4.82$ mm, $h = 0.5$ mm, $\epsilon_r = 2.33$ and $\text{tg}\delta = 0.001$.

Fig. 10 shows the input impedance of this antenna. Both the measured results and the theoretical ones are shown in the normalized form. The reason for the normalization is that, no information about the feed point that is used while performing the measurement is available. For the case of the theoretical results the feed point is chosen at $x_p = 0$ and $y_p = 1.775$ mm. Referring to Fig. 10, the theoretical resonant frequency of $f_{\text{cal}} = 24.25$ GHz can be compared with the measured one which is of $f_{\text{meas}} = 24.4$ GHz. the theoretical error in the resonant frequency is found of the order of 0.61%. The theoretical frequency bandwidth of this structure is evaluated, and it is found equal to 5.516%. This value of the bandwidth indicates that, as the working (resonant frequency) increases the resulting bandwidth increases too.

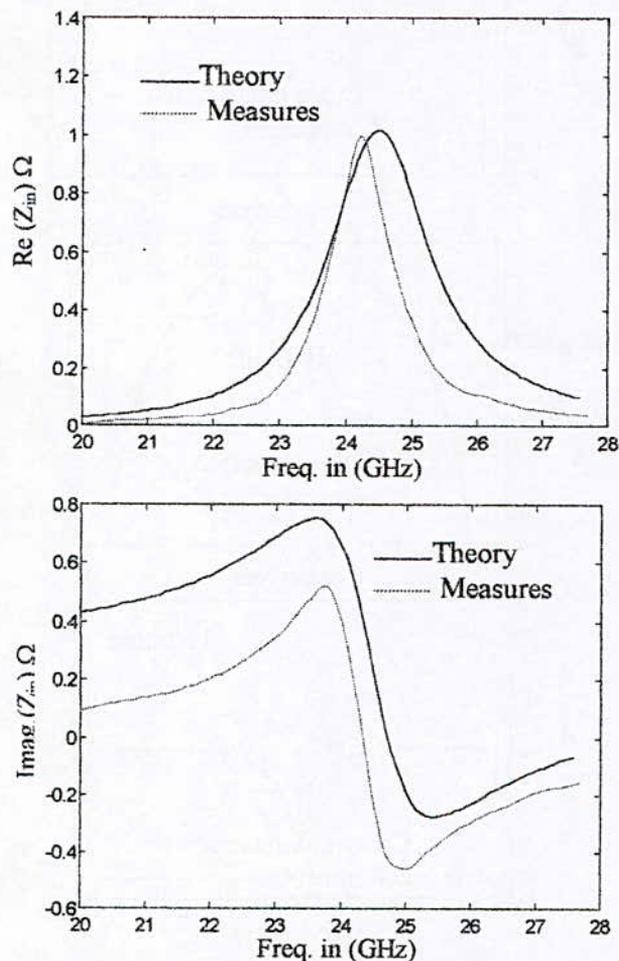


Fig. 10 Normalized input impedance of the antenna working around 24 GHz.

Far field pattern at the frequency of $f = 24.21053$ GHz is drawn in Fig. 11. The -3 dB beam-widths of the antenna in the two planes are 91.32° in the E-plane and 82.50 in the H-plane. So the calculated directivity is 3.38 dB. Again all the obtained results for the far zone properties assume that the ground plane is infinite.

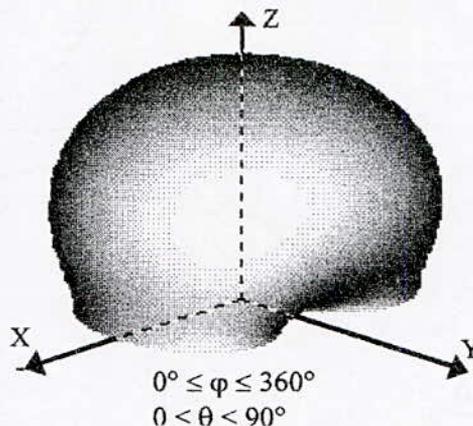


Fig. 11 Three dimensional radiation pattern of the microstrip antenna working around 24 GHz

6.1.2 Circular polarization

The patch elements that we discussed so far and those we will discuss in chapter 4 radiate primarily linear polarized waves if conventional feeds are used with no modification. However, circular polarization can be obtained using various feed arrangement or slight modifications made to the elements.

Circular polarization can be obtained if two orthogonal modes are excited with 90° time shift difference between them. This can be accomplished by adjusting the physical dimensions of the patch and using either single, or two or more feeds [15]. For a square patch element, the way to excite circular polarization is to feed the element at two adjacent edges. The quadrature phase difference is obtained by feeding the element with 90° power divider or 90° hybrid coupler.

To overcome the complexities inherent in dual feed arrangement, circular polarization can also be achieved with a single feed. One way to accomplish this is to feed the patch at a single point to excite two orthogonal degenerate mode (of some resonant frequency) of ideal equal amplitudes [15]. By introducing then a proper asymmetry in the cavity, the degeneracy can be removed with one mode increasing with frequency while the orthogonal mode will be decreasing with frequency by the same amount. Since the two modes will have slightly different frequencies, by proper design the field of one mode can lead by 45° while that of the other can lag by 45° resulting in a 90° phase difference necessary for circular polarization.

To illustrate the procedure, a square patch as shown in Fig. 12 is analyzed. Both the near zone and the far zone properties of the antenna are determined.

Fig. 13 shows the field patterns, in both principle planes, of the circular polarized square patch antenna. 10 basis functions along each direction are considered. At first glance, we see that the co-polarized and the cross-

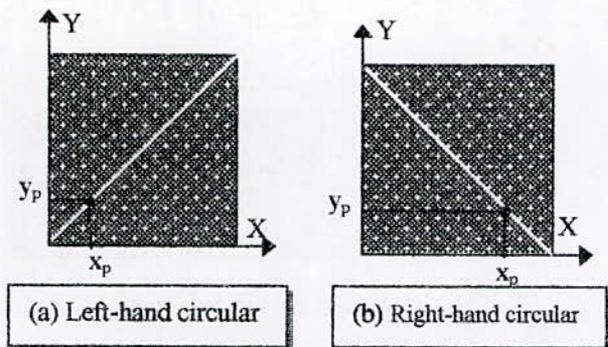


Fig. 12 Single feed arrangement for circular polarization of rectangular microstrip patches

polarized components in both planes are nearly of the same level in the $\theta = 0^\circ$ (maximum radiation direction). If more basis functions are considered, the two components will coincide in $\theta = 0^\circ$. If we deeply investigate the curves shown in Fig. 13 with the help of Equations (24), (25) and (26) we easily conclude that \tilde{E}_x and \tilde{E}_y are identical and this really ensures the circular polarization of the structure.

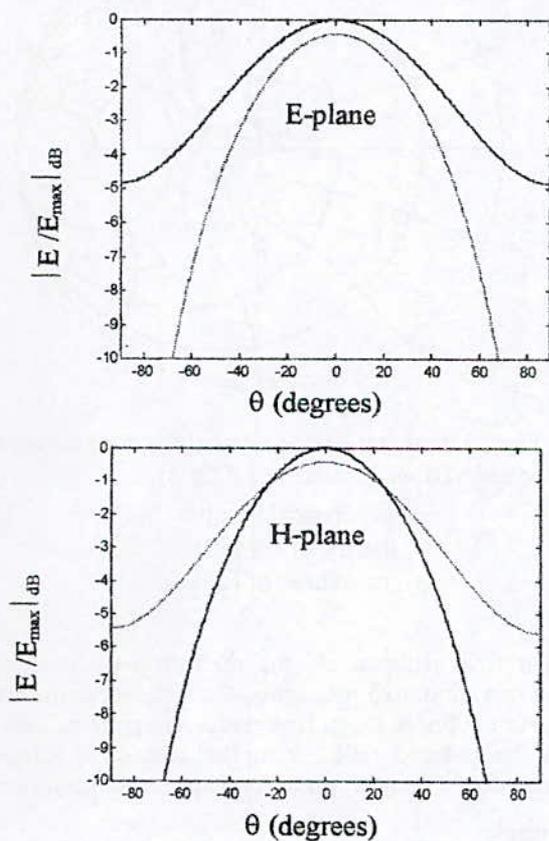


Fig. 13 Far field pattern of the circular polarization patch antenna — Cop. Comp. — Cross. Comp.

Remark If the position of the excitation point on the diagonal is changed, the circular polarization characteristic of the structure remains unchanged, only the level of the input resistance will be reduced as we move toward the center of the patch.

6.2 Antenna with parasitics

6.2.1 Antenna with single parasitic element.

In this section, the additional resonator gap-coupled to the rectangular patch antenna shown in Fig. 14 is considered.

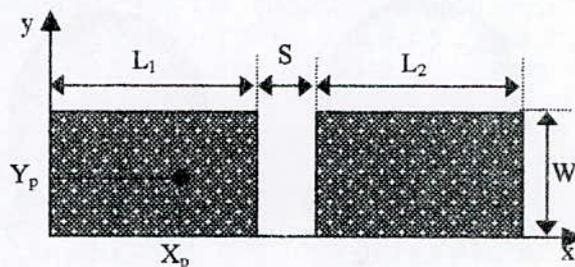


Fig. 14 Top view of the two elements microstrip antenna. $L_1 = 25.0$ mm, $L_2 = 24.5$ mm, $w = 30$ mm, $h = 1.5$ mm. $S = 0.7$ mm, $x_p = 0$, $y_p = 15$ mm. $\epsilon_r = 2.86$, $\text{tg}\delta = 0.001$.

Fig 15 shows the variation of the real and imaginary parts of the input impedance as a function of the frequency. For the purpose of comparison the input impedance of both single patch and two patches antennas are superposed. The figure indicates that the narrow single resonance behavior of a single patch antenna is stretched to approach two resonances or larger bandwidth. The impedance bandwidth is increased from 1.70 % for the single patch to 4.72 % for the two element structure. Physically, the two peaks are due to the resonances of the two patches which combine and form a larger bandwidth.

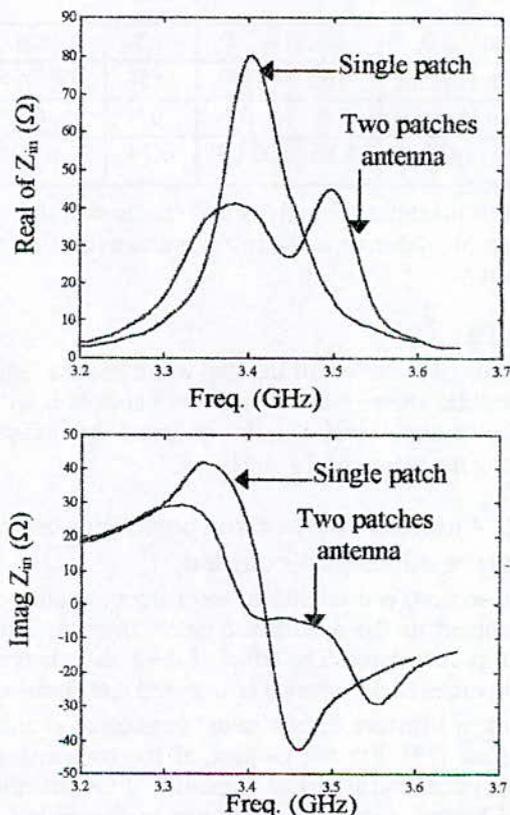


Fig. 14 input impedance of a single and two elements

Far field patterns of the gap coupled microstrip patch antenna, for different frequencies are given in fig. 15.

Unlike the case of the single patch antenna, the radiation pattern of the microstrip antenna employing parasitic element is highly changing with the frequency.

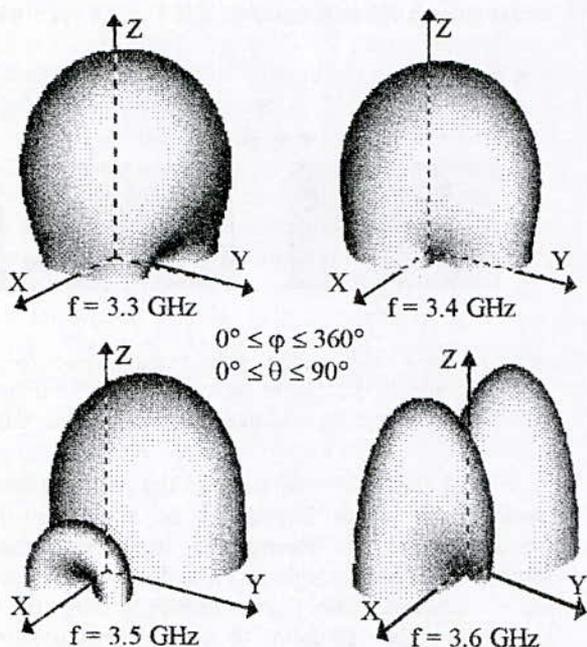


Fig. 15 Three dimensional field patterns of the two elements patch antenna

The effect of the frequency on the pattern is summarized in table below

Freq. GHz	3.3	3.4	3.5	3.6
θ_{max} in E-plane	-10°	-15°	-45°	-50° or 50°
θ_{max} in H-plane	0°	0°	0°	0°
D_{max} (dB)	5.85	6.09	6.14	6.03

Table 1 maximum directivity and direction of the maximum radiation in the two plane as a function of frequency

Remark :

The study of the effect of the gap-width and the length of the parasitic shows that an optimum bandwidth with flat input resistance level can be obtained by judiciously choosing the values of L_2 and S_2 .

6.2.2 Antenna with two parasitic elements radiating edges gap-coupled

In this section, two additional resonators radiating edges gap-coupled to the rectangular patch antenna shown in Fig. 16 is considered. The effect of the added elements on the properties of the antenna is analyzed and discussed.

First, a structure similar to the considered is analyzed by Kumar [25]. For the purpose of the comparison, the same physical and electrical properties of the structure are used. The two additional resonators to the central patch are taken identical ($L_1 = L_2$) and the gaps are of the same lengths ($S_1 = S_2$). The obtained results are shown in Fig. 17. The theoretical bandwidth (VSWR<2) of the simple configuration is 65MHz centered at 3.3 GHz (exactly

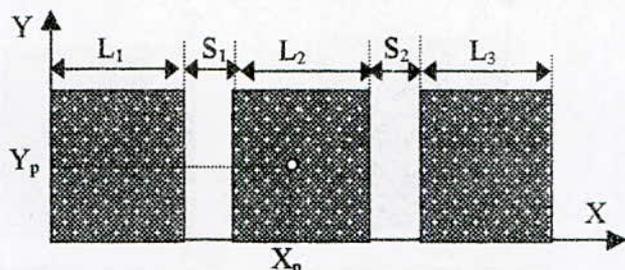


Fig. 16 top view of the three elements patch antenna

identical to the result of [26]) so the equivalent of 1.97%. However, for the structure employing parasitic elements the theoretical bandwidth is found equal to 205MHz while the measures of [26] gives 207 MHz. This bandwidth centered at 3.3 GHz is equivalent to 6.21%. Thus, the bandwidth of the radiating edge gap-coupled microstrip antenna is nearly 3.15 times the bandwidth of the rectangular patch antenna.

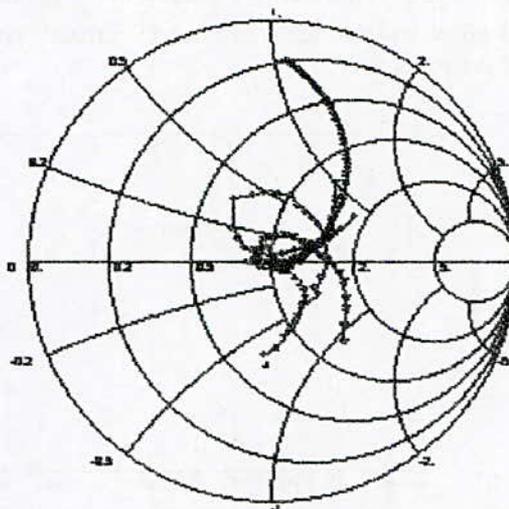


Fig. 17 input impedance locus of the three elements antenna (from 3.2 GHz to 3.5 GHz)

- ++ theoretical results
- ** theory of ref [32]
- oo measures of [25]

Far field patterns of the microstrip patch antenna using two additional resonators, for different frequencies, are given in Fig. 4.18. At first glance, the patterns indicate that, the radiated fields from the considered structure depend on the frequency both in shape and in power level.

Remark

An important increase in the bandwidth can be obtained when we make the parasitic patches of unequal lengths and a good choice of the gap-width.

6.2.3 Nonradiated edges gap-coupled microstrip patch antenna

In the previous section, the radiating edges gap-coupled microstrip antenna was analyzed. The main patch and the parasitics are positioned collinearly along the H-plane. In this section, the microstrip antenna to be studied

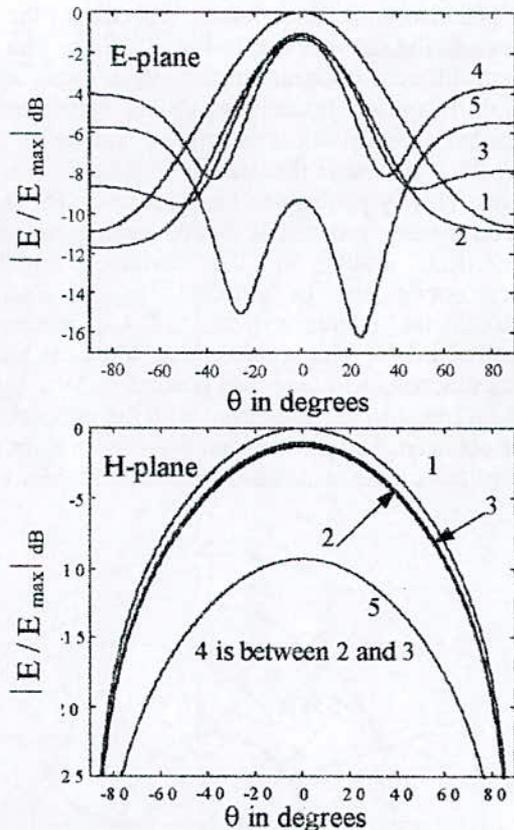


Fig. 18 Two dimensional field patterns of the three elements microstrip antenna.
 $L_1 = L_3 = 2.605$ cm, $L_2 = 2.7$ cm, $W = 3.9$ cm, $S_1 = S_2 = 0.165$ cm, $X_p = 3.22$ cm, $Y_p = 1.95$ cm, $\epsilon_r = 2.55$, $tg\delta = 0.002$.
 (1. $f = 3.33$, 2. $f = 3.38$, 3. $f = 3.41$, 4. $f = 3.44$, 5. $f = 3.47$) GHz.

has its elements positioned collinearly along the E-plane. This arrangement as shown in Fig. 19 is referred to as the E-plane. Both the near-zone and the far zone properties of this structure will be analyzed.

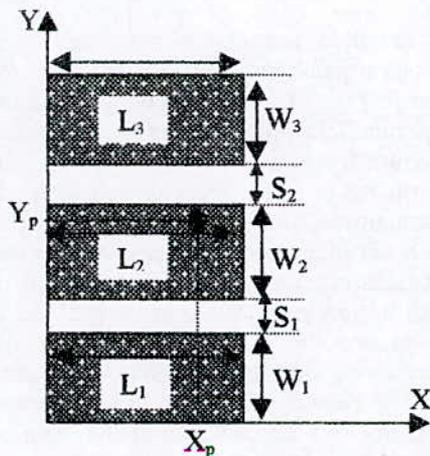


Fig. 19 Top view of the non-radiating edges gap-coupled antenna

Before the study of the obtained input impedance, it is worth to discuss the position by which the patches are disposed. It is a well known results [15] that surface

waves exist and propagate within the dielectric and their excitation is a function of the substrate thickness. In a given direction, the lowest order (dominant) surface mode is TM (odd) and TE (even) modes. For a rectangular microstrip patch, the fields are TM in direction of propagation along the H-plane. Since for the E-plane where the fields in the space between the elements are primarily TM, there is a stronger surface wave excitation (based on the single dominant surface wave mode) between the elements; and the coupling is larger. However for the H-plane arrangement, the fields in the space between the elements are primarily TE and there is not a strong dominant mode surface wave excitation; therefore there is less coupling between the elements. Hence, it is expected that, larger bandwidth will be obtained with the non-radiating edges gap-coupled structure than the radiating edges gap-coupled one.

The input impedance locus for the antenna described previously is shown in Fig. 20. The broadband behavior is clear by the formation of the loop on the smith chart. The theoretical bandwidth is found equal to 5.1%, which is 2.99 Times the bandwidth of the rectangular patch antenna. We see that, eventhough the patches are made of equal lengths, a considerable increase in the bandwidth is obtained and this is attributed to the placement of the patches as already stated.

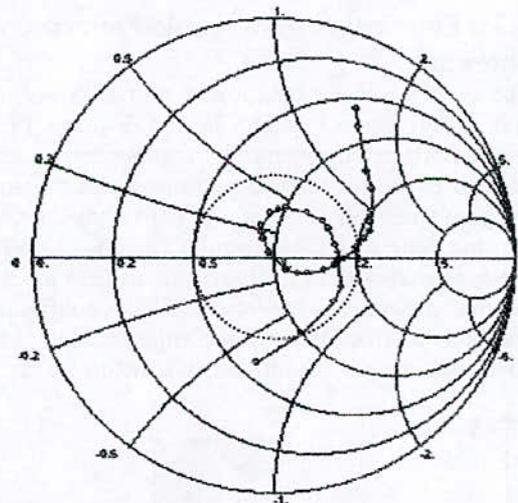


Fig. 4.20 Input impedance locus of the non-radiating edges gap coupled microstrip antenna
 $L_1 = L_2 = L_3 = 2.9$ cm, $W = 4$ cm, $S_1 = S_2 = 0.025$ cm
 $X_p = 1.61$ cm, $Y_p = 6.025$ cm, $\epsilon_r = 2.55$, $tg\delta = 0.002$.
 (From 2.8 GHz to 3.1 GHz)

Far field patterns of the microstrip patch antenna with two additional resonators non-radiating edges gap-coupled, for different frequencies are given in Fig. 4.21. Again, the form of patterns is changing with frequency, but this time in the H-plane as it is expected.

Remark
 important increase in the bandwidth is obtained (more than 20%) when the four additional resonators gap-coupled to the four edges configuration is considered. But the field are also varying with frequency.

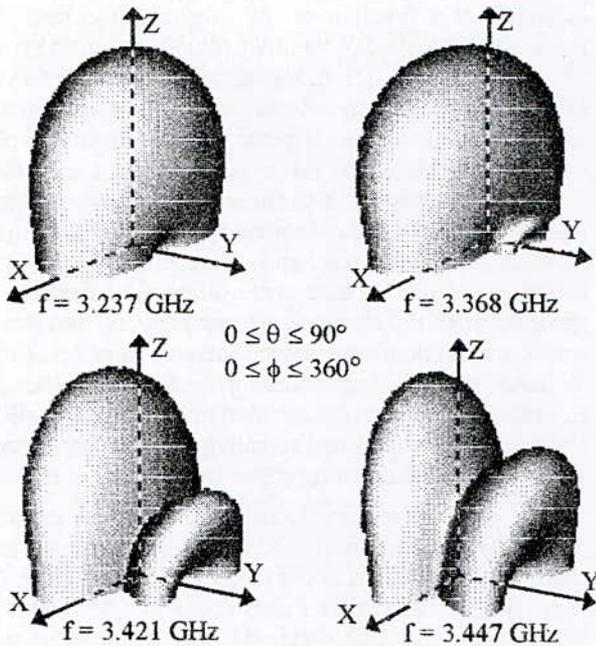


Fig. 21 Three dimensional field patterns of the non-radiating edges gap-coupled microstrip patch antenna $L_1 = L_2 = L_3 = 2.9$ cm, $W = 4$ cm, $S_1 = S_2 = 0.025$ cm, $X_p = 1.61$ cm, $Y_p = 6.025$ cm, $\epsilon_r = 2.55$, $\text{tg}\delta = 0.002$.

6.2.4 Four edges gap-coupled microstrip antenna

The surface waves excitation as well as space, dielectric and leaky waves either in the E-plane or H-plane configurations considered in the previous sections are used to excite one or two parasitic patches associated to the pilot patch. Since the waves exist altogether, then one can use four additional resonators gap-coupled to main patch as shown in Fig. 21 which will lead to an increase in the impedance bandwidth. This configuration, in addition to the bandwidth enhancement, keeps the coplanar nature of the microstrip structure.

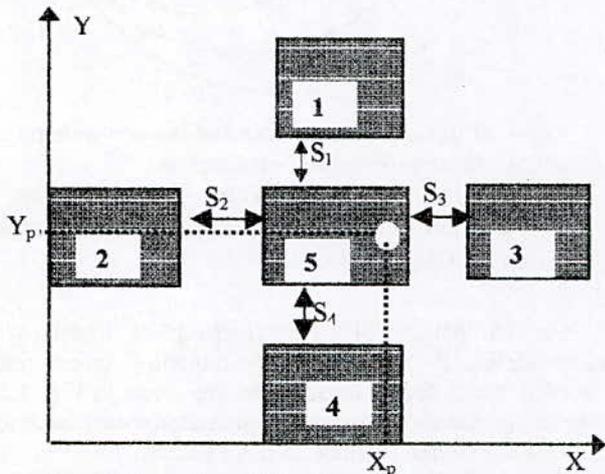


Fig. 22 Top view of the four resonators gap-coupled microstrip antenna. $L_1 = L_4 = 2.4$ cm, $L_2 = 2.49$ cm, $L_3 = 2.755$ cm, $L_5 = 2.9$ cm, $W_1 = W_2 = W_3 = W_4 = W_5 = 3.0$ cm, $h = 0.318$ cm, $X_p = 5.25$ cm, $Y_p = 4.52$ cm, $\epsilon_r = 2.55$ and $\text{tg}\delta = 0.003$, $S_1 = S_4 = 0.02$ Cm, $S_2 = 0.14$ Cm and $S_3 = 0.15$ Cm.

The input impedance versus frequency of the antenna shown in Fig. 22 is drawn in Fig. 23. Since four patches are of different resonant lengths, three loops are found, all of them are inside the $\text{VSWR} = 2$ circle. The impedance bandwidth is calculated and found equal to 20.1 %. It is to note that similar structure is analyzed by Kumar [15] by performing measurements. The results he obtained show some shift as compared to the obtained theoretical results in the current analysis. The discrepancies may be attributed to the a difference between the reference planes of the measured and calculated data. It is observed that when the number of basis functions was increased (from 3 to 5), a significant improvement in the agreement with the measured results was obtained. Still, further improvement is expected with a further increase in the number of basis functions.

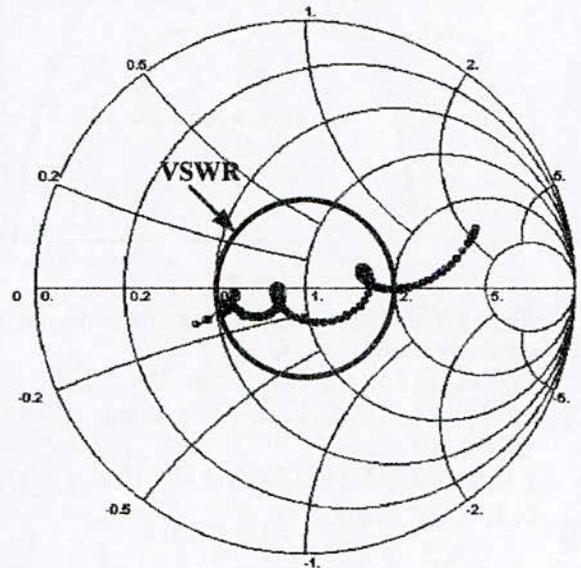


Fig. 23 Locus of the input impedance versus frequency (from 2.85 GHz to 3.6 GHz) of the four edges microstrip antenna.

The radiation patterns of the four edges gap-coupled microstrip patch antennas, for different frequencies, are given in Fig. 24. All the patterns are normalized to the maximum field value in order to emphasize the change in the power levels. We notice that the E-plane patterns vary in form and in power level as well as the direction of the maximum radiation shifts (due to the patches 2 and 3 which are of unequal lengths). As the parasitic patches start radiating, they lead to a creation of the phase error which in turn yields the beginning of the beam splitting. The patterns in H-plane present only variation in the power levels until the frequency where the patches 1 and 4 start to radiate; and at that time the beam start to split. We notice that the direction of the main beam in the H-plane do not shift as the frequency changes because of the fact that patches 1 and 4 are identical.

The cross-polarization components in the E-plane are practically null. However, in the H-plane the cross-polarization component at frequency of $f = 2.935897$ GHz has an important level. This may be attributed to the fact that the central is the only one which is radiating at the

considered frequency, and since its edges are nearly of equal lengths (2.9 cm × 3.0 Cm) then the E_y component has a significant value.

Remark

A further increase in the bandwidth is expected if the five patches are made of different lengths with propre choice of the gap-widths.

7. Conclusion:

The exact dyadic Green's function is used to derive an electric field integral equation with the patch current as an unknown for microstrip patch of arbitrary shape. The application of derived formalism is straightforward. The obtained currents as the moment method solution of the relevant integral equation are used as sources for the radiation from the structures. The obtained radioelectric properties of the antennas show that the single rectangular patch antenna is inherently narrow band-width structure and linear polarized in their broadside radiation. A broader bandwidth can be achieved if parasitic elements are added to the main patch. But the degradation in the radiation pattern field should not be ignored. And by making the patch of equal sides and with just single feed on the main diagonal leads to a circular polarized structure. Also in this analysis, it is shown that the full-wave technique can be used to analyze antennas working around 24 GHz.

During the analysis, it is noted that as the number of the patches increases with an increase in the number of the basis functions, the CPU time needed for the integrals evaluation is very demanding. As a further work, introduction of some analytical simplifications and other techniques in the way of performing the double integrals involved in the solution of the characterizing integral equation may lead to sparse matrices. And at that time, an optimization program is necessary to optimize the choice of the gap-widths and the parasitic resonant lengths for a broader band-width.

References

[1] Pozar D. M." considerations for millimeter-wave printed antennas," *IEEE Trans. Antenna and propagations* Vol. AP. 31, No. 5, Sept 1983, PP. 740 747
 [2] Seashore C. R. and D. R. Singh "Millimeter wave component trade-off for tactical systems" *microwave journal*, Vol. 25, No. 6, June 1982, PP. 41- 62
 [3] Deschamps G. A. "Microstrip microwave antennas", Third U.S.A.F. symposium on antennas, U.S.A. 1953.
 [4] Lewin L. "Radiation from discontinuities in Strip-line", *Proceedings of IEE*, U.K. Vol. 107c, pp. 163-170, 1960.
 [5] J.Q Howell, "Microstrip Antennas", *IEEE AP-S int. Symp. Dig.* 1972 Williamsburg Va, pp. 177-180, Dec. 1972.
 [6] R.E Munson, "Conformal microstrip antennas and microstrip phased arrays", *IEEE Trans. Antenna and propagations* Vol. AP. 22, No. 1, Jan. 1974, PP. 74-78.
 [7] Kai Chang "Handbook of microwave and optical components" Volume 1, John Willey and Sons 1989.
 [8] Sanford G. G. "Conformal mapping phased array for aircraft tests with ATS-6", *IEEE Trans. Antenna and propagations* Vol. AP. 26, 1978, PP. 642-646.

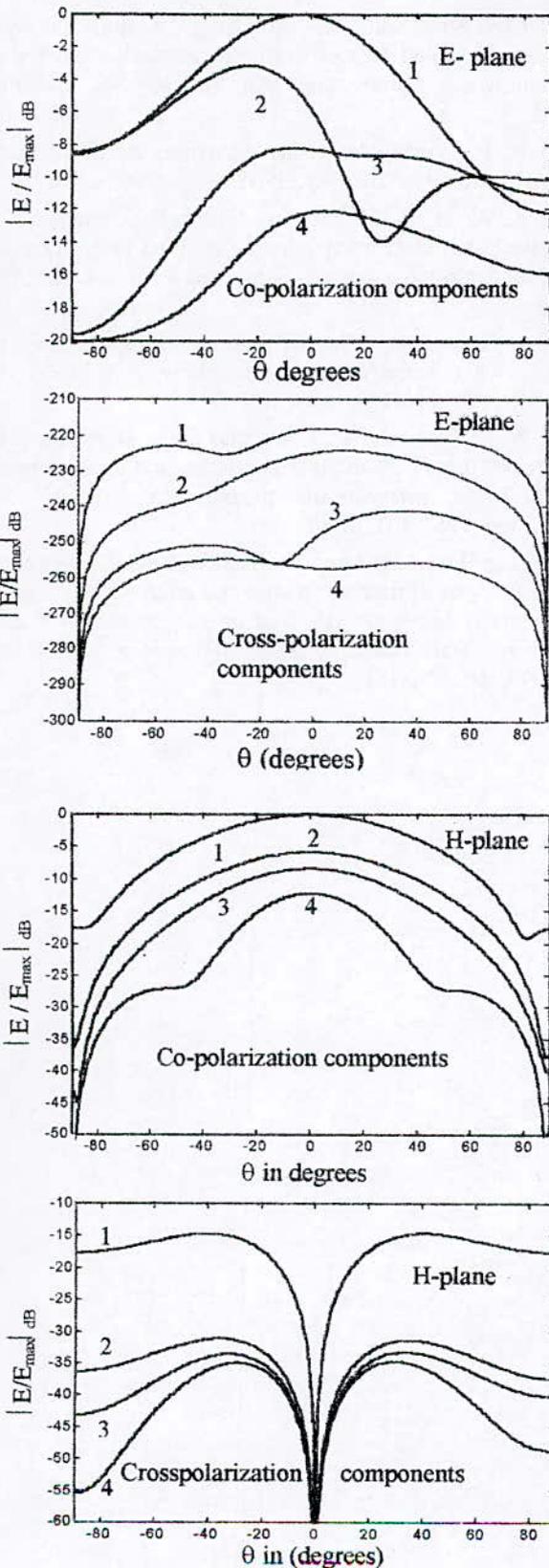


Fig. 24 Radiation patterns of the four edges gap-coupled patch antenna, for different frequencies.
 1. $f = 2.935897$ GHz, 2. $f = 3.151282$ GHz
 3. $f = 3.312821$ GHz. 4. $f = 3.536667$ GHz

- [9] Sterzer F., Paglione R. and Nowogroozski M. "Microwave apparatus for treatment of cancer by hyperthermia", *Microwave Journal*. USA 1980, Vol. 23, No. 1, pp. 39-44.
- [10] Agrawal P. K., Bailey M. C. "An analysis technique for microstrip antenna", *IEEE Trans. Antennas and propagation*, Vol. AP. 25, 1977, PP. 756-759.
- [11] Newman E. H., Tulyathan P. "Analysis of microstrip antenna using moment methods", *IEEE Trans. Antennas and propagation*, Vol. AP. 29, No. 1, 1981, PP. 47-53.
- [12] G wang " application of wavelets on the interval to the analysis of thin-wire antennas and scatsrers", *IEEE Trans. Antennas and propagation*, Vol. AP. 45, No. 5, may 1997.
- [13] R. Aksas "Analyse d'antenne microrubans de forme quelconque basée sur la technique des ligne de transmission. Application aux géométrie circulaires", *doctorat thesis at - Ecole National Polytechnique Alger*, 1995.
- [14] Rumsey, V. H. "Reaction concept in electromagnetic theory", *Phys. Rev.*, 1954, 94, (6), pp. 1483-1491.
- [15] Constantine A. Balanis, "Antenna theory analysis and design (second edition), John Wiley and Sons, In
- [16] Mosig T.R., Gardiol F.E. "Analytical and numerical techniques in the Green's functions treatment of microstrip antenna and scatterers, *IEE Proc* .vol . 130 , Pt .H (MOA) , No. 2 , 1983, pp. 175-182
- [17] R.F. Harrington "Matrix methods for field problems", *IEEE*, vol. 55, No.2, Februry 1967, pp. 136-149.
- [18] Pozar D. M. "Input impedance and mutual coupling of rectangular microstrip antennas", *IEEE Trans. Antennas and propagation*, Vol. AP. 30, 1981, PP. 1191-1196
- [19] K.C. Gupter M.D. Abouzahra " Analysis and design of planar Microwave components". pp 473-493 1994
- [20] J. M. Griffin and J. R. Forrest" Broadband circular disk microstrip antenna.", *Ellectron lett*. Vol. 18, 1982, pp. 267-269.
- [21] H. F. Pues and A. R. vande cappelle "An improved impedance matching technique for increasing the banndwidth of microstrip antenna", *IEEE Trans. Antennas and propagation*, Vol. AP. 37, No. 11, 1989, PP. 1345-1354.
- [22] C. K. Aanandan,P. Mohanan and K. J. Nair "Broad band gap coupled microstrip antenna", *IEEE Trans. Antennas and propagation*, Vol. 38, No. 10, oct 1990.
- [23] A. Sabban "New broad band stacked two layer microstrip antenna", *IEEE Trans. Antennas and propagation*, 1983, PP. 63-66.
- [24] J. Bennegueouche J. P. Damianoand A. Papernik "Antennes imprimées multicouche: choix des fonctions de base dans la méthode des moments", *J. phys. III France* 3, 1993, PP. 553-562.
- [25] G. Kumar and K. Gubta "Broad band microstrip antenna using additionnal resonators gap coupling to the radiation edges", *IEEE Trans. Antennas and propagation*, Vol. AP. 32, 1984, PP. 1375-1379.
- [26] G. Kumar and K. Gubta "Nonradiating edges and four gap-coupled multiple resonators broad band microstrip antennas", *IEEE Trans. Antennas and propagation*, Vol. AP. 33, No. 2, feb 1985.
- [27] Qing-Song and Xue-Xia Zhang, "A study on wide band gap coupled microstrip antenna arrays", *IEEE Trans. Antennas and propagation*, Vol. AP. 43, No. 3, march 1995.
- [28] R. F. Harrington "Time harmonic electromagnetic fields", *Mc Graw-Hill*, Ny, 1961.
- [29] A. Azrar an d R. Aksas, "Full wave analysis of a passive patch electromagnetically coupled to a probe fed rectangular patch antenna", Submitted to the COMAEI at Bejaia.
- [30] A. C. Ludwig, "The definition of crosspolarization", *IEEE Trans. Antennas and propagation*, AP. 21, No. 1, jan 1973, PP.116-119.
- [31] W. F. Richards, Y. T. Lo, and D. D. Harrison, "An improved theory of microstrip antenna and applications", *IEEE Trans. Antennas and propagation*, Vol. AP. 29, No.1, Jan. 1981, PP. 38-46.
- [32] Lale Hayirlioglu and M. Tuncay-Birand, "computer aided analysis of microstrip antennas employing parasitic elements", *International Journal of microwave and millimeter-wave computer aided engineering*, Vol. 3, No. 4, 1993, pp. 404-411.

Contribution à L'Etude d'Optimisation d'Antennes Plaques Microruban

S. SENOUCI *, A. ZERGUERRAS **

ملخص:

هذا العمل عبارة عن امتداد لطريقة خطوط الإرسال الجزئية المتصلة مغناطيسيا والمتناهية، لتصميم وتحديد خصائص الهوائيات المسطحة الميكروشرطية، أحادية، ثنائية و ثلاثية الطبقات. صيغة ثلاثية الطبقات أخذت هنا لأول مرة من زاوية صياغة جديدة تتكفل بالطبقة الواقية، وتغذية تردد واحد أو عدة ترددات. هذه الطريقة المؤهلة بحقل واسع من النتائج التجريبية على أحادية وثنائية الطبقات، قد استعملت في البحث عن الهوائيات المسطحة الميكروشرطية ذات الجداء "ج × ع" العالي، حيث "ج" هو نسبة الاستطاعة المقدمة للهوائي على الاستطاعة المبنوثة، و"ع" هو عرض نطاق الترددات التي يشتغل بها الهوائي. النتائج المحصل عليها في هذا البحث جد مرضية. وتستجيب لتوقعاتنا في إطار هذا العمل علما أن الدراسات لم تعمق بعد لا سيما في اتجاه البحث عن الشكل الخصوصي ذي عرض النطاق الترددي الواسع. نعتقد أن هذه الدراسات ستكون بدون شك مفيدة.

Abstract:

This work is an extension of the elementary finite coupled lines for characterizing and designing the microstrip antennas, with one, two and three layers. The case of three layers is studied here for the first time in a suitable formulation which takes into account the effect of the overlay and the single and multimodes excitation schemes. This method, which is validated with large number of experimental results in the case of one and two layers, is used to look for the microstrip antennas having a maximum gain-bandwidth product. The current obtained results are satisfactory. They fulfill the predictions of the intended work, provided that further investigations are needed especially in a specific form that yields larger bandwidth.

We believe that these investigations will be undoubtedly fruitful.

Résumé:

Ce travail est une extension de la méthode des Lignes Élémentaires Couplées Finies (L.E.C.F) pour la caractérisation et la conception des antennes plaques microruban (A.P.M.), mono, bi, et tri-couches. L'aspect tri-couche est pris ici pour la première fois dans une formulation nouvelle qui prend en charge une couche de protection « overlay », le radôme et l'excitation mono ou multimodes. Cette méthode validée sur un large éventail des résultats expérimentaux mono et bi-couches a été utilisée dans la recherche d'A.P.M. ayant le produit gain-largeur de bande maximal. Les résultats obtenus dans cette recherche sont très satisfaisants et répondent à notre attente dans le cadre de ce travail, tout en sachant que les investigations n'ont pas encore été poussées assez loin en particulier dans la recherche de forme spécifique à très large bande.

Nous pensons que la poursuite de ces investigations sera sans aucun doute très fructueuse.

Key words : Microstrip antenna, Input impedance, bandwidth, Gain, Radiation-pattern, Overlay.

* Article extrait de sa thèse de Magister

** Professeur à l'ENP Alger

1. INTRODUCTION

La méthode de la ligne de transmission utilisée comme outil d'analyse et de conception d'antennes plaques microruban (APM), initiée par le Professeur G. Dubost [5-7], est fondée sur l'équation différentielle de Riccati qui n'admet pas, en général, de solution analytique. Cependant, sachant qu'une application homographique sur la variable d'une équation de Riccati la transforme en une autre équation de Riccati, une résolution numérique devient possible. Cette méthode a été appliquée avec succès à des APM simples ou bi-couches de forme rectangulaire, circulaire, triangulaire, pentagonale, hexagonale et même n-Polygonales [1-3], [8].

En fait, la méthode, sous certaines conditions, s'applique à toute APM simple ou bi-couches de configuration quelconque mais admettant un axe de symétrie.

Les travaux ultérieurs qui ont conduit à une modélisation théorique du coefficient de couplage entre deux lignes stratifiées [4] et la réduction des temps de calcul sur ordinateur grâce à l'introduction de découpage en lignes élémentaires de type angulaire ou adapté ont permis au professeur A. Zerguerras de proposer une modélisation de la méthode des lignes élémentaires couplées finies (L.E.C.F) fondée toujours sur le principe de l'équation différentielle de Riccati, mais cette fois sur une matrice de transfert réciproque avec généralisation aux APM tri-couches comportant un second directeur, toujours de forme quelconque doté quand même d'un axe de symétrie. Cette méthode des LECF prend en compte, en plus, les cas d'APM avec overlay protecteur ou comportant une couche d'air nécessaire à l'élargissement de la bande passante grâce à l'abaissement de la valeur de la permittivité relative ϵ_r du substrat concerné.

Ce travail de thèse est justement une mise en œuvre de cette méthode des LECF pour en déterminer d'une part les performances et les limites de validité et pour l'utiliser, d'autre part,

au développement d'APM simples, bi-couches ou tri-couches dans le but de maximiser le produit caractéristique (G.B) du gain G par la largeur de bande B.

2. Performances des APM

La conception d'une antenne plaque microruban performante se réalise par l'observation des points suivants :

- l'accroissement de la largeur de bande des APM
- l'amélioration de la directivité et du gain de l'APM
- l'amélioration du rendement
- la sélection du substrat
- la réalisation des réseaux d'APM

3. Théorie des lignes élémentaires couplées finies (LECF)

Sachant que la superposition d'éléments parasites (directeurs) à l'élément pilote augmente à la fois la directivité, le gain et la largeur de bande, nous pouvons optimiser donc une structure stratifiée. Car la structure coplanaire à performance égale est trop étendue. Ayant validé la méthode des LECF dans les cas d'antennes simples et bi-couches, nous l'avons utilisée comme outil de conception de l'antenne tri-couches. Nous allons donc exposer cette méthode pour le cas tri-couches. Pour une antenne isolée (sans directeur) pourvue d'un axe de symétrie l'admittance réduite $Y_r(x)$ fonction de x , distance le long de l'axe de symétrie et comptée par rapport au point d'excitation o , est la solution de l'équation de Riccati (3-1).

$$g_c = \frac{G_c(x)}{G_0} \text{ est l'admittance caractéristique}$$

normalisée de la ligne de transmission équivalente qui varie avec x .

ϵ_e est la permittivité relative équivalente variant également avec x .

$H_r = h / \lambda_0$ est l'épaisseur constante du substrat diélectrique normalisée à la longueur d'onde dans le vide λ_0 .

$$\frac{dY_r(x)}{dx} + \left[\frac{4\pi^3}{5\lambda_0} \frac{H_r^2}{\sqrt{\epsilon_e}} g_c(x) + \frac{\pi}{\lambda_0} \sqrt{\epsilon_e} \left(\text{tg}\delta + \frac{D_s}{H_r} \right) + \frac{j2\pi}{\lambda_0} \sqrt{\epsilon_e} \right] \left[g_c(x) - \frac{Y_r^2(x)}{g_c(x)} \right] = 0 \quad (3-1)$$

La permittivité relative équivalente de la ligne ϵ_e peut s'apprécier à l'aide de la formule empirique suivante (3-2a) qui est approchée mais peut être améliorée suivant la procédure donnant (3-2b).

$$\epsilon_e = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left[1 + 12 \frac{H_r}{W_r(x)} \right]^{-1/2} \quad (3-2a)$$

W_r : est la largeur variable de la plaque antenne réduite par rapport à la longueur d'onde λ_0 .

$$\epsilon_e(u, \epsilon_r) = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} F(u, \epsilon_r) \quad (3-2b)$$

où $u = \frac{W}{h}$ et $F(u, \epsilon_r) = \left(1 + \frac{10}{u} \right)^{-a(u)b(\epsilon_r)}$ avec

$$a(u) = 1 + \frac{1}{49} \log \left[\frac{u^4 + \left(\frac{u}{52} \right)^2}{u^4 + 0.432} \right] + \frac{1}{18,7} \log \left[1 + \left(\frac{u}{18,1} \right)^3 \right]$$

$$b(\epsilon_r) = 0,564 \left(\frac{\epsilon_r - 0,9}{\epsilon_r + 3} \right)^{0,053}$$

la précision de ϵ_e calculée par (3-2b) est à mieux de 0,2 % pour $\epsilon_r \leq 128$; $0,01 < u < 100$.

La solution analytique de l'équation de Riccati (3-1) n'est en général pas possible sauf pour l'antenne plaque rectangulaire. Une application homographique sur la variable d'une équation de Riccati la transforme en une autre équation de Riccati. L'exploitation de cette propriété nous permet de décomposer notre antenne de forme quelconque en "tronçons" élémentaires rectangulaires perpendiculaires à son axe de symétrie. Ces tronçons élémentaires sont assimilés à des lignes de transmission de largeur $w(x)$, de longueur $\Delta x = L/N_i$, et d'épaisseur h . L : longueur de l'antenne selon l'axe des x .

N : nombre de découpage en tronçons élémentaires.

G_0 : admittance du vide, $G_0 = \sqrt{\frac{\epsilon_0}{\mu_0}}$.

Les ondes incidentes et réfléchies sur les extrémités d'un tronçon de ligne de transmission sont reliées par une matrice 2x2 appelée matrice

de transfert. Dans le cas de deux tronçons superposés la taille de ladite matrice est 4x4. La présence d'un deuxième directeur donne naissance à une structure tri-couches et par conséquent trois lignes de transmission superposées. Dans ce dernier cas la matrice de transfert a pour dimensions 6x6.

La structure tri-couches comprend trois régions différentes :

1. Région monocouche : région sans lignes couplées (1^{re} et 5^{me} région Fig. (3-2)).
2. Région bi-couches : région à deux lignes couplées (2^{me} et 4^{me} région fig. (3-2))
3. Région tri-couches : région à trois lignes couplées deux à deux (3^{me} région Fig. (3-2)).

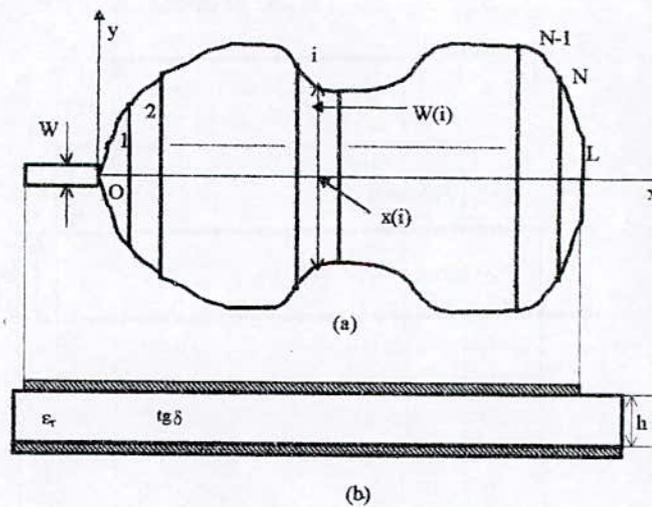
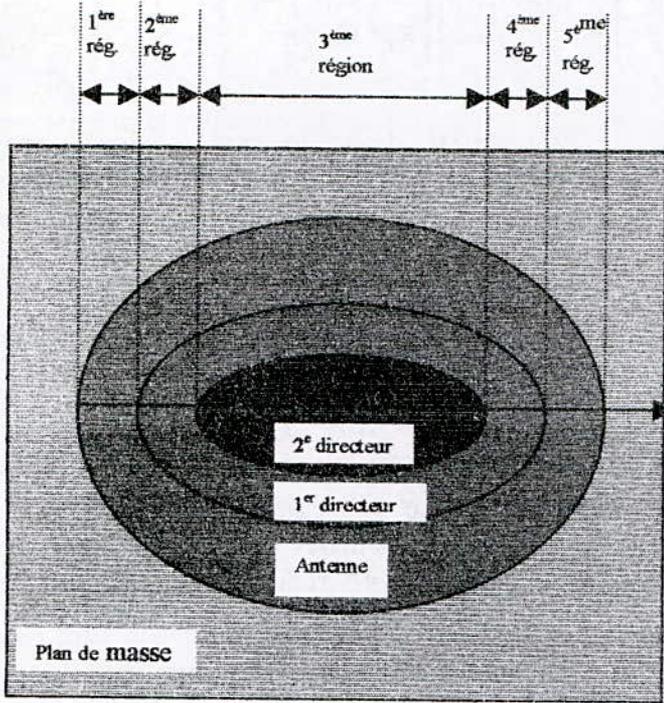
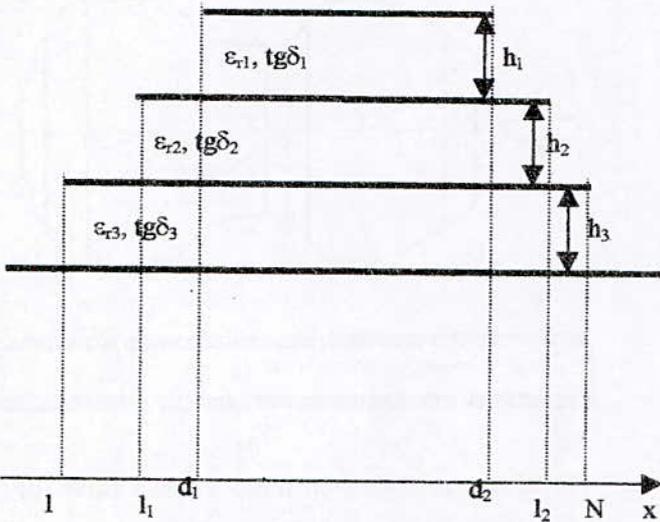


Fig. (3-1) : configuration d'une antenne isolée de forme arbitraire discrétisée en N sections élémentaires.

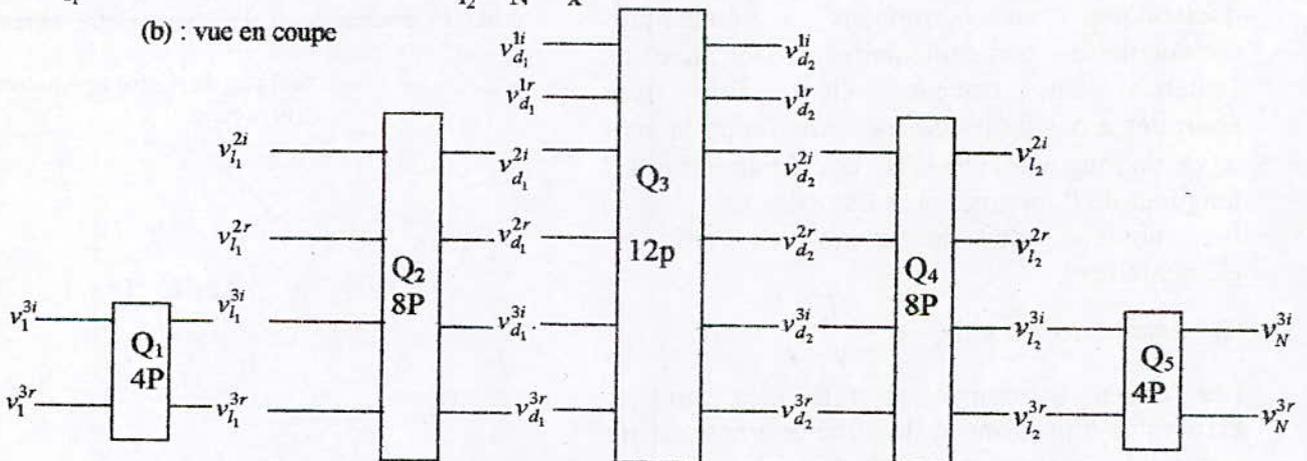
- (a) vue de dessus de l'antenne ;
(b) coupe.



(a) vue de dessus



(b) : vue en coupe



(c) Représentation de la formulation matricielle correspondante

La 1^{re} région forme un quadripôle Q_1 dont les tensions d'entrée et de sortie sont reliées par l'équation matricielle (3-3a).

$$\begin{bmatrix} v_1^{3i} \\ v_1^{3r} \end{bmatrix} = [T_1] \begin{bmatrix} v_{l_1}^{3i} \\ v_{l_1}^{3r} \end{bmatrix} \text{ avec } [T_1] = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \quad (3-3a)$$

$$A = \exp\left(\sum_{i=1}^{l_1} \gamma_{3i} \cdot \Delta x\right); \quad B = \exp\left(-\sum_{i=1}^{l_1} \gamma_{3i} \cdot \Delta x\right)$$

γ_{3i} = Coefficient de propagation du tronçon de ligne d'ordre i sur la plaque antenne (voir annexe).

La deuxième région est équivalente à un octopôle dont les tensions à l'entrée et à la sortie sont reliées par l'équation matricielle (3-3b).

$$\begin{bmatrix} v_{l_1}^{2i} \\ v_{l_1}^{2r} \\ v_{l_1}^{3i} \\ v_{l_1}^{3r} \end{bmatrix} = [T_2] \begin{bmatrix} v_{d_1}^{2i} \\ v_{d_1}^{2r} \\ v_{d_1}^{3i} \\ v_{d_1}^{3r} \end{bmatrix} \text{ avec } d_1 = l_1 + k_1 \quad (3-3b)$$

avec $[T_2] = [t_{l_1+1}^{o1}] [t_{l_1+2}^{o1}] \dots [t_{d_1}^{o1}]$, les $[t_n^{o1}]$ sont les matrices de transfert des octopôles élémentaires formant la 2^{me} région de la structure. En respectant la numérotation du haut vers le bas, les éléments t_{pqn}^{o1} de la matrice élémentaire d'ordre n sont définis comme suit :

$$\text{On pose : } d_{ijn} = \frac{1}{d_{jin}} = \sqrt{\frac{R_{cin}}{R_{cjn}}}; \quad i, j \text{ désignent le}$$

N° de la ligne de transmission, dans ce cas $i=2$ et $j=3$.

Fig (3-2) : Structure tri-couche

Introduisons l'impédance moyenne géométrique : $R_{cijn} = \sqrt{R_{cin} \cdot R_{cjn}}$

$$2Y_{ijn} = 2Y_{jin} = I_n \cdot \omega \cdot R_{ijn} \cdot C_{ijn} = \omega \cdot \tau_{ijn} \quad (3-4)$$

où $\tau_{ijn} = I_n \cdot R_{ijn} \cdot C_{ijn}$ est la constante de temps induite par la capacité distribuée sur le tronçon de longueur $I_n = \frac{L}{N_i} = \Delta x$,

$$C_{ijn} = k_{ijn} \epsilon_0 \cdot \sqrt{\epsilon_{ri} \epsilon_{rj} u_i(n) u_j(n)},$$

$$k_{ijn} = \left[\left(1 + \frac{\epsilon_{rj} u_j(n)}{\epsilon_{ri} u_i(n)} \right) \left(1 + \frac{\sum_{mi} S_{nmi}}{S_{nmi}} \right) \left(1 + \frac{\sum_{mj} S_{nmj}}{S_{nmj}} \right) \right]^{1/2},$$

$$u_i(n) = w_i(n) / h_i, \quad i = 1; 2; 3.$$

où S_{nmi} (S_{nmj}) représentent la surface non métallisée (i.e. : la surface latérale) des plaques i et j , et \sum_{mi} (\sum_{mj}) la surface métallisée.

Afin de renforcer une écriture symétrique des relations on posera :

$$D_{ijn} = \frac{1}{2} (d_{ijn} + d_{jin}) = \frac{1}{2} \left(d_{ijn} + \frac{1}{d_{ijn}} \right) \quad (3-5)$$

$$\text{et } S_{ijn} = \frac{1}{2} (d_{ijn} - d_{jin}) = \frac{1}{2} \left(d_{ijn} - \frac{1}{d_{ijn}} \right) \quad (3-6)$$

L'exposant de transfert de la ligne i au niveau de la tranche d'ordre n dont la définition dynamique précise est en annexe est :

$$\gamma_{in} = (\alpha_{in} + j \frac{2\pi}{\lambda_0} \sqrt{\epsilon_{ein}}) I_n \quad (3-7)$$

α_{in} : est le facteur d'atténuation dû aux pertes diélectriques, ohmiques et par rayonnement.

$$\alpha_{in} = \alpha_{inr} + \alpha_{ind} + \alpha_{inc}$$

$$\alpha_{inr} = \frac{4\pi^3}{5} \cdot \frac{R_0}{R_{in}} \cdot \frac{h_i^2}{\lambda_0^3 \sqrt{\epsilon_{ein}}} \quad (\text{pertes par rayonnement})$$

$$\alpha_{ind} = \frac{\pi}{\lambda_0} \text{tg} \delta_i \sqrt{\epsilon_{ein}} \quad (\text{pertes par conductivité diélectrique})$$

$$\alpha_{inc} = \frac{\pi}{\lambda_0} \cdot \frac{ds_i}{h_i} \cdot \sqrt{\epsilon_{ein}} \quad (\text{pertes par résistance ohmique sur le conducteur})$$

Posons pour le terme couplant de la ligne d'intérêt i à la ligne j :

$$P_{ijn}(\gamma_{in}) = Y_{ijn} [\text{ch}(\gamma_{in}) + d_{ijn} \text{sh}(\gamma_{in})] \quad (3-8)$$

et pour le terme propre de la ligne i , déséquilibrée dans sa symétrie par la présence de la ligne j quand elle lui est différente :

$$\psi_{ijn}(\gamma_{in}) = \text{ch}(\gamma_{in}) + D_{ijn} \text{sh}(\gamma_{in}) \quad (3-9a)$$

Pour les lignes identiques : $d_{ijn} = 1 \Rightarrow$

$$D_{ijn} = 1, S_{ijn} = 0;$$

$$\text{et } \frac{P_{ijn}(\gamma_{in})}{Y_{ijn}} = \psi_{ijn}(\gamma_{in}) = e^{\gamma_{in}}$$

$$\frac{P_{ijn}(-\gamma_{in})}{Y_{ijn}} = \psi_{ijn}(-\gamma_{in}) = e^{-\gamma_{in}}$$

Les 16 éléments de la matrice $[t_n^{o1}]$ sont alors donnés par :

Les quatre éléments de la première diagonale :

$$t_{11n}^{o1} = \psi_{23n}(\gamma_{2n}) + jP_{23n}(\gamma_{2n}) = A_{23n}(\gamma_{2n});$$

$$t_{22n}^{o1} = \psi_{23n}(-\gamma_{2n}) - jP_{23n}(-\gamma_{2n}) = A_{23n}^*(-\gamma_{2n});$$

$$t_{33n}^{o1} = \psi_{32n}(\gamma_{3n}) + jP_{32n}(\gamma_{3n}) = A_{32n}(\gamma_{3n});$$

$$t_{44n}^{o1} = \psi_{32n}(-\gamma_{3n}) - jP_{32n}(-\gamma_{3n}) = A_{32n}^*(-\gamma_{3n}).$$

A_{23n} et A_{32n}^* ne sont pas conjugués l'un de l'autre puisque γ_{in} est complexe. Nous appelons cela une conjugaison impropre.

Les quatre couplages d'ondes progressives et régressives de même ligne (termes adjacents à la première diagonale) :

$$t_{12n}^{o1} = -S_{23n} \text{sh}(\gamma_{2n}) + jP_{23n}(\gamma_{2n}) = B_{23n}(\gamma_{2n});$$

$$t_{21n}^{o1} = -S_{23n} \text{sh}(-\gamma_{2n}) - jP_{23n}(-\gamma_{2n}) = B_{23n}^*(-\gamma_{2n});$$

$$t_{34n}^{o1} = -S_{32n} \text{sh}(\gamma_{3n}) + jP_{32n}(\gamma_{3n}) = B_{32n}(\gamma_{3n});$$

$$t_{43n}^{o1} = -S_{32n} \text{sh}(-\gamma_{3n}) - jP_{32n}(-\gamma_{3n}) = B_{32n}^*(-\gamma_{3n}); \quad B_{ijn} \text{ et } B_{ijn}^* \text{ sont des conjugués impropres.}$$

En général :

$$B_{ijn}(\gamma_{in}) = -S_{ijn} \text{sh}(\gamma_{in}) + jP_{ijn}(\gamma_{in}) \quad (3-9b)$$

Les quatre couplages d'ondes interligne, de la ligne 3 à la ligne 2 :

$$t_{13n}^{o1} = t_{14n}^{o1} = -jP_{23n}(\gamma_{2n}) = E_{23n}(\gamma_{2n}) ;$$

$$t_{23n}^{o1} = t_{24n}^{o1} = jP_{23n}(-\gamma_{2n}) = E_{23n}^*(-\gamma_{2n}) ;$$

Les quatre couplages d'ondes interligne, de la ligne 2 à la ligne 3 :

$$t_{31n}^{o1} = t_{32n}^{o1} = -jP_{32n}(\gamma_{3n}) = E_{32n}(\gamma_{3n}) ;$$

$$t_{41n}^{o1} = t_{42n}^{o1} = jP_{32n}(-\gamma_{3n}) = E_{32n}^*(-\gamma_{3n}) ; \quad E_{ijn} \text{ et } E_{ijn}^*$$

sont des conjugués impropres.

Les ondes incidentes et réfléchies à l'entrée et à la sortie de la région 3 sont reliées par l'équation matricielle (3-10). Appelons le multiport formé par les trois lignes superposées « 12p ».

$$\begin{bmatrix} v_{d_1}^{1i} \\ v_{d_1}^{1r} \\ v_{d_1}^{2i} \\ v_{d_1}^{2r} \\ v_{d_1}^{3i} \\ v_{d_1}^{3r} \end{bmatrix} = [T_3] \begin{bmatrix} v_{d_2}^{1i} \\ v_{d_2}^{1r} \\ v_{d_2}^{2i} \\ v_{d_2}^{2r} \\ v_{d_2}^{3i} \\ v_{d_2}^{3r} \end{bmatrix} \quad \text{avec } d_2 = d_1 + k_2 \quad (3-10)$$

$$\text{avec } [T_3] = [t_{d_1+1}^{12p}] [t_{d_1+2}^{12p}] \cdot [t_{d_2}^{12p}]$$

les éléments des matrices élémentaires $[t_n^{12p}]$ sont définis comme suit :

$$[t_n^{12p}] = \begin{bmatrix} [t_{aa1n}(\gamma_{1n})] & [t_{ab12n}(\gamma_{1n})] & [t_{ab13n}(\gamma_{1n})] \\ [t_{ab21n}(\gamma_{2n})] & [t_{aa2n}(\gamma_{2n})] & [t_{ab23n}(\gamma_{2n})] \\ [t_{ab31n}(\gamma_{3n})] & [t_{ab32n}(\gamma_{3n})] & [t_{aa3n}(\gamma_{3n})] \end{bmatrix}$$

avec Si l'on pose :

$$\Psi_{aain}(\gamma_{in}) = [\Psi_{ijn}(\gamma_{in}) + \Psi_{ikn}(\gamma_{in})] / 2 \quad (3-11)$$

$$\text{avec } i=1, \quad j=2, \quad k=3$$

$$i=2, \quad j=3, \quad k=1$$

$$i=3, \quad j=1, \quad k=2$$

$$P_{aain}(\gamma_{in}) = P_{ijn}(\gamma_{in}) + P_{ikn}(\gamma_{in}) \quad (3-12)$$

$$B_{aain}(\gamma_{in}) = B_{ijn}(\gamma_{in}) + B_{ikn}(\gamma_{in}) \quad (3-13)$$

P_{ijn} , B_{ijn} sont définies par les équations (3-8) et (3-9b). Les matrices 2 x 2 constituant la matrice $[t_n^{12p}]$ sont définies comme suit :

$$[t_{aain}(\gamma_{in})] = \begin{bmatrix} \Psi_{aain}(\gamma_{in}) + jP_{aain}(\gamma_{in}) & B_{aain}(\gamma_{in}) \\ B_{aain}^*(-\gamma_{in}) & \Psi_{aain}(-\gamma_{in}) - jP_{aain}(-\gamma_{in}) \end{bmatrix}, \quad i = 1 ; 2 ; 3. \quad (3-14)$$

* : désigne le conjugué impropre.

$$[t_{abijn}(\gamma_{in})] = \begin{bmatrix} E_{ijn}(\gamma_{in}) & E_{ijn}(\gamma_{in}) \\ E_{ijn}^*(-\gamma_{in}) & E_{ijn}^*(-\gamma_{in}) \end{bmatrix}, \quad i, j=1 ; 2 ; 3 \text{ avec } i \neq j \quad (3-15)$$

Les ondes incidentes et réfléchies sur l'entrée et la sortie de la région 4 sont reliées par l'équation matricielle (3-16).

$$\begin{bmatrix} v_{d_2}^{2i} \\ v_{d_2}^{2r} \\ v_{d_2}^{3i} \\ v_{d_2}^{3r} \end{bmatrix} = [T_4] \begin{bmatrix} v_{l_2}^{2i} \\ v_{l_2}^{2r} \\ v_{l_2}^{3i} \\ v_{l_2}^{3r} \end{bmatrix} \quad \text{avec } l_2 = d_2 + k_2' \quad (3-16)$$

avec $[T_4] = [t_{d_2+1}^{o2}] [t_{d_2+2}^{o2}] \cdot [t_{l_2}^{o2}]$, les matrices élémentaires $[t_n^{o2}]$ sont construites de la même manière que les $[t_n^{o1}]$. Les tensions incidentes et réfléchies à l'entrée et à la sortie de la 5^{me} région sont reliées entre elles par l'équation matricielle (3-17).

$$\begin{bmatrix} v_{l_2}^{3i} \\ v_{l_2}^{3r} \end{bmatrix} = [T_1] \begin{bmatrix} v_N^{3i} \\ v_N^{3r} \end{bmatrix} \quad (3-17)$$

$$\text{avec } [T_1] = \begin{bmatrix} A_1 & 0 \\ 0 & B_1 \end{bmatrix}$$

$$A_1 = \exp\left(\sum_{i=l_2+1}^N \gamma_{3i} \Delta x\right) ;$$

$$B_1 = \exp\left(-\sum_{i=l_2+1}^N \gamma_{3i} \Delta x\right)$$

Le système matriciel donne 6 équations, auxquelles s'ajoutent 5 conditions aux limites et une condition d'excitation. Autrement dit on a 12 équations pour déterminer 12 inconnues les 6 tensions incidentes (3 à chaque extrémité) et les 6 tensions réfléchies correspondantes.

$$v_N^{3i} = v_N^{3r} \quad (3-18)$$

$$v_{l_2}^{2i} = v_{l_2}^{2r} \quad (3-19)$$

$$v_{d_2}^{1i} = v_{d_2}^{1r} \quad (3-20)$$

$$v_{d_1}^{1i} = v_{d_1}^{1r} \quad (3-21)$$

$$v_{l_1}^{2i} = v_{l_1}^{2r} \quad (3-22)$$

En choisissant par exemple le courant d'excitation égal à 1 Ampère, on obtient la 12^{me} équation scalaire. Si l'excitation est par ligne microruban on a $\frac{v_1^{3i} - v_1^{3r}}{Z_c^{31}} = 1A$ (3-23)

Z_c^{31} est l'impédance caractéristique de la ligne de transmission correspondant à la première tranche sur la plaque antenne (voir annexe). Ainsi nous sommes en mesure de déterminer les ondes incidentes et réfléchies sur les extrémités de toutes les lignes de transmission élémentaires sur les trois plaques et par conséquent nous connaissons les coefficients de réflexion sur toutes les interfaces entre tranches. Ceci nous permettra de déterminer l'impédance d'entrée de l'antenne à l'endroit voulu. Si l'antenne est excitée par ligne microruban, le coefficient de réflexion à son entrée est :

$$\Gamma = \frac{v_1^{3r}}{v_1^{3i}} \quad (3-24)$$

L'impédance d'entrée est :

$$Z_{in} = Z_c^{31} \frac{1 + \Gamma}{1 - \Gamma} \quad (3-25)$$

Nous pouvons déduire le taux d'ondes stationnaires et les diagrammes de rayonnement.

Le courant de conduction au niveau de la tranche d'ordre n sur la plaque j est donné par :

$$I_{cn}^j = \frac{v_n^{ji} - v_n^{jr}}{Z_{cin}} \quad (3-26)$$

i désigne tension incidente, r désigne tension réfléchi.

Le courant de polarisation au niveau de la tranche d'ordre n sur la plaque j est donné par :

$$I_{pn} = j\omega\epsilon_0(\epsilon_j - 1)w_n^j [v_n^{ji} + v_n^{jr}] \quad (3-27)$$

ϵ_0 : permittivité du vide

w_n^j : largeur de la tranche d'ordre n sur la plaque j.

Connaissant les distributions des courants de conduction et de polarisation, nous pouvons calculer le gain de l'antenne et en appliquant le théorème d'équivalence en volume nous pouvons établir les expressions du champ rayonné d'où les diagrammes de rayonnement.

Lorsque la plaque supérieure est plus grande que l'intermédiaire et celle-ci plus grande que l'inférieure, la situation se présente comme suit :

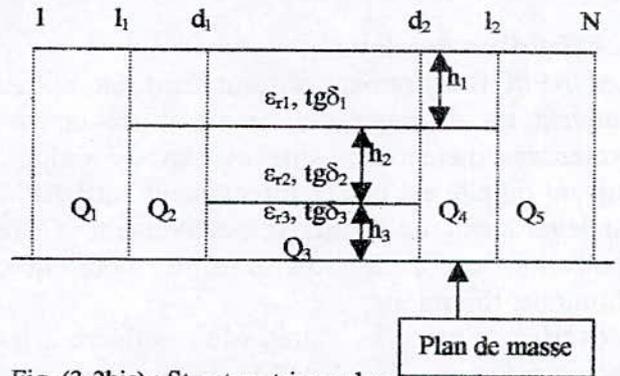


Fig. (3-2bis) : Structure tri-couches

Les lignes de transmission des régions N° 1et 5 formant les quadripôles Q1 et Q5 reposent sur le substrat d'épaisseur équivalente h_{eq1} et de permittivité relative équivalente ϵ_{req1} :

$$h_{eq1} = h_1 + h_2 + h_3$$

$$\epsilon_{req1} = \frac{h_{eq1}}{\frac{h_1}{\epsilon_{r1}} + \frac{h_2}{\epsilon_{r2}} + \frac{h_3}{\epsilon_{r3}}}$$

Les régions N° 2et 4 correspondantes aux octopôles Q2 et Q4 sont formées par deux lignes de transmission superposées. Les paramètres de la ligne inférieure sont h_{eq2} , et ϵ_{req2} où :

$$h_{eq2} = h_2 + h_3$$

$$\epsilon_{req2} = \frac{h_{eq2}}{\frac{h_2}{\epsilon_{r2}} + \frac{h_3}{\epsilon_{r3}}}$$

Le reste demeure inchangé par rapport au cas de la plaque inférieure plus grande. Les ondes incidentes et réfléchies aux interfaces (Q1-Q2) et (Q2-Q3) se divisent conformément à la figure (3-2bis) comme il suit :

$$\left\{ \begin{array}{l} V_{l_1}^{1i} = \frac{h_1/\epsilon_{r1}}{h_{eq1}/\epsilon_{req1}} V_{l_1-1}^{1i} \\ V_{l_1}^{1r} = \frac{h_1/\epsilon_{r1}}{h_{eq1}/\epsilon_{req1}} V_{l_1-1}^{1r} \end{array} \right. \quad \left\{ \begin{array}{l} V_{l_1}^{2i} = \frac{h_{eq2}/\epsilon_{req2}}{h_{eq1}/\epsilon_{req1}} V_{l_1-1}^{2i} \\ V_{l_1}^{2r} = \frac{h_{eq2}/\epsilon_{req2}}{h_{eq1}/\epsilon_{req1}} V_{l_1-1}^{2r} \end{array} \right.$$

de même pour les ondes à l'interface Q₂-Q₃ :

$$\left\{ \begin{array}{l} V_{d_1}^{2i} = \frac{h_2/\epsilon_{r2}}{h_{eq2}/\epsilon_{req2}} V_{d_1-1}^{2i} \\ V_{d_1}^{2r} = \frac{h_2/\epsilon_{r2}}{h_{eq2}/\epsilon_{req2}} V_{d_1-1}^{2r} \end{array} \right. \quad \left\{ \begin{array}{l} V_{d_1}^{3i} = \frac{h_3/\epsilon_{r3}}{h_{eq2}/\epsilon_{req2}} V_{d_1-1}^{3i} \\ V_{d_1}^{3r} = \frac{h_3/\epsilon_{r3}}{h_{eq2}/\epsilon_{req2}} V_{d_1-1}^{3r} \end{array} \right.$$

4. Effet d'un overlay :

Les APM fonctionnant surtout dans un milieu pouvant les dégrader ont besoin d'une couche protectrice diélectrique dite overlay ou radôme suivant qu'elle est placée directement sur l'APM ou légèrement au-dessus respectivement. Cette protection est à la fois d'ordre mécanique, chimique, thermique,...

L'overlay, par sa présence, affecte les caractéristiques de l'APM telles que : la fréquence de résonance, l'impédance d'entrée, les diagrammes de rayonnement etc. Compte tenu de la sélectivité de l'APM, il est important de contrôler tout décalage de la fréquence de résonance pour la maintenir dans la bande de fonctionnement.

Les premiers travaux concernant l'effet de l'overlay sont dus à Bahl [48], [49]. La majorité des théories pour analyser les structures stratifiées avec overlay ou radôme sont fondées sur la méthode des moments Galerkin (M₂ G - DTF) dans le domaine de transformé de Fourier. Ici l'approche adoptée est originale en ce sens qu'après avoir traité rigoureusement l'APM rectangulaire, elle en étend les résultats en se fondant sur les techniques LECF (lignes élémentaires couplées finies) qui lui permettent de traiter les APM de forme quelconque simples ou stratifiées.

Après avoir exposé une formulation du cas statique. On introduit l'overlay dont l'effet sera déterminé par une expression variationnelle de la capacitance. Ensuite une expression empirique sera proposée pour prendre en charge cet effet en tenant compte, à la fois, de l'épaisseur du ruban central et de la nature dispersive du diélectrique.

Du fait que le traitement analytique des limites multiples est plus aisé avec la méthode variationnelle qu'avec la transformation conforme ou autres méthodes numériques, l'effet de l'overlay sur le microruban sera analysé par la méthode variationnelle [49].

L'impédance caractéristique Z et la vitesse de phase v_p de la ligne de transmission fonctionnant en mode TEM sont : $Z = Z_0 / \sqrt{\epsilon_e}$

$$\text{et } v_p = c / \sqrt{\epsilon_e}$$

avec $Z_0 = 1 / C_0 c$, $\epsilon_e = C / C_0$, en outre $\lambda = \lambda_0 / \sqrt{\epsilon_e} = \lambda_0 \sqrt{C_0 / C}$

où C et C_0 sont les capacitances de la ligne avec et sans diélectrique respectivement ; ϵ_e est la constante diélectrique effective qui prend en compte l'effet du champ de bord dans le substrat, la feuille matériel et l'espace libre. c est la vitesse de la lumière dans le vide.

La capacitance, en mode quasi-TEM est déterminée par la méthode variationnelle. Dans le domaine transformé de Fourier (DTF), les conditions aux limites et de continuité donnent [50] pour le potentiel statique $\phi(x,y)$ qui satisfait l'équation de Poisson :

$$\Delta\phi(x,y) = -\frac{1}{\epsilon} \rho(x,y) \quad (4-1)$$

où $\rho(x,y)$ est la distribution de la charge sur la surface du ruban.

$$\bar{\phi} = \iint_s \frac{\rho(x,y)}{Q} \phi(x,y) dx dy, \quad Q \equiv \iint_s \rho(x,y) dx dy$$

$$C = Q / \bar{\phi} \quad (4-2)$$

C : est l'expression variationnelle de la capacitance. Pour un ruban infiniment mince :

$$\rho(x,y) = f(x)\delta(y-b) \text{ où } \delta(y-b) \text{ est la fonction delta de Dirac.}$$

La transformé de Fourier de $\phi(x,y)$ est :

$$\tilde{\phi}(\beta, y) = \int_{-\infty}^{+\infty} \phi(x, y) e^{i\beta x} dx \quad (4-3)$$

$$\tilde{\phi}(-\beta, y) = \tilde{\phi}(\beta, y) \text{ puisque } \phi(x, y) = \phi(-x, y) \quad (4-4)$$

pour $y \neq h$ on a, en transformant l'équation (4-1) par le biais de (4-3) :

$$(-\beta^2 + \frac{d^2}{dy^2})\tilde{\phi}(\beta, y) = 0$$

$$\text{et } \tilde{\phi}(\beta, 0) = \tilde{\phi}(\beta, \infty) = 0 \quad (4-5)$$

$$\tilde{\phi}(\beta, h+0) = \tilde{\phi}(\beta, h-0) = 0 \quad (4-6)$$

$$\epsilon_{ro} \frac{d}{dy} \tilde{\phi}(\beta, h+0) = \epsilon_r \frac{d}{dy} \tilde{\phi}(\beta, h-0) - \frac{\tilde{f}(\beta)}{\epsilon_0} \quad (4-7)$$

$$\tilde{\phi}(\beta, h+d+0) = \tilde{\phi}(\beta, h+d-0) \quad (4-8)$$

$$\frac{d}{dy} \tilde{\phi}(\beta, h+d+0) = \epsilon_{ro} \frac{d}{dy} \tilde{\phi}(\beta, h+d-0) \quad (4-9)$$

où

$$\tilde{f}(\beta) = \int_{-\infty}^{+\infty} f(x)e^{-j\beta x} dx ; \quad \tilde{f}(0) = \int_{-\infty}^{+\infty} f(x) dx ;$$

$$\tilde{f}(\beta)_{normalisée} = \tilde{f}_n(\beta) = \frac{\tilde{f}(\beta)}{\tilde{f}(0)} \quad (4-10)$$

$$\frac{1}{C} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{f}_n(\beta) \tilde{\phi}_n(\beta, h) d\beta ; \quad \text{avec } \tilde{\phi}_n(\beta, h) = \frac{\tilde{\phi}(\beta, h)}{\tilde{f}(0)} \quad (4-11)$$

Pour ce cas il vient :

$$\frac{1}{C} = \frac{1}{\pi\epsilon_0} \int_0^{\infty} \frac{[\tilde{f}_n(\beta)]^2 d(\beta h)}{\left[\frac{\epsilon_{ro} th\beta d + 1/\epsilon_{ro}}{1 + (th\beta d)/\epsilon_{ro}} + \epsilon_r coth \beta h \right] (\beta h)} \quad (4-12)$$

En adoptant une distribution de charge

$$f(x) = \begin{cases} 1 + \left| \frac{2x}{w} \right|^3 & \text{si } |x| \leq w/2 \\ 0 & \text{ailleurs} \end{cases} \quad (4-13)$$

De (4-10) et (4-13) avec $\varphi = \frac{\beta w}{2} = \frac{\beta h}{2} \cdot \frac{w}{h} = \frac{v u}{2}$

où ($v = \beta h$, $u = w/h$), il vient :

$$\tilde{f}_n(\beta) = \frac{\tilde{f}(\beta)}{\tilde{f}(0)} = 1.6 \sin c\varphi + \frac{2.4}{\varphi^2} \left[\cos \varphi - 2 \sin c\varphi + \sin c^2\left(\frac{\varphi}{2}\right) \right] \quad (4-14)$$

Avec (4-14) dans (4-12) l'intégrale peut être évaluée numériquement. Ainsi la constante diélectrique effective d'une ligne microruban couverte d'un overlay est déterminée. Les figures qui suivent montrent les courbes de la constante diélectrique effective en fonction de l'épaisseur de l'overlay pour quelques rapports w/h pour le cas $\epsilon_r = \epsilon_{ro} = 2.5$.

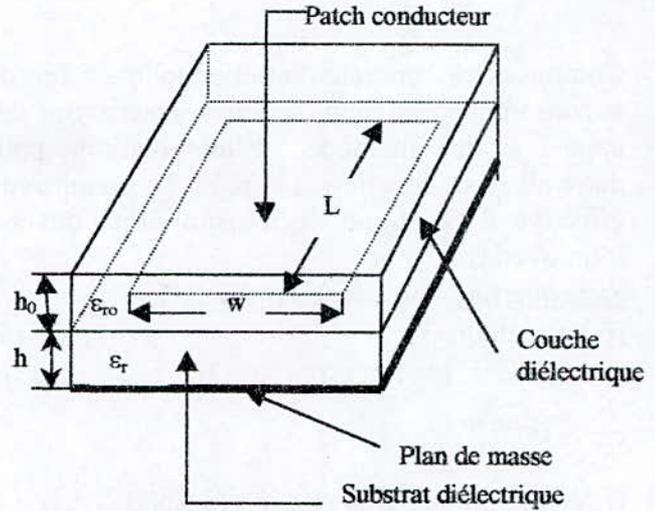
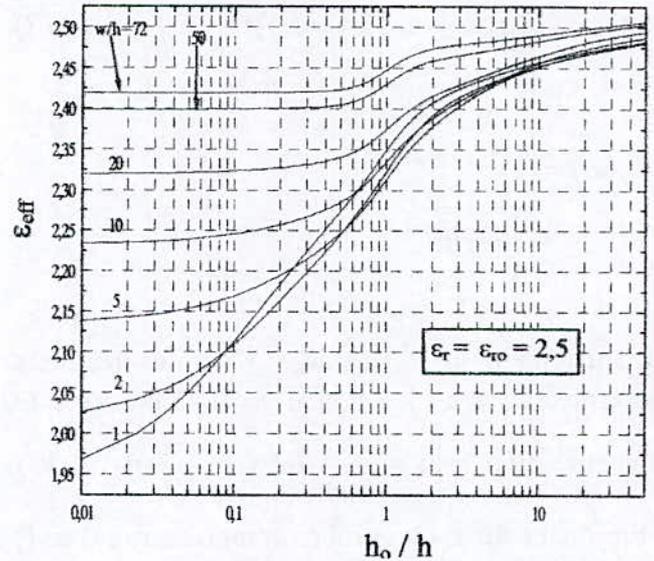
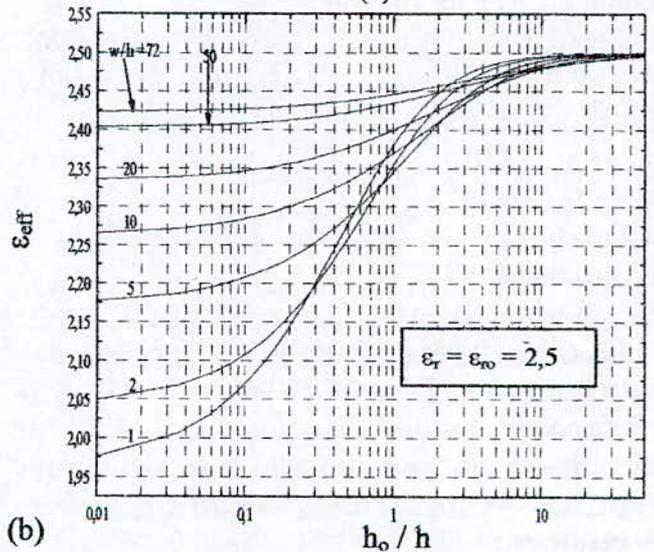


Fig.(4-1) : Configuration de l'antenne plaque microruban avec overlay.



(a)



(b)

Fig. (4-2) : La constante diélectrique effective d'une ligne microruban couverte par une couche diélectrique en fonction de l'épaisseur de la couverture. a) d'après (4-20), b) calcul selon la relation (4-12)

Comme cette approche variationnelle est lourde et consomme beaucoup de temps, nous avons fait appel à la méthode d'interpolation pour déterminer la fonction qui régit la permittivité effective d'une ligne de transmission couverte d'un overlay.

Sans overlay : on a $\epsilon_{re}(u, \epsilon_r) = F_+ \cdot \epsilon_r + F_-$
 $(F_-)(\epsilon_r \coth \beta h + 1)$ (4-15)

avec $2F_+ = 1 + F(u, \epsilon_r)$; $2F_- = 1 - F(u, \epsilon_r)$;
 $\frac{F_-}{F_+} = th \beta h$.

F : étant définie par la relation (3-2bis)

Avec overlay (u_o, ϵ_{ro}): on a
 $\epsilon_{re}(u, \epsilon_r, u_o, \epsilon_{ro}) = F_+ \cdot \epsilon_r + F_- \cdot G_o(u_o, \epsilon_{ro})$ (4-16)

où $G_o(u_o, \epsilon_{ro}) = \epsilon_{ro} [1 - G_o(u_o)] + G_o(u_o)$ (4-17)

$$G_o(u_o) = \begin{cases} \frac{1}{\sqrt{1 + \frac{12}{u_o}}} + 0.04(1 - u_o)^2 & \text{si } u_o = \frac{w}{d} \leq 1 \\ 1 & \text{si } u_o = \frac{w}{d} \geq 1 \end{cases}$$
 (4-18)

L'ajustement de $\epsilon_{re}(u, \epsilon_r, u_o, \epsilon_{ro})$ sur les résultats intermédiaires à ($d = 0$ et $d = \infty$) se fait avec le

facteur $\psi(u, u_o) = \left[1 + \frac{r}{(2\pi)^2} th \left[s \left(\frac{u}{u_o} - \frac{u_o}{u} \right) \right] \right]$ (4-19)

En jouant sur r, un nombre compris entre 0 et 1, en général r est voisin de 1 / 2, et sur s un nombre voisin de 10. Soit :

$\epsilon_{rea} = \psi(u, u_o) \epsilon_{re}(u, \epsilon_r, u_o, \epsilon_{ro})$ (4-20)

Pour $0 < r < 1$ on a $1 - 1 / (2\pi)^2 < \psi < 1 + 1 / (2\pi)^2$ soit $0.975 < \psi < 1.025$. Si $d \rightarrow 0$;

$u_o \rightarrow \infty$; $G_o(u_o) \rightarrow 1$; $G_o(u_o, \epsilon_{ro}) \rightarrow 1$;

$\psi \rightarrow 1 - r / (2\pi)^2$ et

$\epsilon_{re}(u, \epsilon_r, u_o, \epsilon_{ro}) \rightarrow \epsilon_{re}(u, \epsilon_r) = F_+ \cdot \epsilon_r + F_- =$

$(F_-)(\epsilon_r \coth \beta h + 1)$

Si $d \rightarrow \infty$; $u_o \rightarrow 0$; $G_o(u_o) \rightarrow 0.04$;

$G_o(u_o, \epsilon_{ro}) \rightarrow 0.96\epsilon_{ro} + 0.04$; $\psi \rightarrow 1 + r / (2\pi)^2$ et

$\epsilon_{re}(u, \epsilon_r, u_o, \epsilon_{ro}) \rightarrow F_+ \cdot \epsilon_r + F_- (0.96\epsilon_{ro} + 0.04) \cong$

$(F_-)(\epsilon_r \coth \beta h + \epsilon_{ro})$

Si $d \rightarrow h$; $u_o \rightarrow u$; $\psi \rightarrow 1$; $G_o(u_o) \rightarrow G_o(u)$;

$G_o(u_o, \epsilon_{ro}) \rightarrow G_o(u, \epsilon_{ro})$ et $\epsilon_{rea} \rightarrow \epsilon_{re}(u, \epsilon_r, \epsilon_{ro})$.

5. résultats :

Dans cette section, nous nous contentons de rapporter les résultats et les courbes de l'impédance d'entrée, du taux d'ondes

stationnaires et du gain en fonction de la fréquence relatifs aux antennes tri-couches de forme disque normal et disque lissé et les tableaux de comparaison avec les structures simples et bi-couches de mêmes formes.

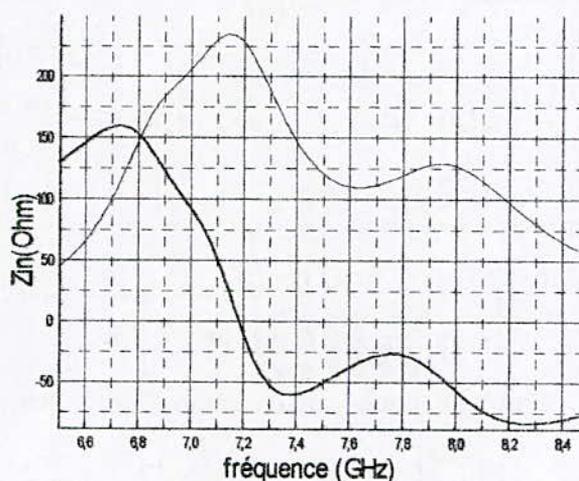
5.1. disque normal (D_N) :

Examinons à titre d'exemple le cas:

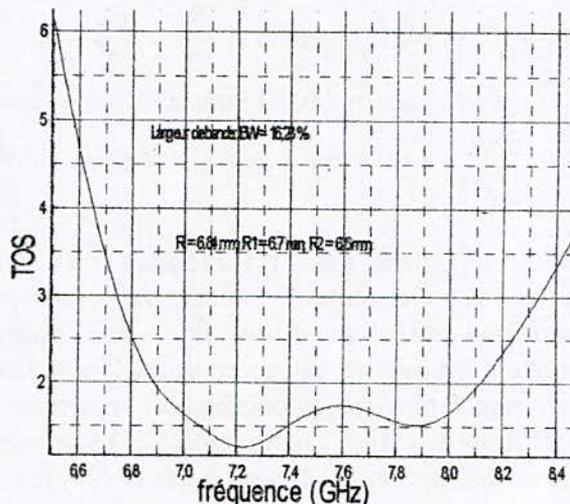
$R = 6.84$ mm, $R_1 = 6.7$ mm, $R_2 = 6.6$ mm, $a_f = 6.15$ mm. a_f : point d'excitation par rapport au centre du disque.

A la fréquence centrale de 7.5 GHz, la largeur de bande est de 16.3 % le gain est de 17.5 dB. Le produit gain x largeur de bande est de 2.85.

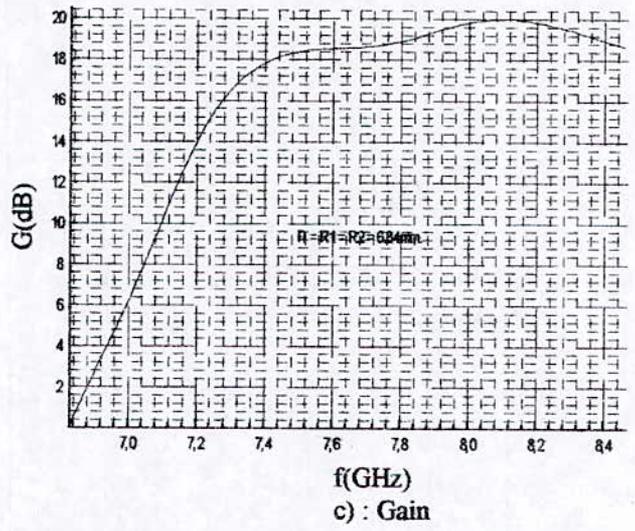
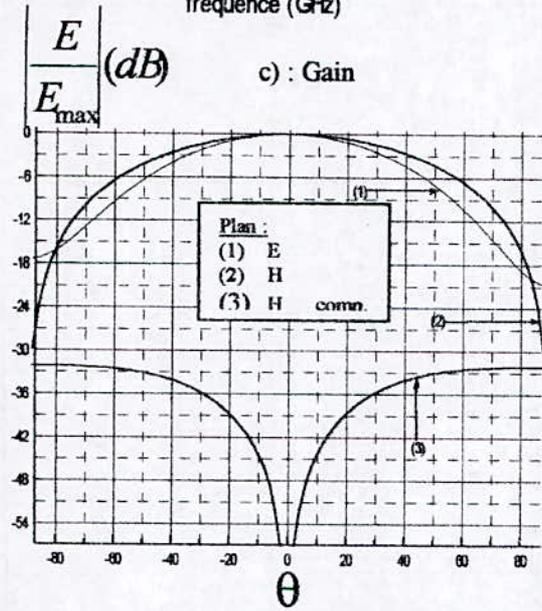
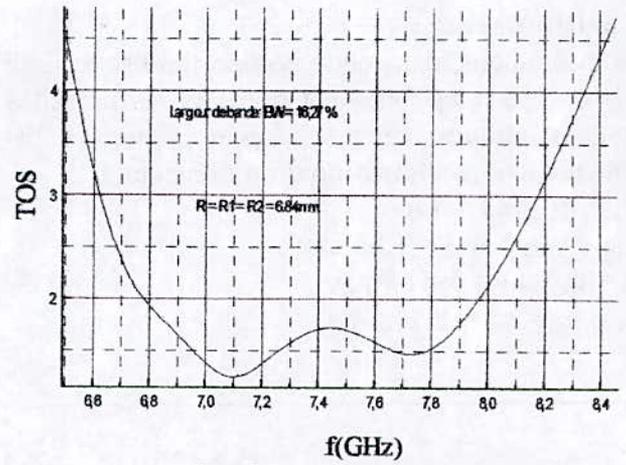
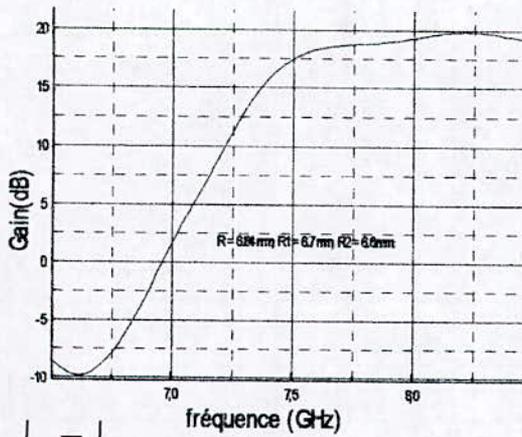
Lorsque $R = R_1 = R_2 = 6.84$ mm, la largeur de bande, le gain à la fréquence centrale et les ouvertures du faisceau à -3 dB dans les plans E et H sont respectivement 16.27%, 17.5 dB, 70° et 84°. Le niveau de la composante croisée dans le plan H est au-dessous de -33 dB.



a) : Impédance d'entrée

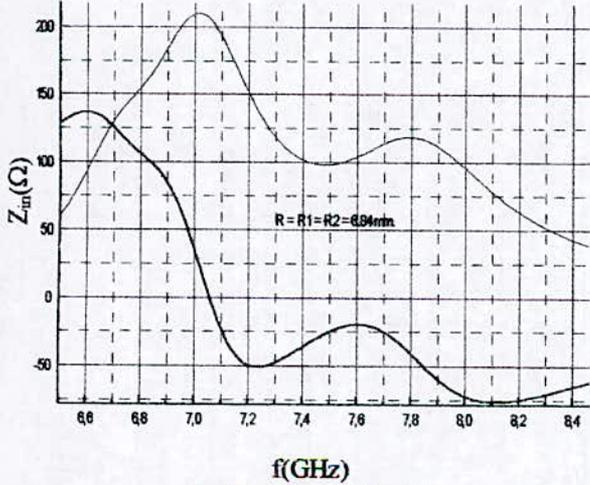


b) : Taux d'ondes stationnaires

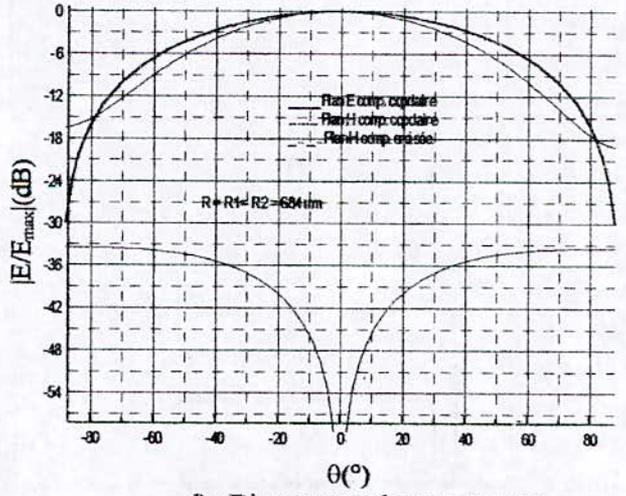


d) : Diagrammes de rayonnement à 7.62 GHz

Fig (4-1) : Antenne disque tri-couches $R_3=6.84\text{ mm}$, $R_2=6.7\text{mm}$, $R_1=6.6\text{ mm}$, $h_1=h_2=h_3=1.6\text{ mm}$,



a) : Impédance d'entrée



d) : Diagrammes de rayonnement
 Fig (4-1bis) : Antenne disque tri-couches $R_3=6.84\text{ mm}$, $R_2=6.84\text{mm}$, $R_1=6.84\text{ mm}$, $h_1=h_2=h_3=1.6\text{ mm}$, $a_f=6.15\text{ mm}$.

5.2. Disque lissé :

Nous étudions, dans cette section, le disque lissé tri-couches dont l'élément pilote et les parasites sont identiques (même structure que la bi-couches avec l'adjonction d'un directeur).

$R_1 = R_2 = R_3 = 3 \text{ mm} ;$

$w_{01} = w_{02} = w_{03} = 1.68 \text{ mm} ;$

$h_1 = h_2 = h_3 = 1.6 \text{ mm} ;$

$\epsilon_{r1} = \epsilon_{r2} = \epsilon_{r3} = 2.17.$

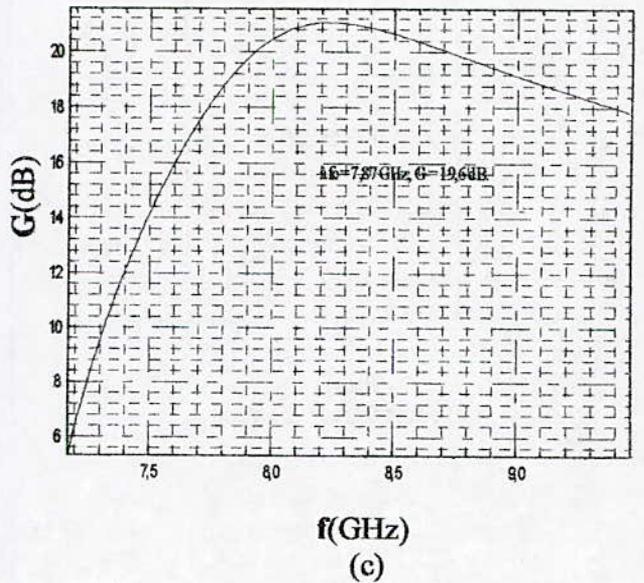
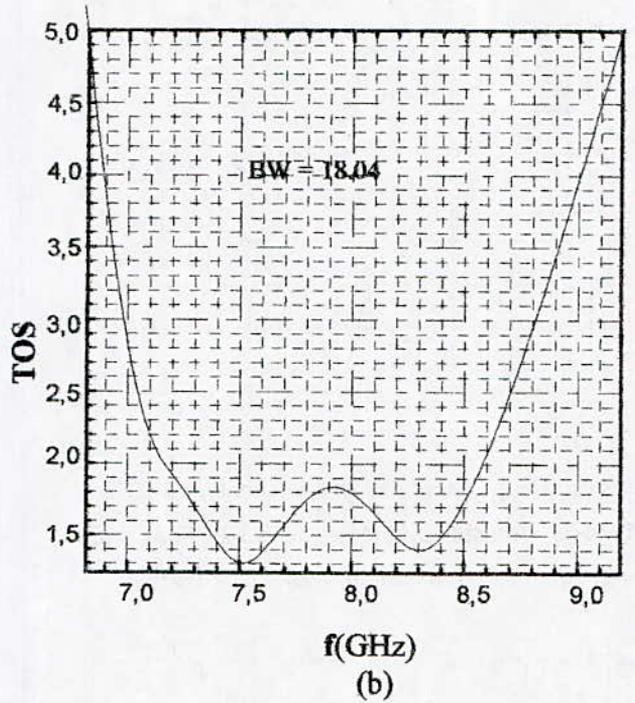
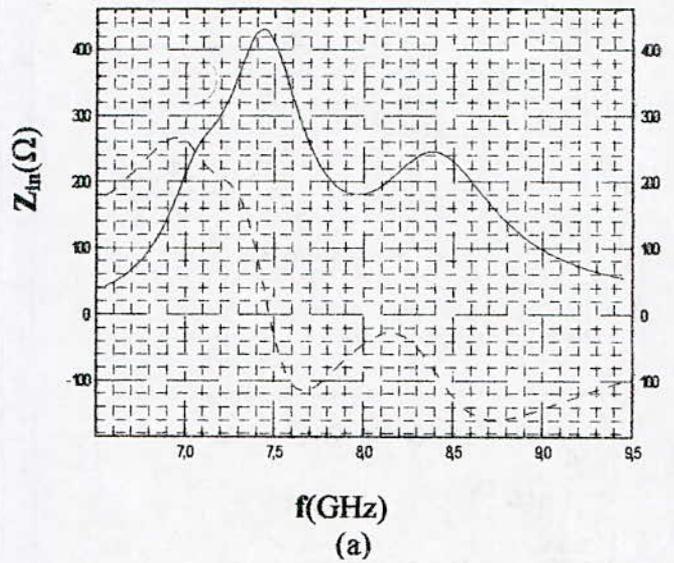
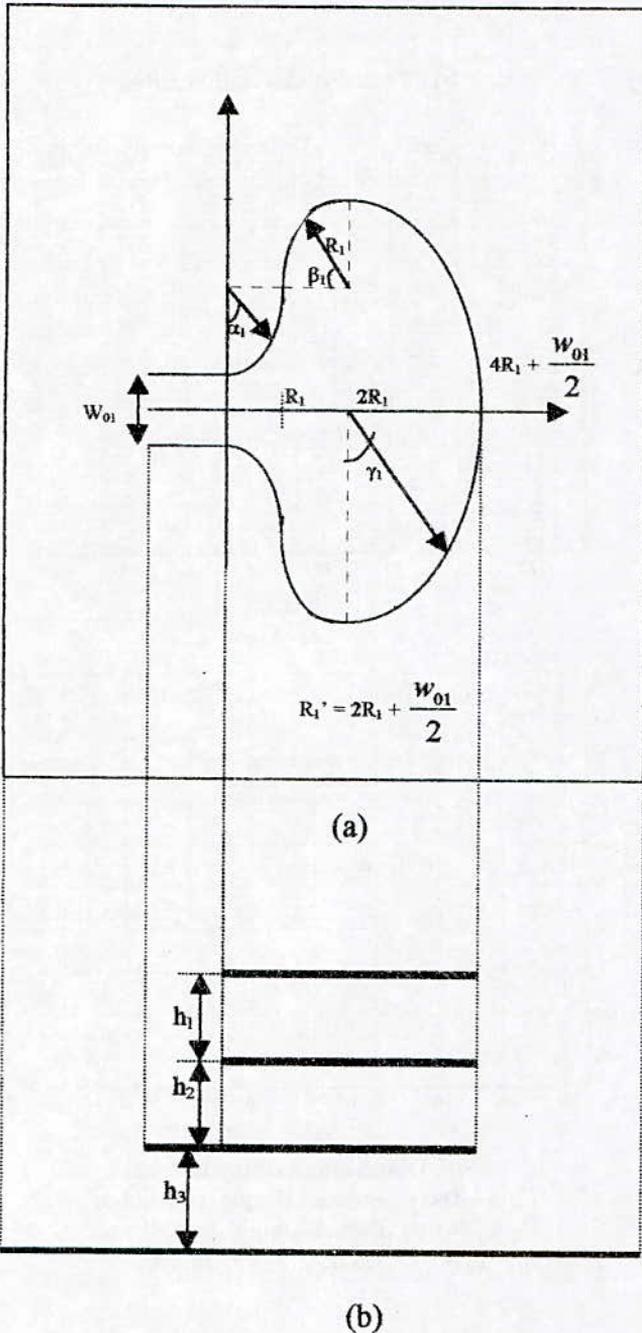


Fig. (4-2) : Antenne tri-couches de forme arbitraire, (a) : vue de dessus, (b) : vue en coupe.

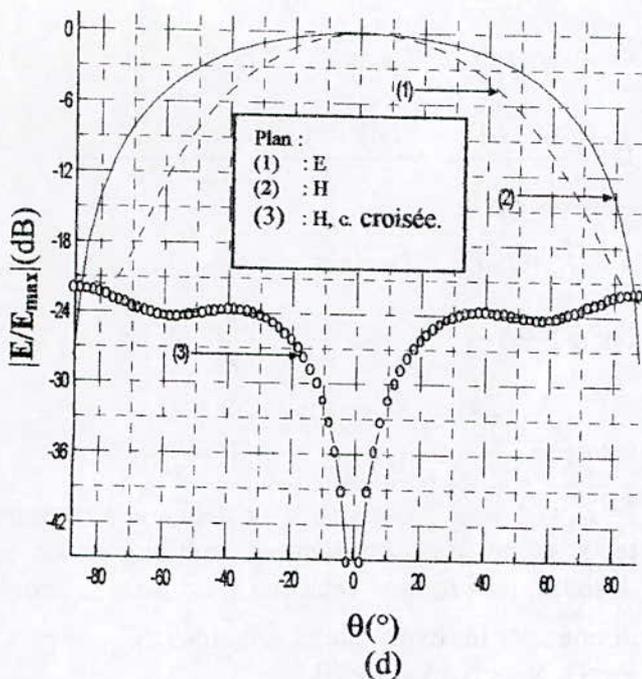


Fig. (4-3) : Antenne de la figure (4-2),
 (a) : Impédance d'entrée, (b) : Taux d'ondes stationnaires,
 (c) : Gain, (d) : Diagrammes de rayonnement à la première
 résonance 7.44 GHz.

5.3. Tableau comparatif :

	Monocouche		Bi-couches		Tri-couches	
	D_N	D_L	D_N	D_L	D_N	D_L
$f_c(\text{cente})(\text{GHz})$	8.00	8.42	7.50	8.02	7.50	7.87
$B(\%)$	7.50	7.65	13.40	16.46	17.50	18.04
$G(\text{dB}) \text{ à } f_c$	7.5	11.8	14.7	17.0	16.3	19.6
$B \times G$	0.60	0.90	1.96	2.80	2.85	3.54

Le produit Gain-Largeur de bande de la structure bi-couches de directeur et antenne de même forme et de mêmes dimensions est de 2.80, alors que la structure tri-couches nous le donne à 3.54 nettement plus grand mais au prix d'un relèvement de la composante croisée, cependant dans des limites tolérables.

5.4. Conclusion :

La méthode des LECF a fait ses preuves dans l'analyse et la conception des antennes plaques microruban, simples ou multicouches. Cette étude de l'antenne plaque microruban, de forme quelconque multicouche à large bande, nous a permis d'analyser, en l'appliquant à des antennes plaques disque normaux et lissés associés à des directeurs de mêmes formes, les caractéristiques radioélectriques de ces antennes ; en particulier l'impédance d'entrée, la

bande de fonctionnement, les diagrammes de rayonnement et le gain. Cette analyse nous a montré qu'il est possible de concevoir des antennes plaques à large bande et à haut gain environ 20 % et 20 dB respectivement. Les Résultats théoriques sont en bon accord avec la théorie et les mesures de la littérature.

La méthode des LECF est facile à mettre en œuvre et exige peu de temps de calcul relativement à d'autres méthodes. Le temps de calcul mis par le programme d'analyse d'une structure tri-couches sur micro-ordinateur de type MMX-200 pour une fréquence (i.e. un point de la courbe de l'impédance d'entrée en fonction de la fréquence) est d'environ 13 mn. Alors que la littérature rapporte des chiffres comme 100 mn pour une fréquence et une structure simple sur station VAX. Ce qui montre l'utilité de la LECF par rapport à d'autres méthodes comme celle des moments M_2 dans le domaine spectral par exemple. En prenant en charge les effets de bord et de peau, elle est devenue très précise, en particulier dans la gamme de fréquence allant de 1 à 30 GHz.

La méthode des LECF peut traiter des structures simples ou multicouches pourvu que leur contour soit régi par une fonction connue.

L'antenne disque lissé s'avère être plus performante que l'antenne disque normal en ce qui concerne le produit Gain-Largeur de bande (B.G). Il passe de 0.6 à 0.9 soit une augmentation de 50% pour les antennes simples. Pour les antennes bi-couches l'augmentation est d'environ 43 %. Pour les antennes tri-couches elle est de 24 %. La méthode des LECF, vu sa simplicité de mise en œuvre et ses faibles exigences en temps de calcul est très recommandée pour la caractérisation et la recherche d'APM de toute spécificité.

La méthode des LECF, avec les améliorations que nous lui avons apportées prend en charge l'overlay soit rigoureusement avec les intégrations variationnelles ou rapidement avec les formules d'interpolation. Elle est en mesure de faire la prédiction des paramètres radioélectriques de l'APM dotée d'une couche de protection ou d'un radôme. Le programme de conception peut donc nous fournir la nature et les dimensions de l'overlay selon le cahier des charges de l'antenne à concevoir. Il peut aussi,

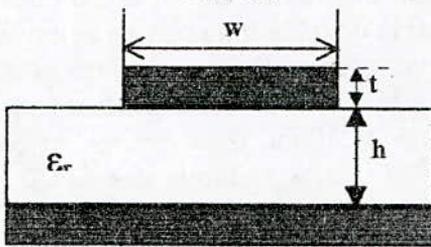
lorsque nous lui introduisons les données d'une antenne déterminée, nous fournir ses caractéristiques radioélectriques.

Pour une bande de fréquence de travail et un diagramme de rayonnement donnés, la conception de l'APM idoine par la méthode des LECF s'avère très efficace.

Malgré les hypothèses de simplification que nous avons adoptées, telles que le plan de masse infini, substrat linéaire, homogène, isotrope, réciproque et non magnétique, les écarts entre les résultats théoriques et les mesures restent dans la marge de $\pm 1,7\%$. Ce qui montre que la méthode des (LECF) utilisée est très compétitive.

L'APM conçue du point de vue de la maximisation du produit G.B maximal correspond essentiellement à la structure tri-couches. Ceci se fait même mieux avec le disque lissé D_L . Cependant le gain est très sensible à la localisation de l'excitation. Si celle ci est au bord, il est élevé, ce qui relève aussi l'impédance d'entrée. La recherche d'une APM omnidirectionnelle nécessite donc une localisation plus proche du centre. Ce qui correspond à une impédance d'entrée de 10 à 20 Ω . Dans tous les cas l'impédance d'entrée doit être adaptée à sa source par des adaptateurs quart d'onde ou stubs suivant le cas désiré.

Annexe



Ligne microruban en coupe

Détermination de l'impédance caractéristique et du coefficient de propagation d'une ligne microruban et calcul du gain d'une antenne
Calcul du gain :

Le gain correspondant à la direction de rayonnement maximal est donné par:

$$G = \eta D \quad \text{avec} \quad \eta = \frac{P_r}{P_t} \quad \text{et}$$

$$D = \frac{\frac{1}{2} \operatorname{Re}(ExH^*) \cdot \hat{r}}{P_r / 4\pi R^2} = \frac{1/2 \operatorname{Re}(E_\theta H_\phi^* - E_\phi H_\theta^*) |_{\theta=0}}{P_r / 4\pi R^2}$$

$$ExH^* = (E_\theta H_\phi^* - E_\phi H_\theta^*) \hat{r}, \quad \text{dans le cas d'une}$$

$$\text{onde TEM } H_\phi^* = \frac{E_\theta^*}{\eta_0} \quad \text{et} \quad H_\theta^* = -\frac{E_\phi^*}{\eta_0} \quad \text{donc}$$

$$G(\theta, \varphi) = \frac{2\pi R^2 (|E_\theta|^2 + |E_\phi|^2)}{\eta_0 P_t}, \quad P_t = \frac{1}{2} Z_{in} I^2,$$

$I = I_{ca} = 1A$, \hat{r} : désigne le vecteur unitaire dans le système de coordonnées sphériques et \times désigne le produit vectoriel. E_θ et E_ϕ sont donnés par les expressions suivantes :

$$E_\theta = E_0 \cdot \cos\theta \cdot \cos\phi \cdot \sin(k_z h).$$

$$\int_0^L Ica(x) \cdot \cos[k_y y_G(x)] e^{jk_x x} dx$$

$$+ 2jE_0 \cdot \cos\theta \cdot \sin\phi \cdot \sin(k_z h).$$

$$\int_0^L Icy(x) \cdot \sin[k_y y_G(x)] e^{jk_x x} dx$$

$$+ jE_0 \cdot \sin\theta \cdot \cos(k_z \frac{h}{2}).$$

$$\int_0^L Ipa(x) \cdot \cos[k_y y_G(x)] e^{jk_x x} dx$$

$$E_\phi = -E_0 \cdot \sin\phi \cdot \sin(k_z h).$$

$$\int_0^L Ica(x) \cdot \cos[k_y y_G(x)] e^{jk_x x} dx$$

$$- 2jE_0 \cdot \cos\phi \cdot \sin(k_z h) \cdot \int_0^L Icy(x) \cdot \sin[k_y y_G(x)] e^{jk_x x} dx$$

$$\text{où } E_0 = \frac{\eta_0}{\lambda_0} \psi(r) \quad \psi(r) = \frac{e^{jk_0 r}}{r} \quad (\text{fonction de}$$

$$\text{Green en espace libre) } \eta_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$$

$$k_x = k_0 \sin\theta \cos\phi \quad k_y = k_0 \sin\theta \sin\phi$$

$$k_z = k_0 \cos\theta \quad k_0 = 2\pi/\lambda_0$$

λ_0 , η_0 , ϵ_0 et μ_0 sont respectivement la longueur d'onde, l'impédance d'onde, la permittivité et la perméabilité en espace libre. Les expressions des courants de conduction I_{ca} et de polarisation I_{pa} sont données respectivement par les équations :

$$I_{ca}(x) = \frac{v^i(x) - v^r(x)}{z_c(x)}$$

$$I_{pa}(x) = j\omega\epsilon_0(\epsilon_r - 1) W(x)[v^i(x) + v^r(x)]$$

Où v^i , v^r , z_c , et W désignent respectivement les ondes de tensions incidente et réfléchie, l'impédance caractéristique et la largeur de la plaque au niveau de l'abscisse x . L désigne la longueur totale de la plaque. Dans le cas d'une ligne microruban élémentaire d'abscisse x et de largeur $W(x)$, la fonction de position du centre de gravité décrivant la ligne de force de part et d'autre de l'axe de symétrie de la plaque est donnée par :

$$Y_G(x) = \frac{7}{25} W(x).$$

L'expression de la composante transverse du courant de conduction est donnée par :

$$I_{cy}(x) = \frac{1}{2} \frac{d(y_G(x))}{dx} I_{ca}(x)$$

η , D , P_r , P_t , E , H sont respectivement le rendement, la directivité, la puissance rayonnée, la puissance totale fournie à l'antenne, et les champs lointains électrique et magnétique.

Impédance caractéristique :

$$Z_c(u) = \frac{\eta_0}{2\pi} \ln \left[\frac{f(u)}{u} + \sqrt{1 + \left(\frac{2}{u}\right)^2} \right], \quad \eta_0 = 120\pi, \quad u = \frac{w}{h}$$

$$f(u) = 6 + (2\pi - 6) \cdot \exp \left[-\left(\frac{30,66}{u}\right)^{0,7528} \right]$$

Le coefficient de propagation d'une ligne microruban avec perte s'écrit

$$\gamma = \alpha + j\beta \tag{A-1}$$

Le terme d'atténuation α , tenant compte des diverses pertes (ohmiques, diélectriques et par rayonnement), et le terme de phase β s'écrivent respectivement

$$\alpha = \alpha_c + \alpha_d + \alpha_r \tag{A-2}$$

$$\beta = k_0 \sqrt{\epsilon_{eff}(f)} \tag{A-3}$$

1. Pertes dans les conducteurs :

Les pertes ohmiques dans les conducteurs sont essentiellement dues à la conductivité finie de ces conducteurs. Cependant l'épaisseur non nulle du conducteur supérieur et la rugosité de la surface des dépôts métalliques peuvent accroître sensiblement ces pertes. Divers modèles [13] donnent une formulation décrivant les pertes ohmiques à l'aide d'expressions fermées. Le modèle le plus précis, dû à Pues (1983) et cité par Van de Capelle [14], donne une valeur des

pertes dans les conducteurs sous forme d'expressions aisément programmables

$$\alpha_c = \alpha_s + \alpha_g \tag{A-4}$$

avec $\alpha_s = \alpha_n R_s F_{\Delta s} F_s = \alpha_n F_s R'_s$

$$\alpha_g = \alpha_n R_g F_{\Delta g} = \alpha_n R'_g$$

$$R_i = \sqrt{\pi f \mu_0 / \sigma_i} = \frac{1}{\sigma_i \delta}, \quad i = s, g$$

$$F_s = 1 + \frac{2}{u_r} \left(1 - \frac{1}{\pi} + \frac{\Delta u_r}{(t/h)} \right)$$

$$F_{\Delta i} = 1 + \frac{2}{\pi} \text{Arctg} \left\{ 1,4 \left(\frac{\Delta_i}{\delta} \right)^2 \right\}, \quad i = s, g$$

$$R'_i = R_i F_{\Delta i}, \quad i = s, g$$

α_n (Nep/unité de longueur)=

$$\begin{cases} \frac{1}{2\pi h z_c(0)} \frac{32 - u_r^2}{32 + u_r^2} \Lambda, & \text{si } u_r < 1 \\ \frac{\sqrt{\epsilon_e}}{2\eta_0 w_{eff}(0)} \left\{ u_r + \frac{0,667 u_r}{u_r + 1,444} \right\} \Lambda, & \text{si } u_r \geq 1 \end{cases}$$

α_c (dB/unité de long.)=

$$\begin{cases} 1,38 \cdot \frac{R'_s}{h z_c(0)} \frac{32 - u_r^2}{32 + u_r^2} \Lambda, & \text{si } u_r \leq 1 \\ 6,1 \cdot 10^{-5} \frac{R'_s Z'_c(0) \epsilon_e}{h} \left[1 + \frac{0,667}{u_r + 1,444} \right] u_r \Lambda, & \text{si } u_r \geq 1 \end{cases}$$

$$w_{eff}(0) = \frac{\eta_0 h}{\sqrt{\epsilon_e z_c(0)}}, \quad u = w/h, \quad u_1 = u + \Delta u_1,$$

$$u_r = u + \Delta u_r,$$

$$\Delta u_1 = \frac{\epsilon}{\pi} \ln \left(1 + \frac{4 \exp(1)}{t \cdot \coth^2 \sqrt{6,517 u}} \right) \quad \text{avec } t^2 = t/h$$

$$\Delta u_r = \frac{1}{2} \left[1 + \frac{1}{\cosh \sqrt{\epsilon_r - 1}} \right] \Delta u_1$$

$$\Lambda = \begin{cases} 1 + \frac{1}{u_r} \left[1 + \frac{5}{4\pi} \left(\frac{t}{w} + \log \left(\frac{4\pi w}{t} \right) \right) \right] & \text{si } u \leq \frac{1}{2\pi} \\ 1 + \frac{1}{u_r} \left[1 - \frac{5}{4} \left(\frac{t}{h} - \log \left(\frac{2h}{t} \right) \right) \right] & \text{si } u \geq \frac{1}{2\pi} \end{cases}$$

$$z_c(0) = \frac{z_{01}(u_r)}{\sqrt{\epsilon_e(u_r, \epsilon_r)}}$$

$$z_{01}(u_r) = \frac{\eta_0}{2\pi} \ln \left[\frac{f(u_r)}{u_r} + \sqrt{1 + \left(\frac{2}{u_r}\right)^2} \right]$$

$$f(u_r) = 6 + (2\pi - 6) \cdot \exp\left[-\left(\frac{30,66}{u_r}\right)^{0,7528}\right]$$

2. Pertes diélectriques :

Elles sont dues à l'énergie dissipée dans le diélectrique qui est proportionnelle au facteur de perte $\text{tg}\delta$. L'expression la plus précise du facteur α_d , tenant compte de la dispersion s'écrit [14] :

$$\alpha_d(\text{Nep/mm}) = \frac{\pi \cdot 10^{-3}}{\lambda_0} \frac{\varepsilon_r}{\sqrt{\varepsilon_e}} \frac{\varepsilon_e - 1}{\varepsilon_r - 1} \text{tg}\delta,$$

$$\alpha_d(\text{dB/unité de long.}) = \frac{27,3}{\lambda_0} \frac{\varepsilon_r}{\sqrt{\varepsilon_e}} \frac{\varepsilon_e - 1}{\varepsilon_r - 1} \text{tg}\delta \quad (\text{A-5})$$

$$\alpha_d(\text{Nep/m}) = 0,115\alpha_d(\text{dB/m})$$

3. Pertes par rayonnement :

En transmission une ligne microruban ne doit pas avoir de pertes excessives par rayonnement et par ondes de surfaces. Par contre pour une APM on favorise le rayonnement. L'expression du facteur de pertes par rayonnement α_r , donné par Dubost [6] s'écrit :

$$\alpha_r(\text{Nep/unité de long.}) = \frac{1}{10} \frac{(k_0 w_e(f))^3}{w_e(f)} \quad \text{ou}$$

$$w_e(f) = w + \frac{w_{\text{eff}}(0) - w}{1 + \frac{f}{f_g}} \quad \text{et} \quad k_0 = \frac{2\pi}{\lambda_0} \quad (\text{A-6})$$

ε_e étant définie par la relation (3-2bis).

Références :

- [1] A. Zerguerras, *Cours de 1^{ère} année, de post-graduation-télécom au Département Electronique de l'E.N.P. intitulé. «Méthodes d'analyse et de conception de dispositifs électromagnétiques»*. (MACE) (1997).
- [2] A. Zerguerras, *Contribution à l'étude d'antennes plaques de forme quelconque multicouches à large bande*. Thèse de doctorat d'état -20 mai 1990- E.N.P. Alger.
- [3] R. Aksas, *Analyse d'antennes microrubans de forme quelconque par une approche basée sur la technique des lignes de transmission. Application aux géométries circulaires*. Thèse de Doctorat d'Etat -3 juin 1995 - E.N.P. Alger.
- [4] A. Zerguerras, R. Aksas, *Etude du coefficient de couplage entre lignes superposées*. Algerian Journal of Technology, série B C D, vol. 9, pp. 22-31, 1992.

[5] G. Dubost, *Méthodes d'analyse et de synthèse de quelques 'microantennes' à large bande en mode quasi transversal électromagnétique*. Annales des télécommunications, vol. 24, N° 9 - 10, pp. 588 - 605, 1978.

[6] G. Dubost, A. Zerguerras, *Transmission Line Model Analysis of Arbitrary Shape Symmetrical Patch Antenna Coupled with a Director*. Electronics Letters, vol. 26, N° 13, pp. 952 - 954, 21st June 1990.

[7] G. Dubost, *Linear Transmission Line Model Analysis of Arbitrary - Shape Patch Antennas*. Electronics Letters, vol. 22, N° 15, pp. 798 - 799, 17th July 1986.

[8] M. Bouzouad, *Etude d'antennes plaques multicouches de formes polygonales par application de la méthode des lignes élémentaires couplées finies (LECF)*. Thèse de Magister, E.N.P. Alger, 1996.

[9] I. J. Bahl, P. Bhartia, *Design of Microstrip Antennas Covered with a Dielectric Layer*. IEEE Trans. A. P., vol. 30, N° 2, March 1982.

[10] I. J. Bahl, P. Bhartia, Stanislaw S. Stuchly, *Design of Microstrip Antenna Covered with a Dielectric Layer*. IEEE Trans. A. P., vol. 30, N° 2, pp. 314 - 318, March 1982.

[11] I. J. Bahl, S. S. Stuchly, *Analysis of Microstrip Covered with a Lossy Dielectric*. IEEE Trans. On MTT, vol. MTT 28, N° 2, pp. 104 - 109, February 1980.

[12] E. Yamashita, R. Mittra, *Variational Methode for the Analysis of Microstrip Lines*. IEEE-MTT, vol. 16, N° 4, pp. 251 - 256, April 1968.

[13] Weeler H. A, *Transmission - line properties of a strip on a dielectric sheet on a plane*. IEEE Trans. Microwave Theory and Tech. Vol. MTT-25, N° 8, p. 631, Aug. 1977.

[14] Van de Capelle A, *Microstrip antennas and arrays*. Int. Research Report of the K. U., pp. 11-12, Oct. 1988. IEEE trans. Microwave Theory and Tech. vol. MTT-36, N° 6, pp. 621-631, Mar. 1978.

UNE MODELISATION DE LA PERMITTIVITE EFFECTIVE SUR UNE LARGE GAMME DE FREQUENCE D'UNE LIGNE MICRORUBAN SUSPENDUE

B. Semmar*, M. Si Moussa*, A. Zerguerras**

ملخص

إن حساب المسامحة التواترية الخاصة بعازل هوائية ميكروشرطية لم يصل إلى حل عن طريق عبارة بسيطة على شرط نافذ واسع بغض النظر عن تطبيق طريقة العزوم ذات الإستعمال الثقيل . العلاقات الشكلية الموجودة لم تعالج إلا الحالة الخطية أو شبه الخطية لحدية أي التواترات فوق 20 جيجا هرتز (الطريقة المقترحة من طرف طومار و باختيار (1986-1987) و التواترات تحت 10 جيجا هرتز (الطريقة المقترحة من طرف ميرشيكار سيأحكال (1983) . هذا العمل يقترح عبارة وحيدة لحساب المسامحة على شرط نافذ يمتد من 1 إلى 100 جيجا هرتز و ذلك بدقة مرضية .

كلمات مفتاحية : ميكروشرط - هيكل متعدد الطبقات - طبقة هوائية - مسامحة حقيقية - هوائية ميكروشرطية - تشتيت

Abstract

The problem of computation of $\epsilon_{\text{eff}}(f)$ the dispersive permittivity of the dielectric of a microstrip antenna including a layer of air didn't have hitherto a solution by a simple formulation on a large band. Outside of the application of the method of moments of heavy use, the formal expressions that exist only treat the extreme linear or almost linear cases, i.e. frequencies superior to 20 GHz (method proposed by Tomar and Bhartia 1986,1987) and the lower frequencies to 10 GHz (method proposed by Mirshakar syahkal 1983). This work proposes an unique formulation to treat $\epsilon_{\text{eff}}(f)$ on a frequency band going of 1 to 100 GHz with a satisfactory precision.

Key words: Suspended microstrip line , Multi layer structure , Air gap , Effective permittivity , Microstrip antenna , Dispersion

Résumé

Le problème de calcul de $\epsilon_{\text{eff}}(f)$ la permittivité dispersive du diélectrique d'une antenne plaque microruban (A.P.M) comportant une couche d'air n'a pas eu jusqu'ici une solution par une formulation simple sur une large bande. En dehors de l'application de la méthode des moments d'emploi lourd, les expressions formelles qui existent ne traitent que les cas linéaires ou quasi linéaires extrêmes, i.e. les fréquences supérieures à 20 GHz (méthode proposée par Tomar et Bhartia 1986,1987) et les fréquences inférieures à 10 GHz (méthode proposée par Mirshkar syahkal 1983). Ce travail propose une formulation unique pour traiter $\epsilon_{\text{eff}}(f)$ sur une bande allant de 1 à 100 GHz avec une précision satisfaisante.

Mots clés : Ligne microruban suspendue , Structure multicouche , Couche d'air , Permittivité effective , Antenne plaque microruban , Dispersion

* Etudiants post - graduants au département Electronique - option télécommunications .

** Professeur à l'Ecole Nationale polytechnique . Département Electronique - Laboratoire Télécommunications.

1. Introduction

En introduisant une couche d'air entre le substrat et le plan de masse dans une antenne microruban, la permittivité de la cavité va changer [1]. La géométrie d'une antenne microruban avec couche d'air est illustrée par la figure (1).

Considérons la cavité se trouvant sous le patch conducteur. Elle est composée de deux couches : un substrat d'épaisseur 'a' et une couche d'air d'épaisseur 'b'. La largeur de bande va croître d'une part à cause de l'augmentation de la hauteur du diélectrique et du fait de l'abaissement de la permittivité effective d'autre part.

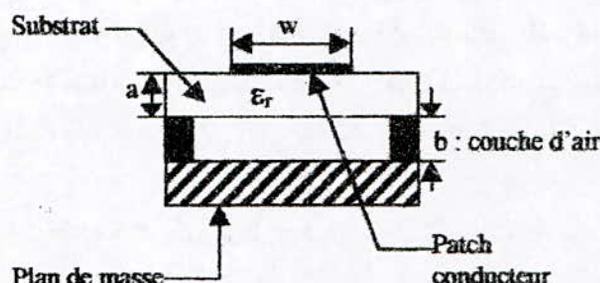


Figure 1 : Géométrie d'une antenne microruban avec couche d'air

2. Analyse

L'évaluation de la permittivité effective statique utilise deux méthodes :

2.1 Méthode T.B

C'est la méthode proposée par Tomar et Bhartia [2] [3] [4] [5], où la permittivité effective statique est donnée par :

$$\varepsilon_{eff}(0) = 1 / (1 - f_1 f_2)^2 \quad (1)$$

où

$$f_1 = 1 - \frac{1}{\sqrt{\varepsilon_r}}$$

$$f_2 = \frac{1}{\sum_{i=1}^4 C_i \left(\frac{w}{b}\right)^{i-1}}$$

tel que

$$C_i = \frac{1}{\sum_{j=1}^4 d_{ij} \left(\frac{b}{a}\right)^{j-1}}$$

où les éléments de la matrice d_{ij} sont donnés en annexe.

L'équation de dispersion est donnée par la formule suivante :

$$\varepsilon_{eff}(f) = \varepsilon_r - \frac{\varepsilon_r - \varepsilon_{eff}(0)}{1 + G(f/f_p)^2} \quad (2)$$

avec $\varepsilon_{eff}(0)$ donné par l'équation (1)

et

$$f_{p\text{GHz}} = 0.398 \frac{Z}{a} \left[\frac{m}{1+m} \right]$$

$$G = \overline{C_0} + \overline{C_1} Z + \overline{C_2} Z^2$$

où f_p est en GHz, a en mm ; $m = a/b$.

tel que

$$Z = \frac{\eta_0}{2\pi\sqrt{\varepsilon_{eff}}} \text{Log} \left[\frac{f(u)}{u} + \sqrt{1 + \left(\frac{2}{u}\right)^2} \right]$$

avec

$$u = \frac{w/b}{1 + a/b}$$

$$f(u) = 6 + (2\pi - 6) e^{(30.666/u)^{0.7528}}$$

où

$\overline{C_0}$, $\overline{C_1}$ et $\overline{C_2}$ sont donnés en annexe

w : Largeur du patch conducteur

a : Epaisseur du substrat

b : Epaisseur de la couche d'air

ε_r : Permittivité du substrat

η_0 : Impédance d'onde

La précision pour Z_0 et $\sqrt{\varepsilon_{eff}}$ est à mieux de 0.6 % pour

$1.0 \leq \varepsilon_r \leq 20$; $0.5 \leq (w/b) \leq 10.0$ et $0.06 \leq (a/b) \leq 1.5$. Ces relations de dispersion sont valables pour $1 < (w/b) < 10$ et $20 < f < 100$ GHz, avec une précision à mieux de 2 %.

2.2 Méthode P.E

C'est la méthode de la permittivité équivalente qui est basée sur le principe selon lequel toute structure bicouche peut être ramenée à une structure monocouche équivalente et ce moyennant quelques transformations.

La permittivité de l'antenne équivalente peut être donnée en fonction de celle de l'air et du substrat dans le cas statique par la formule suivante :

$$\varepsilon_{req} = \frac{\varepsilon_r (1+m)}{\varepsilon_r + m} \quad (3)$$

où $m = a/b$.

La permittivité effective statique est donnée par la relation de Schneider :

$$\varepsilon_{reff}(0) = \frac{\varepsilon_{req} + 1}{2} + \frac{\varepsilon_{req} - 1}{2} f(u) \quad (4)$$

où u est le rapport $(w/(a+b))$

i. e. $u = \frac{w/b}{1 + a/b}$

$$f(u) = \begin{cases} \frac{1}{\sqrt{1 + \frac{12}{u}}} & \text{si } u \geq 1 \\ \frac{1}{\sqrt{1 + \frac{12}{u}}} + 0.04(1-u)^2 & \text{si } u \leq 1 \end{cases}$$

formules valables à mieux de 1% pour $0.05 < w < 20$ et $\epsilon_r \leq 16$ (à mieux de 2% pour $\epsilon_r > 16$).

Il existe des expressions précises (à mieux de 0.2% pour $\epsilon_r < 128$ et $0.01 < u < 100$) comme le montre les formules suivantes :

$$f(u) = \left(1 + \frac{10}{u}\right)^{-a(u)b(u)}$$

avec

$$a(u) = 1 + \frac{1}{49} \text{Log} \left[\frac{u^4 + (u/52)^2}{u^4 + 0.432} \right]$$

$$+ \frac{1}{18.7} \text{Log} \left[1 + \left(\frac{u}{18.1} \right)^3 \right]$$

$$b(\epsilon_r) = 0.564 \left(\frac{\epsilon_r - 0.5}{\epsilon_r + 3} \right)^{0.053}$$

2.2.1. Calcul de ϵ_{req} dynamique

Le calcul de la permittivité effective dynamique de la ligne microruban suspendue nécessite l'introduction de la permittivité équivalente dynamique que l'on

désigne par ϵ_{req} et que l'on formule par :

$$\begin{aligned} \epsilon_{req} &= \frac{\epsilon_{req}(1 - thv) + \epsilon_r(1 + thv)}{2} \\ &= \frac{(\epsilon_{req} + \epsilon_r)}{2} + \frac{(\epsilon_{req} - \epsilon_r)}{2} thv \end{aligned} \quad (5)$$

avec

$$v = \frac{(2f_c - f_a)(2f_c + f_a)}{4f_c \cdot f_a}$$

tel que

$$f_c = f \cdot \frac{(1+m)5}{8\pi m}$$

$$f_a = f_p \frac{1+m}{m}$$

L'introduction de (5) dans l'expression de Mirshekar syahkal [6] donne :

$$\epsilon_{eff}(f) = \tilde{\epsilon}_{req} - \frac{\tilde{\epsilon}_{req} - \epsilon_{req}(0)}{1 + P(f)} \quad (6)$$

avec $\epsilon_{eff}(0)$ donné par la relation (4) et $\tilde{\epsilon}_{req}$ par (5)

$$P(f) = P_1 P_2 [10 \cdot f \cdot (a+b)(0.1844 + P_3 P_4)]^{1.5763}$$

où

$$P_1 = 0.27488 + \left[0.6315 + \frac{0.525}{(1 + 0.157 \cdot f \cdot h_1)^{20}} \right] u$$

$$- 0.065683 \cdot e^{-8.7513 \cdot u}$$

$$P_2 = 0.33622 \cdot [1.0 - e^{-0.03442 \cdot \epsilon_{req}}]$$

$$P_3 = 0.0363 \cdot e^{-4.6 \cdot u} [1.0 - e^{-(f \cdot h_1 / 3.87)^{4.87}}]$$

$$P_4 = 1.0 + 2.75 [1 - e^{-(\epsilon_{req} / 15.916)^2}]$$

P_2 et P_4 ne dépendent que de ϵ_{req} , alors que P_1 et P_3 ne dépendent que de u , f , $h_1 = a+b$, où f est en GHz et h_1 est en cm.

Nous avons deux formulations quasi statiques distinctes pour évaluer la permittivité effective $\epsilon_{eff}(0)$, l'une dite T.B est donnée par la relation (1), l'autre dite P.E est donnée par la relation (4).

$\epsilon_{eff}(0)$ ne dépend que ϵ_r , (a/b) et (w/b) aussi bien dans (1) que dans (4).

L'égalité

$$\epsilon_{eff}(0)_{TB} = \epsilon_{eff}(0)_{PE} \quad (7)$$

implique pour ϵ_r donnée une relation biunivoque entre (w/b) et (a/b) .

Le tracé de la courbe donnant $(w/b)_{exacte}$ solution de (7) en fonction de $(a/b) = m$ partage l'espace en deux parties.

Quand $(w/b) < (w/b)_{exacte}$ on a $\epsilon_{eff}(0)_{TB} > \epsilon_{eff}(0)_{PE}$, au contraire quand on a $(w/b) > (w/b)_{exacte}$ on a $\epsilon_{eff}(0)_{TB} < \epsilon_{eff}(0)_{PE}$.

Quand $(w/b) = (w/b)_{exacte}$ on a $\epsilon_{eff}(0)_{TB} = \epsilon_{eff}(0)_{PE}$.

On peut obtenir cette égalité à mieux de 0.1%. Une fois $(w/b)_{exacte}$ connu, on en déduit le rapport

$$u_0 = \frac{w}{a} = \frac{(w/b)_{exacte}}{(a/b)_{correspondant}}$$

ensuite w , puis la valeur de u de (4)

$$u = u_0 \frac{m}{1+m} \quad (8)$$

et on notera la valeur de $\epsilon_{eff}(0)$ correspondant au couple

$\left(\frac{w}{b}, \frac{a}{b} \right)$ solution de (7) ainsi que celle de ϵ_{req} .

En utilisant la relation donnant

$$\frac{w}{\lambda_r} = \frac{1}{\sqrt{2(\epsilon_{req} + 1)}} \quad (9)$$

selon [7], on peut évaluer λ_r et $f_r = C/\lambda_r$ où C est la vitesse de lumière.

Comme l'on a

$$0.06 < m < 1.5 \tag{10}$$

et

$$1 < \frac{w}{b} < 10 \tag{11}$$

alors

$$\frac{1}{1+m} < u < \frac{10}{1+m} \tag{12}$$

Le domaine d'évolution de u selon (12) en fonction de m est $D_u(m)$ est donné par le tableau suivant

Tableau 1 : Evolution du domaine de $D_u(m)$ en fonction de m

$m = a/b$	0.06	0.10	0.20	0.50	1.00	1.50	2.00	9.00
$D_u(m)$ va de 1 à	9.44	9.09	8.33	6.67	5.00	4.00	3.33	1.00

Sachant que pour une antenne microruban on a :

$$0.01 < \frac{a+b}{\lambda} < 0.125 \tag{13}$$

ce qui donne avec $1.0 \leq \epsilon_r \leq 20$, $b=0$ et (9)

$$0.154 < \frac{w}{\lambda_r} < 0.500 \tag{14}$$

Quand les approches T.B et P.E donnent la même valeur sur tout l'intervalle de fréquence allant de 1 GHz à 100 GHz pour la permittivité effective $\epsilon_{eff}(f)$ à mieux de 2%, elles sont toutes deux bonnes. Par contre, lorsque l'écart entre ces deux méthodes est supérieur à 2%, on a recours dans ce cas à la moyenne comme le montre la figure (3).

Pour chaque valeur de $m = a/b$, il ne correspond qu'une seule valeur de w/b qui satisfait à cette condition. (§ tableau (2) en annexe).

Le fait que la fonction $\left(\frac{w}{b}\right)_{exacte}(m)$ soit une fonction croissante monotone de m , oscillant faiblement autour de sa valeur moyenne, il est possible de la linéariser grossièrement par une ligne droite passant par l'origine $m=0$ et dont la pente donne (w/a)

i.e.
$$\frac{w}{b} = \left(\frac{w}{a}\right) m \tag{15}$$

En réalité, (w/a) oscille autour d'une valeur moyenne

$$\left(\frac{w}{a}\right) = 7.87$$

Quand $b \rightarrow 0$; $m \rightarrow \infty$ et $(w/b) \rightarrow (w/a) \cong 7.895 (a/b)$

Cependant, dès que m dépasse la valeur 1.32, u devient supérieur à la valeur 4.31 qui limite le domaine de u (i.e. $D_u < 4.31$).

Une linéarisation plus précise pour $0.17 < a/b < 1.32$ serait :

$$\left(\frac{w}{b}\right) = 0.81 + 6.92 \left(\frac{a}{b}\right) \tag{16}$$

La courbe tracée à l'aide du tableau (2) élaboré pour $\epsilon_r = e = 2.718\ 2818$ donne les valeurs exactes de w/b pour des valeurs données du rapport $m=a/b$ qui permettent d'avoir la concordance entre les méthodes T.B et P.E au niveau statique (i.e. la solution de (7)).

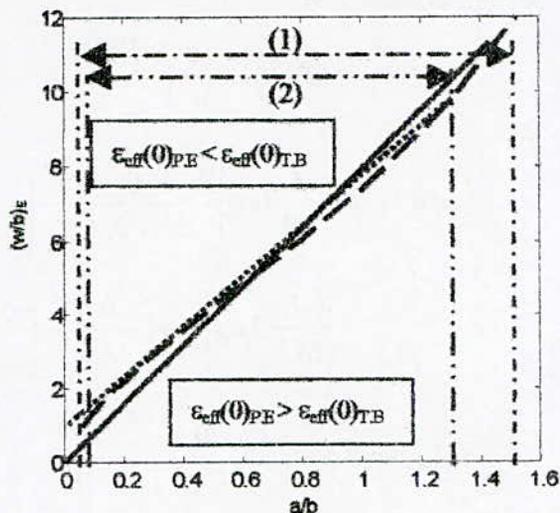


Figure 2 : Concordance de la permittivité effective statique en fonction des dimensions de la ligne microruban suspendue

- Equation (15)
- Equation (16)
- - - - Tableau (2)

- (1) : Domaine de validité de la T.B [3]
- (2) : Domaine de validité de ce travail

3. Effet de la permittivité du substrat

Nous allons maintenant nous intéresser à l'effet de la permittivité du substrat sur les différentes caractéristiques de la ligne microruban suspendue ainsi que sur l'évolution de la permittivité équivalente en fonction de la fréquence en utilisant les méthodes T.B et P.E.

Tableau 3 : Effet de ϵ_r pour $a=b=0.40$ mm

ϵ_r	1.000	1.400	2.718	7.389	20
$\epsilon_{r_{moy}}$	1.000	1.167	1.462	1.762	1.905
$\epsilon_{eff}(0)$	1.000	1.1241	1.3432	1.5743	1.6963
w/λ_r	0.500	0.480	0.451	0.425	0.415
λ_r (mm)	5.864	6.108	6.542	6.899	7.065
f_r (GHz)	51.16	49.12	45.85	43.48	42.46

La fréquence chute de 8.7 GHz en passant de 51.16 GHz à 42.46 GHz lorsque la permittivité croît de 1 à 20. La fréquence est donc faiblement affectée par la permittivité.

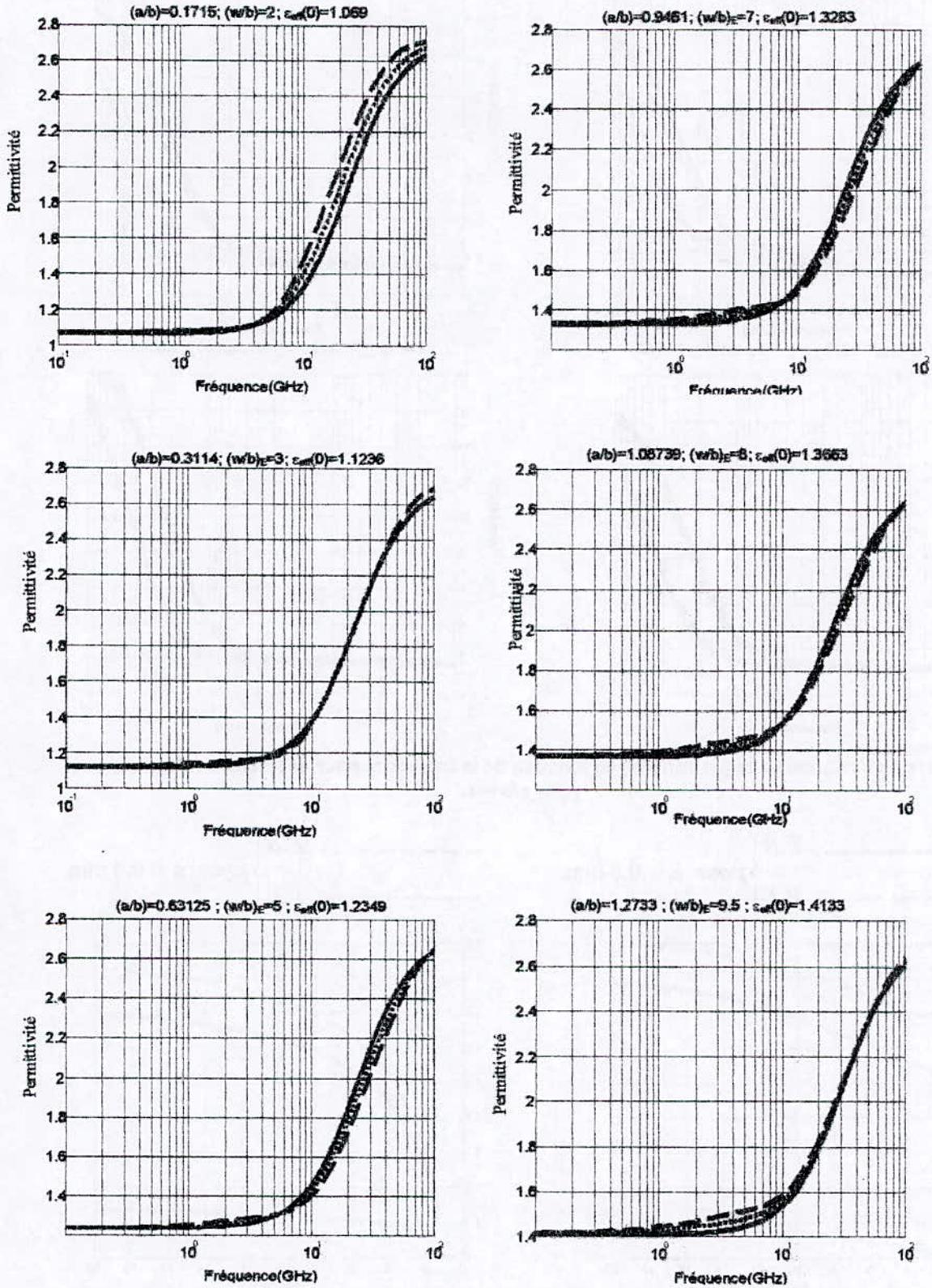


Figure 3 : Variation de la permittivité en fonction de la fréquence pour des dimensions de la ligne microruban suspendue satisfaisant la concordance de la permittivité effective statique ($a = 0.8 \text{ mm} ; \epsilon_r = 2.7$)

- T.B
- - - P.E
- Moyenne = (T.B + P.E)/2

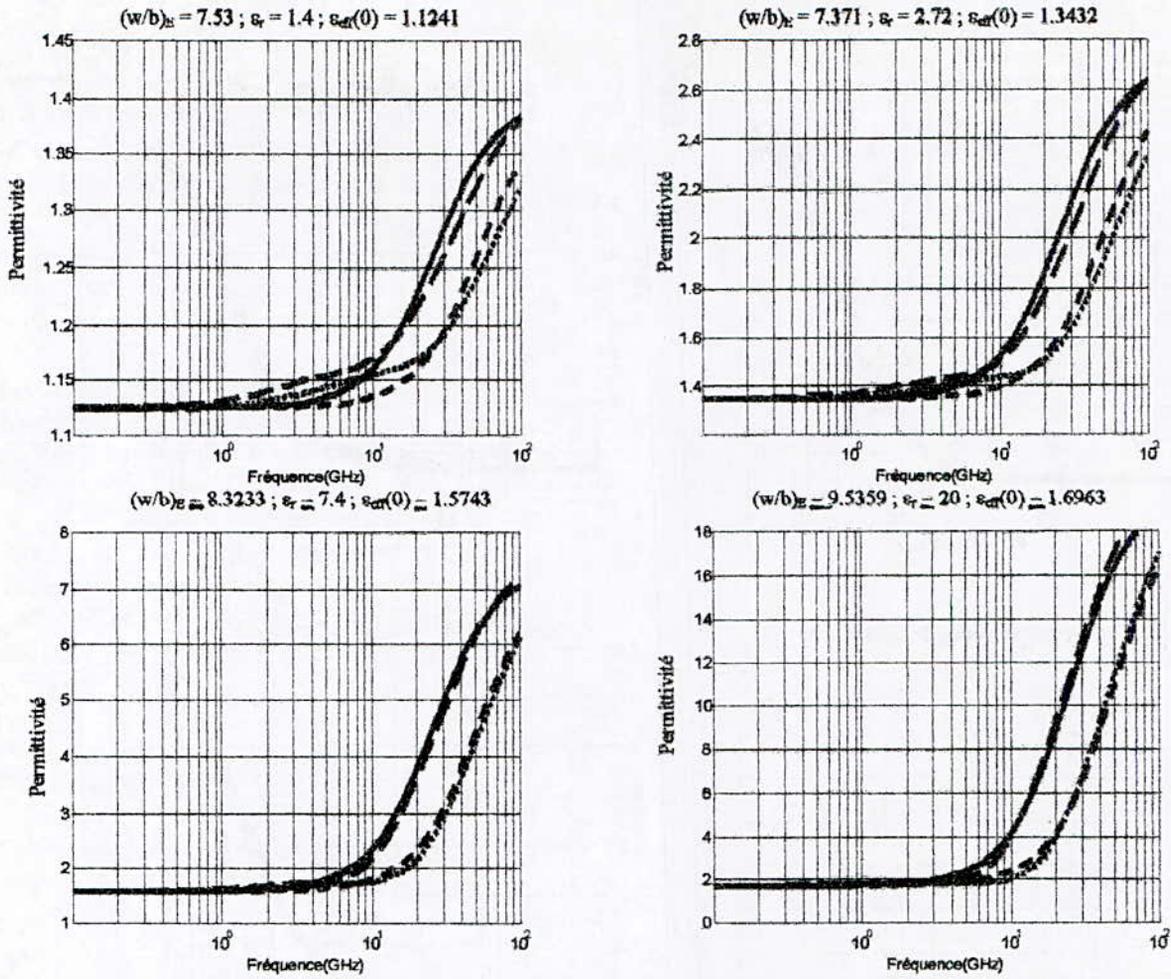


Figure 4 : Variation de la permittivité en fonction de la fréquence pour différentes valeurs de ϵ_r , pour $a/b = 1$

T.B } pour $a = 0.8 \text{ mm}$
 P.E }
 T.B } pour $a = 0.4 \text{ mm}$
 P.E }

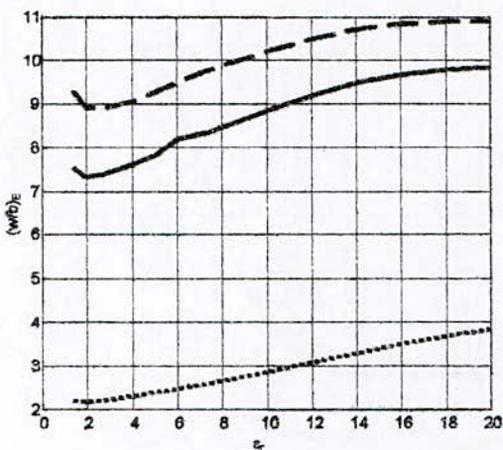


Figure 5-a : variation de $(w/b)_E$ en fonction de la permittivité du substrat

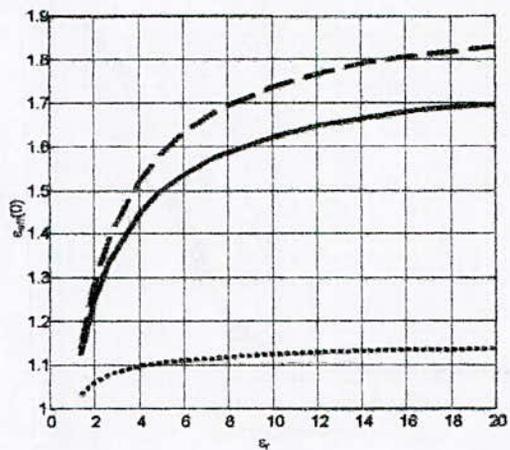


Figure 5-b : variation de la permittivité statique en fonction de la permittivité du substrat

$a/b = 0.2$
 $a/b = 1.0$
 $a/b = 1.2$

La figure (4) montre l'évolution de la permittivité dispersive en fonction de la fréquence pour différentes valeurs de ϵ_r pour $a=0.4$ et 0.8 mm.

Dans chaque cas, il est nécessaire d'ajuster à chaque fois la valeur (w/b) par résolution de l'équation (7) qui assure la concordance entre les deux méthodes utilisées T.B et P.E. La concordance des deux méthodes s'améliore quand ϵ_r croît de 1.4 à 20.

Les figures (5-a) et (5-b) indiquent respectivement les variations de $(w/b)_{\text{exacte}}$ (notée $(w/b)_E$) et $\epsilon_{\text{eff}}(0)$ en fonction de la permittivité du substrat pour $m=0.2, 1.0$ et 1.2 . On remarque l'insensibilité de ces grandeurs aux variations de ϵ_r au delà de la valeur 16, on note cependant leur grande sensibilité aux variations du rapport $m = a/b$.

Ces figures permettent de donner des indications sur le choix des dimensions de la ligne micro ruban suspendue en fonction de la permittivité du substrat utilisé.

Les modèles P.E. et T.B. ont donc été rendu équivalents, alors qu'initialement l'un (P.E.) était plus adapté aux cas avec l'épaisseur $a \geq 0.80$ mm, l'autre (T.B.) l'était pour les cas avec $a \leq 0.80$ mm. Cependant, si l'on veut avoir une précision à micux de 10%, on doit avoir $w/b > 2$, $a/b > 0.1715$ d'une part, à mieux de 5% si $w/b < 10$, $a/b < 1.3282$ d'autre part. Ce qui réduit le domaine de validité de cette théorie de $1 < w/b < 10$ à $2 < w/b < 10$, ou ce qui revient au même l'intervalle $0.06 < a/b < 1.5$ se réduit à $0.1715 < a/b < 1.32$.

4. Conclusion

Dans ce travail, nous avons élargi le domaine de validité fréquentielle de deux méthodes permettant de caractériser la permittivité d'une ligne micro ruban suspendue. La première méthode utilisée étant celle de Tomar - Bhartia (T.B) et la seconde celle de la permittivité équivalente (P.E). A chacune de ces méthodes, nous avons appliqué les deux lois de dispersion (2) et (6).

Bien que nous ne disposons pas de mesures pour valider ces résultats, nous pouvons affirmer que le calcul de la permittivité d'une ligne micro ruban suspendue peut se faire avec ces deux méthodes T.B et P.E et donne des résultats avec une précision satisfaisante (i.e. à micux de 2%) sur toute la gamme de fréquence considérée (1 à 100 GHz), si l'on suit la procédure indiquée ci - dessus. Toutefois, cette modélisation reste perfectible.

Nous remarquons aussi que lorsque nous sommes en présence d'un substrat multicouches, la valeur de la permittivité équivalente tend vers celle dont la permittivité est la plus élevée lorsque la fréquence croît i.e. l'onde électromagnétique tend à être piégée à l'intérieur de cette couche ayant la permittivité la plus élevée.

Bibliographie

- [1] J. R. James, P. S. Hall, " *Handbook of microstrip antenna* ". Vol.1 Peter Peregrinus Ltd. 1989. 813 p.
- [2] R. S. Tomar and P. Bhartia, " *Suspended and Inverted Microstrip Design* ". Microwave Journal. March 1986, pp. 173 - 177
- [3] R. S. Tomar and P. Bhartia, " *Modelling the Dispersion in a Suspended Microstripline* ". IEEE - MTT - S. Symposium Digest 1987. pp. 713 - 715
- [4] R. S. Tomar and P. Bhartia, " *New Quasi - Static Models for the Computer-Aided Design of Suspended and Inverted Microstriplines* ". IEEE - MTT - 35 N° 4 April 1987. pp. 453 - 457 and corrections N° 11, November 1987, p. 1076
- [5] Pramanick, Protap and P. Bhartia, " *An Accurate Description of Dispersion in Microstrip* ". Microwave Journal. December 1983. pp. 89 - 92
- [6] D. Mirshekar - Syahkal, " *An Accurate Determination of Dielectric Loss Effect in Monolithic Microwave Integrated Circuits Including Microstrip and Coupled Microstrip Lines* ". IEEE - MTT - 31 N° 11, November 1983. pp. 950 - 954
- [7] I. J. Bahl, P. Bhartia, " *Microstrip antennas* ". Artech House. 1980. 480 p.

Annexe

En posant $p = \text{Log } \varepsilon_r$, il vient :

$$\begin{aligned} d(1,1) &= (176.2576 - 43.1240 * p + 13.4094 * p^2 - 1.7010 * p^3) * 1e-2; \\ d(1,2) &= (4665.2320 - 1790.4000 * p + 291.5858 * p^2 - 8.0888 * p^3) * 1e-4; \\ d(1,3) &= (-3025.5070 - 141.9368 * p - 3099.4700 * p^2 + 777.6151 * p^3) * 1e-6; \\ d(1,4) &= (2481.5690 + 1430.3860 * p + 10095.5500 * p^2 - 2599.1320 * p^3) * 1e-8; \end{aligned}$$

$$\begin{aligned} d(2,1) &= (-1410.2050 + 149.9293 * p + 198.2892 * p^2 - 32.1679 * p^3) * 1e-4; \\ d(2,2) &= (2548.7910 + 1531.9310 * p - 1027.500 * p^2 + 138.4192 * p^3) * 1e-4; \\ d(2,3) &= (999.3135 - 4036.7910 * p + 1762.4120 * p^2 - 298.0241 * p^3) * 1e-6; \\ d(2,4) &= (-1983.7890 + 8523.9290 * p - 5235.4600 * p^2 + 1145.7880 * p^3) * 1e-8; \end{aligned}$$

$$\begin{aligned} d(3,1) &= (1954.0720 + 333.3873 * p - 700.7473 * p^2 + 121.3212 * p^3) * 1e-5; \\ d(3,2) &= (-3931.0900 - 1890.7190 * p + 1912.2660 * p^2 - 319.6794 * p^3) * 1e-5; \\ d(3,3) &= (-532.1326 + 7274.7210 * p - 4955.7380 * p^2 + 941.4134 * p^3) * 1e-7; \\ d(3,4) &= (138.2037 - 1412.4270 * p + 1184.2700 * p^2 - 270.0047 * p^3) * 1e-8; \end{aligned}$$

$$\begin{aligned} d(4,1) &= (-983.4028 - 255.1229 * p + 455.8729 * p^2 - 83.9468 * p^3) * 1e-6; \\ d(4,2) &= (1956.3170 + 779.9975 * p - 995.9494 * p^2 + 183.1957 * p^3) * 1e-6; \\ d(4,3) &= (62.8550 - 3462.5000 * p + 2909.9230 * p^2 - 614.7068 * p^3) * 1e-8; \\ d(4,4) &= (-35.2531 + 601.0291 * p - 643.0814 * p^2 + 161.2889 * p^3) * 1e-9; \end{aligned}$$

En posant $m = a/b$, il vient :

$$\begin{aligned} \overline{C_0} &= 0.0194 - 0.2398 * m + 0.8977 * m^2 - 0.9924 * m^3 + 0.3468 * m^4; \\ \overline{C_1} &= -0.0008 + 0.0096 * m - 0.0346 * m^2 + 0.0384 * m^3 - 0.0135 * m^4; \\ \overline{C_2} &= 0.0004 * (m^2 - m^3) + 0.0001 * m^4; \end{aligned}$$

Ces relations sont valables pour $1.0 \leq w/b \leq 10$ et $20 \text{ GHz} < f < 100 \text{ GHz}$

pour $a = 0.8 \text{ mm}$				pour $a = 0.4 \text{ mm}$				w/λ_r	$\varepsilon_{\text{eff}}(\theta)$	ε_r	μ	w/a	w/b	a/b
f_r (GHz)	λ_r (mm)	w (mm)	f_r (GHz)	λ_r (mm)	w (mm)	w/λ_r	$\varepsilon_{\text{eff}}(\theta)$							
11.1543	26.8954	13.3245	22.3086	13.4477	6.6622	0.4954	1.0236	1.0372	0.9443	16.6556	1.0010	0.0601		
12.2577	24.4744	12.0907	24.5154	12.2372	6.0453	0.4940	1.0315	1.0488	1.1170	15.1134	1.2000	0.0794		
15.6833	19.1287	9.3294	31.3665	9.5643	4.6647	0.4877	1.0690	1.1020	1.7072	11.6618	2.0000	0.1715		
16.0326	18.7119	9.1106	32.0652	9.3560	4.5553	0.4869	1.0742	1.1092	1.7730	11.3883	2.1000	0.1844		
18.6562	16.0804	7.7071	37.3124	8.0402	3.8536	0.4793	1.1236	1.1766	2.2876	9.6339	3.0000	0.3114		
20.6698	14.5139	6.8383	41.3397	7.2570	3.4192	0.4712	1.1805	1.2524	2.7249	8.5479	4.0000	0.4679		
20.9811	14.2986	6.7149	41.9623	7.1493	3.3574	0.4696	1.1917	1.2671	2.7993	8.3936	4.2000	0.5004		
21.9607	13.6608	6.3366	43.9214	6.8304	3.1683	0.4639	1.2349	1.3238	3.0651	7.9208	5.0000	0.6312		
22.6705	13.2330	6.0553	45.3410	6.6165	3.0276	0.4576	1.2846	1.3879	3.3469	7.5691	6.0000	0.7927		
22.9264	13.0853	5.9190	45.8529	6.5427	2.9595	0.4523	1.3283	1.4436	3.5969	7.3988	7.0000	0.9461		
22.9274	13.0848	5.8966	45.8547	6.5424	2.9483	0.4506	1.3430	1.4621	3.6852	7.3707	7.3700	0.9999		
22.8364	13.1369	5.8857	45.6728	6.5685	2.9428	0.4480	1.3663	1.4910	3.8325	7.3571	8.0000	1.0874		
22.5004	13.3331	5.9264	45.0007	6.6666	2.9632	0.4445	1.3988	1.5308	4.0634	7.4080	9.0000	1.2149		
21.9948	13.6396	6.0232	43.9897	6.8198	3.0116	0.4416	1.4268	1.5640	4.2952	7.5290	10.0000	1.3282		
21.3836	14.0295	6.1620	42.7672	7.0147	3.0810	0.4392	1.4508	1.5918	4.5303	7.7025	11.0000	1.4281		
20.9167	14.3426	6.2794	41.8334	7.1713	3.1397	0.4378	1.4657	1.6085	4.6977	7.8492	11.7000	1.4906		
20.8361	14.3981	6.2926	41.6722	7.1990	3.1463	0.4370	1.4733	1.6177	4.7513	7.8658	12.0000	1.5256		

Tableau 2 : Effet du rapport $w = a/b$ pour $\varepsilon_r = 2.7182818$

Réseaux phasés à balayage électronique de type photonique

F. BENALI* et A. ZERGUERRAS**

ملخص : من أجل إتجاه تفضيلي جيد و فصوص ثانوية ذات مستوى منخفض و إلا معدوم، يجب إستعمال مولد ذو توزيع غير منتظم في شبكة الهوائيات المتوافقة الصفحة، مكونات الحزمة، ذات المسح الإلكتروني. لهذا، تعتبر التقنية الفوتونية وسيلة جد ملائمة ذات ميزات عالية بالنسبة إلى الإمكانيات الكلاسيكية، على الأقل فيما يخص مرونة الإستعمال، الخفة و النممة. في هذا العمل، ينحصر الإهتمام حول الأجهزة الفوتونية من النوع صوت - بصري، و يتم البرهان على تجلياتها، خاصة تلك المطلوبة للتطبيقات داخل أجهزة الرادار الحديثة أو في طور الإنجاز. **كلمات مفتاحية :** شبكة متوافقة الصفحة، جهاز صوت - بصري، مسح إلكتروني.

Résumé : Les réseaux phasés formeurs de faisceaux à balayage électronique, pour avoir une bonne directivité avec des lobes secondaires de niveau très bas sinon nul, passent obligatoirement par une alimentation à distribution non uniforme. Pour obtenir celle-ci, la technique photonique semble un moyen très approprié aux avantages bien supérieurs aux possibilités classiques, du moins en ce qui concerne la flexibilité d'emploi, la légèreté et la miniaturisation.

Dans ce travail, l'accent est mis sur les dispositifs photoniques du type acousto - optique et en démontre les performances particulièrement recherchées pour les applications dans les radars récents ou en cours de développement.

Mots clés : réseaux phasés, dispositif acousto - optique (A.O.), balayage électronique.

Abstract : For high directivity with very low side lobe level otherwise null, the electronic scanning phased array beam former are inevitably supplied with nonuniform distribution. In this way, the photonic technique appears the appropriate mean of advantages well higher than classical possibilities, at least with regard to the use flexibility, lightness and miniaturization.

This work is based on photonic devices of acousto - optic type, and demonstrates their performances especially in great demand for new or in fast expanding radars applications.

Key words : phased array, acousto - optic device (A.O.), electronic scanning.

* Post-graduant au Département Electronique - Option radar

** Professeur à l'E.N.P. - Département Electronique - Laboratoire Télécommunication

Introduction :

La photonique est une technologie en plein développement : une analyse statistique montre que le terme «photonique» est plus fréquemment employé que ceux «virtuel» et «multimédia», dans les titres des articles techniques publiés depuis quelques années, [1], [22], [24], [25], [26]. Cette constatation reflète par delà le simple phénomène de mode, l'impact et le potentiel de la photonique dans un certain nombre de domaines différents allant des produits de consommation aux équipements professionnels.

Un domaine où la photonique réalise des contributions significatives est celui du traitement parallèle du signal à grande vitesse. Peut être, le plus grand bénéficiaire de ces capacités est la communauté du radar, où, par la nature précise des problèmes des radars, le besoin de fonctions sophistiquées de traitement de signal, à grande vitesse est essentiel [7], [13], [29].

Le taux auquel la technologie des composants optiques est mûre, suppose que la prochaine génération des systèmes radars contiendra des processeurs optiques au niveau des sous-systèmes.

1- Rappel et position du problème [12], [23]:

Les antennes à alignements phasés sont constituées d'un certain nombre d'éléments rayonnants. L'amplitude et la phase des courants pilotant chaque élément individuel, sont contrôlées pour générer le faisceau résultant de la superposition de tous les éléments rayonnants.

On considère un alignement phasé de N éléments équidistants, chacun étant équivalent à une source uniformément rayonnante dans toutes les directions.

Pour émettre un signal (énergie) sous un angle θ par rapport à la direction normale au plan de l'alignement, le déphasage des signaux pilotant les éléments

adjacents doit être égal à :

$$\phi = 2\pi \left(\frac{a}{\lambda} \right) \sin(\theta) \quad (1)$$

où ϕ est la phase électrique des courants pilotes des n éléments : $i_n(t) = A \cos(\omega t - n\phi)$ (2), (fig. 1). Dans les alignements phasés conventionnels, le retard (déphasage) introduit au niveau de chaque élément (source de courant) est réalisé par différentes méthodes, on peut en citer :

les circuits déphaseurs (commandés par des programmes), les lignes à retard, ...

Cela rend le réseau plus encombrant, et l'éloigne d'un traitement en temps réel.

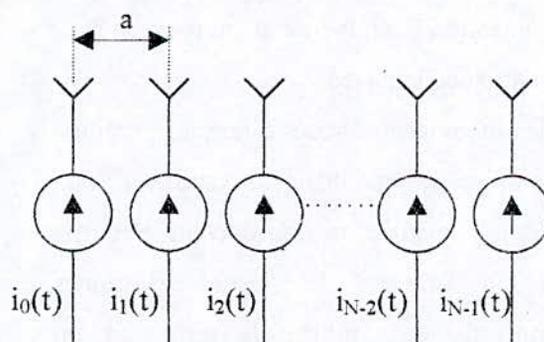


Fig.1 : Configuration d'une antenne à alignements phasés.

Pour y remédier, plusieurs solutions ont été proposées en faisant appel à des fibres optiques pour réaliser le déphasage [2], ou en utilisant des circuits intégrés optiques pour réaliser des processeurs optiques gérant les alignements phasés [9], [11], [15], [33]. On peut aborder le problème autrement par la conception des sources de courant à phases commandées, pilotant les antennes à alignements phasés, par l'utilisation de composants photoniques tels que : le laser, le modulateur de lumière, le photodétecteur, qui se caractérisent par leur dimensions réduites et leur fonctionnement rapide.

2- Conception du système de façonnage de faisceau [3],[11],[18],[29] :

Notre but est de concevoir un système photonique de sources de courants déphasés, alimentant un

alignement d'antennes permettant de faire un balayage électronique. Ces sources de courants doivent être de la forme de (2).

2-1 Le choix des composants :

a) la source d'énergie :

On a opté pour l'utilisation d'un laser [37] qui est un système d'amplification des ondes lumineuses par émission stimulée (induite). Les lasers (à gaz) [41] sont des oscillateurs stables, ce qui permet d'éviter la perturbation du fonctionnement des autres composants dans le système.

b) Le bloc de traitement du signal :

D'après l'équation (1), la variation de la phase ϕ entraîne le balayage électronique des angles θ par le faisceau des antennes, donc le traitement n'est en fait qu'une modulation de la phase.

On choisit des modulateurs acousto-optiques (cellule acousto-optique de Bragg : dispositif constitué d'un matériau cristallin, muni d'un transducteur électro-acoustique), qui utilisent les ondes élastiques (sonores) pour changer l'indice de réfraction du matériau afin de moduler la phase du signal optique qui les traverse [3], [4], [18], [29].

Comme la cellule de Bragg constitue l'élément de base de la conception, une étude générale du phénomène acousto-optique est effectuée [40], [42].

Les modulateurs de phase offrent une grande résolution, une grande vitesse de traitement et une large gamme dynamique de la phase.

c) La réception :

Le signal alimentant les antennes est un courant électrique, cependant, le signal fourni par le bloc de traitement est optique.

On utilise alors un réseau de photodétecteurs à base de photodiodes [29].

Les photodiodes sont utilisées pour convertir un signal optique, en une autre forme d'énergie stockable, plus facilement amplifiable, électrique le plus souvent. La valeur instantanée de l'excitation optique est intégralement recueillie, dans l'aire sensible de la photodiode [28]. Ainsi, le courant électrique obtenu alimentera les antennes.

2-2 La procédure de conception :

Afin d'obtenir des courants de la forme de (2), les recherches effectuées ont montré qu'il faut utiliser deux modulateurs acousto-optiques de phase. De plus, il est nécessaire de disposer de lentilles intercalées entre les différents dispositifs pour raison de couplage et d'orientation des rayons lumineux : (éviter la dispersion de la lumière).

DC : composante continue.

AOD : dispositif acousto-optique (cellule de Bragg).

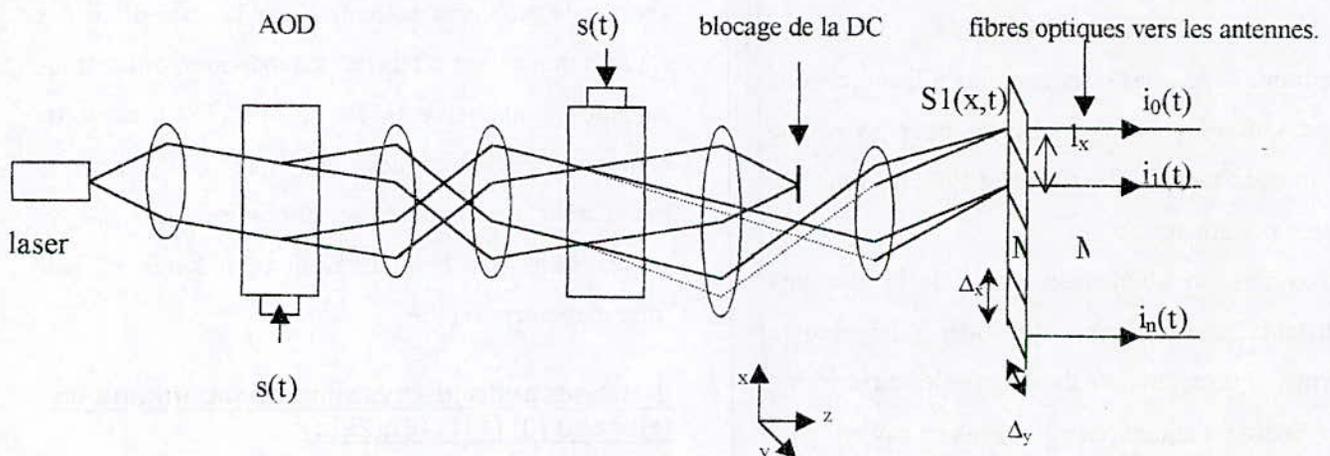


Figure 2-a : Schéma de conception du système photonique 1D

L'application de deux modulateurs acousto - optiques est valable pour un alignement 1D d'antennes. Afin de pouvoir utiliser un rideau de réseaux , la variation du signal doit se faire suivant les axes ox et oy simultanément.

Le schéma de conception 1D peut alors être transformé en 2D par l'ajout de deux autres modulateurs acousto - optiques, orientés suivant oy et le remplacement de l'alignement 1D par un réseau 2D.

En ce qui concerne le choix des dispositifs de conception, on utilise des cellules de Bragg en dioxyde de Tellure (TeO_2), des lentilles sphériques de collimation, des lentilles cylindriques de focalisation pour la première cellule. Pour la deuxième cellule, il faut disposer de lentilles sphériques, et une lentille pour le système d'amplification. Le courant électrique du réseau 1D de photodiodes est conforme à l'équation (2) :

$$i_n(t) = \frac{G\Delta_x\Delta_y}{2} \cos(\omega t - n\phi') \quad (3), \quad \text{où}$$

$$\phi' = \frac{4\pi f_0}{v} l_x. \quad (4)$$

f_0 : fréquence du signal de commande $s(t)$.

l_x : espacement inter - fibres.

v : vitesse acoustique de $s(t)$ dans la cellule

Δ_x, Δ_y : dimensions de chaque photodiode.

3- Application au réseau 2D :

Dans ce cas, il faut effectuer un traitement A.O. du signal à deux dimensions. Le principe est le même que celui du cas 1D, il suffit d'ajouter deux autres cellules de Bragg de sorte que la propagation de l'onde acoustique s'effectue aussi suivant oy (elles sont en quadrature de phase par rapport à celles de l'alimentation 1D). ces cellules sont pilotées par un signal de pulsation ω'_0 . L'alignement 1D est

remplacé par un rideau d'antennes $N \times M$. Soient l_x et l_y les distances inter - fibres (dans le réseau photonique) suivant ox et oy respectivement. On obtient ainsi une alimentation à base de quatre cellules A.O., qui envoie aux photodétecteurs le signal

$$I_{nm}(x, y, t) = a \cdot \cos\left(4\omega_c t - \frac{2\omega_0 x}{v_a} - \frac{2\omega'_0 y}{v_a}\right).$$

D'après les expériences réalisées par Riza [46] , le signal précédent reste théorique mais, en pratique, les lentilles qui orientent et focalisent la lumière en un plan image au niveau des photodétecteurs, sont placées de façon à réaliser une atténuation latérale de l'image (normale à l'axe optique) pour la rendre proportionnelle à la dimension du réseau photodétecteur. Par ce traitement, la totalité de l'énergie contenue dans le signal optique sera reçue sur la surface photodéetectrice [29], [45], [46].

Donc le signal lumineux arrive avec des atténuations M_x suivant ox et M_y suivant oy . Ce traitement influe sur l'angle d'inclinaison du plan image, et modifie donc la phase par les deux facteurs M_x et M_y .

On obtient :

$$I_{nm}(x, y, t) = a \cdot \cos\left(4\omega_c t - \frac{2\omega_0 x}{M_x v_a} - \frac{2\omega'_0 y}{M_y v_a}\right)$$

Soient Δ_{xn} et Δ_{ym} les dimensions de chaque photodiode du réseau ($\sigma_{nm} = \Delta_{xn} \times \Delta_{ym}$: est la surface).

L'intensité générée par les photodiodes est donnée par

$$i_{nm}(t) = \int_{-0.5\Delta_{ym}+ml_y}^{0.5\Delta_{ym}+ml_y} \int_{-0.5\Delta_{xn}+nl_x}^{0.5\Delta_{xn}+nl_x} I_{nm}(x, y, t) dx dy$$

Un calcul analogue à celui du cas 1D donne un courant (après intégration) :

$$i_{nm}(t) = a \frac{\sigma_{nm}}{4} \cos \left(4\omega_c t - \frac{2\omega_0 n l_x}{M_x v_a} - \frac{2\omega'_0 m l_y}{M_y v_a} \right)$$

- a : constante.
- σ_{nm} : surface de la nm^{ième} photodiode.
- ω_c : pulsation de la porteuse.
- ω_0 : pulsation du signal de commande des cellules orientées suivant ox.
- ω'_0 : pulsation du signal de commande des cellules en quadrature de phase (suivant oy).

Par identification de la phase du courant $i_{nm}(t)$ avec celle d'un réseau conventionnel, on arrive à la condition qui permet au réseau photonique de faire un balayage 2D :

$$\frac{2\omega_0 l_x}{M_x v_a} = \frac{2\pi}{\lambda_r} d'_x \sin \theta \cos \psi + 2p\pi$$

$$\frac{2\omega'_0 l_y}{M_y v_a} = \frac{2\pi}{\lambda_r} d'_y \sin \theta \sin \psi + 2q\pi$$

$$\lambda_r = \frac{\lambda_c}{4} \quad (\lambda_c : \text{longueur d'onde porteuse})$$

p, q = 0, +1, +2, ...

d'_x et d'_y : distances inter - éléments dans le réseau conventionnel.

On aura les fréquences de Bragg

suivant ox :

$$f_0 = \left[\left(\frac{d'_x}{\lambda_r} \right) \sin \theta \cos \psi + p \right] \frac{M_x v_a}{2l_x}$$

et suivant oy :

$$f'_0 = \left[\left(\frac{d'_y}{\lambda_r} \right) \sin \theta \sin \psi + q \right] \frac{M_y v_a}{2l_y}$$

f_0 : variation de la fréquence des deux cellules de Bragg orientées suivant ox.

f'_0 : variation de la fréquence des deux cellules de Bragg orientées suivant oy.

q, p = 0, 1, 2, ...

Pour étudier le système, on prend $d'_x = d'_y = \frac{\lambda_r}{2}$

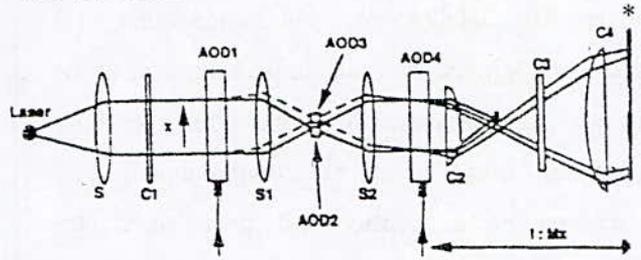
(éviter les lobes d'ambiguïté), et on utilise les données expérimentales de Riza [46] :

$$f_0 = \left[\sin \theta \cos \psi + 2p \right] \frac{M_x v_a}{4l_x} \quad (5)$$

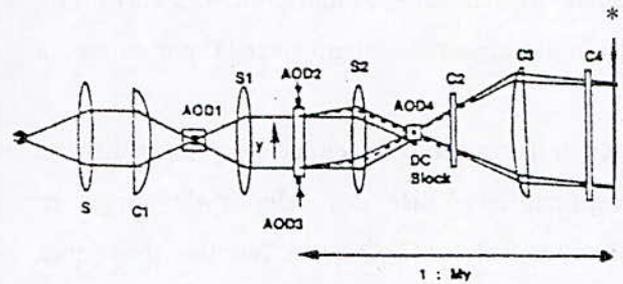
$$f'_0 = \left[\sin \theta \sin \psi + 2q \right] \frac{M_y v_a}{4l_y} \quad (6)$$

La figure 2-b représente un réseau photonique 2D réalisé par Riza [46].

Vue de dessus



Vue de coté



S_i : lentilles sphériques, C_i : lentilles cylindriques, AOD_i : dispositifs acousto - optiques.

* : réseau 2D de fibres optiques.

M_x, M_y : amplifications du plan image sur x et y.

Figure 2-b : Réseau phasé à balayage électronique par la technique acousto - optique.

a) Effet de la fréquence de commande sur l'angle de balayage :

La figure 3 représente cet effet pour des espacements inter - fibres différents. (avec p = q = 0 dans (5) et (6)).

Pour réaliser un balayage site, les fréquences f_0 et f'_0 de 42 Khz et 46 Khz respectivement orientent le faisceau à 0,5 rad par rapport à la normale du réseau (3-a), alors que si on double les espacements inter - fibres, le même angle de balayage est atteint pour des fréquences inférieures de la moitié par rapport aux précédentes (3-b).

b) Effet de la fréquence de commande sur le déphasage :

D'après la figure 4, il y a une linéarité entre la fréquence des signaux pilotes et le déphasage. Si on

veut réaliser un déphasage de $\pi/2$ entre les éléments du réseau, les fréquences f_0 et f'_0 doivent être égales à 45 Khz et 60 Khz respectivement (figure 4-a), ce qui coïncide à quelques centaines de hertz près, aux résultats de Riza [46].

Si on double l'espacement inter – fibres, le même déphasage est obtenu pour la moitié des fréquences précédentes (4-b).

Remarque 1 :

La variation de la phase dépend de f_0 et f'_0 qui ne sont autre que les fréquences des signaux de commande .

Ces fréquences sont indépendantes l'une de l'autre, ce qui implique que le faisceau des réseaux 2D peut être orienté dans les directions site et/ou azimut.

Remarque 2 :

D'après les caractéristiques des modulateurs acousto – optiques, la variation de f_0 et f'_0 implique aussi la déflexion du faisceau optique le long du réseau de photodiodes.

c) Effet de la largeur de bande des signaux de commande sur l'angle de balayage :

Soit le graphe représentant les équations (5) et (6) pour p, q variables (figure 5).

On remarque qu'on peut avoir le même angle de balayage pour différentes valeurs des fréquences de commande.

Ex : $\theta = 0.5$ rad pour $(f_0 - f'_0)_{\text{KHz}} = (43 - 59), (216 - 281), (384 - 504), (551 - 726), \dots$

Cela dote le système d'une immunité contre le brouillage, et plus la gamme de fréquence des signaux de commande est grande, plus le nombre de paires $(f_0 - f'_0)$ croît, ce qui laisse au système une plus grande marge de manœuvre contre les brouilleurs.

Une autre caractéristique importante de la fig.5, est celle de la génération de faisceaux multiples simultanés, favorables à des applications telles que la poursuite de plusieurs cibles.

d) Cas d'un alignement non uniforme :

Application du modèle de Dolph - Tchébychev :

L'amplitude du courant obtenu par le système photonique dépend de la dimension des photodiodes ($\sigma = \Delta_x \Delta_y$).

Pour avoir la pondération des amplitudes nécessaires, il suffit de modifier la dimension des photodiodes (longueur ou largeur).

Exemple :

On considère un alignement conventionnel 1D de quatre antennes colinéaires distantes de $\lambda/2$, dont le diagramme de rayonnement est donné par la figure 6-a. On veut que le rapport R des amplitudes du maximum principal et du premier maximum secondaire soit égal à 9 (19.08 dB).

Après avoir effectuer les calculs, La répartition des amplitudes des courants est la suivante :

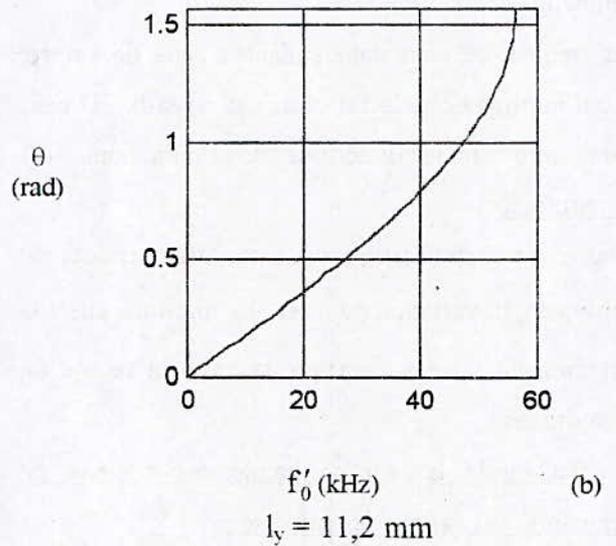
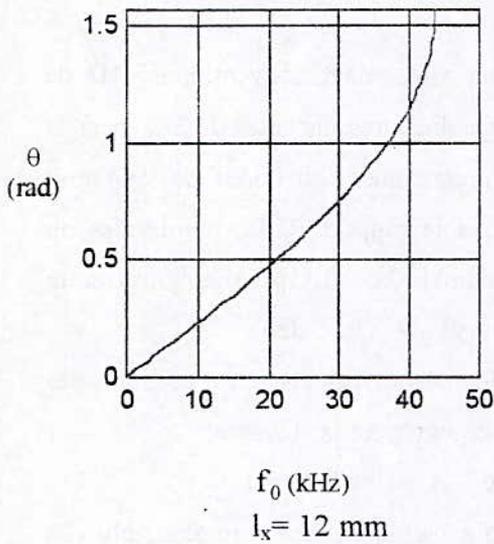
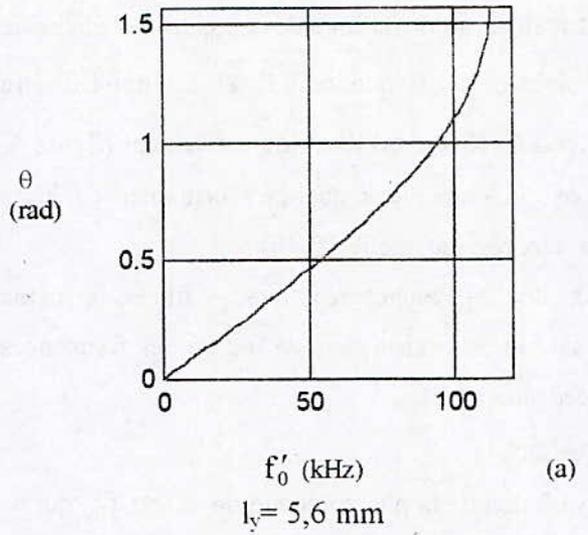
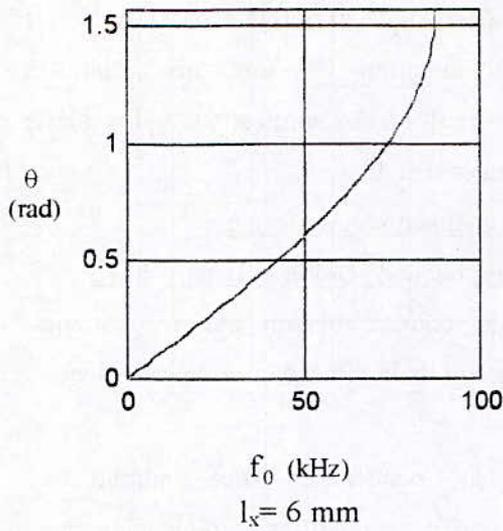
$$A_0 = 1, A_1 = 1.667, A_2 = 1.667, A_3 = 1.$$

Donc dans la fig.2-a, la surface photosensible des deuxième et troisième photodiodes doit être 1.667 fois supérieure à celle des première et quatrième (fig. 6-b).

Dans le cas 1D, l'application de la méthode de Tchébychev annule presque les niveaux des lobes secondaires au prix d'une légère dégradation de la directivité. Ces derniers sont éliminés totalement par le réseau photonique 2D tout en améliorant la directivité, en arrangeant les surfaces des photodiodes selon les pondérations des amplitudes (figure 6-c).

Conclusion :

Le système photonique s'adapte aux alignements uniformes et non uniformes sans avoir à modifier



$$\theta = \arcsin \left(\frac{4f_0 l_x}{M_x v_a \cos \psi} \right)$$

$\psi \rightarrow \frac{\pi}{2}$ réduit l'échelle de variation de f_0

$$\theta = \arcsin \left(\frac{4f'_0 l_y}{M_y v_a \sin \psi} \right)$$

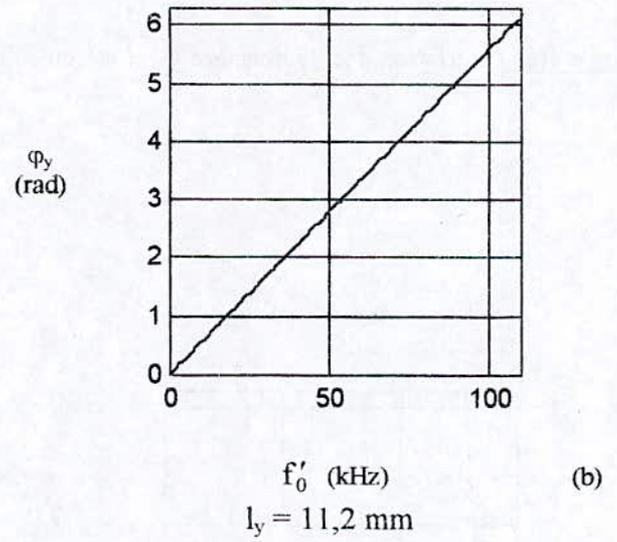
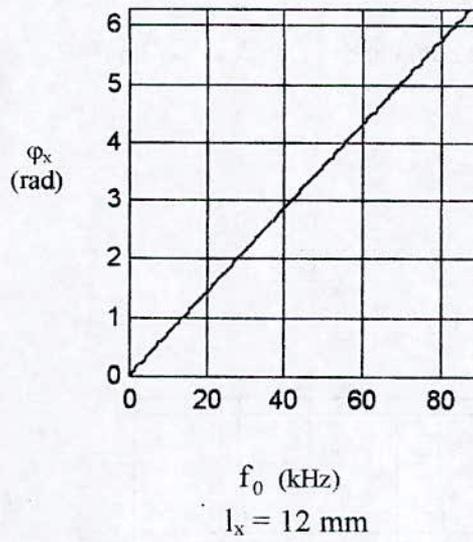
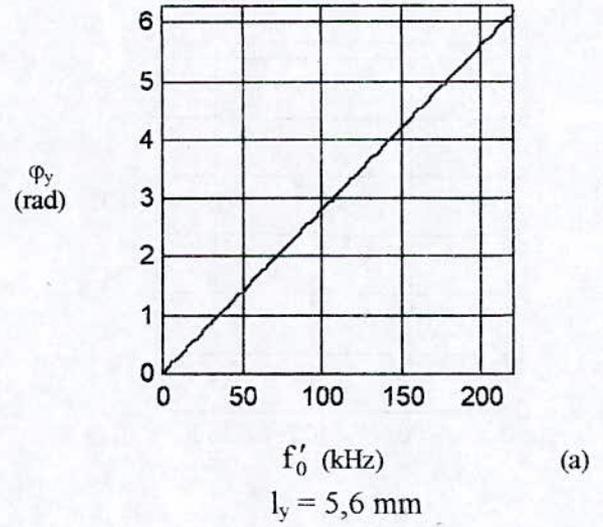
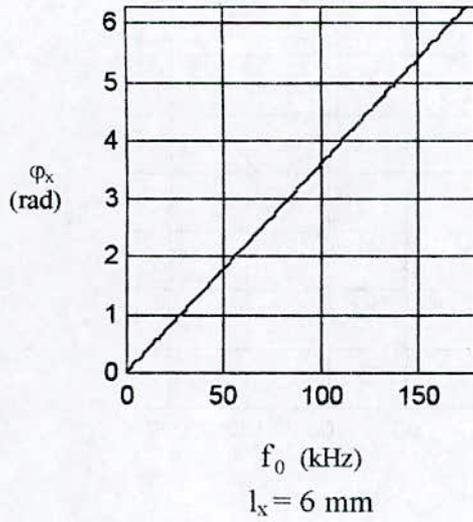
$\psi \rightarrow 0$ réduit l'échelle de variation de f'_0

$M_x = 0.5$, $M_y = 0.6$
 $q = p = 0$, $\psi = \pi / 4$
 $v_a = 4200$ m/s

f_0 : fréquence de commande des cellules de Bragg parallèles à l'axe ox.

f'_0 : fréquence de commande des cellules de Bragg parallèles à l'axe oy.

Figure 3 : Effet de la fréquence de commande sur l'angle de balayage site (θ).



$$\varphi_x = \frac{4\pi f_0 l_x}{v_a M_x} = \pi \sin \theta \cos \psi$$

$$\varphi_y = \frac{4\pi f'_0 l_y}{v_a M_y} = \pi \sin \theta \sin \psi$$

$$\tan \psi = \frac{\varphi_y}{\varphi_x} = \left(\frac{f'_0}{f_0} \right) \left(\frac{l_y}{l_x} \right) \left(\frac{M_x}{M_y} \right) = 0.777 \left(\frac{f'_0}{f_0} \right), \quad \theta = \arcsin \left(\frac{\varphi_x}{\pi \cos \psi} \right) = \arcsin \left(\frac{\varphi_y}{\pi \sin \psi} \right)$$

$$d'_x = d'_y = \frac{\lambda_r}{2} = \frac{\lambda_c}{8}, \quad p = q = 0, \quad \psi = \pi/4.$$

$$M_x = 0.5, \quad M_y = 0.6, \quad v_a = 4200 \text{ m/s}.$$

f_0 : fréquence de commande des cellules de Bragg parallèles à l'axe ox ; $\left(\frac{l_x}{M_x} \right)$ permet d'ajuster la plage de variation de f_0 .

f'_0 : fréquence de commande des cellules de Bragg parallèles à l'axe oy ; $\left(\frac{l_y}{M_y} \right)$ permet d'ajuster la plage de variation de f_0 .

Figure 4 : Effet de la fréquence de commande sur les déphasages φ_x et φ_y .

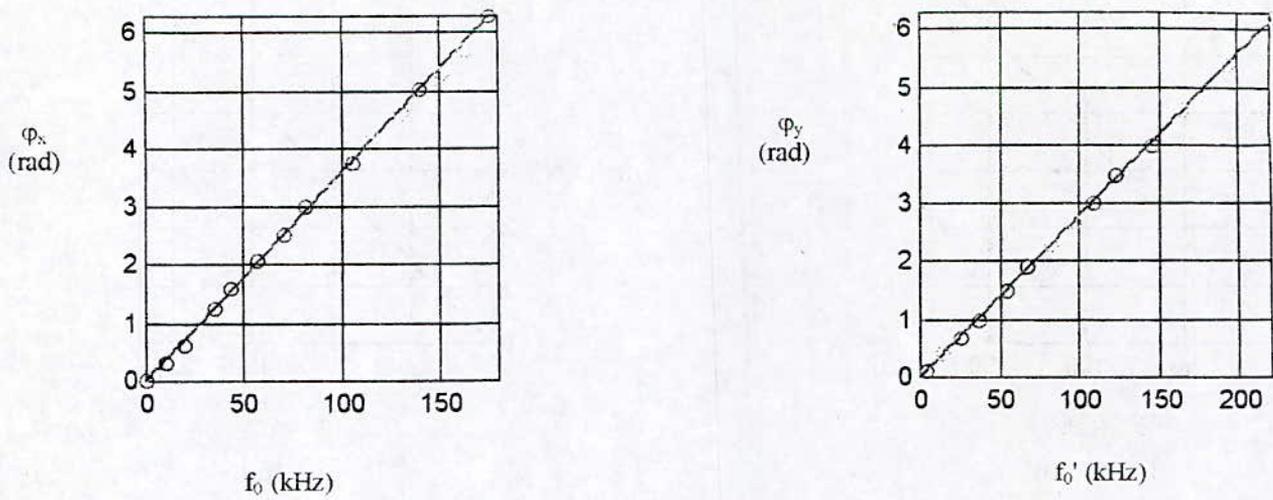
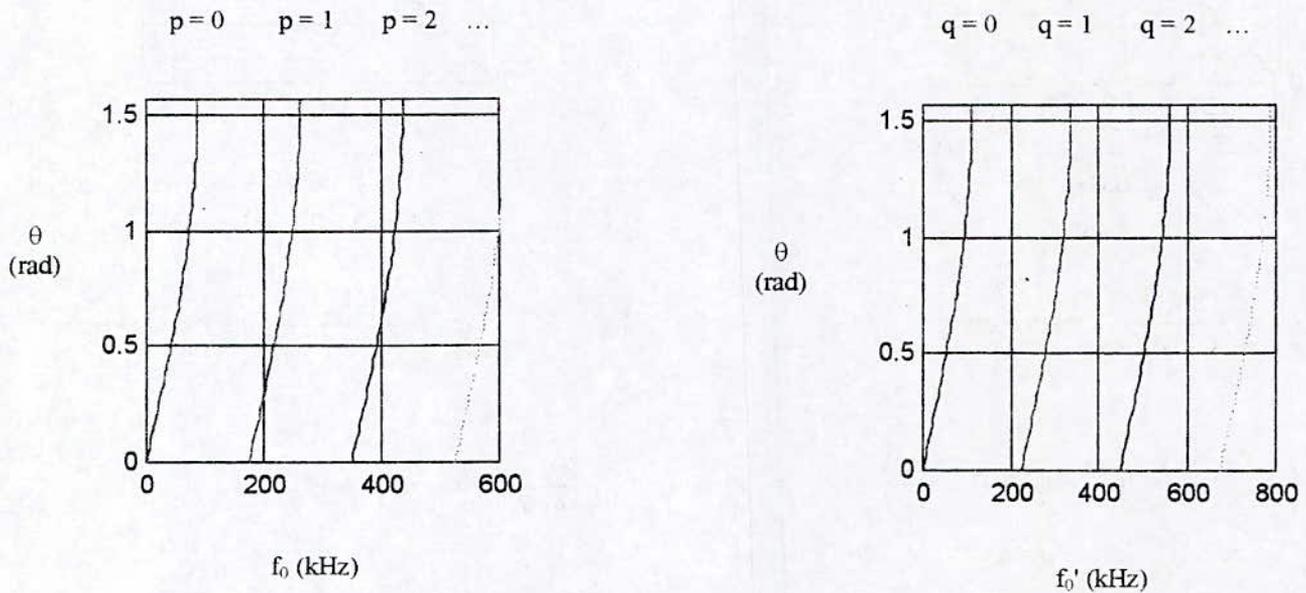


Figure 4(c) : Variation des déphasages φ_x et φ_y en fonction des fréquences de commande selon Riza [46], (1992).



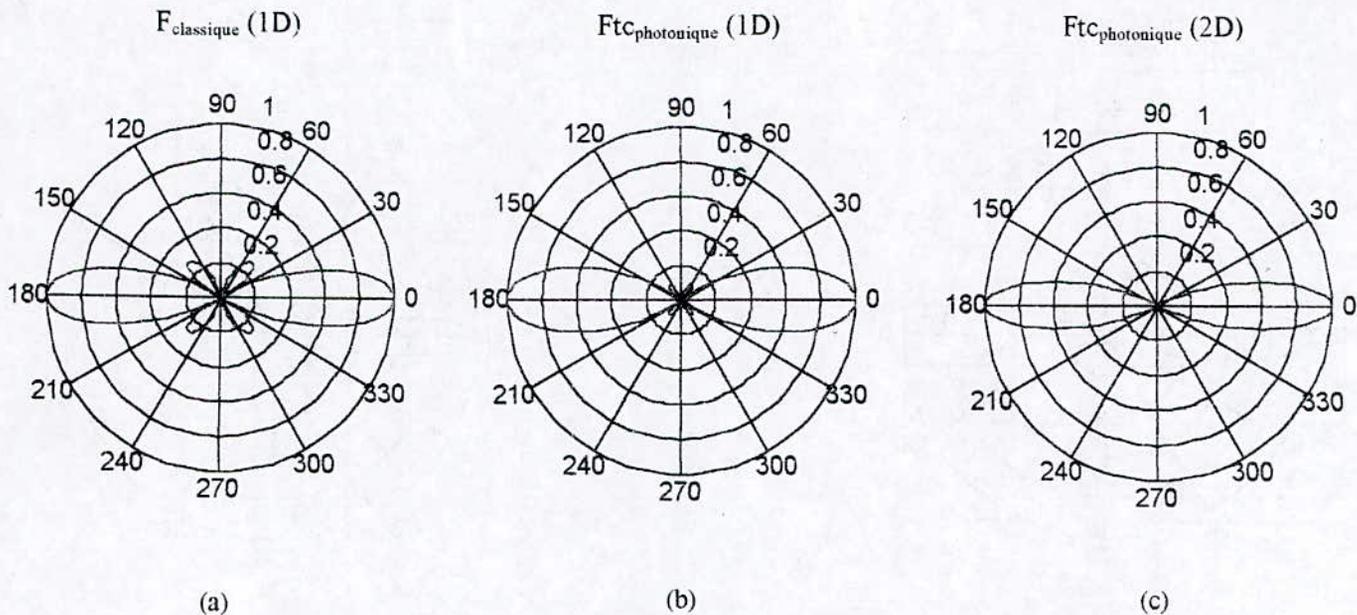
$$\theta = \arcsin \left(\left(\frac{4 f_0 l_x}{M_x v_a} - 2 p \right) \frac{1}{\cos \psi} \right) = \arcsin \left(\left(\frac{4 f'_0 l_y}{M_y v_a} - 2 q \right) \frac{1}{\sin \psi} \right)$$

$l_x = 6 \text{ mm}$, $l_y = 5.6 \text{ mm}$, $\psi = \pi / 4$
 $M_x = 0.5$, $M_y = 0.6$, $v_a = 4200 \text{ m/s}$

f_0 : fréquence de commande des cellules de Bragg parallèles à l'axe ox.

f'_0 : fréquence de commande des cellules de Bragg parallèles à l'axe oy.

Figure 5 : Effet de la largeur de bande des signaux de commande sur l'angle de balayage site (θ).



$$F_{\text{classique}} (1D) = \frac{\sin(n \phi_c / 2)}{n \sin(\phi_c / 2)}, \quad n = 4.$$

$$F_{\text{tc photonique}} (1D) = (1/5.334) * (1 + 1,667 e^{j\phi_p} + 1,667 e^{2j\phi_p} + e^{3j\phi_p})$$

$$F_{\text{tc photonique}} (2D) = (1/5.334)^2 * (1 + 1,667 e^{j\phi_p} + 1,667 e^{2j\phi_p} + e^{3j\phi_p}) * (1 + 1,667 e^{j\phi'_p} + 1,667 e^{2j\phi'_p} + e^{3j\phi'_p})$$

$$\phi_c = \frac{2\pi}{\lambda} d \sin \theta \cos \psi, \quad d = \frac{\lambda}{2}$$

$$\phi_p = \frac{4\pi f_0 l_x}{M_x v_a}, \quad M_x = 0.5, \quad l_x = 6 \text{ mm}$$

$$\phi'_p = \frac{4\pi f'_0 l_y}{M_y v_a}, \quad M_y = 0.6, \quad l_y = 5.6 \text{ mm}, \quad v_a = 4200 \text{ m/s}$$

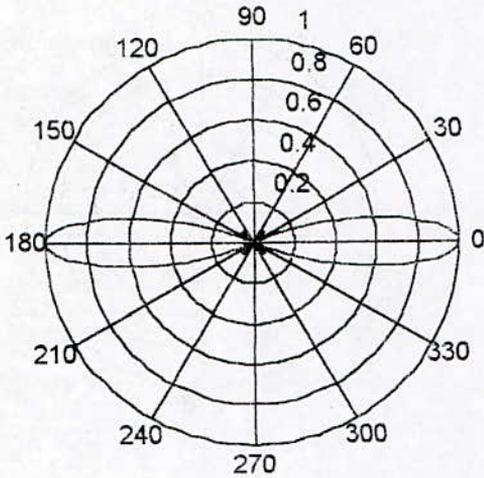
$$f_0 = \frac{M_x v_a}{4 \times l_x} (\sin \theta \cos \psi + 2p), \quad p = 0.$$

$$f'_0 = \frac{M_y v_a}{4 \times l_y} (\sin \theta \sin \psi + 2q), \quad q = 0.$$

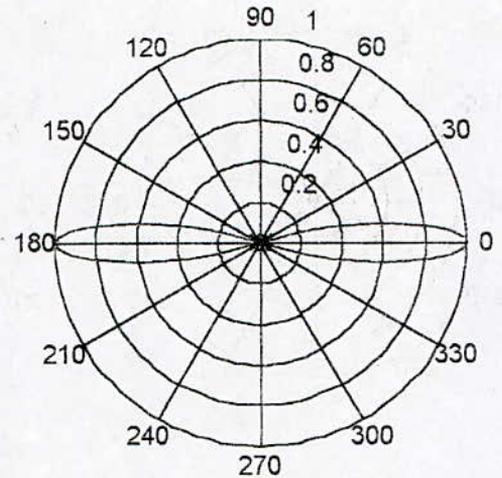
Figure 6 : Diagramme de réseau, rayonnement suivant l'angle de site (coupe horizontale) :

- (a)- Réseau conventionnel uniforme 1D.
- (b)- Réseau photonique non uniforme 1D (Tchébychev).
- (c)- Réseau photonique non uniforme 2D (Tchébychev).

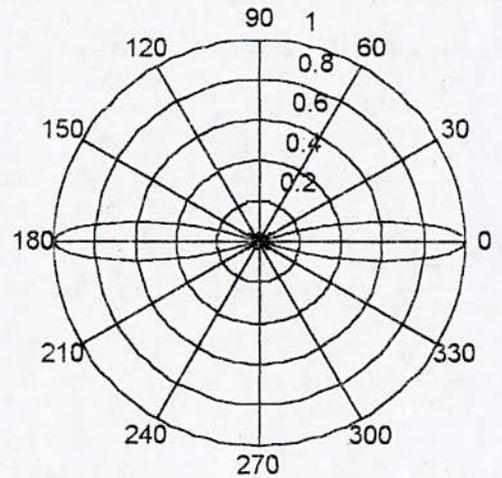
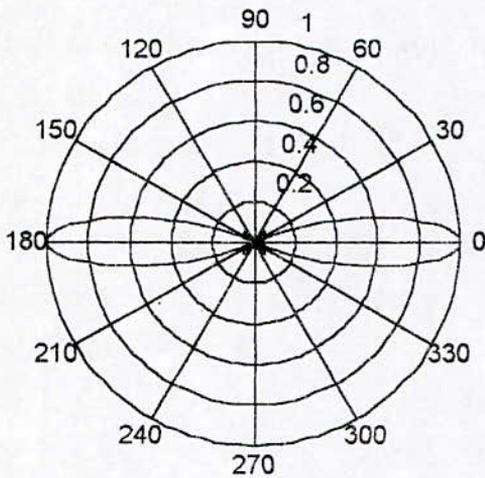
$n = m = 4$



$n = m = 5$



(a)- $d_x = d_y = \frac{\lambda_c}{2}$



(b)- $d_x = d_y = \frac{\lambda_c}{8}$

$$F_c = \frac{\sin(n \phi_c / 2)}{n \sin(\phi_c / 2)} \times \frac{\sin(m \phi'_c / 2)}{m \sin(\phi'_c / 2)} \quad , \quad \phi_c = \frac{2\pi}{\lambda_c} d_x \sin \theta \cos \psi \quad , \quad \phi'_c = \frac{2\pi}{\lambda_c} d_y \sin \theta \sin \psi \quad \text{(classique)}$$

$$F_p = \frac{\sin(n \phi_p / 2)}{n \sin(\phi_p / 2)} \times \frac{\sin(m \phi'_p / 2)}{m \sin(\phi'_p / 2)} \quad , \quad \phi_p = \frac{4\pi f_0 l_x}{M_x v_a} \quad , \quad \phi'_p = \frac{4\pi f'_0 l_y}{M_y v_a} \quad \text{(photonique)}$$

$$M_x = 0.5 \quad , \quad l_x = 6 \text{ mm} \quad , \quad M_y = 0.6 \quad , \quad l_y = 5.6 \text{ mm} \quad , \quad v_a = 4200 \text{ m/s} \quad , \quad \psi = \pi / 4 \quad .$$

$$f_0 = \frac{M_x v_a}{4 \times l_x} (\sin \theta \cos \psi + 2p) \quad , \quad p = 0 \quad , \quad f'_0 = \frac{M_y v_a}{4 \times l_y} (\sin \theta \sin \psi + 2q) \quad , \quad q = 0 \quad .$$

Figure 7 : Diagramme de réseau, rayonnement suivant l'angle de site (coupe horizontale) :
 (a)- Réseau classique ; (b)- Réseau photonique.

l'ensemble de l'alimentation ou la distribution du signal.

e) Effet de la technique photonique sur la dimension du réseau :

Considérons un réseau conventionnel 2D, avec d_x et d_y les espacements inter - éléments.

Si on utilise une porteuse de longueur d'onde λ_c , la phase est donnée par :

$$\phi = \frac{2\pi}{\lambda_c} (d_x \sin \theta \cos \psi + d_y \sin \theta \sin \psi)$$

Pour éviter les lobes d'ambiguïté dans le diagramme de rayonnement, on prend généralement :

$$d_x = d_y = \frac{\lambda_c}{2}$$

En utilisant la même porteuse pour le système photonique 2D, la phase devient :

$$\phi' = \frac{2\pi}{\lambda_r} (d'_x \sin \theta \cos \psi + d'_y \sin \theta \sin \psi)$$

$$\text{où } \lambda_r = \frac{\lambda_c}{4},$$

d'_x et d'_y sont les espacements inter - éléments suivant ox et oy respectivement dans le réseau photonique.

(Dans un réseau 2D (1D), les quatre (deux) cellules de Bragg multiplient la fréquence porteuse par quatre (deux)).

Pour obtenir les mêmes caractéristiques de rayonnement que celles obtenues par le réseau conventionnel, il faut prendre :

$$d'_x = \frac{\lambda_r}{2} = \frac{\lambda_c}{8}$$

$$d'_x = d'_y = \frac{d_x}{4}$$

On constate que pour des caractéristiques de rayonnement données, la dimension du réseau photonique 2D est seize fois plus réduite que celle du réseau purement électronique . La limitation dans ce domaine ne viendra que des couplages mutuelles entre éléments trop proches.

La figure (7) illustre le gain en dimension du cas photonique par rapport au cas conventionnel, et on constate bien que si on augmente le nombre d'éléments, la directivité croît avec une réduction du niveau des lobes secondaires.

- Caractéristiques du système :

La conception électronique et optique du système est relativement simple, elle exige un seul paramètre de contrôle, qui est la fréquence du signal, pour commander la position du faisceau dans l'espace. Ce signal de contrôle étant analogique, permet une réponse graduelle de l'antenne (la réponse suit la variation du paramètre de commande).

On n'a pas besoin de plusieurs signaux de contrôle comme pour les systèmes à déphaseurs numériques.

On remarque qu'il y a une alimentation parallèle directe, à partir de pilotes de courants individuels (photodiodes) vers les éléments respectifs des alignements phasés. Cette alimentation parallèle rend le système plus fiable et augmente sa durée de vie, à l'inverse des systèmes conventionnels utilisant une structure complexe, qui demande des diviseurs micro-ondes de puissance pour la distribution du signal.

4- Conclusion :

L'alimentation photonique se base sur le traitement acousto - optique du signal au niveau des cellules de Bragg. Elles peuvent fonctionner comme modulateurs de phase offrant une bonne résolution, une grande vitesse de traitement (modulation en temps réel), une large gamme dynamique de phase par la variation du signal de commande ainsi qu'un bon rendement (transmission, diffraction), ce qui les rend des dispositifs de choix pour plusieurs applications du traitement optique. Cependant, le traitement A.O peut être sujet au bruit dû à l'effet

de dispersion de la cellule et la diffraction des rayons lumineux indésirables, quand plusieurs fréquences se trouvent simultanément dans la cellule (cas d'une cellule ayant un réseau de transducteurs : Cellule à plusieurs signaux de commande).

Le paramètre unique qui commande la phase facilite la gestion du système par ordinateur, de plus, la possibilité d'intégration sur substrat convenable oriente l'utilisation vers une application aéroportée.

Références :

- [1]- E. Ackerman, M. Marietta « FR fiber optic links » dans « Photonic aspects of modern radar » ch.12, pp 323-350, ed. A-H, 1994.
- [2]- J.L. Corral, J. Marti « True time-delay scheme for feeding optically controlled phased- array antennas using chirped-fiber gratings », IEEE photonics technology letters, vol.9, n°11, novembre 1997, pp 1529-1531.
- [3]- H. Cuhui « Bragg cell receivers », ch.9, pp 327-341, ed. Academic press, 1982.
- [4]- I.C. Chang « Acousto-optic processing for receiver technologies » dans « Photonic aspects of modern radar » ch.16, pp 429-444, ed. A-H, 1994.
- [5]- A. Rihaczek « Principles of high resolution radar » ed. Mc Graw Hill, 1969.
- [6]- L.J. Cutrona « Synthetic aperture radar » ch.23 dans radar handbook, 1^{ère} édition, M.I.Skolnic. ed. New York, 1970.
- [7]- R.D. Esman « Microwave true-time-delay modulator using fiber-optic dispersion » Electronic letters : 24 septembre 1992, vol-28, n° 20, pp 1905-1907.
- [8]- B. Fell « Radar fundamentals for optical engineers » dans « Photonic aspects of modern radar » ch.2, pp 3-7, ed. A-H, 1994.
- [9]- R. Fitterman, Y. Chang « Optically controlled serially fed phased-array transmitter » IEEE microwave and guided wave letters, vol.7, n°3, pp 69-71, mars 1997.
- [10]- M. Gilmore « fiber optic cabling » ed. Newnes, 1994.
- [11]- J.M. Griffin and Wallington, A.R «Optical techniques for signal distribution in phased arrays » journal of light wave technology, n° 2, 1984.
- [12]- L. Harkness, G. Morris « Airborne pulsed doppler radar » ed. A-H, 1996.
- [13]- W.D. Jemison « Optical control of a digital phase shifter », IEEE photonics technology letters, vol.1, n° 1, mars 1988.
- [14]- W.D. Jemison, A. Paoella « Optical waveguide » dans « Photonic aspects of modern radar » ch.8, pp 229-250, ed. A-H, 1994.
- [15]- G. A. Koepf, « Optical processor for phased array antenna beamforming », proc. SPIE, vol. 477, pp 75-81, mai 1984.
- [16]- K. Rossi « Electro-acoustique », ed. pp, Lausanne (Suisse), 1986.
- [17]- K.Y. Lau, « Passive microwave fiber-optic links with gain and very low noise figure », IEEE photonics tech. lett., vol.3, pp.557-559, juin 1991.
- [18]-J.N. Lee « Acousto-optic signal processing and computing », proc.IEEE, vol.77, n°10, pp.1528-1557, octobre 1989.
- [19]- C.H. Lee « Picosecond optoelectronic devices », ed. Academic press, 1984.
- [20]- I. C. Chang, « Acousto-optic channelized receiver » microwave j., p.141, mars 1986.
- [21]-D. Flannery « Fourier optical signal processors », proc. IEEE, vol.77, n°10, pp.1511-1526, octobre 1989.
- [22]-E. N. Leht « Optical processing of synthetic aperture radar data » dans « Photonic aspects of modern radar » ch.14, ed. A-H, 1994.
- [23]-J.P. Hardange, P. Lacomme « Radars aéroportés et spatiaux », ed. Masson, 1995.
- [24]-B. Pedersen « The design of erbium-doped fiber amplifiers », j. lightwave tech. vol.LT.3, pp.1125-1129, octobre 1985.
- [25]-T. Pfeiffer « Output power characteristics of erbium-doped fiber ring laser », IEEE photonics tech.letter, vol.4, pp.847-849, août 1992.
- [26]-A.R. Soref « Silicon-based optical-microwave integrated circuits », microwave j. vol.35, n°5, pp.230-236, mai 1992.
- [27]-R.H. Stolen « Polarization-selective fiber directional coupler », j.lightwave technology, vol.LT.3, pp.1125-1129, octobre 1985.
- [28]-E.N. Touglian, H. Zmuda « Adaptive microwave signal processing: A photonic solution », microwave j., février 1992.
- [29]-E.N. Touglian, H. Zmuda « Photonic aspects of modern radar », ed. A-H, 1994.
- [30]-C.K. Tang « Optical beamsteering using integrated optical modulators », journal of lightwave technology, vol.15, n°12, pp.2259-2263, décembre 1997.
- [31]-Y. Jamani «An active transversal filter MMIC for very high speed lightwave system », IEEE photonics tech.lett.,vol.9, n°6, pp.836-838, juin 1997.
- [32]-K. Po Ho « Efficient photonics mixer with frequency doubling », IEEE photonics tech.lett., vol.9,n°4, pp.511-513, avril 1997.
- [33]-L.B. Lambert, M. Arm, and A. Aimette, « Electro-optical signal processors for phased array antennas » dans Optical information processing, J. T. Tippett, D. A. Berkowitz, L.C.Clapp, eds.(MIT press, Cambridge, Mass., 1965), chap.38.
- [34]-Eyraud, Grange ohanessian « théorie et technique des antennes », Vuibert, 1973.
- [35]- J.Bahl et P.Bhartia « Microstrip antenna »,ed. A-H, 1980, chap.7, pp.247-297.

- [36]- James et Hall « Handbook of microstrip antennas », Peter Perigrinus, 1989.
- [37]-D. Lauvois « L'électronique quantique », P. U.F, 1968.
- [38]-M. Carpentier « Radars base modernes », 6^{ème} édition - 1990-Masson.
- [39]-S. Drabowitch, C. Ancona « Antennes », volume 2 - applications - Masson-1986.
- [40]-Y. Sirotine « Fondements de la physique des cristaux », ed. mir, 1984.
- [41]-H. Maillat « Le laser, principe et techniques d'application », ed. Lavoisier 1990.
- [42]-M. Born « principles of optics », ed. Pergamon press, Oxford, 1970.
- [43]-R.C. Hansen « Phased array antennas », John Wiley & Sons Inc., 1998.
- [44]-Skoudarli et Abdellah « Synthèse des réseaux linéaires d'antennes à lobes secondaires réduits et très réduits », Projet de Fin d'étude - ENP - Janvier, 1989.
- [45]-A. Kumar " Antenna design with fibers optics", chapitre 4, Artech - House, 1996.
- [46]-N. A. Riza " High speed two dimensionnal phased array antenna scanning using acousto - optics", proc. SPIE, vol. 1703, 1992, pp. 460 - 468.
- [47]-D. Dolfi, S. Bann, J.P. Huignard, and J. Roger." Two-dimensional optical architecture for phase and time-delay beam forming in a phased array antenna", Proc. SPIE, Vol. 1703, 1992, pp. 481-489.
- [48]-S.C. Lin, and R.S. Boughton, "A-O multichannel programmable true time delay lines", Proc. SPIE, optical tech. microwave appl. IV, Vol. 1102, 1989, pp. 162-173.
- [49]-Brillouin, L. " Diffusion de la lumière et des rayons X par un corps transparent homogène", Ann. Phys. (Paris), Vol. 17, 1922, pp. 88-122.
- [50]-Dolfi, D., J. P. Huignard, and M. Baril, " Optically controlled true time delays for phased array antenna", Proc. SPIE optical tech. microwave appl. IV, 1102, 1989, pp. 152-161.

Nouvelle Approche des Réseaux Plans Rectangulaires Avec un Niveau des Lobes Secondaire Invariant.

A. CHEMSA* et R. AKSAS*

*Lab. Télécommunications, Ecole Nationale Polytechnique, 10 Avenue Hassen Badi, EL-Harrach, Alger

ملخص:- لقد إقترحنا في عملنا هذا، طريقة جديدة لحساب المميزات الإشعاعية لشبكة مستوية و مستطيلة ذات سعات حث موزونة، وذلك طبقا لقاعدة دولف-تشيبيشيف، بهدف الحصول على أفضل تسوية بين التوجيهية ومستوى الفصوص الثانوية. برهنا في هذا العمل، أن الكتابة النمطية (على الشكل الجيبي) لهذه السعات، جد مفضلة مقارنة مع الكتابات الأخرى المستعملة، سواء من ناحية الوقت اللازم للحساب أو من ناحية معالجة الشبكات الكبيرة.
إن إجتباء شبكة مربعة عن شبكة مستطيلة لها نفس الحجم، من باب التوجيهية، هو الآخر موضح في هذا العمل.

Abstract : We propose a new approach to compute the radiation characteristics of a planar rectangular array with weighted amplitudes excitations using the Dolph-Chebyshev technique, in order to obtain the best compromise between directivity and side lobes level. We showed that the modal form of these amplitudes is more advantageous than the usually used forms in order to reduce the computation time and to treat large scale arrays. We pointed out the amelioration of directivity of a square array with respect to a rectangular array of same dimension.

Résumé : Nous proposons une nouvelle approche du calcul des caractéristiques de rayonnement d'un réseau plan rectangulaire à amplitudes d'excitations pondérées, selon la technique de Dolph-Chebyshev; afin d'obtenir le meilleur compromis directivité-niveau des lobes secondaires. Nous avons montré que la forme modale de ces amplitudes est plus avantageuse que les formes usuellement utilisées et ce aussi bien pour réduire le temps de calcul que pour traiter des réseaux de grande taille. L'amélioration de la directivité d'un réseau carré par rapport à celle d'un réseau rectangulaire, de même taille, a été également mise en évidence.

1. Introduction

Les réseaux d'antennes qui trouvent le plus d'applications aussi bien dans les stations terriennes que marines sont le plus souvent plans. Leur configuration est fréquemment rectangulaire ou circulaire. Par rapport aux alignement (réseaux linéaires), le facteur de réseau de tels aériens comporte d'autres variables qui permettent le contrôle de la forme et de l'orientation du diagramme de rayonnement. On peut ainsi orienter le maximum de ce rayonnement dans n'importe quelle direction de l'espace. Afin d'obtenir des caractéristiques radioélectriques optimales, nous alimentons le réseau à l'aide d'une distribution d'excitation pondérée, de type Dolph-Chebyshev [1]. Une telle distribution offre aux réseaux à espacement uniforme le meilleur compromis directivité-niveau des lobes secondaires.

Dans ce qui suit, nous considérons que les distributions des excitations sont indépendantes dans les deux directions du réseau afin d'utiliser le principe classique de la multiplication des diagrammes de rayonnement. L'objectif essentiel de ce travail est l'élaboration d'une nouvelle formulation du facteur de réseau d'une configuration rectangulaire ($K \times L$) permettant d'englober sous forme d'une expression unique toutes les éventualités que peuvent prendre les valeurs des entiers K et L . Nous montrerons également que par rapport aux expressions classiques des amplitudes pondérées des excitations, un développement modal permet à la fois d'obtenir une réduction du temps de calcul et de meilleurs résultats, particulièrement pour les réseaux de grande taille. Du point de vu directivité, nous mettrons en évidence les meilleures performances d'un réseau carré par rapport à celles d'une géométrie rectangulaire de taille comparable. Les propriétés de symétrie du réseau nous ont amené à développer une nouvelle méthode de détermination de la directivité permettant de calculer aisément et avec précision ce paramètre.

2. Formulation mathématique

Considérons un réseau rectangulaire centré dans le plan xy (Fig. 1). Les $K \times L$ éléments, supposés identiques et à espacement uniforme, sont excités à l'aide d'une distribution à amplitude pondérée, et dont la phase est ajustée à ce que la direction de rayonnement maximum (θ_0, ϕ_0) soit imposée.

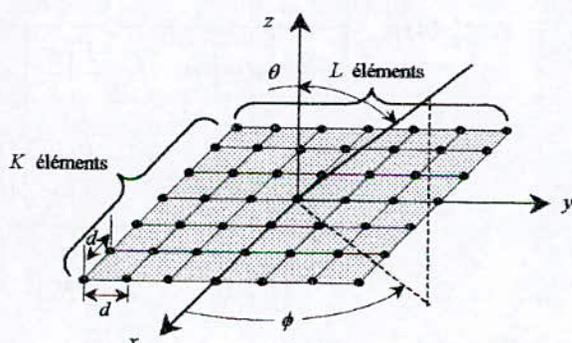


Fig. 1: Réseau rectangulaire à espacement uniforme

2.1. Expression du facteur de réseau

La configuration du réseau permet, afin d'alléger le formalisme mathématique, de considérer la distribution des amplitudes des excitations symétriques par rapport aux axes x et y . Dans ces conditions, nous pouvons écrire (§ annexe A) l'expression du facteur de réseau sous la forme

$$E(\theta, \phi) = F(u, v) = \sum_{m=1}^{M+r_K} \sum_{n=1}^{N+r_L} (2 - r_K \delta_{m-1})(2 - r_L \delta_{n-1}) \cdot I_{mn} \cos[(2m - r_K - 1)u] \cdot \cos[(2n - r_L - 1)v] \quad (1)$$

avec

$$u = \frac{\pi d}{\lambda} (\sin \theta \cos \phi - \sin \theta_0 \cos \phi_0)$$

$$v = \frac{\pi d}{\lambda} (\sin \theta \sin \phi - \sin \theta_0 \sin \phi_0)$$

$\delta_{p-q} = \delta_{pq}$, où δ_{pq} est le symbole de Kronecker, et

$$r_Q = \begin{cases} 0 & \text{si } Q \text{ est pair} \\ 1 & \text{si } Q \text{ est impair} \end{cases} \quad \text{avec } Q = K \text{ ou } L$$

d est la distance entre deux éléments adjacents

λ la longueur d'onde de travail

I_{mn} est l'amplitude de l'excitation de la source d'ordre (m, n) .

Le meilleur compromis, directivité-niveau des lobes secondaires, est obtenu en assimilant ce facteur de réseau à un produit de deux polynômes de Chebyshev d'ordre $(K - 1)$ et $(L - 1)$

$$F(u, v) = f_K(u) \cdot f_L(v) = T_{K-1}(w_0 \cos u) \cdot T_{L-1}(w_0 \cos v) \quad (2)$$

w_0 étant un paramètre (facteur d'échelle) lié au rapport R du niveau du lobe principal par rapport à celui du premier lobe secondaire.

Dans le cas d'une distribution d'amplitude séparable ($I_{mn} = I_m^K \times I_n^L$), les expressions (1) et (2) permettent d'écrire

$$f_K(u) = T_{K-1}(w_0 \cos u) = \sum_{m=1}^{M+r_K} (2 - r_K \delta_{m-1}) I_m^K \cdot \cos[(2m - r_K - 1)u] \quad (3.a)$$

$$f_L(v) = T_{L-1}(w_0 \cos v) = \sum_{n=1}^{N+r_L} (2 - r_L \delta_{n-1}) I_n^L \cdot \cos[(2n - r_L - 1)v] \quad (3.b)$$

Ainsi, on réduit l'étude du réseau rectangulaire à celle de deux alignements indépendantes. Néanmoins, le diagramme de rayonnement obtenu n'est plus optimum

que dans les plans $u = 0$ ou $v = 0$. Dans les autres coupes transverses, ils ne le sont plus car la distribution des amplitudes n'est plus pondérée selon la technique de Dolph-Chebyshev.

L'expression (2), montre que le niveau du lobe principal, situé dans la direction (θ_0, ϕ_0) , et celui de premier lobe secondaire valent respectivement

$$T_{K-1}(w_0) \times T_{L-1}(w_0) \quad (4.a)$$

$$T_{K-1}(w_0) \quad (\text{dans le cas où } K \geq L) \quad (4.b)$$

et on peut donc déduire la relation donnant le facteur d'échelle w_0 en fonction du rapport R par

$$R = T_{L-1}(w_0) \quad (5)$$

2.2. Détermination de la distribution du courant

Le polynôme de Chebyshev

$$T_{Q-1}(w) = \begin{cases} \cosh[(Q-1) \cosh^{-1}(w)] & \text{si } w > 1 \\ \cos[(Q-1) \cos^{-1}(w)] & \text{si } w \leq 1 \end{cases} \quad (6.a)$$

peut être développé [2] sous la forme

$$T_{Q-1}(w) = \sum_{m=1}^{P+r_Q} (-1)^{P-m+r_Q} \frac{2^{2m-r_Q-2} (2P+r_Q-1)}{P+m-1} \cdot C_{P+m-1}^{2m-r_Q-1} \cdot w^{2m-r_Q-1} \quad (6.b)$$

avec
$$\begin{cases} P = M \text{ et } w = w_0 \cos u \text{ pour } Q = K \\ P = N \text{ et } w = w_0 \cos v \text{ pour } Q = L \end{cases}$$

et
$$C_m^n = \frac{m!}{n!(m-n)!}$$

Compte tenu de (6.b), l'expression (2) devient (§ annexe B)

$$F(u, v) = \sum_{m=1}^{M+r_K} \sum_{n=1}^{N+r_L} (2-r_K \delta_{m-1})(2-r_L \delta_{n-1}) \cdot b_m^K \cdot b_n^L \cdot \cos[(2m-r_K-1)u] \cos[(2n-r_L-1)v] \quad (7)$$

avec

$$b_m^K = \sum_{s=m}^{M+r_K} (-1)^{M-s+r_K} \frac{2M+r_K-1}{2(M+s-1)} \cdot C_{M+s-1}^{2s-r_K-1} \cdot C_{2s-r_K-1}^{s-m} \cdot w_0^{2s-r_K-1} \quad (8.a)$$

$$b_n^L = \sum_{s=n}^{N+r_L} (-1)^{N-s+r_L} \frac{2N+r_L-1}{2(N+s-1)} \cdot C_{N+s-1}^{2s-r_L-1} \cdot C_{2s-r_L-1}^{s-n} \cdot w_0^{2s-r_L-1} \quad (8.b)$$

L'identification de (1) avec (7) permet d'obtenir l'amplitude I_{mn} de l'élément d'ordre (m, n) sous la forme

$$I_{mn} = b_m^K \times b_n^L \quad (9)$$

et on peut donc déduire que

$$I_m^K = b_m^K \quad \text{et} \quad I_n^L = b_n^L$$

Généralement dans les problèmes de synthèse des réseaux, la valeur du rapport R est imposée. On en déduit alors le facteur d'échelle w_0 à partir de l'équation (5) pour calculer la distribution des amplitudes I_{mn} à l'aide des relations (8) et (9), et donc l'expression du facteur de réseau pour le tracé du diagramme de rayonnement.

Compte tenu de la symétrie considérée, les résultats obtenus montrent que le nombre d'amplitudes à déterminer passe de $(M+r_K) \times (N+r_L)$ pour une configuration rectangulaire ($K \neq L$) à $(N+r_L)(N+r_L+1)/2$ pour un réseau carré. Le temps de calcul est ainsi réduit de moitié.

Le calcul des amplitudes I_{mn} à partir des relations (8) et (9) est complexe et exige un temps de calcul relativement long. C'est la raison pour laquelle on est conduit à développer (§ annexe C) les amplitudes I_m^K et I_n^L sous la forme modale

$$I_m^K = \sum_{p=1}^{M+r_K} (2-r_K \delta_{p-1}) a_p^K \cdot \cos \left[\frac{2\pi}{K} \left(m - \frac{r_K+1}{2} \right) \left(p - \frac{r_K+1}{2} \right) \right] \quad (10)$$

avec $m = 1, 2, \dots, M+r_K$

qui permet d'écrire l'équation (3.a) comme suit :

$$f_K(u) = T_{K-1}(w_0 \cos u) = \sum_{p=1}^{M+r_K} (2-r_K \delta_{p-1}) a_p^K \psi_p^K(u) \quad (11)$$

où les a_p^K sont des coefficients à déterminer. La fonction ψ_p^K est donnée par

$$\psi_p^K(u) = \frac{1}{2} \left\{ \frac{\sin \left[K \left(u + \left(p - \frac{r_K+1}{2} \right) \frac{\pi}{K} \right) \right]}{\sin \left[u + \left(p - \frac{r_K+1}{2} \right) \frac{\pi}{K} \right]} + \frac{\sin \left[K \left(u - \left(p - \frac{r_K+1}{2} \right) \frac{\pi}{K} \right) \right]}{\sin \left[u - \left(p - \frac{r_K+1}{2} \right) \frac{\pi}{K} \right]} \right\} \quad (12)$$

Compte tenu de fait que

$$\psi_p^K \left[\left(p' - \frac{r_K + 1}{2} \right) \frac{\pi}{K} \right] = \frac{K}{(2 - r_K \delta_{p'-1})} \delta_{p-p'} \quad (13)$$

les coefficients a_p^K et les amplitudes I_m^K deviennent respectivement

$$a_p^K = \frac{1}{K} \cdot T_{K-1} \left[w_0 \cos \left[\left(p - \frac{r_K + 1}{2} \right) \frac{\pi}{K} \right] \right] \quad (14)$$

$$I_m^K = \frac{1}{K} \sum_{p=1}^{M+r_K} (2 - r_K \delta_{p-1}) \cdot T_{K-1} \left[w_0 \cos \left[\left(p - \frac{r_K + 1}{2} \right) \frac{\pi}{K} \right] \right] \cdot \cos \left[\frac{2\pi}{K} \left(m - \frac{r_K + 1}{2} \right) \left(p - \frac{r_K + 1}{2} \right) \right] \quad (15.a)$$

Un développement analogue donne I_n^L sous la forme

$$I_n^L = \frac{1}{L} \sum_{q=1}^{N+r_L} (2 - r_L \delta_{q-1}) \cdot T_{L-1} \left[w_0 \cos \left[\left(q - \frac{r_L + 1}{2} \right) \frac{\pi}{L} \right] \right] \cdot \cos \left[\frac{2\pi}{L} \left(n - \frac{r_L + 1}{2} \right) \left(q - \frac{r_L + 1}{2} \right) \right] \quad (15.b)$$

Il est clair que l'expression des amplitudes I_{mn} , formulée à l'aide de (15.a) et (15.b), est moins complexe que celle obtenue à partir de (8.a) et (8.b). La méthode proposée permet ainsi de réduire le temps de calcul des amplitudes I_{mn} . De plus, elle est plus avantageuse pour les réseaux de grande taille (§ tableau 1).

2.3. Directivité maximale du réseau

La directivité maximale d'un réseau rectangulaire dont le rayonnement maximum est dans la direction (θ_0, ϕ_0) est définie [3] par

$$D_0 = \frac{|E(\theta_0, \phi_0)|^2}{\frac{1}{4\pi} \int_0^\pi \int_0^{2\pi} |E(\theta, \phi)|^2 \sin \theta d\theta d\phi} = \frac{1}{P} |E(\theta_0, \phi_0)|^2 \quad (16)$$

La configuration du réseau considéré nous permet d'obtenir (§ annexe C) l'expression de P sous la forme

$$P = \sum_{m=1}^{M_2+1} \sum_{n=1}^{N_2+1} (2 - \delta_{m-1})(2 - \delta_{n-1}) I_{mn}^{(2)} b_{m-1 n-1} \cdot \cos[(m-1)u_0] \cdot \cos[(n-1)v_0] \quad (17)$$

avec

$$b_{m-1 n-1} = \sin c \left(kd \sqrt{(m-1)^2 + (n-1)^2} \right)$$

$$I_{mn}^{(2)} = I_m^{K_2} \times I_n^{L_2}$$

$$I_m^{K_2} = \frac{1}{K_2} \sum_{p=1}^{M_2+1} (2 - \delta_{p-1}) I_{K-1}^2 \left[w_0 \cos \left[(p-1) \frac{\pi}{K_2} \right] \right] \cdot \cos \left[\frac{2\pi}{K_2} (m-1)(p-1) \right]$$

$$I_n^{L_2} = \frac{1}{L_2} \sum_{q=1}^{N_2+1} (2 - \delta_{q-1}) I_{L-1}^2 \left[w_0 \cos \left[(q-1) \frac{\pi}{L_2} \right] \right] \cdot \cos \left[\frac{2\pi}{L_2} (n-1)(q-1) \right]$$

$$M_2 = K - 1, \quad K_2 = 2M_2 + 1, \quad N_2 = L - 1, \quad L_2 = 2N_2 + 1$$

$$u_0 = kd \sin \theta_0 \cos \phi_0 \text{ et } v_0 = kd \sin \theta_0 \sin \phi_0.$$

Compte tenu de (2), la directivité maximale devient

$$D_0 = \frac{|E(\theta_0, \phi_0)|^2}{P} = \frac{(T_{K-1}(w_0) \times R)^2}{P} \quad (18)$$

Dans le cas d'un espacement de l'ordre de la demi-longueur d'onde ($d \approx \lambda/2$), la quantité $b_{m-1 n-1}$ peut s'écrire, avec une bonne approximation, comme suit :

$$b_{m-1 n-1} \cong \delta_{m-1} \times \delta_{n-1}$$

et l'expression de la directivité maximale est donc

$$D_0 \cong \frac{(T_{K-1}(w_0) \times R)^2}{I_{11}^{(2)}} \quad (19)$$

$$\text{Comme } T_{K-1}(w_0) = \sum_{m=1}^{M_2+1} (2 - \delta_{m-1}) I_m^{K_2} \quad (20.a)$$

$$R = \sum_{n=1}^{N_2+1} (2 - \delta_{n-1}) I_n^{L_2} \quad (20.b)$$

l'expression (19) devient finalement

$$D_0 \cong \sum_{m=1}^{M_2+1} (2 - \delta_{m-1}) I_m^{K_2} / I_1^{K_2} \times \sum_{n=1}^{N_2+1} (2 - \delta_{n-1}) I_n^{L_2} / I_1^{L_2} \quad (21.a)$$

Dans le cas d'un réseau carré ($K = L$), l'expression (21.a) devient

$$D_0 \cong \left(\sum_{n=1}^{N_2+1} (2 - \delta_{n-1}) I_n^{L_2} / I_1^{L_2} \right)^2 = \left(\frac{R^2}{I_1^{L_2}} \right)^2 \quad (21.b)$$

alors que Tseng [3] obtient pour la même structure

$$D_0 \cong 4 \cdot \frac{\left(\sum_{n=1}^N I_n^L \right)^4}{\left(\sum_{n=1}^N (I_n^L)^2 \right)^2} \quad (\text{si } L=2N) \quad (22.a)$$

et

$$D_0 \cong \frac{\left(\sum_{n=1}^{N+1} (2 - \delta_{n-1}) I_n^L \right)^4}{\left(\sum_{n=1}^{N+1} (2 - \delta_{n-1}) (I_n^L)^2 \right)^2} \quad (\text{si } L=2N+1) \quad (22.b)$$

expressions qui donnent, dans les mêmes conditions, des résultats identiques, mais avec un temps de calcul plus long surtout pour une taille L élevée. Cet effet, est illustré dans le tableau 2.

3. Résultats obtenus

Le tableau 1, est un état comparatif des résultats obtenus à l'aide des équations (8) et ceux obtenus par (15). La première colonne représente la taille du réseau carré considéré. Les colonnes 2 et 3 illustrant les temps de calcul (en secondes) des amplitudes I_{mn} en utilisant respectivement (8) et (15), montrent l'avantage évident du développement modal qui nécessite un temps de calcul pratiquement dix fois plus faible. De plus, à l'aide de la relation (20.b), nous avons effectué l'investigation de la validité des approches pour des réseaux de grande taille et, ce en recalculant la valeur du rapport R (fixé au préalable à 30 dB) représentée dans les colonnes 4 et 5. Nous constatons qu'au delà d'une certaine taille ($L=40$), les relations (8) donnent des résultats complètement erronés quoique leur développement mathématique soit rigoureux. Toutefois les résultats obtenus, par l'usage de la fonction factorielle, dépassent les capacités de stockage de l'outil informatique. Même l'usage de la formule approximative de Stirling [5]

$$n! = \sqrt{2\pi n} n^n e^{-n}$$

ne résout pas ce phénomène de débordement [4] car les erreurs s'accroissent lors du calcul, selon (20.b), du rapport R .

Par contre, la méthode modale donne toujours des bons résultats car elle est basée sur l'usage direct des polynômes de Chebyshev sous leur forme (6.a) qui permet de traiter des réseaux carré de n'importe quelle taille. Nous avons testé la validité de cette méthode jusqu'à $L=2000$.

L	Eq. (8)	Eq. (15)	R selon Eq. (8)	R selon Eq. (15)
10	0.11	0.00	30	30
20	0.38	0.06	30	30
30	0.82	0.06	30	30
40	1.43	0.11	30.39	30
50	2.14	0.22	81.81	30
1000	784.89	97.49	∞	30

Tab.1 : Temps de calcul et précision des deux méthodes exposées.

Le tableau 2 permet une étude comparative du temps de calcul (en secondes) nécessaire à l'établissement de la courbe donnant la directivité en fonction de la taille d'un réseau carré dont la valeur de R est fixée à 30 dB.

L_{\max}	t (en sec.) notre approche	t (en sec.) selon [3]
10	0.05	0.05
40	0.27	1.70
70	0.72	8.07
100	1.32	22.14
130	2.08	47.01
160	3.02	85.63
190	4.18	141.32
220	5.61	215.64

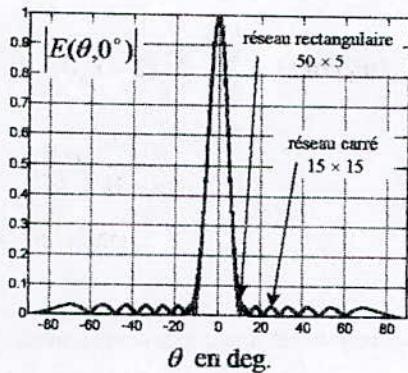
Tab.2 : Temps de calcul de la directivité en fonction de L selon le modèle proposé et celui de Tseng [3].

Nous mettons aussi en évidence l'avantage de l'approche que nous proposons, particulièrement pour les réseaux de grande taille. Ce résultat est d'ailleurs prévisible en analysant les expressions (21.b) et (22). En effet, pour tracer une courbe de directivité en fonction de la taille L d'un réseau carré ($L=2, \dots, L_{\max}$), il faut déterminer $(L_{\max}(L_{\max}+1)(L_{\max}+2)-12)/12$ valeurs dans le cas du modèle élaboré par Tseng [3], et $(L_{\max}-1)(L_{\max}+2)/2$ dans le cas de notre approche, ce qui permet d'obtenir pratiquement un temps de calcul ($L_{\max}/6$) fois plus faible.

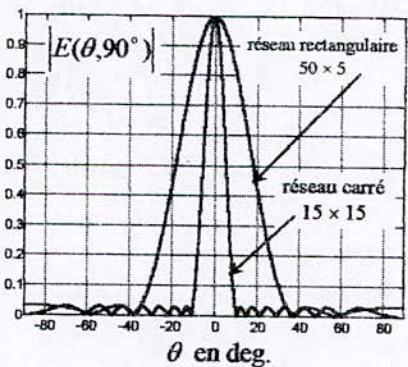
Il est intéressant de comparer les performances de rayonnement d'un réseau carré et celles d'un réseau rectangulaire. Dans cette optique, nous avons effectué les tracés (Fig. 2 et 3) de diagrammes de rayonnement d'un réseau rectangulaire ($K \times L$) et ceux d'un réseau carré ($S \times S$) de taille comparable et ce en prenant S égale à la valeur entière de $\sqrt{K \times L}$. Ces tracés sont obtenus pour les paramètres suivants : $R=30$ dB, $d=\lambda/2$ et $\theta_0=0^\circ$.

Dans le plan xz ($\phi=0^\circ$), les diagrammes de rayonnement sont pratiquement identiques, particulièrement du point de vue de la largeur de lobe principal et de l'angle d'ouverture à demi-puissance.

L'avantage du réseau carré apparaît nettement dans le plan yz ($\phi=90^\circ$) où l'amélioration de la directivité est bien illustrée par la réduction de la largeur de l'angle d'ouverture à -3 dB. Dans le cas de la figure (2.b), cet angle passe de 26° pour une géométrie rectangulaire à 9° pour un réseau carré. Pour la structure de la figure 3, la réduction est moins sensible ; elle passe de 42° à 27° . Ainsi, pour une valeur de K fixée, nous remarquons que cette avantage est d'autant plus prononcé que L est faible.



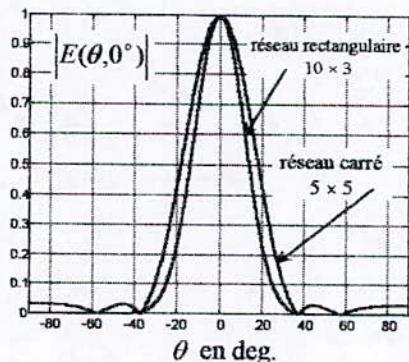
(a)



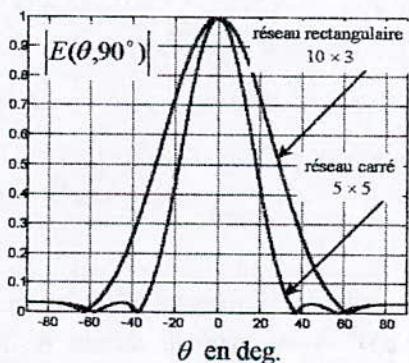
(b)

Fig. 2 Diagrammes de rayonnement du réseau rectangulaire et du réseau carré pour $K = 50$ et $L = 5$.

(a) $\phi = \phi_0 = 0^\circ$ (b) $\phi = \phi_0 = 90^\circ$



(a)



(b)

Fig. 3 Diagrammes de rayonnement du réseau rectangulaire et du réseau carré pour $K = 10$ et $L = 3$.

(a) $\phi = \phi_0 = 0^\circ$ (b) $\phi = \phi_0 = 90^\circ$

La figure 4 illustre bien, que du point de vue de la directivité, le réseau carré offre de meilleures performances qu'un réseau rectangulaire de taille comparable.

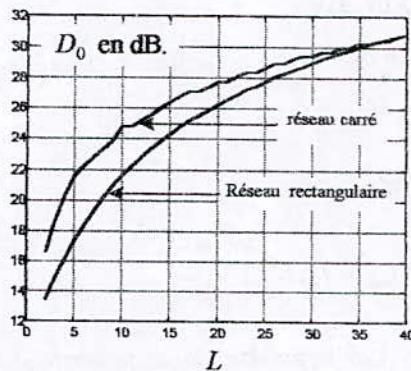


Fig. 4 : Directivité maximale en fonction de L pour $K = 40$

Conclusion

Dans ce travail nous avons élaboré une nouvelle approche de détermination du facteur de réseau en l'écrivant sous une forme plus générale englobant les différents cas possibles de la taille $K \times L$ (pair-pair, pair-impair, impair-pair et impair-impair). L'avantage du développement modal par rapport aux expressions classiques des amplitudes des excitations a été mis en évidence en terme de temps de calcul et de précision. Le modèle que nous proposons permet le tracé de la directivité en fonction de la taille du réseau en un temps d'exécution beaucoup plus réduit que l'approche proposée Tseng [3], particulièrement lorsque la taille de réseau est grande. De plus, nous avons montré que, du point de vue de la directivité, la configuration carrée offre les meilleures performances que celles d'un réseau rectangulaire.

Bibliographie

- [1]. Collin, R. E. and Zucker, F. J., *Antenna theory*, Vol. 7, McGraw-Hill Book Company, New York, 1969, Part1, Chapter 5.
- [2]. Elliott, R. S., *Antenna theory and design*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632, 1981, Chapter 5.
- [3]. Tseng, F-I. and Cheng, D. K., "Optimum Scannable Planar Arrays with an Invariant Sidelobe Level," *IEEE Proc. Antennas and Propagation*, Vol. 56, No. 11, pp. 1771-1778, Nov. 1968.
- [4]. Safaai-Jazi, A., "Modified Chebyshev Arrays" *IEE Proc. Antennas and Propagation*, Vol. 145, No. 1, pp. 45-48, Feb. 1998.
- [5]. Murray Spiegel, R., *Probabilité et Statistiques*, Série Schaum, McGraw-Hill Book Company, New York, traduction de Jacoud R. F, troisième édition, 1983.

Annexe A

L'expression générale du facteur de réseau d'une structure rectangulaire (Fig.1) est

$$E(\theta, \phi) = \sum_{m=-M+1}^{M+r_K} \sum_{n=-N+1}^{N+r_L} i_{mn} e^{jkd \left[\left(m - \frac{r_K+1}{2} \right) \sin \theta \cos \phi + \left(n - \frac{r_L+1}{2} \right) \sin \theta \sin \phi \right]} \quad (A-1)$$

avec

$$i_{mn} = I_{mn} e^{-jkd \left[\left(m - \frac{r_K+1}{2} \right) \sin \theta_0 \cos \phi_0 + \left(n - \frac{r_L+1}{2} \right) \sin \theta_0 \sin \phi_0 \right]} \quad (A-2)$$

Les hypothèse de la symétrie des amplitudes de courant I_{mn} selon les deux axes x et y , se traduit comme suit :

$$I_{-mn} = I_{m+r_K+1, n} \quad I_{m-n} = I_{m, n+r_L+1} \quad (A-3)$$

En décomposant les termes en exponentiel de l'équation (A-1) en cosinus et en utilisant (A-3), on trouve aisément l'expression (1).

Annexe B

La relation (6), s'écrit pour $w = w_0 \cos t$

$$T_{Q-1}(w_0 \cos t) = \sum_{m=1}^{P+r_Q} (-1)^{P-m+r_Q} \frac{2^{2m-r_Q-2} (2P+r_Q-1)}{P+m-1} \cdot C_{P+m-1}^{2m-r_Q-1} \cdot w_0^{2m-r_Q-1} (\cos t)^{2m-r_Q-1} \quad (B-1)$$

avec $Q = 2P + r_Q$ (r_Q est le reste de division de Q par 2).

Le développement de la quantité $(\cos t)^{2m-r_Q-1}$ donne

$$(\cos t)^{2m-r_Q-1} = 2^{-(2m-r_Q-1)} \sum_{l=1}^m (2-r_Q \delta_{m-p'}) C_{2m-r_Q-1}^{p'-1} \cdot \cos[(2m-2p'-r_Q+1)t] \quad (B-2)$$

et l'équation (B-1) devient alors

$$T_{Q-1}(w_0 \cos t) = \sum_{m=1}^{P+r_K} \sum_{p'=1}^m (2-r_Q \delta_{m-p'}) (-1)^{P-m+r_Q} \cdot \frac{2P+r_Q-1}{2(P+m-1)} C_{P+m-1}^{2m-r_Q-1} \cdot C_{2m-r_Q-1}^{p'-1} \cdot w_0^{2m-r_Q-1} \cdot \cos[(2m-2p'-r_Q+1)t] \quad (B-3)$$

En posant dans (B-3) $p = m - p' + 1$, on trouve

$$T_{Q-1}(w_0 \cos t) = \sum_{m=1}^{P+r_Q} \sum_{p=1}^m (2-r_Q \delta_{p-1}) (-1)^{P-m+r_Q} \cdot \frac{2P+r_Q-1}{2(P+m-1)} \cdot C_{P+m-1}^{2m-r_Q-1} \cdot C_{2m-r_Q-1}^{m-p} \cdot w_0^{2m-r_Q-1} \cdot \cos[(2p-r_Q-1)t] \quad (B-4)$$

En permettant les deux opérateurs somme, on obtient

$$T_{Q-1}(w_0 \cos t) = \sum_{p=1}^{P+r_Q} \sum_{m=p}^{P+r_Q} (2-r_Q \delta_{p-1}) (-1)^{P-m+r_Q} \cdot \frac{2P+r_Q-1}{2(P+m-1)} \cdot C_{P+m-1}^{2m-r_Q-1} \cdot C_{2m-r_Q-1}^{m-p} \cdot w_0^{2m-r_Q-1} \cdot \cos[(2p-r_Q-1)t] \quad (B-5)$$

En remplaçant p par m ou n , et m par s , on retrouve (8.a) pour $Q = K$, $P = M$ et $t = u$, et (8.b) pour $Q = L$, $P = N$ et $t = v$.

Annexe C

En substituant la relation (A-2) dans (A-1), et en prenant le carré de $E(\theta, \phi)$, on trouve

$$|E(\theta, \phi)|^2 = (E(\theta, \phi))^2 = \sum_{m_1=-M+1}^{M+r_K} \sum_{m_2=-M+1}^{M+r_K} \sum_{n_1=-N+1}^{N+r_L} \sum_{n_2=-N+1}^{N+r_L} I_{m_1 n_1} I_{m_2 n_2} e^{j2[(m_1+m_2-r_K)u + (n_1+n_2-r_L)v]} \quad (C-1)$$

En posant

$$m = m_1 + m_2 - r_K \quad \text{et} \quad n = n_1 + n_2 - r_L$$

(C-1) devient

$$|E(\theta, \phi)|^2 = \sum_{m=-M_2+1}^{M_2+1} \sum_{n=-N_2+1}^{N_2+1} I_{mn}^{(2)} e^{j2[(m-1)u + (n-1)v]} \quad (C-2)$$

avec $M_2 = 2M + r_K - 1$, $N_2 = 2N + r_L - 1$;

$$I_{mn}^{(2)} = \sum_{m_2=\max(m-M, -M+1)}^{\min(m+M+r_K-1, M+r_K)} \sum_{n_2=\max(n-N, -N+1)}^{\min(n+N+r_L-1, N+r_L)} I_{m-m_2+r_K, n-n_2+r_L} I_{m_2 n_2} \quad (C-3)$$

A partir de la relation (C-2), on peut assimiler $(E(\theta, \phi))^2$ à un nouveau facteur de réseau d'une géométrie rectangulaire de taille $K_2 \times L_2$, avec $K_2 = 2M_2 + 1$ et $L_2 = 2N_2 + 1$, excitée à l'aide d'une distribution séparable d'amplitudes $I_{mn}^{(2)}$ ($I_{mn}^{(2)} = I_m^{K_2} \times I_n^{L_2}$) également symétriques par rapport

aux axes x et y . La méthode modale appliquée à $(E(\theta, \phi))^2$ donne alors

$$I_m^{K_2} = \frac{1}{K_2} \sum_{p=1}^{M_2+1} (2 - \delta_{p-1}) I_{K-1}^2 \left[w_0 \cos \left[(p-1) \frac{\pi}{K_2} \right] \right] \cdot \cos \left[\frac{2\pi}{K_2} (m-1)(p-1) \right] \quad (\text{C-4.a})$$

$$I_n^{L_2} = \frac{1}{L_2} \sum_{q=1}^{N_2+1} (2 - \delta_{q-1}) I_{L-1}^2 \left[w_0 \cos \left[(q-1) \frac{\pi}{L_2} \right] \right] \cdot \cos \left[\frac{2\pi}{L_2} (n-1)(q-1) \right] \quad (\text{C-4.b})$$

L'utilisation de la relation (C-2) permet d'obtenir la quantité P , introduite dans (16), sous la forme

$$P = \sum_{m=-M_2+1}^{M_2+1} \sum_{n=-N_2+1}^{N_2+1} I_{mn}^{(2)} b_{m-1 \ n-1} \cdot e^{-jkd \sin \theta_0 [(m-1) \cos \phi_0 + (n-1) \sin \phi_0]} \quad (\text{C-5})$$

avec

$$b_{m-1 \ n-1} = \sin c \left(kd \sqrt{(m-1)^2 + (n-1)^2} \right) \quad (\text{C-6})$$

La propriété de symétrie des $I_{mn}^{(2)}$ et $b_{m-1 \ n-1}$ par rapport aux valeurs $m=1$ et $n=1$, entraîne l'expression (17).

SEGMENTATION NON SUPERVISEE D'IMAGES SAR PAR APPLICATION DE L'ALGORITHME EM.

A. BOUDAIEB* et A. ZERGUERRAS**.

ملخص: يهدف هذا العمل إلى تطبيق طريقة التقطيع بدون مراقبة على صور رادار ذو مجال تركيبي. و لأجل هذا قمنا بتركيب إشارة الرادار مع نموذج أصناف التقطيع. هذا الأخير، أفترض أنه حقل عشوائي لمركوف لكي يسمح لنا بالحصول على دالة توزيع تالية للصورة. الطريقة المستعملة تعتمد أساسا على خوارزمية الأمل الأعظمي، و عليه تم تقدير مختلف الوسائط بوجهة متناوبة. أخيرا، قدمت أمثلة لتوضيح مدى فعالية المتوال المدروس.

كلمات مفتاحية: حقل عشوائي لمركوف، صور رادار ذو مجال تركيبي، خوارزمية الأمل الأعظمي، طريقة التقطيع بدون مراقبة، خوارزمية التسخين التمثيلي، خوارزمية الأنماط الشرطية المتناوبة.

Résumé: Un modèle statistique est développé pour la segmentation non supervisée des images SAR (Synthetic Aperture Radar). A cet effet, la représentation des données SAR est combinée avec le modèle statistique des classes, fondé sur le champ aléatoire de Markov (MRF) afin d'obtenir une distribution a posteriori. La méthode de segmentation utilisée s'appuie sur l'algorithme de l'espérance maximale EM (Expectation-Maximisation), dans la mesure où l'estimation des paramètres s'obtient d'une façon itérative. Des exemples illustrent l'efficacité de cette stratégie et sa robustesse au bruit speckle.

Mots clés: Champ aléatoire de Markov, radar à ouverture synthétique, algorithme EM, segmentation non supervisée, algorithme RS(recuit simulé), algorithme des modes conditionnels itérés.

Abstract: A statistical model is developed for an unsupervised segmentation of SAR images (Synthetic Aperture Radar). To this effect, the representation of the SAR data is combined with class model. This last model is considered to be a Markov random field in order to get a posteriori distribution to the image. The used segmentation is based on the EM algorithm (Expectation-Maximization), that leads to an iterative procedure for parameters estimation. Finally, experiments results have been provided to illustrate the efficiency of this strategy.

Keywords: Markov random field, synthetic aperture radar, EM algorithm, unsupervised segmentation, simulated annealing algorithm, iterated conditionnels modes algorithm (ICM).

* Post-graduant au Département Electronique- Option Radar.

** Professeur à l'Ecole Nationale Polytechnique- Département Electronique- Laboratoire Télécommunications.

1. INTRODUCTION

L'imagerie radar à visée latérale et à ouverture synthétique (SAR) a considérablement pris de l'importance depuis qu'il est devenu possible d'envoyer des satellites ou avions d'observations autour de la terre. En effet, cette technique de télédétection d'hyperfréquences a l'avantage, par rapport à la télédétection optique, d'être insensible à la luminosité et aux conditions d'environnement d'ordre météorologique [1]. L'imagerie radar a trouvé de très nombreuses applications, que ce soit dans le domaine civil pour la cartographie géologiques(sol et sous-sol), l'hydrologie des surfaces continentales, l'océanographie (étude des vagues et des courants), l'agriculture(suivie de croissance de culture), et l'urbanisme ou dans le domaine militaire pour la surveillance des territoires.

L'information portée par le signal SAR est fortement affectée par un bruit multiplicatif appelé *speckle*. Il est dû à la sommation cohérente de plusieurs signaux provenant d'un nombre très grand de points diffuseurs. A cause de sa nature multiplicative, le *speckle* masque les frontières séparant les différentes parties(textures ou régions) de l'image. Ainsi, il rend difficile la segmentation et par suite l'interprétation des images acquises par le radar. En conséquence, plusieurs techniques ont été élaborées pour modéliser le *speckle* et réduire son effet. Un moyen de réduire le *speckle* est d'utiliser la méthode multilook(multivue), qui consiste à moyenner M images vues sous des angles différents, ce qui améliore d'autant le rapport signal sur bruit et dégrade la résolution de l'image[1]. Ces dernières années, des algorithmes ont été proposés pour le traitement optimal de l'image SAR, en se basant sur le modèle statistique du *speckle*. Plus précisément, ces méthodes supposent que la distribution du *speckle* suit la loi Gamma pour l'intensité de l'image alors que la loi de Rayleigh est celle de l'amplitude de l'image. Cette distribution a été utilisée pour développer les algorithmes de détection de contours [2] et filtrage [3, 4] des images SAR.

L'intérêt récent portant sur l'usage des techniques de segmentation conduit à modéliser les régions de l'image par un champ aléatoire de Markov (MRF) [5] dont le principal avantage est de fournir une distribution jointe de l'image, qui permet de prendre en compte les différentes contraintes et les données a priori(modèle du *speckle*) au moyen de la règle de Bayes. Le problème de segmentation dans ce cas consiste à assigner à chaque pixel une étiquette de $\{1,2,\dots,K\}$ pour laquelle le

logarithme de la distribution jointe est maximal. Les meilleurs résultats sont obtenus en utilisant le recuit simulé qui est une technique de relaxation stochastique [6, 7], dont la séquence d'images générée converge vers l'estimateur MAP(*Maximum a Posteriori*). Le passage d'un état à un autre se fait par un tirage aléatoire [6, 8]. Pour éviter les maxima locaux, le recuit simulé exige la diminution lente d'un paramètre, appelé température, attaché aux densités de probabilités conditionnelles locales. L'inconvénient du recuit simulé est qu'il nécessite un temps de calcul élevé. Une solution à ce problème, est d'utiliser l'algorithme des modes conditionnels itérés (ICM)[9] qui est à relaxation déterministe. Cette méthode converge vers un maximum local, mais donne encore de bons résultats [5].

Dans le cas où le nombre de régions et les paramètres des modèles sont connus a priori, ou font l'objet d'une estimation préalable en se basant sur des échantillons d'apprentissage, la segmentation est dite supervisée. Les résultats de cette technique sont de bonne qualité [10, 11]. Dans les applications réelles, le nombre de régions et de paramètres sont souvent inconnus, on doit donc les déterminer à partir de l'image observée. Ainsi, Cette situation conduit au problème des données incomplètes. En effet, pour estimer les paramètres on doit disposer d'une réalisation de l'image segmentée et pour faire la segmentation, les paramètres doivent être connus. Pour résoudre ce problème on fait appel aux algorithmes de segmentation non supervisée. Ces algorithmes réalisent l'estimation des paramètres et la segmentation récursivement. Plusieurs méthodes basées sur cet aspect ont été proposées dans [12, 13, 14]. Dans plusieurs cas pratiques, l'algorithme EM s'est avéré efficace par essence[14].

La segmentation non supervisée adoptée ici est basée sur le champ aléatoire de Markov, pour pouvoir estimer au sens du maximum a posteriori. Pour cela, un modèle de distribution représentant les données SAR dans une classe est combiné avec un modèle de Markov représentant la distribution des classes. L'estimation des paramètres est réalisée par l'algorithme de l'EM et la mise à jours des étiquettes est faite par une technique similaire à l'ICM. Les testes effectués sur des images synthétiques ont montré l'efficacité de cette stratégie dans l'obtention d'une bonne classification des pixels.

2. LE CHAMP MARKOVIEEN DES REGIONS

L'image du champ des régions est prise comme une réalisation discrète x d'un champ aléatoire X , défini sur une maille L de $N1 \times N2$ sites (pixels): $X = \{X_{i,j}; (i,j) \in L\}$. Chaque variable aléatoire $X_{i,j}$ est définie sur l'ensemble des étiquettes $Q = \{1, 2, \dots, K\}$ où K est le nombre de classes typant les régions. Alors, $X_{ij} = x_{ij} = k$ dénote le fait que le pixel (i,j) dans la réalisation x appartient à la région de type k . Donc, une réalisation $x = X(\omega)$ est simplement une partition de la maille L en K régions types, pouvant chacune occuper plusieurs emplacements dans la maille L . A chaque région dans une image correspond une zone de la surface illuminée par le radar. Ainsi, il existe une continuité spatiale dans l'image, en laquelle chaque pixel a une forte probabilité d'avoir le même type de région que ses voisins immédiats. Cette contrainte est incorporée dans le modèle en supposant que X est un champ aléatoire de Markov (MRF) par rapport au système de voisinage $\eta = \{\eta_{ij}; (i,j) \in L, \eta_{ij} \subset L\}$ où η_{ij} , le voisinage du pixel (i,j) , est tel que $(k,l) \in \eta_{ij}$ alors $(i-k)^2 + (j-l)^2 \leq m$, où m est l'ordre de η [6, 7]. Par définition, le MRF X a la propriété que la probabilité conditionnelle pour qu'un pixel appartienne à un type région donné connaissant toutes les valeurs de X dans les autres pixels, égale la probabilité conditionnelle connaissant les valeurs de X au voisinage de ce pixel. Ceci est exprimé par :

$$P(X_{ij} = x_{ij} / X_{kl} = x_{kl}, (k,l) \neq (i,j)) = P(X_{ij} = x_{ij} / X_{kl} = x_{kl}, (k,l) \in \eta_{ij}) \tag{1}$$

Avec la condition de positivité suivante:

$$P(X(\omega) = x) > 0 \text{ pour tout } \omega \in \Omega \tag{2}$$

où Ω est l'ensemble des configurations possibles de X (i.e. $\text{card}(\Omega) = K^{N1 \times N2}$).

Le théorème d'Hammersley-Clifford [6], énonce qu'un champ aléatoire de Markov X suit une loi de probabilité de Gibbs donnée par :

$$P(X = x) = \frac{\exp(-U(x))}{Z} \tag{3}$$

où Z est la fonction de répartition définie par :

$$Z = \sum_{\omega \in \Omega} \exp(-U(x)) \tag{4}$$

$U(x)$ est appelée la fonction d'énergie, et représente une somme de fonctions locales $V_c(x)$, dites fonctions de potentiel :

$$U(x) = \sum_{c \in C} V_c(x) \tag{5}$$

Où c est un ensemble de pixels, appelé clique, dont chaque pair de pixels distincts sont voisins au sens de Markov. C désigne l'ensemble de toutes les cliques engendrées par le système de voisinage η à travers la maille L . Les types de cliques associés avec le système de voisinage de premier ordre ($m=1$) et du second ordre ($m=2$) sont donnés à la figure 1.

Pour notre application nous avons utilisé un système de second ordre, générant trois types de cliques, parmi lesquels nous avons retenu que les cliques d'ordre deux (deux pixels voisins). Nous avons choisi pour la fonction de potentiel $V_c(x)$ le Modèle MLL (Multilevel-logistic), donné par :

$$V_c(x) = \begin{cases} -\beta & \text{Si tous les } x_{ij} \text{ dans } c \text{ sont égaux.} \\ +\beta & \text{Dans le cas contraire} \end{cases} \tag{6}$$

β est un paramètre positif, son rôle est de favoriser l'homogénéité spatiale des pixels. En effet, sa valeur influe sur la taille et la forme des régions dans l'image.

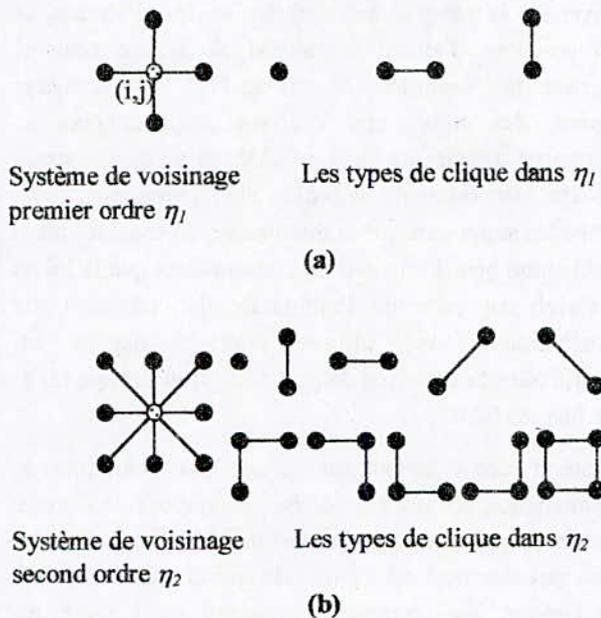


Figure 1. systèmes de voisinage premier ordre et second ordre et leurs types de cliques.

3. STATISTIQUES DES DONNEES SAR DANS UN TYPE REGION

La segmentation correcte de l'image (reflectivité de la surface), réalise le champ de régions $x=X(\omega)$, mais elle n'est pas observable. Par ailleurs, on observe une image specklée(bruitée), dans laquelle la quantité utilisée pour coder le niveau de gris de chaque pixel, est l'amplitude du signal rétrodiffusé par la surface étudiée. La zone imagée est supposée composée d'un très grand nombre de diffuseurs élémentaires. A un site(pixel) l'écho radar est donc la somme des signaux réfléchis par les diffuseurs élémentaires contenus dans une cellule de résolution. Ainsi, l'amplitude complexe du signal reçu est :

$$Z = Y e^{j\phi} = \sum_{k=1}^N a_k e^{j\phi_k} \quad (7)$$

où a_k et ϕ_k sont respectivement l'amplitude et la phase associées avec le k^{ieme} diffuseur et N et le nombre des diffuseurs dans une cellule de résolution.

La répartition aléatoire des diffuseurs au sein d'une cellule de résolution, ainsi que les variations de leurs caractéristiques électromagnétiques, font que a_k et ϕ_k sont des variables aléatoires. Par conséquent, les interférences entre ces signaux résultent en un motif granulaire au niveau de l'image, connu sous le nom de chatoiement ou speckle. Si on admet que pour chaque diffuseur a_k et ϕ_k sont indépendants, les ϕ_k sont uniformément distribuées sur $[-\pi, \pi]$ et si le nombre de diffuseurs est grand lorsque ($N \rightarrow \infty$), on peut montrer par l'utilisation du théorème centrale limite [12] que la partie réelle de Z , $Re[Z]=Z_R$, et la partie imaginaire de Z , $IM[Z]=Z_I$, sont des variables aléatoires gaussiennes non corrélées ($E(Z_R Z_I)=0$) et de moyennes nulles. Dans l'image amplitude, le niveau de gris observé à un pixel est le module de l'amplitude complexe Z

$$Y = \sqrt{Z_R^2 + Z_I^2} \quad (8)$$

On peut montrer que conditionnellement à une région type donné ($X=k$), Y suit une loi de probabilité de Rayleigh donnée par:

$$P_Y(y / X = k) = \begin{cases} \frac{y}{\sigma_k^2} e^{-\frac{y^2}{2\sigma_k^2}} & \text{si } y \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (9)$$

où σ_k est l'écart type de Z_R et Z_I dans la région de type k .

Si on considère une " image intensité ": $I=Y^2$, le signal reçu suit alors une distribution exponentielle de la forme:

$$P_I(i / X = k) = \begin{cases} \frac{1}{2\sigma_k^2} e^{-\frac{i}{2\sigma_k^2}} & \text{si } i \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (10)$$

avec $E(I) = (2\sigma_k^2)$ $Var(I) = (2\sigma_k^2)^2$.

L'image radar dont on vient de faire l'étude est appelée image 1-look ou monovue. En pratique, on est amené parfois à faire le moyennage de plusieurs images de la même surface, vu sous divers angles(image multivues : M-looks), afin de réduire la variance du speckle. Si M réalisations indépendantes sont moyennées, alors à chaque pixel on observe la quantité :

$$V = \frac{1}{M} \sum_{l=1}^M I_l \quad (11)$$

Où I_l est l'image intensité 1-look. Dans ce cas, la probabilité conditionnelle V relative à un type de région $X=k$ est une distribution Gamma[12], définie par :

$$P_V(v / X = k) = \begin{cases} \frac{v^{M-1} e^{-\frac{v}{2\sigma_k^2/M}}}{(M-1)! (\frac{2\sigma_k^2}{M})^M} & \text{si } v \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (12)$$

sachant que pour une image multivue on a

$$E(V) = (2\sigma_k^2) \quad Var(V) = \frac{(2\sigma_k^2)^2}{M}$$

Donc $E(V)/[Var(V)]^{1/2} = M^{1/2}$, ainsi l'écart type est proportionnel à la valeur moyenne. Par conséquent, le speckle est un bruit multiplicatif.

L'image amplitude M-looks s'écrit :

$$Y_M = \sqrt{V} = \left(\frac{1}{M} \sum_{l=1}^M I_l \right)^{1/2} \quad (13)$$

La probabilité de Y_M relative à un type de région k est la distribution :

$$P_{Y_M}(y_M / X = k) = \begin{cases} 2 \frac{y_M^{2M-1} e^{-\frac{y_M^2}{2\sigma_k^2/M}}}{(M-1)! (\frac{2\sigma_k^2}{M})^M} & \text{si } y_M \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (14)$$

Il à noter que la distribution(de Rayleigh ou exponentielle) de l'image 1-vue (en amplitude ou en

intensité) est un cas particulier de celle de l'image M -vues pour $M=1$.

4. LE PROBLEME DE LA SEGMENTATION

Notre approche de segmentation consiste à attribuer à chaque pixel de l'image une étiquette parmi K étiquettes possibles. Le but est que pour une réalisation d'une image specklée y donnée, on doit déterminer la réalisation des champs des régions x qui a donné naissance à y . La réalisation x est non mesurable et ne peut être trouvée directement à partir de y . Dans ce cas le problème se résume à trouver l'estimée $\hat{x} = \hat{X}(y)$ de la scène X , en se basant sur la réalisation y . Ayant formulé le modèle statistique, on adopte l'estimation de maximum a posteriori (MAP) comme un critère statistique. Ainsi, l'objectif maintenant est d'avoir un algorithme d'estimation, lequel fournira l'estimée \hat{x} qui maximise la distribution a posteriori $P(X=x/Y=y)$ pour y donnée. Dans la section suivante on donne un algorithme qui réalise cette tâche.

5. ALGORITHME DE SEGMENTATION MAP

Notre but maintenant est de déterminer l'estimée \hat{x} au sens du critère MAP :

$$\hat{x} = \arg \max_x P(X = x / Y = y) \quad (15)$$

Par l'utilisation du théorème de Bayes, on aura :

$$P(X = x / Y = y) = \frac{P(Y = y / X = x)P(X = x)}{P(Y = y)} \quad (16)$$

Comme $P(Y=y)$ est indépendant de X ceci implique que :

$$\hat{x} = \arg \max_x P(Y = y / X = x)P(X = x) \quad (17)$$

Si on suppose que les échantillons du bruit sont indépendants et que X et Y sont indépendants, le problème de segmentation se réduit à maximiser à chaque pixel la distribution conditionnelle locale suivante [12] :

$$\begin{aligned} &P\{X_{ij} = x_{ij} / X_{kl} = x_{kl}, Y = y, (k,l) \neq (i,j)\} = \\ &P\{Y_{ij} = y_{ij} / X_{ij} = x_{ij}\} P\{X_{ij} = x_{ij} / X_{kl} = x_{kl}, (k,l) \in \eta_{ij}\} \end{aligned} \quad (18)$$

Où $P\{Y_{ij} = y_{ij} / X_{ij} = x_{ij}\}$ est donnée par (12-14) et

$$P\{X_{ij} = x_{ij} / X_{kl} = x_{kl}, (k,l) \in \eta_{ij}\} \text{ est donnée par (3).}$$

L'algorithme du RS [6] est souvent utilisé pour trouver une solution itérative pour un tel problème. Quoique cette technique de relaxation stochastique nécessite un temps de calcul élevé. Dans ce travail, on utilise un algorithme de relaxation déterministe ICM (modes conditionnels itérés) qui nécessite un temps de calcul très réduit[9]. La solution de l'ICM est obtenue en affectant à chaque pixel l'étiquette qui maximise l'expression (18). Le schéma de cet algorithme est donné comme suit :

- (1) **initialiser le champ des régions X pour chaque pixel (i,j) , en choisissant la valeur de x_{ij} dans $\{1,2,\dots,K\}$ qui maximise (12 ou 14).**
- (2) **Pour $i=1,\dots,N_1, j=1,\dots,N_2$ mise à jours de \hat{x}_{ij} par l'étiquette qui maximise (18).**
- (3) **Répéter (2) Jusqu'à ce que le nombre de pixels qui changent devienne inférieur à un seuil fixé.**

6. LE PROBLEME DES DONNEES INCOMPLETES ET L'ALGORITHME EM

Dans le but d'implémenter la procédure de segmentation de la section précédente, d'une part on a besoin des paramètres du modèle $(\beta, \sigma_k^2)_{k=1,\dots,K}$. Cependant, dans la plupart du temps, ces paramètres sont inconnus et doivent être estimés. D'autre part pour les estimer, on doit disposer d'une réalisation du champ des régions X (image segmentée). Une solution raisonnable à ce problème serait l'estimation des paramètres et la segmentation de l'image récursivement. Utilisons l'algorithme EM pour calculer les estimateurs de type maximum de vraisemblance (ML) des paramètres à partir des données incomplètes[14]. En effet, soit $W=(Y, X)$ un champ aléatoire composé des champs Y et X respectivement observé et non observé (régions d'image). Par convention, Y est appelé données incomplètes et W est appelé données complètes.

Le problème des données incomplètes consiste à estimer l'ensemble des paramètres en se basant seulement sur les données incomplètes Y . Ce problème est très compliqué, en utilisant l'approche ML conventionnelle par le fait que la partie des données relative à x est absente. Ainsi, l'algorithme EM est une technique efficace pour résoudre ce problème. Soit $P(W/\theta) = P(Y, X/\theta)$ la densité de probabilité des données complètes W (la densité conjointe de X et Y), où θ est le vecteur de paramètre à déterminer, et donnant un estimateur initial $\hat{\theta}^{(0)}$. Alors, à la $p^{\text{ème}}$ itération, l'algorithme EM procède en deux étapes :

Etape E) Calculer l'espérance conditionnelle :

$$Q(\theta / \hat{\theta}^{(p)}) = E[\log(P(W / \theta)) / Y, \hat{\theta}^{(p)}] \quad (19)$$

Etape M) Trouver $\hat{\theta}^{(p+1)}$ qui maximise $Q(\theta / \hat{\theta}^{(p)})$:

$$\hat{\theta}^{(p+1)} = \arg \max_{\theta} Q(\theta / \hat{\theta}^{(p)}) \quad (20)$$

L'obstacle principal lors de l'implémentation de l'algorithme EM réside dans la difficulté de trouver une expression analytique de l'espérance conditionnelle $Q(\theta / \hat{\theta}^{(p)})$ dans l'étape E. Ceci n'est facile que pour le cas des modèles simples. Dans la section qui suit, on propose une procédure qui produit récursivement les estimateurs des paramètres ainsi que la segmentation, qui correspond au modèle statistique de l'image SAR, en se basant sur l'algorithme EM.

7. SEGMENTATION AVEC ESTIMATION DES PARAMETRES

Dans cette partie, on adopte un dénombrement simple des sites de la maille L en assignant le nombre $t=j+Nl x(i-1)$ au site (i,j) . Cette classification dénombre les sites ligne par ligne de 1 à $n=NlxN2$, en commençant par le site(1,1). De cette manière, l'image observée sera représentée par un vecteur de n composantes, $y=\{y_1, y_2, \dots, y_n\}$ et similairement l'image régions, $x=\{x_1, x_2, \dots, x_n\}$. Dans ce cas les x_i sont des vecteurs unitaires de K composantes, indiquant la classe d'appartenance du site i. Si x_i est attribué à la classe k alors on écrit $x_i=e_k$ où e_k est le vecteur unitaire dont le k^{ème} composant est égal à 1. On peut alors formuler le problème de la segmentation non supervisée en utilisant l'algorithme EM de la manière suivante.

L'ETAPE E

Calculer l'espérance conditionnelle $Q(\theta / \hat{\theta}^{(p)})$ pour le modèle statistique choisi. On a $P(W / \theta) = P(Y / X, \theta_Y) P(X / \theta_X)$ où θ_Y est le vecteur de paramètres attaché au modèle des données et θ_X est celui du modèle des régions(i.e. $\theta = (\theta_Y, \theta_X)$), Donc :

$$\log(P(W / \theta)) = \log[P(Y / X, \theta_Y)] + \log[P(X / \theta_X)] \quad (21)$$

Si les échantillons du champ Y sont indépendants alors $P(Y / X, \theta_Y)$ peut s'écrire :

$$P(Y / X, \theta_Y) = \prod_{i=1}^n P(Y_i = y_i / X_i = x_i, \theta_Y) \quad (22)$$

Pour une implémentation efficace de l'algorithme de segmentation non supervisée on va utiliser l'approximation de la pseudo-vraisemblance concernant la distribution du champ de régions $P(X / \theta_X)$ [9] :

$$P(X / \theta_X) \approx \prod_{i=1}^n P(X_i = x_i / X_j = x_j, j \in \eta_i, \theta_X) \quad (23)$$

où η_i est le voisinage immédiat du site i dans la présente représentation de la maille L.

Cela signifie la distribution jointe de X égale au produit des distributions conditionnelles locales $P(X_i = x_i / X_j = x_j, j \in \eta_i, \theta_X)$.

Ainsi (21) aura la forme suivante:

$$\begin{aligned} \log(P(W / \theta)) &= \sum_{i=1}^n \log[P(Y_i = y_i / X_i = x_i, \theta_Y)] \\ &+ \sum_{i=1}^n \log[P(X_i = x_i / X_j = x_j, j \in \eta_i, \theta_X)] \end{aligned} \quad (24)$$

Puisque x_i est un vecteur dont les composantes sont toutes nulles sauf une composante(i.e. $x_i = [0, 0, \dots, 1, 0, \dots, 0]^T$) donc on aura:

$$\log[P(Y_i = y_i / X_i = x_i, \theta_Y)] = x_i^T U(y_i / \theta_Y) \quad (25)$$

ainsi que

$$\log[P(X_i = x_i / X_j = x_j, j \in \eta_i, \theta_X)] = x_i^T V(\eta_i, \theta_X) \quad (26)$$

Où

$$U(y_i / \theta_Y) = [\log[P(y_i / e_1, \theta_Y)], \dots, \log[P(y_i / e_K, \theta_Y)]]^T \quad (27)$$

et

$$V(\eta_i, \theta_X) = [\log[P(e_1 / x_j, j \in \eta_i, \theta_X)], \dots, \log[P(e_K / x_j, j \in \eta_i, \theta_X)]]^T \quad (28)$$

Par conséquent, l'expression de la fonction $Q(\theta / \hat{\theta}^{(p)})$ devient:

$$\begin{aligned} Q(\theta / \hat{\theta}^{(p)}) &= \sum_{i=1}^n E[x_i^T / Y, \hat{\theta}^{(p)}] U(y_i / \theta_Y) \\ &+ \sum_{i=1}^n E[x_i^T / Y, \hat{\theta}^{(p)}] V(\eta_i, \theta_X) \end{aligned} \quad (29)$$

On peut monter facilement que les composantes de l'espérance conditionnelle $E[x_i^T / Y, \hat{\theta}^{(p)}]$ sont données par:

$$E[x_{ik} / Y, \hat{\theta}^{(p)}] = \frac{\pi_k^{(p)} P(y_i / x_i = e_k, \hat{\theta}_Y^{(p)})}{\sum_{l=1}^K \pi_l^{(p)} P(y_i / x_i = e_l, \hat{\theta}_Y^{(p)})}; \quad k=1,2,\dots,K. \quad (30)$$

Où x_{ik} est la $k^{ième}$ composante du x_i . et $\pi_k^{(p)}$ est donnée par:

$$\pi_k^{(p)} = P(x_i = e_k / x_j, j \in \eta_i, \hat{\theta}_X^{(p)}) \quad (31)$$

On doit signaler que l'espérance conditionnelle donnée par(30) représente l'estimation quadratique moyenne de x_i^T [16], ainsi :

$$\hat{x}_i^T = E[x_i^T / Y, \hat{\theta}^{(p)}] \quad (32)$$

\hat{x}_i^T est appelé l'estimée soft de x_i^T .

Si on suppose que $\hat{x}_j^{(p-1)}$ est l'estimateur de x_j à l'itération (p-1) alors l'estimée de $\pi_k^{(p)}$ sera:

$$\hat{\pi}_k^{(p)} = P(x_i = e_k / \hat{x}_j^{(p-1)}, j \in \eta_i, \hat{\theta}_X^{(p)}) \quad (33)$$

L'ETAPE M

Dans cette étape on cherche le vecteur $\hat{\theta} = (\hat{\theta}_y, \hat{\theta}_X)$ des paramètres qui maximise la fonction $Q(\theta / \hat{\theta}^{(p)})$. La solution est obtenue donc, par l'annulation du gradient partiel de l'expression(29)par rapport à $\hat{\theta}$. Ce qui permet d'avoir les équations suivantes :

$$\sum_{i=1}^n E[x_i^T / Y, \hat{\theta}^{(p)}] \frac{\partial}{\partial \theta_Y} [U(y_i / \theta_Y)] = 0 \quad (34)$$

et

$$\sum_{i=1}^n E[x_i^T / Y, \hat{\theta}^{(p)}] \frac{\partial}{\partial \theta_X} [V(\eta_i, \theta_X)] = 0 \quad (35)$$

Pour le modèle utilisé dans ce travail, les paramètres inconnus sont le vecteur $\theta_Y = [\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2]^T$ caractérisant les distributions a priori $P(y_i / x_i = e_k, \hat{\theta}_Y^{(p)})_{k=1,2,\dots,K}$ et $\theta_X = \beta$

caractérisant la distribution de Gibbs. Si on utilise pour les distributions a priori le modèle de l'image amplitude M-Look, c'est à dire:

$$P(y_i / x_i = e_k, \hat{\theta}_Y^{(p)}) = 2 \frac{y_i^{2M-1} e^{\frac{-y_i}{2\sigma_k^2/M}}}{(M-1)! \left(\frac{2\sigma_k^2}{M}\right)^M} \quad k=1, 2, \dots, K \quad (36)$$

Le remplacement dans l'équation (34), permettent d'estimer $\hat{\sigma}_k^{2(p+1)}$ pour $k=1,2,\dots,K$, soit :

$$\hat{\sigma}_k^{2(p+1)} = \frac{\sum_{i=1}^n E[x_{ik} / Y, \hat{\theta}^{(p)}] y_i^2}{M \sum_{i=1}^n E[x_{ik} / Y, \hat{\theta}^{(p)}]} \quad k=1, 2, \dots, K. \quad (37)$$

Du fait que X est un champ de Markov alors la distribution conditionnelle locale par rapport à une classe k est:

$$P(e_k / x_j, j \in \eta_i, \theta_X) = \frac{e^{\beta \sum_{j \in \eta_i} v_{kj}}}{\sum_{k=1}^K e^{\beta \sum_{j \in \eta_i} v_{kj}}} \quad (38)$$

ou v_{ki} est donné par:

$$v_{kj} = \begin{cases} +1 & \text{si } \hat{x}_j^{(p-1)} = e_k \\ -1 & \text{si } \hat{x}_j^{(p-1)} \neq e_k \end{cases} \quad (39)$$

Et η_i représente le vecteur de voisinage du site i donné par :

$$\eta_i = [u_1, u_2, u_3, u_4, v_1, v_2, v_3, v_4]^T \quad (40)$$

où les voisinages u_i et v_i sont montrés à la figure 2

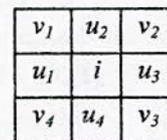


Figure. 2. Le système de voisinage η_i du site i

En substituant (38) dans (35) et en effectuant la dérivation par rapport β , on obtient :

$$\sum_{i=1}^n \left[\sum_{k=1}^K \sum_{j \in \eta_i} E(x_{ik} / Y, \hat{\theta}^{(p)}) v_{kj} - \frac{\sum_{k=1}^K \sum_{j \in \eta_i} v_{kj} e^{\beta \sum_{j \in \eta_i} v_{kj}}}{\sum_{k=1}^K e^{\beta \sum_{j \in \eta_i} v_{kj}}} \right] = 0 \quad (41)$$

Remarquant que cette équation est fortement non linéaire. Ainsi, sa résolution en $\hat{\beta}^{(p+1)}$ nécessite des méthodes numériques comme celle de Newton-Raphson. L'utilisation brute de (41) encombre beaucoup l'algorithme de segmentation non supervisée, à moins de trouver une forme plus adéquate que (41) pour le cas $K=2$ ou $K=3$. En fait, pour le cas $K=2$, on trouve :

$$U^{(p)} + 8c_1^{(p)} \operatorname{tgh}(8\beta) + 6c_2^{(p)} \operatorname{tgh}(6\beta) + 4c_3^{(p)} \operatorname{tgh}(4\beta) + 2c_4^{(p)} \operatorname{tgh}(2\beta) = 0 \quad (42)$$

où tgh est la tangente hyperbolique et $U^{(p)}$ est donné par :

$$U^{(p)} = \sum_{i=1}^n \sum_{k=1}^K \sum_{j \in \eta_i} E(x_{ik} / Y, \hat{\theta}^{(p)}) v_{kj} \quad (43)$$

et $c_i^{(p)}$ sont des constantes données par :

$$\begin{aligned} c_1^{(p)} &= \operatorname{card} \left\{ \eta_i, i = \overline{1:n} : \operatorname{abs} \left(\sum_{j \in \eta_i} v_{kj} \right) = 8, \text{ pour } k \in Q \right\} \\ c_2^{(p)} &= \operatorname{card} \left\{ \eta_i, i = \overline{1:n} : \operatorname{abs} \left(\sum_{j \in \eta_i} v_{kj} \right) = 6, \text{ pour } k \in Q \right\} \\ c_3^{(p)} &= \operatorname{card} \left\{ \eta_i, i = \overline{1:n} : \operatorname{abs} \left(\sum_{j \in \eta_i} v_{kj} \right) = 4, \text{ pour } k \in Q \right\} \\ c_4^{(p)} &= \operatorname{card} \left\{ \eta_i, i = \overline{1:n} : \operatorname{abs} \left(\sum_{j \in \eta_i} v_{kj} \right) = 2, \text{ pour } k \in Q \right\} \end{aligned} \quad (44)$$

et pour le cas $K=3$ on aboutit à l'équation suivante :

$$U^{(p)} - [C_1^{(p)} f_1(\beta) + C_2^{(p)} f_2(\beta) + C_3^{(p)} f_3(\beta)] - [C_4^{(p)} f_4(\beta) + C_5^{(p)} f_5(\beta) + C_6^{(p)} f_6(\beta) + C_7^{(p)} f_7(\beta)] = 0 \quad (45)$$

Avec les constantes $C_i^{(p)}$ et les fonctions f_i données en annexe.

Pour $K > 3$ le nombre des constantes $C_i^{(p)}$ et des fonctions f_i augmente. Par conséquent l'implémentation de l'algorithme devient difficile. Dans ce cas, il est montré que les résultats de segmentation sont insensibles pour certain intervalle de β [12,14]. Ainsi, la valeur adéquate de β peut être choisie à la base de la règle essai et erreur.

SEGMENTATION

Dans cette étape, on cherche à assigner à chaque pixel ou site une étiquette. On remarque que l'étape E de l'algorithme EM, voir (32), fournit à l'itération p , l'estimée soft de x_i : $\hat{x}_i^{(p)} = E(x_i / Y, \hat{\theta}^{(p)}) = [\hat{x}_{i1}^{(p)}, \hat{x}_{i2}^{(p)}, \dots, \hat{x}_{iK}^{(p)}]^T$; le vecteur indicateur de la classe d'appartenance au site i . Les éléments $\hat{x}_{ik}^{(p)}$ sont compris dans l'intervalle $[0, 1]$ et ne sont pas forcément égale à 0 ou 1. Par conséquent, on utilise l'estimée hard qui correspond à l'élément maximum de $\hat{x}_i^{(p)}$. En d'autre termes, la classe affectée au site i correspond à l'élément maximale de $\hat{x}_i^{(p)}$. L'estimée hard est alors donné par :

$$\tilde{x}_i^{(p)} = \arg \max_k \hat{x}_{ik}^{(p)} \quad k = 1, 2, \dots, K.$$

D'après l'expression de $\hat{x}_{ik}^{(p)}$ (30), on déduit que cette dernière opération représente la segmentation au sens de l'estimateur MAP. Plus précisément :

$$\tilde{x}_i^{(p)} = \arg \max_k \hat{\pi}_k^{(p)} P(y_i / x_i = k, \hat{\theta}_Y^{(p)}) \quad k = 1, 2, \dots, K.$$

Ici, on utilise pour $\hat{\pi}_k^{(p)}$ l'expression suivante :

$$\hat{\pi}_k^{(p)} = P(x_i = k / \tilde{x}_j^{(p-1)}, j \in \eta_i, \hat{\theta}_X^{(p)}) \quad (46)$$

où $\tilde{x}_j^{(p-1)}$ est l'estimateur hard obtenu à l'itération $(p-1)$.

De ce fait cette stratégie de segmentation ressemble à l'algorithme ICM vu dans le paragraphe(4) sauf qu'ici chaque segmentation est suivie par l'estimation des paramètres du modèle de l'image.

Finalement, il faut noter que cette approche de segmentation nécessite une segmentation initiale $\tilde{x}_j^{(0)}$ qui peut être trouvée par la méthode du maximum de vraisemblance en supposant les $\hat{\pi}_k^{(0)}$ équiprobables (i.e. $\hat{\pi}_k^{(0)} = 1 / K$).

Le schéma block de cette stratégie de segmentation non supervisée avec l'algorithme EM est illustré à la figure 3.

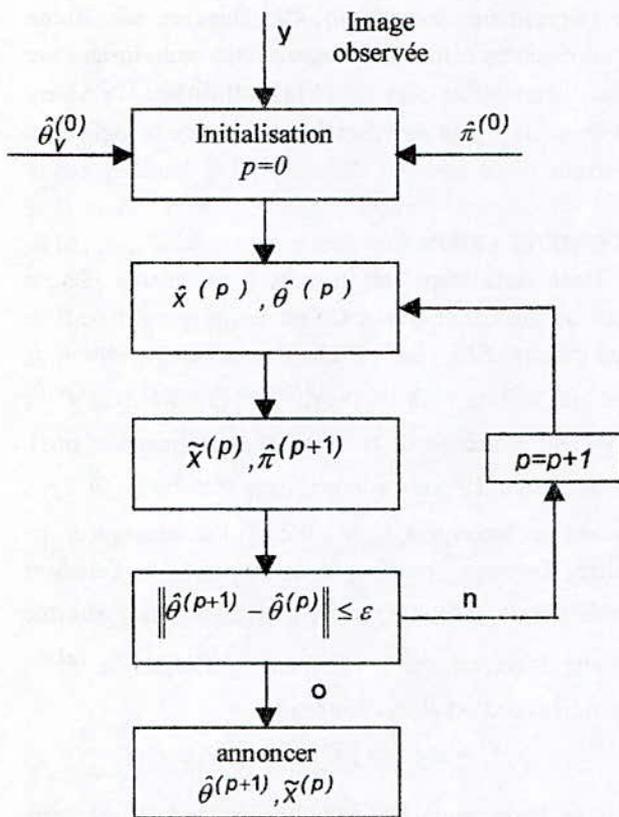


Figure. 3. Algorithme de segmentation non supervisée

8. RESULTATS EXPERIMENTAUX

On présente dans cette section les résultats de segmentation obtenus par application de la stratégie décrite par la figure(3). L'efficacité de la stratégie est testée d'abord en utilisant des images synthétiques dont on connaît leurs segmentations. Ensuite, on applique cette technique sur des images réelles. Pour tous les tests, le vecteur initial des paramètres $\hat{\theta}_Y^{(0)} = [\hat{\sigma}_1^{2(0)}, \hat{\sigma}_2^{2(0)}, \dots, \hat{\sigma}_K^{2(0)}]^T$ est déterminé en estimant la variance σ^2 sur des fenêtres non chevauchées de taille 10×10 . Ces estimations sont ensuite regroupés dans K classes en appliquant l'algorithme K-moyennes[13].

Considérons maintenant deux images intensité de référence : une image 128×128 , composée de deux types de régions (figure. 4.a) tel que le rapport entre l'intensité claire et l'intensité sombre soit de 12dB. Une autre image 256×256 à trois types régions est présentée à la (figure. 5.a). Dans ce cas le rapport entre deux intensités successives est 6dB. Ces images sont générées par

l'échantillonneur de Gibbs[6, 7] en considérant le modèle décrit dans le paragraphe 3. Ensuite, les images amplitudes specklées correspondantes sont obtenues en multipliant l'image de référence par un speckle simulé 1-look (figure. 4.b) et speckle simulé 4-looks (figure. 5.b) respectivement. Les figure 6.a et 6.b montrent les résultats de segmentation de l'approche non supervisée appliquée sur les deux images. Les résultats en pourcentage de l'erreur de classification sont donnés au tableau I. On remarque que l'erreur de classification des pixels est meilleure pour le cas de l'image composée de deux classes relativement à celle de l'image composée de trois classes. Cela est dû au fait que les intensités de référence dans ce dernier cas sont proches (6dB). En effet, il est fort possible que lors de l'introduction de speckle, deux pixels ayant des intensités voisines soient confondus. En plus, la taille des régions y est plus petite comparée à celle de la première image, ce qui conduit à une estimation non consistante.

L'évolution du paramètre (β) estimé à chaque itération des deux images est donnée à la figure 7. On observe que ce paramètre commence à se stabiliser après l'itération 18 pour l'image deux régions 1-look et l'itération 28 pour l'image trois régions 4-look. Le temps de calcul correspondant à ces deux cas est donné au tableau I.

Enfin on présente les résultats de segmentation concernant une image réelle de Laghouat figure 8.a. Sa segmentation en deux classes est montrée à la figure. 8.b. le paramètre β dans ce cas (figure 7.c) atteint le régime permanent après 18 itérations. Le temps de calcul correspondant est illustré au Tableau I.

9. CONCLUSION

Dans ce travail on a présenté le modèle statistique des images SAR ainsi que des notions concernant les champs de Markov. Ensuite, on a développé une stratégie de segmentation non supervisée basée sur l'algorithme EM qui réalisera la segmentation de l'image et l'estimation des paramètres récursivement. Enfin, des résultats expérimentaux concernant des images synthétiques illustrent l'efficacité de cette stratégie. On a testé par ailleurs la robustesse de cette procédure sur une image radar réelle. Généralement, la capacité du traitement et le temps de calcul dépendent de l'image à traiter. La procédure est semi-automatique dont la mesure où elle nécessite la connaissance du nombre des classes.

ANNEXE

Pour déterminer une forme simplifiée de l'équation (41) qui permettra de trouver la valeur de $\hat{\beta}^{(p+1)}$ d'une manière efficace, on doit tester toutes les configurations possibles de η_i . Ce qui, conduira à l'équation (45) avec :

$$f_1(x) = \frac{8(-e^{8x} + 2e^{-8x})}{e^{8x} + 2e^{-8x}}$$

$$f_2(x) = \frac{-6e^{6x} + 6e^{-6x} + 8e^{-8x}}{e^{6x} + e^{-6x} + e^{-8x}}$$

$$f_3(x) = \frac{-4e^{4x} + 4e^{-4x} + 8e^{-8x}}{e^{4x} + e^{-4x} + e^{-8x}}$$

$$f_4(x) = \frac{-4e^{4x} + 12e^{-6x}}{e^{4x} + 2e^{-6x}}$$

$$f_5(x) = \frac{-2e^{2x} + 2e^{-2x} + 8e^{-8x}}{e^{2x} + e^{-2x} + e^{-8x}}$$

$$f_6(x) = \frac{-2e^{2x} + 4e^{-4x} + 6e^{-6x}}{e^{2x} + e^{-4x} + e^{-6x}}$$

$$f_7(x) = \frac{8e^{-4x}}{2e^{-4x} + 1}$$

et

$$C_i^{(p)} = \text{card} \left\{ \eta_i, i = \overline{1:n} : \text{abs} \left(\sum_{j \in \eta_i} v_{kj} \right) = \{n\}_i \text{ pour } k \in Q \right\}$$

où $\{n\}_i$ est donné par :

<i>l</i>	1	2	3	4	5	6	7
$\{n\}_l$	8	6	4 et 8	4 et 6	2 et 8	2,4,6	4 et 0

Image	Erreur de classification en %	Temps de calcul en seconde
à 02 régions	4.0222	13
à 03 régions	5.4565	118
Image réelle		34

Tableau. I. Erreur de classification

BIBLIOGRAPHIE

- [1] FAO/ESA joint study, "Principales of Radar Imagery", RSC Series 46, Food and Agriculture Organisation of United Nations, Rome, 1989.
- [2] C. Bovik and D. C. Munson, "On detecting edges in speckle imagery," *IEEE Trans. Acoustics, Speech, and signal Processing*, vol. 36 pp. 1618-1627, October 1988
- [3] D. T. Kuan and P.Chavel, "Adaptive restoration of images with speckle," *IEEE Trans. Acoustics, Speech, and signal Processing*, vol. 35 pp. 373-383, March 1987.
- [4] N. D. A. Mascarenhas and A. C. Frery, "SAR image filtering with ICM algorithm," in *Proc. IEEE. Int. Conf. Geoscience and Remote Sensing Symposium, IGRASS'94, Pasadena, CA, 1994*, vol. IV, pp. 2185-2187.
- [5] J.-M Cocquerez, S. Phillip. *Analyse d'Images : Filtrage et Segmentation*, Masson, Paris, France. 1996.
- [6] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6 pp. 721-741, November 1989.
- [7] R. C. Dubes and A. K. Jain, "Random field models in images analysis," *Journal of Applied Statistics*, vol. 16, pp. 131-164, 1989.
- [8] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller E. Teller, "Equations of state calculations by fast computing machine," *Journal of chemistries and Physics*, vol. 21 pp. 1087-1091, 1953
- [9] J. Besag, "Towards Bayesian image analysis," *Journal of Applied Statistics*, vol. 16, pp. 395-407, 1989.
- [10] H. Derin and H. Elliot, "Modeling and segmentation of noisy and textured images using Gibbs random fields," *IEEE Trans. Pattern*

Analysis and Machine Intelligence, vol. 9 pp. 39-55, January 1987.

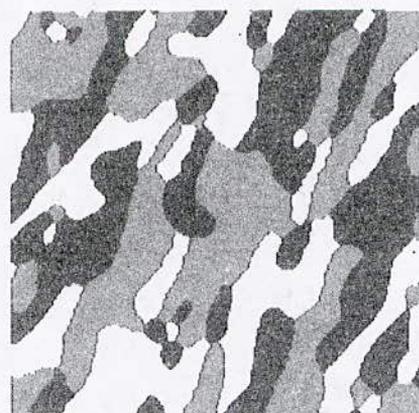
- [11] E. Rignot and R. Chellapa, "Segmentation of polarimetric synthetic aperture radar data," *IEEE Trans. Image Processing*, vol. 1 pp. 281-299, July 1992.
- [12] H. Derin, P. A. Kelly and K. D. Hartt, "Adaptive segmentation of speckled images using a hierarchical random field model," *IEEE Trans. Acoustics, Speech, and signal Processing*, vol. 36 pp. 1628-1641, October 1988
- [13] M.Mignotte, C. Collet, P. Pérez and B. Bouthemy, "Unsupervised Markovian segmentation of sonar images," in *Proc. IEEE. Int. Conf. Acoustics Speech, and Signal Processing, ICASP'97*, New York, 1997, pp. 2781-2784.
- [14] J. Zhang and D. Modestino, "Maximum likelihood parameter estimation for unsupervised stochastic model-based image segmentation," *IEEE Trans. Image processing*, vol. 3, pp. 404 - 420, July 1994
- [15] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 5, pp. 25-39, January 1983.
- [16] A. Papoulis. *Probability, Random Variables and Stochastic Process*, McGraw-Hill. New York, USA. 1984.



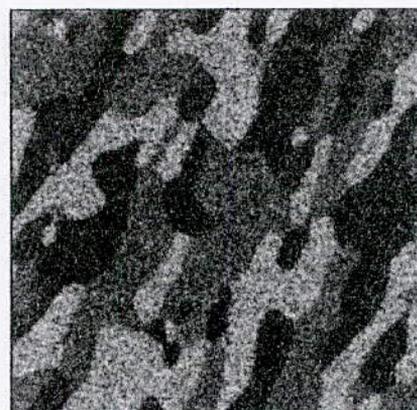
(a)

(b)

Figure. 4. Image synthétique deux régions.

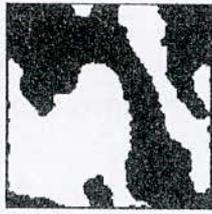


(a)

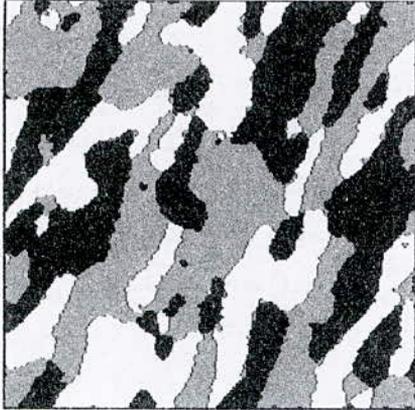


(b)

Figure. 5. Image synthétique trois régions.



(a)

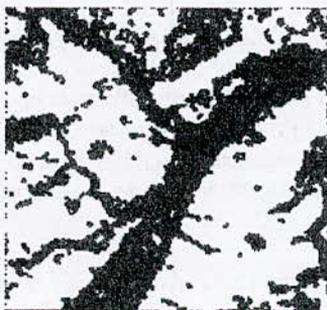


(b)

Figure. 6. Résultats de segmentation des Image (4.b) et (5.b)

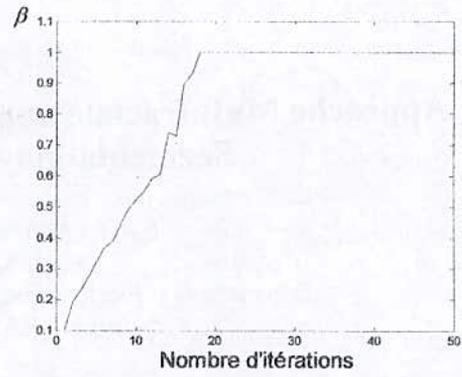


(a)

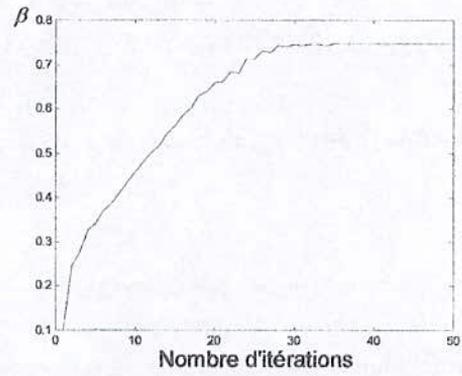


(b)

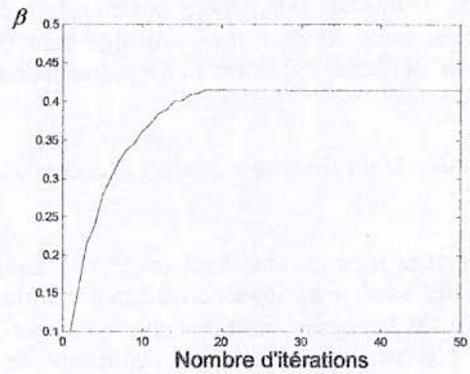
Figure. 8. Image réelle de LAGHOUAT et sa Segmentation en deux classes.



(a) Image synthétique de 02 régions



(b) Image synthétique de 03 régions



(c) Image réelle de LAGHOUAT

Figure 7. La variation de β en fonction de nombre d'itérations

Approche Multifractale appliquée à l'Extraction de Contours et à la Segmentation d'images biomédicales

L. HAMAMI et N. LASSOUAOU

Ecole Nationale Polytechnique

Département d'Electronique. Laboratoire Signal & Communications.

E-mail : L_HAMAMI @HOTMAIL.COM

ملخص: قياسات متعددة فراكتال قادرة على إعطاء معلومات كثيرة على نسيج الصور. قليلة الاستعمال بسبب صعوبة حساب بطريقة ودية ودقيقة

قياسات متعددة فراكتال كالعزم D_q و $\tau(q)$ وخاصة الطيف $f(\alpha)$.

الطيف الأحادي $f(\alpha)$ يمكن تقديمه ككمية تعطي نسبيًا بصفة دقيقة عن درجة نظام و تجانس لقياس فراكتال.

الهدف الأساسي لهذه المقالة هو تقديم مناهج حساب أبعاد متعددة فراكتال لاستخراج نطاق. في هذا الأخير، الخاصة المميزة في خوارزمية استخراج

النطاق هي تقسيم الطيف الأحادي $f(\alpha)$.

خوارزمية حساب المركبات الأحادية في صورة معطاة و قدمنا كذلك تطبيق مباشر لهذه المركبات في تقسيم الصورة. ان الخوارزميات مطبقة على

صور بيوطية .

مفاتيح: حسابات متعددة فراكتال، الطيف الأحادي $f(\alpha)$ ، استخراج النطاق، تقسيم الصورة، صور خليات بيوطية.

Abstract: Multi-fractals measures are susceptible to bring a big information on the texture of an image, but used little by reason of difficulties to calculate of reliable manner moments D_q and $\tau(q)$ of multi-fractal measure, and especially specter $f(\alpha)$. Specter $f(\alpha)$ oddness can be introduced like a quantity that gives a relatively precise characterization of degree of regularity and homogeneity of a fractal measure. This article concerns methods of multi-fractals parameters evaluation and multi-fractal algorithm of edge extraction. In this last, the property constituting the criteria of detection is chosen in order to do a segmentation by the specter of oddness $f(\alpha)$. An algorithm of count of oddness exhibitors in an image and a direct application of these exhibitors in the segmentation are presented. These algorithms are applied on biomedical image.

Keys words: Multi-fractals measures, Specter $f(\alpha)$, edge detection, image segmentation, biomedical cells images.

Résumé : Les mesures multifractales sont susceptibles d'apporter une grande information sur la texture d'une image. Elles sont peu utilisées en raison des difficultés qu'elles ont à calculer, de manière fiable, les moments D_q et $\tau(q)$ de la mesure multifractale et surtout le spectre $f(\alpha)$. Le spectre des singularités $f(\alpha)$ peut être introduit comme une quantité qui donne une caractérisation relativement précise du degré de régularité et d'homogénéité d'une mesure fractale.

Le but essentiel de cet article est de présenter des méthodes d'estimation des paramètres multifractals et donner un algorithme multifractal d'extraction de contours. Dans ce dernier cas, la propriété constituant le critère de détection est choisie de manière à effectuer une segmentation par le spectre de singularité $f(\alpha)$. Un algorithme de calcul des exposants de singularités dans une image est donné ainsi qu'une application directe de ces exposants à la segmentation. Ces algorithmes sont appliqués aux images biomédicales.

Mots clés: Mesures multifractales, spectre $f(\alpha)$, détection de contours, segmentation d'images, images de cellules biomédicales.

1. Introduction :

Le concept de mesure fractal est apparu dans les travaux de Mandelbrot avec l'élaboration des premiers modèles hiérarchiques de répartition du taux de transfert d'énergie dans les flots turbulents 3 D.

Le terme multifractal et la notion de spectre des singularités ont été introduits pour la première fois par Frisch et Parisi. A l'instar de la dimension fractale pour les ensembles irréguliers, le spectre $f(\alpha)$ des singularités peut être introduit comme une quantité qui donne une caractérisation relativement précise du degré de régularité et d'homogénéité d'une mesure fractale.

L'analyse multifractale est adaptée à l'imagerie, elle considère l'intensité lumineuse associée à chaque pixel comme une mesure, et la description multifractale se fait par le calcul des paramètres multifractals, i.e. par le calcul des spectres $f(\alpha)$ et $\tau(q)$.

Dans ce qui suit, nous utiliserons le spectre de singularité pour la détection de contours et nous verrons une application du calcul des exposants de singularité dans la segmentation des images biomédicales (images de cellules).

**2. Le Formalisme Multifractal
Définition du Spectre des Singularités**

Soient : μ une mesure quelconque et $\text{supp}\mu$ son support. (Nous considérons le support d'une mesure comme étant l'ensemble sur lequel la mesure est concentrée.)

On appelle exposant de singularité au point $x_0 \in \text{supp}\mu$, la limite :

$$\alpha(x_0) = \lim_{\varepsilon \rightarrow 0^+} \frac{\ln \mu(B_{x_0}(\varepsilon))}{\ln \varepsilon} \tag{1}$$

Où $B_{x_0}(\varepsilon)$ désigne une boule centrée en x_0 et de taille ε .

Le spectre $f(\alpha)$, des singularités associées à la mesure μ , est la fonction qui à tout α associe la dimension fractale de l'ensemble des points x_0 tel que $\alpha(x_0) = \alpha$:

$$f(\alpha) = d_r(\{x_0 \in \text{Supp}\mu / \alpha(x_0) = \alpha\}) \tag{2}$$

L'exposant de singularité en un point rend compte du degré local de régularité de la mesure considérée. La définition (1) de $\alpha(x_0)$ indique simplement que de manière générale, on pourra écrire :

$$\mu(B_{x_0}(\varepsilon)) \approx C\varepsilon^{\alpha(x_0)} \tag{3}$$

avec $\alpha \geq 0$, quelconque.

Le spectre $f(\alpha)$ des singularités décrit la répartition statique des exposants α sur le support de la mesure. Ainsi, si l'on pave le support de la mesure de "boîtes" de taille ε , alors le nombre de ces boîtes dont la mesure varie comme ε^α pour un α donné, se comporte de la façon suivante :

$$N_\alpha(\varepsilon) \approx \varepsilon^{-f(\alpha)} \tag{4}$$

Le spectre $f(\alpha)$ décrit donc le comportement de l'histogramme $N_\alpha(\varepsilon)$ quand $\varepsilon \rightarrow 0$. Remarquons que le spectre $f(\alpha)$, ainsi défini, pourrait être considéré comme la capacité d'un ensemble; il convient cependant d'être prudent quant à cette interprétation car $N_\alpha(\varepsilon)$ dans l'équation (4) réfère au recouvrement d'un ensemble qui évolue avec ε .

Une mesure sera qualifiée d'homogène, si son spectre des singularités est concentré en un seul point : un seul "type" de singularité permet de caractériser la mesure. Par contre si le support de $f(\alpha)$ est large, la mesure considérée n'est pas homogène ; l'exposant $\alpha(x)$ fluctue d'un point à l'autre du support de μ , on parle dans ce cas de mesures multifractales.

Pour obtenir une description multifractale de μ on calcule d'abord l'ensemble des exposants α possibles puis on évalue la «taille», du sous-ensemble E_α de $[0,1]$, associée à α par calcul de sa dimension de Hausdorff $f(\alpha)$.

Cette description, en terme de $(\alpha, f(\alpha))$, est ainsi locale (via α) et globale (via $f(\alpha)$). Elle est désignée sous le nom de *spectre multifractal* de μ .

Remarque : Une mesure permet d'associer des poids relatifs aux différentes parties d'un ensemble. On peut penser par exemple à des niveaux de gris d'une image, des répartitions de charges, de masses, d'énergie.

Définition du Spectre $\tau(q)$

Considérons une mesure μ et un recouvrement de son support par des boules $B_i(\varepsilon)$ de rayon ε . Soit $N(\varepsilon)$, le nombre de boules nécessaires. On considère, pour tout $q \in \mathbb{R}$, la fonction de partition $Z [2]$:

$$Z(q, \varepsilon) = \sum_{i=1}^{N(\varepsilon)} \mu_i^q(\varepsilon) \tag{5}$$

Où l'on a posé $\mu_i(\varepsilon) = \mu(B_i(\varepsilon))$. On définit alors le spectre $\tau(q)$ à partir du comportement, en loi de puissance, de $Z(q, \varepsilon)$ quand $\varepsilon \rightarrow 0^+$:

$$Z(q, \varepsilon) \approx \varepsilon^{\tau(q)} \quad (6)$$

et le spectre des dimensions fractales généralisées comme le rapport :

$$D_q = \frac{\tau(q)}{q-1} \quad (7)$$

$\tau(q)$ caractérise le comportement global de la mesure quand la taille des intervalles tend vers zéro, et est relié à une notion de dimension généralisée. En effet: D_0 est la dimension fractale du support μ , $D_1 = \lim_{q \rightarrow 1} D_q$ est la dimension d'information, D_2 la dimension de corrélation.

Contrairement aux exposants de singularités $\alpha(x)$, les exposants $\tau(q)$, ou les dimensions fractales généralisées, décrivent le comportement de quantités globales, $Z(q, \varepsilon)$, qui peuvent être considérées comme des valeurs moyennes sur le support de la mesure. De ce fait, on peut s'attendre à ce que le spectre $\tau(q)$ se prête mieux à une détermination numérique stable que le spectre $f(\alpha)$.

L'Approche Multifractale

Un des résultats fondamentaux qui justifie le calcul du spectre $\tau(q)$ est que celui-ci peut être relié au spectre $f(\alpha)$ des singularités. Si l'on considère α comme une variable continue de densité (à l'échelle ε) $\rho(\alpha)\varepsilon^{-f(\alpha)}$ et que l'on substitue dans l'équation (5) le comportement $\mu_r^q(\varepsilon) \sim \varepsilon^{q\alpha}$, alors $Z(q, \varepsilon)$ se réécrit sous la forme intégrale :

$$Z(q, \varepsilon) \cong \int \rho(\alpha) \varepsilon^{-f(\alpha)} \varepsilon^{q\alpha} d\alpha \quad (8)$$

A la limite $\varepsilon \rightarrow 0^+$, l'intégrale ci-dessus est dominée par le terme $\varepsilon^{\min(q\alpha - f(\alpha))}$. En identifiant ce comportement et celui qui définit $\tau(q)$ (Eq.6), on en déduit :

$$\tau(q) = \min_{\alpha} (q\alpha - f(\alpha)) \quad (9)$$

$\tau(q)$ est donc relié au spectre $f(\alpha)$ des singularités par une transformation de Legendre. Remarquons que si $f(\alpha)$ est continûment dérivable, cette relation peut se réécrire sous la forme :

$$\begin{cases} q = \frac{df}{d\alpha} \\ \tau(q) = q\alpha - f(\alpha) \end{cases} \quad (10)$$

Nous appelons *formalisme multifractal* l'approche qui consiste à considérer le spectre

$f(\alpha)$ des singularités comme la transformée de Legendre (inverse) du spectre $\tau(q)$:

$$f(\alpha) = \min_q (q\alpha - \tau(q)) \quad (11)$$

que l'on peut réécrire lorsque $\tau(q)$ est continûment dérivable, sous la forme :

$$\begin{cases} \alpha = \frac{d\tau}{dq} \\ f(\alpha) = q\alpha - \tau(q) \end{cases} \quad (12)$$

Remarquons que si $f(\alpha)$ est C^2 , la condition (9) implique que $d^2f/d\alpha^2 < 0$ quelque soit α ; f convexe.

3. Méthodes de calcul des Paramètres Multifractals :

Calcul du Spectre Multifractal $f(\alpha)$

Une mesure pour un ensemble Y est définie par :

$$m(Y) = \int_Y N(dt) dt \quad (13)$$

où dt représente la mesure de Lebesgue, $N(dt)$ le nombre de points sur dt .

On a l'hypothèse suivante :

Localement, la mesure n'est pas proportionnelle à la mesure de Lebesgue dt de l'objet, mais est équivalente à dt^α quand $dt \rightarrow 0$. Si α existe en un point (x,y) d'une image alors nous avons :

$$m(I(x,y)) \approx (1/r)^{-\alpha(x,y)} \text{ quand } r \rightarrow 0, \quad (14)$$

Sachant que :

- $m(x,y)$: est la mesure associée au point (x,y) de niveau de gris $I(x,y)$.
- r : est la largeur de boîte entourant le point (x,y) .
- α : est l'exposant de Hölder au point (x,y) .

Le calcul du spectre se fait en deux étapes :

- Pour chaque point, on calcule $\alpha(x,y)$ dans un voisinage de taille r autour de ce point. On obtient ainsi un nuage de points $\left(-r \ln 2, \ln \sum_{j \in V_r(i,j)} I(i,j) \right)$, où $V_r(i,j)$ est le voisinage de taille r autour de (i, j) , avec r variant de 1 à r_{max} . Il est important de garder r_{max} petit (pas plus de 7) afin de conserver l'aspect local de α . L'estimation aux moindres

carrés de la pente de nuage de points donne l'estimation de $\alpha(i,j)$.

- On discrétise l'ensemble des valeurs de α en intervalles $I_\alpha = [\alpha - \varepsilon, \alpha + \varepsilon]$. On groupe tous les points ayant un exposant de Holder α le plus proche, i.e. appartenant à un même intervalle I_α , donnant ainsi des ensembles de points E_α , qui ont un α le plus proche.
 $E_\alpha = \{(i,j) \in \text{Image}, \alpha \in I_\alpha\}$.

On obtient le spectre $f(\alpha)$ en calculant la dimension fractale de E_α , par la méthode de boîte ou de Voss [3].

Le spectre $(\alpha, f(\alpha))$ contient un grand nombre d'informations. Les α donnent une mesure locale des singularités, $f(\alpha)$ mesure le comportement fractal des points de même α . Il a généralement la forme d'une cloche.

Calcul du Spectre $\tau(q)$ et les Dimensions Fractales Généralisées

Sachant que :

$$D_q = \lim_{l \rightarrow 0} \frac{1}{q-1} \frac{\log \sum p_i^q(l)}{\log l} \quad (15)$$

où $p_i(l)$ est la mesure du sous-ensemble d'indice i et de rayon l .

Pour $q=1$: la dimension d'information est de

$$D_1 = - \lim_{l \rightarrow 0} \frac{\sum p_i \log p_i}{\log l}$$

D_q et $\tau(q)$ sont reliés par la simple relation : $\tau(q) = (q-1)D_q$.

Alors l'algorithme de calcul se fait comme suit :

En recouvrant l'image de taille L par un quadrillage de boîtes distinctes de taille r , on obtient une distribution de mesures normalisée p_i définies par $p_i = m_i/m$, où m_i est la mesure de la $i^{\text{ème}}$ boîte et m la mesure totale de l'image. Il faut estimer $\sum p_i^q$. L'estimation de D_q et $\tau(q)$ suppose

de calculer la limite quand $r \rightarrow 0$ de $\frac{\log \sum p_i^q}{\log r / L}$ de la

mesure, nous estimons cette limite par la pente de régression linéaire du nuage de points coordonnées $(\log \sum p_i^q, \log r / L)$.

Application à l'Imagerie

En utilisant les algorithmes décrits précédemment, le calcul des paramètres multifractals sur l'image 00069031.bmp donne les courbes de la figure 1.

4. Algorithme Multifractal pour la Détection de Contours :

Cette approche procède par le calcul direct du spectre des singularités $f(\alpha)$. Cette approche est particulièrement délicate, dans la mesure où elle nécessite un bon algorithme d'estimation de la dimension fractale d'un ensemble, et ce même sur une image comportant peu de points. L'algorithme de détection est le suivant :

Algorithme

- Calculer en chaque point de l'image l'exposant de singularité α (voir plus haut), et ceci par l'estimation de la limite du rapport $\frac{\log p_i}{\log \frac{1}{r}}$

une régression linéaire. Ce qui nous permet d'obtenir une image de l'exposant. (p_i est la mesure d'un sous-ensemble). α est donc calculé comme étant l'estimation de la pente du nuage de points $\{\log(\mu(V(r)), \log(1/r))\}$, avec $\mu(V(r))$ la mesure du voisinage de taille r autour du pixel concerné.

- Diviser l'intervalle $[\alpha_{\min}, \alpha_{\max}]$ en M parties.
- Extraire séparément les pixels correspondant à chaque intervalle.
- Affecter à chaque pixel appartenant à l'intervalle, la valeur 255, 0 pour les autres. Nous obtenons ainsi M images binaires.
- Calculer la dimension fractale de chaque sous image par la méthode de boîtes. Affecter cette valeur à chaque pixel significatif comme étant $f(\alpha(x,y))$. Nous obtenons alors une image de $f(\alpha)$.
- Opérer une segmentation selon les valeurs de $f(\alpha)$.

La dernière étape s'effectue par l'extraction des pixels dont le $f(\alpha)$ est supérieur à 1 et proche de 2, c'est à dire : $1.4 \leq f(\alpha) \leq 1.8$. On peut extraire les pixels dont le $f(\alpha)$ est proche de 1, ce qui donne une image de contours réguliers. On peut aussi extraire tous les points dont le $f(\alpha)$ est proche de 2 ($1.8 \leq f(\alpha) \leq 2$), ces points correspondent aux régions homogènes. Mais dans notre cas, on s'intéresse aux points dont $1.4 \leq f(\alpha) \leq 1.8$.

Remarques

- Pour obtenir une meilleure localisation possible, nous utilisons un voisinage pas plus grand que 7 afin de conserver l'aspect local de α . Nous utilisons un voisinage $V(3)$ c'est à dire 3×3 . Le calcul de α portera seulement sur deux valeurs de r à savoir 1 et 3 ce qui nous évite le calcul de la régression ; il en découle un gain

en temps de calcul considérable. α est obtenue par la relation suivante :

$$\alpha = \frac{\log \mu(V(3)) - \log \mu(V(1))}{\log 3 - \log 1}$$

- $f(\alpha)$ est obtenue par la méthode de boîtes au lieu de celles de Voss [3]. Car la méthode de Voss est trop gourmande en temps de calcul.
- Le choix de la mesure μ : Un choix naturel consisterait à définir cette mesure comme la somme des intensités des pixels de la région mesurée. Cette définition respecte l'axiome de la théorie des mesures. Mais on peut généraliser la notion d'exposant de Holder en introduisant des définitions pour μ qui ne respectent pas nécessairement cet axiome. Des exemples de ces mesures pour une région Ω peuvent être donnés par les définitions suivantes :

$$\mu_{sum}(\Omega) = \sum_{i \in \Omega} p(i),$$

$$\mu_{max}(\Omega) = \max_{i \in \Omega} p(i),$$

$$\mu_{min}(\Omega) = \min_{i \in \Omega} p(i),$$

où Ω^* est l'ensemble de Ω où l'intensité est non nulle, $p(i)$ est l'intensité au point i .

Les exposants calculés avec ces mesures donneront des informations différentes quant à la singularité rencontrée.

La mesure \min n'est pas utilisée, puisqu'elle donne des résultats médiocres, on peut aisément le voir après l'avoir testé sur plusieurs images, un exemple est donné par la figure 2, qui est le résultat obtenu avec l'images 2des1.bmp. Dans la suite de ce travail cette mesure ne sera pas utilisée.

Les résultats d'application de cet algorithme multifractal sur plusieurs images, par utilisation des mesures sum et max , sont donnés en figure 3.

La première colonne présente l'image originale; la deuxième, donne la détection de contours par l'approche multifractale avec la mesure sum et la troisième, la détection de contours avec le même algorithme en utilisant la mesure max .

D'après les images précédentes, on constate que l'extraction des contours par utilisation du spectre des singularités (α , $f(\alpha)$) est bien possible et ceci, même sur des images très complexes et manufacturées, telles que les images biomédicales. On constate que la mesure max donne de meilleurs résultats que la mesure sum . De même, on voit bien l'intérêt de cet algorithme qui

s'adapte à tous types d'images et ceci sans paramètres à adapter ni de seuil, donc tous les problèmes de seuillage ou de binarisation que l'on trouve dans la plupart des méthodes de détection de contours tel que les approches frontières [7, 8], ne figurent pas dans cet algorithme multifractal. Il est à noter que cet algorithme n'est pas coûteux en temps de calcul.

5. Algorithme Multifractal pour la Segmentation

Dans cette méthode, nous adaptons le paramètre multifractal α à la segmentation des images. Pour cela, nous étudions le comportement des exposants de singularité locaux, c'est à dire calculés en chaque point de l'image, nous obtenons alors une image des exposants de singularité.

L'algorithme de segmentation est le suivant :

- calculez les exposants de singularité en chaque point de l'image, et trouvez la singularité minimale α_{min} et la singularité maximale α_{max} .
- divisez le spectre de singularité en plusieurs parties égales, c'est à dire divisez l'intervalle $[\alpha_{min}, \alpha_{max}]$ en N sous intervalles.
- pour chaque sous intervalle, trouvez l'ensemble des points de l'image ayant l'exposant de singularité appartenant à l'intervalle, ce qui nous permet d'avoir N sous images, chacune correspondant à un intervalle iso-alpha. Sachant que pour un intervalle $[\alpha_1, \alpha_2]$ donné, les pixels retenus seront ceux dont le niveau de gris sera compris dans l'intervalle $\left[p\left(\frac{1}{L}\right)^{\alpha_2}, p\left(\frac{1}{L}\right)^{\alpha_1} \right]$.
- pour chaque image d'un intervalle iso-alpha, les points en noir sont ceux dont la valeur α est dans l'intervalle souhaité.

Dans cet algorithme, le calcul de l'exposant de singularité s'effectue comme suit, selon la formule déjà vu :

$$m(I(x,y)) \approx (1/r)^{-\alpha(x,y)}$$

quand $r \rightarrow 0$,

avec :

- $m(x,y)$: est la mesure associée au point (x,y) de niveau de gris $I(x,y)$.
- r : est la largeur de boîte entourant le point (x,y) .
- α : est l'exposant de Hölder au point (x,y) .

Ce qui nous permet d'écrire l'équation suivante :

$$\frac{p(x,y)}{p} = \left(\frac{r}{L}\right)^{\alpha(x,y)}$$

soit à l'échelle du pixel

$$p(x,y) = p \left(\frac{1}{L}\right)^{\alpha}$$

avec p est la somme des niveaux de gris de toute l'image et L la taille de l'image.

L'application de cet algorithme sur plusieurs images biomédicales, permet d'avoir les résultats des figures 4 à 6.

La figure 4d présente les contours de l'image 4a. Les figures 4b, 5b et 6b, présentent des séparations des images entre les contours des figures 4a, 5a et 6a respectivement. Les figures 4c, 5c et 6c présentent des images de cytoplasme des figures 4a, 5a et 6a respectivement. Les figures 5d et 6d présentent des images noyaux des figures 5a et 6a respectivement.

6. Conclusion

Le problème central de la théorie multifractale est de relier les deux descriptions $(\alpha, f(\alpha))$ et $(q, \tau(q))$. Ceci a d'importantes applications. En effet, $\tau(q)$ est généralement beaucoup plus facile à calculer sur des données expérimentales que $(\alpha, f(\alpha))$: $\tau(q)$ est obtenu par un moyennage sur un grand nombre d'intervalles suivi par un passage à la limite. α est plus sensible au bruit, puisqu'il est calculé indépendamment en chaque point. En ce qui concerne $f(\alpha)$, son calcul implique l'évaluation d'une dimension de Hausdorff, tâche généralement difficile. Sous des hypothèses assez générales, il a été prouvé que :

$$f(\alpha) \leq \inf_q \{ q\alpha - \tau(q) \}$$

Pour les mesures multinomiales nous avons une égalité et la dimension de Hausdorff est obtenue par la transformée de Legendre de $\tau(q)$; $D(\alpha)$ a une forme de cloche. Cette forme est aussi observée pour un grand nombre de phénomènes naturels. Cependant, elle n'est en aucun cas une propriété générale des spectres multifractals, puisqu'il a été prouvé que toute fonction réglée peut être le spectre d'une mesure multifractale.

Les paramètres multifractals ont été déjà utilisés pour modéliser des phénomènes physiques chaotiques. Par exemple, Meneveau les utilise pour des turbulences, Ait-Kheddache et Rajala ont tenté de classifier des textures à l'aide des dimensions généralisées.

Dans cet article, nous avons vu l'utilisation du spectre de singularités dans la segmentation par détection de contours et nous avons donné des résultats de segmentations par région en utilisant toujours le spectre de singularités. Nous remarquons que certaines valeurs de α correspondent à une présence de contours vifs. Et d'autres valeurs intermédiaires séparent l'image entre les contours.

Les résultats confirment donc l'hypothèse que le paramètre α a un rôle important à jouer en segmentation des images. D'après ces images, nous remarquons qu'il y a des images représentant des noyaux et d'autres des cytoplasmes. Nous obtenons donc une segmentation de cellules en noyaux et cytoplasmes, ce qui est très important pour la reconnaissance et la détection des anomalies car les tailles du noyau et du cytoplasme présentent un critère de présence ou non du cancer.

Références

- [1] J.P. Berroir. « Application des mesures multifractales à l'analyse et la segmentation d'images médicales », Rapport de recherche INRIA, 1991.
- [2] A. Arneodo, F. Argoul, E. Bacry, J. Elezgaray, J-F. Muzy. « Ondelettes, multifractales et turbulences de l'ADN aux croissances cristallines », Arts et Sciences, 1995.
- [3] J.M.K. Keller et S. Chen. « Texture Description and Segmentation through Fractal Geometry », Computer Vision, Graphic, and Image Processing, Vol 45, N°2, pp : 150-166, 1989.
- [4] Yuxin Liu, Yandan Li. « Image Feature Extraction and Segmentation using Fractal Dimension », IEEE Information Communications and Signal Processing, pp :975-979, Sep 1997.
- [5] J.M. Blackledge and E. Fowler. « Fractal Dimension Segmentation of Synthetic Aperture Radar Images », Image Processing and its Applications, pp : 445-449, 1992.
- [6] B. Mandelbrot. « Fractals Form, Chance, And Dimension », W.H. Freeman And Company, 1977.
- [7] Ph. Bolon, J-M. Chassery, J-P. Cocquerez, D. Demigny, C. Graffigne, A. Montanvert, S. Philipp, R. Zéboudj, J. Zérubia. « Analyse d'images : filtrage et segmentation », Masson, 1995.
- [8] R.C. Gonzalez. « Digital Image Processing », 2^{ème} Edition, Addison-Wesley, 1987.
- [9] B. Mandelbrot. « Les objets fractals », Flammarion, 1995.

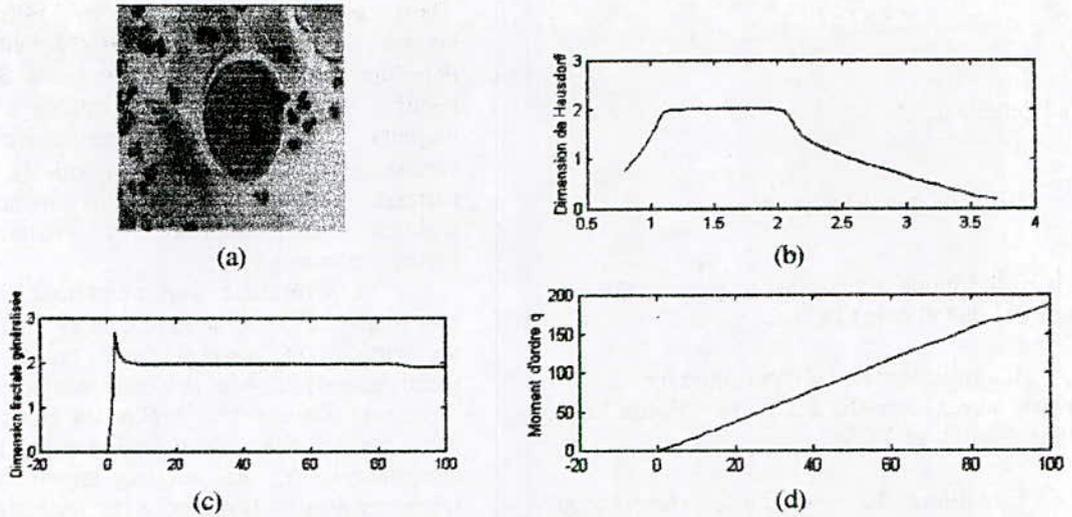


Figure 1: (a) Image 00069031.bmp, (b) Le spectre multifractal ($\alpha, f(\alpha)$), le calcul donne $\alpha \in [0.884, 3.662]$, (c) La courbe des dimensions fractales généralisées (q, D_q), (d) Les moments d'ordre q .

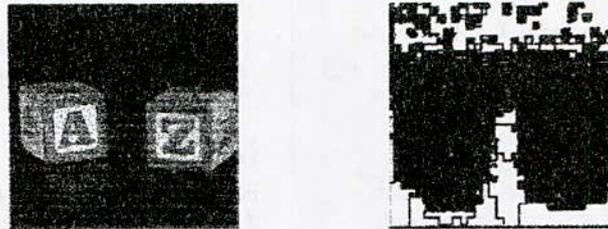


Figure 2: Détection de contour par utilisation de la mesure min.

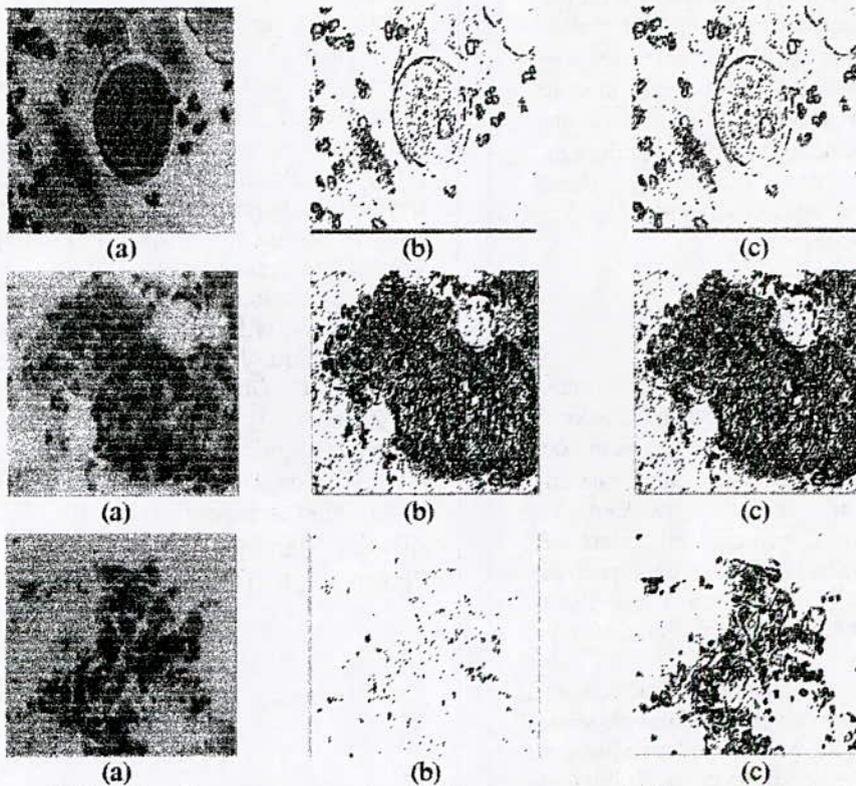


Figure 3 Détection de contours pour les différentes mesures, (a) image originale, (b) avec une mesure sum, (c) avec une mesure max.

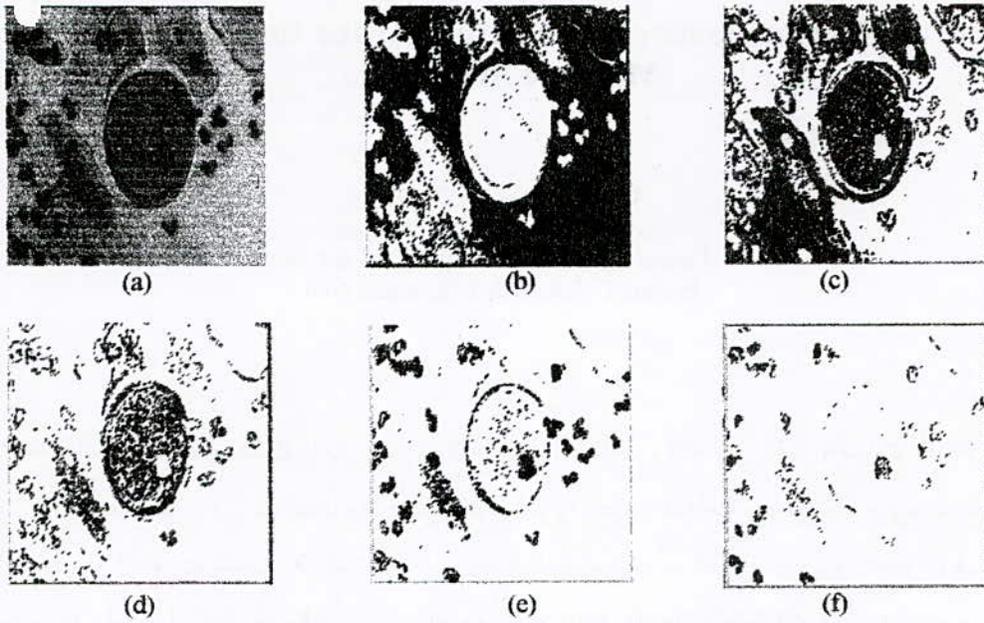


Figure 4: Les images iso- α , avec $\alpha_{\min}=1.899$ et $\alpha_{\max}=2.699$, (a) image originale, (b) $\alpha \in [1.899, 2]$, (c) $\alpha \in [2, 2.1]$, (d) $\alpha \in [2.1, 2.2]$, (e) $\alpha \in [2.2, 2.4]$ et (f) $\alpha \in [2.4, 2.699]$.

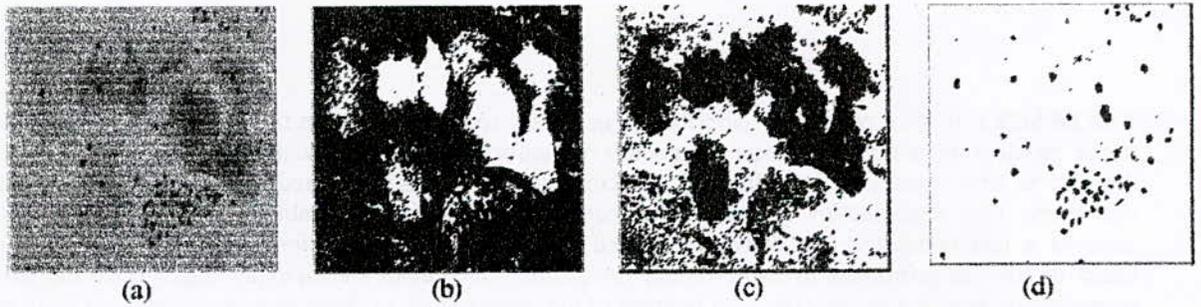


Figure 5: Les images iso- α , avec $\alpha_{\min}=1.951$ et $\alpha_{\max}=2.807$, (a) image originale, (b) $\alpha \in [1.951, 2]$, (c) $\alpha \in [2, 2.1]$, (d) $\alpha \in [2.1, 2.4]$.

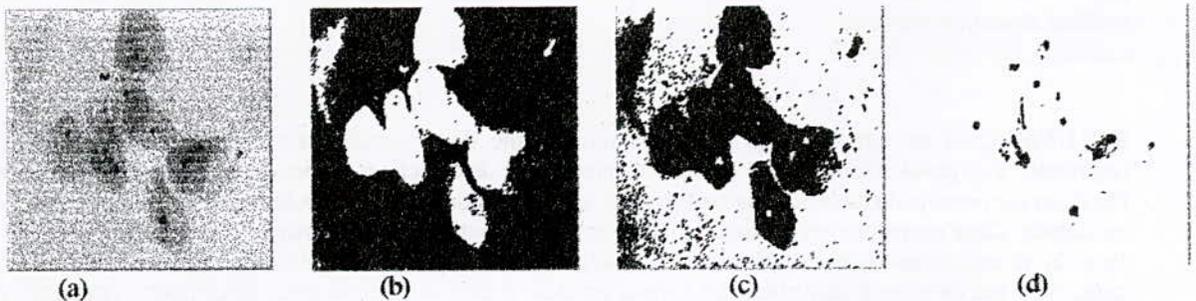


Figure 6: Les images iso- α , avec $\alpha_{\min}=1.971$ et $\alpha_{\max}=2.420$, (a) image originale, (b) $\alpha \in [1.971, 2]$, (c) $\alpha \in [2, 2.1]$, (d) $\alpha \in [2.1, 2.420]$.

Système de Reconnaissance de l'écriture Arabe Imprimée Utilisant une Méthode Statistique

L. HAMAMI MITICHE *
Ecole Nationale Polytechnique.
Département d'électronique, laboratoire Signal et Communications
E-mail: L_HAMAMI@hotmail.com

ملخص: العمل الذي قمنا به يهدف إلى إنجاز نظام تعرف ضوئي للحروف، قادر على القيام بعمليات المعالجة والتعرف للحروف العربية بعد تقسيمها والتي هي صادرة من نص ممسوح ضوئيا بالنسبة للمعالجة نستعمل مرشح وسيط. فيما يخص استخراج الحروف من الكلمات نستعمل طريقة تعتمد على بنية الكتابة العربية، أما الطريقة المتبعة للتعرف فهي طريقة إحصائية. تعتمد على استخراج خاصيات التي تتمثل في العزوم السبع الثابتة. المرحلة الأخيرة للتعرف وهي التقرير، تستعمل طريقة البعد الإقليدي.

مفاتيح: تعرف ضوئي للحروف، كتابة عربية، حروف مطبوعة، تقسيم النص، طريقة إحصائية، العزوم الثابتة.

ABSTRACT : In this paper, we describe an automatic recognition system for printed Arabic characters. The major problem in automatic cursive characters recognition being the segmentation in constituted components, the text is first segmented into lines of characters, then the line is segmented in connected parts and finally in characters. This segmentation is based upon contextual properties of Arabic writing. The segmentation part receives a text paragraph, and delivers separated characters for the recognition part. The method under study is based on the observation of line and column histograms. All methods known to day (e.g.: [1],[2]) for Arab text segmentation have not solved. The total problem of the segmentation of characters, thence some of them perform a sub-segmentation for some characters and others perform an over segmentation. In contrast to all these previous proposed approaches, our method succeed in performing a 100% segmentation of Arab characters. A statistical method of recognition is used based on seven invariant moments. The last procedure consists of a decision stage that utilizes the Euclidean distance principle.

Key Words: Optical Character Recognition, Arabic writing, Printed characters, text segmentation, statistical method, invariant moments.

RESUME : Dans cet article, nous décrivons un système de reconnaissance automatique de l'écriture arabe imprimée. Le problème majeur de la lecture automatique de caractères arabes cursifs étant la segmentation d'un tracé en éléments constituants, nous segmentons le texte d'abord en lignes, puis en parties connexes et ensuite en caractères. Cette segmentation est basée sur les propriétés contextuelles propres à l'écriture arabe.

Pour la reconnaissance, nous utilisons une méthode statistique qui consiste à décrire chaque forme, qu'est le caractère, par un vecteur d'attributs. Ce dernier est dans notre cas un descripteur de contours représenté par une série de sept moments. Avant la reconnaissance proprement dite, il est au préalable nécessaire d'extraire les caractéristiques de tous les caractères : Ce qui constitue la phase d'apprentissage. La dernière opération qui consiste à identifier le caractère est l'étape de décision, elle utilise le principe de la distance euclidienne.

Mots clés : Reconnaissance de caractères optiques, écriture arabe, caractères imprimés, segmentation de texte, méthode statistique, moments invariants.

* Chargée de cours à L'ENP

1. Introduction

Le but de la reconnaissance de l'écriture est de transformer un document de texte écrit sur un support papier, en une représentation compréhensible par un ordinateur et facilement reproductible par un traitement de texte ; c'est une interface permettant de faciliter la communication homme-machine. C'est aussi un problème d'identification d'une forme donnée qu'on affecte à une classe obtenue par apprentissage [1]. Cette technique n'est pas facile vu que les mots possèdent une multitude de représentations dues aux nombreuses polices de caractères existant avec de nombreux styles (gras, italique, ...).

L'intérêt suscité par la reconnaissance des caractères est justifié par ses nombreuses applications dont l'archivage de documents, la saisie automatisée en entreprise, la lecture automatique de chèques, ... [13]

Des recherches ont été entreprises de part le monde, depuis plusieurs décennies, pour reconnaître l'écriture imprimée, dactylographiée et même manuscrite. Des systèmes très performants arrivent à des taux de reconnaissance très élevés pour l'écriture latine, chinoise, indienne, etc ; quant au manuscrit, il pose encore quelques problèmes vu la cursivité des lettres et leur variabilité d'une personne à une autre.

L'écriture arabe, objet de notre étude, a bénéficié de quelques recherches. C'est en 1980 que A.Amin a (parmi les premiers chercheurs comme : ElSheikh, Abuhaiba, HajHassan...) débuté ses travaux sur les OCR arabes et réalisa le système IRAC (Interactive Recognition of Arabic Characters) [6]. Depuis, plusieurs recherches ont succédé sur l'écriture arabe en France, en Syrie, en Egypte, en Algérie, en Tunisie... utilisant différentes approches (statistiques, structurelles et neurales).

Le problème de segmentation de l'écriture arabe a été résolu partiellement car la majorité des chercheurs s'y intéressant ont rencontré un problème de sur-segmentation de quelques caractères et que nous avons résolu dans notre système ; par contre le problème de sous-segmentation de quelques caractères ligatures, a été contourné en les considérant comme un seul et en lui affectant, après la reconnaissance, les codes ASCII des caractères ligaturés dans l'ordre de leur écriture.

Il existe de nos jours, des produits industrialisés concernant l'écriture arabe mais leurs performances doivent être améliorées.

Dans ce domaine, si on se penche sur l'écriture cursive (qu'est l'Arabe ou le latin manuscrit), on rencontre certaines difficultés en arabe que l'on ne trouve pas en latin et qui sont :

- a) Il existe différents types de caractères très utilisés.
- b) Les caractères qui s'attachent en se chevauchant sont très difficiles à séparer.
- c) Les caractères arabes peuvent être reliés selon trois types ou non reliés de manière que chaque caractère puisse avoir quatre formes différentes dans le même jeu.

d) Certains caractères ont une ponctuation dont la taille est très faible par rapport au reste, ce qui la rend difficile à reconnaître.

e) Les caractères arabes n'ont pas la même hauteur ce qui rend difficile la détection du bruit.

Nous présentons dans cet article un système de reconnaissance de texte arabe imprimé utilisant une méthode statistique qui consiste à représenter chaque caractère par une suite de sept moments.

2. Constitution d'un système OCR

Les différentes phases d'un processus de reconnaissance peuvent être représentées selon le schéma de la Fig.1 où chaque phase représente un module avec une entrée et une sortie.

L'entrée de notre système est un texte, résultat d'une phase d'analyse et de reconnaissance de documents. Les problèmes d'inclinaison ont donc été corrigés dans cette phase.

2.1. Acquisition du texte :

Dans notre cas l'acquisition du document se fait à l'aide d'un scanner. L'image texte est obtenue sous forme de bit-map. Le texte que nous traitons ne doit pas comporter de voyelles.

2.2. Filtrage :

Cette étape consiste à éliminer les bruits dans l'image binaire obtenue à la sortie du scanner. Plusieurs types de filtres ont été implémentés (Filtre moyen, filtre de Gauss, filtre de Deriche, Filtre de Gabor et le filtre médian) ; nous avons opté pour le filtre médian non-linéaire qui s'adapte bien à notre application vu que le texte présente dans la majorité des cas un type de bruit impulsif. Le filtre médian considère chaque point de l'image et lui donne la valeur médiane de ses huit voisins.

Pour l'implémentation de ce filtre, nous procédons de la façon suivante (voir Fig.2) :

- Nous classons les pixels voisins du pixel central par valeurs croissantes.
- Nous affectons au pixel courant (central) la valeur médiane des pixels classés.

La Fig.3 donne un exemple d'utilisation de ce filtre.

Remarque :

Cette étape de filtrage peut être optionnelle et prend en considération l'état du document et du système d'acquisition.

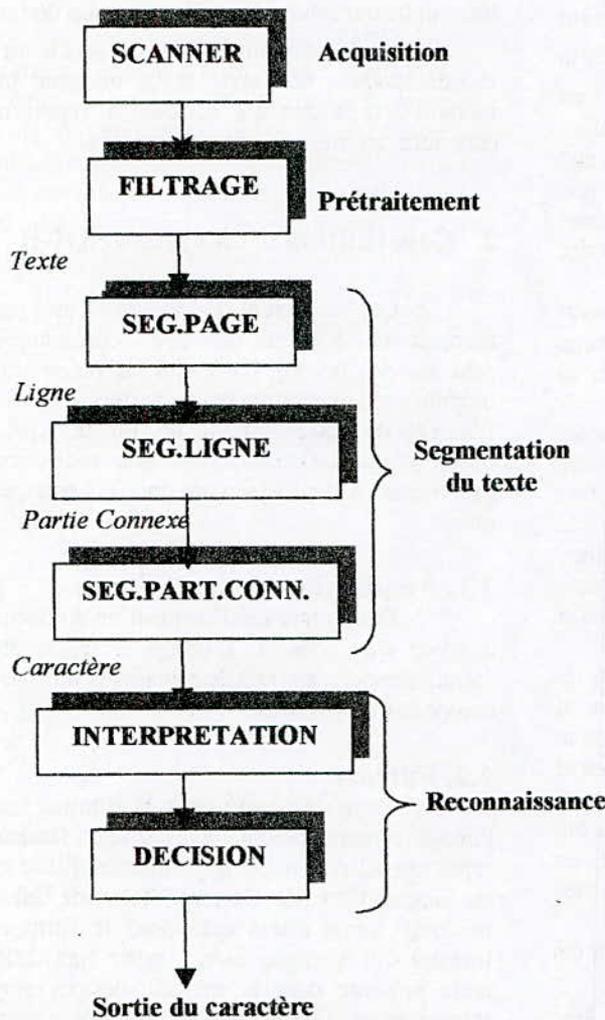


Fig.1 : Block diagramme du système OCR

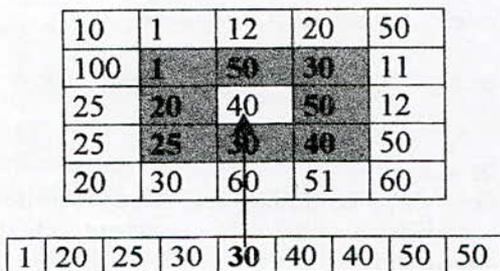


Fig.2 : Principe du filtre médian.

ربنا إفتح بيننا وبين قومنا بالحق
وأنت خير الفاتحين

(a)

ربنا إفتح بيننا وبين قومنا بالحق
وأنت خير الفاتحين

(b)

Fig.3: (a) image bruitée, (b) image filtrée.

2.3.Segmentation d'un texte :

Après l'étape d'acquisition, l'image obtenue est une matrice binaire contenant des zéros (0) et des uns (1). Le module de segmentation reçoit en entrée un paragraphe de texte et fournit en sortie des caractères séparés destinés au module de reconnaissance. La méthode étudiée est basée sur l'observation des histogrammes des lignes et des colonnes et tient compte des caractéristiques propres à l'écriture arabe.

Les projections horizontale et verticale sont définies comme suit :

$$h(i) = \sum_j g(i, j) \text{ projection horizontale}$$

$$v(j) = \sum_i g(i, j) \text{ projection verticale}$$

Où $g(i, j)$ vaut 0 ou 1, i désignant les lignes et j les colonnes.

La segmentation se fait séquentiellement en trois étapes : [5, 6]

- a) Localisation des lignes de texte.
- b) Séparation des lignes en parties connexes.
- c) Séparation des parties connexes en caractères.

La première étape constitue la segmentation horizontale et les deuxième et troisième étapes les segmentations verticales.

2.3.1. Segmentation horizontale :

Elle consiste à détecter les lignes de texte (voir Fig.4). Son principe est le suivant :

- Procédant à un balayage horizontal, on détermine le début de la ligne de texte qui correspond à la première ligne de la matrice binaire qui possède au minimum un pixel allumé.
- On détermine ensuite la fin de la ligne de texte qui correspond à la matrice binaire qui ne possède aucun pixel allumé. Le balayage horizontal s'effectue de droite à gauche car la langue arabe s'écrit dans ce sens.

Algorithme de la segmentation horizontale :

Début
Tant que non fin image (matrice binaire)
faire
Tant que la ligne de texte ne contient aucun point noir
faire
 - Saut vers la ligne suivante
fait
 Sauvegarder le début de la ligne de texte
Tant que la ligne contient au minimum un point noir
faire
 - Saut vers la ligne suivante
fait
 Sauvegarder la fin de la ligne du texte
 Saut vers la ligne suivante
fait
Fin



Figure 4: Segmentation horizontale

On considère que deux lignes consécutives sont séparées par au moins une ligne vide, pour laquelle la valeur de l'histogramme est nulle. On peut ainsi détecter les limites supérieure et inférieure de chaque ligne du texte. Si l'espace entre deux lignes de texte consécutives contient quelques points (chevauchement entre lignes), on ne peut plus détecter les limites d'une ligne.

A défaut de détecter les limites d'une ligne, on peut détecter la ligne médiane pour laquelle le nombre de pixels allumés est maximum : lignes à densité maximum (voir fig.5).

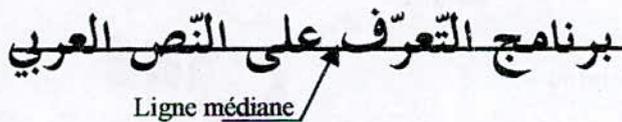


Figure 6 : Détection d'une ligne de texte.

2.3.2. Segmentation verticale :

Elle consiste à déterminer le début et la fin de la partie connexe. Chaque ligne de texte obtenue dans la

segmentation horizontale sera subdivisée en parties connexes qui ne sont pas forcément des mots (voir Fig.6). Le balayage vertical se fait de haut en bas.

Son principe est le suivant :

- Procédant par un balayage vertical, on détermine le début de la partie connexe qui correspond à la première colonne de la matrice binaire qui contient au minimum un pixel allumé.
 - On détermine ensuite la fin de la partie connexe qui correspond à la première colonne de la matrice binaire qui ne contient aucun pixel allumé.

Algorithme de la segmentation verticale :

Début
Pour chaque ligne de texte
faire
Tant que la colonne ne contient aucun point noir
faire
 - Saut vers la colonne suivante
fait
 Sauvegarder la colonne début de la partie connexe
Tant que la colonne contient un point noir au minimum
faire
 - Saut vers la colonne suivante
 Sauvegarder la colonne fin de la partie connexe
fait
Fin

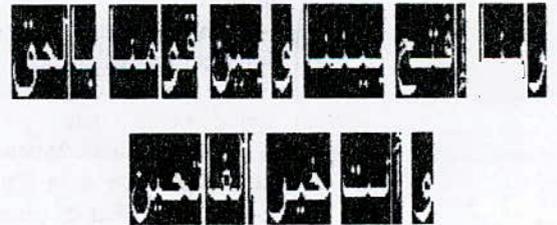


Figure 6 : Segmentation verticale

2.3.3. Segmentation des parties connexes en caractères :

Son but est d'extraire les caractères de chaque partie connexe obtenue dans la segmentation verticale (voir Fig.7).

Pour extraire un caractère de la partie connexe, il faut chercher :

- i- La ligne de jonction ou médiane : celle ci constitue la région qui a la plus forte concentration de pixels en ligne.
- ii- La ligne de haut pour chaque colonne de la partie connexe : Par un balayage vertical commençant par la première ligne de haut, on détermine la ligne max. qui

correspond au numéro de la ligne qui contient le premier point noir rencontré.

iii- La ligne de bas pour chaque colonne de la partie connexe : Par un balayage vertical commençant par la dernière ligne, on détermine la ligne min. qui correspond au numéro de la ligne qui contient le premier point noir rencontré.

iv- L'histogramme vertical pour chaque colonne de la partie connexe.

v- Le seuil (ou zone de constance) qui correspond à la valeur la plus répétitive dans l'histogramme des colonnes.

vi- Le nombre de transitions verticales (0 - 1) ou (1 - 0) pour chaque colonne de la partie connexe.

L'extraction d'un caractère se fait après avoir déterminé son début et sa fin.

Dans notre méthode nous avons introduit quelques conditions permettant la détection du début et de la fin d'un caractère. La colonne qui correspond au début d'un caractère doit vérifier la condition (a) :

(a) L'histogramme vertical de cette colonne doit être supérieur au seuil.

La colonne fin de caractère correspond à la dernière colonne qui précède le début du caractère suivant, de plus cette colonne fin doit vérifier les conditions (b) suivantes:

- (b) {
- 1- La ligne haut de cette colonne doit être inférieure ou égale à la ligne médiane.
 - 2- La ligne bas de cette colonne doit être supérieure ou égale à la ligne médiane.
 - 3- La différence entre la ligne bas et la ligne haut doit être inférieure ou égale au seuil.
 - 4- L'histogramme vertical doit être inférieur ou égal au seuil.
 - 5- Le nombre de transitions verticales doit être égal à deux.
 - 6- La ligne haut de cette colonne doit être supérieure à la ligne haut de la colonne de début du caractère.

L'algorithme suivant permet l'extraction d'un caractère de la partie connexe après avoir déterminé son début et sa fin. Si le caractère est en position médiane, son début correspond à la colonne qui suit immédiatement la fin du caractère précédent.

Algorithme :

deb : variable booléenne (deb = vrai correspond au début du caractère).

Début

deb := vrai ;

Pour chaque partie connexe

faire

Tant que non fin de la partie connexe **faire**

Si deb est vrai alors

- chercher la première colonne qui vérifie la condition (a);

- Sauvegarder la colonne début du caractère courant, soit $j1$;

- deb := faux ;

Si non

- $j1 = j2 + 1$;

- Saut vers la colonne suivante.

- Chercher la colonne de début du caractère suivant, soit $j1$;

Si elle existe alors :

- Faire un balayage vertical de droite à gauche (de $j'1$ à $j1$) ;

- Déterminer la première colonne qui vérifie les conditions(b), soit $j2$;

Si non dans ce cas on a atteint la fin de la partie connexe, soit $j'2$;

fait

Fin.

$j1$ représente la colonne début de caractère.

$j2$ représente la colonne fin de caractère.

ربنا إفتح بيننا وبين قومنا بالحق
وأنت خير الفاتحين

Figure 7 : Segmentation en caractères

2.3.4. Problèmes de la segmentation et solutions apportées :

Nous avons testé cette méthode sur plusieurs textes arabes contenant 500 caractères. Les résultats comme attendus, n'étaient pas une segmentation à 100 %. En effet :

1- Le caractère " sin " demeure sursegmenté. Ce dernier est coupé en trois portions comme le montre l'image suivante :



Nous avons trouvé une solution à ce problème. A la base des constatations observées au niveau du caractère " sin ", nous avons pu déterminer quelques

caractéristiques topologiques de ce caractère; ce caractère diffère des autres par :

- Il se compose de trois parties dont les débuts sont identiques.
- Le nombre de transitions verticales de chaque colonne du caractère " sin " est égal à deux.

Il faut donc modifier l'algorithme de segmentation en caractères comme suit :

Algorithme de segmentation modifié :

Début

- Déterminer le début du caractère courant ; soit j1.
- Déterminer le début du caractère suivant ; soit j1'.
- Déterminer le début du deuxième caractère suivant ; soit j1''.

Si leurs histogrammes verticaux sont égaux et si les colonnes inclues entre ces trois colonnes (j1, j1' et j1'') ont le nombre de transitions verticales égal à deux.

Alors Sauter vers colonne qui suit j1''.

Si non Sauter vers la colonne qui suit j1.

- Déterminer la fin du caractère courant.

Fin.

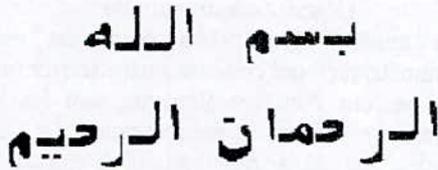


Fig.8: Segmentation du caractère " sin "

2-Il existe aussi certains caractères sous-segmentés (cas des caractères qui se chevauchent et que l'on appelle ligatures comme : ...أحمد، محمد، الملك...). Ces problèmes ont été contournés en grande partie en considérant les caractères qui se chevauchent comme étant un seul.

Cette méthode a été testée sur les mêmes textes que précédemment, tous les caractères ont été correctement segmentés. Nous arrivons donc à un taux de segmentation très proche de 100 %, contrairement aux méthodes utilisées jusque là ([1], [6], [7]).

2.4.Interpretation et Décision :

Elle se fait en deux étapes :

2.4.1 Extraction des caractéristiques :

L'extraction des caractéristiques est le dernier processus de réduction de l'information, il est fortement lié à la structure de la bibliothèque des objets que l'on désire pouvoir reconnaître. En effet, ces caractéristiques doivent avoir la propriété d'effectuer une bonne discrimination d'un objet à l'autre tout en restant en nombre limité afin de réduire les temps de calcul tant au niveau de leur extraction qu'au niveau de leur analyse

par le système de reconnaissance. Ceci explique pourquoi les systèmes actuels de vision sont capables de reconnaître un nombre limité d'objets pour une application donnée tout en gardant des performances honorables.

On peut classer les caractéristiques suivant trois familles :

- *Les caractéristiques topologiques.
- *Les caractéristiques fonctionnelles.
- *Les caractéristiques géométriques.

Dans notre cas, nous avons à faire à des caractères arabes isolés (segmentés). Ces derniers sont décrits par une série de moments, qui font partie des caractéristiques fonctionnelles. On obtient ainsi par décomposition de l'image dans une base de fonctions (moments) une représentation des caractères sous forme de série. La dimension de la base est sept du fait qu'on s'est limité aux sept premiers termes du développement.

Définition des Moments :

Soit I(i , j) l'image traitée que l'on veut reconnaître. Le moment généralisé d'ordre (p+q) d'une telle image s'écrit : [2 , 3]

$$m_{pq} = \sum_{i=0}^{max\ lin} \sum_{j=0}^{max\ col} i^p j^q I(i, j)$$

pour p, q=0,1,2,3.

Dans une image blanche sur fond noir, le moment d'ordre zéro représente la surface du caractère :

$$m_{00} = \sum_{i=0}^{max\ lin} \sum_{j=0}^{max\ col} I(i, j)$$

Les moments d'ordre 1 sont définis par :

$$m_{10} = \sum_{i=0}^{max\ lin} \sum_{j=0}^{max\ col} iI(i, j)$$

$$m_{01} = \sum_{i=0}^{max\ lin} \sum_{j=0}^{max\ col} jI(i, j)$$

Ils définissent le centre de gravité (\bar{X}, \bar{Y}) du caractère considéré :

$$\bar{X} = \frac{m_{10}}{m_{00}} \quad \bar{Y} = \frac{m_{01}}{m_{00}}$$

Les **moments centrés** d'ordre (p, q) sont donnés par :

$$\mu_{pq} = \sum_{i=0}^{max\ lin} \sum_{j=0}^{max\ col} (i - \bar{X})^p (j - \bar{Y})^q I(i, j)$$

Les moments centrés jusqu'à l'ordre 3 sont donnés par :

$$\begin{aligned} \mu_{00} &= m_{00} \\ \mu_{10} &= \mu_{01} = 0 \\ \mu_{20} &= m_{20} - \bar{X} m_{10} \\ \mu_{02} &= m_{02} - \bar{Y} m_{01} \\ \mu_{11} &= m_{11} - \bar{Y} m_{10} \end{aligned}$$

$$\begin{aligned} \mu_{30} &= m_{30} - 3\bar{X} m_{20} + 2m_{10} \bar{X}^2 \\ \mu_{21} &= m_{21} - 2\bar{X} m_{11} - \bar{Y} m_{20} + 2\bar{X}^2 m_{01} \\ \mu_{12} &= m_{12} - 2\bar{Y} m_{11} - \bar{X} m_{02} + 2\bar{Y}^2 m_{10} \\ \mu_{03} &= m_{03} - 3\bar{Y} m_{02} + 2m_{01} \bar{Y}^2 \end{aligned}$$

Ces moments sont invariants par translation de l'image.
Les **moments centrés normés** notés η_{pq} se définissent par l'expression :

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad \text{avec} \quad \begin{cases} p + q = 2, 3, \dots \\ \gamma = \frac{p + q}{2} + 1 \end{cases}$$

Ces moments engendrent des moments généralisés invariants par translation, rotation et homothétie.
Nous en représentons ci-dessous un groupe de sept :

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} + \eta_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \\ &\quad [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03}) \\ &\quad (\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 &= (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \\ &\quad [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30}) \\ &\quad (\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned}$$

Ces moments ont la particularité d'être invariants par translation, rotation et changement d'échelle [2].

Apprentissage :

Cette étape consiste à établir une correspondance entre le caractère et son modèle. Notre système va extraire les moments de chaque caractère de l'alphabet de manière à constituer une bibliothèque de moments ; cette notion de bibliothèque montre bien qu'on évolue dans un univers limité de caractères reconnaissables. La base de données que nous avons élaborée possède une structure matricielle de 61 lignes par 7 colonnes. Chaque ligne présente les caractéristiques (moments) d'un des caractères constituant la base de données.

Nom Moments Code

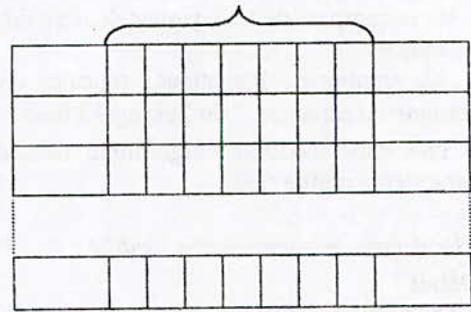


Fig.9 : Organisation du dictionnaire

2.4.2 Reconnaissance et Décision :

L'étape finale de la vision constitue l'interprétation des caractéristiques extraites de l'image ; c'est à dire la reconnaissance des formes. Le choix d'un processus de reconnaissance de formes est lié à la base au type de description que l'on désire pour caractériser un objet. Il existe grossièrement deux :

- Description paramétrique.
- Description structurelle.

Le modèle choisi dans notre cas est un modèle paramétrique qui consiste à caractériser un caractère par un vecteur dont les éléments sont les sept moments invariants décrits précédemment. Le choix d'un tel modèle est dû à sa simplicité de construction. La seule difficulté consiste à choisir une liste de caractéristiques qui permettent une bonne discrimination entre les différents objets de la bibliothèque.

Règle de Décision :

Nous avons vu que chaque caractère de la base de données est représenté par un vecteur S_{ij} où $i=1..61$ et $j=1..7$, i désignant l'indice du caractère et j celui relatif aux caractéristiques (ordre du moment).

La règle de décision utilisée s'appuie sur la notion de distance euclidienne. On définit cette distance entre deux points repérés par leurs vecteurs r_1 et r_2 de dimension N par :

$$D(r_1, r_2) = \left\| r_1 - r_2 \right\| = ((r_1 - r_2)(r_1 - r_2)^T)^{1/2}$$

Où le facteur $(^T)$ indique le vecteur transposé qui est utilisé pour mesurer la ressemblance entre deux objets. La reconnaissance d'un caractère représenté par un vecteur R consiste à trouver la distance minimale $D(S_i, R)$ définie comme suit :

$$D(S_i, R) = \text{Min} \{D(S_i, R)\}_{i=1..61}$$

Donc la $i^{\text{ème}}$ ligne de la base de données représente les moments de l'élément le plus voisin du caractère x que l'on veut reconnaître.

3. Conclusion:

L'objectif visé est largement atteint : il s'agissait d'étudier les performances des moments sur les caractères arabes isolés (segmentés) dans leurs différentes positions (début, milieu, fin et isolés). Les sept moments invariants ont été suffisants pour modéliser les caractères avec une bonne discrimination. En effet, le taux de reconnaissance global qui est de 98% est acceptable avec un temps de calcul largement dans les normes. La méthode s'est avérée efficace pour les caractères ayant subi des translations, rotations et symétries, mais moins performante pour l'homothétie à cause des erreurs d'arrondi qui interviennent dans le calcul de cette transformation. Ce qui n'est pas le cas pour une homothétie réelle (non simulée).

Dans le cas des méthodes statistiques, le problème de fonte et de taille des caractères reste non résolu (car les moments représentent, entre autres, la surface du caractère et son centre de gravité). Ceci implique que nous devons apprendre à notre système de reconnaissance toutes les fontes de caractères que nous voudrions reconnaître, ce qui exige une grande taille mémoire pour la réalisation du dictionnaire

Nous avons pour cela, introduit une routine de normalisation de la taille du caractère juste après l'étape de segmentation ; ce qui nous permet d'avoir une bibliothèque réduite et aussi une taille uniforme du caractère à l'entrée du module de reconnaissance. Nous avons appris à notre système trois fontes.

Le logiciel a été réalisé en langage DELPHI.4.

4. Références:

[1] M.B Kurdy and A.Joukhadar, Multifont Recognition System for Arabic Characters, Proceedings of the 3rd International Conference and Exhibition on Multilingual Computing, Durham U.K. Dec.1992.

[2] R.C Gonzalez and P. Wintz, Digital Image Processing, Addison Wesley 1987

[3] M.K Hu, Visual pattern recognition by moments invariant, I.R.E Trans. Inf. Theory. IT, 8 Feb. 1962

[4] L. Hamami-Mitiche et D. Berkani, Système de Reconnaissance de l'écriture Arabe Imprimée par utilisation d'une Méthode Statistique, 2^{ème} Conférence Internationale d'Electronique sur les Signaux, Systèmes et Automatique, SSA2'99, 10-12 Mai 1999 Blida - Algérie, Volume I pp.112-116.

[5] L. Hamami-Mitiche, Segmentation of an Arabic text printed in characters, Proceedings of the 8th International Conference on Computer Theory and Applications, ICCTA'98 IEEE (Alexandria Chapter), 15-17 September 1998 Alexandria, Egypt, pp. IV.6-IV.8.

[6] A.Amin and H.B.Al-Sadoun, A new segmentation technique of Arabic text, Proceeding IEEE , p.441-445,1992.

[7] K.Bouhlila, M.K Hamrouni and N.Ellouze, Method of segmentation of Arabic text image into characters, First Kuwait Conference, Mars 1989.

[8] H. Yamada, K. Yamamoto and T. Saito, A Non-linear Normalization Method for Hand-printed Kanji Character Recognition - Line Density Equalization -, CH2614-6/88/0000/0172\$01.00,1988 IEEE, pp.172-175.

[9] S. El-Dabi, R.Ramsis and A.Kamel, Arabic Character Recognition System : A Statistical Approach for recognizing cursive typewritten text, Pattern Recognition, pp.485-495, 1990.

[10] A. Khotanzad and Yaw Hua Hong, Invariant Image Recognition by Zernike Moments, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.12, n°5, May 1990, pp.489-497.

[11] T.S. El-Sheikh and R.M. Guindi " Computer Recognition of Arabic Cursive Script." Pattern Recognition, Vol.21, No.4, 1988, pp 293-302.

[12] H.Al-Yousefi and S.S.Udpa, « Recognition of arabic characters »IEEE transaction on P.A.M.I, Vol.14, n°8, Aug.92, pp. 853-857.

[13] Abdel.Belaid and Y.Belaid, « Reconnaissance de formes, Inter-Edition, Paris, 1992, 412p.

[14] I.S.I Abuhaiba, S.A. Mahmoud and R.J. Green, Recognition of handwritten cursive arabic characters, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.16, n°6, 1994, pp.664-672.

[15] T.S. El-Sheikh and R.M. Guindi, Automatic recognition of isolated Arabic characters, Signal Processing, vol.14, 1988, pp. 177-184.

[15] T.S. El-Sheikh and S.G. El-Taweel, Real-time Arabic handwritten character recognition, Pattern Recognition, vol.23, n°12,1990, pp.1323-1332.

[16] S.A. Mahmoud, I.S.I Abuhaiba, and R.J. Green, Skeletonization of Arabic characters using clustering based skeletonization algorithm (CSBA), Pattern Recognition, vol.24, n°5, 1991, pp.453-464.

[17] Al-Emami and M. Usher, On-line recognition of handwritten Arabic characters, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.12, n°7, 1990, pp.704-710.

[18] M.S. El-Wakil and A.A. Shoukry, On-line recognition of handwritten isolated Arabic characters, Pattern Recognition, vol.22, n°2, 1989, pp.97-106.

CONTRIBUTION A L'ETUDE ET LA REALISATION D'UN SYSTEME DE RECONNAISSANCE AUTOMATIQUE DE LA PAROLE (SRAP) CONTINUE MULTILOCUTEUR PAR LES TECHNIQUES D'INTELLIGENCE ARTIFICIELLES

B. BOUSSEKSOU, M^{me} K. DJOUADI, M. D. KONE
Laboratoire Signal et Communications
Ecole Nationale Polytechnique BP. 182, 16200 ALGER, ALGERIE
Fax : (213) 02.52.24.65, Tel : (213) 02.52.24.65

ملخص

نقدم في هذا العرض جهاز يكشف آليا القول المتعدد الناطقين. وهذا الجهاز الذي أقيم على أساس معارف الخبراء في الميدان ، يستعين بالتجول القوية للذكاء الصناعي. وبما أن المعاينة في كشف القول غير واضحة ومعرضة للتطور ، فإن البرامج المقيمة بـ " ليسب " و " ولوق " أعدت لكي أعدت لكي تحمل كل المعارف وتنبأ بتطوراتها. وإن هذا المفهوم يعطي للجهاز إمكانيات تيسر له التوسع والتغنية. وفيما يلي ، نقدم وصفا للأجزاء المختلفة التي تكون جهازنا .
المصطلحات الأساسية المستعملة :

تعرف القول ، تلصيق ، تعريف لفظي ، تحليل جدولي ، جهاز خبير ، محرك الاستدلال

ABSTRACT

We present an expert system of an automatic recognition of a continuous speech multispeakers. This system, based on the experts knowledges in the field, use the powerful performances of artificial intelligence [1]-[2]. The expertise in recognition of speak being insufficiently and exposed to evolution, developed softwares in Lisp and Prolog, are conceived to support the set of knowledges and previous their modification. This conception gives to the system possibilities of extension and enrichment. In the following, we describe the different modules which compose our system.

Key words : Speech recognition, labelling, segmentation, phonetic identification, lexical analysis, expert system, inferences engine, scheme generator, lexical search, syntactic search, flexible words, generative grammar, morphemes, Hidden Markov Models (HMMs), dynamic time warpping, dynamic programming.

RESUME

Nous présentons un système expert de Reconnaissance Automatique de la Parole continue multilocuteur. Ce système, basé sur les connaissances des experts du domaine, fait appel aux puissantes performances de l'intelligence artificielle [1]-[2]. L'expertise en reconnaissance de la parole étant insuffisamment explicite et sujette à évolution, les logiciels développés en Lisp et Prolog sont conçus pour supporter l'ensemble des connaissances et prévoir leur modification. Cette conception donne au système des possibilités d'extension et d'enrichissement Dans ce qui suit, nous décrivons les différents modules qui composent notre système.

Mots clés : Reconnaissance de la parole, segmentation, identification phonétique, analyse lexicale, système expert, moteur d'inférences, schème générateur, recherche lexicale, recherche syntaxique, mots fléchis, grammaire générative, morphèmes, Modèles de Markov Cachés (HMM), alignement temporel dynamique (DTW), programmation dynamique (DP)

I. INTRODUCTION

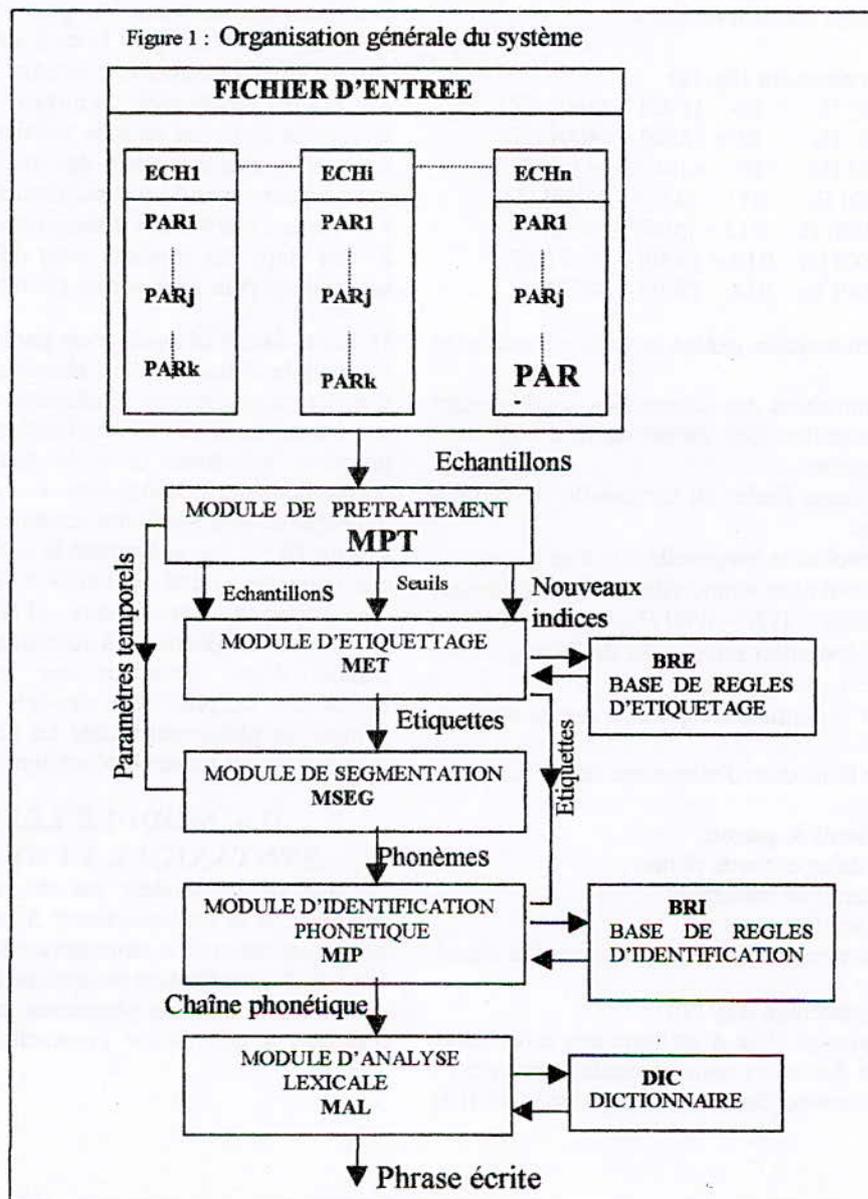
La Reconnaissance Automatique de la Parole (RAP), qui vise comme objet ultime la communication en langage naturel avec la machine, est un problème non encore résolu totalement à l'heure actuelle. Les principaux problèmes posés par le domaine sont liés au processus de la communication orale et aux propriétés du signal de parole. En effet, la parole présente une très grande variabilité. De plus, elle véhicule plusieurs niveaux d'information (phonétique, lexical, sémantique, etc.) insuffisamment explicites pour être modélisée par des méthodes algorithmiques [3,4,5,6,7,8,9,10,11].

Notre travail s'inscrit dans une contribution à cette délicate tâche en tenant de mettre au service de la Reconnaissance Automatique de la Parole les puissantes techniques de l'intelligence artificielle.

I. PRESENTATION DU SYSTEME (fig.1)

Se basant sur les caractéristiques acoustiques des différents phonèmes d'une langue, notre système tente de faire correspondre à un signal acoustique, décrit par des paramètres, une suite d'unités phonétiques identifiées. La difficulté de ce traitement tient au fait qu'il n'existe pas de relations entre une portion de signal et un phonème donné. Notre système [12] a donc pour fonctions principales :

- L'extraction des paramètres pertinents ;
- La segmentation du signal en unités bien définies;
- L'identification de ces unités ;
- La reconstitution, à partir de ces dernières, de la phrase à reconnaître.



II.1. Paramètres retenus

Pour des raisons de nonstationnarité, le signal vocal est divisé en échantillons centisecondes représentés par des vecteurs de paramètres spectraux. Nous avons retenu la représentation suivante :

ECSi = (E1 E2 E3 E4 E5 E6 E7 E8 E9 E10 E11 E12 E13 E14 IF0 E1 IF1 IF2 TPZ).

Chaque échantillon de signal est donc représenté par 20 paramètres correspondant à :

E1 Energie dans la bande de fréquence **B1**

E2 Energie dans la bande de fréquence **B2**

.....

.....

Ej Energie dans la bande de fréquence **Bj**

.....

.....

E14 Energie dans la bande de fréquence **B14**

IF0 Intensité du fondamental

F1, F2 Fréquence des deux premiers formants

IF1,IF2 Intensités des deux premiers formants

TPZ Taux de passage par zéro du signal

II.2. Module de traitement (fig.2a)

B1 = [250 - 450] Hz **B8** = [1900 - 2500] HZ

B2 = [450 - 650] Hz **B9** = [2500 - 3400] HZ

B3 = [650 - 850] Hz **B10** = [3400 - 4300] HZ

B4 = [850 - 1050] Hz **B11** = [4300 - 5000] HZ

B5 = [1050 - 1300] Hz **B12** = [5000 - 6000] HZ

B6 = [1300 - 1600] Hz **B13** = [6000 - 7000] HZ

B7 = [1600 - 1900] Hz **B14** = [7000 - 8000] HZ

Le module de prétraitement génère le fichier d'entrée en assurant :

- La sortie des paramètres des échantillons centisecondes successifs qu'il organise sous forme de liste ordonnée, facilement manipulable.

- La création de toute évolution temporelle de certains paramètres tels que :

- La courbe d'évolution temporelle de l'énergie totale,
- La courbe d'évolution temporelle de l'énergie basses fréquences : **BBF** = [250 - 650] Hz,
- La courbe d'évolution temporelle de l'énergie dans les bandes [250 - 2500] Hz,
- La courbe d'évolution temporelle des centres de gravité;

- L'interface avec le module d'étiquetage des échantillons centisecondes;

- La détection du seuil de parole;

- La détection du début et fin de phrase ;

- La détection du seuil de voisement ;

- La détection du seuil de friction ;

- Le calcul de l'énergie maximale contenue dans le signal.

II.3. Module d'étiquetage (fig. 2b)

Le module d'étiquetage tente d'attribuer aux échantillons centisecondes des étiquettes nous permettant de savoir à quelles macro-classes ou classes ils appartiennent [13].

Ces étiquettes sont inférées par un ensemble de règles contenues dans la "base de règles d'étiquetage". Les règles opèrent sur des indices acoustiques déterminés par les paramètres de l'échantillon pour identifier les traits phonétiques distinctifs [14].

Les indices retenus sont :

- macro - classe voyelle
- macro - classe consonne
- voyelle postérieure
- voyelle antérieure
- voyelle centrale
- trait d'ouverture des voyelles
- indice de voisement, etc.

II. 4. Module de segmentation (fig. 3a)

Le module de segmentation opère en 4 étapes :

1^{ère} étape : segmentation en pseudo-syllabes

Sachant que la syllabe est une unité facilement localisable, car généralement comprise entre deux minima d'énergie totale, nous construisons dans une première étape des unités pseudo-syllabiques. Chaque segment ainsi constitué contient un maximum local d'énergie totale correspondant généralement à un noyau vocalique.

2^{ème} étape : réduction de segments excédentaires

les segments "pseudo-syllabiques" construits par la première étape sont contrôlés pour vérifier s'ils présentent bien des caractéristiques liées à l'énergie totale et à la durée compatibles avec une syllabe.

3^{ème} étape : localisation du noyau vocalique

la voyelle contenue dans la syllabe est détectée dans la zone stationnaire autour du maximum. Le reste du segment correspond à une réalisation consonantique.

4^{ème} étape : attribution d'étiquettes aux phonèmes

à cette étape, les résultats issus du module d'étiquetage sont utilisés pour localiser les phonèmes et les étiqueter.

II. 5. Module d'identification phonétique (fig. 3b)

Le module d'identification phonétique [15,16,17] a pour fonction de reconnaître le phonème qui lui est présenté à son entrée. Le signal est segmenté en phonème et chaque phonème est donné avec la liste des étiquettes des échantillons qui le composent. Les étiquettes d'un même phonème ne sont pas toutes identiques car nous avons pris le soin de ne pas segmenter le signal par regroupement des étiquettes. Cette opération a été prise pour pouvoir tenir compte des erreurs d'étiquettes éventuelles. L'étiquette du phonème à identifier sera dans notre cas déduite d'une statistique sur les étiquettes qui le composent. De plus, cette stratégie permet de prendre en compte les phénomènes dont les réalisations acoustiques comportent des phases discontinues tels que les plosives.

II.6. MODULE LEXICAL ET SYNTAXIQUE ET SYNTAXIQUE

Le rôle de ce module est de reconstituer la phrase candidate à la reconnaissance à partir de la chaîne de phonèmes (ou treillis de phonèmes) donnée à la sortie du module d'identification phonétique [18,19,20]. Il s'agit de passer d'une suite de phonèmes probablement entachée d'erreurs à une phrase éventuellement incomplète (ou

multiples) sous la forme écrite, c'est à dire, en lui faisant correspondre une orthographe. Ceci nécessite de traiter le passage de la forme orale à la forme écrite. Notre système correspond au modèle suivant :

Figure 2 : (a) Organisation générale du module de prétraitement

(b) Fonctionnement du module d'étiquetage

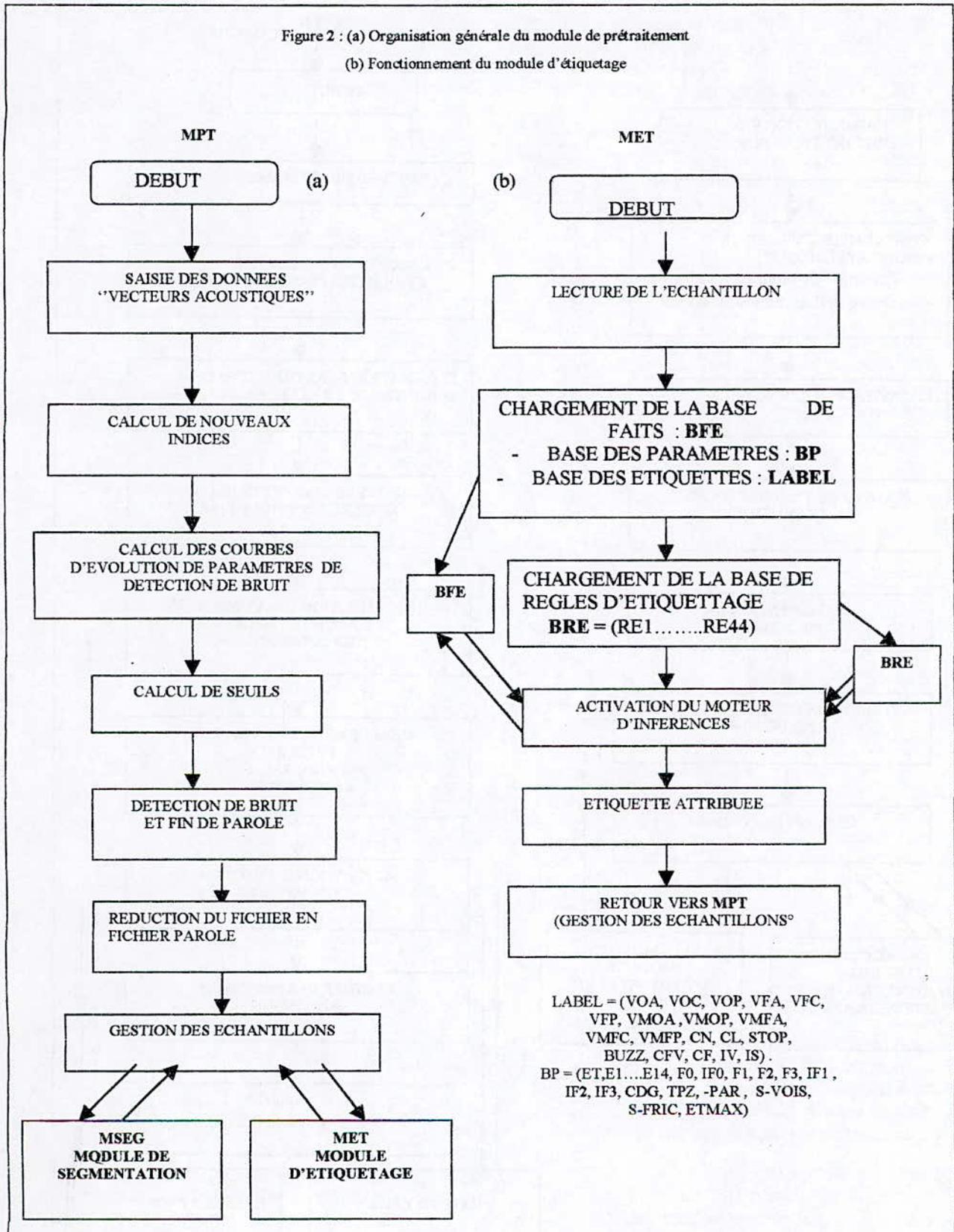
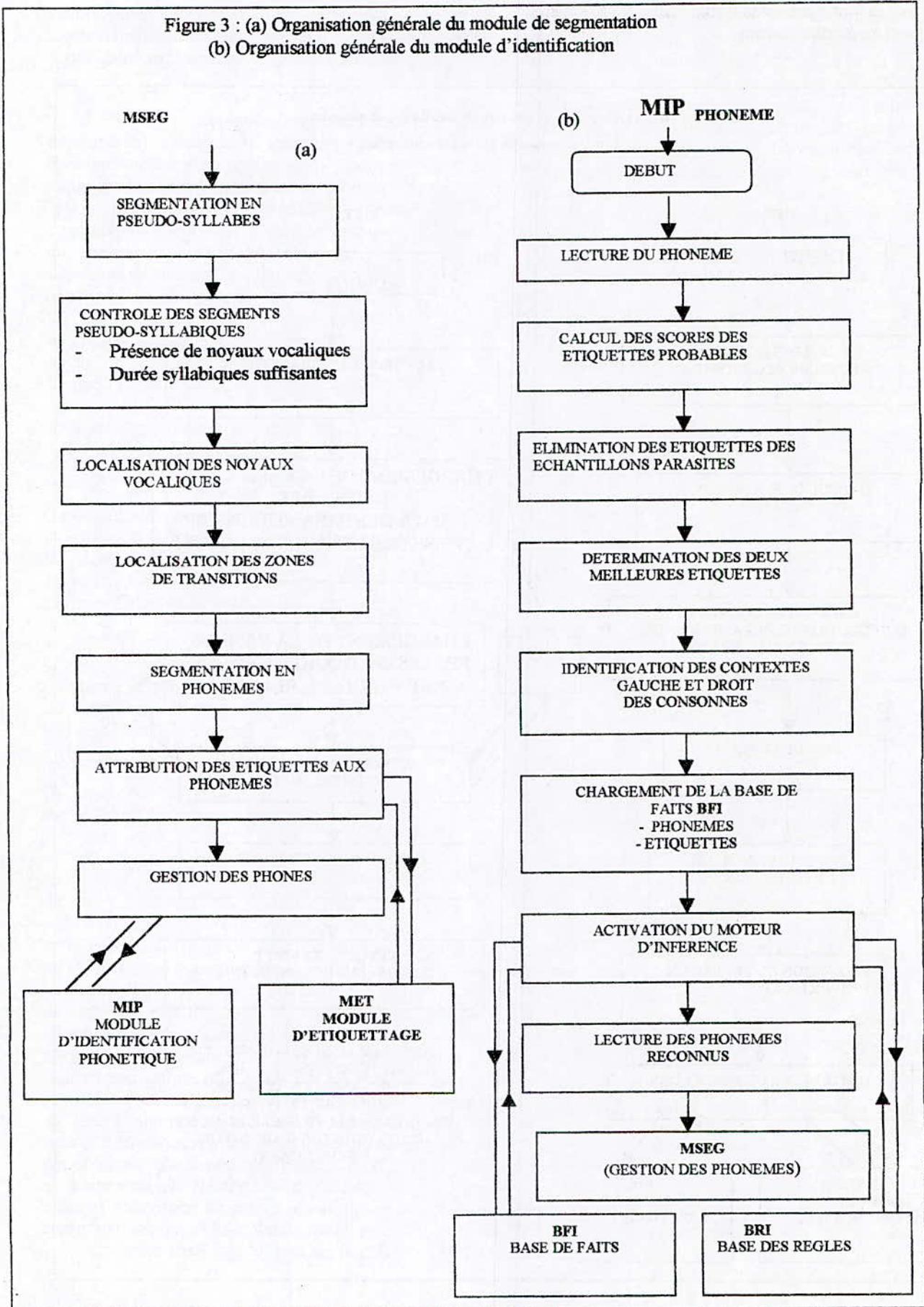


Figure 3 : (a) Organisation générale du module de segmentation
 (b) Organisation générale du module d'identification



I.6.1. Organisation de la composante lexicale

II.6.1.1. Base de données lexicales

Intégrer un corpus lexical important, en vue de l'utiliser dans un traitement informatique, pose de véritables problèmes de taille mémoire, de complexité de mesure de gestion et d'accès aux données. De plus, on doit envisager la possibilité de partager le lexique entre plusieurs utilisateurs ou plusieurs applications. Pour ce faire, nous avons opté pour un lexique dynamique. La saisie des mots utilisés pour une application se fait indépendamment des autres applications, avec le choix de garder ou pas l'ancien lexique (le lexique saisi ultérieurement).

II.6.1.2. Natures des entrées lexicales

Il s'agit de déterminer le type d'éléments qui va constituer le lexique. Pour se faire une idée des différentes possibilités qui existent, on peut examiner les formes d'organisation utilisées dans les dictionnaires destinés à la distribution grand public, et dans les lexiques à vocation généralement plus linguistique. Il existe deux conceptions différentes pour le lexique, celle des lexicographes et celle des lexicologues [21].

II.6.1.2.1. Organisation lexicographique

Cette organisation propose des lexiques comportant des faits de tout ordre, destinés à être consultés par des usagers, locuteurs d'une langue mais spécialistes de celle-ci. Les mots qui composent la langue française sont divisés en deux catégories : les mots grammaticaux et les mots lexicaux [21]. Les premiers sont limités par leur nombre et ne posent pas de problèmes d'organisation, ils se présentent sous une seule forme ; c'est le cas des pronoms en général. Par contre, les mots lexicaux sont les noms, les adverbes, les adjectifs et les verbes. A l'exception des adverbes invariables, tous les mots lexicaux possèdent des formes fléchies qui correspondent pour les noms et les adjectifs à la flexion en genre et en nombre, et pour les verbes à la conjugaison. Dans la perspective de l'implantation informatique d'une composante lexicale, il faut que l'organisation choisie soit sans ambiguïté vis - vis de ce type de problème.

II.6.1.2.1.1. Forme fléchie

Un tel type d'organisation implique que toutes les formes propres à un mot de la langue étudiée, soient représentées dans le lexique, ce qui équivaut à disposer d'un lexique en extension [22]. Ceci présente un certain intérêt dans la mesure où on est dispensé de la mise en œuvre des processus flexionnels permettant d'obtenir les formes fléchies. Mais surtout, cela permet de traiter de manière simple les exceptions et les formes présentant des irrégularités. Un tel schéma d'organisation n'est convenable que pour une langue comme l'Anglais qui présente un nombre de formes fléchies, par opposition aux langues latines telles que l'Italien, l'Espagnol et le Français, qui sont très riches en formes fléchies.

II.6.1.2.1.2. Forme canonique

Ce mode d'organisation [18], conventionnellement utilisé par la plus part des dictionnaires, consiste à ne présenter les formes verbales, nominales ou adjectivales que sous un seul aspect, respectivement celui de la forme infinitive,

de la forme au féminin singulier et de la forme au masculin singulier. Il s'agit de factoriser les formes fléchies d'un même mot modulo la flexion genre/nombre ou la conjugaison et d'en faire une classe d'équivalence qui sera représentée dans le lexique par une forme [22,23,24].

II.6.1.2.2. Organisation lexicologique

Basés sur des faits purement linguistiques, les linguistes ont conçu un modèle linguistique du fonctionnement du langage, puis proposé un modèle de lexique plutôt qu'un outil réellement consultable. D'après Ferrane [25], Corbin [27] a cité à ce propos une phrase empruntée à Rey [26] "*le lexique des linguistes, composante du système abstrait de la langue, est un modèle théorique cohérent, mais dérisoire par rapport au lexique, objet historique et anthropologique énorme et confus.*" [27].

Cette organisation présente ainsi des formes différentes suivant toujours l'unité lexicale choisie. Ici, suivant les liens morphologiques, on rencontre deux opinions divergentes :

Certains linguistes sont partisans d'une morphologie du mot, d'autre d'une morphologie du morphème [21].

II.6.1.2.2.1. Morphologie du mot [20]

C'est une politique partagée par la majorité des linguistes. Dans ce cas, on a considéré que l'unité morphologique minimale doit être le mot. En effet, ce choix est convainquant du fait de l'unicité des caractères grammaticaux pour un mot. On peut donc bien l'accepter comme unité de base.

II.6.1.2.2.2. Morphologie du morphème

Le morphème est une unité morphologique inférieure au mot. Il possède, selon les théories, une partie plus ou moins étendue des propriétés du mot, à l'exception des traits syntaxiques.

II.6.1.2.3. Unité lexicale

L'unité lexicale est la composante de base dans un SRAP continue. Il faut donc choisir la plus proche de l'unité de traitement dans de tels systèmes. Par ailleurs, dans ce contexte, on se ramène toujours à la reconnaissance des mots. En effet, en se référant à l'introduction de ce chapitre, on remarque que le mot répond bien aux interrogations posées. En gros, l'unité lexicale doit alors véhiculer toutes les informations nécessaires au traitement. Vu la complexité de ce type de traitement, on peut se poser la question de savoir s'il faut considérer le mot sous son aspect fléchi ou pas. De plus, du point de vue informatique, on peut dire que quel que soit la capacité mémoire des machines, elle est toujours limitée. De ce fait, la forme de présentation à choisir ne doit pas favoriser les risques d'atteindre cette limite. Parmi les représentations citées ultérieurement (canonique, fléchie), nous avons choisi une organisation basée sur la forme canonique complétée par un système de génération de formes fléchies. Vu que notre travail rentre dans le cadre de la RAP ; nous avons couplé à cette représentation une autre représentation qui est plus proche du phonème.

Cette dernière sera d'un apport très important pour le calcul du taux de reconnaissance d'un mot. En effet, au début on représente le mot sous sa forme phonétique la plus simple, complétée par ses traits syntaxiques et quelques traits sémantiques avec la représentation orthographique.

A partir de là, on va représenter chaque phonème de chaque mot dans le contexte où ce dernier est utilisé. On sait qu'en parlant, un son particulier peut être confondu avec un autre, donc pour avoir un traitement complet il faut représenter toutes les particularités d'un phonème suivant l'application.

II.6.1.2.4. Présentation du lexique choisi

La forme fléchie augmente considérablement le risque d'atteindre une limite de la capacité mémoire. Pour palier à ce problème, on mettra au point un système de renvoi, permettant d'obtenir les formes fléchies à partir de la forme canonique

En résumé, notre lexique est basé sur la forme canonique des mots à laquelle on a ajouté un système qui génère la conjugaison des verbes. Chaque mot du lexique est représenté par :

- Sa transcription phonétique,
- Son type (verbe, adjectif, nom,..etc.)
- Ses paramètres suivant le type,
- Indication concernant la première lettre des mots pouvant suivre le mot courant,
- Sa transcription orthographique,
- Son rang dans le lexique.

C'est ce qui est appelé dans notre travail : lexique de base (schème générateur).

En parallèle à cela et pour résoudre le problème de la conjugaison des verbes, une représentation des terminaisons des verbes est introduite suivant la classification de Beschrelle [28].

II.6.2. ORGANISATION SYNTAXIQUE

Le niveau syntaxique joue un rôle important dans un SRAP. La syntaxe définit les séquences de mots autorisés. Elle permet de limiter le nombre de solutions qui peuvent interpréter le treillis phonétique. Son objectif est donc d'extraire parmi l'ensemble des combinaisons possibles. Les modèles syntaxiques, issus des travaux en linguistique dans le cadre des langues écrites constituent une base intéressante pour la parole. Dans notre travail, nous avons opté pour le mode des grammaires génératives de Chomsky [29].

II. 6.2.1. Grammaire générative

Ce type de grammaire [20] se présente comme une suite de règles de réécriture. Le mécanisme de dérivation obtenu en appliquant, les unes après les autres, ces différentes règles permettent d'engendrer l'ensemble des phrases syntaxiquement correctes du langage. Nous illustrons ici deux méthodes d'analyse.

II. 6. 2. 2. 1. Analyse syntaxique descendante

L'analyse syntaxique descendante consiste à appliquer un algorithme qui, partant de l'axiome, construit le chemin

de proche en proche en suivant les règles de production de la grammaire sur une phrase.

Exemple : soit G une grammaire dont les règles de production sont :

<PHRASE> := <GN><GV> (1)

<GN> := Pronoms personnels (2)

<GV> := Verbe (3)

{PHRASE, GV, GN} : ensemble du vocabulaire non terminal.

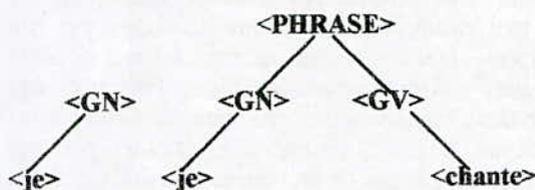
{Pronoms personnels, Verbe} : ensemble du vocabulaire terminal.

Soit à analyser la phrase W suivante : "je chante". L'analyse descendante de cette phrase se fait comme suit :

$\langle \text{PHRASE} \rangle \xrightarrow{(1)} \langle \text{GN} \rangle \langle \text{GV} \rangle \xrightarrow{(2)} \text{je} \langle \text{GV} \rangle \xrightarrow{(3)} \text{je chante}$

II. 6.2.2.2. Analyse syntaxique ascendante

L'analyse syntaxique ascendante, quant à elle construit le chemin vers l'axiome à partir de la phrase elle-même ; Reprenons l'exemple précédent :



II.6.2.2.3. Solution retenue (fig. 4)

L'analyse syntaxique agit directement sur le treillis de phonèmes. Il dirige même la recherche lexicale. Cette opération se fait en respectant l'algorithme suivant : Soit à reconnaître une phrase composée d'un sujet, d'un verbe et d'un complément. Pour cela, il faut traiter sa représentation phonétique comme suit :

Algorithme

- (1) – lancer la recherche lexicale suivant le premier phonème du treillis.
- (2) – Si la recherche lexicale est satisfaite alors calculer le taux de reconnaissance du mot sinon émission d'hypothèses.
- (3) – lancer la recherche lexicale suivant le premier phonème du reste du treillis.
- (4) – Si la recherche lexicale est satisfaite alors vérifier si le mot satisfait une des règles de la grammaire choisie sinon émission d'hypothèses;
- (5) – Si l'étape (4) est vérifiée alors calculer le taux de reconnaissance du mot.
- (6) – Aller à (3) jusqu'à la fin du treillis.

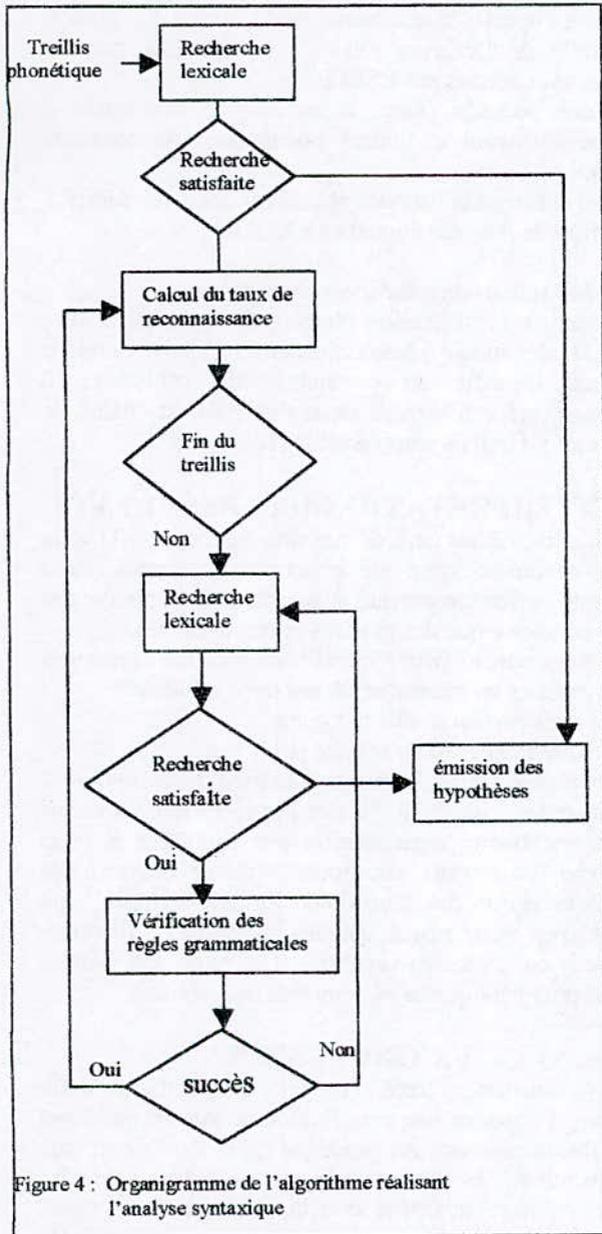
II.6.3. Traitements utilisés et présentation du logiciel

II. 6. 3.1. Introduction

Dans un SRAP, les traitements nécessaires sont de deux types :

- les traitements de recherches,
- les traitements de filtrages.

Les traitements de recherches sont directement liés au lexique de l'application. Ils sont regroupés en gros dans l'analyse lexicale. Par contre, les traitements de filtrages dépendent plus de la grammaire choisie et de la syntaxe de la langue traitée. Ils sont directement liés à l'analyse syntaxique.



II. 6. 3. 2. Analyse lexicale

II. 6. 3. 2. 1. Objectif

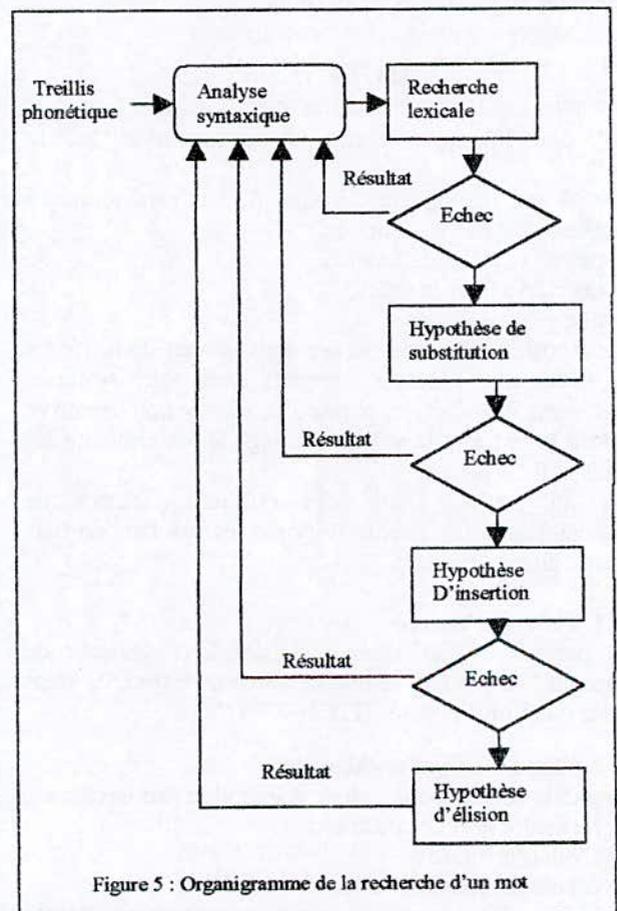
L'analyse lexicale joue un rôle important pour le traitement de la parole. Son objectif est de localiser les mots suivant un certain argument de recherche. Dans notre cas, l'argument de recherche est le premier phonème d'une chaîne phonémique pour chaque niveau de recherche. Cette étape est dirigée par l'analyse syntaxique.

II. 6. 2. 2. Recherche lexicale (fig. 5)

Au cours de l'analyse d'une certaine phrase, l'analyseur syntaxique sollicite le niveau lexical et lui demande de lui envoyer un mot. Si le mot existe, il sera immédiatement traité syntaxiquement. Ce processus s'exécutera jusqu'à l'obtention de la phrase complète.

Plus en détail, la recherche d'un se fait en respectant l'algorithme suivant :

- (1) - Chercher dans le lexique le mot qui commence par le premier phonème de la chaîne phonémique à traiter.
- (2) - Si succès alors aller à (8)
- (3) - Hypothèse de substitution.
- (4) - Si succès alors aller à (8)
- (5) - Hypothèse d'insertion.
- (6) - Si succès alors aller à (8)
- (7) - Hypothèse d'élimination.
- (8) - Renvoyer le résultat à l'étape précédente.



II. 6. 3. 3. Analyse syntaxique

L'analyse syntaxique consiste à réaliser une procédure indéterministe permettant le traitement d'une phrase en un temps minimum (pour obtenir des résultats proche du temps réel), et une programmation dans le langage choisi : PROLOG. L'objectif essentiel de cette procédure est de réduire l'indéterminisme.

II. 6. 3. 3. 1. Réduction de l'indéterminisme [19]

Parmi les causes de l'indéterminisme on a :

- la définition des terminaux,

- la possibilité d'obtenir plusieurs résultats, même dans le cas d'une grammaire non ambiguë.

La procédure d'analyse syntaxique traite ces deux problèmes à deux niveaux :

- Par des filtres syntaxiques qui dépendent de la linguistique,
- Par un calcul de score de reconnaissance (HMMs ou DTW) [18].

II. 6. 3. 3. 2. Réduction de l'indéterminisme par des filtres syntaxiques

Il s'agit, en fait, de traiter le choix entre les différentes règles de la grammaire vus aux paragraphes précédents pour les différents terminaux possibles.

Soit à traiter une certaine phrase. Après la reconnaissance d'un mot, il existe des filtres syntaxiques au niveau de la procédure qui peuvent provoquer l'échec de l'analyse si le prochain mot reconnu n'est pas valide.

La validité d'un mot repose uniquement sur des contraintes linguistiques de la langue.

III. TESTS

Nous avons testé notre système avec la phrase "ce chat subi" dont l'écriture phonétique correspond à "sø fa sybi".

Cette phrase, incomplète du point de vue sémantique, a été choisie car elle regroupe :

- quatre voyelles différentes,
- deux fricatives sourdes,
- une plosive voisée.

La composition de cette phrase nous permet de tester les indices les plus robustes contenus dans notre système. Nous avons introduit des phrases de silence non dépourvu de bruit pour tester la validité du logiciel déterminant les débuts et fin de parole.

Pour un meilleur suivi des différentes étapes de reconnaissance, le système imprime les résultats partiels issus de chaque module.

III. 1. Fichier d'entrée

Les paramètres de chaque échantillon centiseconde composant la phrase candidate ont été introduits dans fichier que l'on a appelé "FICH-ENT".

III. 2. Module de pré traitement

Ce module calcule pour chaque échantillon centiseconde :

- Le centre gravité spectral ;
- L'énergie totale ;
- L'énergie dans la bande [250 - 650] Hz.
- L'évolution temporelle de ces paramètres est donnée respectivement par :

LCDG, LET, LEBF (voir listing)

Le module de pré traitement calcule aussi :

- L'énergie totale maximale **ETMAX**.
- le seuil de parole **S-PAR**,
- Le début de parole **DEB** ;
- La fin de parole **FIN** ;
- Le seuil de voisement **S-VOIS** ;
- le seuil de friction **S-FRIC**.

Il réduit le fichier d'entrée en un fichier ne comportant plus que le signal de parole effective **FICH-PAR**.

III. 3. Module d'étiquetage

Le module d'étiquetage attribue à chaque échantillon centiseconde une étiquette. La liste des étiquettes attribuées est contenue dans **LETIQ**.

III. 4. Module de segmentation

Dans une première étape, il segmente le signal en pseudo-syllabes en opérant une découpe sur la courbe d'évolution temporelle de l'énergie totale. Les segments pseudo-syllabes sont donnés par **LSEG**.

Dans une seconde étape, le module de segmentation découpe le signal en unités phonétiques données par **LPHON**.

Enfin, il construit la liste des phonèmes étiquetés **LPHET** en utilisant la liste des étiquettes **LEQIT**.

III. 5. Module d'identification phonétique

Le module d'identification phonétique construit la liste **LTREIL** des treillis phonétiques en attribuant à chaque phonème identifié un coefficient de confiance. Il détermine, grâce à la base de règles **BRI**, la chaîne de phonèmes **CHAINE** composant la phrase test.

IV. INTERPRETATION DES RESULTATS

Ce test a donné un taux de reconnaissance de 100%, la phrase candidate ayant été entièrement reconnu. Cette apparente perfection est due, à la nature de la phrase test qui ne comporte que des plosives et des fricatives.

Les indices retenus pour l'identification de ces consonnes et des voyelles en général semblent donc robustes.

Ce test nous permet d'affirmer que :

- le voisement est bien détecté par le système ;
- les plosives et les fricatives sont bien discriminées.

La phrase le "canard lit" testée par le système a donné lieu à une bonne segmentation par étiquettes mais la recherche des noyaux vocaliques a donné des résultats erronés en raison des transitions "liquide-voyelle" qui ne sont pas aussi nettes que les transitions "fricative-voyelle" ou "plosive-voyelle". De plus, les indices retenus pour les liquides ne semblent pas robustes.

V. CONCLUSION

Nous avons tenté, à travers l'élaboration de notre système, d'apporter une contribution au délicat problème de la Reconnaissance Automatique de la Parole continue multilocuteur. Les résultats obtenus, variables selon les phrases testées, montrent que la faiblesse du système réside dans le manque de robustesse de certaines utilisées par le système. Néanmoins, l'indépendance totale des logiciels vis-à-vis des connaissances expertes permet un enrichissement, voire même une modification complète des bases de règles. Pour des domaines d'application limités avec des contraintes sur le langage et un vocabulaire de petite taille, le système s'avère performant. Nous constatons que le niveau local porte essentiellement sur la mise en œuvre d'un modèle local (dictionnaire) en se basant sur des caractéristiques purement linguistiques. Pour le niveau syntaxique, le moteur du logiciel s'occupe de l'analyse syntaxique et dirige tous les traitements.

L'insuffisance dont peut souffrir ce logiciel, est la capacité mémoire des machines très limitée.

En perspective, ce logiciel peut être connecté à un système acoustico-phonétique en mettant en correspondance les paramètres en commun. Il peut être sujet pour une situation au niveau des traitements, en utilisant une grammaire complexe. Comme il peut être directement connecté à un analyseur sémantico-pragmatique. Ce dernier est un maillon nécessaire dans un Système de Compréhension Automatique de la Parole et la gestion du dialogue homme - machine. Mais nous sommes encore très loin du rêve de la communication homme- machine en langage naturel. Pour augmenter les performances de notre système, nous avons besoin de la collaboration d'experts du domaine.

VI. PERSPECTIVES

Une étude plus détaillée des descriptions des sons de la parole, en collaboration avec un expert, permettrait d'élargir les performances du système. Un module d'analyse lexicale plus élaboré permettrait de prendre en compte un vocabulaire plus large et un langage moins contraint. Le système peut être étendu à un système de compréhension de la parole par l'implémentation de modules d'analyse syntaxique, sémantique et pragmatique.

VII. REFERENCES.

- [1] BONNET. A., "Application de l'intelligence artificielle" : les systèmes experts. Rairo Informatique, Vol. 15, N° 4, pp. 325-341, 1981.
- [2] BONNET. A., "L'intelligence artificielle, promesse et réalité", Inter-édition, 1984.
- [3] CORDIER M. O., "Les systèmes experts", La recherche, N° 151, pp. 60-70, Janv. 1984.
- [4] CALLAIRE H., "Les systèmes experts. Une percée de l'intelligence artificielle", La recherche, Vol. 13, N° 133, pp. 644-645, 1982.
- [5] NILSON N., "Problem-solving method in artificial intelligence", MC Graw-Hill Compagny, 1971.
- [6] NILSON N., "Principles of artificial intelligence", Troga Publishing Compagny, 1980.
- [7] FARRENY H., "Les systèmes-experts : Principes et exemples", Cepadués Editions, 1985.
- [8] PASTED, "Un système de démonstration automatique de théorèmes utilisant les connaissances et métaconnaissances en mathématiques", Thèse d'Etat Univ. de Paris VI, 1984.
- [9] RICH E., "Artificial intelligence", Mc Graw Hill, 1983
- [10] SEYDEN E., "systèmes-experts : Simuler l'intelligence", Sciences et techniques, N° 4, pp. 46-63, mai 1984.
- [11] MERCIER G., "Expert-systems, Approach to acoustic-phonetic decoding and word recognition", CNET, Centre Lannion A., pp. 201-221, 1990.
- [12] M^{me} DJOUADI K., "Etude des techniques d'intelligence artificielle : application à la réalisation d'un système expert de Reconnaissance de la Parole par les techniques d'intelligence artificielle", thèse de Magister en Electronique appliquée, Option : acquisition et traitement de l'information, ENP, Alger, 1996.
- [13] BOUCHEFFRA. K., "Contribution à la Reconnaissance Automatique de la Parole continue : Etude et réalisation d'un système de reconnaissance acoustico - phonétique", thèse de Magistère, ENP, Alger, 1995.
- [14] BONNEAU A., "Indices de reconnaissance des consonnes occlusives sourdes du français", Thèse de 3° cycle, Université d'Aix Marseille I, 1984.
- [15] CARBONNEL N., "APHODEX, Design and implementation of an acoustic phonetic decoding expert system", ICASSP, pp. 1201-1204, 1986.
- [16] STERNE P. E., ESKENASI M., MEMMI D., "An expert system for speech spectrogram reading", ICASSP, Tokyo, 1986.
- [17] MELONI H., GISPERT J., GUIZOL J., "An expert system for analytic word identification in continuous speech", 5^{ème} J.I. : Les systèmes experts et leurs applications, Avignon, mai 1985.
- [18] BOUSSEKSOU. B., "Reconnaissance Automatique de la Parole par les méthodes globales : Etude de certaines particularités linguistiques de l'Arabe Standard", thèse de Magistère, ILP, Alger, 1984.
- [19] BOUSSEKSOU. B., BOUCHEFFRA. K., "Contribution à la Reconnaissance Automatique de la Parole continue : Etude et réalisation d'un système de reconnaissance acoustico - phonétique", A.J.O.T. International Publication, Série A, Vol. 13, 1997, ENP, Alger.
- [20] MAHAMI. R., OUHAROU. M., "Composante lexico-graphique dans un Système Reconnaissance Automatique de la Parole continue", Projet de Fin d'Etude (PFE), Univ. Tizi-Ouzou, 1994.
- [21] FERRANE. I., "Base de données et de reconnaissances lexicales morphosyntaxique", Thèse de Doctorat de l'Univ. Paul Sabatier, 1991.
- [22] PERROUNOU. G., "Le projet BDEX de base de données et de connaissances lexicales et phonologiques", Paris, pp. 117-140, 1990.
- [23] HATON J. M., FFOHR D., "Paramétrisation acoustique et décodage phonétique fondé sur des connaissances pour la parole continue multilocuteurs", Acte du séminaire <<Gréco-PRC>> Communication Homme-Machine, 1988.
- [24] HATON J. P., PIERREL J. M., RERENOU G., CAELEN J., GAUVAIN JL., "Reconnaissance automatique de la parole", Edition Dunod Informatique, 1991, édition, 1984.
- [25] PIERREL. J.M., "Lexique et compréhension automatique de la parole", Lexique N° 8, Presse Universitaire de Lille, pp. 137-165, 1989.
- [26] CORBIN. D., "Morphologie dérivationnelle et structure du lexique", "Thèse de Doctorat d'état de l'Université de Paris VIII, 1986.
- [27] BESCHERELLE, "Le nouveau Bescherelle, l'art de conjuguer", Hatier, Paris, 1980.
- [28] CHOMSKY. N., "On certain formal properties of grammars" information and control, vol. 2, 1980.

MODELISATION DE LA COARTICULATION DES OCCLUSIVES SONORES

Mhania GUERTI

Laboratoire : Signal et Communications ENP Alger

E-mail : mhaniag@hotmail.com

ملخص : في تليغ الكلام، الأصوات تكون متتابعة حدودها غير واضحة لأنها ناتجة عن حركات متواصلة للجهاز الصوتي و ليس من مجموعة متعاقبة من أوضاع مختلفة. من بين المشاكل التي تجعل معالجة الكلام الإصطناعي حساسة تلك المتعلقة بالتغيير الكلي في الكلام الطبيعي الناتج عن حركة ثابتة. المفهوم الذي يدل على أن صوتا واحدا يمكن أن يكون من نفس مجموعة تشكلات مختلفة مهما جدا. التطبيق و الإستفادة من أي نموذج يشتمل على إختبار تفسير أو مقارنة تغييرات الخصوصيات الناتجة عن وحدات التنفيذ. من اجل تحقيق هذا الهدف قمنا بمعدونة $[V_1 C V_2]$ $[V_1 = a, i, u]$ و $[V_2 = \text{العشر مصوتات الفهوية و } C = [b, d, g]]$ قد سجل لكي نستنتج نماذج للحركة المشتركة. كتصويب الصوامت بالرنين ابتداء من المصوتات المجاورة. سنقدم قواعد نموذجية أستنتجناها إستفادة بالحاسوب و بإختبار متغير اللغة.

مفاتيح الكلام : تغيير إشارة الصوت، تمثيل الرنين، نموذج الحركة المشتركة.

ABSTRACT : In the speech spoken sounds are linked ones to others, their limits are fuzzy : the speech results of the continuous movement of the vocal apparatus, and not of following positions. Among problems that complicated the Automatic treatment of the speech, difficulties linked to the omnipresent variability in continuous speech. These result of the feat that the articulation is in constant evolution and depends even of the context in which sound is pronounced. The notion that some sound can characterised by set of different formants is important.

The stake to test then the exploitation of a model consist to test, to compare and to interpret variations of features owed to variations of command parameters. In this goal a corpus $[V_1 C V_2]$ ($V_1 = [a, i, u] \neq [V_2] \neq$ oral vowels et $C = [b, d, g]$) has been recorded in order to study the coarticulation modelling. We will infer consonantic targets in terms of resonances when leaving the adjacent vocalic targets like anchrage points follow-up of affiliation.

Our study is based on the resonances representation. Modelling rules will be given . These last are implemented and tested by a rule compiler.

KEY WORDS : Speech signal variability, Resonances representation , Coarticulation modelling.

RESUME

Dans le message parlé les sons sont enchaînés les uns aux autres, leurs limites sont floues : la parole résulte de mouvement continuel de l'appareil vocal, et non d'une suite de positions. Parmi les problèmes qui rendent le Traitement Automatique de la Parole délicat, les difficultés liées à la variabilité omniprésente en parole continue. Celles-ci résultent du fait que l'articulation est en constante évolution et dépend même du contexte dans lequel le son est prononcé. La notion qu'un même son peut être caractérisé par une même série de formants différents est importante.

La mise à l'épreuve puis l'exploitation d'un modèle va consister à tester, interpréter, comparer les variations des caractéristiques de sortie dues aux variations des paramètres de commande. Dans ce but un corpus $[V_1 C V_2]$ ($V_1 = [a, i, u]$ et $V_2 = 10$ voyelles orales et $C = [b, d, g]$) a été enregistré afin d'étudier la modélisation de la coarticulation. Nous inférerons des cibles consonantiques en termes de résonances en partant des cibles vocaliques adjacentes comme points d'anchrage d'un suivi d'affiliation.

Notre étude est basée sur la représentation en résonances. Des règles de modélisation seront données. Ces dernières sont implémentées et testées par un compilateur de règles.

MOTS CLES : Variabilité d'un signal vocal- Représentation en résonances- Modélisation de la coarticulation.

1. INTRODUCTION

Dans la théorie quantique de la parole, il y a une antinomie entre le discret et le continu et entre le simple et le complexe [1]. La production de la parole représente un procès phonatoire finalisé par la cible auditive et les autres cibles auditives sont choisies de façon à maximiser la stabilité acoustique par rapport à la variation articulaire [2]. La phonation et l'audition se trouvent liées par une boucle de réaction (feed-back). L'interaction entre les sons dépend de leur nature. Par exemple, les sons alvéolaires ont une influence énorme sur les voyelles nasales qui suivent. On peut l'expliquer par le mouvement rapide de la langue. Par contre, un son labial n'affecte que peu le son qui suit, car la langue bouge à peine. Il n'impose aucune contrainte sur les postures linguales.

La coarticulation consonantique est le résultat d'une négociation entre les termes de la tâche phonologique (qui impose soit une constriction dans un certain intervalle spatial ou la réalisation de cette constriction par un articulateur donné), et un lissage articulaire qui colore la consonne selon l'articulation vocalique adjacente. Une variation du lieu d'articulation et l'aperture labiale des consonnes, par exemple, se traduit par des loci dépendant du contexte.

Les indications que nous donnons pour le locus et les transitions formantiques s'appliquent bien à la parole synthétique et la stricte observance de ces lois suffit à obtenir une parole hautement intelligible. Des données spectrographiques suggèrent qu'il est possible d'estimer la fréquence de résonance de la cavité avant à partir d'une information sur le signal de la parole [3, 4]. Il semble que des contributions de F_2 , F_3 et des bursts des consonnes occlusives (figure 1) au lieu de la perception dynamique sont de 10 % par locus et 90 % par transition [5].

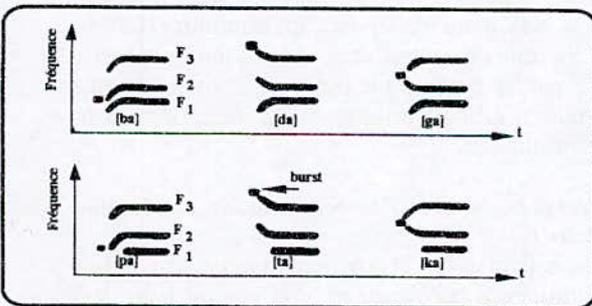


Figure 1 : Transitions formantiques et bursts pour les syllabes [CV] (C = occlusives ; V = a),

2. RAPPEL SUR LA THEORIE DU LOCUS

Les recherches entreprises sur le premier synthétiseur de parole : le pattern playback, un système permettant de transformer les spectrogrammes en sons, soit dans leur forme originale, soit dans une forme volontairement modifiée [6, 7]. Ce synthétiseur a été considéré à l'origine de la théorie du locus [8], qui a connu un grand succès en phonétique, et de la théorie motrice (tentative d'inférence du geste vocalique et du geste consonantique) de la perception selon laquelle le

message parlé est perçu par reconnaissance des mouvements articulaires.

Les chercheurs des Laboratoires Haskins ont révélé que les indices essentiels utilisés dans la perception de la plupart des consonnes sont les modifications que subissent les fréquences des formants lors des transitions [CV] ou [VC].

On donne le nom de locus au point de convergence de toutes les transitions de formants qui sont reliées à la perception d'un même lieu d'articulation consonantique ou de tout autre trait pertinent, quelle que soit la voyelle qui précède ou qui suit (figure 1). Ce lieu est donc d'après Delattre et al., un point virtuel de la trajectoire formantique (puisque masqué par l'occlusion). Le locus était déterminé par extrapolation d'environ 50 ms à partir de l'observation des transitions formantiques pour une consonne donnée avant différentes voyelles.

Delattre et al. [8] attestent : il y a pour chaque consonne, des positions fréquentielles caractéristiques, ou des loci dans lesquels les transitions formantiques peuvent être vues simplement comme des mouvements de formants à partir de leurs loci respectifs aux niveaux fréquentiels appropriés pour le phonème suivant, partout où ces niveaux pouvaient être.

L'hypothèse de départ est la suivante : étant donné que le lieu d'articulation de chaque consonne est la plupart du temps fixe, nous devrions nous attendre à trouver qu'il existe en rapport à ceci, une position de fréquence fixe ou -locus- pour son F_2 . Nous pourrions alors décrire de façon relativement simple les diverses transitions du F_2 comme des mouvements depuis ce locus acoustique jusqu'à son niveau stable dans la voyelle, et ceci quelle que soit cette voyelle.

La connaissance des loci pour les trois premiers formants de toutes les consonnes va permettre de construire des règles au niveau phonémique. Les consonnes ayant le même point d'articulation présentent la même transition de F_2 , et ce sont alors les informations sur le spectre basse fréquence qui distinguera les divers modes d'articulation.

On voit alors qu'à l'aide de 6 règles, une pour chaque point d'articulation et une pour chaque mode, on peut définir 9 phonèmes (figure 2).

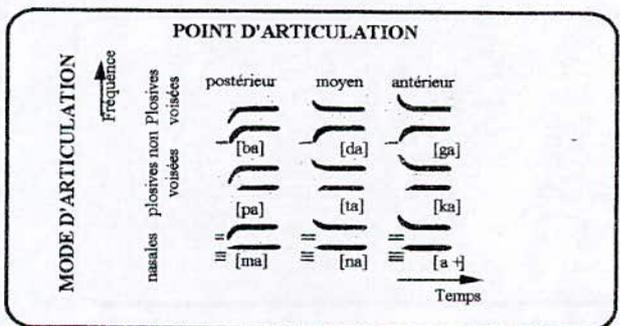


Figure 2 : Relations entre l'articulation et les transitions formantiques des consonnes occlusives et nasales.

Des résultats bien connus ont montré que le F_2 contribue hautement à la perception des lieux

d'articulation des consonnes occlusives dans les syllabes [CV] [8]. Néanmoins dans une parole réelle voisée, pour la voyelle [i] la région de F_3 peut être approximativement associée à la cavité avant plutôt qu'avec le F_2 . Pour les voyelles [a, u] dans une articulation hautement constrictive, la région de F_2 est prédite pour être approximativement associée à la cavité avant plutôt qu'avec celle de F_3 .

3. VARIABILITE DU SIGNAL. DE PAROLE

Dans le système de la communication parlée, le processus de l'encodage-décodage passe par une négociation production-audition ou une divergence-convergence :

- ♦ divergence due à la variabilité des comportements articulatoires intra-individuels et aux différentes stratégies contextuelles du locuteur (les phénomènes de la coarticulation) ;

- ♦ convergence bien obtenue, puisque la variabilité est finalement réduite pour que la communication soit assurée par l'auditeur.

Nous pouvons distinguer trois sources de variabilité liées à des différences physiologiques entre les locuteurs, aux latitudes variables de réalisation au plan linguistique et aux effets de la coarticulation.

3.1. Variabilité d'origine physiologique

Le conduit vocal d'une femme est en moyenne de 15 % plus court que celui d'un homme. En première approximation, la théorie acoustique indique que l'augmentation des fréquences de résonance est proportionnelle à la diminution de la longueur. Pour une forme de conduit vocal ayant subi une déformation uniforme, les formants sont par conséquent d'environ 15% plus élevés. Cette hypothèse de déformation uniforme est infirmée par l'observation sur des sujets réels : un adulte et un enfant (figure 3). Nous remarquons sur cette figure que les cavités avant ont presque la même taille et la même forme pour chaque voyelle prononcée par les deux sujets. Par contre les cavités arrières diffèrent complètement, car elles subissent des distorsions qui sont dues aux diminutions de la longueur disponible.

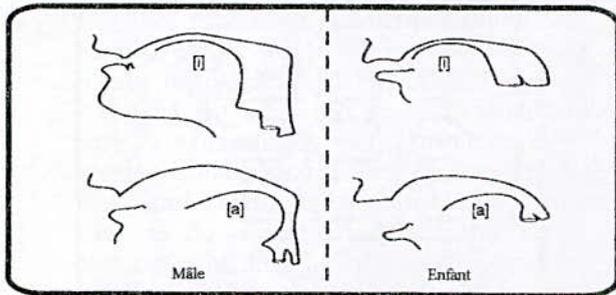


Figure 3 : Coupes sagittales de conduits vocaux des voyelles [i] et [a] réalisées par un adulte et un enfant.

3.2. Variabilité d'origine linguistique

Pour un même locuteur prononçant plusieurs fois un

même son, la dispersion des formants est déjà importante, car l'articulation n'est jamais répétée de façon identique. Pour plusieurs locuteurs provenant d'une même région linguistique, la dispersion augmente. Plusieurs auditeurs peuvent percevoir différemment un même son à l'état isolé (statique), prononcé par un même locuteur. Un même auditeur peut percevoir comme étant différents deux sons physiquement identiques (ceci dépend du système phonologique). Donc le sexe, l'âge, le contexte, les caractéristiques du conduit vocal, les origines du locuteur et de l'auditeur, etc. jouent un rôle très important. En Français standard, par exemple la distinction entre les voyelles nasales [ɛ̃, œ̃] n'est pas faite par un nombre important de locuteurs.

3.3. Les effets de la coarticulation

3.3.1. Qu'est-ce-qu'une coarticulation ?

La coarticulation peut être définie dans le sens large comme l'influence du contexte phonétique sur un segment déterminé. Par conséquent elle est plus marquée dans le discours, car le son prononcé isolément n'a pas de contexte qui peut l'influencer. On a généralement estimé que les causes de la coarticulation sont à chercher dans l'inertie des muscles et des organes articulatoires [9]. Autrefois, la coarticulation était vue comme la conséquence d'une limitation d'ordre mécanique de l'appareil phonatoire. Elle est à présent considérée comme le reflet de l'organisation en unités de programmation dans la production du parlé. Elle reste un problème fondamental de la recherche phonétique. En se groupant les sons s'influencent mutuellement et se modifient de diverses façons. Les consonnes sont soumises à l'influence acoustique des voyelles et les spectres vocaliques sont modifiés au contact des consonnes.

En prononçant les sons du langage, l'homme a tendance à obtenir le maximum d'effet avec un minimum d'effort. Une description de trajectoires acoustiques peut être simplifiée par le choix d'un espace qui interface une transformation articulatoire-acoustique avec des non-linéarités minimales.

3.3.2. Quelques modèles de coarticulation sur le plan articulatoire

Sur le plan articulatoire, la coarticulation peut se définir comme l'influence qu'exerce un son sur un autre son contigu. Cette modification contextuelle de l'articulation résulte sans doute d'un effet d'inertie mécanique mais également d'une réorganisation du geste articulatoire, liée au concept de minimisation de l'effort articulatoire. Par exemple, pour articuler le mot [dut], le locuteur part d'une articulation antérieure apico-dentale, recule rapidement la langue vers la zone vélaire, puis doit ramener la langue en position antérieure (figure 5). En débit rapide, le locuteur dispose de peu de temps pour accomplir ce déplacement important. La vitesse de déplacement de la langue doit être grande. Si le locuteur ne veut (ou ne peut) accroître l'énergie articulatoire, donc la vitesse, il lui reste la

possibilité de réduire la distance à parcourir en articulant un [u] moins extrême, plus antérieur (figure 4). Ce phénomène est appelé en Anglais un

undershooting [10]. La limite de ce phénomène est sans doute d'origine perceptive [11].

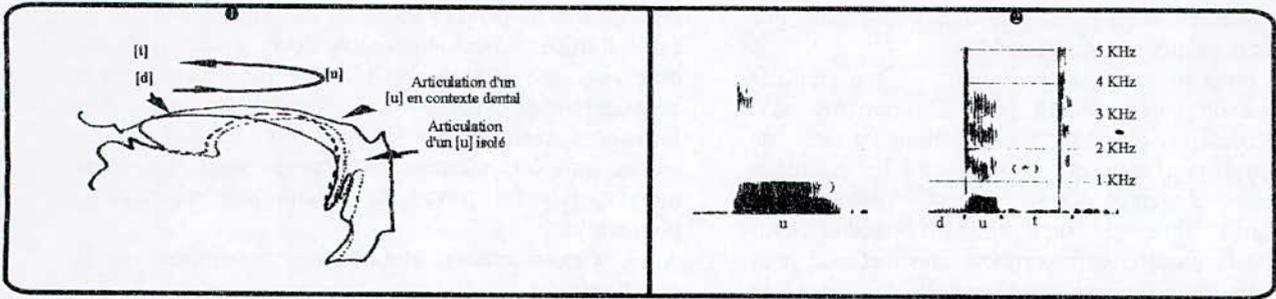


Fig. 4 ● et : Articulation d'un [u] hors contexte et en contexte dental, dans le mot [dut] et spectrogrammes de la voyelle [u] hors contexte et en contexte dental. (★) représente la fréquence la plus élevée de F₂ en contexte dental, due à la coarticulation vers l'avant

En étudiant la coarticulation en Suédois et en Anglais américain, Öhman élabore un modèle mathématique pour rendre compte des mouvements articulatoires en coarticulation [12]. Il propose une équation décrivant les traits essentiels de la coarticulation dans les occurrences [V₁CV₂]. Dans son modèle, basé sur des relevés de films cinéradiologiques, les voyelles et les consonnes sont produites de façon différente.

L'auteur suggère de décrire les mouvements de la langue par trois systèmes mécaniques contrôlables indépendamment et correspondant à des ensembles de muscles pouvant être partiellement identiques ou partiellement différents. Il attribue les noms suivants à ces trois systèmes articulatoires : apical, dorsal et le corps de la langue. Les deux premiers articulatoires permettent la production des phénomènes consonantiques, alors que le dernier est réservé aux phénomènes vocaliques. Pour modéliser les voyelles, il utilise les positions extrêmes du conduit vocal [i], [a], [u] ; par la suite toutes les voyelles seront représentées par des combinaisons linéaires de ces trois positions. L'équation proposée est la suivante :

$$S(x, t) = V(x, t) + k(t) [C(x) - V(x, t)] W C(x) \quad (1)$$

- ♦ x : distance longitudinale entre les lèvres et la glotte ;
- ♦ $S(x, t)$: définit la forme du conduit vocal à l'instant t durant l'occurrence VCV ;
- ♦ $V(x, t)$: composante vocalique qui est une combinaison linéaire des 3 positions extrêmes des voyelles cardinales, avec les coefficients variables dans le temps,
- ♦ $C(x)$: représente la cible idéale de la consonne ;
- ♦ $W C(x)$: fonction de coarticulation ;
- ♦ $k(t)$: facteur qui varie dans le temps. Il représente le degré d'écart entre la cible idéale et la réalisation concrète sur le plan articulatoire.

Dans une analyse spectrographique des séquences [V₁CV₂] (C est une occlusive sonore) Öhman a pu montrer que [V₁] subit l'effet de la coarticulation de

[V₂] à travers la consonne intervocalique. La deuxième voyelle était en quelque sorte programmée pendant la réalisation de la première voyelle. Il suggère que les mouvements linguaux sont la base pour produire les voyelles et que les mouvements nécessaires pour produire les consonnes sont superposés au substrat vocalique avec lequel ils coarticulent d'une manière plus ou moins claire. Le mouvement pour produire les voyelles est constant et interrompu et le mouvement consonantique est simplement superposé au mouvement vocalique qui ne s'arrête jamais. Autrement dit, le geste vocalique est lent et le geste consonantique est rapide. Ceci correspond à deux canaux différents : voyelles, canaux en raideur et les consonnes, canaux en force.

Les conclusions tirées semblent pour l'essentiel, confirmées par l'étude de Mac Neilage et de Declerck qui constatent que le système articulatoire ne travaille pas à partir d'instructions données pour chaque phonème, mais plutôt à partir d'instructions pour la syllabe entière [13]. Leurs résultats confirment le point de vue que la syllabe [CV] est le seul type universellement attesté et que les effets de la coarticulation fournissent à l'auditeur des indices qui facilitent l'identification des segments.

En Français la syllabe [CV] est le cadre de la coarticulation. Cette dernière est fréquente à travers la consonne alvéolaire dans les syllabes [VtV], et beaucoup moins fréquente à travers une vélaire dans les [VkV]. Aussi la voyelle [i] est fréquemment source d'une coarticulation dans les deux sens (de gauche à droite et inversement). La conclusion la plus intéressante est que le degré de la coarticulation semble varier selon le lieu d'articulation. Etant donné que jusqu'à présent on n'est pas arrivé à cerner avec précision les limites de la coarticulation, on admet tout de même que son rôle est celui de garantir le caractère naturel du langage et que les transitions articulatoires d'une cible à l'autre se réalisent avec la plus grande facilité possible, sans heurts ni sauts abrupts.

La zone du conduit vocal où se situe la constriction reste relativement invariante tandis que les autres zones sont susceptibles de varier largement en fonction du

contexte [14]. Ces dernières assimilent les formes des voyelles adjacentes. Les phénomènes de la coarticulation sont donc en général fonction des différentes aires constitutives du conduit vocal. Certaines de ces aires sont très affectées tandis que d'autres sont relativement insensibles.

Öhman propose des améliorations à son modèle numérique de coarticulation [12]. Il confirme qu'il serait nécessaire d'établir les fonctions d'aires du conduit vocal pour que l'on puisse obtenir les équations acoustiques correspondant aux mouvements articulatoires. Il ajoute qu'il faudrait pouvoir tenir compte de la nasalité et incorporer une méthode pour contrôler les types de consonnes (sourdes ou sonores) et localiser la source. Les inconvénients majeurs de ce modèle demeurent dans son impuissance à générer des séquences Consonnes Consonnes et surtout à générer des occlusives faisant intervenir l'articulateur dorsal. En effet, pour celles-ci, il paraît difficile de localiser le lieu d'articulation. Selon les entourages vocaliques de ces consonnes, les lieux d'occlusion sont très différents et en tout cas trop éloignés par les écarts prévus par le modèle. Après une période d'enthousiasme, on s'est donc rendu compte que les trois loci étaient insuffisants pour caractériser correctement les consonnes [k] et [g]. Delattre a défini un certain nombre d'indices acoustiques des consonnes françaises, qu'il a classés en indices de lieu¹ (loci de F₂ et de F₃, fréquence de bruit pour les fricatives) et de mode d'articulation (forme et rapidité des transitions formantiques, locus de F₁, présence ou absence de bruit de friction, etc) [6]. Depuis on a découvert d'autres indices comme le VOT (Voice Onset Time : délai d'établissement de voisement). Ainsi une consonne est caractérisée par l'ensemble des valeurs que prennent ces indices au cours d'une transition, de même qu'une voyelle est caractérisée par les fréquences formantiques F₁, F₂ et F₃. Comme il y a une grande variabilité des transitions, Öhman fut amené à mettre en doute la théorie du locus. Ses remarques critiques donnent l'occasion à Delattre de souligner que le concept de locus était un concept perceptuel et ne pouvait être affecté par la coarticulation [15].

En réalité la théorie du locus souffre de très nombreuses exceptions. Nous pouvons dire, en première approximation que ces variations dans une transition [CV] ou [VC] dépendent de la consonne et de la voyelle. Toutefois, en parole spontanée, le contexte phonémique qui conditionne la transition est plus étendu [16]. Cette théorie suppose que la consonne est issue d'une position articulatoire idéale, bien fixe quel que soit l'entourage vocalique. Cette conception s'est avérée très riche mais trop simplificatrice. On a dû admettre que le locus de F₂ pour les vélares est largement variable, et se situe à environ 3 kHz pour les voyelles antérieures et écartées, et à environ 1000-1500 Hz pour les voyelles postérieures et arrondies. Ceci

signifie que le point d'articulation peut être différent pour une même consonne, et en particulier lorsque l'articulation est palatale : il est décalé vers l'avant pour les voyelles antérieures et vers l'arrière pour les postérieures (dépend de l'articulateur mis en jeu).

Dans l'étude radiocinématographique la coarticulation observée sur le plan articulatoire ne se reflète pas nécessairement dans les spectrogrammes. Les influences exercées par les phonèmes les uns sur les autres sont certainement explicables, dans une large mesure, par les propriétés dynamiques du système phonatoire.

3.3.3. Conséquences acoustiques maximales de la coarticulation

Les recherches menées dans les autres disciplines : production, perception, phonétique, etc. montrent sans ambiguïté que la parole est spécifique, donc adaptée à la communication orale. En conséquence de la continuité de la parole, l'influence d'un phonème s'étend plus ou moins sur ses voisins, phénomènes dits de l'assimilation (anticipation, rétention), de la coarticulation, de l'harmonie vocalique (assimilation à distance entre voyelles). Ces phénomènes limitent les performances des synthétiseurs et des reconnaisseurs actuels.

Le système consonantique français est très riche en phénomènes d'assimilation. C'est surtout au point de vue de la sonorité que les consonnes agissent les unes sur les autres quand elles sont combinées dans le mot ou dans la phrase. Les liquides s'assourdissent plus ou moins au contact de l'occlusive sourde précédente.

Les principaux cas d'assimilation de voisement sont les suivants :

- ♦ le dévoisement d'une voyelle haute antérieure, si elle est entourée de fricatives sourdes, exemple : [ĕstity] ;
- ♦ la voyelle peut être complètement dévoisée dans un entourage de consonnes dévoisées. On peut trouver un [a] dévoisé dans [swasât] comme une conséquence du dévoisement de [w] ;
- ♦ une consonne occlusive ou fricative sourde suivie immédiatement d'une occlusive ou fricative sonore se sonorise. Parallèlement, une sonore suivie d'une sourde s'assourdit

[mōdōsĕ] → [mōtōsĕ] ; [anekdote] → [anōgdōt] ;

- ♦ lorsqu'une consonne sonante [R, l, m, n, ŋ, j, w, Ô] intrinsèquement sonore précède ou suit immédiatement une consonne sourde, elle a tendance à se dévoiser totalement, ou partiellement dans sa partie la plus proche de la sourde, exemple [pRi] ;

Les principaux cas d'assimilation dûs aux mouvements d'anticipation des articulateurs de configuration du conduit vocal sont les suivants :

- ♦ une occlusive se nasalise entre une voyelle nasale (contexte gauche) et une consonne fricative (contexte droit)

[p, b] > [m] [bōbpōisât] [bōmpōisât] (bombe puissante) ;

[t, d] > [n] [pâtōkō:t] [pānkō:t] (pentecôte) ;

[k, g] > [ŋ] [lōgōmā] lōŋmā (longuement)

- ♦ devant [i], plus rarement [y] et [e], les occlusives [k] et [g] peuvent être palatalisées (réalisation beaucoup plus antérieure).

¹ En Français l'articulation de la consonne est toujours plus en avant que celle de la voyelle.

La théorie du locus a permis de simplifier et de systématiser la recherche d'éléments acoustiques pertinents dans les premiers essais de parole synthétique. Mais dès lors qu'elle était présentée comme une véritable théorie d'une perception consonantique idéale elle se heurtait à des controverses parfaitement justifiées, en particulier par Klatt pour la théorie du locus modifié [7]. Ce dernier a trouvé que cette conception du locus est très riche mais trop simplificatrice, car elle s'applique au mieux à F₂. Klatt avait des difficultés pour synthétiser les occlusives anglaises. Il a essayé de déterminer si un concept de locus modifié pouvait être créé ou si une liste est nécessaire pour tabuler les débuts des fréquences pour les trois premiers formants avant chaque voyelle. Quand tous les points des données restent sur une ligne droite, c'est-à-dire on peut prédire le commencement de la fréquence F₂ voyelle par une équation de la forme :

$$F2_{onset} = F2_{locus} + k (F2_{voyelle} - F2_{locus}) \quad (2)$$

k : degré de coarticulation

Plusieurs facteurs complexes causent l'échec de la théorie du locus. Une transition peut avoir à la fois une composante rapide et lente, due au relâchement rapide de l'obstruction suivie par des mouvements graduels du corps de la langue ; une voyelle qui suit peut influencer le F_{2onset} d'une transition [CV]. Le F₂ peut être relativement insensible aux constriction orales quand il est affecté à une résonance de la cavité avant comme le cas de [i]. Klatt a fait l'hypothèse sur les influences avant et arrière du corps de la langue et de l'arrondissement des lèvres. Il a divisé l'ensemble des voyelles en [+avant, +arrondies] et [-avant, -arrondies] et trouva à l'intérieur de chaque ensemble que les données sont suffisamment régulières pour être approximées par des lignes droites (figure 5●)

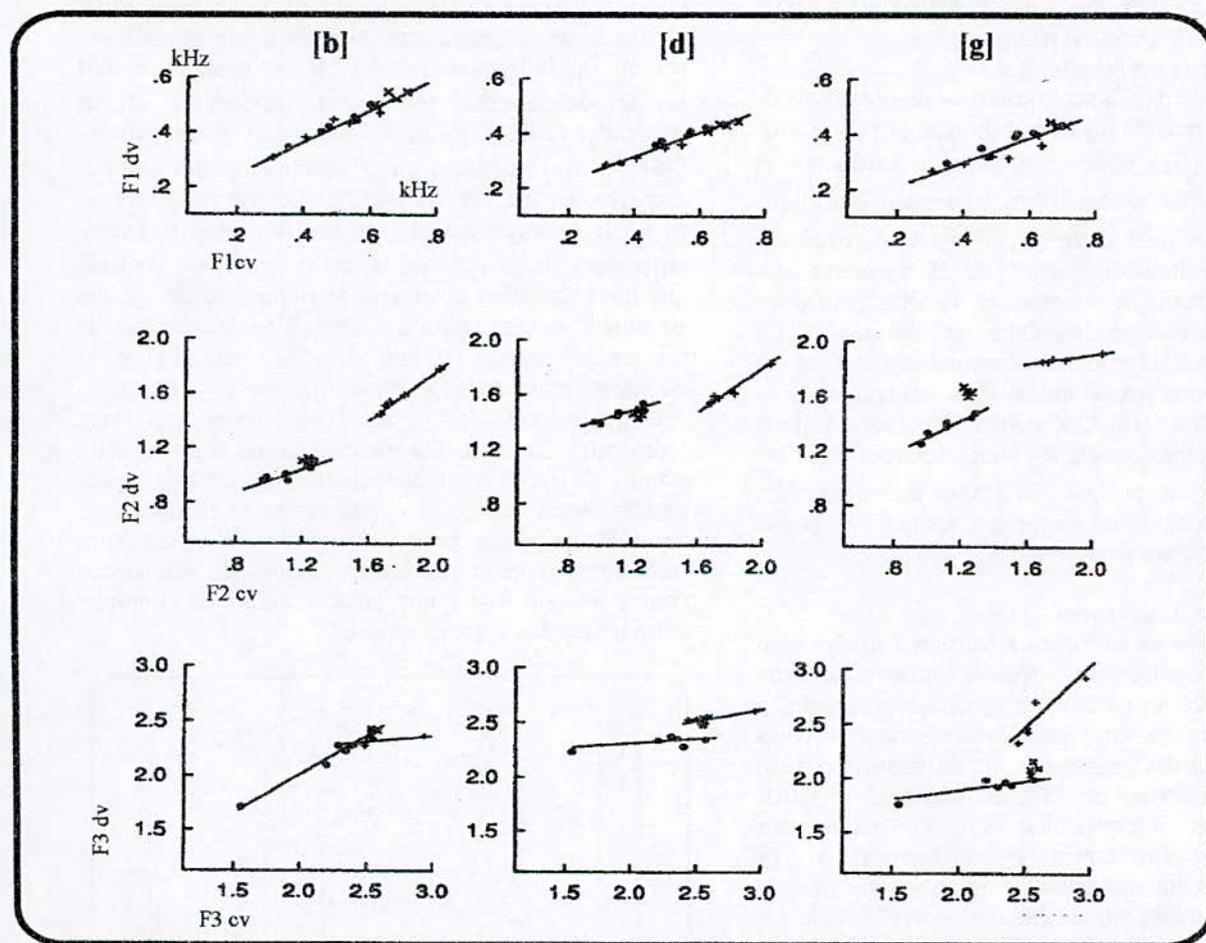


Figure 5 ● : Fréquences du début de la transition de F₁, F₂, F₃ en fonction de celles de la voyelle cible, pour les 16 voyelles américaines après les consonnes [b, d, g]

Nous avons superposé les tracés des fréquences F₂, F₃ concernant les [gV] en les remettant à la même échelle (fig. 5 ●). Nous remarquons que les portions de droites correspondant aux voyelles intermédiaires se

superposent parfaitement (fig. 5 ●). Pour les voyelles basses, nous observons que R₂ = F₂ alors que pour les voyelles hautes R₂ = F₃. La courbe de coarticulation R₂([g]) en fonction de R₂ de la voyelle adjacente ne

donne alors pas de saturation. Nous pouvons faire de même pour [d].

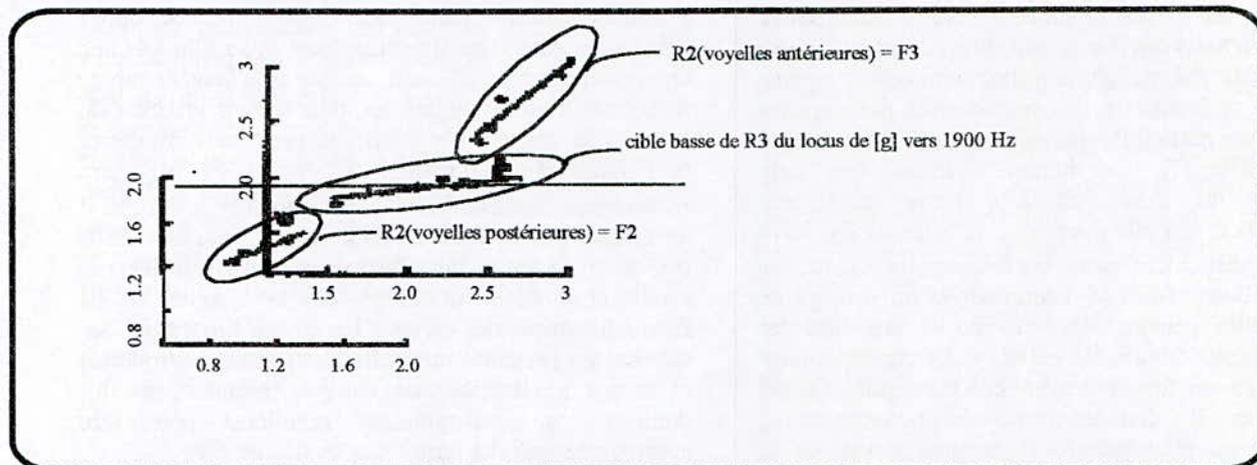


Figure 5 : Réinterprétation des schémas de coarticulation $F_2([g])$ et $F_3([g])$ en termes de résonances.

4. MODELISATION DE LA COARTICULATION DES OCCLUSIVES SONORES

4.1. Motivations et méthodologie

Dans l'objectif de modéliser la coarticulation nous avons constitué un corpus de logatomes du type $[V_1CV_2]$ avec $C = [b, d, g]$, $[V_1]$ = les trois voyelles cardinales et $[V_2] \neq 10$ voyelles orales. L'environnement vocalique est choisi parce qu'il donne à la fois l'information : fermeture et relaxation autour de la consonne. La structure théorique de résonances va être confirmée avec des données expérimentales de ces stimuli. Le nombre total des différentes combinaisons est de 90 réalisations. Nous avons utilisé deux techniques : la technique cepstrale et la LPC à Pitch Synchrone à glotte fermée. Les fréquences de R_2 sont mesurées dans les parties stables de chaque voyelle et au milieu des consonnes occlusives, de manière à assurer une bonne continuité des trajectoires.

4.2. Mesures en résonances

La première étape de ce travail a consisté à analyser un ensemble d'enregistrements effectués sur deux locuteurs afin d'en extraire les paramètres spectraux pertinents, à savoir la valeur des cinq premières résonances. Nous avons procédé à des regroupements de manière à nous trouver en présence de classes limitées, à partir desquelles une interprétation sur la nature des consonnes peut être tentée. Les fréquences de ces résonances ont été mesurées en un point médian de l'état stable : la fréquence cible de la voyelle. Pour les transitions nous avons relevé la différence de fréquence par rapport à la partie stable de la voyelle

4.3. Etude des trois occlusives sonores

Les consonnes occlusives fournissent un type de constriction extrême que peut produire une cavité avant bien définie. Pour étudier les consonnes, Stevens considère un modèle idéalisé constitué de deux résonateurs, séparés par une constriction : un postérieur

(fermé) et l'autre antérieur (ouvert) [1]. Il suppose qu'il n'y a pas de couplage entre eux. Si la constriction est très étroite, la fréquence de R_1 est très basse et ce sont les fréquences des résonances supérieures et en particulier celles de R_2 et de R_3 qui sont déterminantes. Celles-ci correspondent aux fréquences propres de ces deux résonateurs. Or, lorsque l'on fait varier le rapport de l_p / l_a (respectivement longueur de la cavité arrière ou postérieure et celle de la cavité avant) on constate que les résonances changent de cavités en des points singuliers correspondant à l'intersection des courbes de fréquences propres (figure 6). Sur cette figure la longueur totale du tube est de 16 cm et celle de la constriction est de 3 cm. Les lignes en tirets représentent les deux plus basses résonances de la cavité avant ; les lignes continues représentent les quatre plus basses résonances de la cavité arrière. Les lignes en pointillés près des points de convergence des deux résonances représentent les fréquences de résonances pour le cas où il y a une faible quantité de couplage entre les cavités avant et arrière.

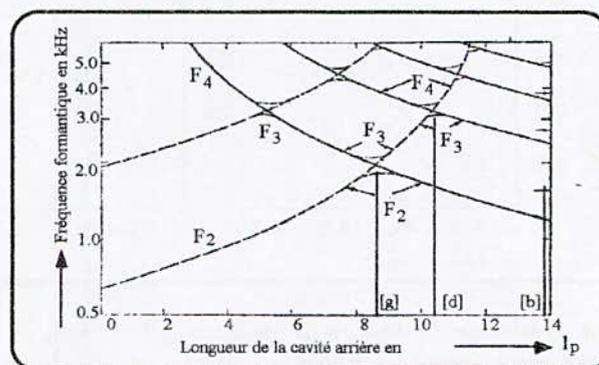


Figure 6 : Relations entre les fréquences naturelles et la position de la constriction.

4.3.1. L'occlusive labiale [b]

La réalisation d'une labiale consiste en une fermeture

des lèvres. L'ouverture de sortie du résonateur buccal devient plus petite ou même supprimée totalement, il est évident que sa fréquence de résonance est plus basse que quand il est ouvert². Mais pour prononcer n'importe quelle labiale suivie d'une voyelle, on n'adopte aucune articulation que celle de la voyelle, par conséquent les transitions R_2 - R_3 sont toujours montantes à des degrés divers³. Si nous avons une [VC] (C = labiale), c'est l'inverse qui se produit, les transitions des labiales étant réversibles. Cette montée universelle de R_2 est faible pour les voyelles graves à médianes, devient nettement plus accentuée avec les voyelles aiguës. Dans les expériences de Delattre et al. [8], le locus de [b] a une fréquence plus basse que le R_2 le plus bas de toutes les qualités vocaliques utilisées (soit 700 Hz). Quant aux transitions résonantiques de R_1 , elles dépendent surtout de la nature sourde ou sonore de ces consonnes. Compte tenu de ces particularités de détail, elles sont dans un [CV] toujours montantes et dans un [VC] descendantes, quelle que soit la voyelle qui précède ou qui suit (du fait de l'occlusion partielle ou complète de tout le conduit vocal).

4.3.2. L'occlusive apico-dentale [d]

Lors de l'explosion de [d] dans les séquences [dV], le résonateur buccal est hors service du fait de l'articulation dentale. Cette dernière est très largement indépendante de celle des voyelles qui suivent, le résonateur buccal doit donc nécessairement, avant d'être à même de produire ces voyelles, passer momentanément par des configurations géométriques différentes. Ce qui explique aisément cet aspect de divergence des transitions de R_2 à partir d'un point unique : locus, qui est une notion bien réelle, du fait qu'il possède une véritable corrélation articuloacoustique. Dans ces cas les transitions de :

- ♦ R_1 sont montantes à degrés divers ;
- ♦ R_2 sont descendantes en contexte postérieur et montantes en contexte antérieur ;
- ♦ R_3 sont toutes descendantes sauf pour le [i].

Donc, avec toutes les voyelles, même si les transitions résonantiques ne correspondent pas absolument avec les loci de R_2 - R_3 , les transitions de ces résonances semblent toujours provenir de ces valeurs de base.

La lecture de sonagrammes peut faire naître des confusions avec des [dV], [V] = voyelles aiguës, la différenciation de lecture se faisant alors surtout pour la transition de R_3 à montée très abrupte après une labiale.

L'anticipation prononcée avec les voyelles médianes, limite les pentes des transitions de R_2 . Elle devient nettement faible avec les voyelles aiguës, ce qui amplifie les pentes de R_2 , les rendant généralement plus

abruptes pour [bi] que pour [di], le tout résultant en un semblant de divergences des R_2 de toutes les voyelles depuis un locus commun, au moins dans des [bV] artificiels soumis au choix forcé où une pente R_2 plus accentuée suffit à différencier auditivement [di] de [bi]. Cependant, dans des [CV] naturels présentant une assez forte anticipation vocalique, il y a souvent si peu de différence entre les transitions de R_2 de [di] et de [bi], que la différenciation acoustique ne se fait plus, pratiquement que par la nature du bruit d'explosion.

La transition de F_3 est importante pour le [d] dans [di] mais non pas pour le [d] dans [du]. Ce résultat peut être dû au fait que la résonance fondamentale de la cavité avant est fortement associée à F_3 de [i] et à F_2 du [u]. Donc la résonance de la cavité avant peut jouer un rôle dans la perception de [d].

4.3.3. L'occlusive vélaire [g]

Pendant l'élocution d'une vélaire suivie d'une voyelle, l'ouverture de sortie de la cavité buccale ne joue aucun rôle puisqu'elle reste absolument inchangée. Cependant pour l'élocution de la consonne, le dos de la langue est accolé au palais et ce, dans des positions très diverses suivant la voyelle qui suit. Mais quelle que soit l'articulation, quand on passe de la consonne à la voyelle, le dos de la langue se décolle du palais, ce qui produit nécessairement un agrandissement du volume du résonateur buccal. Naturellement, il s'ensuit un abaissement de F_2 en cours de transition, et ce dans tous les cas, quelle que soit la voyelle attenante (sauf pour [i, e] et [y]). Le point d'articulation (point de contact du dos de la langue avec le palais) change selon la voyelle qui suit à cause du phénomène de la coarticulation. Par contre le volume de cette cavité avant, qui est au moment de la réalisation de la consonne du fait de l'occlusion palatale, forcément plus petit que lors de l'élocution de la voyelle suivante et vice-versa, dans ce cas la transition de R_2 est toujours descendante à des degrés divers (exception faite pour les voyelles antérieures). La pente des transitions de R_2 suit grossièrement le degré d'ouverture des voyelles adjacentes. Elle est forte avec les voyelles ouvertes : grande augmentation de volume du résonateur buccal en passant de la consonne à la voyelle et devient progressivement plus faible au fur et à mesure qu'on a affaire à des voyelles plus fermées aiguës ou graves (figure 7). Autrement dit, cette pente est peu prononcée pour [gi] et [gu], et nettement plus prononcées pour [ga], par exemple.

Avec toutes les qualités intermédiaires entre [a] et [o], le [g] est toujours nettement perçu ce qui infirme la validité d'un locus double pour les vélares. Les pentes des transitions de R_2 suivent le degré d'ouverture des voyelles attenantes. Nous savons que lors de la prononciation d'une vélaire la taille de la cavité avant varie beaucoup en fonction de la forte coarticulation linguale. Ce phonème devient palato-vélaire quand il est associé à des voyelles ou des consonnes à

² La proportionnalité de cette montée dépend en grande partie de l'anticipation vocalique (mise en position des organes comme pour prononcer la voyelle qui suit).

articulation antérieure ou bien vélaire dans les autres contextes (figure 7). Ceci provoque d'importantes variations dans les fréquences de résonances de R_2 et de R_3 . Par exemple, dans les séquences [aga] et [agØ], le R_2 passe d'environ 1250 Hz pour le [a] à 1600 Hz pour le [Ø]. Les deux voyelles ne sont pas antérieures, donc il n'y a pas de changement d'affiliation entre les cavités lors de ce passage. Par contre pour [iga], le R_2 de [i] passe d'environ 3000 Hz (voire 3500 Hz) à 1250 Hz et pour [igØ] à 1600 Hz. Ceci explique la forte descente de R_2 et le changement d'affiliation que nous remarquons, le F_3 de [i] devient égal à R_2 (F_2) de [a] ou bien de [Ø]. La figure 6, illustre bien ce changement d'affiliation. Pour $l_p < 8.5$ cm, le F_2 est produit par le résonateur antérieur et le F_3 par le résonateur postérieur. A la traversée de $l_p = 8.5$ cm, il y a échange, F_2 devient produit par le résonateur postérieur et F_3 par le résonateur antérieur. Si l'on élargit alors la constriction (relâchement de l'occlusion), F_2 et F_3 se séparent. Leurs transitions résonantiques T_2 et T_3 sont convergentes ensuite elles divergent ($T_2=T_3$ après $T_2 \neq$

T_3). Ce fait caractérise bien sûr les vélaire. Notons que si les transitions des résonances R_2-R_3 ont une grande valeur perceptive de reconnaissance avec les voyelles médianes (où ces transitions sont accentuées), cette valeur devient beaucoup plus faible avec les voyelles aiguës et surtout graves, où la concentration spectrale de l'explosion joue un rôle acoustique prépondérant. Le fait de supprimer F_3 dans des [gV] oblige à relever la transition de F_2 en lui donnant une pente très supérieure à sa pente réelle. Ceci infirme la valeur perceptuelle des directions des transitions telles que celles pilotées par des loci théoriques. Delattre et al. [8] cherchent à appliquer aux vélaire la théorie perceptuelle des loci. Cependant ils emploient une grande restriction, en imaginant deux loci virtuels distincts, l'un dont semblent diverger les F_2 des voyelles graves, placé vers 1 kHz, et l'autre dont semblent diverger les F_2 des voyelles aiguës à médianes, placé à 3 kHz. La seule catégorisation possible de [g] est obtenue avec un F_2 droit à 3 kHz (une sorte de [i] très fermé confinant à une qualité semi-consonantique de [j]).

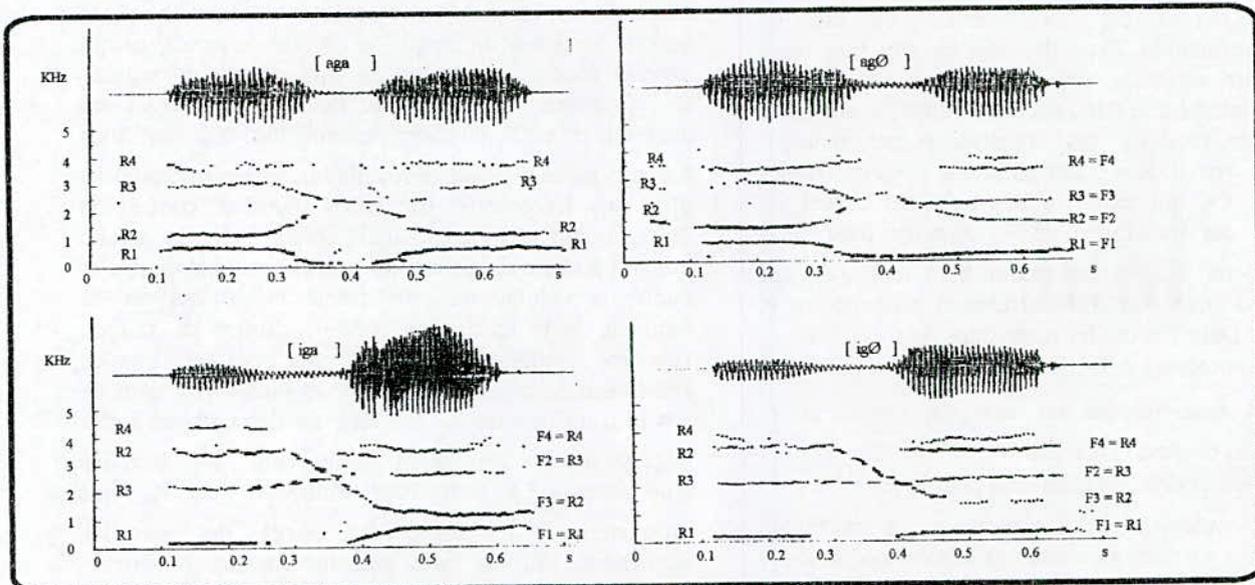


Figure 8 : Suivi de résonances dans les séquences [agV] et [igV] [V = a, Ø]

4.4. Equations des loci dans les séquences $[V_1CV_2]$

Conformément aux résultats de Delattre et al. [8], nous avons trouvé dans le cas du [g], deux loci : un à une fréquence basse d'environ 1200 Hz pour des contextes impliquant des voyelles postérieures et un deuxième locus haut aux alentours de 3500 Hz pour les voyelles Antérieures [17, 18, 19]. Ces résultats indiquent clairement qu'en Français la voyelle subséquente $[V_2]$ a une influence sur les transitions de la voyelle précédente $[V_1]$. La voyelle [a] a presque toujours des transitions plus longues que celles de [i]. Sa variation de fréquence est plus importante au contact des consonnes vélaire qu'au contact des labiales ou alvéodentales. Des

trajectoires de résonance utilisent un critère de continuité et un son-une trajectoire. Des résonances sont définies pour chaque phonème, par un ensemble de trajectoires connectées par des fonctions d'interpolation: X_1, X_2D, X_2G, X_3 . Basées sur cette représentation, des règles sont données, pour décrire une coarticulation des trois occlusives sonores dans des contextes $[V_1CV_2]$. En utilisant les données acoustiques comme un point de départ, nous présentons les règles de coarticulation de $[ibV_2], [idV_2], [V_1gV_2]$:

$$R_2(C) = f(R_2(V_2))$$

$$R_2[b] = 70\sqrt{(R_2(V_2) - 1841)} \quad (2)$$

$$R_2[d] = -8 \sqrt{(R_2(V_2) + 2500)} \quad (3)$$

$$R_2[g] = 2300 + 1500 * \text{Sig}[0.004 * R_2(V_2) - 6]$$

$$\text{Avec Sig} = \frac{(e^X - 1)}{(e^X + 1)} \quad (4)$$

Ces règles ont été implémentées et testées à l'aide d'un compilateur de règles. Autrement dit, la fréquence d'une résonance au milieu d'une voyelle tend, suivant une fonction exponentielle de sa durée, vers une valeur fixe (cible).

5. DISCUSSION

L'examen des spectrogrammes et de suivi de résonances des séquences [V₁CV₂] prononcées par deux sujets, suggère un type de coarticulation sensiblement différent en Français de celui constaté par Öhman [12]. Selon les résultats d'Öhman : *The articulatory system prepares for the medial consonant during all of the initial vowel*. Ce qui est confirmé pour le Français, ce que l'influence de la voyelle subséquente ne se fait pas sentir avant la consonne. Certaines composantes de cette voyelle se réalisent déjà pendant la voyelle précédente, d'où des transitions variables en fonction de la voyelle transconsonantique. Lorsqu'on examine les séquences où les deux voyelles sont identiques [aga, igi, ...], nous constatons que les transitions de départ et d'arrivée de la même voyelle varient légèrement en étendue. La tendance est à ce que les transitions de départ soient plus longues que celles d'arrivée, ce qui semble pouvoir s'expliquer si nous supposons que les instructions sont à chaque fois renouvelées pour le segment le plus proche. Nous pouvons également supposer que l'arrondissement labial coarticule indépendamment de la position linguale. Tout ceci reste à vérifier et constitue autant de directions de recherche pour l'avenir.

6. CONCLUSIONS

Cette étude montre l'importance du phénomène de la coarticulation dans les systèmes de synthèse par règles durant les perturbations consonantiques (variabilité du locus). La théorie du locus qui tente de définir chaque consonne par le point de convergence des formants vocaliques n'est pas vérifiée dans les faits. D'ailleurs, nos données confirment la proposition de Klatt pour une théorie de locus variable selon la voyelle attenante. Les déformations dues au contexte n'ont pas à notre connaissance fait l'objet d'études systématiques d'une certaine importance. La segmentation de la parole continue en dépend. Il est d'ailleurs probable qu'elles soient un obstacle à la reconnaissance automatique de la parole. Il existe d'importantes variations individuelles dans l'étendue et le détail des faits de la coarticulation. Mais, il est clair que l'effet est d'autant plus marqué que le débit est plus grand. Nous pensons que sans être indispensable à l'intelligibilité de la parole synthétique, la prise en compte de ces inflexions en contexte doit en améliorer la compréhension et surtout le caractère naturel. Quelle que soit l'explication que nous donnons à la coarticulation, il faut admettre qu'elle constitue une

caractéristique importante du langage humain dans tous ses aspects : acoustique, physiologique, articulatoire et perceptuel.

Les découvertes d'Öhman ne semblent pas être universelles. Il est probable que chaque langue ou famille de langues possède ses propres règles de coarticulation.

L'interprétation des spectrogrammes comme des structures de résonances permet une inversion acoustico-articulatoire plus claire et permet ainsi une voie facile de la modélisation des trajectoires formantiques. Des loci consonantiques ont été révisés pour tenir compte de cette nouvelle architecture. Nous pensons que cette représentation acoustique (en résonances) du signal de la parole peut suggérer une nouvelle perspective pour un nombre de problèmes-clés, tel que le calcul du formant effectif F₂'.

7. REFERENCES

- [1] Stevens K.N. (1972). Bases for universals in the properties of the speech production and perception systems. Proc. of the IXth ICPhS. Vol 2, Copenhagen, 53-59.
- [2] Lindblöm B.E.F. (1972). Some phonetic null hypotheses for biological theory of language. IX^{ème} ICPhS, 2, Copenhagen, 33-40.
- [3] Kuhn G.M. (1979). Stop consonant place perception with single formant stimuli : evidence for the role of the front cavity resonance. J. Acoust. Soc. Am. 65 (3), 774-788.
- [4] Hermansky H. & Broad D. (1989). The effective second formant F₂' and the vocal tract front cavity. ICASSP, 480-483.
- [5] Stevens K.N. & Blumstein S.E. (1981). The search for invariant acoustic correlates of phonetic features. In Eimas P. & Miller J. (Eds.), Perspectives on the study of speech. Hillsdale, New Jersey : Lawrence Erlbaum Associates.
- [6] Delattre P. (1966). Les attributs acoustiques de la nasalité vocalique et consonantique, les indices acoustiques de la parole. Studies in French and Comparative Phonetics, Mouton, the Hague.
- [7] Klatt D.H. (1987). Review of Text-To-Speech conversion for English. J. Acoust. Soc. Am. 82 (3), 737-1221.
- [8] Delattre P., Liberman A.M., & Cooper F.S. (1955). Acoustic loci and transitional cues for consonants. J. Acoust. Soc. Am. 27 (4), 769-774.
- [9] Hammarberg R. (1982). On redefining coarticulation. Journal of Phonetics, N°10, 123-137.
- [10] Guerti M. & Bailly G. (1990). Anticipation et rétention dans les mouvements vocaliques du Français. XVIII^{ème} JEP. 28-31 Mai, Montréal-Canada, 292-295.
- [11] Callopie (1989). La parole et son traitement. Edition Masson, 718 pages.
- [12] Öhman S.E.G. (1967). Numerical model of coarticulation. J. Acoust. Soc. Am. 41 (2), 310-320.
- [13] Strange W. (1989). Dynamic specification of coarticulated vowels spoken in sentence context. J. Acoust. Soc. Am. 85 (5), 2135-2153.
- [14] Boë L.J., Perrier P. & Bailly G. (1992). The geometric vocal tract variables controlled for vowel production : proposals for constraining acoustic-to-articulatory inversion. Journal of Phonetics, Vol. (20) N°1, 27-38.
- [15] Delattre P. (1969). Coarticulation and the locus theory. Studia Linguistica 23, 1-26.
- [16] Recasens D. (1984). Vowel-to-vowel coarticulation in Catalan VCV sequences. J. Acoust. Soc. Am. 76 (6), 1624-1635.
- [17] M. Guerti, Speech synthesis by rule, 8th International Conference on Computer Theory and Applications ICCTA'98, IEEE Alexandria Chapter, 15-17 September 1998, Alexandria - EGYPT, III.12-III.15
- [18] M. Guerti, Speech synthesis by diphones, 9th International Conference on Computer Theory and Applications, ICCTA'99, IEEE Alexandria Chapter, 27-30 August 1999, Alexandria - EGYPT, vol.2, pp. 278-281.
- [19] M. Guerti, Résonance du triangle vocalique, Séminaire National sur l'Automatique et les Signaux, SNAS'99, 9 et 10 Novembre 1999, Annaba Algérie, pp. 110-116.

DUREES INTRINSEQUES DES SONS SPECIFIQUES A L'ARABE STANDARD

M.L. BENZAOUÏ* et M. GUERTI**.

الملخص

إن لمعرفة العناصر العروضية أهمية كبرى عند تركيب الصوت أو التعرف عليه [1],[2]. فلهذا السبب أردنا دراسة أحد هذه العناصر ألا وهي المدة الزمنية للصوت و هذا بالاعتماد على النموذج التنبؤي ل : KLATT [5],[6] والذي يعطي المدة الزمنية للصوت معتمدا على قيمة أساسية تسمى القيمة الذاتية. في البدء قمنا بدراسة القيمة الذاتية للأصوات الخاصة باللغة العربية باستعمال كلمات مصطنعة بهدف أخذ كل التركيبات الصوتية الممكنة. ثم استخرجنا قواعد التنبؤ لهذه الأصوات بهدف استعمالها في تركيب الصوت من الأصوات الثنائية [5].

الكلمات الخاصة : العروض المدة الصوتية المدة الذاتية التنبؤ للمدة الأصوات الثنائية.

Abstract

The knowledge of phoneme prosodic parameters has a very big importance in the speech recognition and synthesis [1], [2]. So we are interested to make a survey of phonemic duration which is one of the prosodic parameters, and this by inspiring it of a predictive model of KLATT [3], [4], which proposes a predicted duration from duration of basis called intrinsic duration.

We first measured the intrinsic duration of the Standard Arabic specific sounds, using artificial words in order to take all possible combinations of phonemes, after this we extract some prediction rules of this duration in order to use it in the synthesis by diphone [5].

Key words : prosody, phonemic duration, intrinsic duration, predicted duration, diphone.

Résumé

La connaissance des paramètres prosodiques a une grande importance en synthèse et en reconnaissance de la parole [1],[2]. De ce fait nous nous sommes intéressés à l'étude d'un de ces paramètres pertinents, à savoir la durée phonémique. En s'inspirant du model prédictif de KLATT [4],[5], qui prédit la durée du phonème à partir d'une durée de base appelé durée intrinsèque.

En premier lieu, nous avons mesuré les durées intrinsèques des sons spécifiques à l'Arabe Standard, en utilisant des mots artificiels afin de prendre toutes les combinaisons des phonèmes possibles, puis nous avons déterminé les règles de prédiction de ces durées pour les utiliser dans la synthèse par diphtongues [5].

Mots clés : prosodie, durée phonémique, durée intrinsèque, durée prédite, diphtongues.

* Institut d'Electronique, U.S.T.H.B. B.P. 32, Bab Ezzouar 16111 Alger, Algérie.

** Ecole Nationale Polytechnique B.P. 182 EL HARRACH 16200 Alger, Algérie.

Fax (213) 2 522973 E-mail mhaniag@hotmail.com

1. INTRODUCTION

L'homme cherche toujours à communiquer avec la machine à l'aide de la voix humaine. D'où la nécessité de systèmes de synthèse de la parole de haute qualité.

La synthèse de la parole a pour rôle de convertir les messages écrits en une forme d'ondes acoustiques et ceci en utilisant toutes les possibilités qu'offre la technologie électronique : des supports de stockage des messages "mémoires" aux circuits de Conversion Numérique Analogique. Mais jusqu'à présent cette technologie ne permet pas d'obtenir des tailles mémoires très grandes ce qui limitent les textes des messages.

Pour remédier à ce problème, les chercheurs ont trouvé que l'utilisation des unités plus petites constituant le signal parole "Phonèmes" permet de réduire la taille de dictionnaire (vocabulaire du message). Cette technique a permis d'avoir des systèmes simples, de synthèse de la parole mais l'intelligibilité reste mauvaise, d'où on a choisi une autre unité "le diphonème" ou diphone ou encore dyade (ensemble contenant deux phonèmes). En utilisant ces unités nous devons appliquer des règles de variations des caractéristiques physiques du signal parole, qui varient en fonction de la langue de la hauteur de la voix et du débit d'élocution. Ces paramètres sont appelés paramètres prosodiques (fréquence fondamentale ou F_0 ou pitch, intensité ou énergie et la durée ou rythme).

Nous présentons dans ce travail une étude d'un de ces paramètres à savoir la durée phonémique pour une application à une synthèse par règles basée sur les diphones de la langue Arabe Standard (A.S.).

Pour cela nous avons élaboré un corpus afin de prendre les combinaisons de diphones de l'A.S. Nous avons choisi des logatomes (mots artificiels) contenant les diphones à étudier et après l'enregistrement par quatre locuteurs nous avons analysé les signaux afin de dégager un modèle de prédiction des durées phonémiques suivant différents contextes et ceci en s'inspirant du modèle prédictif de D.KLATT [5], [6].

2. TECHNIQUES DE SYNTHESE DE LA PAROLE A PARTIR DU TEXTE

La synthèse de la parole à partir du texte TTS (Texte To Speech) est une opération de génération automatique d'un signal de parole à partir d'un message écrit [1], [7]. Pour cela les systèmes de synthèse à partir du texte ont l'architecture suivante (figure 1) :

- l'analyse linguistique regroupe l'analyse morphologique, syntaxique et sémantique, etc ;
- le transfert phonologique permet de générer l'arbre prosodique ;
- l'étape de génération est caractérisée par la méthode de synthèse employée (synthèse par unités stockées ou bien, synthèse par règles) ;
- l'étape de synthèse passage de la représentation paramétrique au signal acoustique.

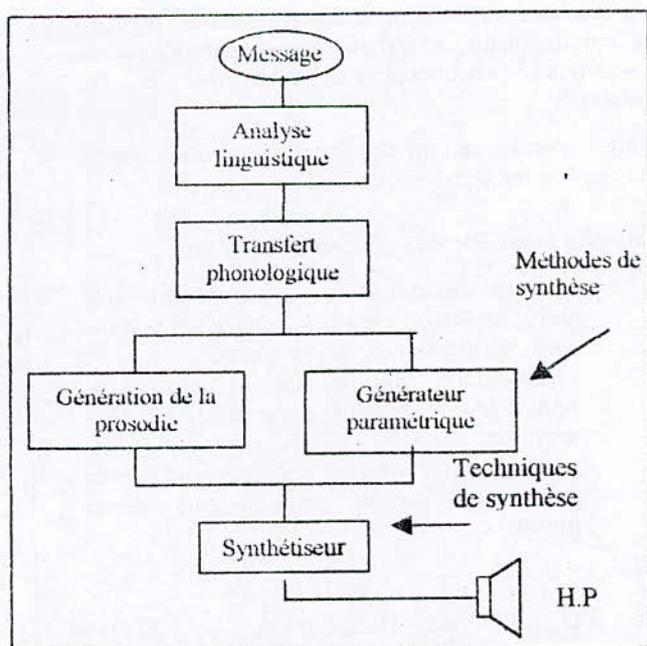


Fig.1 Architecture générale d'un système de synthèse à partir du texte (TTS).

Dans notre travail nous nous sommes intéressés aux techniques de synthèse (l'étape de génération) et spécialement aux paramètres prosodiques, pour une synthèse par règles.

3. QU'EST-CE-QUE LA PROSODIE ?

Lorsque nous parlons, nous ne sommes en général pas conscients des mouvements complexes des muscles de la phonation, et il en va de même en particulier pour le contrôle de la hauteur et de l'intensité de la voix lors des vibrations des cordes vocales. Ces deux paramètres auxquels on joint habituellement les durées successives des segments syllabiques constituent en leur évolution la prosodie de la phrase [1], [8].

Dans le Traitement Automatique de la Parole (TAP), les paramètres prosodiques prennent une importance particulière. En synthèse, ils améliorent le naturel et l'intelligibilité du signal synthétique en signalant les grandes articulations de la phrase ; en reconnaissance, ils peuvent servir d'indices pour l'identification d'éléments segmentaux déterminés et également signaler le type syntaxique de la phrase [1], [7], [9].

Les paramètres prosodiques sont :

- la mélodie (fréquence fondamentale) ou F_0 correspond à la fréquence des vibrations des cordes vocales lors de la production des phonèmes ;
- l'intensité traduit l'importance énergétique d'un phonème. Pour les sons spécifiques à l'Arabe Standard, elle permet de déterminer l'emphase (son plus énergétique que son opposé non-emphatique), utilisée aussi pour la détection des sons voisés (périodiques) ;
- la durée ou rythme, contient un temps de phonation et un temps de silence. Elle permet la perception de l'accent, et dépend du débit d'élocution (lent, rapide, normal, etc...).

La durée phonémique, dépend des paramètres linguistiques, extralinguistiques, physiologiques, phonologiques, inter-locuteurs, extra-locuteurs, ainsi que la co-articulation⁽¹⁾, ...

Ce qui a permis de prédire la durée d'un segment phonémique suivant des modèles de prédiction :

- D'après le modèle de D. KLATT [4], [5],[6] :
 - chaque segment phonémique possède une durée intrinsèque qui est considérée comme une spécification de ce phonème ;
 - chaque règle doit prendre un pourcentage additif ou soustractif dans la durée d'un segment ;
 - les segments ne peuvent être compressés plus d'un certain seuil appelé durée minimale.

le modèle est donc :

$$Dur = Dur_{min} + (Dur_{int} - Dur_{min}) * k\% \quad (1)$$

Avec :

- Dur : durée prédite ;
- Dur_{min} : durée minimale ;
- Dur_{int} : durée intrinsèque ;
- k% : Pourcentage de variation de la durée en fonction des effets prosodiques des mots.

- Suivant le modèle multiplicatif de SORIN [4],[9]:

$$Dur = Dur_{int} * k \quad (2)$$

dont :

- Dur : durée prédite ;
- Dur_{int} : durée intrinsèque ;
- K : coefficient multiplicatif dépendant des paramètres cités.

En ce qui concerne notre travail nous avons mesuré la durée intrinsèque des sons spécifiques à l'Arabe Standard, et déterminé les règles de prédiction pour les appliquer dans la synthèse par règles.

4. METHODES DE SYNTHÈSE

Pour synthétiser une phrase, on peut concaténer les mots qui la constituent. Ces derniers sont donc stockés pour former un dictionnaire de synthèse. Mais si on s'intéresse à produire un texte long le dictionnaire augmente de taille. Ce qui a obligé les chercheurs à trouver une méthode pour réduire la taille du dictionnaire [6], [7]. Cette méthode se base sur des unités plus petites "les Phonèmes" qui sont les éléments essentiels de production de la parole.

La synthèse devient donc plus facile et le dictionnaire plus réduit ; mais l'intelligibilité reste mauvaise et ceci à cause de la zone de transition entre deux phonèmes successifs, ou encore le phénomène de coarticulation,

dont il faut tenir compte lors des variations prosodiques dans la phrase à synthétiser [1], [2].

De ce fait on prend en considération ces transitions par le choix d'une nouvelle unité stockée qui est le diphone.

Le **diphone** étant défini comme « le segment qui s'étend de la zone stable d'une réalisation phonémique à la zone stable de la réalisation suivante et qui protège en son centre toute la zone de transition » [3], [10] (figure 2).

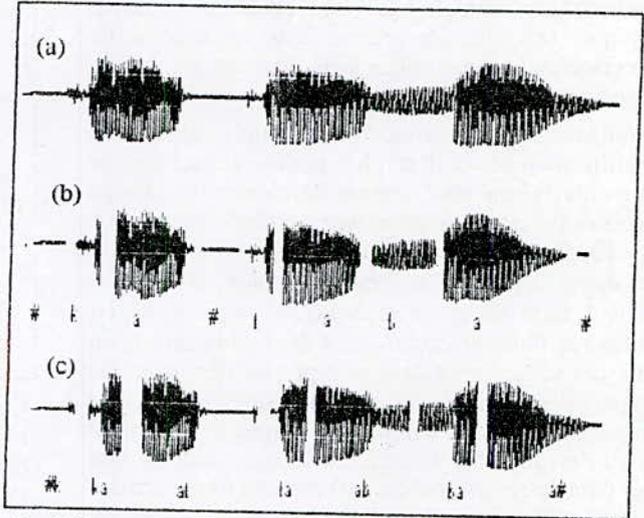


Fig.2 Exemple d'unité segmentée du mot [kataba].

- (a) : forme temporelle ;
- (b) : segmentation en phonèmes ;
- (c) : segmentation en diphones.

5. LES TECHNIQUES D'ANALYSE

Il existe plusieurs techniques d'analyse du signal de parole : l'analyse temporelle, spectrale, FFT, LPC (Lineaire Predictive Coding), Cepstrale (ou homomorphique),...

Ces techniques peuvent être utilisées pour extraire les indices prosodiques (fréquence fondamentale ainsi que les paramètres pertinents du signal vocal), mais pour l'extraction de la durée phonémique on fait appel à d'autres techniques, par exemples : le VoCodeur à canaux (Voice Coder) et le sonographe [2],[3],[11].

Concernant notre étude nous avons utilisé la méthode d'analyse par sonagrammes numériques (figure3) pour une segmentation des mots du corpus en phonèmes, et l'extraction de leurs durées intrinsèques. [2],[3],[10].

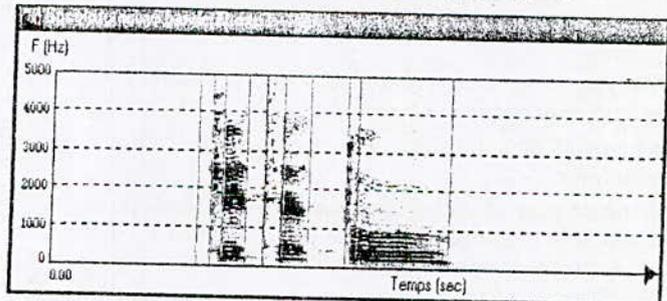


Fig.3 le sonogramme numérique du mot kataba.

(1) La coarticulation est l'effet d'un son sur celui qui le précède

Un sonagramme [1] étant une représentation à trois dimensions (temps-fréquence et intensité ou énergie du son) sur laquelle se projette le spectre tridimensionnel de la parole. Son principe est fondé sur le filtrage hétérodyné :

Soit $S(t)$ le signal de parole et $\check{S}(t)$ le signal modulé par une porteuse sinusoidale.

$$\check{S}(t) = S(t) * \cos(2\pi ft) \quad (3)$$

Le spectre d'amplitude est décalé vers les fréquences positives (figure 4) et défile sous un filtre fixe de bande passante B . Selon la largeur du filtre on obtient un sonagramme large bande ou en bande étroite dans lequel l'intensité (énergie) est représentée par le degré de noircissement.

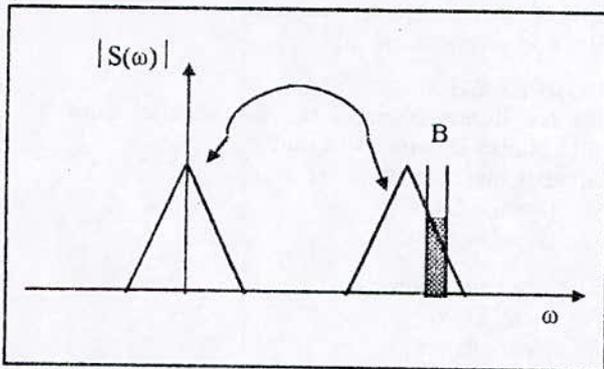


Fig.4 : spectre de $S(t)$ modulé

On remarque que si B tend vers zéro, le spectre tend vers la Transformée de Fourier de $S(t)$.

Cette technique numérique permet aux traiteurs du signal de la parole suivant sa représentation tridimensionnelle de déterminer les sections des sons et leurs fréquences ainsi que leur énergie.

6. MESURES ET RESULTATS

Afin d'extraire les paramètres prosodiques nous avons élaboré un corpus de mots contenant les diphtonges à étudier et suivant différents contextes.

Puis nous avons choisi quatre locuteurs professionnels : Deux masculins et deux féminins, pour enregistrer le corpus. L'enregistrement a été fait à une vitesse d'élocution normale entre 4 et 7 syllabes par seconde en utilisant une carte « OROS-AU21 » spécifique au traitement du signal vocal, avec une fréquence d'échantillonnage de 10 kHz. Les échantillons ont été codés sur 16 bits par échantillon, dans des fichiers format de données.

Le corpus choisi est pris suivant le dictionnaire de diphtonges en Arabe Standard élaboré par M.GUERTI [12], [13].

Les notations que nous avons prises sont :

- [#] : représente le silence ;
- [c] : // la consonne ;
- [v] : // une voyelle brève ;
- [v:] : // une voyelle longue ;
- [c_e] : // une consonne emphatique ;
- [c_{ne}] : // une consonne non-emphatique ;
- [-] suivi d'un phonème représente le début du diphtongème à étudier.

Les phonèmes étudiés sont représentés dans le code du tableau 1.

Phonème	Code API
ظ	[d.]
ط	[t]
ء	[?]
ق	[q]
ض	[d]
ح	[h]
ع	[ε]
ص	[S]

Tableau 1 : représentation des consonnes de l'Arabe Standard suivant le code API. (Alphabet Phonétique international).

Les locuteurs qui ont fait l'enregistrement du corpus ont à peu près le même âge et ils n'ont pas d'influence dialectale dans leurs prononciations de l'Arabe Standard. Les mesures des durées se font après une segmentation manuelle des phonèmes en utilisant (figures 5 de 5a jusqu'à 5d) un exemple du logatome [katat_eu:].

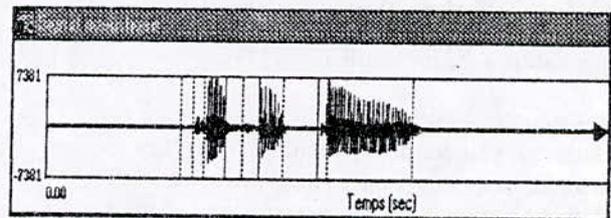


Fig. 5 a : forme temporelle du logatome [katat_eu:].

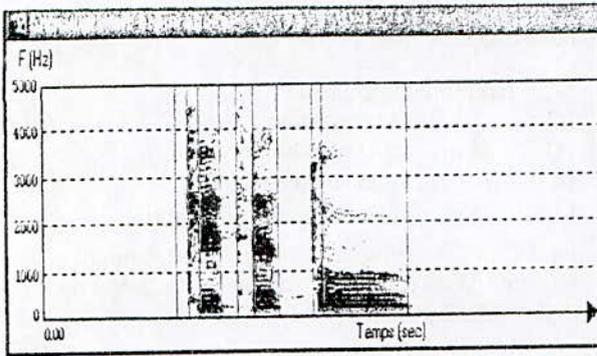


Fig.5 b : sonagramme du logatome [katat,u:].

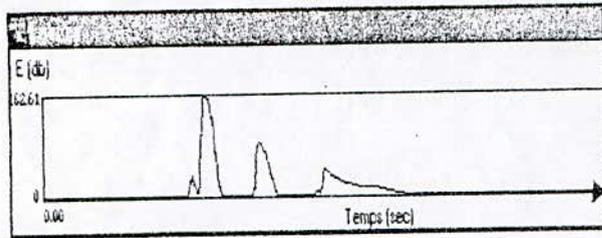


Fig.5 c : Energie.

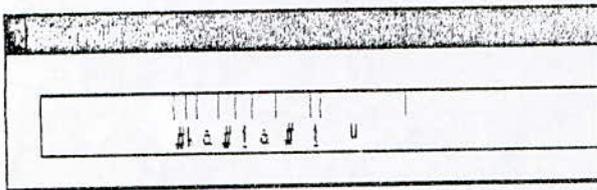


Fig.5 d : segmentation en phonèmes.

Fig.5 Exemple d'analyse sonographique du logatome [katat,u:] et sa segmentation en phonèmes.

Les figures précédentes de 5 a jusqu'à 5 d, présentent successivement :

- le signal temporel numérique ;
- le sonagramme numérique appliqué sur le signal temporel du mot enregistré. Pour voir les fréquences formantiques et déterminer les zones de parole à segmenter ;
- l'énergie on l'utilise avec le sonagramme pour déterminer les sons voisés et non voisés. Dans le cas de la segmentation, l'énergie permet de limiter le début et la fin d'un segment. Avec toutefois une vérification par réécoute pour s'assurer que la segmentation a été bien faite [2],[3],[4].

Le corpus d'étude élaboré permet :

- d'extraire les durées phonémiques des voyelles ;
- d'extraire les durées phonémiques des consonnes ;
- de voir les variations des durées suivant le contexte.

Les sons choisis sont :

- les voyelles :
 - brèves [a], [u], [i] ; fatha, damma et kasra.
 - longues [a:], [u:], [i:]. Huru:f el mad.

- les consonnes :
 - occlusifs non voisés : [t.], [ʔ], [q].
 - occlusifs voisés : [d.].
 - fricatifs emphatiques : [s], [d].
 - fricatifs pharyngales : [ɛ], [h].

Ces phonèmes sont placés dans des logatomes ou mots artificiels afin de prendre toutes les combinaisons possibles.

7. PREDICTION DES DUREES PHONEMQUES DES VOYELLES

La langue Arabe Standard comporte trois types de voyelles (haraka) appelées voyelles brèves : [a],[u],[i] et des voyelles longues qui sont réalisées par un allongement des voyelles brèves (El mad) : [a:], [u:], [i:]. En présence de consonnes emphatiques toutes ces voyelles sont influencées énergétiquement d'où l'obtention de variantes : voyelles emphatisées et voyelles non-emphatisées[13].

7.1 CORPUS CHOISI

L'étude des durées phonémiques des voyelles nous a conduit à choisir le corpus suivant :

- contexte non emphatique :
 - [#cat_{ne}at_{ne}-v:#] ;
 - [#cat_{ne}at_{ne}-v#].
- contexte emphatique :
 - [#cat_{ne}at_e-v#] ;
 - [#cat_{ne}at_e-v:#].

7.2 MESURES DES DUREES INTRINSEQUES

Ce corpus nous permet d'extraire une durée qu'on l'appelle durée intrinsèque, et les lois de variation des durées phonémiques pour obtenir un modèle de prédiction des durées des voyelles de l'Arabe Standard. Les tableaux 2 et 3 représentent les résultats obtenus pour les phonèmes utilisés.

Phonèmes	Locuteurs			
	1	2	3	4
[a]	115	090	103	143
[a:]	180	163	154	256
[u]	110	103	122	110
[u:]	197	170	178	240
[i]	095	097	105	121
[i:]	210	160	162	260

Tableau 2: Durées en ms des voyelles brèves (longues) dans le contexte non-emphatique [#katat_{ne}-v#] ([#katat_{ne}-v:#]), c'est-à-dire les diphtones [v#] et [v:#].

Phonèmes	Locuteurs			
	1	2	3	4
[a]	102	085	105	132
[a:]	180	156	195	250
[u]	117	098	120	128
[u:]	217	155	167	257
[i]	102	098	110	128
[i:]	226	197	210	247

Tableau 3: Durées en ms des voyelles brèves (longues) dans le contexte emphatique [#katat_e-v#] ([#katat_e-v:#]), diphtones [v#] et [v:#].

Les tableaux 2 et 3 donnent les valeurs des durées des voyelles dans la position fin du mot (contexte [#katat_{ne}-v#] et [#katat_e-v#]).

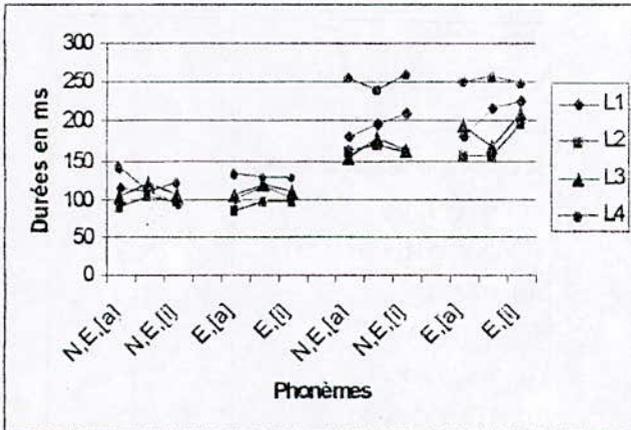


Fig.6 : Comparaison des durées phonémiques des voyelles de l'Arabe Standard.

Pour une comparaison plus complète, le graphe de la figure 6 donne pour chaque locuteur les durées des voyelles dans le contexte emphatique ou non-emphatique avec une voyelle brève ou une longue, et ceci pour les trois types de voyelles [a], [u], [i].

Les valeurs des durées phonémiques pour les quatre locuteurs sont très proches dans le cas des voyelles brèves, mais la distance augmente dans le cas des voyelles longues.

Nous pouvons conclure d'après ce graphe et ces tableaux, que les valeurs des durées intrinsèques des voyelles sont proches et de valeurs données par les tableaux ci-dessous :

Voyelles	Contexte emphatique	Contexte non emphatique
[a]	112	106
[u]	111	115
[i]	104	109

Tableau 4 : durées intrinsèques des voyelles.

Phonèmes	Locuteurs			
	1	2	3	4
[a:]/[a]	1.56	1.81	1.49	1.93
[u:]/[u]	1.79	1.65	1.45	2.18
[i:]/[i]	2.21	1.64	1.54	2.14

Tableau 5: Rapport [v:]/[v] dans le contexte non-emphatique [#katat_{ne}-v#] ([#katat_{ne}-v:#]).

Phonèmes	Locuteurs			
	1	2	3	4
[a:]/[a]	1.76	1.83	1.85	1.89
[u:]/[u]	1.85	1.58	1.36	2.00
[i:]/[i]	2.21	2.01	1.90	1.92

Tableau 6: Rapport [v:]/[v] dans le contexte emphatique [#katat_e-v#] ([#katat_e-v:#]).

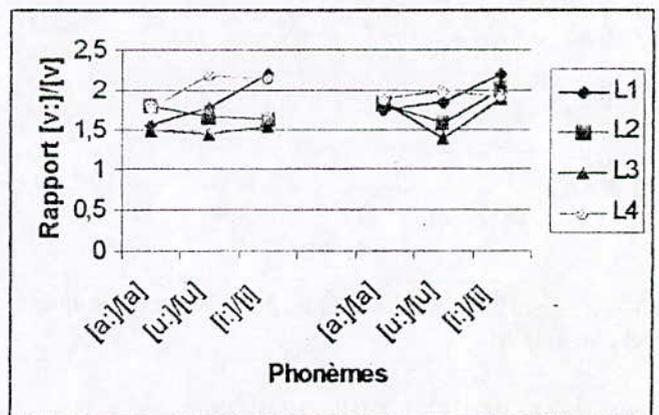


Fig.7 : Rapport [v:]/[v] (contexte non-emphatique puis contexte emphatique).

D'après le graphe de la figure 7, et d'après les tableaux 3, 4,5 et 6 nous pouvons dire que les durées des

voyelles longues peuvent être prédites à partir de celles des voyelles brèves, tels que dans le contexte :

- contexte non-emphatique :

$$\begin{aligned} [a:] &: D_{i[a]} = D_{i[a]} + 69\%D_{i[a]} \\ [u:] &: D_{i[u]} = D_{i[u]} + 76\%D_{i[u]} \\ [i:] &: D_{i[i]} = D_{i[i]} + 88\%D_{i[i]} \end{aligned}$$

- contexte emphatique :

$$\begin{aligned} [a:] &: D_{i[a]} = D_{i[a]} + 83\%D_{i[a]} \\ [u:] &: D_{i[u]} = D_{i[u]} + 69\%D_{i[u]} \\ [i:] &: D_{i[i]} = D_{i[i]} + 101\%D_{i[i]} \end{aligned}$$

Phonèmes	Locuteurs			
	1	2	3	4
[a]	1.12	1.05	0.98	1.08
[a:]	1.00	1.04	0.78	1.02
[u]	0.94	1.05	1.01	0.85
[u:]	0.90	1.09	1.06	0.93
[i]	0.93	0.98	0.95	0.94
[i:]	0.92	0.81	0.77	1.05

Tableau 7: Rapport[v] [#katat_{ne}-v#] ([#katat_{ne}-v:#]) sur [v] dans le contexte [katat_e-#] ([#katat_e-v:#]).

La figure 7 et le tableau 7 nous donnent une idée sur les variantes qui existent dans les contextes emphatiques et non-emphatiques.

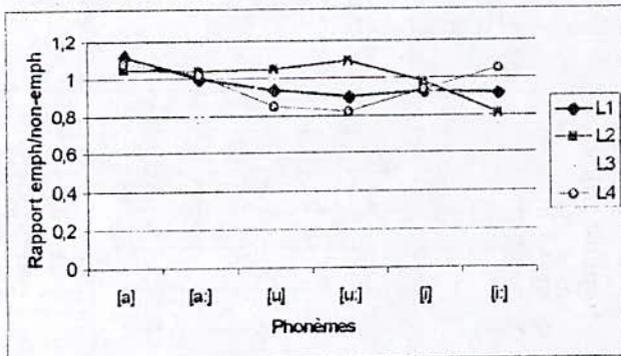


Fig.8 rapport des durées dans les contextes emphatiques et non-emphatiques.

7.1 REMARQUES ET DISCUSSION

Nous remarquons que les voyelles réalisées dans des contextes emphatiques ont des durées légèrement supérieures à celles réalisées dans des contextes non-emphatiques, et d'après les statistiques nous trouvons que les valeurs du rapport des durées entre les voyelles dans le contexte non emphatique et celles dans des

contextes emphatiques, sont autour d'une moyenne de 0.96 et d'une variance $\sigma=0.09$, ce qui justifie les résultats trouvés.

Finalement nous proposons un algorithme et un organigramme (figure 9), pour une prédiction des durées phonémiques des voyelles, et ceci suivant les équations dont nous utilisons des variables booléennes :

la durée D_i de la voyelle est calculée par prédiction à partir de la durée intrinsèque D_i :

$$D_i = K_e * D_e + K_{e'} * D_{ne} \tag{4}$$

Avec :

D_i : durée intrinsèque ;

K_e : variable booléenne définissant l'emphase de la Voyelle ;

$K_{e'}$: inverse de K_e ;

D_e (D_{ne}) : durée de la voyelle dans le contexte emphatique(non-emphatique).

$$D_i = D_i + L * (K_{ie} * K_e + K_{ine} * K_{e'}) * D_i \tag{5}$$

Avec :

D_i : durée totale prédite ;

L : variable booléenne définissant la voyelle longue ;

K_{ie} et K_{ine} : coefficient définissant la variation additive sur la durée.

Algorithme de prédiction:

```

Début
Initialisation des variables
Lecture de l'unité C1V1
Si C1 est emphatique alors (Ke=1)
sinon (Ke=0);
Si V1 est long alors (L=1)
sinon (L=0);
Si V1 = [a] alors
(De=106, Kie=0.83,
Dne=112, Kine=0.69)
sinon
Si V1 = [u] alors
(De=115, Kie=0.69,
Dne=111, Kine=0.76)
sinon
Si V1 = [i] alors
(De=109, Kie=1.01,
Dne=104, Kine=0.88)
sinon (erreur).
calcul de la durée intrinsèque:
Di = Ke * De + Ke' * Dne
calcul de la durée prédite:
Di = Di + L * (Kie * Ke + Kine * Ke') * Di
Fin.
    
```

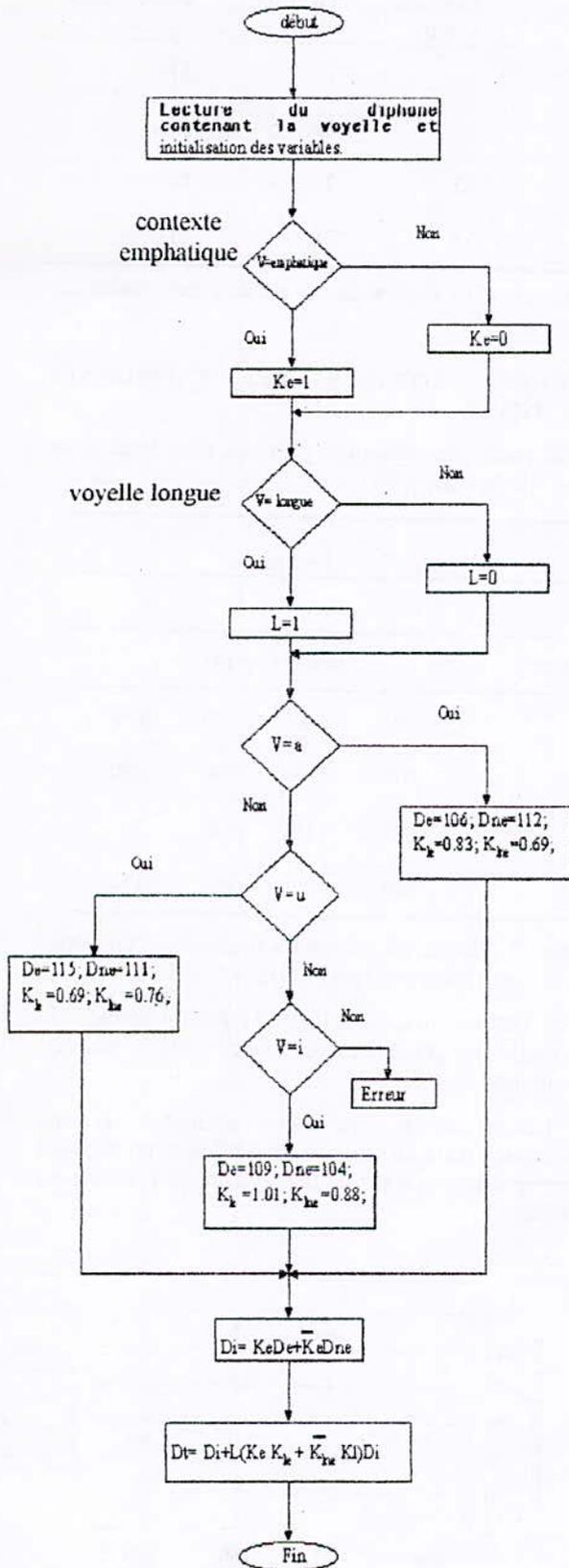


Fig. 9 organigramme de prédiction des durées phonémiques des voyelles de l'Arabe Standard.

8. PREDICTION DES DUREES PHONEMQUES DES CONSONNES

Comme nous avons vu, les consonnes de l'Arabe Standard ont différentes caractéristiques : consonnes plosives, consonnes fricatives,... et comme particularité de cette langue : les consonnes emphatiques (son plus énergétique) qui sont une caractéristique des langues sémitiques [3], [12], [14].

Nous avons élaboré le corpus de la manière suivante :

- [#c-ata #]: pour l'étude des durée intrinsèque consonnes.
- [#kata-c #]: pour voir l'effet de la position finale sur la durée de la consonne.

En premier lieu nous avons choisit les consonnes occlusifs : [t.], [ʔ], [q], [d], en suite nous avons étudié les autres cas de consonnes.

8.1 DUREES DE LA PLOSION CONSONANTIQUE

Etant donnée que les sons plosifs sont caractérisés par un temps de silence (tenue consonantique) et un temps de phonation appelé plosion, nous avons préféré mesurer en premier lieu les durées des plosions et de voir leur variation, par la suite nous avons étudié les durées des tenues consonantiques et finalement les durées finales et la proposition du modèle de prédiction.

Phonèmes	Locuteurs			
	1	2	3	4
[d.]	15	18	16	12
[t]	22	22	21	22
[ʔ]	25	24	21	22
[q]	25	23	25	27

Tableau 8: durées des consonnes plosives dans le contexte [#kata-c#], plosion du diphone [c#].

Phonèmes	Locuteurs			
	1	2	3	4
[d.]	18	18	20	20
[t]	22	27	25	25
[ʔ]	28	26	25	27
[q]	28	29	27	27

Tableau 9: durées des consonnes plosives dans le contexte [#c-kata], plosion du diphone [c-].

A partir de ces résultats, (Tableau 8 et 9), nous remarquons que la durée de la plosion est plus grande aux positions finales du mot qu'à celles de début.

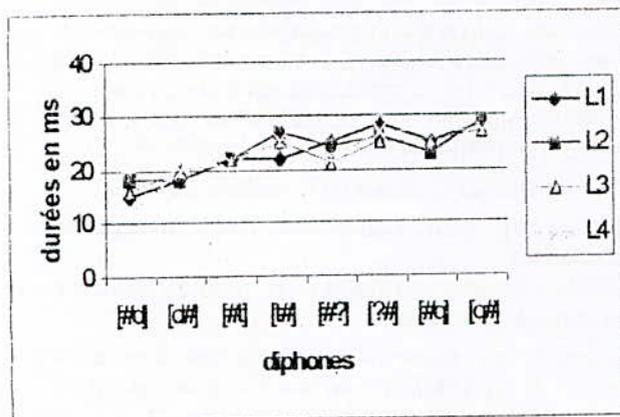


Fig. 10 Comparaison des durées des plosions des consonnes plosives de l'Arabe Standard.

Nous avons, calculé la proportion entre ces durées pour prendre la relation entre les durées, d'après leurs positions (tableau 10).

Proportion de la durée de l'explosion entre le début et la fin du mot (explosion de[-c#]/explosion de[#c-).

Phonèmes	Locuteurs			
	1	2	3	4
[d.]	1.20	1.00	1.25	1.66
[t]	1.00	1.22	1.13	1.19
[ʔ]	1.12	1.08	1.19	1.22
[q]	1.12	1.26	1.08	1.00

Tableau 10 : rapport des durées des phonèmes suivant les positions début et fin du mot.

De ces résultats et par un choix, nous avons défini les durées intrinsèques de l'occlusion comme la durée du phonème dans la position début du mot, du fait que ces durées et plus petite que celle de la position finale du mot.

Donc, la durée dans la position finale du mot est donnée par l'équation :

$$D_{pf} = D_{pd} + k_p * D_{pd} \quad (6)$$

Avec:

- D_{pf}: durée de la plosion dans la position finale ;
- D_{pd}: durée de la plosion dans le début du mot ;
- K_p : pourcentage déterminé par la proportion du tableau 9.

Ainsi nous regroupons les résultats dans le tableau 10. Les durées de l'explosion.

Phonèmes	Position début	Position finale	pourcentage (kp %)
[d.]	15	19	27
[t]	22	25	13
[ʔ]	23	26	15
[q]	25	28	11

Tableau 11 : les durées de la plosion des consonnes plosives.

8.3 DUREES TOTALES DES CONSONNES PLOSIVES

La durée totale des consonnes plosives, dans le contexte [#kata-c #], diphone [c #].

Phonèmes	Locuteurs			
	1	2	3	4
[d.]	153	130	130	194
[t]	200	182	180	270
[ʔ]	208	116	130	272
[q]	195	157	183	124

Tableau 12 : durées phonémiques totales des consonnes plosives spécifiques à la longue Arabe Standard.

Dans ce Tableau et dans la Figure 11 nous présentons la durée totale des consonnes occlusives dans la position finale du mot.

Nous devons noter cette durée occlusive est une concaténation entre un segment de silence, et un segment appelé explosion où il y a de la phonation et l'énergie est importante.

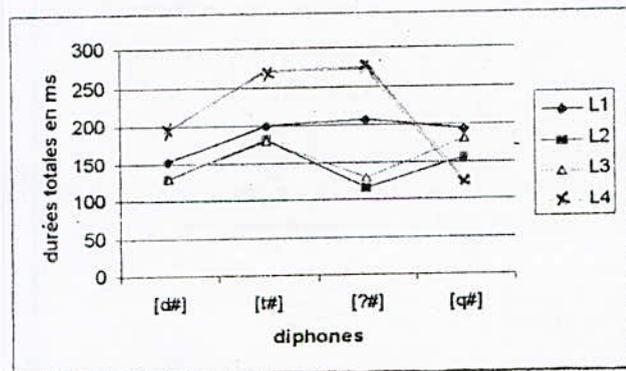


Fig 11 La durée totale des consonnes plosives.

La proportion de la durée de la consonne occlusive entre (durée de l'explosion de [-c #]) et (durée de la consonne totale de [-c #]) dans la même position dans le mot.

Phonèmes	Locuteurs			
	1	2	3	4
[d.]	11.76	13.84	15.38	10.31
[t]	11.00	14.83	13.88	09.25
[ʔ]	13.46	22.41	19.23	09.92
[q]	14.35	18.47	14.75	21.77

Tableau 13 : proportions en % entre la durée de la plosion et la durée totale du phonème.

Du tableau 13, nous pouvons dire que la durée de la plosion est prédite à partir de la durée totale par l'équation:

$$D_p = p * D_T \quad (7)$$

Avec :

- D_p durée de la plosion.
- D_T durée totale.
- P coefficient de plosion.

Et les durées du silence D_s qui sont données par le tableau 12, seront prédites par l'équation :

$$D_s = D_T - D_p \quad (8)$$

Ce qui nous permet d'écrire le tableau 14 comme conclusion de prédiction des durées des tenues consonantiques.

Phonèmes	Durées dans la position initiale en ms	Durées dans la position finale en ms
[d.]	100	133
[t]	160	183
[ʔ]	120	155
[q]	122	137

Tableau 14 : durées des tenues consonantiques.

Si nous utilisons E comme une variable Booléen indiquant la position du phonème dans le mot (si #c E = 1 ; si c #E = 0).

Nous pouvons prédire la durée intrinsèque totale de l'équation :

$$D_T = D_{Td} + E * K_E * D_{Td} \quad (9)$$

Où K_E est le pour cent de durée de la position initiale, et D_{Td} la durée dans le début du mot.

Nous devons noter que la durée intrinsèque dans notre cas est D_{Td} la durée de la plosion dans le début du mot. Ainsi nous regroupons les coefficients dans le tableau ci-dessous :

Les durées de la prédiction.	D _{Tb} ms	K _E %	P %
Locuteurs phonèmes			
[d.]	115	32	13
[t]	182	14	12
[ʔ]	143	26	16
[q]	147	12	17

Tableau 15 paramètres de prédiction des durées phonémiques des consonnes.

Dans le tableau 15 résumes les paramètres de prédiction des durées des consonnes plosives.

Ans nous proposons l'algorithme de prédiction :

```

Début
  Lire le diphone ([#C -] ou [-C #])
  Initialisation des variables
Si C est dans le début du mot [-C #] alors
  E = 0 sinon E = 1
Lire DTd et KE
Calculer la durée totale
DT = DTd + E * KE * DTd
Calculer la durée de l'explosion
  Dp = Kp * DT
Calculer la durée du silence
  Ds = DT - Dp
La fin
    
```

9.3 DUREES TOTALES DES CONSONNES NON PLOSIVES

De la même manière nous obtenons pour les consonnes non plosives les résultats suivants.

Phonèmes	Locuteurs			
	1	2	3	4
[s]	134	140	100	145
[d]	115	089	086	094
[ɛ]	112	088	088	123
[h]	125	087	128	115

Tableau 15 : Durées en ms des consonnes dans le contexte #c-ata#, diphone #c.

Phonèmes	Locuteurs			
	1	2	3	4
[s]	244	215	182	272
[d]	158	112	105	124
[ɛ]	141	139	122	165
[h]	196	168	160	150

Tableau 16 : Durées en ms des consonnes dans le contexte #kata-c#, diphone c#.

Phonèmes	Locuteurs			
	1	2	3	4
[s]	1.82	1.53	1.82	1.87
[d]	1.37	1.26	1.22	1.32
[ɛ]	1.26	1.58	1.38	1.34
[h]	1.56	1.93	1.25	1.30

Tableau 17: Rapport des durées des consonnes dans la fin du mot sur ceux du début du mot. (contexte #kata-c# et #c-ata#)

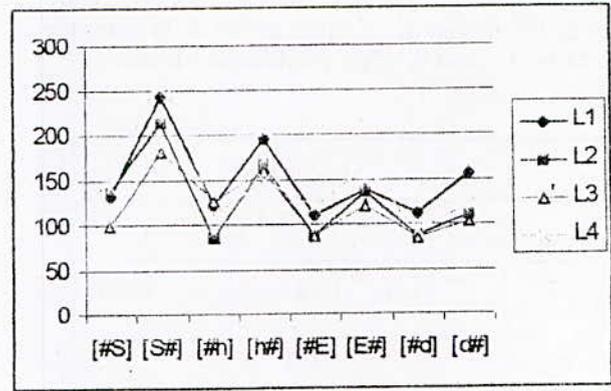


Fig.12 représentation graphique des durées des consonnes fricatives.

D'après le graphe de la figure 12 et les résultats des tableaux 15,16 et 17 nous pouvons extraire les durées intrinsèques et les règles de prédiction des consonnes: durée intrinsèque:

$$Di_{[s]} = 129 \text{ ms.}$$

$$Di_{[d]} = 96 \text{ ms.}$$

$$Di_{[ɛ]} = 102 \text{ ms.}$$

$$Di_{[h]} = 113 \text{ ms.}$$

nous pouvons donc prédire les durées dans la position finale à partir de celles de la position initiale par le modèle :

$$Df_{[s]} = Di_{[s]} + 76\% Di_{[s]} .$$

$$Df_{[d]} = Di_{[d]} + 29\% Di_{[d]} .$$

$$Df_{[ɛ]} = Di_{[ɛ]} + 38\% Di_{[ɛ]} .$$

$$Df_{[h]} = Di_{[h]} + 48\% Di_{[h]} .$$

Finalement nous proposons le même algorithme de prédiction des durées des consonnes plosives pour celles non plosives avec $Ds = 0$.

Et la relation (6) reste valable.

9. CONCLUSIONS

Pour obtenir une synthèse TFS de haute qualité nous sommes amenés à étudier les caractéristiques du signal de parole à savoir : la fréquence fondamentale, l'intensité et la durée phonémique. Ces paramètres varient en fonction des effets externes et internes à la langue.

Nous avons donc choisi d'étudier pour la synthèse de l'Arabe Standard, les durées phonémiques des sons spécifiques à cette langue telles que les emphatiques et la glottale, et de voir leurs variations en fonction du contexte.

Les mesures que nous avons faites permettent de proposer des durées intrinsèques des phonèmes étudiés. Les résultats obtenus sont proches de ceux trouvés par d'autres chercheurs.

Et en s'inspirant du modèle prédictif de D.KLATT [8], nous avons proposé des règles de prédiction de ces durées en fonction du contexte. Ces règles concernent les voyelles brèves et longues et les consonnes peuvent être utilisées dans la synthèse par règles appliquée aux Diphones de l'Arabe Standard.

10. REFERENCES

- [1] CALLIOPE "la parole et son traitement automatique" (Masson 1989).
- [2] J. Hatton, J. M. Pierrel, G. Perrenou, J. Caelen, J. L. Gauvain, "La reconnaissance automatique de la parole". Dunod, Paris, France 1991.
- [3] A. Amrouche "Contribution à la synthèse de la parole en Arabe Standard, modèle de prédiction des phonèmes". thèse de magistère. USTHB, Algérie. mai 1995.
- [4] A. Amrouche, B. Boudraa, J. M. Rouvaen "Modèle de prédiction des durées des phonèmes pour la synthèse automatique de la parole en Arabe standard". COMAEL/96 vol. No2. pp12-16. Télémeccen 3-5 déc 1996 Algérie.
- [5] D. O'shaughnessy, L. Barbeau, D. Bernardi D. Archambault. "Diphone Speech Synthesis". Speech Communication. No 7, pp 55-65. 1988.
- [6] D. H. Klatt. "Linguistic uses of segmental duration in English: Acoustic and perceptual evidence". J.Acoust. Soc. Am., Vol.59. No.5 pp.1208-1221. 1976
- [7] T. H. Barbe. "méthodologie et outils pour la mise en oeuvre automatique d'une synthèse de la parole de haute qualité". Thèse de doctorat institut nationale polytechnique de GRENOBLE, France nov1988.
- [8] M. Guerti "Méthodes et techniques d'analyse et de synthèse de la parole." S.S.A. pp 135-151. BLIDA. Algérie, 13-15 Déc 1992.
- [9] K. Bartkhova. "Nouvelle approche dans le modèle de prédiction de la durée segmentale". Actes 14-ème J.E.P. S.E.F. PARIS. France pp188-191. 1985.
- [10] F. Emerard. "Synthèse par Diphones et traitement de la Prosodie". Thèse de docteur 3ème cycle Grenoble mars 1977.
- [11] R. Boite "Traitement de la parole". Presse Polytechniques Romandes (Lausanne 1987).
- [12] M. Guerti "Contribution à la synthèse de la parole en Arabe Standard, synthèse par diphones et technique de prédiction Linéaire". Thèse de Magister, ILP-ALGER Algiers.
- [13] M.L.BENZAOU, M.GUERTI "Durées intrinsèques des voyelles de l'Arabe Standard". Pp120-125. SNAS'99. 9-10/11/99 Annaba, Algérie.
- [14] M.L.BENZAOU, M.GUERTI "Phonemic Duration of Plosive Consonants Specifics to the Standard arabic". CATAEE'99. 19-20/10/99 Jordanie.

Etude et Réalisation d'un Convertisseur AC/DC à Facteur de Puissance Unitaire

M.O. Mahmoudi^(*), O. Stihi , S. Hamdani

Laboratoire de Commande des Processus, Ecole Nationale Polytechnique
BP 182, Hassen Badi, El Harrach, 16200, Alger.

ملخص :

نقدم في هذا العمل دراسة مفصلة لمحوّل (متناوب / مستمر) أحادي الطور له معامل استطاعة أعظمي ، وكيفية التحكم المستعملة والتي تتبنى استراتيجيتين لتعديل عرض النبضة :

1. استراتيجية التعديل الجيبي
2. استراتيجية ديلتا (Δ)

نبدأ بتوضيح مبدأ التشغيل وكيفية وضع النموذج والقيام بالتحكم في المحوّل ، ثم نعطي النتائج النظرية التي حصلنا عليها عن طريق المحاكاة الرقمية وباستعمال الاستراتيجيتين السابقتين ، كما نتطرق بالتحليل لتأثير معاملات التعديل على تشغيل المحوّل.

تمت هذه الدراسة عن طريق المحاكاة ثم قيمناها تجريبيا بإعداد وتركيب نموذج تطبيقي في المعبر.

Résumé :

Ce travail est consacré à l'étude d'un convertisseur AC/DC monophasé à facteur de puissance unitaire commandé par deux stratégies de modulation de largeur d'impulsions, à savoir :

- 1- la stratégie triangulo-sinusoidale SPWM,
- 2- la stratégie delta DPWM.

Ainsi, après avoir décrit le principe de fonctionnement, sa modélisation et la commande du convertisseur, nous présentons les résultats théoriques obtenus par simulation numérique en appliquant les deux stratégies, ensuite nous analysons l'influence des paramètres de modulation sur le comportement du convertisseur vis à vis de ces stratégies. L'étude est menée par simulation puis validée expérimentalement sur un banc d'essai conçu et réalisé au laboratoire.

Abstract :

This work deals with a study of AC/DC converter with unit power factor controlled by two PWM strategies, that are:

- 1- the SPWM sinusoidal modulation,
- 2- the DPWM delta strategy.

Thus, after having described the principle of working, its modeling and the control of the converter, we present the theoretical results given by simulation while applying the two strategies, we analyze the influence of modulation parameters on the behavior of converter relative to these strategies. The study is conducted by simulation and then validated by an experimental prototype realized in our laboratory.

^(*) e-mail: momahmoudi@yahoo.fr

Mots clés

Convertisseur AC/DC, Stratégie triangulo-sinusoidale, Stratégie delta, Facteur de puissance.

1. Introduction

L'alimentation des systèmes à vitesse variable nécessite l'emploi des convertisseurs AC/DC à diodes ou thyristors comme étage d'alimentation des onduleurs. Cependant, ces redresseurs génèrent un fort taux d'harmonique dans le réseau entraînant des distorsions dans l'onde de tension, ce qui conduit à la détérioration du facteur de puissance du coté réseau [1],[2].

Pour éviter ces perturbations, on s'oriente de plus en plus vers le remplacement des redresseurs classiques par des convertisseurs à modulation de largeur d'impulsions capables de :

- imposer une forme de courant sinusoïdale quel que soit le type de charge ;
- contrôler le facteur de puissance du réseau ;
- assurer la réversibilité fonctionnelle de l'installation sans recourir à un pont auxiliaire contrairement aux convertisseurs classiques [3], [4].

Dans ce travail, nous proposons l'étude d'un convertisseur AC/DC monophasé à MLI permettant de satisfaire ces exigences. Ainsi, après avoir modélisé le fonctionnement du convertisseur, nous avons étudié l'influence de deux stratégies de modulation sur son comportement. Il s'agit des stratégies triangulo-sinusoidale et à hystérésis.

L'étude est menée par simulation puis validée expérimentalement sur un banc d'essai conçu et réalisé au laboratoire.

2. Modélisation du système étudié

La figure 1 montre le schéma de principe et indique les notations adoptées. Le convertisseur est formé de quatre interrupteurs K_i ($i=1,4$) commandés en courant. Chacun contient un transistor bipolaire T_i et une diode D_i montée en antiparallèle avec le transistor.

Le convertisseur est monté directement à la sortie du transformateur d'entrée.

On désigne par L_r l'inductance totale du réseau et des fuites de transformateur ramenée au secondaire de celui-ci. Dans ces conditions, v_r est la tension supposée sinusoïdale et constante du réseau ramenée au secondaire. Une capacité C est placée à la sortie du convertisseur permettant ainsi le filtrage de la tension de sortie. Le convertisseur est alimenté par une source de courant à cause de l'inductance L_r , et débite sur un récepteur de tension à cause de la capacité C .

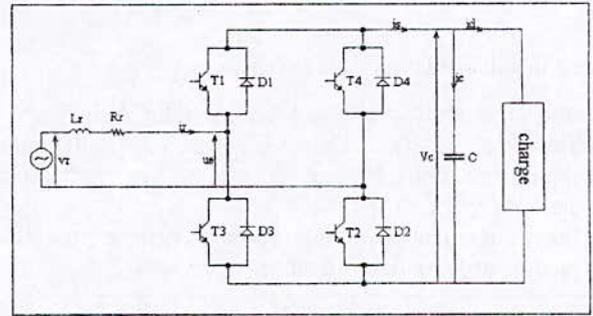


Figure 1: Schéma de principe du convertisseur AC/DC

Soit i_{r1} et u_{e1} les fondamentaux du courant i_r et de la tension u_e . En négligeant $R_r i_{r1}$ devant $L_r \omega i_{r1}$, on peut écrire :

$$\hat{v}_r = \hat{u}_{e1} + jL_r \omega \hat{i}_{r1} \tag{1}$$

\hat{v}_r , \hat{u}_{e1} , \hat{i}_{r1} sont des représentations complexes de v_r , u_{e1} et i_r respectivement. L'équation (1) traduit le diagramme vectoriel de la figure 2-a.

D'ordinaire, On cherche à mettre le fondamental du courant d'entrée i_{r1} en phase avec la tension v_r , le diagramme vectoriel est alors celui de la figure 2-b.

Dans ce cas, on peut écrire :

$$U_{e1} = \sqrt{V_r^2 + (L_r \omega I_{r1})^2} \tag{2}$$

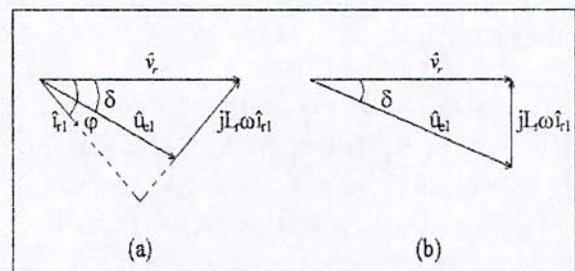


Figure 2 : Diagramme vectoriel.

avec :

$$\text{tg } \delta = \frac{L_r \omega I_{r1}}{V_r} \tag{3}$$

δ désigne le déphasage du fondamental de la tension d'entrée u_{e1} par rapport à la tension v_r . Pour une commande en modulation de largeur d'impulsions δ est le déphasage de la référence servant à déterminer les instants de commutation des interrupteurs, par rapport à la tension d'alimentation.

Le sens de déphasage dépend du sens d'écoulement de puissance :

- Il s'agit d'un déphasage arrière si la source fournit de la puissance à la charge à travers le convertisseur.
- Il s'agit d'un déphasage avant lorsque la charge renvoie de la puissance à la source (fonctionnement onduleur).

le fondamental de la tension d'entrée peut être donné par la relation :

$$U_{e1} = r \frac{V_{c0}}{\sqrt{2}} \quad (4)$$

V_{c0} est la valeur moyenne de la tension v_c .

Comme le rapport cyclique r est inférieur à 1, V_{c0} est supérieure à $\sqrt{2} U_{e1}$. Contrairement aux redresseurs classiques, ce convertisseur fonctionne en redresseur élévateur de tension [5].

Le fonctionnement réversible du convertisseur peut être représenté par le modèle suivant :

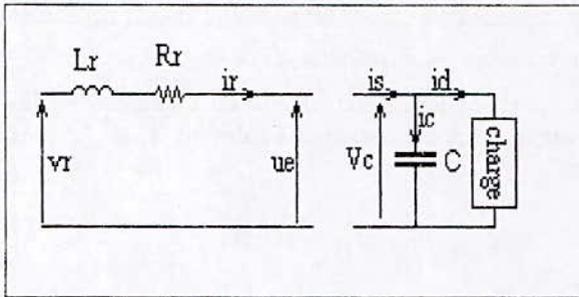


Figure 3 : Modèle du convertisseur.

Cela nous permet d'exprimer la relation entre les grandeurs d'entrée et celles de sortie du convertisseur par les équations suivantes :

$$\begin{cases} u_e = K_f v_c \\ i_s = K_f i_r \end{cases} \quad (5)$$

K_f est la fonction de commande du convertisseur, elle dépend des séquences de fermeture et d'ouverture des interrupteurs K_i .

$$K_f = \begin{cases} 1 & \text{si } (K_1, K_2) \text{ fermés et } (K_3, K_4) \text{ ouverts} \\ 0 & \text{si } (K_1, K_4) \text{ fermés et } (K_3, K_2) \text{ ouverts} \\ & \text{ou } (K_1, K_4) \text{ ouverts et } (K_3, K_2) \text{ fermés} \\ -1 & \text{si } (K_1, K_2) \text{ ouverts et } (K_3, K_4) \text{ fermés} \end{cases} \quad (6)$$

Les équations des courants et des tensions à l'entrée et à la sortie du convertisseur sont :

$$\begin{aligned} \frac{di_r}{dt} &= \frac{v_r - R_r i_r - u_e}{L_r} \\ \frac{dv_c}{dt} &= \frac{i_s - i_d}{C} \end{aligned} \quad (7)$$

i_d est le courant absorbé par la charge, son équation dépend de la nature de celle-ci.

Pour une charge résistive R_d :

$$i_d = \frac{v_c}{R_d} \quad (8)$$

Pour une charge inductive passive (L_d, R_d) :

$$\frac{di_d}{dt} = \frac{v_c - R_d i_d}{L_d} \quad (9)$$

Pour une charge inductive active (L_d, R_d, E) :

$$\frac{di_d}{dt} = \frac{v_c - R_d i_d - E}{L_d} \quad (10)$$

3. Application des stratégies de modulation

Afin de montrer les performances du convertisseur en mode redresseur et ses réponses vis à vis de sa commande, nous avons fait une simulation numérique en appliquant deux techniques de modulation de largeur d'impulsion, il s'agit de la stratégie triangulo-sinusoidale et de la stratégie delta.

3.1. La stratégie triangulo-sinusoidale

Le principe de cette stratégie consiste à la comparaison d'un signal triangulaire v_p de fréquence f_p dit onde porteuse, avec un signal v_m dit onde de référence ou modulante qui porte les caractéristiques de forme et de fréquence de l'onde désirée à l'entrée du convertisseur soit u_e . Cette stratégie permet d'avoir une valeur moyenne de la tension variable en fonction de la référence.

La technique de commande triangulo-sinusoidale telle qu'elle est définie permet un réglage par les deux paramètres :

- L'indice de modulation noté m qui influe sur le contenu harmonique de l'onde désirée;
- Le rapport de réglage r qui influe directement sur la valeur efficace de la tension u_e ;

La figure 4 montre le principe de la stratégie triangulo-sinusoidale et la tension u_e obtenue à l'entrée du convertisseur.

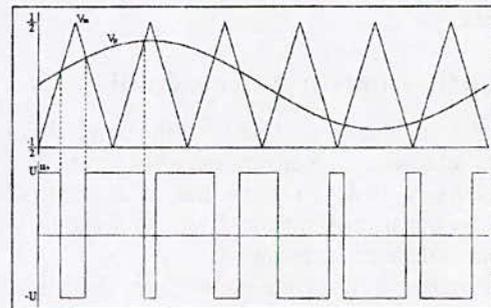


Figure 4: Principe de la stratégie triangulo-sinusoidale

Nous avons simulé le fonctionnement du convertisseur débitant dans une charge inductive de paramètres: $r=0.7$, $m=18$, $f=50\text{Hz}$, $V_{ref}=50\text{V}$, $C=4500\mu\text{F}$, $R_r=5\Omega$, $L_r=0.02\text{H}$, $R_d=80\Omega$, $L_d=0.5\text{H}$. Les résultats obtenus par simulation numérique sont illustrés par la figure 5.

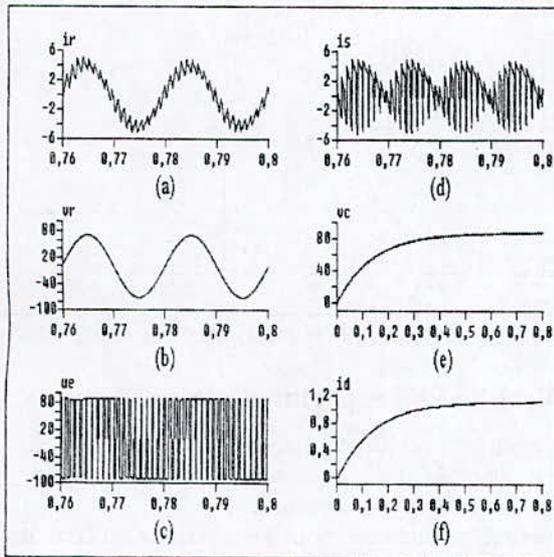


Figure 5 : Résultats de simulation obtenus par la stratégie triangulo-sinusoidale.

La figure 5-a montre le courant i_r au régime permanent, nous remarquons que ce dernier est sinusoïdal et en phase avec la tension v_r , après le décalage de l'onde porteuse v_p d'un angle δ par rapport à la tension v_r . Il présente une forme ondulée selon les séquences de commutation des interrupteurs.

La figure 5-c illustrent la tension u_e à l'entrée du convertisseur, on voit qu'elle a une forme de peigne qui prend les deux valeurs $+v_c$ et $-v_c$.

La figure 5-d montre le courant à la sortie du convertisseur, nous remarquons qu'il peut prendre des valeurs négatives, donc le convertisseur peut renvoyer de la puissance de la charge vers le réseau d'alimentation, ce qui justifie sa réversibilité fonctionnelle.

Les figures 5-e et 5-f montrent la tension v_c et le courant absorbé par la charge. Nous voyons que la tension augmente pour atteindre une valeur constante. Le temps de montée de la tension dépend de la capacité C placée à la sortie du convertisseur, il croît avec l'augmentation de cette dernière. Le courant absorbé par la charge à pratiquement la même allure que la tension v_c .

3.2. La stratégie delta

La figure 6 montre le convertisseur AC/DC commandé par cette technique.

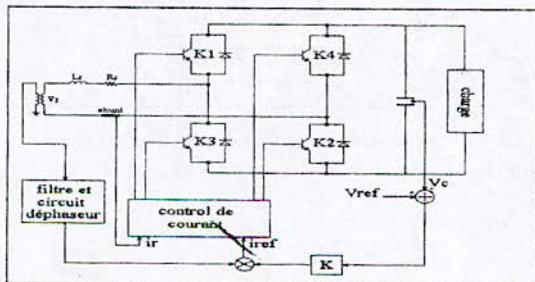


Figure 6 : Schéma de principe du convertisseur commandé par la stratégie delta.

La consigne du courant sinusoïdal est prise à travers un transformateur abaisseur. Un déphaseur permet d'avoir un facteur de puissance variable en introduisant un déphasage ϕ à la consigne. Pour assurer et contrôler le débit de puissance demandée par la charge, la tension v_c est mesurée et comparée avec une tension de référence V_{ref} . L'erreur obtenue est multipliée par un gain K et la consigne sinusoïdal, pour former le courant de référence i_{ref} donné par :

$$i_{ref} = K(V_{ref} - v_c) \sin(\omega t + \phi) \quad (11)$$

La technique consiste à forcer le courant d'entrée i_r , à osciller autour du courant de référence i_{ref} dans une fenêtre de largeur h réglée appelée bande d'hystérésis [6].

La figure 7 montre le principe de cette technique.

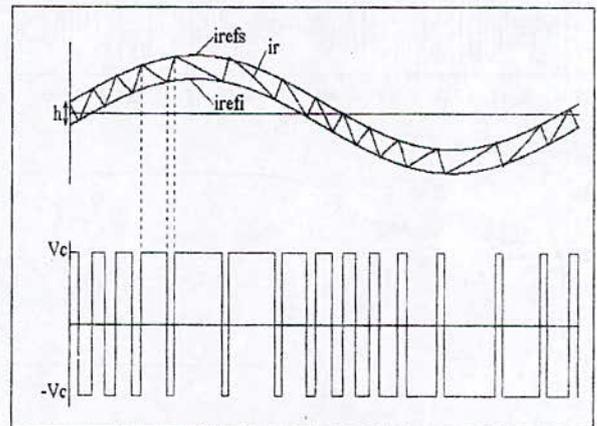


Figure 7 : Principe de la stratégie delta

Les séquences de fermeture et d'ouverture des interrupteurs sont fonction des états actuels et précédents du courant i_r , par rapport aux limites de la bande hystérésis définies par :

$$\begin{aligned} i_{refh} &= i_{ref} + \frac{h}{2} \\ i_{refl} &= i_{ref} - \frac{h}{2} \end{aligned} \quad (12)$$

La forme de la tension d'entrée u_e est déterminée par l'intersection du courant i_r avec les limites de la bande. Nous avons simulé le fonctionnement du convertisseur commandé par cette stratégie dont les paramètres sont : $V_{ref}=95V$, $h=0.5A$, $L_r=0.01H$, $R_r=0.5\Omega$, $C=10000\mu F$, $R_c=50\Omega$, $L_c=0.5H$, $f=50Hz$, $K=1$, $\phi=0$.

La figure 8 illustre les formes d'ondes des différentes grandeurs d'entrée et de sortie obtenues par simulation numérique.

Dans la figure 8-a, nous remarquons que le courant i_r est sinusoïdal, en phase avec la tension d'alimentation v_r , il suit parfaitement la référence du courant imposé et il oscille à l'intérieur de la bande d'hystérésis.

Les résultats obtenus pour les autres grandeurs sont identiques à ceux obtenus par la stratégie triangulo-sinusoidale.

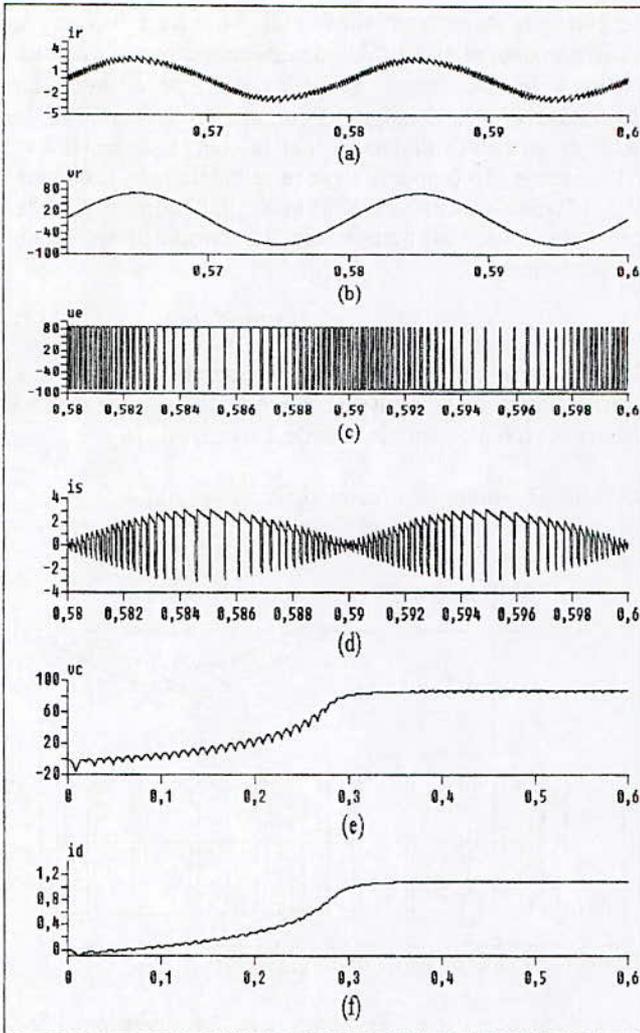


Figure 8 : Résultats de simulation pour la stratégie delta

Pour ce type de commande, la charge initiale du condensateur influence énormément sur le transitoire du courant i_r . Les figure 9-a et 9-b montrent respectivement le courant i_r et la tension v_c dans le cas d'un condensateur non chargé initialement.

Pendant le régime transitoire, le courant i_r atteint des valeurs très élevées. Cela peut être expliqué par le fait qu'au début, la tension v_c est nulle ce qui rend l'erreur ($V_{ref} - v_c$) très grande. Le courant i_r suivant une référence d'amplitude élevée, atteint des surintensités considérables. Au fur et à mesure que la tension v_c augmente, l'amplitude de la référence diminue et le courant i_r tend à s'amortir pour prendre finalement une forme sinusoïdale.

Le choix d'un condensateur chargé initialement permet de réduire les surintensités du courant comme le montre les figure 9-c et 9-d.

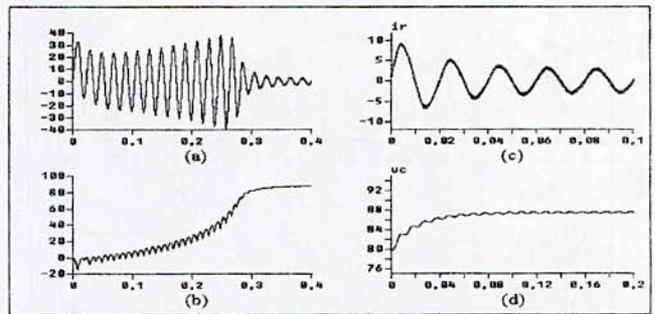


Figure 9 : Influence de la charge initiale de condensateur

4. Vérification expérimentale

Pour mettre en évidence l'étude théorique faite sur le convertisseur AC/DC, nous avons conçu et réalisé les différents circuits qui le constituent.

Le montage pratique que nous avons réalisé au laboratoire contient deux principaux circuits.

- Un circuit de puissance formé de quatre bras, chacun contient un montage Darlington constitué de deux transistors bipolaires et de quatre diodes montées en antiparallèle. On mène chaque bras du convertisseur par ses circuits d'aide à la commutation permettant de réduire les surintensités à la fermeture et les surtensions à l'ouverture.

Les signaux de commande générés seront amplifiés grâce à des circuits de bases convenables liés aux transistors.

- Pour la commande du convertisseur, nous avons réalisé deux circuits : le premier pour la stratégie triangulo-sinusoidale et le second pour la technique delta. Le principe de réalisation de ces circuits est illustré par les figures 10 et 11 respectivement.

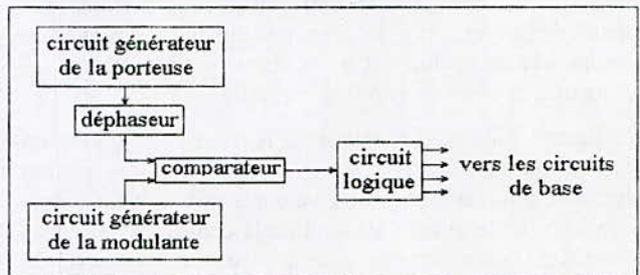


Figure 10 : Principe de circuit de commande de la stratégie triangulo-sinusoidale.

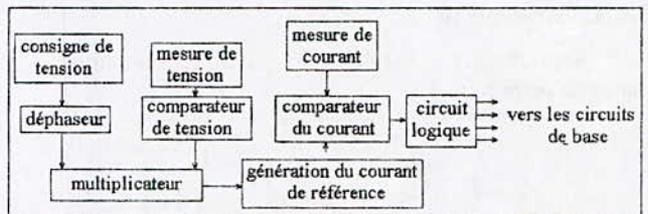


Figure 11 : Principe de circuit de commande de la stratégie delta.

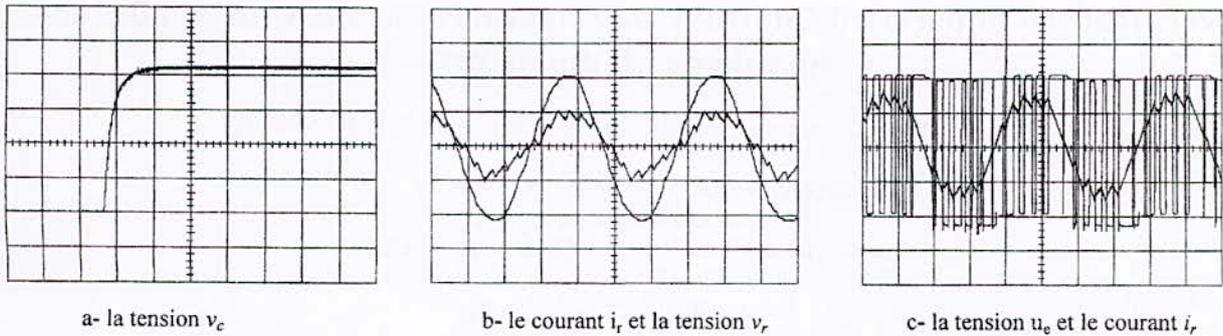
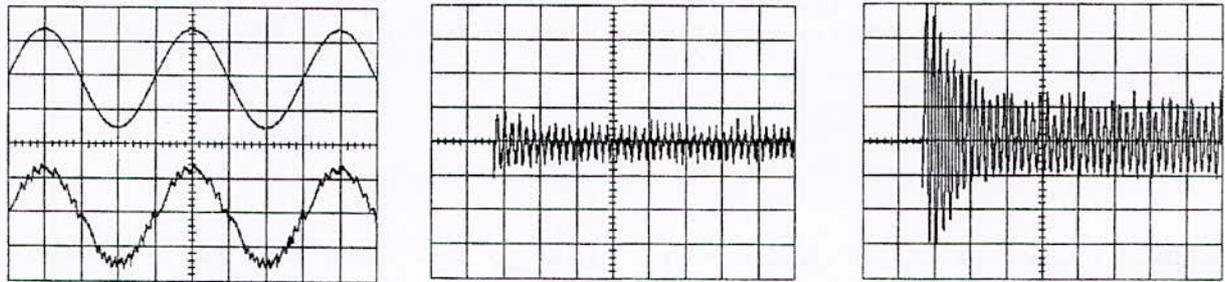
a- la tension v_c b- le courant i_r et la tension v_r c- la tension u_e et le courant i_r

Figure 12 : Résultats expérimentaux obtenus par l'application de la stratégie triangulo-sinusoidale

a- le courant i_r , mesuré et la référence

b- démarrage avec condensateur chargé

c- démarrage avec condensateur vide

Figure 13 : Résultats expérimentaux obtenus par l'application de la stratégie delta

Après réalisation des différents circuits, le convertisseur a été testé au laboratoire, il fonctionne en redresseur en débitant sur une charge passive formée d'une résistance et d'une inductance. Les résultats expérimentaux illustrés par les figures 12 et 13 ont été relevés sur un oscilloscope numérique (TEK 2211) lié à un PC.

Pour la stratégie triangulo-sinusoidale, nous avons pris un rapport cyclique égale à 0.7 et un indice de modulation égal à 12. La valeur efficace de la tension délivrée par l'autotransformateur est prise égale à 50V. la capacité de filtrage à la sortie du convertisseur est de 4500 μ F.

La figure 12-b montre clairement que le courant i_r est en phase avec la tension v_r , donc le convertisseur fonctionne à facteur de puissance unitaire. La tension d'entrée suit parfaitement la commande imposée, La tension v_c croit de zéro à une valeur finale constante, le temps de montée dépend des valeurs des éléments du filtrage, il augmente lorsque la capacité C et l'inductance L_r placées prennent des valeurs élevées.

Pour la même charge, les résultats obtenus par l'application de la technique delta montrent que le courant i_r est en phase avec la tension d'alimentation en suivant fidèlement le courant de référence.

L'influence de la charge initiale du condensateur a été vérifiée également. Les résultats illustrés par la figure 12-b et 12-c montrent que le courant i_r présente des surintensités appréciables pour un condensateur vide initialement, ces surintensités peuvent être diminuées en travaillant avec un condensateur chargé.

5. Conclusion

Un convertisseur AC/DC monophasé a été étudié et réalisé. Ce convertisseur permet un fonctionnement à un facteur de puissance unitaire en introduisant un déphasage

de la porteuse par rapport à la tension d'alimentation pour la stratégie triangulo-sinusoidale ou en choisissant une référence de courant en phase avec la tension d'alimentation pour la technique delta. Les résultats obtenus par simulation ou par des essais pratiques sont très proches, ils montrent que :

- Les amplitudes des grandeurs d'entrée et de sortie du convertisseur dépendent du rapport cyclique pour la stratégie triangulo-sinusoidale et du courant de référence pour la stratégie delta.
- La formes d'onde de ces grandeurs sont affectées par l'indice de modulation pour la stratégie triangulo-sinusoidale et par la largeur de la bande d'hystérésis pour la stratégie delta.
- Pratiquement pour éviter des surintensités au régime transitoire, il faut prévoir un condensateur chargé initialement.
- La réversibilité fonctionnelle du convertisseur est assurée pour les deux commandes appliquées.

6. Références

- [1] J. W. Dixon, A. B. Kulkarni et M. Nishimoto, " Characteristics of a controlled-current PWM rectifier-inverter link", IEEE Trans. Indus. Appl, Vol. IA-23, No.6, Nov./Dec. 19 88, pp. 1022-1028.
- [2] O. Stihl et B. T. Ooi, "A single phase controlled-current PWM rectifier", IEEE Trans. in Power Electronics. Vol. 3, No 4, Oct. 1988, pp. 453-459.
- [3] E. P. Wiechmann, P. D. Ziogas et V. R. Stefanivic "A novel bilateral power conversion scheme for variable frequency static power supplies", IEEE trans. Indus. Appl, Vol IA-21, No. 5, Sept./Oct. 1985, pp. 1226-1233.
- [4] B. T. Ooi, J. C. Salmon, J. W. Dixon et A. B. Kulkarni. "A three-phase controlled-current PWM converter with leading power factor", IEEE Trans. Indus. Appl. Vol. IA-23, No:1, jan/fev 87, pp 78-84
- [5] G. Segulier, F. Labrique. "Les convertisseurs de l'électronique de puissance. Tome IV la conversion continu - alternatif". Edition : TEC & DOC, Lavoisier, 1989 chapitre 2 pp 119-177.

Decentralized Supervised Control Using Neural Networks with Application to Robotic Manipulators

D. Boukhetala, F. Boudjema, M. Djebiri, M. S. Boucherit & N. K. M'Sirdi*

Laboratoire de Commande des Processus
Ecole Nationale Polytechnique,
10 avenue Hassen Badi, BP 182 El-Harrach, ALGERIE.
Tél: (213)(02) 52 14 94 Fax: (213) 02 52 29 73.
Email : dboukhetala@yahoo.fr

*Laboratoire de Robotique de Paris
Université de Versailles, Saint Quentin,
10-12 Avenue de l'Europe, Velizy, 78140, France

ملخص— يخصص هذا المقال التحكم عن طريق المعاينة باستعمال الشبكات العصبونية الاصطناعية لقانون تحكم موجود. في هذا العمل تمت معاينة قانون تحكم تلاؤمي لامركزي وطبقت هذه الطريقة على ذراع آلية حيث تمّ تبيان قدرة المنظم العصبوني على معاينة قانون تحكم موجود رغم أن وسائطه متلائمة ومعاملات الشبكات العصبونية ثابتة. الطريقة المستعملة أثبتت ظلاعة التحكم عند حدوث اضطرابات خارجية.
كلمات مفاتيح— شبكات عصبونية اصطناعية، تحكم عن طريق المعاينة، تحكم لامركزي، تحكم تلاؤمي

Abstract—This paper presents a control method using neural networks based on the supervision of an existing control law. In our work, an adaptive controller is supervised and a decentralized scheme is used. Application of this control technique to robotic manipulators shows the effectiveness of the neural controller to mimic the existing controller despite that this later has adjustable parameters and the neural network has fixed weights. The neural controller demonstrates also robust behavior toward external disturbances.

Key words—Artificial neural networks, supervised control, decentralized control, adaptive control .

Résumé: Cet article concerne la commande supervisée basée sur les réseaux de neurones artificiels d'une loi de commande existante. Dans ce travail, une commande adaptative est supervisée en utilisant un schéma décentralisé. Une application de cette technique au robot manipulateur montre la faisabilité du régulateur neuronale pour imiter un régulateur donné. Malgré le caractère adaptatif des paramètres du régulateur, le réseau de neurone utilisé est à poids fixes. L'approche neuronale proposée a montré la robustesse de la commande vis à vis des perturbations externes.

Mots clés— Réseaux de neurones artificiels, Commande supervisée, commande décentralisée, commande adaptative.

I. INTRODUCTION

THE RECENT INTEREST in artificial neural networks with their attractive features, such as nonlinearity, parallel processing, learning and generalization capabilities, has given solution to many control problems especially those related to nonlinear systems. In fact, traditional techniques for analysis and synthesis of nonlinear controllers exist (phase plane method, linearization techniques,...) but they are limited to specific classes of nonlinear systems [4]. The ability of neural networks to approximate arbitrary nonlinear mappings has become a powerful alternative to overcome restrictions of traditional techniques.

Application of neural networks to control field can be divided into some major methods, such as supervised control, inverse control, neural adaptive control, backpropagation through time and adaptive critics methods [7].

In supervised control, an already established controller (human or automated) must be provided at first. Then, a neural network learns the mapping from input to output of this controller. In other words, supervised control is nothing else than copying an existing controller.

Widrow's early broom balancer in the 1960's was the first working example of supervised control [3]. Widrow refers to this control as a method of constructing a neural expert system by acquiring skills from an existing expert [7]. Guez and Selinsky [2] provided a more sophisticated updated version of supervised control. They demonstrated the capability of a feedforward neural network to learn how to control a dynamic system from examples generated by linear and nonlinear control laws and a human teacher. Recently, Hamzi et al. [6] resumed the work of Guez and Selinsky but, instead of using a static feedforward neural network, they used a dynamical neural network. As a result, they obtained a neural controller with better robustness and reduced dimension and the number of iterations in the training phase was less than that required by the static network in [2]. Jorgenson [7] also gave an excellent example of supervised control in aviation.

But one might question the utility of this method on the grounds that if there already exists an effective controller, why would it be useful to have another one in the form of a network? Several answers can be evoked [2], [7]:

1) The existing controller can be a device impracticable to use such as human operator. Guez and Selinsky obtained excellent results by supervising a control law generated by a human teacher.

2) As computations in a neural network are done in parallel, decision making and control are very fast.

3) Control laws need not be explicitly stated as learning can be accomplished through examples.

4) Distributed representation of information in neural networks provides better robustness towards disturbances and prevents information degradation.

5) Natural robustness to unmodeled parameters due to network generalization capability.

6) The neural controller may be able to form a control law based on a state representation of the controlled system that is easier to measure than the representation required by the existing controller.

Supervised control has not been widely used and studied as it has been the case with neural adaptive control and other techniques. In this paper a new approach to control multi-joint manipulators by using supervised control is described. It is based on a decentralized scheme where the system to be controlled is formed of interconnected subsystems. For each subsystem an independent supervised control is synthesized and each neural controller has access only to local information. Simulation results of the decentralized supervised control applied to a CLASS 4 manipulator ([5], are given. A comparison between the neural controller and the existing controller is shown and the effectiveness of the neural controller is clarified.

II. SUPERVISED CONTROL

The supervised control is generally implemented in two stages:

-Training stage: in this stage, the neural network learns the function input/output of the existing controller. When this later controls the system, training examples are generated. Then a supervised learning algorithm is used in order to train the neural network.

-Control stage: after training, the network controls the system and the existing controller is removed.

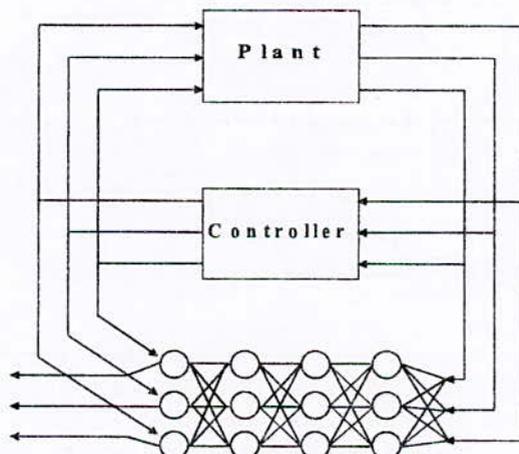


Fig. 1. Supervised control.

III. DECENTRALIZED ADAPTIVE CONTROL

In this paper, the process to be controlled is a multi-joint robot described by the dynamic equation

$$M(q)\ddot{q} + N(q, \dot{q}) + G(q) + H(\dot{q}) = T \quad (1)$$

Where

$q(t)$ is the $n \times 1$ vector of joint positions

$\dot{q}(t)$ is the $n \times 1$ vector of joint velocities

$T(t)$ is the $n \times 1$ vector of applied joint torques

$M(q)$ is the $n \times n$ manipulator inertia matrix

$N(q, \dot{q})$ is the $n \times 1$ coriolis and centrifugal torque vector

$G(q)$ is the $n \times 1$ gravity loading vector

$H(\dot{q})$ is the $n \times 1$ frictional torque vector.

A decentralized adaptive controller is to be supervised here. The adaptive independent joint controller dedicated to the i th joint is illustrated in Fig.2 and described by the following equation [1]

$$T_i(t) = f_i(t) + [k_{i0}(t)e_i(t) + k_{i1}(t)\dot{e}_i(t)] + [p_{i0}(t)q_{ri}(t) + p_{i1}(t)\dot{q}_{ri}(t) + p_{i2}(t)\ddot{q}_{ri}(t)] \quad (2)$$

where

$e_i(t) = q_{ri}(t) - q_i(t)$ is the position tracking error of joint i

$q_{ri}(t)$ desired trajectory

$q_i(t)$ position of joint i

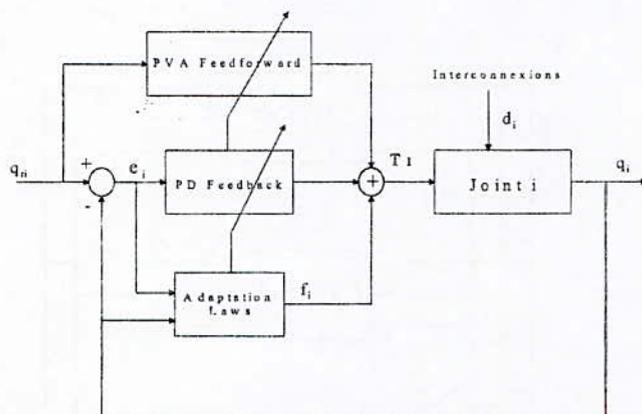


Fig. 2. Independent adaptive control for joint i

This control law is composed of three terms:

1) The term $f_i(t)$ is an auxiliary signal synthesized by the adaptation scheme to improve the tracking performance and partly compensate for the disturbance d_i .

2) The term $[k_{i0}(t)e_i(t) + k_{i1}(t)\dot{e}_i(t)]$ which is due to the adaptive PD feedback controller.

3) The term $[p_{i0}(t)q_{ri}(t) + p_{i1}(t)\dot{q}_{ri}(t) + p_{i2}(t)\ddot{q}_{ri}(t)]$ which is due to the adaptive position-velocity-acceleration feedforward controller.

The controller adaptation laws, which ensure asymptotic tracking of the desired trajectory, are based on the weighted error $r_i(t)$ as follows

$$r_i(t) = w_{pi}e_i(t) + w_{vi}\dot{e}_i(t) \quad (3)$$

to adjust the auxiliary signal

$$f_i(t) = f_i(0) + \delta_i \int_0^t r_i(t) dt + \rho_i r_i(t) \quad (4)$$

to adjust the feedback controller gains

$$k_{ij}(t) = k_{ij}(0) + \alpha_{ij} \int_0^t r_i(t) e_i^{(j)}(t) dt + \beta_{ij} r_i(t) e_i^{(j)}(t), \quad j = 0, 1 \quad (5)$$

to adjust the feedforward controller gains

$$p_{ij}(t) = p_{ij}(0) + \gamma_{ij} \int_0^t r_i(t) q_{ri}^{(j)}(t) dt + \lambda_{ij} r_i(t) q_{ri}^{(j)}(t), \quad j = 0, 1, 2 \quad (6)$$

where

$\delta_{ij}, \alpha_{ij}, \gamma_{ij}$ are any positive scalar gains

$\rho_i, \beta_{ij}, \lambda_{ij}$ are zero or any positive gains

w_{pi}, w_{vi} are positive scalars weighting factors, which reflect the relative significance of the position and velocity errors in forming $r_i(t)$.

By using the equations (3), (4), (5) and (6) the control law can be written as the combination of a PID feedback controller and PD² feedforward controller as

$$T_i(t) = T_i(0) + \left[\bar{k}_{ip} + \bar{k}_{il} \int_0^t dt + \bar{k}_{iv} \frac{d}{dt} \right] e_i(t) + \left[p_{i0} + p_{i1} \frac{d}{dt} + p_{i2} \frac{d^2}{dt^2} \right] q_{ri}(t) \quad (7)$$

where

$$\begin{cases} \bar{k}_{ip} = k_{i0} + \rho_i w_{pi} + \delta_i w_{vi} \\ \bar{k}_{il} = \delta_i w_{pi} \\ \bar{k}_{iv} = k_{il} + \rho_i w_{vi} \end{cases} \text{ are the adjustable gains of}$$

the PID

$$U_i(0) = f_i(0)$$

IV. NEURAL NETWORK CONTROLLER

In order to supervise the control law presented above we have to use a neural network for each joint i . The local neural controller will be trained to mimic the adaptive independent joint controller shown in Fig.2 and which is described by (7).

The structure of the neural network is shown in Fig.3. It is a three-layered network consisting of 5 inputs, 10 hidden units and 1 output. According to (7), the inputs are chosen to be : $e_i(t), \dot{e}_i(t), q_{ri}(t), \dot{q}_{ri}(t), \ddot{q}_{ri}(t)$. This will provide the network with sufficient amount of information in order to perform his task.

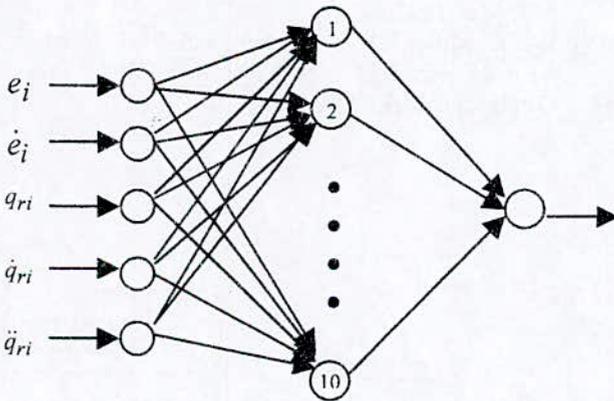


Fig. 3. Neural network structure.

Fig.4 shows the control scheme for joint i when the training of the network has been accomplished. A constant gain K_i is introduced on the feedforward loop in order to adjust, if needed, the output of the neural controller.

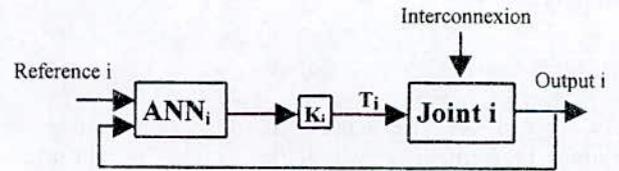


Fig. 4. Independent supervised control of the joint i .

V. SIMULATION RESULTS

The manipulator used in simulations is a three-joint CLASS 4 manipulator [5]. The decentralized adaptive control is first applied and training examples are generated. Table I lists the different gains used in this first stage.

TABLE I
ADAPTATION GAINS

	w_p	w_v	δ	α_0	α_1	γ_0	γ_1	γ_2	β_0	β_1	λ_0	λ_1	λ_2
Joint 1	400	200	50	100	200	5	5	5	1	1	1	1	1
Joint 2	80	40	50	60	100	10	10	10	1	1	1	1	1
Joint 3	80	40	50	60	100	3	3	3	1	1	1	1	1

Three neural networks are used, all with the same structure (Fig. 3). Training has been performed by using the backpropagation algorithm over 150 examples. The learning rate η was 0.2 for the first layer and 0.1 for the second layer. Weights were arbitrary initialized between -0.5 and 0.5. The constant gain K_i mentioned above (see Fig. 4) were taken as 4, 1 and 1 for the first, second and third joint respectively.

Training took 1200, 3600 and 600 trial for the ANN 1, 2 and 3 respectively.

Fig. 5 illustrates a comparison between the decentralized adaptive controller and the decentralized neural controller when the CLASS 4 manipulator pick up a $m_0=10$ Kg load with its end effector and drop it at $t=1.5$ s.

As we can see, the neural controller shows good tracking performances which are close to those obtained by the adaptive controller, despite that this later has adjustable parameters and the neural network has only fixed weights. As a result, the neural controller will significantly reduce the number of computations, and then computation time, by removing the adaptation step of the adaptive controller.

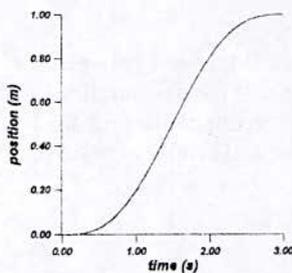
In addition to this, neural network demonstrates robust behavior toward external perturbations. In fact, by changing the loads supported by the end effector, performances of the control scheme remains satisfactory (despite that the network was trained with examples where the changing load was not taken into account).

VI. CONCLUSIONS

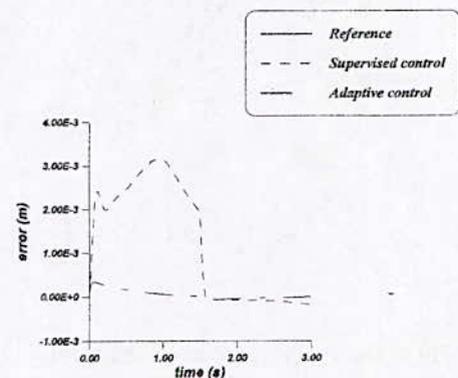
A decentralized supervised neural controller has been presented in this paper. The neural network not only performs the control task as the existing controller does, but also reduces considerably the computation time. This is very important in some control situations where time constraints subsist. The generalization capability of the neural network is also demonstrated in this paper. The neural controller knew how to control the manipulator when the model parameters changed significantly because of the change of the loading.

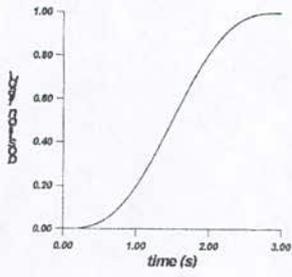
REFERENCES

- [1] H. Seraji, "Decentralized adaptive control of manipulators: theory, simulation and experimentation," *IEEE Trans. on Robotics and Automation*, vol. 5, no. 2, pp. 183-200, Apr. 1989.
- [2] Guez and J. Selinsky, "A trainable neuromorphic controller," *J. Robotic Sys.*, vol. 5, pp. 363-388, Apr. 1988.
- [3] T. Fukuda and T. Shibata, "Theory and applications of neural networks for industrial control systems," *IEEE Trans. Indust. Electronics*, vol. 9, no. 6, pp. 472-489, Dec. 1992.
- [4] K. J. Hunt, D. Sbarbaro, R. Zbikowsky, and P. J. Gawthrop, "Neural networks for control systems-A survey," *Automatica*, vol. 28, no. 6, pp. 1083-1112, 1992.
- [5] D. P. Stoten, "Generalized manipulator dynamics with regard to model reference adaptive control," *Int. J. Control*, vol. 50, no. 6, pp. 2249-2268, 1989.
- [6] Hamzi, S. Labiod, D. Boukhetala, and F. Boudjema, "Supervised control using dynamical neural networks with application to an inverted pendulum," *IFAC-IFIP-IMACS Conference on Control of Industrial Systems*, vol. 2, pp. 570 - 575, May 1997.
- [7] W. T. Miller, R. S. Sutton, and P. J. Werbos, *Neural networks for control*, MIT Press, Cambridge, MA, 1992.

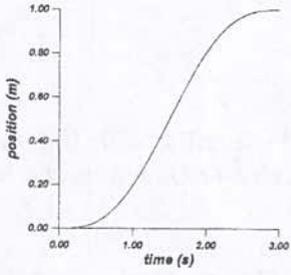
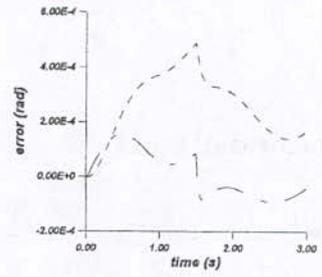


(a)





(b)



(c)

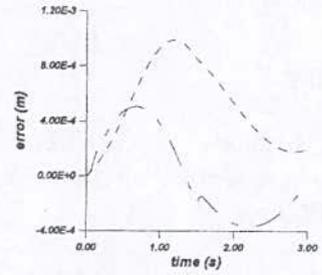


Fig. 5. Comparison between the decentralized adaptive control and the decentralized supervised control
(a) joint 1, (b) joint 2, (c) joint 3

THESAURIA 1998

Thèses de Doctorat d'Etat :

En Electronique

- 1- Baba Ali Ahmed Riadh (18/03) DT : A. Farah
Optimisation des performances temporelles des circuits V.L.S.I.

Thèses de Magister :

En Electronique :

- 1- Djeddou Mustapha (11/02) DT : Mlle Guerti et Mlle Halimi (CDTA)
Conception et réalisation d'un codeur/décodeur de parole à 16 kbits/s à faible retard et à bande étroite (300-3400 Hz).
- 2- Kaidali Zoheir (24/06) DT : B. Derras et Belouchrani
Analyse d'écho lointain dans les systèmes de télécommunications.
- 3- Djouadi Mohamed Said (26/06) DT: D. Berkani
Design d'un quantificateur vectoriel hybride pour le codage d'images.
- 4- Azrar Arab (30/06) DT : R. Aksas
Analyse d'APM avec parasites coplanaires à l'aide de l'équation intégrale résolue dans le domaine spectral.
- 5- Laidi Kamel (27/09) DT : A. Farah
Implémentation orienté Hard ware des réseaux de neurones artificiels. Application à la navigation d'un robot autonome.
- 6- Dahmani Rabea (12/10) DT : D. Akhrib (U. de Blida)
Acquisition et transmission de données numériques-Télémétre.
- 7- Kadri Mohamed (13/10) DT : N. Louam et B. Bouzouia (CDTA)
Elaboration de stratégies de commande neuro-dynamiques adaptatives pour robots manipulateurs.
- 8- Boubakir Chabane (25/10) DT : D. Berkani
Rehaussement de la parole hyperbare.

En Automatique/Electronique de puissance

- 1- Bouchhida Ouahid (23/02) DT : MS. Boucherit et L. Nezli
Réglage d'état multivariable échantillonné et avec observateurs d'une MSAP alimentée en tension.
- 2- Bachir Bouyadjra Rochdi (27/04) DT : MS. Boucherit et D. Boukhetala
Commande adaptative décentralisée des robots manipulateurs.
- 3- Labiod Salim (17/05) DT : MS. Boucherit et H. Chekireb
Commande adaptative par les systèmes Flous : Application aux robots manipulateurs.

- 4-Yeddou Yazid M'hamed (09/06) DT : MS. Boucherit
Etude de synthèse sur les réseaux de neurones et leurs applications.
- 5-Guedouani Rabéa (20/06) DT : EM. Berkouk et Boucherit
Commande d'un onduleur de tension multiniveaux. Application à une MSAP.
- 6- Bouhali Omar (19/10) DT : F. Boudjema et D. Boukhetala
Etude et développement de méthodes de commande adaptative centralisée et décentralisée en utilisant les réseaux de neurones artificiels. Application au robot PUMA 560.
- 7-Bouhassoun El-Hadi (24/10) DT : MO. Mahmoudi et MS. Boucherit
Contribution à la commande par mode de glissement d'une machine synchrone à aimants avec pilotage vectoriel.
- 8- Djehbar Abdelkader (04/11) DT : MO. Mahmoudi et B. Hemic
Sur l'application de commandes adaptatives à une machine asynchrone alimentée en courant avec pilotage vectoriel.

En Electrotechnique (Machines électriques, Réseaux et HT)

- 1- M^{me} Namane (née Sator Dalila) (11/02) DT : A. Mekhaldi et S. Bouazabia
Effet de la discontinuité de la couche de pollution sur le comportement d'isolateurs haute tension, sous fréquence 50 Hz.
- 2- Larbani Méziani (09/06) DT : S. Mekhtoub
Etude des régimes dynamiques du moteur asynchrone triphasé en saturé. Application au court-circuit et au déclenchement.
- 3- Boumaza Salim (22/06) DT : A. Boubakeur
Amorçage aux chocs de foudre des intervalles d'air tige-plan avec terre hétérogène.
- 4- Benkaci Ali Hassen (01/07) DT : A. Hellal et M.Ghezaili
Evaluation de la stabilité statique en temps réel. Application à la chaîne de calcul TRE de la SONELGAZ.
- 5- Mecheri Yacine (26/10) DT : A. Boubakeur
Influence du vieillissement thermique continu sur les propriétés diélectriques et mécaniques du PRC utilisé dans l'isolation des câbles moyenne tension.

THESAURIA 1999

Thèses de Doctorat d'Etat :

En Electronique

- 1- Chohra Amine (16/03) DT : A. Farah et C. Ben mehrez
Planification et contrôle de la navigation des véhicules autonomes intelligents (V.A.I) en environnements dynamiques.
- 2- Benazzouz Djamel (27/06) DT : A. Farah
Modélisation et évaluation des réseaux d'interconnexions dynamiques par les réseaux de Pétri stochastiques.

En Génie Electrique/Electrotechnique

- 1- Mekhaldi Abdelouahab (29/09) DT : A. Beroual
Etude des phénomènes de conduction et de décharge électrique sur les surfaces isolantes polluées sous tension alternative 50 Hz.
- 2- Djemai Zoheir (26/12) D.T : A. Beroual
Détermination de la dimension fractale des décharges électriques radiales sur les surfaces isolantes.

En Génie Electrique/Electronique de puissance

- 1- Mahmoudi Mohand Oulhadj (08/12) DT: F. Boudjema
Sur la commande de la machine asynchrone alimentée en tension. Etude de l'étage d'alimentation de la machine et commande par mode de glissement avec orientation du champ.

Thèses de Magister :

En Electronique

- 1- M^{lle} Izeboudjen Nouma (06/05) DT : A. Farah
Conception et implémentation en FPGA d'un classificateur neuronal d'arythmies cardiaques.
- 2- Kamri Djekidel (12/05) DT : Ait Cheikh
Commande à relais dans les systèmes à retard.
- 3- Senouci Saddok (26/05) DT : A. Zerguerras
Contribution à l'étude d'optimisation d'A.P.M
- 4- Bensaid Hocine (22/05) DT : A. Belouchrani et B. Derras
Synthèse des filtres numériques R.I.I par la méthode des moindres carrés vrais.
- 5- Ould Cheikh Mourad (23/06) DT : Mlle M. Guerti et Mr. M. Halimi
Conception et réalisation d'un codeur/décodeur de la parole à large bande (50-7000 Hz) et à faible débit (13 kbits/s).

- 6- Akroum Hamza (28 /06) DT : D. Berkani et Mme Hamami
Prétraitement et segmentation d'image de document.
- 7- Tounsi Mohamed (29/06) DT : A. Farah
Conception d'architectures digitales unifiées TCD/TCDI-2D. Etude comparative du mode de calcul Half-line de l'arithmétique distribuée.
- 8- Moussaoui Aicha (04/07) DT : M. Mehenni
Etude et réalisation d'une chaîne de télémétrie.
- 9- Flihi Farid (27/07) DT : D. Berkani
Codeur harmonique à bas débit.
- 10- Bouali Nacereddine (26/06) DT : D. Berkani et R. Zergui
Compression du signal électrocardiographique.
- 11- Guerid Abdelkader (27/11) DT : A. Zerguerras et B. Derras
Détection de cibles radar corrélées et noyées dans du clutter.
- 12- Kirat Fouad (07/12) DT : M. Haddadi
Etude et réalisation d'un contrôleur photovoltaïque autonome universel.

En Automatique/Electronique de puissance

- 1- D. Lalili (27/02) DT : MO. Mahmoudi et B. Hemici
Sur les commande non linéaires adaptatives de la machine asynchrone.
- 2- Z. Tidjani (07/03) DT : D. Boukhetala et MS. Boucherit
Commande décentralisée floue appliquée aux robots manipulateurs.
- 3- R. Chibani (27/06) DT : EM. Berkouk et MS. Boucherit
Modélisation de la commande d'un onduleur de tension à cinq niveaux à structure NPC : Application à une conduite d'une machine synchrone à aimants.
- 7- A. Kerboua (03/07) DT : B. Hemici et MS. Boucherit
Commandes et observateurs par mode de glissement. Application à une machine asynchrone.
- 8- A. Tlemçani (22/12/99) DT : H. Chekireb et MS. Boucherit
Sur la commande adaptative par les systèmes Flous. Application à la machine synchrone à aimants.

En Electrotechnique (Machines électriques, Réseaux et HT)

- 1- S. Hamdani (15/05) DT : MO. Mahmoudi et O. Stihi
Etude et réalisation d'un convertisseur AC/DC à facteur de puissance unitaire.
- 2- H. Gheraia (13/06) DT : EM. Berkouk
Modélisation et commande d'un onduleur NPC à sept niveaux. Application à une conduite d'un moteur asynchrone.
- 7- Mezouar (04/07) DT : O. Touhami et R. Ibtouen
Analyse et contrôle des systèmes singulièrement perturbés : Application aux modèles des machines à courant alternatif.
- 4- T. Khaldi (06/07) DT : S. Mekhtoub
Retour réseau sur un auto-amorçage accidentel de la machine asynchrone.

- 5- M. Ladjal (29/11) DT : MEH. Zaim
Contribution à l'étude de la machine asynchrone à rotor massif.
- 6- E. Miliari (07/12) DT : O. Touhami
Commande non linéaire d'une machine asynchrone : Différentes techniques de réglage.
- 7- A. Chouchou (20/12) DT : A. Boubakeur
Détermination des zones d'attraction d'un paratonnerre vertical dans le cas d'une terre discontinue.
- 8- A. Abimouloud (25/12) DT : A. Mekhaldi et M. Tegar
Comportement d'une surface isolante partiellement polluée sous tension alternative 50 Hz.
- 9- A. Nouar (08/12) DT : A. Boubakeur
Influence des cavités (gaz -eau) sur la distribution du champ électrique, de la pression électromécanique et des pertes diélectriques dans l'isolation des câbles HT.

Abonnements - diffusion

Dans le cadre de la diffusion de ses résultats de recherche, l'E.N.P produit depuis 1986, une revue scientifique intitulée «Algerian Journal Of Technology» (AJOT).

Plusieurs formules d'abonnement à cette revue sont disponibles :

1- **Abonnement Annuel** (2 numéros d'une série – Frais d'envoi inclus-)

Algérie : 1000 DA.

Etranger : 400 FF

2- **Collection Complète** de la revue depuis 1986 (Volume 1 à 14 des trois séries) (Voir Appel aux Sponsors)

(Compte tenu des disponibilités)

Algérie : 10.000 DA.

Etranger : 4000 FF

3- **Abonnement de Soutien** à la revue avec des commandes en 3 exemplaires ou plus.

4- **Echanges** avec d'autres établissements et centres de recherche conformément à la circulaire N° 17/50 du 13/05/1998 du Secrétaire Général du Ministère de l'Enseignement Supérieur et de la Recherche Scientifique .

Pour vous abonner à l'A.J.O.T, il vous suffit de remplir le bon de commande ci-joint accompagné d'un chèque libellé.

□ Au nom de la Société Algérienne de Technologie (S.A.T) société savante à but non lucratif (Agrément N° 38 du 14 juin 1995 (MICREA) conformément à la loi 90/91).

Compte BADR : Hacem Badi 635.200.000.586/78

Adresser vos commandes à Monsieur -C- Larbes – Responsable de la diffusion et des abonnements, en précisant la nature de celle – ci (Abonnement annuel, de soutien, d'échange etc...) accompagnée des modalités de règlement correspondantes ainsi que l'adresse à laquelle doit se faire l'envoi.

Je désire souscrire un abonnement annuel à L'Algerian Journal Of Technology

NOM

ADRESSE

Formule d'abonnement retenue* : A S C E

Signature

*- Rayer la mention inutile

A Annuel

S Soutien

C Collection complète E Echange

Appel aux sponsors

La revue AJOT existe depuis 1986, avec l'ambition de la sortie d'un volume par an, à raison d'un numéro par semestre. Les différentes filières technologiques sont regroupées sous trois séries en l'occurrence :

- Série A** : Pour la Mécanique et couvre le Génie Civil, la Mécanique, l'Energétique, L'Hydraulique, les Mines, le Génie des Matériaux,...
- Série B** : Pour l'Electronique, l'Automatique, le Génie Electrique, l'Informatique,...
- Série C** : Pour le Génie Chimique, le Génie de l'environnement, le Génie industriel, le Génie des Procédés,...

Afin de pouvoir améliorer la qualité de la revue AJOT dans sa présentation, dans son contenu et dans sa diffusion la plus large possible, d'autres ressources de financement que la ligne budgétaire accordée à l'ENP à cet effet lui sont nécessaires.

Pour cette raison la revue AJOT fait appel aux sponsors éventuels intéressés. Le sponsoring peut se faire soit directement par des dons (50 000.00 DA minimum) soit par abonnement de 50 exemplaires d'un numéro au minimum.

Les sponsors, s'ils le demandent, leur sigle commerciale figurera en bonne place dans la revue.

Pour avoir d'autres informations, ou explications à ce sujet, contacter le responsable de la publicité et du sponsoring de la revue M. M. Bouabdallah, qui se chargera de les fournir avec plaisir, à l'adresse ci-dessous :

Ecole Nationale Polytechnique
B.P .N° .182 , Hassan Badi El – Harrach, 16200 , Alger, Algérie
Tel : (213) (0) 21 52 59 29 ou 52 53 01/03 ou 52 14 94
Fax : (213) (0) 21 52 29 73
e-mail : mbouabdallah@hotmail.com

RECOMMENDATION TO AUTHORS

The Algerian journal of technology (AJOT) magazine publishes original articles of researches fitting scientific and technologic dimension. Theory, design, applications and development are greatly taken in to account by type subjects dealt with by the issues A, B, and C for the contribution, they appear as following :

1. By articles presenting a meaningful and significant supply in research, development in application of concepts. (Less than 20 pages)
2. By technical notes and short correspondences commenting or correcting type earlier published articles (Not more than 10 pages)
3. By syntheses or monographs (Less than 20 pages)
4. By reports of conferences colloquy, symposium and seminars, gathered in a special issue of The AJOT.

The submission of the manuscript means that it has already been protected by a copyright, and not published, or submitted or accepted for publication elsewhere.

The manuscript must be written preferably in English. The AJOT does not accept the manuscripts written in Arabic or French unless they exceed 50% of their global content as competition regarded.

The whole manuscripts must be concise at great extent. The editor has the right to refuse the articles which exceed the needed number of pages. In case of acceptance the overstepping will be invoiced on the authors.

A. SUBMISSION PROCEDURE OF MANUSCRIPT

1. All manuscripts should be submitted in 5 complete copies, one of item must be original. This is to facilitate the work for the reading committee.
2. The authors whom their articles are accepted, if not done before, have to give a short biography (Less than 100 words) and two photographs of passport.
3. The manuscripts will be transmitted within 2 copies of the application for publication letter.

The address must be shown on the latter and if possible the phone-number and / or telex / fax, etc. . The director is to be informed about any change or alteration in due time.

B. COPYRIGHT

By its rule of procedure, the AJOT adjudges 50% of its authors' rights to compensate the supply of editors. So the application letter must include the signed convention Of authors' on this procedure.

C. PRESENTATION OF THE MANUSCRIPT

1. The manuscripts must be typewriter in double columns, on white paper of A4 (21.29.1 cm) and on one side only, with a 1.5 cm margin on both side. The magazine is off-printed recto-verso. The characters font are times-roman. The title in 14 pts, the authors names and address in 12 pts and the text in 10 pts.
2. What should be written in the first page
 - The title of the article (without symbols) well condensed.
 - Authors names (in limited number, avoid articles of more than three authors) which should be proceeded by their names initials. A slash must mark the name of the authors who will be communicated of any correspondence related to the article.
 - The names and addresses of concerned laboratories.
 - The key words according to norms admitted by the information bank.
 - The summaries in three languages : Arabic, English, and French, in this order - 300 words in maximum for the original articles or synthesis 75 words for technical notes and correspondences. These summaries should not contain equations, references, or notes of acknowledgements.
 - The articles which are exclusively written in English should contain abstracts in Arabic and French.
3. An introduction that indicate the state of the subject and the contribution of the article will be supplied.
4. Indicate, eventually, the advantages, limitations and possible applications at the conclusions.
5. The references should be numbered between brackets []. They must, also, be placed at the end of the article. The references should precise in order the following :
 - The names of all the authors preceded by their last names initials.
 - The title at the article.
 - The abbreviation of the magazine's title. It is a book give the editor's name.
 - The numbers of inclusive pages (or Chapter) after the volume indication, the month and the year.
6. The equations must be numbered in the text in Arabic numbers (between parentheses). Avoid using simultaneous indexes.
7. The references in the text should appear numbered, put between brackets and placed after the name of one or two authors.
8. The list of references should be printed in double spaces at the of text, in the order they given in the text.
9. The nomenclature and the abbreviations must be represented in a glossary at the end of a text and after the list of references.
10. The International System (I.S) has to be used even if a particular practice needs the use of an other system of units. The equivalence in the International System must appear in brackets.
11. The impression of the black colour must be well contrasted (laser impression for example). Avoid computer printers that gives a difficult impression to reproduce.
12. The chosen characters of Times New Roman type must be neither too narrow nor too spindly. The letters and signs must have at least 2 mm of height. Avoid too much bold characters and symbols.
13. As for the titles of paragraphs, use capital letters or bold characters, but not underlined, the titles of alineas must be underlined.
14. All the titles of paragraphs or alineas start too the left margin of the page.
15. The width of columns must be 85 mm. The spacing between columns must be 5 mm. The text should be written in Times New Roman 10 pts. The space between the lines must be 4 mm.

D. CURVES, TABLES AND PHOTOS

1. The authors are asked to submit professional illustration of high quality. The AJOT has not the means to do it.
2. The original drawings should be done with black Chinese ink on white tracing paper. The figures scale must be 7 x 7 cm. It is not possible to insert the figures, tables and photos in the columns. The needed size is limited to 13 x 21 cm. The printing on photograph paper of the illustrations is also possible.
3. The whole subscriptions must be large enough to be legible after the printing.
4. The figures should not be overloaded by legends and text. Give only what is necessary.
5. The figures and tables (charts) must be presented in the text in the desired positions.
6. The numbering of figures and photographs is the same. The tables are numbered apart.
7. All figures and tables must have a title. Their numbering should be dealt with in the increasing order.

E. PRINTING ASIDE

Whenever possible, 30 copies without covers would be printed and given to the authors. In this case the corresponding costs will be paid off immediately after the acceptance of the manuscript according to the estimate of the printer.

F. MAILING OF ARTICLES

The articles and manuscripts comforts to the above recommendation (notes) should be sent to :

Mr. le Directeur de la rédaction de la revue AJOT

Ecole Nationale Polytechnique

B.P. No. 182, Hassan Badi El-Harrach, 16200 Alger, Algérie.

Tel : (213) (0) 21 52 59 29 ou 52 53 01/03 ou 52 14 94 Poste : 252

Télex : 64.147 ENP/DZ

FAX : (213) (0) 21.52.29.73.

Contents/ Sommaire

Avant propos.		1
B.1- Antennes et Méthodes Numériques connexes		
1	A.Azrar and A.Aksas, "Full wave analysis of microstrip antennas employing coplanar parasitic elements based on moment method solution in spectral domain".	4
2	S.Senouci et A.Zerguerras, "Contribution à l'étude d'optimisation d'antennes plaques microruban".	19
3	A.Semmar, M.Si Moussa et A.Zerguerras, "Une modélisation de la permittivité effective sur une large gamme de fréquence d'une ligne microruban suspendue".	35
4	F.Benali et A.Zerguerras, « Réseaux phasés à balayage électronique de type photonique ».	44
5	A.Chemsa et R.Aksas, « Nouvelle approche des réseaux plans rectangulaires avec un niveau de lobes secondaires invariant ».	58
B-2- Segmentation d'Images		
1	A.Boudaib et A.Zerguerras, « Segmentation non supervisée d'images S.A.R par application de l'algorithme E.M ».	66
2	L.Hamami et N.Lassouaoui « Approche multifractale appliquée à l'extraction de contours et à la segmentation d'images biomédicales ».	78
B-3- Systèmes de Reconnaissance		
1	L.Hamami "Système de reconnaissance de l'écriture arabe imprimé utilisant une méthode statistique".	86
2	B.Bousseksou, K.Djouadi et M.D.Kone « Sur un système de reconnaissance automatique de la parole (S.R.A.P) continu multilocuteur ».	94
B-4- Synthèse de la Parole		
1	M.Guerti "Modélisation de la coarticulation des occlusives sonores"	104
2	M.L.Benzaoui et M.Guerti « Durées intrinsèques des sons spécifiques à l'arabe standard ».	114
B-5- Automatique		
1	M.O.Mahmoudi, O.Stihi et S.Hamdani « Etude et réalisation d'un convertisseur AC/DC à facteur de puissance unitaire ».	126
2	D.Boukhetala, F.Boudjema, M.Djebiri, M.S.Boucherit et N.K.M'Sirdi « Decentralized supervised control using neural networks with application to robotic manipulators ».	132
Thesauria 1998		138
Thesauria 1999		140