

57/84

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

ECOLE NATIONALE POLYTECHNIQUE D'ALGER

2x

DEPARTEMENT D'ELECTRONIQUE



PROJET DE FIN D'ETUDES

Sujet :

**ETUDE DE PROGRAMME POUR LA RECONSTITUTION
DE LA PAROLE PAR PREDICTION LINEAIRE**

(UTILISATION DE L'APPLE II)

Sujet proposé par :

Mlle A. MOUSSAOUI

Etudié par :

Mlle SIAD Zineb

Mr HAMDJ Rachid

Promotion Janvier 1984

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

ECOLE NATIONALE POLYTECHNIQUE D'ALGER

DEPARTEMENT D'ELECTRONIQUE

PROJET DE FIN D'ETUDES

المكتبة الوطنية للعلوم الهندسية
المكتبة
Sujet :
ECOLE NATIONALE POLYTECHNIQUE
BIBLIOTHEQUE

ETUDE DE PROGRAMME POUR LA RECONSTITUTION
DE LA PAROLE PAR PREDICTION LINEAIRE

(UTILISATION DE L'APPLE II)

Sujet proposé par :

Mlle A. MOUSSAOUI

Etudié par :

Mlle SIAD Zineb

Mr HAMDI Rachid

Promotion Janvier 1984

— REMERCIEMENTS —

- L'étude que nous présentons a été proposée par mademoiselle MOUSSAOUI qui a supervisé notre projet, ce pourquoi nous voulons l'assurer de notre profonde reconnaissance.
- Nous tenons également à assurer les professeurs de l'école polytechnique d'Alger et ceux de l'université d'Annaba, de nos sentiments de profonde gratitude pour l'enseignement qu'ils nous ont prodigué tout au long de notre formation.
- Nous remercions aussi monsieur FARAH et monsieur PERRIER professeur au centre de linguistique de Ben Aknoun pour l'aide précieuse qu'il nous a fournie en nous procurant une documentation qui nous a été très utile.
- Nous sommes très sensible à l'aide précieuse que nous a apporté monsieur HANACHI Tahar, ce pourquoi nous le remercions.
- Nos remerciements vont à tous ceux qui ont contribué de près ou de loin à notre formation.

DEDICACES

- A la mémoire de mon père
- A ma mère et mon frère
- A mon oncle KADDOUR mabrouk et ses fils brohim et ali
- A toute ma famille
- A tous mes amis (ies)

Je dédie ce travail

H. Rachid

DEDICACES

A la mémoire de mon père

A ma mère ; A son sacrifice

A tous mes frères , sœurs et mes beaux frères

A tous mes amies (is)

A tous mes neveux et nieces

Je dédie ce travail

S. Zineb

SOMMAIRE

PARTIE I

INTRODUCTION

CH.1. RAPPELS SUR LE MECANISME REEL DE LA PRODUCTION DE LA VOIX

CH.2. PRINCIPE D'ANALYSE - SYNTHÈSE DE LA PAROLE

2.1 Analyse de la parole

2.2 Synthèse de la parole

2.3 Fenêtrage

CH.3 - MODELE DE LA PRODUCTION DE LA PAROLE

3.1 Détermination du modèle de production de la parole

3.2 Détermination du nombre de pôles

3.3 Principales méthodes de prédiction linéaire .

3.3.1. Méthode exacte

3.3.2 Méthode de covariance

3.3.3 Méthode d'autocorrelation.

3.4 Choix des méthodes

3.5 Etablissement des équations de covariance et d'autocorrelation

3.5.1 Equations de Covariance

3.5.2 Equations d'autocorrelation

3.6 Erreur quadratique totale minimale et calcul du facteur de gain A

3.6.1 Erreur quadratique totale minimale de prédiction.

3.6.2 Calcul du facteur de gain A

CH4 - RESOLUTION DES EQUATIONS DE PREDICTION LINEAIRE

4.1 Méthodes de résolution des équations

4.2 Résolution du système d'équations de la méthode de covariance

4.3 Résolution du système d'équations de la méthode d'autocorrelation

4.2 Stabilité du prédicteur linéaire

CH.5_ CALCUL DE LA PERIODE FONDAMENTALE DU SIGNAL VOCAL

DECISION VOISE / NON-VOISE

5.1 Calcul de la période fondamentale

5.2 Décision voisé / non-voisé

PARTIE II

- PROGRAMMATION

CH1 - PRESENTATION DE L'APPLE II PLUS

1.1 Introduction

1.2 Le Hardware

1.2.1 L'unité centrale

1.2.2 La mémoire centrale

1.23 Les unités périphériques

1.3 Le software

1.3.1. Le software de base (ou support logiciel)

1.3.2 Langages disponibles dans l'Apple II plus

CH.2 - ETUDE DE PROGRAMMES

Exemple

Conclusion finale

Bibliographie

PARTIE I

ANALYSE - SYNTHÈSE DE LA PAROLE

INTRODUCTION.

Si l'on dit que les jeunes enfants apprennent à parler en écoutant, il en est de même des machines parlantes. C'est à dire que l'obtention du son de parole synthétique n'est possible qu'après l'analyse de ces mêmes sons prononcés par un locuteur humain. Cependant le rôle des synthétiseurs de parole n'est pas de reproduire intégralement des sons préalablement enregistrés, mais de reproduire des signaux en simulant des mécanismes réels du canal Vocal.

Parmi les méthodes d'analyse de parole, on trouve essentiellement les méthodes classiques de traitement (transformé de Fourier, convolution, corrélation, etc...). L'application de ces méthodes souffre cependant de sérieuses limitations à cause du caractère non stationnaire et pseudo-périodique de la parole. Nous présentons ici l'application d'une nouvelle méthode d'analyse du signal de parole, cette méthode est appelée "prédiction linéaire", dont la première application date de moins de dix ans. Son principe théorique est connu depuis longtemps (méthode d'approximation d'une fonction par un polynôme). On peut considérer la prédiction linéaire comme une analyse temporelle aussi bien que comme une méthode d'analyse spectrale et en ce sens elle souffre des mêmes limitations que les méthodes classiques d'analyse. Sa supériorité provient

du fait qu'elle est fondée sur un modèle simple (filtre numérique ne possédant que des pôles) de production de parole qui constitue une bonne approximation du système phonatoire. Ce modèle de notre système vocal permet d'obtenir, par interpolation, l'enveloppe du spectre de signal de parole entre les points de cette enveloppe qui ont été fournis par l'analyse.

CH. 1. RAPPELS SUR LE MECANISME REEL DE LA PRODUCTION DE LA VOIX

Notre organe phonatoire comprend pour l'essentiel : les poumons, la trachée, le larynx, le pharynx et les cavités buccale et nasale. L'ensemble du canal vocal constitue un résonateur ou filtre acoustique, à l'entrée duquel est appliquée une excitation mécanique; l'énergie du système étant fournie par les poumons lorsqu'ils expulsent l'air.

Nous distinguons deux types d'excitations :

- Le premier type d'excitation est produit par la vibration des cordes vocales sous l'action de la pression de l'air en provenance des poumons. L'onde produite par cette vibration a une forme approximativement triangulaire que l'on peut assimiler à un train d'impulsions. La fréquence de récurrence de ces impulsions, appelé fréquence fondamentale ou fréquence de mélodie (pitch en anglais) varie en moyenne de 100 à 150 Hz pour les hommes et de 200 à 250 Hz pour les femmes. Les sons produits par ce type d'excitation sont dit voisés, ce sont par exemple les voyelles.

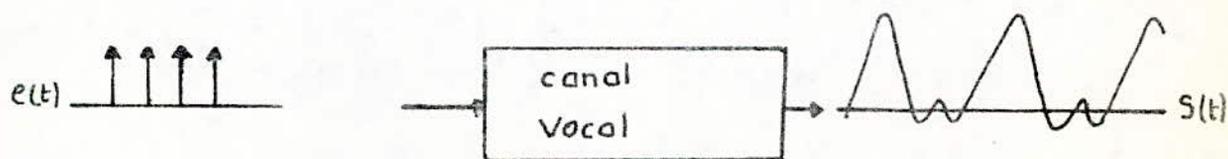


fig-1- Production d'un son voisé.

- Le deuxième type d'excitation est provoqué par la génération d'un écoulement d'air turbulent dans le canal vocal, l'entrée du filtre peut alors être assimilée à un bruit acoustique approximativement blanc (spectre plat dans la bande de fréquence audio). Les sons ainsi produits sont appelés non-voisés.

CH2. PRINCIPE D'ANALYSE-SYNTHESE DE LA PAROLE

L'idée essentielle de l'analyse-synthèse de la parole est de considérer les signaux de parole humaine comme la réponse d'un filtre linéaire (le canal vocal) à une excitation mécanique (les vibrations des cordes vocales par exemple).

Le canal vocal est défini par certains paramètres qui varient dans le temps.

- Ouverture de la bouche
- Position de la langue
- Position des dents etc ...

Son comportement est modifié par celui d'un filtre à coefficients évolutifs. A l'entrée de ce filtre, on envoie une excitation reproduisant le mieux possible pour chaque type de son, l'excitation réelle du canal vocal. On obtient ainsi en sortie le signal de parole synthétique.

2.1. L'analyse de la parole.

Soit un signal de parole $S(t)$ tel que $t \in [0, T]$.

La première opération d'analyse consiste à tronçonner le temps en intervalles d'environ 10 à 20 ms. Ces intervalles de temps sont appelés "fenêtres temporelles". En effet le signal $S(t)$ pris dans son ensemble n'est pas stationnaire, mais sa restriction à chacun de ces petits intervalles de

temps peut être considérée comme telle. L'analyse sera faite successivement sur chacune des fenêtres. Le signal est ensuite échantillonné et l'on obtient ainsi pour chaque fenêtre une suite d'échantillons $S(n)$, $n \in [0, N]$, N est généralement de l'ordre de 200. La prédiction linéaire consiste à exprimer la valeur du n ème échantillon en fonction d'un nombre fini P d'échantillons antérieurs et de l'entrée $u(nT)$ du système. Il s'agit d'obtenir une formule du type :

$$\hat{S}(nT) = \sum_{k=1}^P a(k) \hat{S}(nT - kT) + A u(nT) \quad 3.1$$

où T est la période d'échantillonnage

Les coefficients $a(k)$ et A reproduisent le mieux possible le comportement du système considéré, qui est le canal vocal, compte tenu de l'entrée $u(nT)$ appliquée.

On obtient ainsi une formule de type (3.1) pour chaque fenêtre ce qui signifie que les coefficients $a(k)$ et A doivent être recalculés sur chaque fenêtre.

Pour simplifier la formulation nous substituons $\hat{S}(nT)$ par $\hat{S}(n)$, d'où la relation (3.1) devient :

$$\hat{S}(n) = \sum_{k=1}^P a(k) \hat{S}(n-k) + A u(n) \quad (3.2)$$

où encore en transformé en Z

$$\hat{S}(Z) = \sum_{k=1}^P a(k) S(Z) Z^{-k} A u(Z)$$

$$\hat{S}(Z) = \left(1 - \sum_{k=1}^P a(k) Z^{-k} \right) A u(Z)$$

Nous posons :

$$H(Z) = 1 - \sum_{k=1}^P a(k) Z^{-k}$$

d'où
$$\hat{S}(Z) = \frac{A}{H(Z)} \cdot u(Z)$$

Le signal \hat{S} apparaît donc comme la réponse du filtre $A/H(Z)$, ne possédant que des pôles, à l'entrée u .

Une fois $A/H(Z)$ connu, la phase d'analyse est terminée.

2.2 La synthèse .

Principe :

On dispose pour chaque fenêtre temporelle des paramètres définissant le filtre $A/H(Z)$ qui, repellons-le est censé imiter le comportement du canal vocal. Il reste maintenant à exciter ce filtre de façon appropriée, afin d'obtenir en sortie des échantillons de parole synthétique $\hat{S}(n)$.

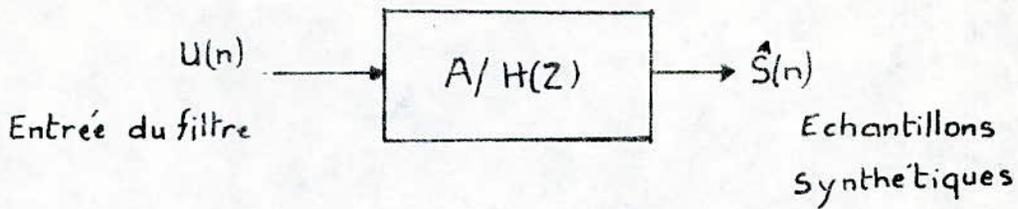


fig 3.1- Production du signal synthétique.

Ayant déterminées les deux types d'excitations $u(n)$ du filtre qui, rapellons-les, sont un train périodique d'impulsions et un bruit acoustique blanc.

Nous disposons maintenant de tous les paramètres nécessaires à l'obtention du signal synthétique.

Ces paramètres nécessaires sont:

- Les coefficients du polynôme $A/H(Z)$
- Un paramètre de selection " Voisé/nonVoisé " que nous définissons par la suite .
- La période pitch dans le cas des sons voisés .
- Un paramètre d'amplitude correspondant au gain A à l'entrée du filtre $1/H(Z)$.

2.3. Fenêtrage

Dans la pratique pour améliorer les performances du modèle de la production de la parole, il convient de traiter le signal de parole avant de l'analyser, par exemple en lui faisant un fenêtrage de Hamming.

Le lissage du spectre par fenêtrage de Hamming :

Considérons une fenêtre temporelle $[0, T]$, telle que nous l'avons envisagée au départ, l'analyse d'un signal $S(t)$ sur cette fenêtre est faite en échantillonnant le signal et en ne prenant en compte que les échantillons situés à l'intérieur de la fenêtre. Ceci revient à multiplier $S(t)$ par une fonction rectangulaire, nulle à l'extérieur de $[0, T]$ et valant 1 à l'intérieur.



fig 4.1- fenêtre rectangulaire.

Le spectre du signal ainsi tronqué est obtenu par convolution du spectre initial par une fonction sinus cardinal, cette opération introduit des irrégularités dues à la nature du sinus cardinal

dont les lobes secondaires contiennent une énergie non négligeable. Pour palier à cet inconvénient on multiplie $S(t)$ par une fonction "plus douce" qu'une fenêtre rectangulaire que l'on appelle fenêtre de Hamming et que l'on note $w(t)$.

L'analyse est alors faite à partir du signal :

$$\hat{S}(t) = S(t) \cdot w(t)$$

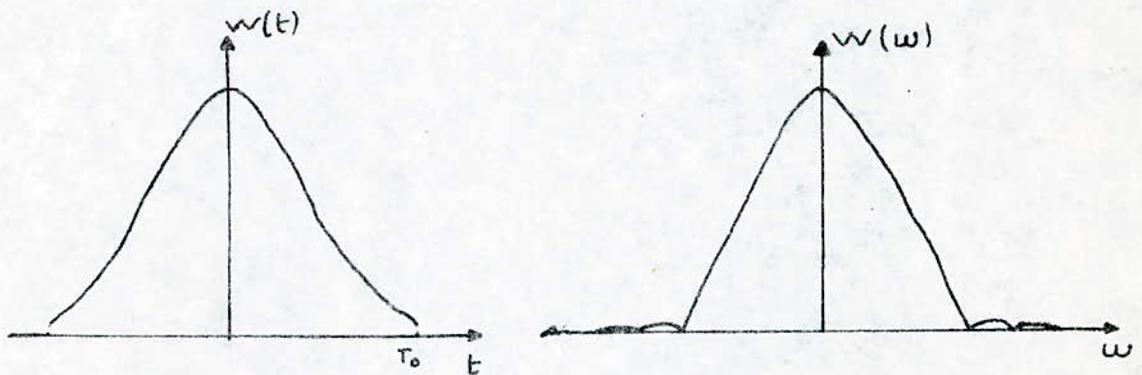
C'est à dire on considère les échantillons :

$$\hat{S}(n) = S(n) \cdot w(n)$$

Pour une fenêtre comportant N éléments on a :

$$w(n) = 0.54 - 0.46 \cos \pi(n-1)/(N-1)$$

Soient $S(f)$, $\hat{S}(f)$ et $w(f)$ les spectres des signaux $S(t)$, $\hat{S}(t)$ et $w(t)$ respectivement.



domaine temporel

4.a

domaine fréquentiel

4.b

fig 4.2 Fenêtre de Hamming

Comme il apparaît sur la figure 4.b, la majeure partie de l'énergie de $w(f)$ (plus de 99 %) est concentrée dans le lobe principale, les lobes secondaires pouvant être négligés ; l'opération de convolution se traduit par un lissage de $S(f)$.

But du fenêtrage :

Pour calculer la fonction d'autocorrelation instantanée d'un signal $S(t)$ relativement à une séquence de durée T de ce signal, il faut que $S(t)$ soit stationnaire. Mais le signal de parole est non stationnaire. Et de ce fait, on admet généralement, comme nous l'avons indiqué, on peut considérer le signal de parole comme étant stationnaire sur des intervalles de temps d'une durée maximale égale à 20 ou 30 ms. Ainsi on est conduit à faire un fenêtrage.

CH3. MODELE DE LA PRODUCTION DE LA PAROLE

3.1 Détermination du modèle de production de la parole.

Le principe de la méthode de prédiction linéaire est fondé sur l'hypothèse selon laquelle un échantillon du signal de parole peut être prédit approximativement par une somme, pondérée linéairement, d'un certain nombre d'échantillons le précédant immédiatement.

On désigne par $S'(n)$ la valeur prédite de $S(n)$ qui s'écrit :

$$S'(n) = \sum_{k=1}^P a(k) S(n-k) \quad 3.1$$

Où $a(k)$, $1 \leq k \leq P$ est un ensemble de coefficients réels, que l'on appelle généralement coefficient du prédicteur linéaire, P est l'ordre du prédicteur.

L'erreur de prédiction entre le signal original et le signal prédit est représentée par :

$$e(n) = S(n) - S'(n) \quad 3.2$$

$$= S(n) - \sum_{k=1}^P a(k) S(n-k) \quad 3.3$$

Cela revient à dire :

$$S(n) = e(n) + \sum_{k=1}^P a(k) S(n-k) \quad 3.4$$

L'équation 3.4 tient compte aussi bien de la contribution des cavités buccale et nasale que de la source.

Cependant, dans un but de simplification et afin de pouvoir déterminer facilement les paramètres du filtre du signal original, nous faisons les approximations suivantes:

Evoquant l'approximation principale: on suppose que la séquence $e(n)$ des échantillons de l'erreur de prédiction peut être approchée, pendant une quelconque stationnarité, par la séquence $\hat{e}(n)$ suivante:

$$\hat{e}(n) = \begin{cases} A & n=0 \\ 0 & n \neq 0 \end{cases} \quad 3.5$$

A étant une constante (facteur de gain) à déterminer.

Nous allons appliquer maintenant la transformée en Z à l'équation (5.3):

$$E(Z) = S(Z) - \sum_{k=1}^P a(k) Z^{-k} \cdot S(Z) \quad 3.6$$

$$= S(Z) \left(1 - \sum_{k=1}^P a(k) Z^{-k} \right) \quad 3.7$$

$$\bar{=} S(Z) \cdot H(Z) \quad 3.8$$

où $E(Z)$ et $S(Z)$ sont les transformés en Z de $e(n)$ et $s(n)$ respectivement.

nous avons
$$H(Z) = 1 - \sum_{k=1}^P a(k) Z^{-k} \quad 3.9$$

avec $1 \leq k \leq P$

$H(Z)$ désigne le filtre inverse.

De l'équation 3.8 nous déduisons :

$$S(Z) = E(Z) / H(Z) \quad 3.10$$

D'où l'approximation faite ci-avant dans le domaine temporel se traduit ici par une approximation sur la fonction de transfert $S(Z)$.
Nous désignons par $G(Z)$ la fonction de transfert approximée :

$$G(Z) = \frac{A}{H(Z)} \quad 3.11$$

$$= \frac{A}{1 - \sum_{k=1}^P a(k) Z^{-k}} \quad 3.12$$

Cette fonction de transfert est celle d'un filtre ne possédant que des pôles. Nous voyons maintenant que l'approximation que nous avons faite consiste à admettre que la fonction de transfert du système phonatoire ne comporte que des pôles.

Remarque : notre appareil auditif est plus sensible à un pôle qu'à un zéro de la fonction de transfert.

En résumé, le modèle de production de la parole utilisé dans les méthodes de prédiction linéaire est fondé sur les hypothèses suivantes :

1. Pendant des intervalles de temps courts, de l'ordre de 10 à 25 ms,

la source vocale et le conduit vocal peuvent être considérés comme constants.

2- A l'intérieur de chacun de ces intervalles de temps, on suppose que l'on peut approximer la fonction de transfert de l'ensemble source vocale, conduit vocal (cavités buccale et nasale) et rayonnement, par la fonction de transfert d'un filtre linéaire, invariable dans le temps et ne possédant que des pôles, ce filtre est par ailleurs, excité par une séquence d'impulsions à période de mélodie ou par une séquence de bruit blanc.

3- Le signal de parole de synthèse est constitué par le signal de sortie d'un tel filtre dont on modifie les coefficients $a(k)$, $1 \leq k \leq P$ à des intervalles de temps déterminés (10 à 25 ms).

Le schéma d'un tel modèle de production de la parole est représenté par la figure 3.1

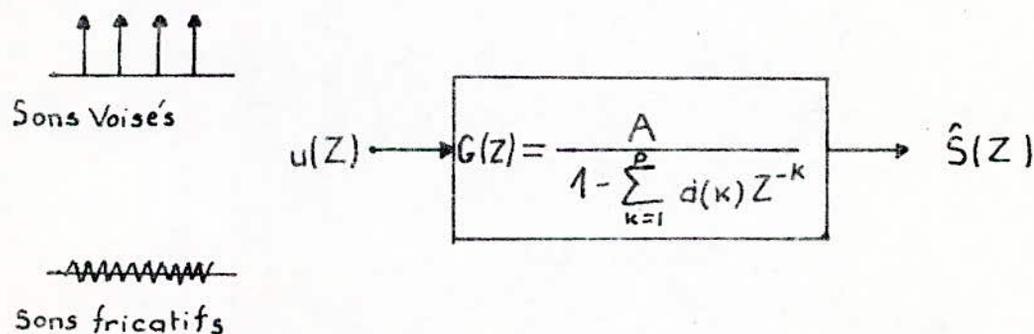


fig 3.a - Modèle dans le domaine fréquentiel.

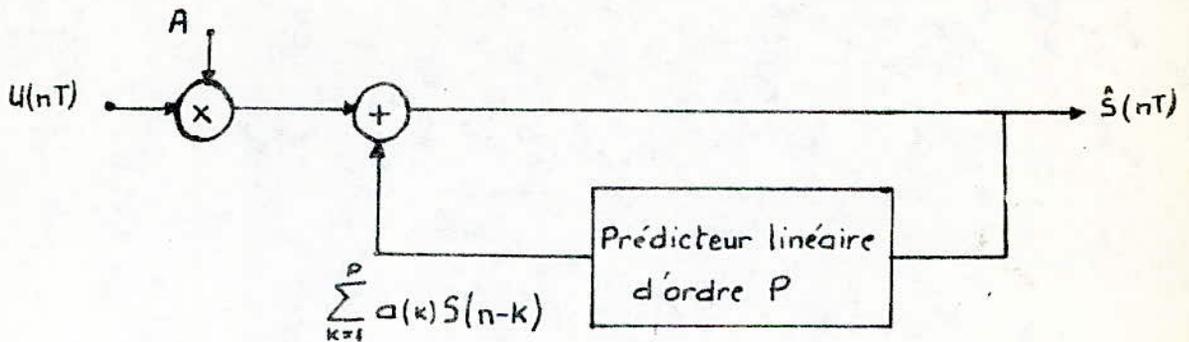


fig 3.b Modèle dans le domaine temporel .

FIGURE 3.1 Modèle de production de la parole utilisé dans les méthodes de prédiction linéaire .

A partir de la figure 3.1 , on constate que les échantillons du signal de parole en sortie du filtre $G(Z)$ sont notés $\hat{S}(nT)$ et les échantillons du signal de source , à l'entrée du filtre sont notés $U(nT)$.

On peut écrire de la figure 5.a que :

$$\hat{S}(Z) = U(Z) \cdot G(Z) = \frac{A U(Z)}{1 - \sum_{k=1}^P a(k) Z^{-k}} \quad 3.13$$

où $U(Z)$ et $S(Z)$ sont respectivement les transformés en Z de $U(nT)$ et $\hat{S}(nT)$.

De l'équation 3.13 nous pouvons écrire :

$$\hat{S}(Z) = \hat{S}(Z) \cdot \sum_{k=1}^P a(k) Z^{-k} + A U(Z) \quad 3.14$$

En prenant la transformée inverse de la relation 3.14 en aboutit à la représentation temporelle du modèle de production de la parole dont le schéma équivalent est indiqué par la figure 3. b

$$\text{d'où} \quad \hat{S}(n) = \sum_{k=1}^P a(k) \hat{S}(n-k) + AU(n) \quad 3.15$$

Ayant déterminé le modèle de production de la parole utilisé dans les méthodes de prédiction linéaire, alors il reste à déterminer les paramètres du prédicteur qui sont déjà cités précédemment.

3.2 Détermination du nombre P de pôles.

Ayant vu que l'on peut approcher la fonction de transfert de l'ensemble source vocale, conduit vocal et rayonnement par la fonction de transfert d'un filtre linéaire qui ne possède que des pôles. Il nous faut donc déterminer le nombre de pôles nécessaires à l'obtention d'une bonne approximation de la fonction de transfert. Ce nombre P de pôles, ou nombre de coefficients du prédicteur, permettant de représenter correctement toute séquence du signal de parole est fonction du nombre de résonances et d'antirésonances du conduit vocal dans la bande de fréquence considérée, de la nature de la fonction d'excitation ainsi que des caractéristiques de rayonnement au niveau des lèvres.

Il a été montré (ATAL.1971) qu'afin de parvenir à une représentation correcte des pôles de la fonction de transfert du conduit vocal, en faisant abstraction des effets de la source et de rayonnement, la durée de mémorisation $p.T$ nécessaire au calcul de l'estimée $\hat{S}(n)$ de $S(n)$ dans (3.1) doit correspondre à deux fois le temps mis par l'onde de parole pour se propager depuis la glotte jusqu'aux lèvres. Ainsi par exemple si le conduit vocal a une longueur de 17 cm (valeur moyenne), le temps équivalent au nombre d'échantillons contenus dans la mémoire doit être égale à 1 ms (qui est le double du temps mis par l'onde sonore à la vitesse de 340 m/s pour traverser le conduit vocal). Si la fréquence F_e d'échantillonnage du signal de parole est 10 KHZ, la valeur correspondante de p est alors égale à 10. Par ailleurs, on ajoute deux pôles qui sont nécessaires pour représenter l'influence de la source et du rayonnement, ce qui porte à 12 la valeur moyenne du nombre P dans le cas de $F_e = 10$ KHZ. On peut formuler l'expression de P de la façon suivante :

$$P = F_e (\text{KHZ}) + N_r \quad 3.16$$

où F_e est la fréquence d'échantillonnage en KHZ et N_r le nombre nécessaire à la représentation des effets de la source et du

rayonnement (N_r est égale à 2 ou 3)

Remarque : cette valeur de P est une valeur indicative car le nombre de pôles du système phonatoire varie suivant les sons analysés et aussi en fonction du locuteur.

3.3 Principales méthodes de prédiction linéaire.

Toute méthode de prédiction linéaire est fondée sur l'hypothèse qu'à un instant donné, on peut prédire un échantillon de signal de parole $S(n)$ par une somme des P échantillons précédents.

Le signal prédit étant :

$$S'(n) = \sum_{k=1}^P a(k) S(n-k) \quad 3.17$$

Le problème posé est de déterminer les coefficients $a(k)$ du prédicteur qui soient tel que la relation de l'échantillon prédit demeure optimale pendant une période de temps déterminée. La relation 3.17 s'adapte bien au signal de parole car comme on l'a vu il est possible de l'appliquer à un modèle de production de la parole. Ils existent plusieurs possibilités de déterminer les paramètres $a(k)$, nous étudierons que 3 principales méthodes.

3.3.1 - Méthode exacte.

Nous citons les hypothèses fondamentales sur lesquelles est fondée

la méthode :

- a - Le signal est défini exactement pour $2P$ échantillons consécutifs.
- b - Un échantillon de signal de parole peut être prédit exactement (sans erreur) à partir des P échantillons précédents
- c - L'hypothèse b. est valable pour les P échantillons consécutifs qui suivent.

L'ensemble de ces hypothèses est représenté par les équations suivantes:

$$\sum_{k=1}^P a(k)S(n-k) = S(n) \quad 3.18$$

avec $n = P, P+1, \dots, 2P-1$

3.3.2 Méthode fondée sur la covariance

Nous citons les hypothèses fondamentales sur lesquelles est fondée

la méthode :

- a - Le signal est défini pour $P+N$ échantillons consécutifs, avec N entier.
- b - Un échantillon de signal de parole peut être prédit approximativement à partir des P échantillons précédents.
- c - L'hypothèse b. est valable pour les N échantillons consécutifs qui suivent.
- d - L'écart quadratique total, entre le signal original et sa valeur prédite, est minimisée pour l'ensemble des N échantillons consécutifs. Ici la détermination des coefficients $a(k)$, $1 \leq k \leq P$ est obtenue par

minimisation de l'erreur quadratique totale E que nous établirons par suite, ce qui permet d'établir l'ensemble d'équations suivant :

$$\sum_{k=1}^P a(k) C(i, k) = C(i, 0) \quad 1 \leq i \leq P \quad 3.19$$

$$C(i, k) = \sum_{n=0}^{N-1} S(n-i) S(n-k) \quad 3.20$$

Nous disposons de P équations à P inconnues que nous pouvons résoudre pour obtenir les coefficients $a(k)$, $1 \leq k \leq P$. Les coefficients $C(i, k)$ qui sont les échantillons de la fonction de covariance du signal forment la matrice de covariance.

$$\begin{bmatrix} C(1,1) & C(1,2) & \dots & C(1,P) \\ C(2,1) & C(2,2) & \dots & C(2,P) \\ \vdots & \vdots & \ddots & \vdots \\ C(P,1) & C(P,2) & \dots & C(P,P) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \vdots \\ a(P) \end{bmatrix} = \begin{bmatrix} C(1,0) \\ C(2,0) \\ \vdots \\ C(P,0) \end{bmatrix} \quad 3.21$$

3.3.3 Méthode fondée sur l'autocorrélation

Nous citons les hypothèses fondamentales sur lesquelles est fondée la méthode :

a. Le signal est défini pour toutes les valeurs du temps, il est identiquement nul en dehors d'une séquence de N échantillons, où N est un entier, ceci équivaut à multiplier le signal de parole par une fenêtre finie

correspondant à N échantillons

b. Chaque échantillon peut être prédit approximativement à partir de P échantillons précédents.

c. L'hypothèse b. est valable pour toutes les valeurs du temps (pour tous les échantillons du signal).

d. L'écart quadratique total entre le signal original et le signal prédit est minimisé sur toute l'échelle du temps.

La détermination des coefficients $a(k)$, $1 \leq k \leq P$ est obtenue par minimisation de l'erreur quadratique, ce qui permet d'établir l'ensemble d'équations suivant :

$$\sum_{k=1}^P a(k) R|i-k| = R(i) \quad i=1, 2, \dots, P \quad 3.22$$

$$R(i) = \sum_{n=0}^{N-1-i} S(n) S(n+i) \quad 3.23$$

L'ensemble de ces équations constitue P équations à P inconnues que l'on peut résoudre pour obtenir les coefficients $a(k)$.

Ici les coefficients $R(i)$ sont les échantillons de la fonction d'autocorrelation du signal et les coefficients $R|i-k|$ constituent une matrice dite d'autocorrelation.

3.4 Choix des méthodes.

On a vu que pour le modèle de Production de la parole, le signal

original peut être exprimé en fonction du signal prédit $\hat{S}(n)$ et de l'erreur de prédiction.

$$S(n) = e(n) + \sum_{k=1}^P a(k) S(n-k)$$

Mais la méthode exacte suppose que l'erreur $e(n)$ est identiquement nulle à chaque instant. Cela entraîne que dans le modèle de production de parole, $u(n)$ est nul pour tout n . Alors cette méthode implique qu'il n'y a pas d'impulsions de mélodie pendant l'intervalle de temps correspondant aux $2P$ échantillons de parole nécessaires à l'analyse. D'où l'utilisation de cette méthode ne s'applique pas à notre modèle. Les deux autres méthodes ont été utilisées dans plusieurs travaux de recherche sur l'application de la prédiction linéaire à l'analyse-synthèse de la parole. De ce fait nous étudierons ces deux méthodes pour l'application au modèle de production de parole.

3.5 Etablissement des équations de Covariance et d'autocorrelation.

Comme nous l'avons vu précédemment on désigne l'erreur $e(n)$ entre la valeur du signal original de parole et le signal prédit par :

$$\begin{aligned} e(n) &= S(n) - \hat{S}(n) \\ &= S(n) - \sum_{k=1}^P a(k) S(n-k) \end{aligned}$$

Pour déterminer l'ensemble des coefficients $a(k)$, $1 \leq k \leq P$ on procède à la minimisation de l'erreur quadratique pour l'ensemble des échantillons.

de la séquence considérée. Dans chacune des deux méthodes celle de covariance et d'autocorrelation, on utilise le critère "des moindres carrés". Nous désignons par E l'erreur quadratique totale :

$$E = \sum_n e^2(n) = \sum_n (s(n) - \hat{s}(n))^2 \quad 3.24$$

$$= \sum_n \left(s(n) - \sum_{k=1}^P a(k) s(n-k) \right)^2 \quad 3.25$$

Nous allons minimiser l'erreur quadratique E sans nous préoccuper du domaine de sommation qui en fait un facteur très important. Le minimum de E est obtenu lorsque toutes les dérivées partielles par rapport à $a(k)$ sont nulles.

$$\frac{\partial E}{\partial a(i)} = \sum_n 2 \left(s(n) - \sum_{k=1}^P a(k) s(n-k) \right) \cdot s(n-i) = 0 \quad 3.26$$

avec $1 \leq i \leq P$

$$\text{ou} \quad \sum_n \sum_{k=1}^P a(k) s(n-k) \cdot s(n-i) = \sum_n s(n) \cdot s(n-i) \quad 3.27$$

avec $1 \leq i \leq P$

on obtient encore :

$$\sum_{k=1}^P a(k) \sum_n s(n-k) s(n-i) = \sum_n s(n) \cdot s(n-i) \quad 3.28$$

avec $1 \leq i \leq P$

Les équations 3.28 constituent un ensemble de P équations à P inconnues. A partir de ces équations nous pouvons déduire les

équations de Covariance et d'autocorrelation.

3.5.1 Equations de covariance.

A partir des hypothèses faites précédemment sur la méthode de covariance, nous pouvons déduire le domaine de sommation de la relation 3.28, qui est de 0 à N-1

$$\sum_{k=1}^P a(k) \sum_{n=0}^{N-1} S(n-k) S(n-i) = \sum_{n=0}^{N-1} S(n) \cdot S(n-i) \quad 3.29$$

Nous posons :

$$C(i, k) = \sum_{n=0}^{N-1} S(n-i) S(n-k) \quad 3.30$$

Des relations 3.29 et 3.30 nous obtenons :

$$C(i, 0) = \sum_{k=1}^P a(k) C(i, k) \quad 1 \leq i \leq P \quad 3.31$$

Nous avons aboutit aux relations 3.19 et 3.20 qui définissent la méthodes de covariance. A partir de ces relations nous déduisons que les valeurs du signal $S(n)$ doivent être définies pour $n = -p, \dots, 0, 1, \dots, N-1$

3.5.2 Equations d'autocorrelation

De même, à partir des hypothèses faites précédemment sur la méthode d'autocorrelation, nous définissons le signal comme suit :

$$S(n) = 0 \quad \text{pour } n \notin [0, N-1]$$

$S(n)$ égale à un certain signal échantillonné pour $n = 0, 1, \dots, N-1$

Et nous pouvons déduire le domaine de sommation de la relation 3.28 qui cette dernière sera :

$$\sum_{k=1}^p a(k) \sum_{n=-\infty}^{+\infty} S(n-k) S(n-i) = \sum_{n=-\infty}^{+\infty} S(n) S(n-i) \quad 3.32$$

Nous posons $m = n - i$ dans la relation 3.32

$$\text{d'où} \quad \sum_{k=1}^p a(k) \sum_{m=-\infty}^{\infty} S(m) S(m+i-k) = \sum_{m=-\infty}^{\infty} S(m) S(m+i) \quad 3.33$$

Nous donnons par définitions la fonction d'autocorrélation $R(i)$ du signal $S(n)$ comme suit :

$$R(i) = \sum_{n=-\infty}^{+\infty} S(n) \cdot S(n+i) \quad 3.34$$

avec $R(i) = R(-i)$

Par conséquent, de la relation 3.33 nous obtenons :

$$\sum_{k=1}^p a(k) R(i-k) = R(i) \quad 1 \leq i \leq p \quad 3.35$$

Etant donné que le signal est limité par une fenêtre :

$$S(n) = 0 \quad \text{pour } n \notin [0, N-1]$$

L'équation 3.34 devient :

$$R(i) = \sum_{n=0}^{N-1-i} S(n) S(n+i) \quad 3.36$$

D'où nous avons retrouvé les relations 3.22 et 3.23 citées précédemment.

Réécrivons la relation 3.35 sous forme matricielle :

$$\begin{bmatrix} R(0) & R(1) & R(2) & \dots & R(P-1) \\ R(1) & R(0) & R(1) & & R(P-2) \\ \vdots & & & & \vdots \\ R(P-1) & R(P-2) & \dots & & R(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \vdots \\ a(P) \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(P) \end{bmatrix} \quad 3.37$$

On constate que la matrice (P, P) d'autocorrelation est symétrique et que les éléments situés sur la diagonale sont identiques. Cette forme matricielle est appelée forme de TOEPLITZ.

3.6 Erreur quadratique totale minimale et calcul du facteur de gain A.

Pour déterminer les coefficients $a(k)$ du prédicteur linéaire, on utilise la minimisation de l'erreur quadratique totale E . Le calcul des coefficients $a(k)$, nous permet de déterminer l'erreur quadratique minimale, notée $E(P)$, car celle-ci dépend du nombre P de coefficients de prédiction utilisés.

3.6.1 Erreur quadratique totale minimale de prédiction.

Nous avons :

$$\begin{aligned}
 E &= \sum_n \left(S(n) - \sum_{k=1}^P a(k) S(n-k) \right)^2 \\
 &= \sum_n \left(S(n)^2 - 2S(n) \sum_{k=1}^P a(k) S(n-k) + \sum_{k=1}^P \sum_{i=1}^P a(k) a(i) S(n-k) S(n-i) \right) \\
 &= \sum_n S(n)^2 - 2 \sum_{k=1}^P a(k) \sum_n S(n) S(n-k) + \sum_{k=1}^P a(k) \sum_{i=1}^P a(i) \sum_n S(n-k) S(n-i)
 \end{aligned}$$

Nous avons la relation 3.28 qui s'écrit :

$$\sum_{k=1}^P a(k) \sum_n S(n-k) S(n-i) = \sum_n S(n) S(n-i)$$

de même
$$\sum_{i=1}^P a(i) \sum_n S(n-k) S(n-i) = \sum_n S(n) S(n-k)$$

En remplaçant la relation 3.28 dans l'expression de E nous obtenons :

$$E(p) = \sum_n S(n)^2 - \sum_{k=1}^P a(k) \sum_n S(n) S(n-k) \quad 3.38$$

Dans la méthode de covariance, nous avons obtenu les résultats suivants :

$$C(i, k) = \sum_{n=0}^{N-1} S(n-i) S(n-k)$$

d'où
$$C(0, 0) = \sum_{n=0}^{N-1} S(n)^2$$

et
$$C(0, k) = \sum_{n=0}^{N-1} S(n) \cdot S(n-k)$$

En remplaçant ces dernières équations dans l'expression de E(p), nous obtenons l'erreur quadratique totale minimale pour la méthode de covariance qui s'écrit :

$$E(p) = C(0, 0) - \sum_{k=1}^P a(k) C(0, k) \quad 3.39$$

Pour la méthode d'autocorrelation, nous avons obtenu les résultats suivants :

$$R(k) = \sum_{n=0}^{N-1-|k|} S(n) S(n+|k|)$$

$$\text{d'où } R(0) = \sum_{n=0}^{N-1} S(n)^2$$

En remplaçant toujours ces dernières relations dans l'expression de $E(p)$, nous obtenons :

$$E(p) = R(0) - \sum_{k=1}^p a(k) R(k) \quad 3.40$$

L'erreur normalisée $V(p)$ est définie comme suit :

$$V(p) = \frac{E(p)}{R(0)} \quad 3.41$$

$$\text{d'où } V(p) = 1 - \sum_{k=1}^p a(k) r(k) \quad 3.42$$

$$\text{avec } r(k) = \frac{R(k)}{R(0)}, \quad \forall k \quad 3.43$$

L'erreur normalisée $V(p)$ est définie comme étant le rapport de l'énergie représentée par les échantillons de l'erreur de prédiction à l'énergie représentée par les échantillons du signal de parole pour la séquence considérée .

Les coefficients $r(k)$ sont appelés coefficients normalisés de la fonction d'autocorrelation .

3.6.2. Calcul du facteur de gain A

Pour calculer le facteur de gain A de la fonction de transfert donnée par la relation 3.12, nous faisons appel au critère suivant :

L'énergie totale contenue dans la séquence du signal de synthèse doit être égale à l'énergie totale de la séquence correspondante du signal analysé. En d'autres termes, le signal en sortie du filtre $1/H(Z)$ doit avoir la même énergie que le signal original $S(n)$.

Maintenant nous faisons une simplification qui consiste à supposer que le signal est analysé sur une période du fondamental; ce qui signifie qu'à la synthèse, il n'y aura qu'une seule impulsion d'excitation du filtre pendant le temps correspondant. Par conséquent le critère précédent peut être formulé comme suit: l'énergie totale contenue dans la réponse impulsionnelle de $G(Z)$ doit être égale à l'énergie totale du signal analysé.

L'énergie totale contenue dans la réponse impulsionnelle de $G(Z)$ est donnée par la fonction d'autocorrelation $\hat{R}(0)$ de cette réponse impulsionnelle. La réponse impulsionnelle de $G(Z)$ est obtenue à partir de l'équation 3.15, en posant:

$$\hat{S}(n) = h(n) \quad \text{et} \quad u(n) = \delta(n, 0)$$

$$\text{d'où} \quad h(n) = \sum_{k=1}^p a(k) h(n-k) + A \delta(n, 0) \quad 3.44$$

$$\text{avec} \quad \delta(n, m) = \begin{cases} 1 & , \quad n = m \\ 0 & , \quad n \neq m \end{cases} \quad 3.45$$

Ainsi d'après les relations 3.44 et 3.45, nous déduisons que:

$$h(n) = 0, \quad n < 0 \quad 3.46$$

$$h(0) = A \quad 3.47$$

$$h(n) = \sum_{k=1}^P a(k) h(n-k), \quad n \geq 1 \quad 3.48$$

Par définition, la fonction d'autocorrelation $\hat{R}(i)$ est donnée par :

$$\hat{R}(i) = \sum_{n=-\infty}^{+\infty} h(n) h(n+i), \quad \forall i \quad 3.49$$

A partir des relations 3.46 à 3.48 nous pouvons déduire que :

$$\hat{R}(i) = \sum_{n=0}^{+\infty} h(n) h(n+i) \quad i \geq 0 \quad 3.50$$

Si on prend les valeurs de i tel que $i \geq 1$, alors $n+i \geq 1$ dans la relation 3.49, puis en remplaçant n par $n+i$ dans la relation 3.48 nous obtenons :

$$h(n+i) = \sum_{k=1}^P a(k) h(n+i-k) \quad 3.51$$

$$\text{d'où} \quad \hat{R}(i) = \sum_{n=0}^{+\infty} h(n) \sum_{k=1}^P a(k) h(n+i-k) \quad 3.52$$

$$= \sum_{k=1}^P a(k) \sum_{n=0}^{+\infty} h(n) h(n+i-k) \quad 3.53$$

Ainsi nous obtenons :

$$\hat{R}(i) = \sum_{k=1}^P a(k) \hat{R}|i-k|, \quad 1 \leq i \leq \infty \quad 3.54$$

car
$$\hat{R}(i-k) = \sum_{n=0}^{+\infty} h(n)h(n+i-k)$$

On détermine $\hat{R}(0)$ à partir des relations 3.46 jusqu'à 3.49.

$$\hat{R}(0) = \sum_{n=0}^{+\infty} h^2(n) \quad 3.55$$

$$= h^2(0) + \sum_{n=1}^{+\infty} h(n) \sum_{k=1}^P a(k) h(n-k)$$

$$= A^2 + \sum_{k=1}^P a(k) \sum_{n=1}^{+\infty} h(n) h(n-k) \quad 3.56$$

Nous posons $n = m+k$ dans $\hat{R}(0)$

$$\hat{R}(0) = A^2 + \sum_{k=1}^P a(k) \sum_{m=1-k}^{+\infty} h(m) h(m+k)$$

Et comme $h(m) = 0$ pour $m < 0$, alors :

$$\hat{R}(0) = A^2 + \sum_{k=1}^P a(k) \sum_{m=0}^{+\infty} h(m) h(m+k) \quad 3.57$$

$$= A^2 + \sum_{k=1}^P a(k) \hat{R}(k) \quad 3.58$$

Nous pouvons remarquer que comme les équations 3.35 et 3.54 sont de même type, on peut déterminer les coefficients $a(k)$ uniquement à partir de $\hat{R}(i)$, $0 \leq i \leq P$. En fait, pour une valeur de A fixé il existe une relation bi-univoque entre la relation impulsionnelle de $G(z)$ et la fonction d'autocorrelation correspondante. La différence entre les équations 3.35 et 3.54 consiste à ce que, pour les équations 3.35, le domaine de variation de l'indice i est limité. Cependant elles

sont de formes identiques et nous pouvons dire que les fonctions d'autocorrelation $\hat{R}(i)$ et $R(i)$ sont représentées par la même équation matricielle 3.37. Ainsi nous pouvons conclure que les fonctions d'autocorrelation $\hat{R}(i)$ et $R(i)$ doivent satisfaire l'équation suivante:

$$\hat{R}(i) = C R(i) \quad 0 \leq i \leq P \quad 3.59$$

C est une constante que l'on détermine au cours du calcul du gain. Afin de conserver l'égalité des énergies de la réponse impulsionnelle de $G(z)$ et du signal original, il faut que $\hat{R}(0) = R(0)$. Ainsi d'après la relation 3.59 nous pouvons conclure que la constante C doit être égale à l'unité.

Donc :

$$\hat{R}(i) = R(i) \quad 0 \leq i \leq P \quad 3.60$$

Et la relation 3.58 devient :

$$R(0) = A^2 + \sum_{k=1}^P a(k) R(k)$$

Nous obtenons aussi :

$$A^2 = R(0) - \sum_{k=1}^P a(k) R(k) \quad 3.61$$

On constate que le deuxième membre de l'équation 3.61 est identique au deuxième membre de l'équation 3.40, donc

$$E(P) = A^2 = R(0) \quad V(P) = R(0) \left(1 - \sum_{k=1}^P a(k) r(k) \right) \quad 3.62$$

Pour $k=0$, $A^2 = E(0) = R(0)$

A^2 est donc égale à l'erreur quadratique totale minimale .

CH 4. RESOLUTION DES EQUATIONS DE PREDICTION LINEAIRE

4.1 Méthodes de résolution des équations.

Comme on l'a vu, pour chacune des trois méthodes de Prédiction linéaire présentées précédemment, nous pouvons calculer les coefficients de prédiction $a(k)$, $1 \leq k \leq P$ en résolvant un système de P équations à P inconnues du type $AX = B$. Ils existent plusieurs méthodes de résolution des systèmes d'équations linéaires, telles que les méthodes itératives du type Gauss-Seidel et Jacobi, mais ont l'inconvénient que leur temps de calcul est a priori indéterminé. Par contre, ils existent des méthodes directes, telles que la triangularisation de Gauss et la méthode Gauss-Jordan dont la première méthode utilise $P(2P^2 + 6P - 2)/6$ opérations. Ainsi le nombre d'opération est très grand, cela nous amène à chercher d'autres méthodes qui utilisent moins d'opérations et moins d'encombrement mémoire. Les coefficients $C(i, k)$ constituent une véritable matrice de covariance tandis que les coefficients $R|i-k|$ constituent une matrice de covariance du type particulier. Du fait que ces matrices de covariance sont symétriques et définies positives, on peut résoudre les équations 3.31 et 3.35 par une méthode rapide appelée méthode de décomposition de Cholesky ou de la racine carrée. Cette méthode, lors de sa programmation sur calculateur nécessite un encombrement mémoire plus petit que celui qui est nécessaire avec les autres méthodes. Du fait de la forme particulière de la matrice d'autocorrelation, on peut

résoudre les équations d'autocorrelation par une méthode récursive appelée "Méthode de Levinson" qui est rapide et occupe une place réduite en mémoire. Mais dans cette méthode aucune hypothèse n'est faite sur la nature de la matrice colonne du deuxième membre de l'équation 3.37. On remarque que cette matrice colonne est constituée par les mêmes éléments de la matrice d'autocorrelation $R_{(i-k)}$. De ce fait, nous résolvons notre système d'équations linéaires par une méthode plus rapide que celle de Levinson, cette méthode est appelée "méthode de Durbin". Cette méthode utilise $P(P+1)$ opérations et une taille mémoire très réduite. Il est intéressant de comparer les quantités d'opérations nécessaires à la résolution des équations (Tableau 4.1).

Méthode	Taille mémoire	nombre d'opérations
Triangularisation de Gauss	P^2	$P(2P^2+6P-2)/6$
Cholesky	$P(P+1)/6$	$P(P^2+6P+11)/6$
Durbin	$2P$	$P(P+1)$

TABLEAU 4.1 : Le nombre d'opérations et taille mémoire nécessaire aux principales méthodes de prédiction linéaire.

4.1.1 Résolution du système d'équations de la méthode de Covariance.

Méthode de Cholesky :

Nous avons le système d'équations suivant :

$$\sum_{k=1}^p a(k) C(i,k) = C(i,0) \quad , \quad 1 \leq i \leq p \quad 3.31$$

Nous écrivons ce système d'équation sous la forme matricielle suivante:

$$C.A = B \quad 4.1$$

où C est la matrice (p, p) constituée des éléments $C(i, k)$.

A est la matrice colonne constituée d'éléments $a(k)$, $1 \leq k \leq p$.

B est la matrice colonne constituée d'éléments $C(i, 0)$, $1 \leq i \leq p$.

La matrice C est une matrice symétrique, définie positive, alors elle peut s'écrire sous la forme suivante :

$$C = V D V^T \quad 4.2$$

où V est une matrice triangulaire inférieure dont les éléments de la diagonale principale sont égaux à l'unité et D est une matrice diagonale.

$$V = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ v(2,1) & 1 & & & 0 \\ v(3,1) & v(3,2) & 1 & & \\ \vdots & \vdots & & \ddots & \\ v(p,1) & v(p,2) & \dots & & 1 \end{bmatrix}, \quad D = \begin{bmatrix} d(1) & 0 & 0 & \dots & 0 \\ 0 & d(2) & 0 & & 0 \\ 0 & 0 & d(3) & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & 0 & 0 & \dots & 0 & d(p) \end{bmatrix}$$

A partir de l'équation 4.2 nous obtenons le système d'équations suivant

$$C(i, j) = \sum_{k=1}^j v(i, k) d(k) v(j, k) \quad 1 \leq j \leq i-1 \quad 4.3$$

Nous déterminons les éléments des matrices V et D en résolvant le système d'équations 4.3.

Nous pouvons déduire de la relation 4.3 que :

$$C(i, j) = \sum_{k=1}^{j-1} V(i, k) d(k) V(j, k) + V(i, j) d(j) V(j, j)$$

Or $V(j, j) = 1$

D'où $V(i, j) d(j) = C(i, j) - \sum_{k=1}^{j-1} V(i, k) d(k) V(j, k)$ 4.4

avec $1 \leq j \leq i-1$

Et pour les éléments de la diagonale $C(i, i)$, nous déduisons que :

$$\begin{aligned} C(i, i) &= \sum_{k=1}^i V(i, k) d(k) V(i, k) \\ &= \sum_{k=1}^i V(i, k)^2 d(k) \\ &= \sum_{k=1}^{i-1} V(i, k)^2 d(k) + V(i, i)^2 d(i) \end{aligned} \quad 4.5$$

Or $V(i, i) = 1$

D'où $d(i) = C(i, i) - \sum_{k=1}^{i-1} V(i, k)^2 d(k) \quad i \geq 2$ 4.6

Avec la condition initiale :

$$d(1) = C(1, 1) \quad 4.7$$

Connaissant $d(1)$, nous pouvons calculer les éléments $V(i, 1)$ avec l'équation 4.4. Puis on détermine $d(2)$ avec l'équation 4.6.

Ayant déterminés tout les éléments des matrices V et D , il est simple de calculer récursivement les éléments de la matrice colonne A , cela en passant par deux procédés.

A partir des équations 4.1 et 4.2 nous pouvons déduire que :

$$VDV^T A = B \quad 4.8$$

Nous posons :

$$DV^T A = Y \quad 4.9$$

Donc

$$VY = B \quad 4.10$$

De l'équation 4.9 nous obtenons :

$$V^T A = D^{-1} Y \quad 4.11$$

A partir de l'équation 4.10 et de la matrice V nous pouvons déduire que :

$$Y(i) = B(i) - \sum_{j=1}^{i-1} v(i,j) Y(j), \quad P \geq i \geq 2 \quad 4.12$$

Pour $i=1$, nous obtenons la condition initiale suivante :

$$Y(1) = B(1) \quad 4.13$$

L'équation 4.11 nous permet de trouver la relation récursive suivante :

$$a(i) = \frac{Y(i)}{d(i)} - \sum_{j=i+1}^P v(j,i) a(j), \quad 1 \leq i \leq P-1 \quad 4.14$$

Pour $i=P$, nous obtenons la condition initiale suivante :

$$a(P) = \frac{Y(P)}{d(P)} \quad 4.15$$

Ainsi nous obtenons tous les coefficients $a(k)$ en utilisant les équations 4.4, 4.6, 4.12 et 4.14.

Erreur quadratique de prédiction

L'erreur quadratique totale pour cette méthode s'écrit :

$$E(P) = C(0,0) - \sum_{k=1}^P a(k) C(0,k) \quad 4.16$$

Sous forme matricielle, l'équation 4.16 s'écrit :

$$E = C(0,0) - A^T B \quad 4.17$$

Nous allons chercher l'erreur quadratique en fonction des éléments des matrices Y et D .

L'équation 4.11 peut s'écrire sous la forme suivante

$$(V^T A)^T = (D^{-1} Y)^T$$

Du fait que V est une matrice inférieure et D une matrice diagonale nous pouvons écrire :

$$\begin{aligned} A^T V &= Y^T (D^{-1})^T \\ &= Y^T D^{-1} \end{aligned}$$

Nous obtenons la relation suivante :

$$A^T = Y^T D^{-1} V^{-1}$$

Nous substituons A^T par sa valeur dans l'équation 4.17 :

$$E = C(0,0) - Y^T D^{-1} V^{-1} B \quad 4.18$$

De même nous substituons B par VY :

$$E = C(0,0) - Y^T D^{-1} Y \quad 4.19$$

A partir de l'équation 4.19 nous pouvons déduire que :

$$E(P) = C(0,0) - \sum_{k=1}^P Y(k)^2 / d(k) \quad 4.20$$

4.1.2 Résolution du système d'équations de la méthode d'autocorrelation

Méthode ou algorithme de Durbin

Cette méthode suit la procédure suivante

$$E(0) = R(0) \quad 4.21$$

$$K(i) = \left(R(i) - \sum_{j=1}^{i-1} a^{(i-1)}(j) R(i-j) \right) / E(i-1) \quad 1 \leq i \leq P \quad 4.22$$

$$a^{(i)}(i) = K(i) \quad 4.23$$

$$a^{(i)}(j) = a^{(i-1)}(j) - K(i) a^{(i-1)}(i-j) \quad 1 \leq j \leq i-1 \quad 4.24$$

$$E(i) = (1 - K(i)^2) E(i-1) \quad 4.25$$

La solution est obtenue après P itérations des relations 4.21 à 4.25 avec l'indice variant de 1 jusqu'à P . Les coefficients du prédicteur sont donnés par :

$$a(j) = a^{(P)}(j)$$

On relève que dans la résolution du système d'équations pour obtenir les coefficients $a(j)$ d'un prédicteur d'ordre P , on calcule les coefficients de tous les prédicteurs d'ordre inférieur à P .

Les coefficients $K(i)$ qui apparaissent dans l'algorithme de Durbin sont appelés "coefficients de corrélation partielle" ou "coefficients de réflexion". Dans les applications de la prédiction linéaire à l'analyse du signal vocal, ces coefficients sont très importants

$$|K(i)| < 1 \quad 4.26$$

La relation 4.26 est une condition nécessaire et suffisante pour que

le filtre modèle soit stable.

Un autre élément du procédé de Durbin est le calcul de l'erreur quadratique totale minimale $E(i)$ à chaque itération. À partir des relations 4.25 et 4.26, nous pouvons vérifier que $E(i)$ décroît (ou reste invariable) avec l'accroissement du prédicteur. $E(i)$, qui est une erreur quadratique ne peut être négative. Par conséquent la relation suivante doit être vérifiée :

$$0 \leq E(i) \leq E(i-1)$$

4.2. Stabilité du prédicteur linéaire.

Toutes les méthodes décrites précédemment ont pour objet de déterminer les valeurs optimales des paramètres $a(k)$ de la fonction $G(Z)$ qui ne possède que des pôles. Or, il est important de savoir si le filtre $G(Z)$ est stable ou non. La condition que doit satisfaire $G(Z)$ pour sa stabilité soit assurée, est que tout ses pôles soient situés à l'intérieur du cercle unité dans le plan des Z . Ces pôles sont donnés par les racines du dénominateur de $G(Z)$, ce dénominateur est le polynôme $H(Z)$ qui s'écrit :

$$H(Z) = 1 - \sum_{k=1}^P a(k) Z^{-k}$$

Des trois méthodes de prédiction linéaire décrite précédemment, seule la méthode d'autocorrélation garantit la stabilité. Cette propriété résulte de la théorie des polynômes orthogonaux. Ainsi, l'analyse du signal de parole par la méthode d'autocorrélation en vue d'effectuer une synthèse

est très avantageuse, il suffit d'utiliser directement en synthèse les coefficients $a(k)$ qui constituent le prédicteur sans avoir à tester auparavant la stabilité du prédicteur. Dans le cas des méthodes fondées sur la covariance, la stabilité ne peut être garantie en général. Pour vérifier la stabilité, on fait appel à la relation 4.26. Pour cela, il faut calculer les coefficients de réflexion $K(i)$ à partir des relations récurrentes suivantes :

$$K(i) = a_{(i)}^{(i)}$$

$$a_{(j)}^{(i-1)} = (a_{(j)}^{(i)} - a_{(i)}^{(i)} a_{(i-j)}^{(i)}) / (1 - K(i)^2)$$

$$1 \leq j \leq i-1$$

où l'indice i prend successivement les valeurs $p, p-1, \dots, 1$

CH5. CALCUL DE LA PERIODE FONDAMENTALE DU SIGNAL VOCAL

DECISION VOISE / NON-VOISE

L'extraction de la fréquence fondamentale ou mélodie du signal vocal est un problème auquel on s'intéresse depuis longtemps. Qualitativement la mélodie est cette caractéristique subjective de la voix à laquelle on attribue une valeur dans une gamme qui s'étend des voix dites basses aux voix dites hautes. La mélodie est un phénomène quasi-périodique, ce sont non seulement la période et l'amplitude de l'onde de source qui varient mais également sa forme. D'où ces propriétés suffisent à dire que la mesure de cette période fondamentale est rendue si difficile.

5.1 Calcul de la période fondamentale.

La méthode est fondée sur la détection des maximums de la fonction d'autocorrélation. Nous savons que la fonction d'autocorrélation d'un signal périodique est elle-même périodique de même période que le signal. Elle présente, dans ce cas, des maximums d'amplitude comparable à la valeur de la fonction d'autocorrélation pour un retard nul, chacun d'eux étant séparé du précédent, sur l'axe des retards τ , par un intervalle de temps égal à la période du signal.

Par ailleurs, si le signal ne présente pas de périodicité notable, la fonction d'autocorrélation de ce signal ne possède alors aucun maximum en dehors de celui obtenu pour un retard nul. Cette propriété permet, dans une

certaines mesures, de déterminer si le signal est périodique ou non.

Ainsi, nous calculons la période du signal, en mesurant l'intervalle de temps séparant le maximum $R(0)$ de la fonction d'autocorrélation du premier maximum d'amplitude comparable.

5.2 Décision "voisé/non-voisé"

On peut estimer que le signal de parole pour les sons non-voisés est considéré comme un bruit blanc gaussien de valeur moyenne nulle. Ainsi théoriquement quand la séquence de signal analysé devient grande, la fonction d'autocorrélation $R(n)$ tend vers zéro pour n différent zéro. Et pour une longueur finie de la séquence du signal analysé, on peut déterminer un seuil tel que, avec une certaine probabilité et pour un certain intervalle de confiance, aucun échantillon de la fonction d'autocorrélation (excepté celui de l'origine) ne dépasse la valeur de ce seuil.

Nous pouvons obtenir ce seuil, noté α , qui satisfait la relation suivante

$$P_r(r \leq \alpha) = 0,01B$$

Ce seuil est obtenu à partir d'un graphe de la fonction $\alpha = f(N)$ pour des intervalles de confiance différents

Où N est le nombre d'échantillons du signal, B est l'intervalle de confiance.

$r = R(n)/R(0)$, avec $R(n)$ est le premier maximum de la fonction d'autocorrélation à partir de $R(0)$, dans la séquence considérée.

D'où la valeur de r comparée à la valeur du seuil α , nous permet de déduire

si la séquence du signal analysé est voisé ou non-voisé.

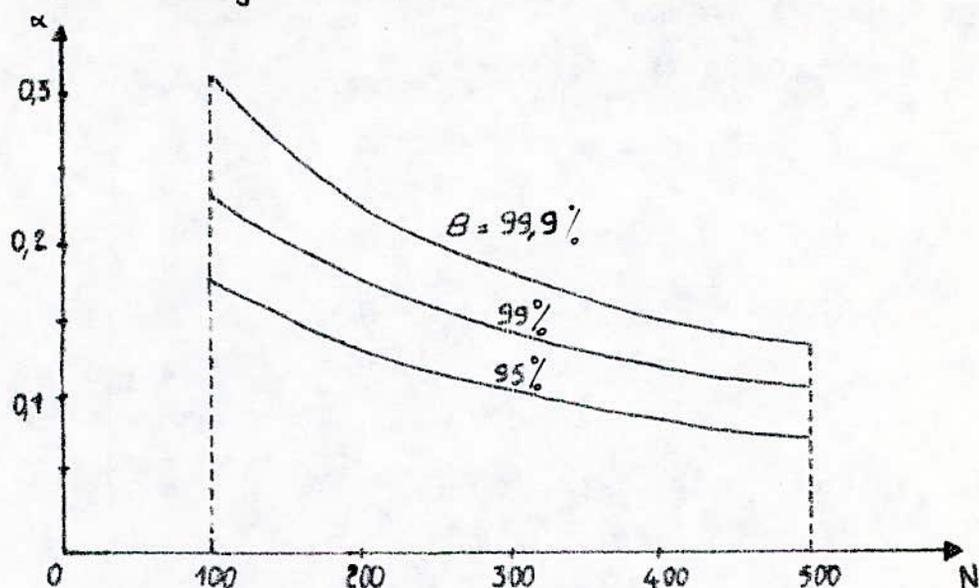


FIGURE 5.1 - Courbe du seuil α en fonction du nombre d'échantillons du signal pour différentes valeurs de l'intervalle de confiance B .

A partir de ce graphe, nous pouvons constater, par exemple que pour $\alpha = 0,2$ et $N \geq 250$, la probabilité d'avoir une erreur (c'est à dire qu'un maximum de la fonction d'autocorrelation normalisée r soit supérieur au seuil α pour des échantillons de bruit) est inférieur à 0,001. Cependant, ce graphe est rigoureusement valable pour un bruit blanc gaussien. Le signal de parole, dans le cas des sons non-voisés, peut quelque fois être assez différent d'un bruit gaussien, alors on est conduit à relever ce seuil α , pour ne pas commettre d'erreurs dans la décision "voisé/non-voisé".

Le nombre d'échantillons utilisé pour calculer la fonction d'autocorrelation à court terme a été fixé à 200. D'après la figure 5.1, on constate que la valeur α , correspondant à ce nombre d'échantillons ($N=200$) pour un intervalle de confiance β supérieur à 99% est alors environ 0,2. Comme indiqué précédemment, on relève ce seuil; il a été déterminé par expérience qu'une bonne valeur de ce seuil est d'environ 0,5.

Dans cette discrimination des sons voisés et non-voisés, on constate que ce principe tend à renforcer la décision en faveur des sons non-voisés. En effet la figure 5.1 indique que l'on est sûr qu'il y aura moins de 1% des signaux de bruits gaussien dont la fonction d'autocorrelation à court terme normalisée, calculée dans les conditions indiquées ci-dessus, présentera un maximum (en dehors de sa valeur pour un retard nul) supérieur à cette valeur α du seuil. Ce cas se présente en effet à la fin des segments de parole voisés pour lesquelles la période de mélodie varie dans une large proportion. Les périodes du signal sont alors faiblement corrélées et la valeur du maximum de la fonction d'autocorrelation qui correspond à une valeur moyenne de la période de mélodie, est de ce fait notablement réduite; elle peut devenir inférieure à la valeur α du seuil et la séquence de parole sera déterminée comme étant non-voisée. Pour éliminer ce type d'erreur, on est amené à considérer un second seuil dont la valeur est inférieure à d'environ 15% à celle du seuil α et à corriger la décision voisé / non-voisé en fonction de la

valeur du maximum de la fonction d'autocorrélation relativement à ce seuil.

Au cours du test de la décision voisé/non-voisé d'une séquence K , on est conduit à faire des corrections sur la décision de la séquence précédente.

Cette correction suit la procédure suivante :

Si la séquence $K-2$ est voisé et la séquence $K-1$, non-voisé, alors la séquence $K-1$ est redéfinie comme voisé si la séquence K est déterminée comme voisé. La valeur de la période fondamentale T_0 pour cette séquence est prise égale à la valeur moyenne des périodes pour les séquences K et $K-2$. Ce cas peut se présenter occasionnellement pour des transitions rapides de la mélodie car l'analyse avec l'aide d'une fenêtre assez large (environ 20 ms) fournit seulement une valeur moyenne de la période de mélodie.

Nous donnons un algorithme du calcul de la période fondamentale et de la décision voisé/non-voisé.

- entrée de la séquence K
- Si $n > \alpha$ alors la séquence K est considérée comme voisé.
- La période fondamentale est égale à nT_e (n est le numéro du premier maximum de la fonction d'autocorrélation de cette séquence K).
- Si la séquence $K-2$ est voisé et la séquence $K-1$ est non-voisé alors la séquence $K-1$ est redéfinie comme voisé.
- La période fondamentale T_0 pour la séquence $K-1$ serait :

$$T_0(k-1) = \frac{T_0(k) - T_0(k-2)}{2}$$

- Si $r < \alpha$ et les deux séquences $k-2$ et $k-1$ sont définies comme non voisées alors la séquence k est définie comme non-voiséé .
- Si $r < \alpha$ et les deux séquences $k-2$ et $k-1$ sont voisée alors on réduit la valeur du seuil α de 15% .
- Si $r > 0,85 \alpha$ alors la séquence k est voisée , dans le cas inverse la séquence k est non-voiséé .

PARTIE II

PROGRAMMATION

CH1. PRESENTATION DE L'APPLE II PLUS

1.1 Introduction :

L'Apple II plus est un microordinateur transportable de fabrication américaine. Il a été conçu en 1979 par STEVEN JOBS et STEPHEN WORNIAK. Ses domaines d'utilisation sont nombreux (Scientifiques, Gestion applications domestiques etc...)

1.2 Le Hardware

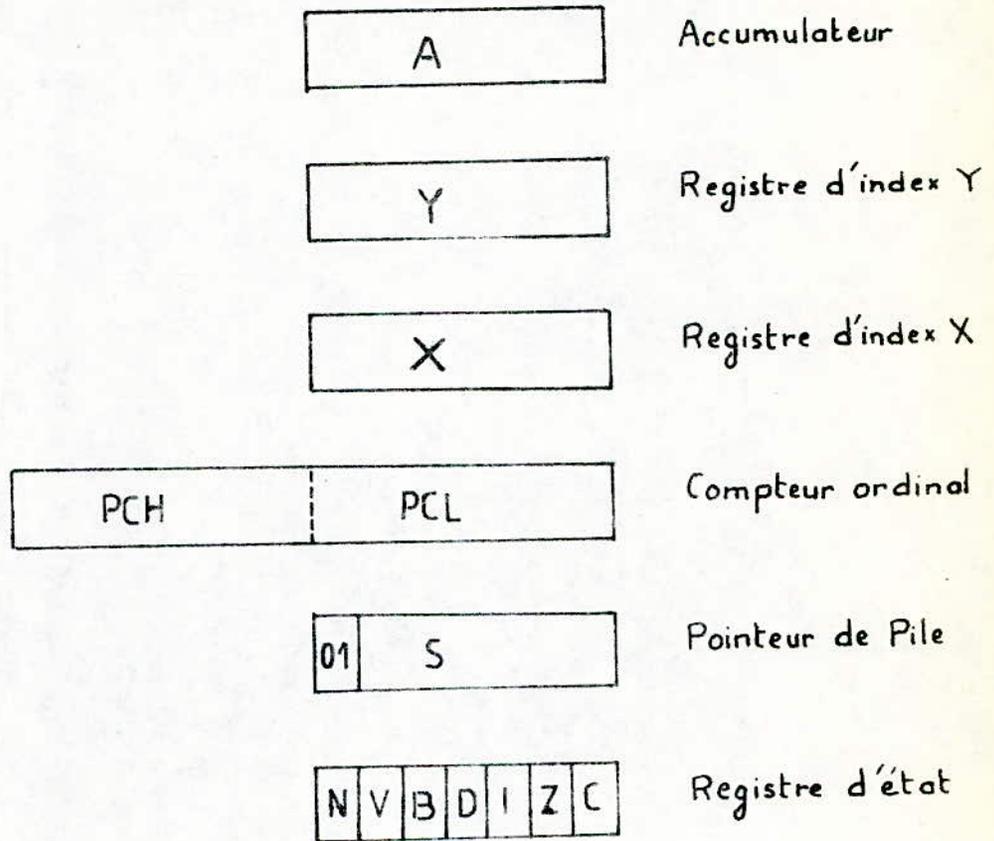
1.2.1 L'unité centrale.

L'unité centrale de l'Apple II est composée autour du microprocesseur MCS 6502 / SY6502.

Fiche technique du MCS 6502

- Modèle MCS 6502 / SY6502
- Fabriqué par MOS Technologie, Inc. Synerlek, Rockwell.
- nombre d'instructions 56.
- Modes d'adressage 13.
- Voltage : 5 Volts
- Puissance : 0,25 watt
- Horloge : 1,023 MHz

ARCHITECTURE INTERNE DU MCS 6502 / SY 6502



- Indicateurs

N : signe

V : dépassement

B : arrêt

D : décimal arithmétique

I : interruption

Z : zéro

C : retenue

- Les bus

Bus d'adresses 16 bits (parallèles)

Bus de données 8 bits (parallèles, bidirectionnels).

1.2.2 La mémoire centrale.

Après le microprocesseur, la mémoire centrale est l'organe le plus important dans un ordinateur.

La capacité mémoire de l'Apple II est de 64 K octets réparties en RAM et ROM.

- ROM (mémoire morte) occupe une place mémoire de 12 K octets, répartis en 6 circuits intégrés de 2 K chacun

- RAM (mémoire vive)

Les mémoires RAM utilisées par le microordinateur Apple II plus sont des mémoires RAM dynamiques de 16 K bits (type "4116") à lecture et écriture. L'Apple II utilise 24 boîtiers de ce type, donc une capacité totale de ces mémoires de 48 K Octets places mémoires.

- Les I/O (entrée/sortie) occupent 4 K Octets places mémoires, dont 2 K Octets sont placés en ROM.

Organisation mémoire :

N° de Page		
0 1 2 ...	RAM (48K)	
190 191		
192 ...		I/O (2K)
199 200 ...		I/O ROM (2K)
207	ROM (12K)	
208		
209 ...		
254		
255		

La mémoire centrale est constituée de 256 "pages mémoires" répartis en :

RAM → 192 pages

I/O → 16 pages

ROM → 48 pages

Une page mémoire occupe 256 "locations mémoires".

1.2.3 Les unités périphériques.

Les unités périphériques comprennent des dispositifs d'entrée, de sortie et de stockage.

1.2.3.1 Le clavier :

Il comprend 52 touches à caractère alphanumérique dont cinq touches spéciales qui sont les suivantes :

- CTRL est la touche de contrôle.
- RESET : commande la "mise à zéro".
- REPT : est la touche de répétition.
- ESC : espace (échappement) commande d'effacement de l'écran.
- ← → : commande le déplacement à gauche ou à droite.

1.2.3.2 La visualisation :

Elle est assurée par un T.V N/B "Sonelec" qui reçoit directement la sortie vidéo du microordinateur sur ampli vidéo.

1.2.3.3 Les connecteurs d'interfaces :

Ils sont au nombre de 8 et sont reliés au Bus de l'Apple II. Ils

peuvent recevoir différentes cartes d'interfaces pour adopter des unités de disques, imprimantes, etc...

1.2.34 Les drivers de disques

Ils sont au nombre de deux et sont pilotés par un seul contrôleur. La mémoire disponible n'étant que de 48 K octets places mémoires, celle-ci ne répond pas aux besoins de l'utilisateur. De ce fait, il faut recourir à d'autres espaces mémoires (mémoire auxiliaire ou de masse); d'où l'utilité des disquettes.

La capacité mémoire d'une disquette est de 140 K Octets.

1.3 Le software

1.3.1 Le software de base (ou support logiciel)

La tâche du programmeur est de réaliser et de charger des programmes que le microprocesseur exécute. Ces programmes sont appelés "programme d'application". Pour faciliter la réalisation de ces programmes, un ensemble de programme d'aide est nécessaire. Cet ensemble constitue le support logiciel.

Cet ensemble comprend :

1.3.1.1 Le programme d'exploitation

1.3.1.1.1 Le programme moniteur.

Il gère l'enchaînement des tâches et détecte les commandes

de l'opérateur tout en y répondant correctement. Il permet ainsi l'accès aux disques, à organiser la mémoire etc ...)

1.3.1.1.2 Le programme éditeur

C'est un programme permettant de composer, modifier et stocker sur une mémoire de masse, un texte, à partir d'un terminal conversationnel (clavier et écran). Le texte ainsi composé peut être un programme, un fichier de données, un message, etc ...).

1.3.1.1.3 Les programmes de traduction :

Ils assurent la transcription d'un programme source en un programme objet, suivant le langage utilisé (assembleur ou évolué).

Le programme est appelé assembleur ou interpréteur.

Pour l'Apple II, on dispose d'un mini-assembleur, d'un interpréteur pour l'utilisation de l'Applesoft Basic et d'un compilateur pour le pascal.

1.3.1.1.4 Les programmes de diagnostic (DEBUG) :

Ils permettent de détecter des erreurs éventuelles de logique. Dans l'Apple II cette fonction est remplie par les programmes suivants :

- SYSTEM SYNTAX : pour les langages évolués.
- 6500 ERRORS

1.3.2 Langages disponibles dans l'Apple II PLUS

- Le langage machine
- Le langage assembleur
- Le Basic Applesoft (interprété resident)

- Le Basic integer (interprété sur disque).
- Le Pascal (compilé sur disque).

La chaîne de microprogrammation incorporée est l'AppleSoft. Les autres langages nécessitent une carte de programmation.

1.3.2.1 Le langage Basic Applesoft :

Nous avons choisi pour notre programmation le langage basic Applesoft. Ce choix a été fait pour les raisons suivantes :

- Simplicité du langage
- Sa présence directe dans la mémoire de l'Apple II plus. Cela permet d'avoir un espace plus dans la disquette.
- Son large éventail d'utilisation. En effet il permet des traitements sur fichier et présente également des possibilités étendues en mathématique, logique et graphisme

1.3.2.1.1 Principales caractéristiques.

- Utilisation d'un vocabulaire restreint.
- Adjonction systématique d'une étiquette à chaque instruction sous forme de numéro de ligne (0 - 63999)
- Utilisation d'un seul type de variables pour les calculs (variables réelles représentées en virgule flottante).

La numération de chaque ligne d'instruction est obligatoire, elle définit strictement comment vont s'enchaîner les opérations à exécuter.

1.3.2.1.1 Quelques commandes relatives au basic Applesoft.

Commandes d'entrée/sortie

INPUT A\$: affiche ? sur l'écran et attend qu'une chaîne de caractères soit tapée au clavier.

GET A\$: attend un seul caractère du clavier sans "RETURN"

DATA X,Y : établit une liste de données à exploiter par READ

READ A\$: affecte à A\$ la prochaine donnée de DATA.

PRINT "X="X : affiche X = suivie de la valeur de la variable X.

Commandes de branchement.

GOTO XY : branchement à la ligne XY

GO SUB : branchement à un sous-programme.

RETURN : marque la fin d'un sous-programme, renvoie à la prochaine instruction suivant GO SUB.

IF ... THEN ... ELSE : si la condition est vraie, l'ordre suivant THEN est exécuté. Si la condition est fausse, c'est l'ordre qui suit ELSE qui est exécuté.

Autres commandes :

RUN : lance l'exécution d'un programme.

NEW : efface la mémoire.

LIST : affiche les données rangées en mémoire centrale.

Quelques commandes utiles des disquettes.

* INIT (nom de programme)

Cette commande permet :

- d'initialiser une disquette vierge et la rend "disquette esclave".
- d'affecter éventuellement un numéro de volume V à la disquette
 $1 \leq V \leq 254$
- de sauvegarder sur la disquette le programme écrit en Basic, actuellement en mémoire sous le nom de programme.

* CATALOG, (S6), (D1)

Cette commande permet :

- L'affichage de la liste des fichiers enregistrés sur la disquette.
- L'affichage du numéro de volume de la disquette présente dans le lecteur numéro 1, connecté au slot numéro 6.

* SAVE (nom de programme).

Cette commande permet :

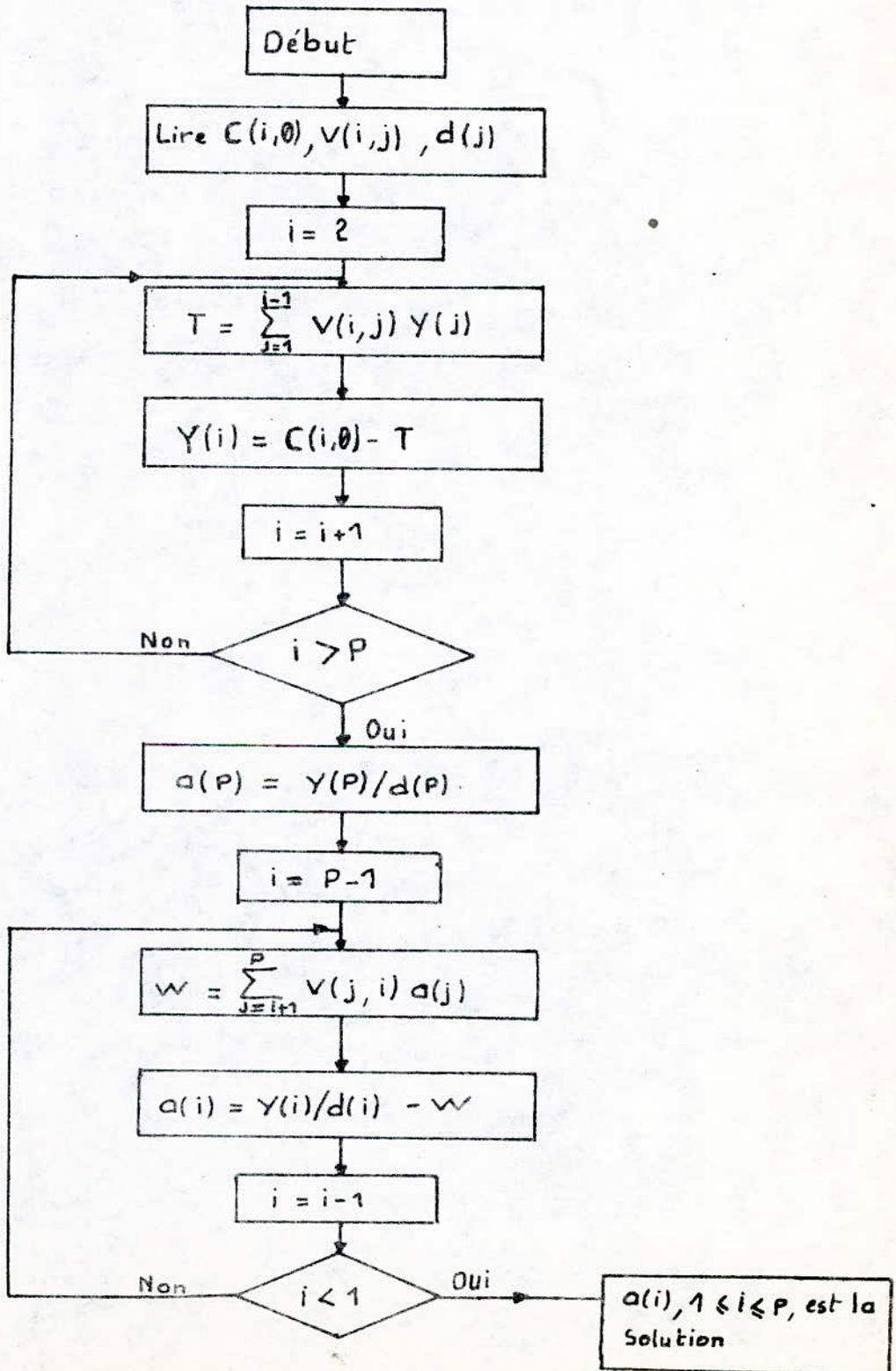
- La sauvegarde ou l'enregistrement sur la disquette du programme.
- L'effacement de la disquette de tout programme portant le même nom et de même type.

* LOAD (nom du programme)

Cette commande permet le chargement du programme, depuis la disquette vers la mémoire RAM

CH.2. ETUDE DE PROGRAMMES

Organigramme de calcul des coefficients de Prédiction de la méthode de covariance.



Commentaire de l'organigramme.

Dans cet organigramme, nous avons cherché les coefficients de prédiction $a(i)$, $1 \leq i \leq P$, en calculant les éléments $\gamma(i)$ de la matrice Y à partir des éléments $C(i,0)$ et $V(i,j)$ des matrices B et V respectivement. Puis dans un second procédé nous calculons les coefficients $a(i)$ à partir de $\gamma(i)$, $d(j)$ et $V(i,j)$.

Nous remarquons que les éléments $C(i,0)$, $d(j)$ et $V(i,j)$ sont connus dans cet organigramme.

Programme de calcul des coefficients de prédiction $a(i)$ et de l'erreur quadratique de la méthode de Covariance.

```

10  REM RECHERCHE DES ELEMENTS DE LA MATRICE C(I, J)
20  FOR I = 1 TO P
30  FOR K = 0 TO P
40  C(I, K) = 0
50  FOR M = P TO N-1
60  C(I, K) = C(I, K) + S(M-I) * S(M-K)
70  NEXT M
80  NEXT K
90  NEXT I
100 REM RECHERCHE DES ELEMENTS DE LA MATRICE V(I, J)
110 REM ET DE LA MATRICE D(J)
120 D(1) = C(1, 1)
130 J = 1
140 V(J, J) = 1
150 FOR I = J+1 TO P
160 T = 0
170 FOR K = 1 TO J-1
180 T = T + V(I, K) * D(K) * V(J, K)
190 NEXT K

```

```

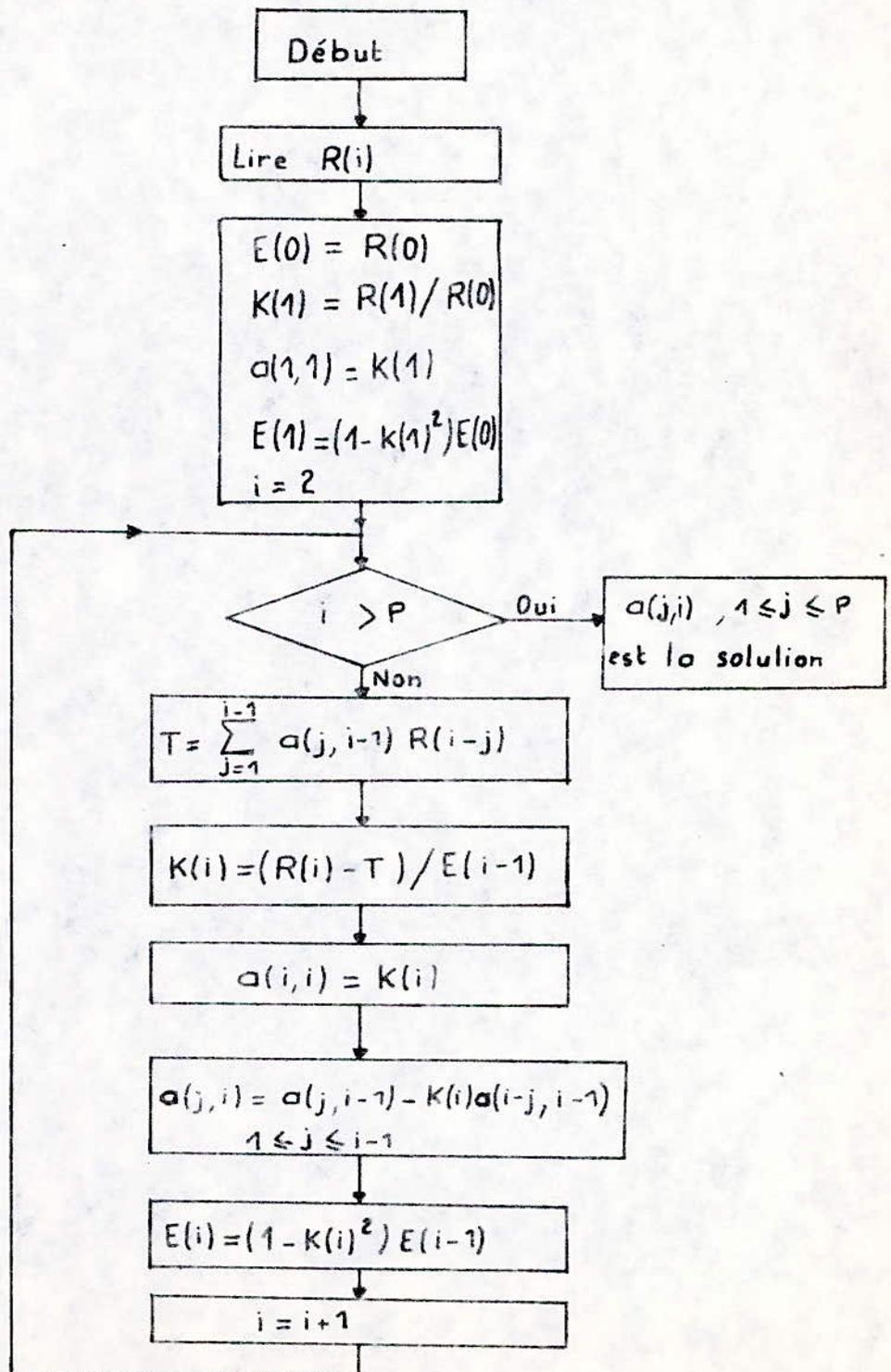
200  V(I, J) = (C(I, J) - T) / D(J)
210  NEXT I
220  J = J + 1
230  F = 0
240  FOR K = 1 TO J - 1
250  F = F + (V(J, K) ^ 2) * D(K)
260  NEXT K
270  D(J) = C(J, J) - F
280  IF J <= P THEN 140
300  REM RECHERCHE DES ELEMENTS DE LA MATRICE Y
310  Y(1) = C(1, 0)
320  FOR I = 2 TO P
330  G = 0
340  FOR J = 1 TO I - 1
350  G = G + V(I, J) * Y(J)
360  NEXT J
370  Y(I) = C(I, 0) - G
380  NEXT I
390  REM RECHERCHE DES COEFFICIENTS DE PREDICTION A(I)
400  A(P) = Y(P) / D(P)
410  I = P - 1
420  H = 0

```

```
430 FOR J = I+1 TO P
440 H = H + V(J, I)* A(J)
450 NEXT J
460 A(I) = ( Y(I)/D(I) ) - H
470 I = I - 1
480 IF I < 1 THEN 500
490 GOTO 420
500 REM RECHERCHE DE L'ERREUR QUADRATIQUE E(P)
510 C(0,0) = 0
520 FOR M = 0 TO N-1
530 C(0,0) = C(0,0) + S(M)^2
540 NEXT M
550 W = 0
560 FOR K=1 TO P
570 W = W + ( Y(K)^2 ) / D(K)
580 NEXT K
590 E(P) = C(0,0) - W
600 REM RECHERCHE DES COEFFICIENTS DE REFLEXION K(I)
610 FOR I = 1 TO P
620 B(I, P) = A(I)
630 NEXT I
640 I = P
```

```
650 K(I) = B(I,I)
660 FOR J = 1 TO I-1
670 B(J,I-1) = (B(J,I) - B(I,I) * B(I-J,I)) / (1 - K(I)^2)
680 NEXT J
690 I = I - 1
700 IF I < 1 THEN 720
710 GOTO 650
720 END
```

Organigramme de calcul des coefficients de Prédiction pour la méthode d'autocorrelation.

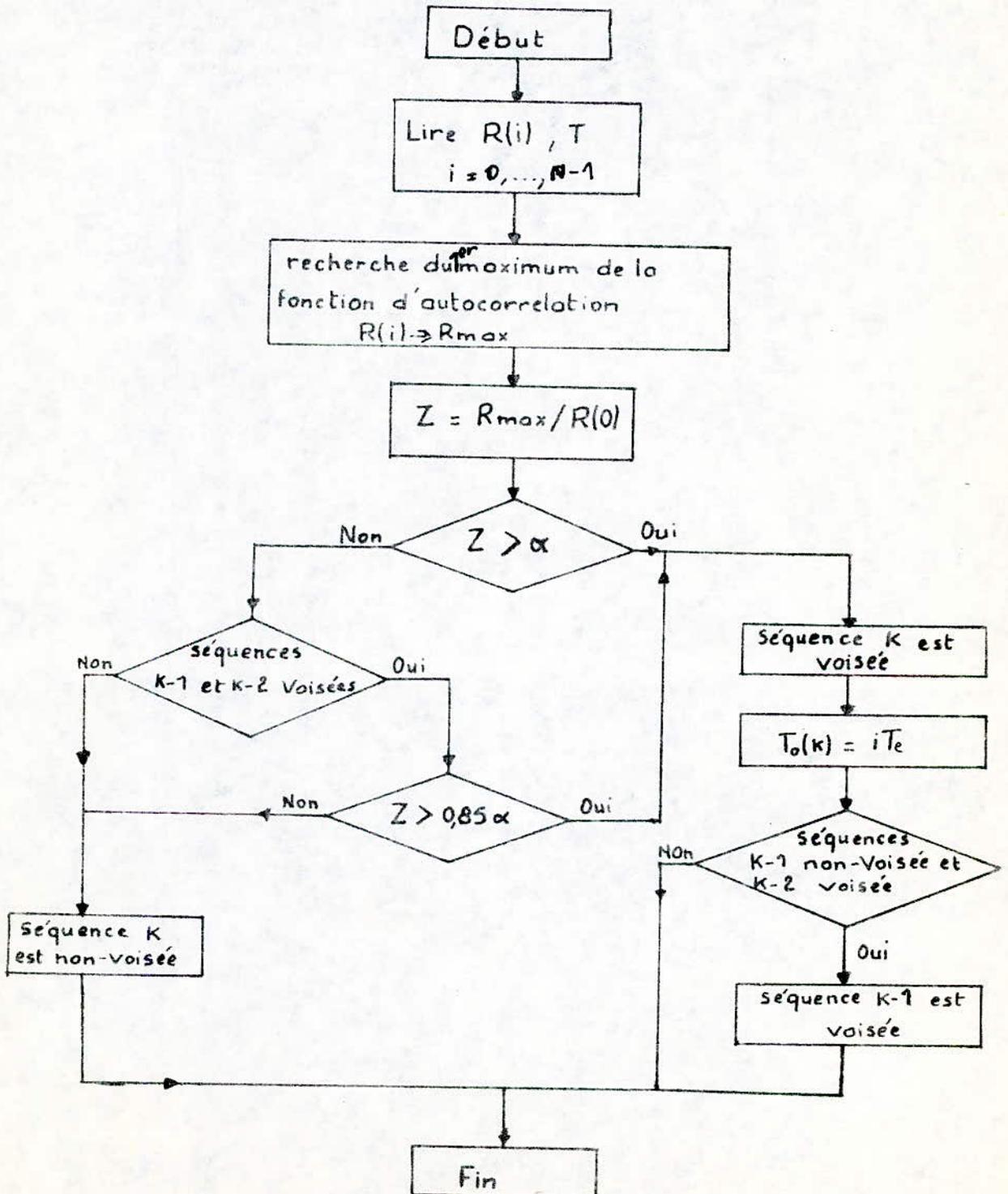


Commentaire de l'organigramme

Cette organigramme nous donne les coefficients de prédiction $\alpha(j,P)$ tel que $1 \leq j \leq P$, du prédicteur d'ordre P . Aussi nous pouvons avoir tous les coefficients des prédicteurs d'ordre inférieur à P .

Nous pouvons trouver toutes les valeurs de l'erreur quadratique et tous les coefficients de reflexion de chaque prédicteur.

Organigramme de décision voisé/nonvoisé et calcul de la période fondamentale de la séquence k.



Commentaire sur l'organigramme

Nous avons essayé dans cet organigramme de faire une discrimination entre les sons voisés et non-voisés pour une séquence K du signal de parole.

Les résultats que nous fournit cet organigramme sont la séquence K voisée ou non-voisée, la valeur de la période fondamentale T_0 dans le cas où la séquence est voisée et la correction de la séquence $K-1$ si elle a lieu.

Dans la pratique, on utilise la valeur de α égale à 0,5 comme il a été indiqué précédemment. Aussi dans le programme nous prenons les notations suivantes :

- Une séquence K voisée sera notée $V(K) = 1$
- Une séquence K non-voisée sera notée $V(K) = 0$

T_e est la période d'échantillonnage

Programme de calcul des coefficients de la fonction d'autocorrelation, des coefficients de réflexions $K(i)$, de coefficients de prédiction $A(j, I)$ et de l'erreur quadratique minimale.

```

10  REM RECHERCHE DES COEFFICIENTS D'AUTOCORRELATION R(k)
20  DATA S(0), S(1), ..., S(N-1)
30  FOR I = 0 TO N-1
40  READ S(I)
50  NEXT I
60  FOR K = 0 TO N-1
70  R(K) = 0
80  FOR I = 0 TO N-1-K
90  R(K) = R(K) + S(I) * S(I + K)
100 NEXT I
110 NEXT K
120 REM RECHERCHE DES COEFFICIENTS DE REFLEXIONS K(I)
130 E(0) = R(0)
140 K(1) = R(1)/R(0)
150 A(1,1) = K(1)
160 E(1) = (1 - K(1)^2) * E(0)
170 FOR I = 2 TO P
180 T = 0

```

```

190 FOR J=1 TO I-1
200 T = T + A(J, I-1) * R(I-J)
210 NEXT J
220 K(I) = (R(I) - T) / E(I-1)
230 REM RECHERCHE DES COEFFICIENTS DE PREDICTIONS A(J, I)
240 A(I, I) = K(I)
250 FOR J=1 TO I-1
260 A(J, I) = A(J, I-1) - K(I) * A(I-J, I-1)
270 NEXT J
280 REM RECHERCHE DE L'ERREUR QUADRATIQUE E(P)
290 E(I) = (1 - K(I)^2) * E(I-1)
300 NEXT I

```

Nous poursuivons ce programme, en ajoutant le programme de décision 'voisé / non-voisé' et calcul de la période fondamentale.

```

310 REM PROGRAMME DE DECISION VOISE / NON-VOISE ET
320 REM CALCUL DE LA PERIODE FONDAMENTALE
330 REM RECHERCHE DU MAXIMUM DE LA FONCTION
340 REM D' AUTOCORRELATION
350 I = 0
360 FOR J = 2 TO N-1

```

```

370 IF R(J) / R(I) > 0.5 THEN 390
380 NEXT J
390 I = J
400 RMAX = R(I)
410 REM RECHERCHE SI LA SEQUENCE K EST VOISEE OU NON-VOISEE
420 T = 0,0001
430 Z = RMAX / R(0)
440 IF Z > 0,5 THEN 510
450 IF V(K-1) = 1 AND V(K-2) = 1 THEN 490
460 V(K) = 0
470 PRINT SEQUENCE K NON-VOISEE
480 GOTO 590
490 IF Z > 0,425 THEN 510
500 GOTO 590
510 V(K) = 1
520 PRINT SEQUENCE K VOISEE
530 REM CALCUL DE LA PERIODE FONDAMENTALE
540 T0 = I * T
550 REM CORRECTION DE LA SEQUENCE K-1
560 IF V(K-1) = 0 AND V(K-2) = 1 THEN 580
570 GOTO 590
580 PRINT SEQUENCE K-1 VOISEE
590 END

```

Exemple

Nous donnons un exemple avec lequel on utilise les tests des programmes de la méthode de covariance et la méthode d'autocorrelation. Le signal $S(n)$ est défini par :

$$S(n) = 2 \cdot (0,99)^{(n-4)} - (0,99)^{2(n-4)}$$

Nous utilisons le nombre N d'échantillons et le nombre P du prédicteur égales à $N = 24$ et $P = 4$ respectivement.

Résultats obtenus :

Methode de Covariance

i	Coefficients A(i)	Coefficients K(i)
1	0.0739428895	1.01277108
2	1.39187482	11.361593
3	0.866666667	-1.24104495
4	-1.33333333	-1.33333333

$$E(4) = 3,99373601$$

Methode d'autocorrelation.

i	Coefficients A(i)	Coefficients K(i)
1	0,979734152	0,958898808
2	-9,63731632E -05	-0,0228517706
3	-9,44832117E -05	-0,0232028484
4	-0,0235732024	-0,0235732024

$$E(4) = 1,89430281$$

Commentaire

Dans cet exemple, nous remarquons que dans la méthode de Covariance, les coefficients de reflexion $K(i)$ sont supérieurs, en valeur absolue, à l'unité. Par conséquent le filtre modèle de cette méthode est instable. Aussi nous constatons que l'erreur quadratique est grande. Par compte, nous remarquons que dans la méthode d'autocorrelation les coefficients de reflexion $K(i)$ sont tous inférieurs, en valeur absolue, à l'unité. Ainsi, le filtre résultant de cette analyse est stable.

Alors le filtre modèle est caractérisé par les coefficients $A(i)$ trouvés et aussi par la valeur du gain A égale à :

$$A = \sqrt{E(4)} = 1,3763364$$

Par conséquent, le filtre nous donne approximativement, à une excitation donnée, le signal analysé.

CONCLUSION.

Le but de notre travail était l'étude de programme pour reconstituer la parole. Pour parvenir à cette fin nous avons fait une étude de l'analyse et de la synthèse par prédiction linéaire. Seulement deux principales méthodes ont été l'objet de notre étude d'analyse, qui sont la méthode de covariance et la méthode d'autocorrelation. Nous avons constaté que la méthode d'autocorrelation est très avantageuse du fait de la bonne stabilité du filtre résultant. Aussi cette méthode, lors de la programmation, s'est révélée très rapide au cours de la résolution des équations de prédiction linéaire.

La méthode de prédiction linéaire est une méthode très rapide. Il a été déjà envisagé d'utiliser cette méthode dans les systèmes de synthèse, de reconnaissance et de transmission de la parole.

Nous nous sommes limités ici à la synthèse par une méthode de prédiction linéaire. Bien d'autres méthodes existent et en particulier, des méthodes utilisant le domaine fréquentiel comme par exemple, la synthèse par formants.

En fin nous avons fait l'étude de programme qui s'est révélée très intéressante. Nous constatons que la méthode d'autocorrelation utilise un programme occupant une place mémoire très réduite. Aussi cette étude, nous était profitable du fait qu'elle nous a permis l'accès au microordinateur APPLE II PLUS.

BIBLIOGRAPHIE

1. EL MALLAWANY (I): Dialogue homme - machine :
Le traitement du signal
2. LIFERMANN (J) : Principes du traitement statistique du signal. T.I
Les méthode classiques .
3. MARHOUL (J) : Linear Prediction : A Tutorial Review
4. MARKEL / GRAY : Linear Prediction of speech
5. RABINER / SCHAFFER : Digital Processing of Speech Signals
6. LILEN . (H) : La pratique de l'Apple II