

## UNE CARTE DE TRAITEMENT DU SIGNAL

Patrick CHENAIS

BENDIX ELECTRONICS (\*) - Avenue du Mirail - BP 11-49 31036 TOULOUSE CEDEX.

CERFIA - Université Paul Sabatier - 118 route de Narbonne 31062 TOULOUSE CEDEX.

### RESUME

### SUMMARY

Nous décrivons la réalisation d'une carte de traitement numérique du signal adaptée à la reconnaissance automatique de la parole et la commande d'actionneurs.

Pour parvenir à un coût minimal, l'architecture de la carte est étudiée sans dissocier les deux objectifs de reconnaissance et de commande. Après avoir rappelé les structures possibles du système, nous dégagons l'architecture optimale à notre application.

Le système est basé autour de deux processeurs : un microcontrôleur (8031 INTEL) et un processeur de signal (TMS 32010 TEXAS INSTRUMENTS). Nous exposons le principe du partage de la mémoire de données. Le microcontrôleur a un accès aléatoire à la mémoire et le processeur de signal a un accès séquentiel par ses ports d'entrée/sortie. Le temps d'accès des mémoires utilisées est déterminé par la fréquence de lecture/écriture du processeur de signal et non pas par la durée du cycle. L'utilisation de mémoires peu rapides, de faible consommation et de faible prix est donc rendu possible.

Le système est peu complexe et d'un coût de fabrication réduit.

*Le travail présenté ici a fait l'objet d'un contrat de recherche entre la société BENDIX ELECTRONICS et le laboratoire CERFIA de l'université Paul Sabatier de TOULOUSE.*

We describe the realization of a digital signal processing board destined to speech recognition and environment control.

In order to obtain a minimal cost, the architecture of the board is planned without dissociation of the two aims : recognition and command.

After a brief overview over the possible structures of the system, we explain the most convenient architecture for our application.

The system is based on two processors, a microcontroller (8031 INTEL) and a digital signal processor (TMS 32010 TEXAS INSTRUMENTS). We expose the principle of shared memory. The microcontroller has random acces to the memory whereas the signal processor has sequential acces via his input/output ports. The time of memory acces is determined by the read/write frequency of the signal processor and not by the memory acces time. Thus it is possible to use slow memories. They are cheap and have little consumption.

The system is not complex and has a small fabrication cost.

(\*) Le département *petite série* où ont été réalisées les études s'est constitué en société indépendante : Société ACTIA, Avenue du Général Eisenhower, TOULOUSE CEDEX 31047.



## Introduction

Dans une application de contrôle d'environnement, la société BENDIX ELECTRONICS (ex RENIX ELECTRONIQUE) a développé un produit construit autour d'une carte de reconnaissance vocale O.E.M. (Original Equipment Manufacturing). La figure ci-dessous fait ressortir les différents modules de ce système.

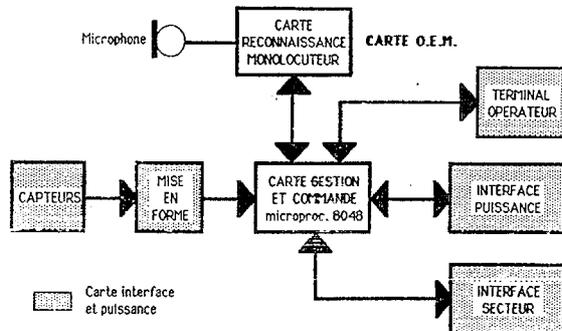


Figure 1

Nous distinguons la carte O.E.M. de reconnaissance vocale (VECSYS RMI 50), la carte de gestion et de commande des actionneurs et la carte interface et puissance. La difficulté de commercialisation d'un tel produit a deux causes techniques principales : le coût trop élevé de fabrication et les possibilités limitées de la reconnaissance.

Notre objectif a été d'obtenir une forte réduction du coût matériel par l'intégration de façon économique de la fonction de commande des actionneurs et de la fonction de reconnaissance sur la même carte. Cette intégration va permettre une réduction de la redondance du matériel.

## Fonction

La fonction unique du système est la transduction d'un message parole en informations transmises à l'environnement. Cette fonction est assujettie à l'exécution de trois tâches principales :

- le traitement du signal
- l'exécution des algorithmes de reconnaissance
- la commande des actionneurs et le dialogue avec l'environnement.

### Traitement de la parole

Les deux premières tâches sont spécifiques à notre application de reconnaissance vocale. Nous avons implanté le projet MILOS de reconnaissance de mots isolés développé au Laboratoire CERFIA.

Le traitement du signal consiste principalement à calculer cinq indices acoustiques [1]. Les indices sont significatifs du rapport d'énergie dans des bandes fréquentielles particulières et liées à la répartition statistique des résonances du conduit vocal. Ils sont calculés toutes les 10 ms. Leur détermination nécessite un filtrage par onze filtres résonateurs du second ordre

à la fréquence de 10 kHz [2]. La méthode de reconnaissance utilisée est basée sur la programmation dynamique [3].

### Commande des actionneurs et dialogue avec l'environnement

L'application étant le contrôle d'environnement, nous devons prévoir la commande d'appareils en tout ou rien (sorties logiques, relais etc.), le dialogue avec un microordinateur (liaison série) et la commande à distance (télécommande infrarouge, télécommande secteur).

## Structure

Le choix de la structure consiste à déterminer les principaux composants autour desquels sera construite la carte ; c'est attribuer les tâches à des modules matériels.

Pour le traitement du signal nous avons deux solutions envisageables : le filtrage analogique et le traitement numérique. Nous avons rejeté la solution du filtrage analogique pour son aspect figé : nécessité de modifier le matériel pour toute modification des paramètres de filtrage, et ce sans un gain certain au niveau du prix comparativement aux processeurs de signaux. Pour le traitement numérique du signal nous pouvons envisager l'utilisation d'un processeur spécialisé au traitement de la parole (processeur de LPC, vocodeur à canaux, etc.) ; ces circuits sont chers car peu distribués et ne sont pas compatibles avec le calcul des indices acoustiques. Dernière alternative économique : les circuits calculateurs programmables (microprocesseurs, microcontrôleurs et processeurs de signaux).

L'exécution de l'algorithme de reconnaissance peut être réalisée à l'aide d'un circuit de calcul spécialisé ou d'un circuit calculateur programmable. La première solution ne permet pas d'évolution à la stratégie de reconnaissance ; par ailleurs ces circuits sont chers.

Le choix d'un type de circuit est beaucoup plus limité pour la dernière tâche à remplir. Cette tâche de commande, contrôle et de dialogue réclame presque impérativement l'utilisation d'un microprocesseur ou d'un microcontrôleur, tant ces circuits sont adaptés à ces applications et sont bon marchés.

Ainsi le système comportera nécessairement un microcontrôleur ou un microprocesseur. Partant de cette constatation le choix des structures se réduit aux possibilités indiquées sur la figure suivante. (v. fig. 4)

Nous nous intéressons aux structures 1, 2 et 3 ; les possibilités 4 et 5 ont précédemment été éliminées, nous les citons pour mémoire.

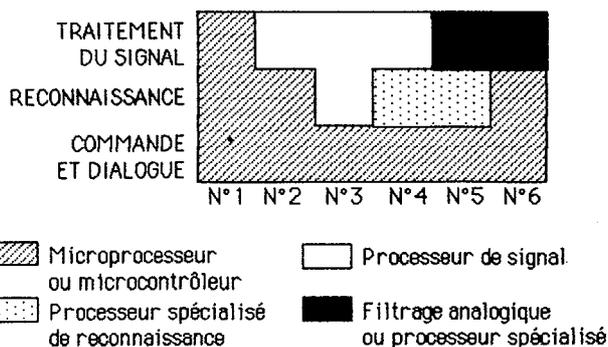


Figure 2

**Evaluation des structures**

L'évaluation des structures se fait sur la base du critère d'optimalité suivant : nous disons que le système est optimal si à tout moment, chaque module exécute une tâche. Ce critère revient à minimiser les temps d'attente. Avec ce critère chaque structure a un domaine d'application particulier.

- Structure N°1 : Un microprocesseur ou un microcontrôleur assure toutes les tâches. Cette structure très simple est optimale. Cependant les performances actuelles des circuits ne permettent pas d'effectuer un traitement du signal très élaboré. La phase de reconnaissance débute après l'acquisition complète du mot. Le vocabulaire sera très réduit pour obtenir un temps de réponse court. Cette structure convient particulièrement aux applications bas de gamme et à faible vocabulaire (marché du jouet).

- Structures N°2 (et N°4) : Un processeur de signal effectue le traitement du signal. C'est un composant relativement cher qu'il ne faut pas sous-employer. Pour être optimale la structure N°2 devra avoir un microprocesseur assez puissant pour que la reconnaissance soit faite en temps réel (ce problème de puissance de calcul est résolu dans la structure N°4). Les systèmes de reconnaissance de la parole continue peuvent avoir cette structure. Ces systèmes (s'ils sont optimaux) sont puissants mais aussi relativement chers.

- Structure N°3 : Le processeur de signal assure seul les calculs importants. Un microprocesseur ou un microcontrôleur peu puissant suffit pour le contrôle. Cette structure est bien adaptée à la reconnaissance de mots isolés sur un vocabulaire moyen (100 mots environ). Dans ce cas l'exécution différée des tâches de traitement de signal et de reconnaissance permet d'avoir sur chaque phase la totale puissance de calcul du processeur de signal. L'inconvénient est cependant l'impossibilité d'atteindre le temps réel. Pour notre application un délai de réponse bref (moins d'une 1/2 s) est tolérable ; cette structure est donc optimale.

**Architecture**

La structure ne fait pas apparaître le flux des informations, entre les modules et avec la mémoire de donnée, ni l'aspect temporel du traitement. L'architecture doit prendre en compte ces caractéristiques.

Pour déterminer l'architecture correspondant à une structure donnée nous remarquons :

- qu'un échange d'informations entre deux tâches se traduit sur le plan matériel par un dialogue entre deux modules,
- qu'il est permis de représenter une tâche par le module qui la réalise

Les échantillons d'informations ou de données entre les tâches sont représentées par des barres reliant les différents modules. Le sens des informations est dans notre représentation de haut vers le bas (les modules pouvant échanger des données sont situés sur le même niveau).

L'architecture correspondante à la structure N°3 est déterminée en remarquant que parmi les trois tâches, seuls le traitement du signal et la reconnaissance nécessite l'accès à la mémoire de données. Ainsi la figure suivante représente la complexité minimale de cette structure.

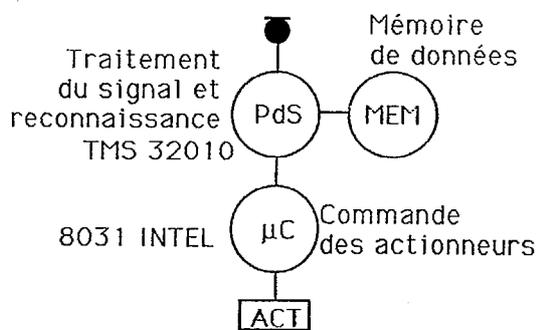


Figure 3

Nous avons choisi comme processeur de signal un TMS 32010 TEXAS INSTRUMENTS qui offre un rapport performances/prix intéressant et dont la vitesse de calcul est suffisante pour le traitement du signal que nous réalisons. Nous utilisons un microcontrôleur 8031 INTEL, peu cher et disposant suffisamment d'interfaces (ports série et parallèle) pour notre application de base. Il est à noter que ces deux circuits disposent d'une mémoire interne.

Les microprocesseurs et les microcontrôleurs sont orientés pour gérer une mémoire de données, il peut être intéressant de modifier l'architecture de la figure 3 pour profiter de cette facilité. L'architecture suivante a été finalement adoptée.

Le microcontrôleur gère entièrement le système. Le processeur de signal est considéré comme une unité de calcul auxiliaire du premier. Nous avons opté pour un

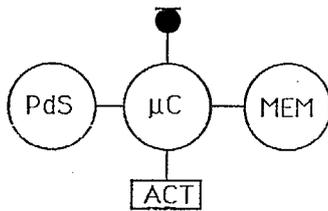


Figure 4

partage complet de la mémoire entre les deux processeurs ; le fait de partager entièrement la mémoire n'augmente que peu la complexité du système par rapport à la mise en place de mémoires tampons locales.

### Principe du partage mémoire

L'accès à la mémoire se fait de façon aléatoire pour le microcontrôleur et séquentiellement pour le processeur de signal. L'intérêt du procédé utilisé est de permettre l'utilisation d'une mémoire "lente" par un processeur "rapide". Le temps d'accès de la mémoire ne dépend que de la fréquence des lectures/écritures et non pas de la durée du cycle d'écriture/lecture du processeur utilisé.

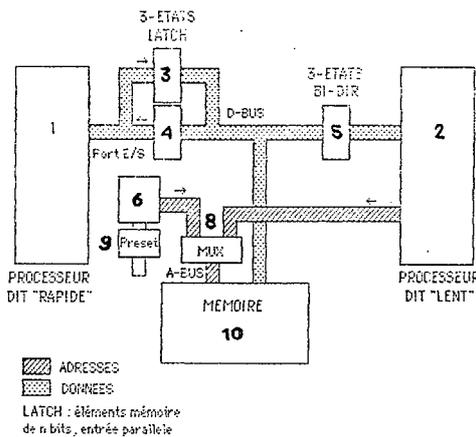


Figure 5

### Principe de fonctionnement

Les bus d'adresses et de données sont multiplexés. Lors de l'utilisation de la mémoire par le processeur 2, ses bus d'adresse et de données sont en communication avec la mémoire 10 respectivement par 8 (MULTIPLEXEUR) et 5 (BUFFER BIDIRECTIONNEL). L'environnement normal du processeur est ainsi recréé. Lorsque le processeur 1 accède à la mémoire, le bus d'adresse A-BUS est connecté à un compteur 6 ; les latches 3 et 4 permettent l'accès à la mémoire.

Le principe de fonctionnement est le suivant :

- En écriture, la donnée issue de 1 est mémorisée dans le latch 3 puis mémorisée dans la mémoire 10 à l'adresse pointée par le compteur 6. Un ensemble simple de circuits crée ensuite le timing nécessaire à l'écriture en temps masqué. Le compteur 6 est alors post-incrémenté (ou décrémenté).

- En lecture, le processeur 1 lit la donnée préalablement mémorisée dans le latch 4, le compteur 6 est post-incrémenté (ou décrémenté). Comme précédemment, un jeu de circuits provoque l'écriture dans 4 de la prochaine valeur qui sera lue, son adresse est le contenu du compteur 6.

- Le choix de l'adresse de lecture ou d'écriture est effectué en chargeant le compteur 6, ceci peut être fait par l'un ou l'autre des processeurs (le microcontrôleur dans notre réalisation).

- Si plusieurs pages mémoire doivent être utilisées en alternance rapide par le processeur 1, on pourra multiplier les compteurs 6 et les latches 3 et/ou 4 en conséquence. Nous avons choisi deux compteurs 6, deux latches 4 et un latch 3.

### Performances

Nous obtenons avec ce procédé un débit mémoire -TMS 32010 de 2500 kmots/sec. (1 lecture/écriture toutes les 400 ns, soit son débit maximum) avec une mémoire C-MOS statique de temps d'accès de 150 ns (au lieu des 85 ns normalement nécessaires).

### Conclusion

Le premier intérêt de l'analyse de la fonction du système est d'avoir permis la définition d'une architecture spécialisée optimale à notre application.

Avec le principe du partage mémoire décrit, le coût matière de la carte unique remplaçant la carte gestion et la carte de reconnaissance (voir fig. 1) est de moins de 1000F. La carte dispose des ressources d'un TMS 32010, d'un microcontrôleur 8031 INTEL, d'un plan mémoire commun de 64 koctets et d'un convertisseur A/N et N/A (12bits) ; avec ces ressources la reconnaissance de 100 mots codés sur 100 octets chacun réclame environ 350 ms [2]. La vitesse de traitement du processeur de signal permet le calcul en temps réel des 5 cinq indices à la fréquence de 10 kHz (14 kHz maximum).

Le second intérêt de cette étude est de nous avoir permis d'envisager dès la conception du premier produit l'évolution de son architecture.

### Références

- [1] CAELEN J. et CAELEN G. : Indices et propriétés dans le projet ARIAL, Actes du séminaire GALF-CNRS, Les processus d'encodage et de décodage phonétique ; TOULOUSE Septembre 1981 ; pp 129-143.
- [2] CHENAIS P. : Une carte de traitement et de reconnaissance de la parole - Etude de cibles acoustiques. Th. de Docteur Ingénieur. Université Paul Sabatier TOULOUSE , 1987 ; 280 pages.
- [3] SAKOE H. et CHIBA S. : Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. on ASSP, Vol ASSP-26, N° 1, Feb. 1978, pp. 43-49.