

COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 26 au 30 AVRIL 1977

IMPLEMENTATION DE MICRO PROCESSEURS POUR LE TRAITEMENT DU SIGNAL
VOCAL EN TEMPS REEL

P.Y. CAZENAVE D. DOURS R. FACCA

C.E.R.F.I.A. - 118, route de Narbonne - 31077 TOULOUSE CEDEX -

RESUME

La mise en oeuvre des systèmes programmés de traitement du signal vocal sur des ordinateurs séquentiels ne permet pas de rendre compte de façon explicite du parallélisme existant entre les différentes tâches du système.

Ceci se traduit par une impossibilité de fonctionnement en temps réel qui déprécie fortement tous les avantages des systèmes programmes (adaptativité, souplesse, modifications, etc...).

Les caractéristiques des technologies nouvelles, en particulier leur faible coût et leur grande facilité d'assemblage et de mise en oeuvre permettent d'envisager à court terme une implémentation efficace de tels systèmes sur une architecture spécialisée en vue d'un traitement en temps réel.

Nous nous proposons de décrire ici de façon schématique une structure à micro-processeurs adaptée à un système particulier de traitement du signal vocal (projet A.R.I.A.).

Nous évoquerons surtout les problèmes liés au parallélisme dans les machines, et la façon de les traiter sans entrer dans le détail du fonctionnement des micro-processeurs eux-mêmes.

SUMMARY

The setting up of programmed systems of the treatment of the vocal signal on sequential computers does not allow us to show explicitly the parallelism existing between the different tasks of the system.

This results in the impossibility of functioning in real time which strongly depreciates the advantages of the programmed systems (adaptativity, flexibility, modifications, etc...)

The characteristics of the new technologies, especially their low cost and their great assembling and setting up capacity allow us to consider in the near future an efficacious implementation of such systems on a specialized architecture for treatment in real time.

We intend to describe here in a simplified manner a micro-processors structure adapted to a peculiar system of vocal signal treatment (A.R.I.A. project). We shall mainly recall the problems attached to the parallelism in the machines and how to treat them without entering into the details of the functioning of the micro-processors themselves.



IMPLEMENTATION DE MICRO PROCESSEURS POUR LE TRAITEMENT DU SIGNAL VOCAL
EN TEMPS REEL - P.Y. CAZENAVE - D. DOURS - R. FACCA

INTRODUCTION

Le traitement du signal vocal, que ce soit l'analyse, la reconnaissance ou la synthèse, doit s'effectuer en temps réel. Ceci impose au système de traitement d'élaborer ses résultats dans des temps cohérents avec l'évolution du signal.

Les systèmes câblés, tels que banc de filtres, etc, sont les premiers systèmes d'analyse ayant fonctionné en temps réel. Ils présentent toutefois certains inconvénients. Les résultats fournis par ces systèmes ne sont pas toujours très précis, surtout lors d'une évolution rapide du signal, ce qui est souvent le cas en parole continue. Seules des procédures très élaborées permettent d'obtenir des résultats satisfaisants dans ce cas là. Mais ces procédures sont difficilement réalisables à partir de systèmes câblés. Les systèmes programmés, permettent le développement et la mise en oeuvre rapide de telles procédures. Mais, leur multiplicité et leur complexité font que le temps réel n'est pas envisageable avec une machine de type séquentiel. Il faut donc envisager une organisation logique des systèmes de traitements et se tourner en particulier vers les machines de type parallèle sur lesquelles plusieurs processus de traitement peuvent se dérouler simultanément. L'apparition des micro-processeurs permet d'envisager la réalisation de telles machines.

PRESENTATION DU PROJET A.R.I.A.

L'étude du message parlé, qu'elle qu'en soit l'approche, s'articule selon un processus d'analyse synthèse basé sur le cycle suivant :
Emission (phonation) → Audition → Perception → Compréhension

Une façon classique d'aborder cette étude est d'admettre que l'on peut séparer les problèmes en trois groupes :

- Ceux inhérents à la structure physique du signal, qui peut se décomposer en une suite de segments chacun de ces segments représentant la réalisation d'une unité élémentaire, le phonème.

- La deuxième catégorie traite de l'aspect linguistique. Les syllabes, combinaisons de phonèmes s'articulent en mots, eux-mêmes regroupés en syntagmes. Ces syntagmes, énonçant différentes propositions, constituent la phrase.

- Les autres problèmes sont liés à la compréhension et font appel à des techniques analogues à celles développées en intelligence artificielle.

Cette étude s'inscrit dans le cadre du projet A.R.I.A.

Le projet A.R.I.A., développé au laboratoire C.E.R.F.I.A., vise à mettre au point un système de dialogue oral avec une machine selon la procédure décrite ci-dessus.

Les différents modules ont été conçus dans l'optique d'un système hiérarchisé de reconnaissance de phrases construites à partir d'un vocabulaire non limité et prononcées de façon naturelle par un ensemble de locuteurs quelconques. Ces modules organisés autour d'un processeur central (module parole), sont les suivants :

- Le module d'acquisition et de codage fournit un signal discrétisé que l'on mémorise.

- Le module de segmentation découpe le signal en segments, tels que chacun d'eux corresponde à la réalisation d'un phonème, et les classe dans un des trois catégories : plosives, sons voisés, sons fricatifs (1)

- Le module d'analyse prend en charge les différents segments et les aiguille vers l'un des trois sous modules d'analyse, correspondant chacun à un type de segment. Il fournit les vecteurs paramètres au module de reconnaissance (2)

- Le module de reconnaissance est constitué de quatre sous-modules qui dépendent du type d'analyse effectuée. Il permet l'identification des différents segments en tant que phonèmes et la génération de la chaîne phonémique (2)

- Le module de syllabation, à partir d'une classification des phonèmes par niveau énergétique, analyse les variations caractéristiques de cette énergie et découpe la chaîne phonémique en syllabes (3).

- Le module d'analyse lexicale, basée sur une méthode ascendante, effectue le traitement spécifique des élisions, des liaisons, des insertions, des enchaînements, etc. Il tient compte des marqueurs prosodiques et phonologiques. Un ensemble de règles et de filtres permettent d'éliminer les suites lexicales ayant une trop faible consistance syntaxique et sémantique (3).

- Le module de reconnaissance des phrases prend en charge les chaînes lexicales. Une interprétation de la chaîne lexicale est fournie pour un "Univers" donné. Ce module dépend donc de l'application envisagée.

- Le module de génération des phrases, qui est l'inverse de l'analyseur lexical, prend en charge une phrase graphique en entrée et génère une phrase phonétique de prononciation courante en tenant compte de la liaison, de l'élision, de la dénasalisation et des pauses (4).

IMPLEMENTATION DE MICRO PROCESSEURS POUR LE TRAITEMENT DU SIGNAL VOCAL EN TEMPS REEL

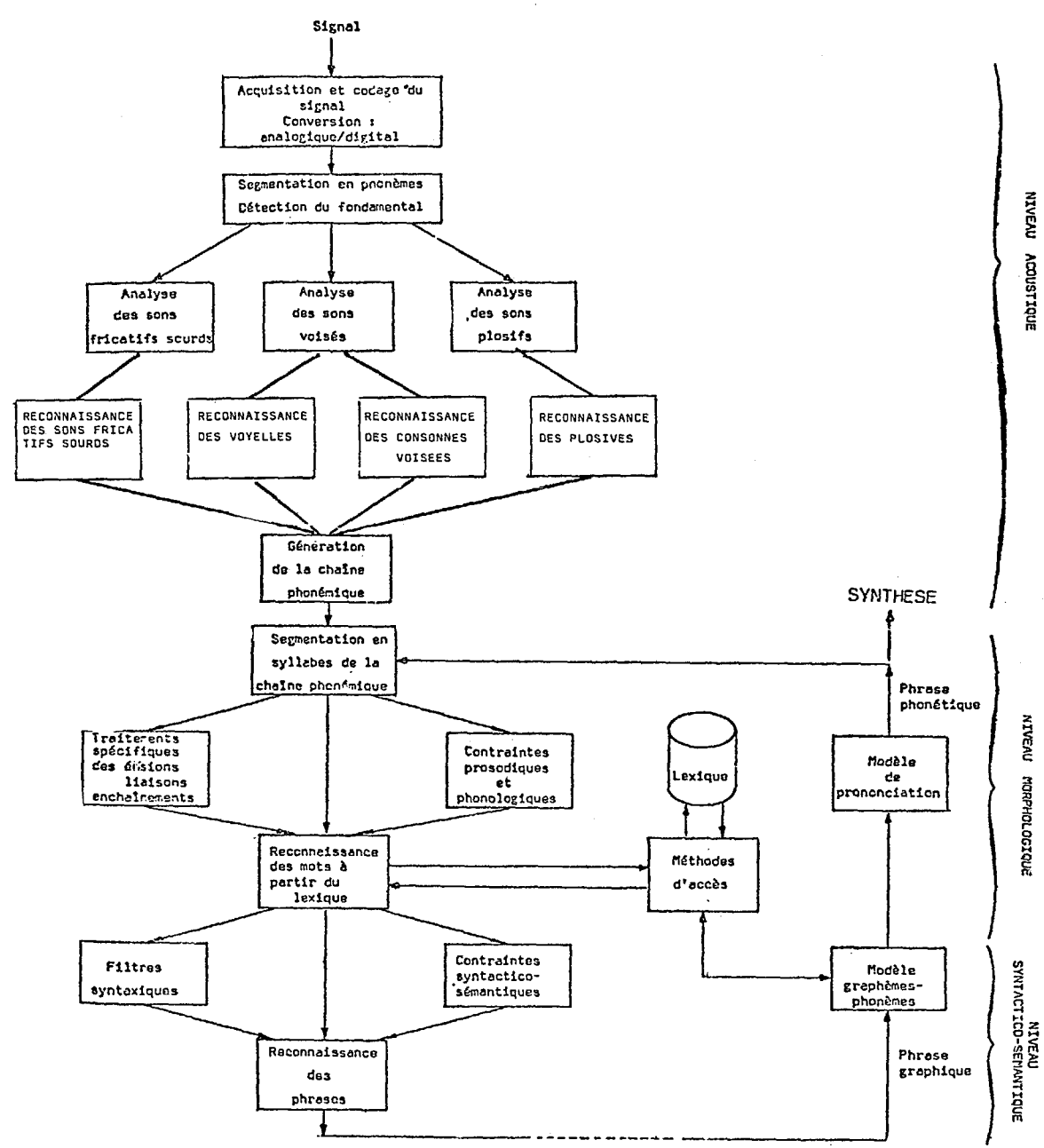


Figure 1 - PROJET A.R.I.A. : Les différents modules

Le module de synthèse, génère un signal synthétique à partir de la phrase phonétique. Ce signal peut être écouté sur un haut-parleur.

Nous ne nous intéresserons par la suite qu'aux modules du niveau acoustique.

ORGANISATION GENERALE DU SYSTEME

Une machine parallèle effectuant les traitements prévus dans les différents modules du projet A.R.I.A., peut être réalisée à l'aide d'un système à structure mixte centralisée - Figure 2 -

Les modules de traitement (acquisition, segmentation...) sont organisés autour d'un module de commande (Parole) qui les active en tenant compte de leur disponibilité et de l'ordre dans lequel les différentes tâches doivent être exécutées. La hiérarchisation de ces tâches impose une structure pipe-line. Seules des



IMPLEMENTATION DE MICRO PROCESSEURS POUR LE TRAITEMENT DU SIGNAL VOCAL
EN TEMPS REEL

tâches de même niveau peuvent être exécutées en parallèle. Plusieurs tâches sont exécutées simultanément dans différents modules. Les modules de traitement échangent des informations entre eux par l'intermédiaire d'une mémoire commune. Ils informent le module de commande lorsqu'une tâche est terminée, chacun ayant son propre contrôle.

Le fonctionnement d'une machine ainsi organisée ne va pas sans poser un certain nombre de problèmes que nous classerons en deux catégories : le contrôle de l'accès aux informations, d'une part, le réglage des conflits de dépendance d'autre part. Nous nous proposons de décrire l'acquisition des données et le principe de fonctionnement des différents modules de traitements (segmentation, analyse, reconnaissance), en nous intéressant tout particulièrement aux deux problèmes précités.

Acquisition des données

- Le module d'acquisition effectue l'échantillonnage et le codage du signal vocal à l'aide d'un convertisseur A/D ainsi que la mémorisation du signal discrétisé dans la mémoire de données.

- La structure pipe-line de notre système nous a conduit à une organisation de la mémoire en blocs, ceci afin d'éviter au maximum les conflits d'accès à

la mémoire et de faciliter les calculs d'adresses. En effet, pour que le système puisse fonctionner il est nécessaire que le temps maximum de traitement d'un bloc de données soit inférieur ou égal au temps d'acquisition de ce bloc. Il faut donc que plusieurs modules puissent avoir accès aux données simultanément. Dans notre cas trois modules peuvent avoir accès aux données. Nous avons choisi une mémoire de 64 K mots organisée en quatre blocs de 16 K mots, ce qui permet de mémoriser environ quatre à cinq secondes de signal parole.

- Le module de contrôle d'accès mémoire doit gérer la mémorisation du signal et les demandes de lecture provenant des modules de traitements. L'accès à la mémoire pose deux problèmes importants : l'aiguillage des informations lues dans la mémoire, la gestion des conflits d'accès au niveau d'un même bloc.

Dans le premier cas le problème est de diriger toute information lue en mémoire vers le module de traitement auquel elle est destinée. Dans ce but on associe à chaque adresse un indicateur désignant le destinataire, cet indicateur est envoyé au module de contrôle d'accès en même temps que l'adresse. Celui-ci le mémorise pendant le temps d'accès et le restitue avec l'information lue sur le bus de lecture.

La gestion des conflits d'accès au niveau des blocs mémoires se fait par l'intermédiaire d'une file d'attente. En effet, il faut conserver au niveau du contrôleur de mémoire les adresses des références qui, bien qu'acceptées, n'ont pu être servies, le bloc adresse étant occupé.

Modules de traitements

Les traitements que doivent effectuer les différents modules sont de deux types :

- Calculs de paramètres par des processus différents et indépendants sur des données identiques. Ces processus sont exécutés par des processeurs en parallèle et constituent le premier niveau.

- Prise en compte de ces paramètres et élaboration des résultats par un ou plusieurs processeurs en pipeline, constituant les niveaux suivants.

Le nombre de processeurs dépend de l'importance du traitement, chaque processeur possède une mémoire programme (PROM) et une mémoire de travail (RAM).

L'ensemble est géré par un processeur maître qui a pour fonctions principales, l'allocation des ressources aux différents processeurs, le contrôle et l'enchaînement des différentes tâches, le dialogue avec le module de commande (Parole).

L'allocation des ressources consiste, d'une part

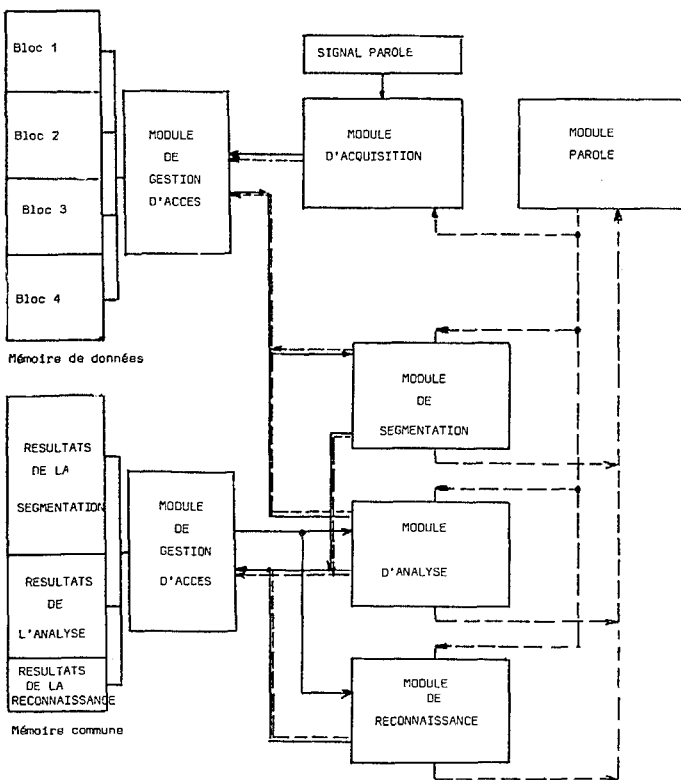


FIGURE 2 - STRUCTURE GENERALE DU SYSTEME

— Bus de données
--- Bus de commandes



IMPLEMENTATION DE MICRO PROCESSEURS POUR LE TRAITEMENT DU SIGNAL VOCAL
EN TEMPS REEL

à transférer à chaque processeur les informations qui lui sont nécessaires pour effectuer sa tâche, d'autre part à mémoriser les résultats qu'il obtient. Cette mémorisation se fait dans la mémoire du processeur maître qui sert dans ce cas de station tampon. Excepté pour le premier niveau, ces résultats serviront de données d'entrées pour les processeurs du niveau suivant. Les données nécessaires au premier niveau sont soit le signal échantillonné, soit les résultats issus du module de traitement antérieur, soit les deux à la fois. Le signal échantillonné provient de la mémoire des données, les résultats sont mémorisés dans la mémoire commune à laquelle accèdent les processeurs maîtres des différents modules de traitements. Dans le cas de processeurs en parallèle, le bus commun de sortie est alimenté par tous les processeurs dont on attend des résultats et alimente le processeur maître qui mémorise ces résultats. Le bus commun transporte, outre les résultats, un indicateur désignant sa source.

Le contrôle et l'enchaînement des tâches se font à partir d'un système d'indicateurs.

A chaque niveau p , sont associés deux indicateurs I_1^p indique que la tâche de niveau p est active, $I_1^p = 1$ ou terminée $I_1^p = 0$.

I_2^p indique que les résultats produits par le ou les processeurs de niveau p ont été pris en compte par le module maître et mémorisés $I_2^p = 0$, $I_2^p = 1$ sinon.

Le processeur maître doit : soit activer des tâches, soit mémoriser des résultats.

Pour activer une tâche de niveau p , il doit s'assurer que les trois conditions suivantes sont remplies :

- $I_1^p = 0$, processeurs de niveau p libres,
- $I_2^{p-1} = 0$, résultats de niveau $p-1$ disponibles
- $I_2^p = 0$, résultats du traitement précédent pris en compte par le processeur maître

Dès qu'une tâche de niveau p peut être activée, le processeur maître transfère les données mémorisées dans la station-tampon vers les processeurs de niveau p . Les indicateurs I_1^p et I_2^p sont alors mis à 1.

Lorsque le processeur maître est averti que la tâche de niveau p est terminée, il prend en compte les résultats et remet les indicateurs I_1^p et I_2^p à zéro.

La prise en compte des résultats issus du niveau p ne pourra être effective que si $I_2^{p+1} = 0$. C'est-à-dire, si les résultats mémorisés dans la station-tampon ont été transférés vers le processeur de niveau $p+1$.

Lorsque la tâche du dernier niveau est terminée, le processeur maître prend en compte les résultats et les mémorise dans la mémoire commune afin qu'ils puissent être exploités par le module de traitement suivant.

Le dialogue avec le module de commande est de deux sortes : celui initialisé par le processeur maître indiquant que des résultats sont disponibles dans la mémoire commune pour le module de traitement suivant, et celui initialisé par le module de commande indiquant au processeur maître que des résultats issus du module précédent sont disponibles dans la mémoire commune.

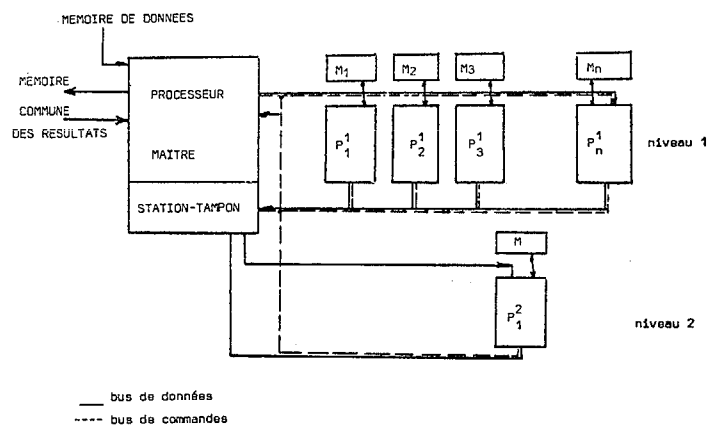


Figure 3 - SCHEMA D'UN MODULE DE TRAITEMENT

L'utilisation de station-tampon ainsi que la technique retenue pour la dialogue permet un enchaînement efficace des différents traitements en évitant au maximum les situations de conflits d'accès aux informations ou de dépendance.

Dans ces conditions l'ensemble des traitements à effectuer ne se trouve pas trop ralenti par les opérations de contrôle nécessaires et un fonctionnement en temps réel de l'ensemble peut être obtenu.



IMPLEMENTATION DE MICRO PROCESSEURS POUR LE TRAITEMENT DU SIGNAL VOCAL
EN TEMPS REEL

(1) - D. DOURS - R. FACCA

"Analyse temporelle du signal vocal comparée
à l'analyse fréquentielle du point de vue de
la reconnaissance".

Projet SESORI IV/137 n° 74-94

(2) - D. DOURS - R. FACCA - G. PERENNOU

"Analyseur d'un signal fortement structuré :
le signal vocal".

5ème Colloque National sur le traitement du
signal - GRETSI - Nice du 16 au 21 Juin 1975.

(3) - B. CAUSSE - D. DOURS - R. FACCA - G. PERENNOU

"Evaluation d'une méthode ascendante d'analyse
lexicale dans le discours continu"

7ème J.E.P. - NANCY - 19 au 21 mai 1976.

(4) - B. CAUSSE - G. GOUARDERES - G. TEP

"L'utilisation du lexique pour la génération
de phrases parlées".

Rapport C.E.R.F.I.A. - Octobre 1976 -