

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي و البحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

ال مديرية العامة للبحث العلمي و التطوير التكنولوجي

Direction Générale de la Recherche Scientifique et du Développement Technologique



Projet national de recherche : rapport final

SYSTÈME D'AIDE AUX NON-VOYANTS PAR COMMANDE VOCALE

Chef de projet : MESSADEG Djemil

Affiliation: Laboratoire des Signaux et Images (LASA)

Organisme de domiciliation : Université BADJI Mokhtar de Annaba

Organisme pilote : CERIST

Programme national de Recherche : 2011-2013





BILAN D'EVALUATION DU PROJET PNR Technologie & Industrie

1. DOMICILIATION DU PROJET

Organisme de domiciliation

**Laboratoire des Signaux et Images (LASA).
Université Badji Mokhtar de Annaba.**

Organisme pilote

CERIST

2. IDENTIFICATION DU PROJET

Intitulé du projet:	Système d'aide aux non-voyants par commande vocale
Intitulé du domaine	Domaines émergents
Intitulé de l'axe	Traitement de la parole
Intitulé du thème	Communication homme/machine

3. IDENTIFICATION DE L'EQUIPE DE RECHERCHE

Chef de projet		
Nom et prénom	Grade	Etablissement de rattachement
Dr. Messadeg Djemil	M.C. (A)	Université Badji Mokhtar de Annaba.

Equipe de recherche			
Nom et prénom	Grade	Etablissement de rattachement	Observation
Pr. Bousbia Salah Mounir	Prof.	Université Badji Mokhtar d'Annaba.	
Dr. Hamdi Rachid	M.C. (A)	Université Badji Mokhtar d'Annaba.	
Dr. Djemili Rafik	M.C. (A)	Université du 20 Août 1955 Skikda.	
Dr. Bourouba Hocine	M.C. (B)	Université du 08 Mai 1945 Guelma.	
Dr. Amara Korba Med Cherif	M.C. (B)	Université de Souk Ahras	

Republique Algerienne Démocratique & Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Badji Mokhtar de Annaba.

Laboratoire des Signaux et Images (LASA)

RAPPORT FINAL PNR

Intitulé du domaine : Domaines émergents

Intitulé de l'axe : Traitement de la parole

Intitulé du thème : Communication homme/machine

Intitulé du projet:

**Système d'aide aux non-voyants par commande
vocale**

Table de matières

I.	Introduction	4
I.1	Le logiciel d'aide au non-voyant	5
a)	La première partie	5
b)	La deuxième partie	5
I.2	La carte électronique de commande	5
II	Principe de fonctionnement du système	5
III.	Description du système de reconnaissance vocale	8
III.1	Extraction des paramètres	8
III.1.1	Chaîne de prés-traitement	8
III.1.2	Les coefficients MFCC	9
III.1.3	Les coefficients différentiels	12
III.2	Moteur de reconnaissance	13
III.2.1	Description de la Méthode DTW	13
III.2.2	Distance globale	15
IV	Logiciel de reconnaissance vocale	16
V	Présentation de la carte de commande	20
V.1	Description détaillée du microcontrôleur pic16f877	21
V.2	Caractéristiques du pic16f877	21
V.3	Brochage du pic16f877	22
V.4	Architecture interne du pic 16f877	23
V.5	Organisation de la mémoire du pic 16f877	24
V.6	Jeu d'instructions du pic16f877	25
V.7	Schéma électronique de la carte de commande	26
V.8	Circuit imprimé de la carte de commande	27
V.9	Schéma de la carte d'alimentation	29
	Conclusion	30
	Bibliographie	31
	Information financière	37

Liste des tables

Table 1. Caractéristiques du pic16f877	21
Table 2. jeu d'instructions du pic16f877	25
Table 3. Nomenclature des dépenses	37

Liste des figures

Figure 1.1	Blocs et signaux du système d'aide au non-voyant	6
Figure 1.2	Schéma synoptique du système d'aide au non-voyant	7
Figure 1.3	Schéma bloc d'un système de reconnaissance de la parole	8
Figure 1.5	Chaîne de prétraitement du signal parole	9
Figure 1.6	Les filtres triangulaires passe-bande en Mel-Fréq ($B(f)$) et en fréquence (f)	10
Figure 1.7	Schéma en blocs de l'analyse acoustique permettant le calcul des vecteurs MFCC	11
Figure 1.8	Visualisation du cheminement de l'alignement temporel pour des formes de la base de référence	14
Figure 1.9	Les transitions autorisées entre les points du graphe	15
Figure 1.10	Schéma typique d'une fonction de recalage en alignement temporel	16
Figure 1.11	Représentation des différents blocs du logiciel d'aide au non voyant	18
Figure 1.12	Interface graphique du logiciel d'aide au non-voyant	18
Figure 1.13	Commande par interrupteurs	19
Figure 1.14	Graphique du système d'apprentissage du logiciel d'aide au non-voyant	19
Figure 1.15	Schéma bloc de la carte de commande	20
Figure 1.16	Boitier Dip 40 broches pour pic16f877	22
figure 1.17	Architecture interne du pic 16f877	23
figure 1.18	Organisation de la mémoire RAM du pic16f877	24
Figure 1.19	Schéma de la carte PC de commande	25
Figure 1.20	Circuit imprimé de la carte pc de commande	27
Figure 1.21	Dispositif de commande	28
Figure 1.22	Dispositif de commande sans couvercle	28
Figure 1.23	Schéma électrique de la carte d'alimentation	29
Figure 1.24	Carte d'alimentation	29

II. Introduction

Des millions de personnes dans le monde sont atteintes de déficience visuelle ou de cécité, Ces individus sont soumis à de nombreuses difficultés de déplacement. Un réel besoin est présent afin d'améliorer leur confort de vie et minimiser au maximum leur handicap, de nombreuses solutions répondant à cette problématique sont actuellement sur le marché.

Notre projet vise à apporter des solutions techniques pour répondre aux besoins de confort des personnes non voyantes dans leur domicile.

Ce projet peut aussi s'inscrire dans le cadre de la domotique, nous tentons par ce travail de rendre les tâches quotidiennes effectuées par le non-voyant, exécutées par de simples commandes vocales tel que '*ouvrir la porte*' ou '*monter le volet*', ...etc. Pour ce faire, nous avons élaboré un système d'aide au non voyant intelligent, essentiellement basé sur la commande vocale. Une commande infrarouge est prévue aussi.

Notre système est capable de s'adapter aux besoins de chaque personne par :

- 1- Le grand choix de commandes vocales laissées à l'utilisateur pour piloter ses équipements, afin qu'il soit très à l'aise lors des échanges avec le système.
- 2- L'interaction du système avec l'utilisateur qui est informé en permanence par des messages vocaux annoncés par le système d'aide.
- 3- Le nombre d'entrée/sortie prévus pour Le raccordement (équipements et capteurs) est important ce qui offre une multitude de solutions pour s'adapter à l'environnement de l'utilisateur.
- 4- la carte électronique développée est programmable assurant des échanges en temps réel avec le logiciel d'aide installé sur le PC, conçue autour du microcontrôleur PIC, ce qui permet des acquisitions de plusieurs grandeurs physiques : température, humidité, position, vitesse ...etc., et avec une bonne précision.

Notre projet est composé de deux parties :

I.1 Le logiciel d'aide au non-voyant : composé à son tour de deux parties :

a) La première partie représente la phase d'apprentissage, elle est cruciale est utilisée dans un premier temps par le système, ce dernier devra apprendre les commandes désirées par l'utilisateur, par la répétitions d'un ensemble de commande par l'utilisateur pour constituer la base de données d'apprentissage.

L'apprentissage est effectué par l'intermédiaire d'une interface interactive qui répond aux exigences de l'utilisateur, elle signal le début et la fin de chaque opération par des messages vocaux et des bips sonores.

b) La deuxième partie représente la phase d'utilisation, dans cette partie le système exécute les commandes prononcées par le non voyant, une confirmation au cours et à la fin de chaque opération.

Cette partie utilise la base de données d'apprentissage et un algorithme de reconnaissance vocal bien adapté à cette application, les prochaines sections abordent une description détaillée.

I.2 La carte électronique de commande :

La carte est conçue au tour du microcontrôleur PIC16F877A développé par Microchip de la famille 'mid range', ce dernier communique avec le PC via une communication série asynchrone en mode 'full duplexe', il contrôle un clavier composé de 7 boutons et décode jusqu'à 32 commandes infrarouge, il peut afficher la température interne, la température externe en temps réel, l'heure et la date et l'état de chaque périphérique.

II Principe de fonctionnement du système

Le système d'aide au non-voyant développé au niveau du laboratoire d'automatique et des signaux d'Annaba (LASA) est conçu pour assurer un maximum de confort et soulager l'utilisateur des tâches quotidiennes répétitives, le principe de fonctionnement est le suivant :

Le non voyant doit porter un casque audio sans fil avec un microphone d'une portée de quelques centaines de mètres. Un système audio sans fil est connecté à un ordinateur portable, reçoit le signal sonore sensé contenir les commandes vocales. (voir figure 1.1)

Le logiciel de reconnaissance installé sur l'ordinateur reconnu chaque commande et transmis l'ordre à la carte électronique, par le moyen de communication série. Après réception de l'ordre du PC la carte active ou désactive l'actionnaire approprié. (voir figure 1.2)

Nous avons utilisé 5 commandes lors du développement du prototype :

- 1- Ouvrir la porte/ fermer la porte
- 2- Monter le volet / descendre le volet
- 3- Allumer la lumière / éteindre la lumière
- 4- Allumer la télévision / éteindre la télévision
- 5- Demander la température

La carte électronique de commande reçoit aussi les ordres (commandes) par l'intermédiaire d'une commande infrarouge (CRI), chaque bouton de la CRI représente une action (activer/désactiver), les confirmations des commandes sont toujours signalées par des messages vocaux et des bips sonores.

On peut utiliser aussi les boutons sur la carte électronique de commande pour commander les actionneurs des équipements sans se déplacer, un bouton est prévu pour chaque commande (bouton/commande).

Le schéma ci-dessous nous donne une idée globale sur le système d'aide au non-voyant, dans notre conception nous avons laissé un grand choix à l'utilisateur pour commander les équipements, Il peut utiliser la commande infrarouge lorsqu'il est près de quelques mètres du dispositif de commande, le système assure toujours l'exécution de la commande par des messages vocaux.

Si le non voyant est tout près du dispositif ou un problème survient à la commande infrarouge ou le casque audio il peut commander facilement les périphériques par de simples appuis sur les boutons.

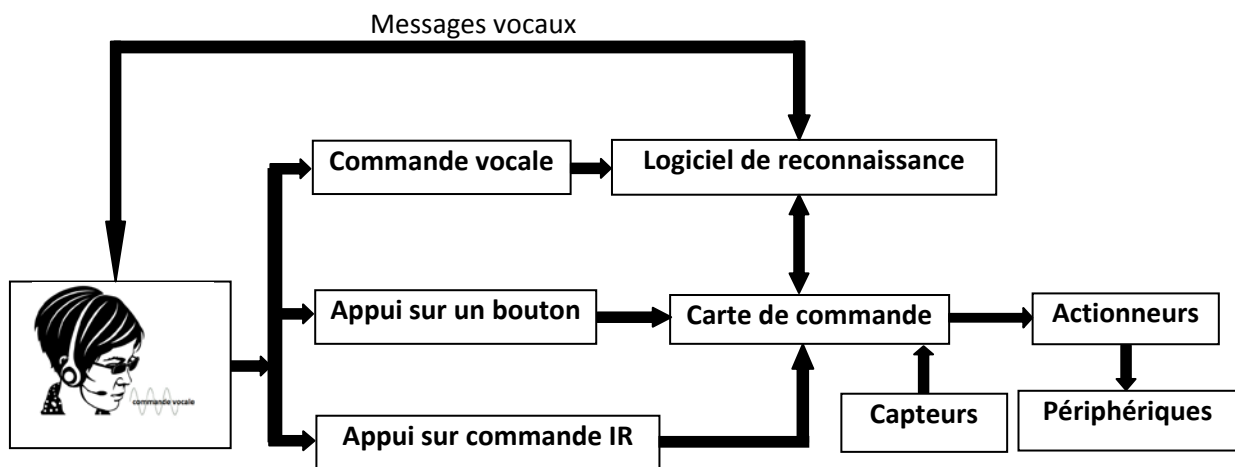


Figure 1.1 blocs et signaux du système d'aide au non-voyant

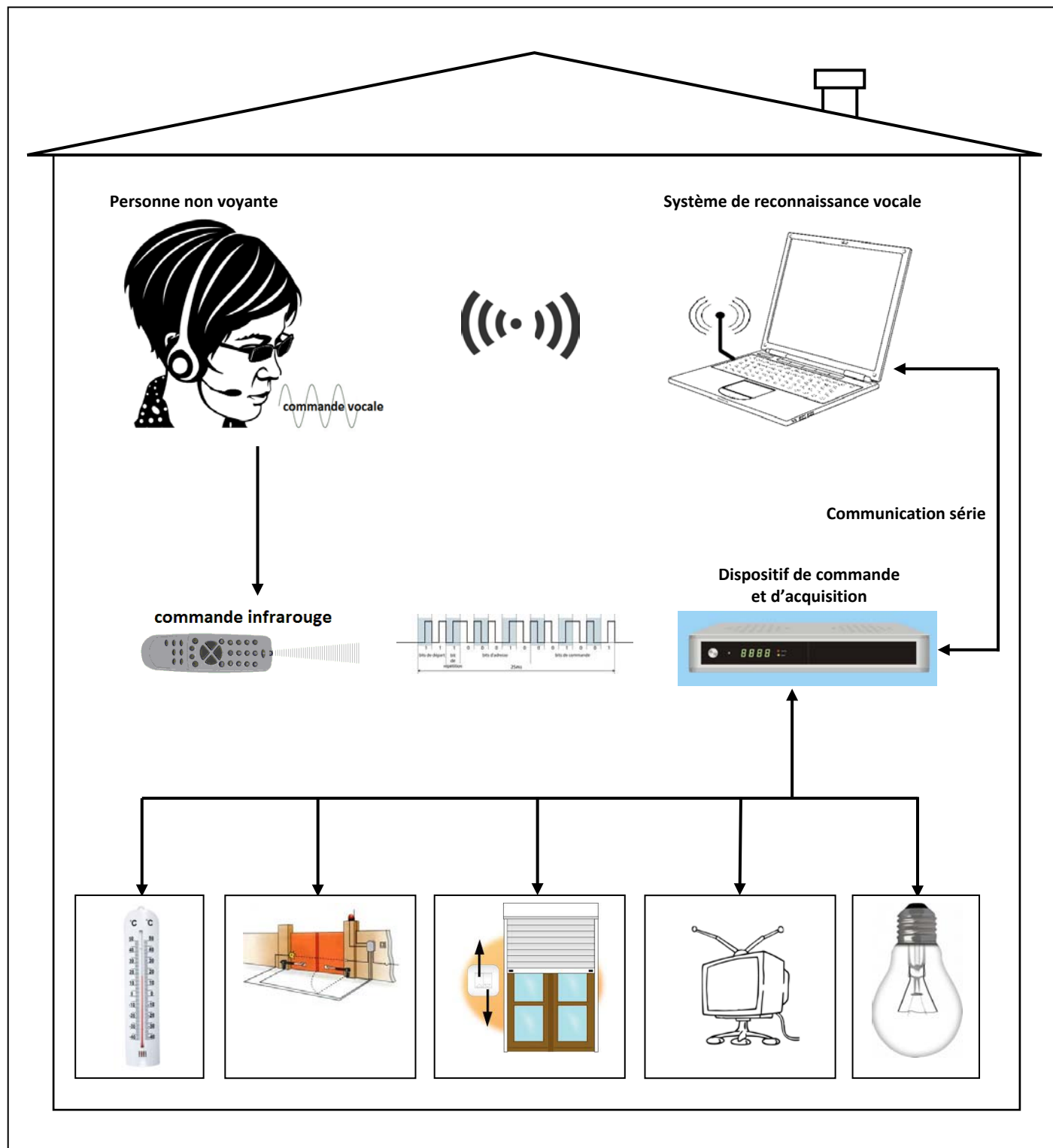


Figure 1.2 Schéma synoptique du système d'aide au non-voyant

III. Description du système de reconnaissance vocale

Tout Les système de reconnaissance vocale sont divisés en deux parties, une première partie qui représente la phase d'extraction des paramètres, et une deuxième partie qui est le moteur de reconnaissance. Les performances des systèmes de reconnaissance vocale dépendent de façon considérable des paramètres acoustiques utilisés.

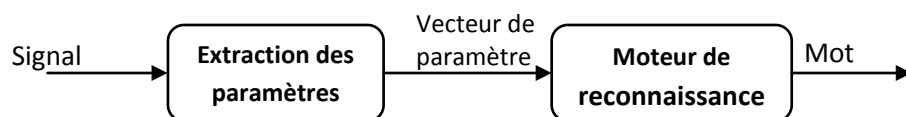


Figure 1.3 Schéma bloc d'un système de reconnaissance de la parole

III.1 Extraction des paramètres

Pour résoudre les problèmes liés à la complexité de la parole, il est possible de calculer des coefficients représentatifs du signal traité. Ces coefficients sont calculés à l'intervalle temporel régulier. En simplifiant les choses, le signal de parole est transformé en une série de vecteurs de coefficients.

Ces coefficients doivent représenter au mieux le signal qu'ils sont censés modéliser, et extraire le maximum d'informations utiles pour la reconnaissance.

Un système de paramétrisation du signal, se décompose en deux blocs (figure 1.4), le premier de mise en forme (figure 2.3) et l'autre de calcul de coefficients (figure 2.5).

Le signal analogique est fourni en entrée et une suite discrète de vecteurs, appelée trame acoustique est obtenue en sortie.

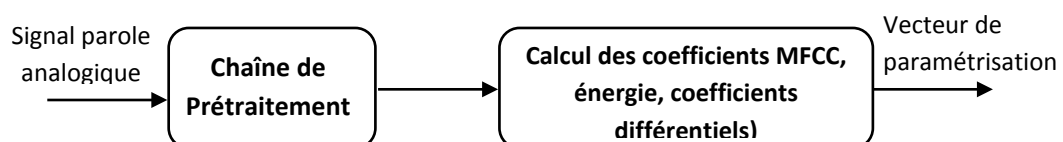


Figure 1.4 Phase de paramétrisation acoustique

III.1.1 Chaîne de pré-traitement

Il est nécessaire de mettre en forme le signal de parole. Pour cela, quelques opérations sont effectuées avant tout traitement. La (figure 2.6) illustre l'ensemble de ces

opérations. Le signal est tout d'abord filtré puis échantillonné à une fréquence donnée. Une pré-accentuation est effectuée afin de relever les hautes fréquences. Qui sont moins énergétiques que les basses fréquences; la pré-accentuation s'_n de l'échantillon s_n à l'instant n est calculée pour une valeur α comprise entre 0,9 et 1 comme :

$$S'_n = S_n - \alpha S_{n-1} \quad (1)$$

Puis le signal est segmenté en trames. Chaque trame est constituée d'un nombre N fixe d'échantillons de parole. En général, N est fixé de telle manière que chaque trame corresponde à environ 25 ms de parole (durée pendant laquelle la parole peut être considérée comme stationnaire). Enfin une multiplication par une fenêtre de pondération W_n est effectuée, afin de réduire les effets de bords. Le choix se porte sur la fenêtre de Hamming:

$$Ham \min g(n) = 0,54 - 0,64 \cos\left(2\pi \frac{n}{N-1}\right), \quad 0 \leq n \leq N-1 \quad (2)$$

avec

$$S''_n = W_n S'_n \quad (3)$$

Après cette mise en forme du signal (commune à la plupart des méthodes d'analyse de la parole), une transformée de Fourier discrète DFT en particulier FFT (Transformée de Fourier Rapide) est appliquée pour passer dans le domaine fréquentiel.

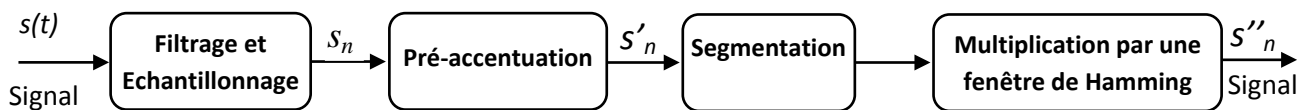


Figure 1.5 Chaîne de prétraitement du signal parole

III.1.2 Les coefficients MFCC

L'analyse acoustique MFCC est l'une des techniques les plus utilisées pour la paramétrisation du signal en segmentation markovienne de parole.

Cette technique est basée sur deux idées clés [65] [47] [39]. La première consiste à exploiter les propriétés du système auditif humain par la transformation de l'échelle linéaire des fréquences en échelle de Mel. Et la deuxième consiste à effectuer une transformation cepstrale qui permet la décorrélation des composantes spectrales du signal de parole.

Pour transformer une fréquence linéaire en une fréquence Mel, on utilise la formule de transformation suivante:

$$B(f) = 2595 \log_{10} \left(1 + \frac{f}{700} \right) \quad (4)$$

où f est la fréquence en Hz, $B(f)$ est la fréquence mel-échelle de f . Les bandes-passantes sont de même taille dans l'échelle Mel.

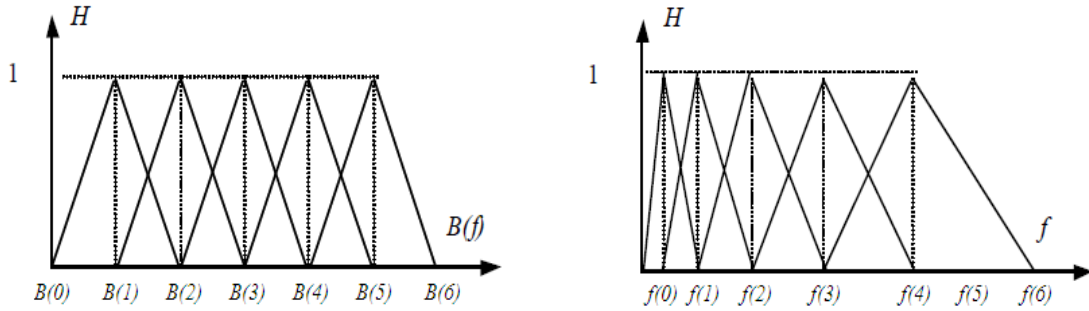


Figure 1.6 Les filtres triangulaires passe-bande en Mel-Fréq ($B(f)$) et en fréquence (f)

On peut calculer les points frontières $B(j)$ des filtres en mel-fréquence ainsi :

$$B(j) = B(f_l) + j \frac{B(f_h) - B(f_l)}{J + 1} \quad 0 \leq j \leq N + 1 \quad (5)$$

N est le nombre de filtres ($N = 22$).

On doit calculer les points $f(j)$ correspondants dans le domaine de fréquence réelle :

$$f(j) = \frac{N}{F_s} B^{-1} B(j) \quad (6)$$

Puis on détermine tous les coefficients de chaque filtre :

$$H_j(k) = \begin{cases} 0 & k \leq f(j-1) \\ \frac{k - f(j-1)}{f(j) - f(j-1)} & f(j-1) \leq k \leq f(j) \\ \frac{f(j+1) - k}{f(j+1) - f(j)} & f(j) \leq k \leq f(j+1) \\ 0 & k \geq f(j+1) \end{cases} \quad (7)$$

L'analyse MFCC comporte plusieurs étapes représentées dans la (figure 1.7). Le prétraitement consiste à effectuer sur le signal de parole, échantillonné à 16000 Hz et quantifié sur 16 bits, les opérations suivantes :

- Toutes les 10ms (160 échantillons), une trame acoustique de 25ms (400 échantillons) est

extraite du signal.

- La composante continue des échantillons constituant cette trame est enlevée.
- Afin de compenser l'atténuation naturelle du spectre du signal de parole, la séquence des échantillons constituant la trame subit une pré-accentuation avec le filtre du premier ordre

$$H(Z) = 1 - 0,97Z^{-1} \quad (8)$$

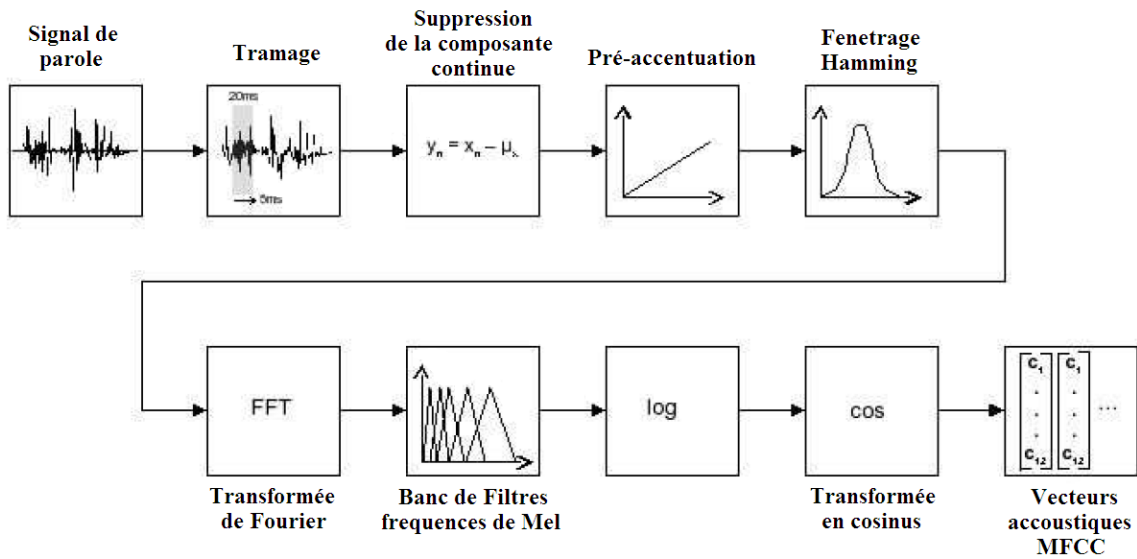


Figure 1.7 Schéma en blocs de l'analyse acoustique permettant le calcul des vecteurs MFCC.

L'analyse MFCC proprement dite consiste à effectuer sur chacune des trames résultantes du prétraitement les opérations suivantes :

- La transformation de Fourier permet de calculer le spectre d'amplitude de la trame.
- Pour chacun des 22 filtres triangulaires répartis sur l'échelle des fréquences de Mel, l'énergie du spectre d'amplitude en sortie de ce filtre est calculée. Cette opération donne un vecteur de 22 valeurs énergétiques E_j .

$$E_j = \sum_{k=0}^{N-1} |S(k)|^2 H_j(k) \quad (9)$$

- Les logarithmes de ces 22 valeurs sont alors transformés en 12 coefficients MFCC par l'inverse de la transformée en cosinus discrète :

$$c_i = \sqrt{\frac{2}{N}} \sum_{j=1}^N \log_{10}(E_j) \cos\left(\frac{\pi i}{N}(j+0,5)\right) \quad (10)$$

où c_i est le $i^{ème}$ coefficient mel-cepstral, E_j est l'énergie du spectre calculée sur la bande passante du $j^{ème}$ filtre, et N est le nombre de filtres ($N = 22$).

- Afin d'augmenter la robustesse de ces coefficients pour le calcul des distances cepstrales, une pondération en sinus (liftering) est appliquée sur les coefficients MFCC c_i [46] :

$$\hat{c}_i = \left(1 + \frac{L}{2} \sin \frac{i\pi}{L}\right) c_i \quad 1 \leq i \leq 12 \quad (11)$$

où \hat{c}_i est le $i^{ème}$ coefficient mel-cepstral liftré et L est le coefficient du liftering ($L = 22$). Ces pondérations corrigent la décroissance rapide des coefficients MFCC d'indice élevé et permet l'utilisation d'une distance euclidienne.

III.1.3 Les coefficients différentiels

Les systèmes de reconnaissance classiques ne modélisent pas l'évolution du signal. Il est possible d'introduire dans ces systèmes une information sur la dynamique temporelle du signal en utilisant, en plus des paramètres initiaux, des coefficients différentiels du premier ordre issus des coefficients cepstraux ou de l'énergie. soit $C_k(t)$ le coefficient cepstral d'indice k de la trame t , alors le coefficient différentiel $\Delta C_k(t)$ correspondant est calculé sur $2n_\Delta + 1$ trames d'analyse par :

Ces coefficients appelés aussi coefficient delta, estiment la pente de la régression linéaire

$$\Delta C_k = \frac{\sum_{i=-n_\Delta}^{n_\Delta} i \cdot C_k(t+i)}{\sum_{i=-n_\Delta}^{n_\Delta} i^2} \quad (12)$$

d'un coefficient à un instant donné. La dérivée de l'énergie ΔE_0 est calculée de la même façon.

Les expériences en montrant que les parties du signal de parole concentrant les plus fortes transitions spectrales correspondent à des 'des points perceptifs critiques' nécessaires à l'identification des syllabes, tandis que les zones stationnaires des silences et des voyelles semblent moins cruciales.

Des coefficients de second ordre peuvent aussi contribuer à l'amélioration du système surtout dans le cas de la parole bruitée soumise à l'effet lombard. Ces coefficients $\Delta \Delta C_k$ et $\Delta \Delta E_0$ sont calculés par la régression linéaire des coefficients delta sur $n_{\Delta\Delta}$

III.2 Moteur de reconnaissance

Nous avons utilisé la méthode de l'alignement temporel, plus connu sous l'acronyme de Dynamic Time Warping (DTW)

III.2.1 Description de la Méthode DTW

C'est une méthode fondée sur un principe de comparaison d'un signal à analyser avec un ensemble de signaux stockés dans une base de référence. Le signal à analyser est comparé avec chacune des références et est classé en fonction de sa proximité avec une des références stockées. Le DTW est en fait une application au domaine de la reconnaissance de la parole de la méthode plus générale de la programmation dynamique. Elle peut ainsi être vue comme un problème de cheminement dans un graphe.

Ce type de méthode pose deux problèmes : la taille de la base de référence, qui doit être importante, et la fonction de calcul des distances, qui doit être choisie avec soin. La taille de la base contenant les signaux de référence est directement liée aux capacités, variables, de reconnaissance du système d'alignement temporel. Chacun des signaux de référence est en effet stocké dans son état brut, sans compression d'aucune sorte. Ce stockage permet de disposer d'un vocabulaire dont la taille correspond au nombre de mots du vocabulaire multiplié par le nombre de locuteurs et le nombre des éventuelles répétitions des mots. Cette base de référence permet d'effectuer une mise en correspondance entre le signal stocké, d'une part, et sa retranscription symbolique d'autre part.

La taille de la base de référence est importante et implique une charge de travail non négligeable puisque la classification de chaque forme à analyser impose de la comparer à chaque forme de la base de référence. Donc, si la constitution de la base de référence est assez rapide et si le processus d'apprentissage est inexistant dans la méthode de l'alignement temporel, la phase d'utilisation nécessite une puissance de calcul non négligeable pour chaque référence atomique de signal à analyser. Le schéma de principe de la méthode est présenté dans la figure 1.8.

Comme le montre le schéma de la figure 1.8, la forme choisie sera celle pour laquelle le chemin de mise en correspondance est le plus court, cette taille minimale marquant le peu de différences entre la forme à analyser et la forme de référence.

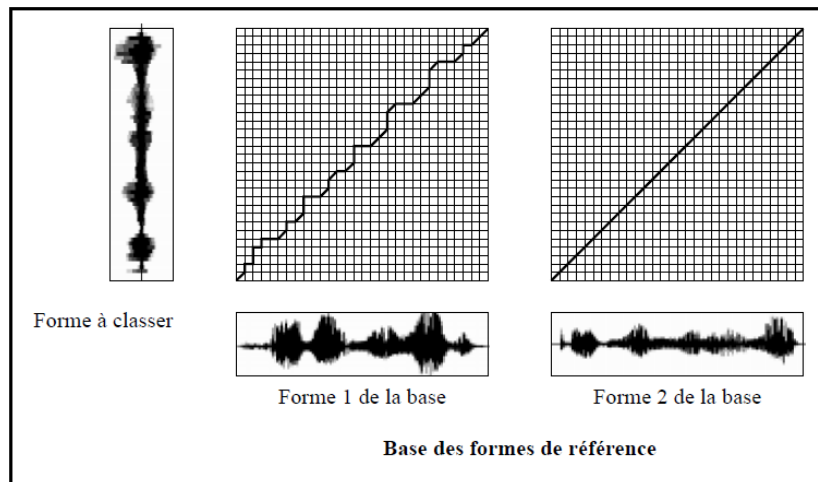


Figure 1.8 Visualisation du cheminement de l'alignement temporel pour des formes de la base de référence.

L'autre partie importante de l'alignement temporel est la définition de la fonction de recalage qui permet de calculer, selon certaines contraintes, la distance entre la forme à comparer et la forme de référence. La forme à analyser est mise en correspondance dans le plan temporel par l'algorithme d'alignement qui essaie de trouver le plus court chemin dans le graphe ainsi constitué. Cette fonction de mise en correspondance définit une valeur pour chaque arc du graphe, ces valeurs favorisant l'axe médian qui correspond à une parfaite mise en relation de la forme à analyser et d'une forme de référence comme le montre la figure 1.8.

La fonction de recalage suit typiquement le schéma présenté dans la figure 1.9. La fonction $d(i,j)$ est la fonction de calcul de la distance entre deux points successifs du graphe. Les valeurs α , β et γ permettent de définir une partie du comportement de la fonction d qui peut être soit symétrique ($\alpha = \gamma$) soit asymétrique ($\alpha \neq \gamma$). Ce calcul de distance entre deux noeuds successifs du graphe n'est cependant pas suffisant pour calculer la longueur totale du chemin parcouru dans le graphe. Une fonction supplémentaire, G , calcule une longueur totale qui permettra, après le calcul de cette longueur des chemins sur toutes les formes de la base de référence, de savoir à quel mot du vocabulaire préenregistré correspond la forme à classer. D'un point de vue mathématique, M et N étant les longueurs respectives de la forme à classer et de la forme de référence, on cherche sur l'ensemble du corpus la distance globale minimale $D(M,N)$.

III.2.2 Distance globale

Les distances cumulées représentent la dissemblance entre les références et les tests. Alors que les distances locales représentent la dissemblance entre les deux signaux en un instant donné, la distance cumulée en un point est la somme des distances locales depuis l'origine en suivant le chemin optimal, c'est à dire de moindre coût. Pour préserver une certaine cohérence dans le calcul du chemin optimal, les transitions autorisées entre les points du graphe de coïncidence sont limitées à quelques uns des points les plus proches.

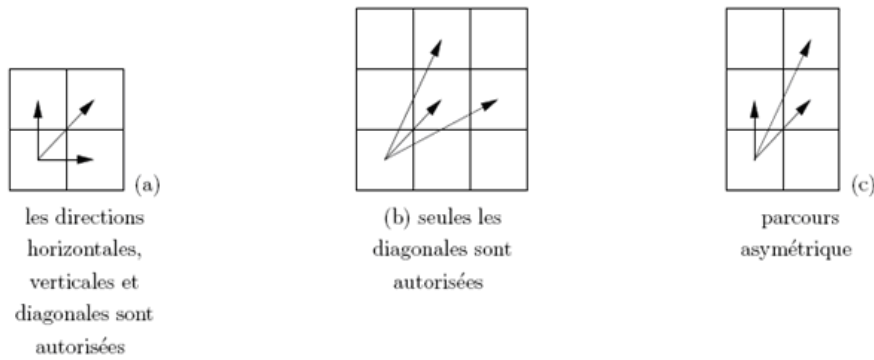


Figure 1.9 les transitions autorisées entre les points du graphe.

La distance cumulée au point (i,j) est obtenue de manière récursive par la formule suivante :

$$G(i, j) = \min \left\{ \begin{array}{l} G(i-1, j) + d(i, j) \\ G(i-1, j-1) + d(i, j) \\ G(i-1, j-2) + d(i, j) \end{array} \right\} \quad (13)$$

La distance globale entre l'observation T et la référence R est alors donnée par

$$D(R, T) = G(M, N) / (M + N) \quad (14)$$

Où M et N sont respectivement les nombres de trames des signaux R et T .

Le calcul de cette fonction G répond au même principe que le principe général énoncé par Bellman pour la programmation dynamique : toute sous-partie du chemin optimal est lui-même un chemin optimal. Des exemples de fonctions d et G de calcul de distance, qui peuvent être bien plus complexes que la fonction de recalage présentée en figure 1.10, pourront être trouvées dans [70] ou [69]. Dans ces références, les fonctions présentées peuvent analyser jusqu'à 9 chemins différents pour d , la fonction G étant de complexité égale à celle de d .

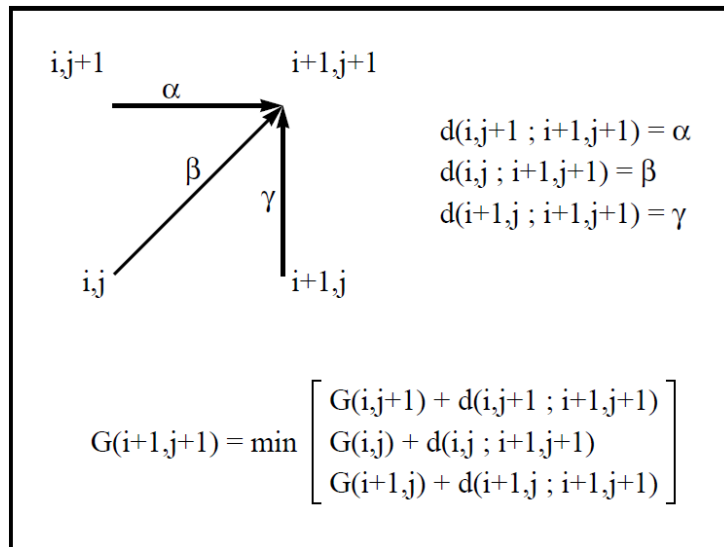


Figure 1.10 Schéma typique d'une fonction de recalage en alignement temporel.

Cette méthode de reconnaissance des formes est, initialement, bien adaptée à la reconnaissance de mots isolés mais des extensions ont été développées pour permettre de l'appliquer à la parole continue.

D'autres méthodes complémentaires ont par ailleurs été développées pour tenter de réduire la taille de la base des formes de référence par sélection optimale des formes à conserver. Ces méthodes reposent surtout sur une exploration statistique de la base des formes de référence et permettent d'obtenir une caractérisation des différents ensembles la constituant, ces ensembles correspondant aux différents symboles référencés dans la base. Une des techniques qu'il est possible d'employer pour ce faire est, par exemple, la méthode des plus proches voisins. Certaines méthodes permettent de réduire ce temps de calcul à l'utilisation par apprentissage a priori de coefficients qui permettent de compacter la connaissance présente dans la base de référence qui devient ainsi un corpus d'apprentissage.

IV Logiciel de reconnaissance vocale

Le logiciel effectue la paramétrisation acoustique et la reconnaissance vocale, il communique avec le non-voyant par l'émission des messages vocaux au cours et après chaque opération effectué par l'utilisateur, il établie des échanges avec la carte de

commande, qui à son tour transmet les valeurs de la température acquise par le moyen d'un convertisseur analogique numérique.

Le logiciel comporte une interface graphique ergonomique qui peut être utilisé par les gens qui ont une cécité visuel partiel, elle permet de commander un périphérique par un simple click de souris et voir l'état actuel de ce dernier, elle affiche la température de l'intérieur et de l'extérieur du domicile, elle affiche la date et l'heure.

Un menu avec aide est prévu afin de faciliter la tâche aux utilisateurs.

L'interface d'apprentissage permet à l'utilisateur d'introduire les commandes désirées au système. Le système lance un message vocal qui lui permette de prononcé la commande après le bip sonore, un enregistrement est déclenché, à la fin de cette enregistrement le système sauvegarde la commande dans un fichier WAV.

Le système effectuera plusieurs fois cette opération pour créer des commandes de références dès que le nombre est atteint le système ne demande plus à l'utilisateur de répéter.

Par la suite une extraction acoustique du type MFCC est effectuée sous forme de suite de vecteurs, les vecteurs sont ensuite sauvegardés dans la base de données d'apprentissage, ils vont être utilisés comme des commandes de références par le moteur de reconnaissance.

Lors de la phase de reconnaissance (utilisation quotidienne du système), le non-voyant d'éclanche la communication avec le système par appui sur sa commande RF, le système demande à l'utilisateur de prononcé la commande après un bip sonore, une fois terminé, le système tente de reconnaître la commande, et confirme par un message vocal si la commande reconnue est la même que la commande désiré, le non-voyant répond par (oui ou non), si la réponse est oui la commande est envoyée à la carte de commande, à son tour elle exécute la commande et renvoie l'état du périphérique et confirme par un message vocal.

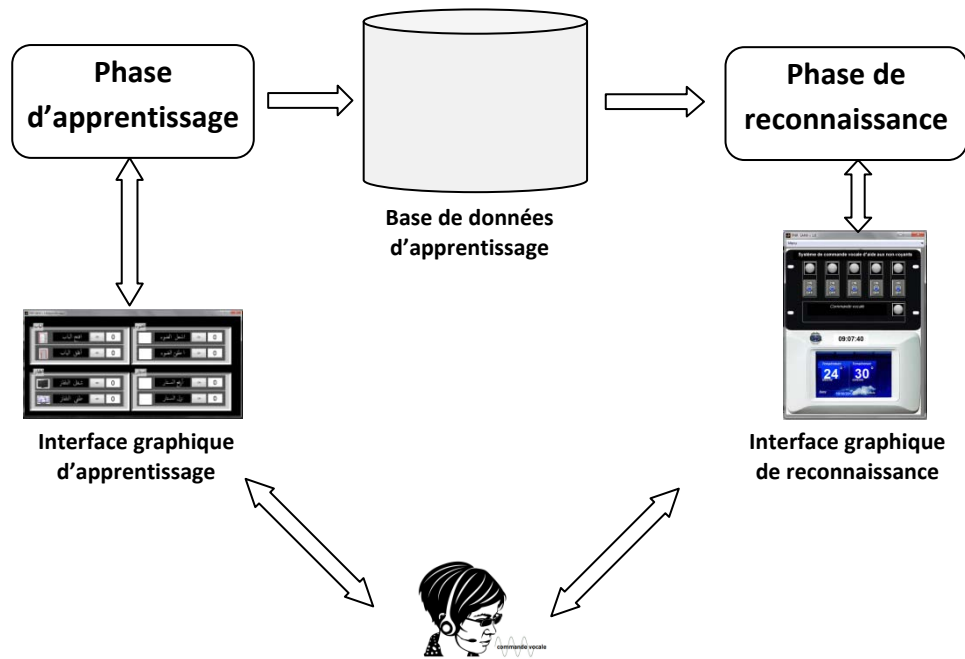


Figure 1.11 Représentation des différents blocs du logiciel d'aide au non voyant



Figure 1.12 interface graphique du logiciel d'aide au non-voyant

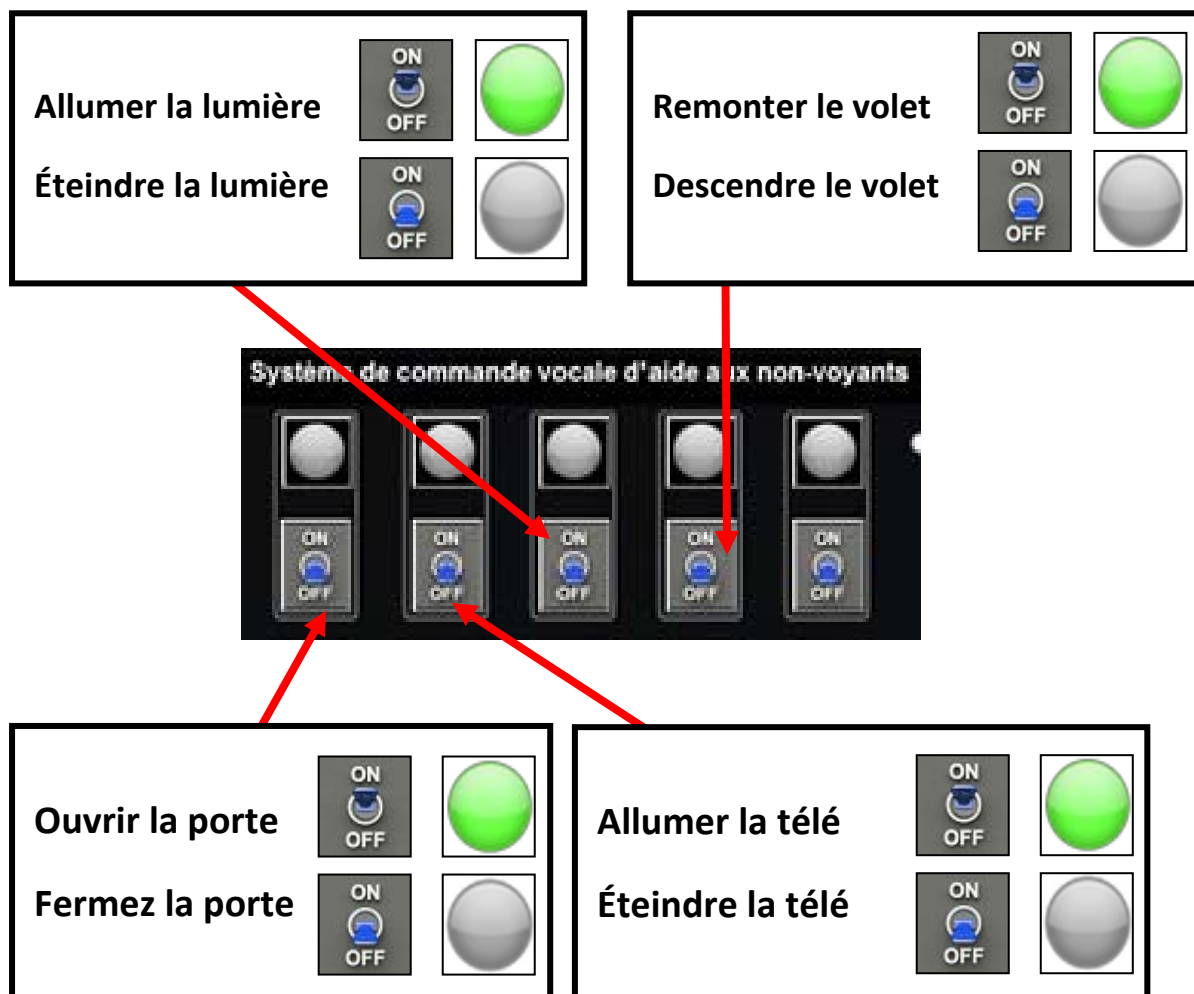


Figure 1.13 commande par interrupteurs



Figure 1.14 graphique du système d'apprentissage du logiciel d'aide au non-voyant

V Présentation de la carte de commande

Le cœur de notre carte est le microcontrôleur pic16F877A développé par Microchip, il permet un l'affichage 7 segments varié selon le mode (numéro du périphérique active, température interne ou externe, signaler l'enregistrement...etc.). (Voir figure xx)

Le microcontrôleur assure les échanges de données entre la carte et l'ordinateur via la communication série en mode asynchrone full duplexe.

L'acquisition de la température à l'intérieur et à l'extérieur du domicile est assurée par deux capteurs de température, qui sont conditionnés et liés aux entrées analogiques du convertisseur analogique numérique du microcontrôleur.

Un clavier composé de 7 boutons poussoirs permet une commande directe des périphérique, il est attaché au microcontrôleur ou à chaque bouton en attribut une action.

Un module récepteur infrarouge qui démodule le signal et le transforme en une trame binaire, cette trame est décodée par le microcontrôleur à chaque code on attribue une action à faire.

Un étage de puissance composé de triacs commandés à travers des opto-tric pour une bonne isolation galvanique de l'ensemble des composants de la carte, les trics ont été choisie de manière à assuré une commande rapide, silencieuse et efficace.

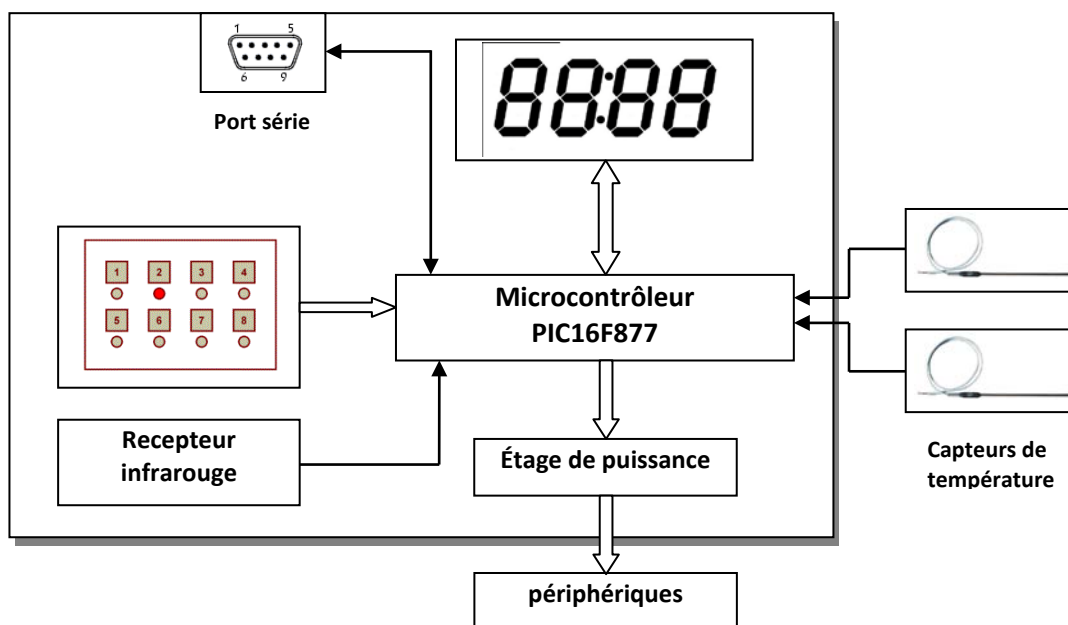


Figure 1.15 Schéma bloc de la carte de commande

V.1 Description détaillée du microcontrôleur pic16f877

Dans cette section nous allons décrire en détaille les différentes caractéristiques du microcontrôleur, l'espace mémoire, le brochage, l'architecture interne et le jeu d'instructions.

V.2 Caractéristiques du pic16f877

Mémoires
8Ko de mémoire Flash interne pour le programme
368 octets de RAM
256 octets de d'EEPROM
Entrées / sorties
33 entrées/sorties numériques configurables individuellement
Les Ports
Port A (06 entrées/sorties)
Port B (08 entrées/sorties)
Port C (08 entrées/sorties)
Port D (08 entrées/sorties)
Port E (03 entrées/sorties)
communication
1 port I2C / SPI
1 port série (RS232)
Timers
Timer 0 (08 bits)
Timer 1 (16 bits)
Timer 2 (08 bits)
ADC (convertisseur analogique/numérique)
Supporte jusqu'à 08 entrées analogiques
10 bits
1 Watchdog (chien de garde)
Interruptions
15 sources d'interruptions
35 instructions (composant RISC)
Instructions réduites pour un minimum de temps d'exécution
Fonctionne à 20 MHz maximum
5 cycles machine/ 1us
un mode SLEEP
Mode économique le réveil peut se faire par événement

Table 1 Caractéristiques du pic16f877

V.3 Brochage du pic16f877

Le pic 16f877 comporte 40 broches, les majores broches peuvent être configurées en entrée et en sortie, la figure ci-dessous montre le brochage du microcontrôleur

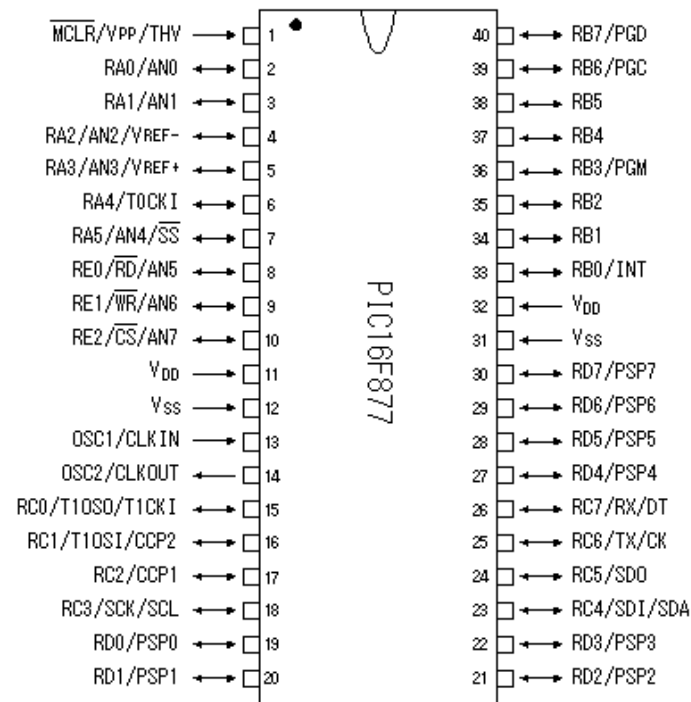
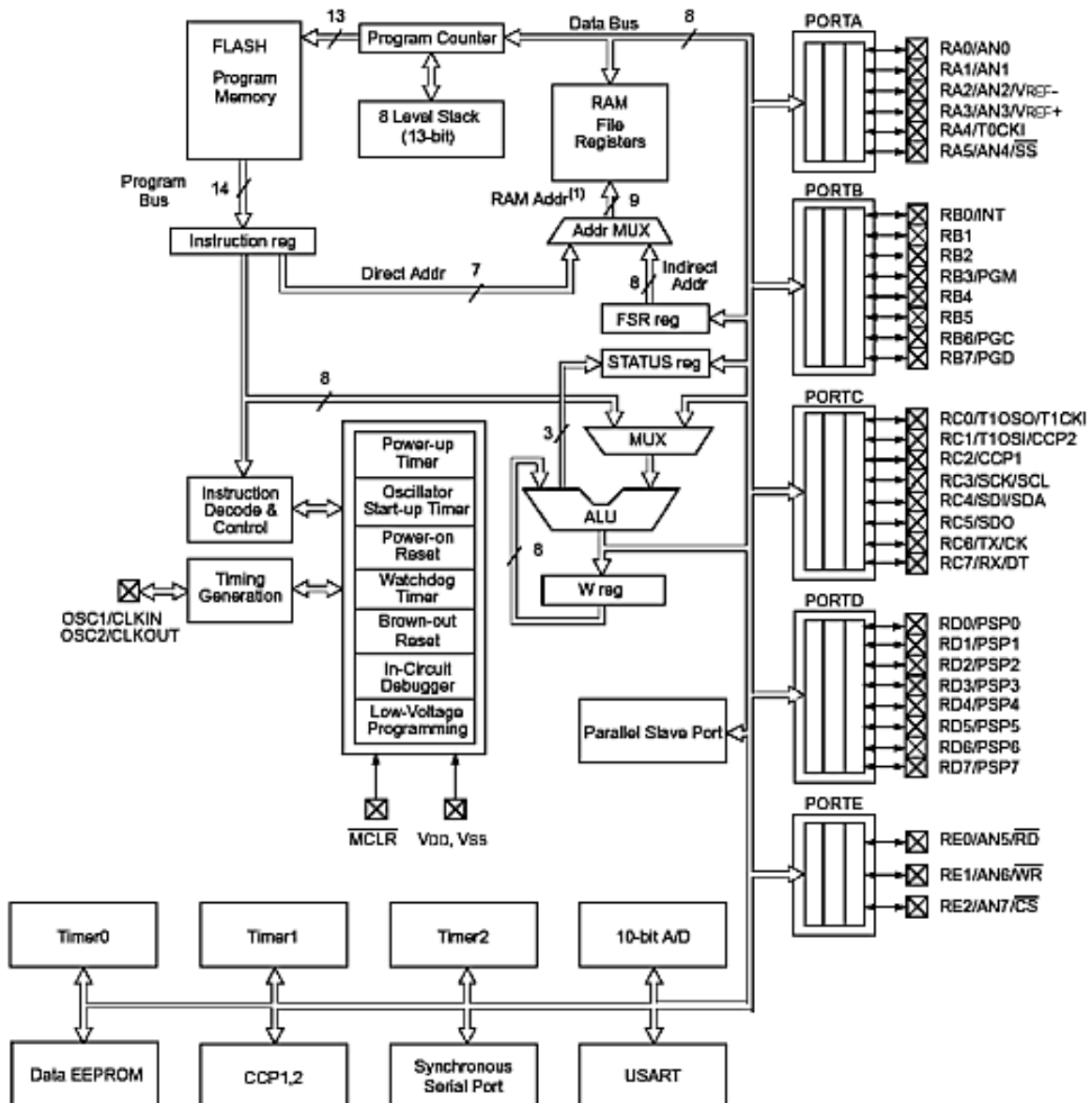


Figure 1.16 Boitier Dip 40 broches pour pic16f877

V.4 Architecture interne du pic 16f877

Device	Program FLASH	Data Memory	Data EEPROM
PIC16F874	4K	192 Bytes	128 Bytes
PIC16F877	8K	368 Bytes	256 Bytes



Note 1: Higher order bits are from the STATUS register.

Figure 1.17 Architecture interne du pic 16f877

V.5 Organisation de la mémoire du pic 16f877

La RAM apparaît alors organisée en 4 banques de 128 octets chacune, Les zones hachurées ne sont pas implantées physiquement.

source : PIC16F877A
data
manual,microchip.inc

FIGURE 2-3: PIC16F876A/877A REGISTER FILE MAP

File Address	File Address	File Address	File Address
Indirect addr. ^(*) 00h	Indirect addr. ^(*) 80h	Indirect addr. ^(*) 100h	Indirect addr. ^(*) 180h
TMR0 01h	OPTION_REG 81h	TMR0 101h	OPTION_REG 181h
PCL 02h	PCL 82h	PCL 102h	PCL 182h
STATUS 03h	STATUS 83h	STATUS 103h	STATUS 183h
FSR 04h	FSR 84h	FSR 104h	FSR 184h
PORTA 05h	TRISA 85h	105h	185h
PORTB 06h	TRISB 86h	PORTB 106h	TRISB 186h
PORTC 07h	TRISC 87h	107h	187h
PORTD ⁽¹⁾ 08h	TRISD ⁽¹⁾ 88h	108h	188h
PORTE ⁽¹⁾ 09h	TRISE ⁽¹⁾ 89h	109h	189h
PCLATH 0Ah	PCLATH 8Ah	PCLATH 10Ah	PCLATH 18Ah
INTCON 0Bh	INTCON 8Bh	INTCON 10Bh	INTCON 18Bh
PIR1 0Ch	PIE1 8Ch	EEDATA 10Ch	EECON1 18Ch
PIR2 0Dh	PIE2 8Dh	EEADR 10Dh	EECON2 18Dh
TMR1L 0Eh	PCON 8Eh	EEDATH 10Eh	Reserved ⁽²⁾ 18Eh
TMR1H 0Fh	8Fh	EEADRH 10Fh	Reserved ⁽²⁾ 18Fh
T1CON 10h	90h	110h	190h
TMR2 11h	SSPCON2 91h	111h	191h
T2CON 12h	PR2 92h	112h	192h
SSPBUF 13h	SSPADDD 93h	113h	193h
SSPCON 14h	SSPSTAT 94h	114h	194h
CCPR1L 15h	95h	115h	195h
CCPR1H 16h	96h	116h	196h
CCP1CON 17h	97h	117h	197h
RCSTA 18h	TXSTA 98h	General Purpose Register 16 Bytes 118h	General Purpose Register 16 Bytes 198h
TXREG 19h	SPBRG 99h	119h	199h
RCREG 1Ah	9Ah	11Ah	19Ah
CCPR2L 1Bh	9Bh	11Bh	19Bh
CCPR2H 1Ch	CMCON 9Ch	11Ch	19Ch
CCP2CON 1Dh	CVRCON 9Dh	11Dh	19Dh
ADRESH 1Eh	ADRESL 9Eh	11Eh	19Eh
ADCON0 1Fh	ADCON1 9Fh	11Fh	19Fh
20h	A0h	120h	1A0h
General Purpose Register 96 Bytes	General Purpose Register 80 Bytes	General Purpose Register 80 Bytes	General Purpose Register 80 Bytes
7Fh	EFh	16Fh	1EFh
Bank 0	accesses 70h-7Fh F0h	accesses 70h-7Fh 170h	accesses 70h - 7Fh 1F0h
	FFh	17Fh	1FFh
Bank 1		Bank 2	Bank 3


 Unimplemented data memory locations, read as '0'.
 * Not a physical register.

figure 1.18 Organisation de la mémoire RAM du pic16f877

V.6 Jeu d'instructions du pic16f877

Une instruction nécessite 1 cycle, ou bien 2 cycles dans le cas d'une instruction de branchement (ratern, retlw ...).

Avec une horloge à quartz de 20 MHz, un cycle correspond à $4/(20 \cdot 10^6) = 200$ nanosecondes.

Le microcontrôleur peut donc exécuter jusqu'à 5 millions d'instructions par seconde (5 MIPS)

Mnemonic, Operands		Description	Cycles	14-Bit Opcode				Status Affected	Notes
				MSb		LSb			
BYTE-ORIENTED FILE REGISTER OPERATIONS									
ADDWF	f, d	Add W and f	1	00	0111	dfff	ffff	C,DC,Z	1,2
ANDWF	f, d	AND W with f	1	00	0101	dfff	ffff	Z	1,2
CLRF	f	Clear f	1	00	0001	1fff	ffff	Z	2
CLRWF	-	Clear W	1	00	0001	0xxx	xxxx	Z	
COMF	f, d	Complement f	1	00	1001	dfff	ffff	Z	1,2
DECf	f, d	Decrement f	1	00	0011	dfff	ffff	Z	1,2
DECFSZ	f, d	Decrement f, Skip if 0	1(2)	00	1011	dfff	ffff		1,2,3
INCF	f, d	Increment f	1	00	1010	dfff	ffff	Z	1,2
INCFSZ	f, d	Increment f, Skip if 0	1(2)	00	1111	dfff	ffff		1,2,3
IORWF	f, d	Inclusive OR W with f	1	00	0100	dfff	ffff	Z	1,2
MOVF	f, d	Move f	1	00	1000	dfff	ffff	Z	1,2
MOVWF	f	Move W to f	1	00	0000	1fff	ffff		
NOP	-	No Operation	1	00	0000	0xx0	0000		
RLF	f, d	Rotate Left f through Carry	1	00	1101	dfff	ffff	C	1,2
RRF	f, d	Rotate Right f through Carry	1	00	1100	dfff	ffff	C	1,2
SUBWF	f, d	Subtract W from f	1	00	0010	dfff	ffff	C,DC,Z	1,2
SWAPF	f, d	Swap nibbles in f	1	00	1110	dfff	ffff		1,2
XORWF	f, d	Exclusive OR W with f	1	00	0110	dfff	ffff	Z	1,2
BIT-ORIENTED FILE REGISTER OPERATIONS									
BCF	f, b	Bit Clear f	1	01	00bb	bfff	ffff		1,2
BSF	f, b	Bit Set f	1	01	01bb	bfff	ffff		1,2
BTFSC	f, b	Bit Test f, Skip if Clear	1 (2)	01	10bb	bfff	ffff		3
BTFSS	f, b	Bit Test f, Skip if Set	1 (2)	01	11bb	bfff	ffff		3
LITERAL AND CONTROL OPERATIONS									
ADDLW	k	Add literal and W	1	11	111x	kkkk	kkkk	C,DC,Z	
ANDLW	k	AND literal with W	1	11	1001	kkkk	kkkk	Z	
CALL	k	Call subroutine	2	10	0kkk	kkkk	kkkk		
CLRWDT	-	Clear Watchdog Timer	1	00	0000	0110	0100	$\overline{TO}, \overline{PD}$	
GOTO	k	Go to address	2	10	1kkk	kkkk	kkkk		
IORLW	k	Inclusive OR literal with W	1	11	1000	kkkk	kkkk	Z	
MOVLW	k	Move literal to W	1	11	00xx	kkkk	kkkk		
RETFIE	-	Return from interrupt	2	00	0000	0000	1001		
RETLW	k	Return with literal in W	2	11	01xx	kkkk	kkkk		
RETURN	-	Return from Subroutine	2	00	0000	0000	1000		
SLEEP	-	Go into standby mode	1	00	0000	0110	0011	$\overline{TO}, \overline{PD}$	
SUBLW	k	Subtract W from literal	1	11	110x	kkkk	kkkk	C,DC,Z	
XORLW	k	Exclusive OR literal with W	1	11	1010	kkkk	kkkk	Z	

- Note** 1: When an I/O register is modified as a function of itself (e.g., MOVF PORTB, 1), the value used will be that value present on the pins themselves. For example, if the data latch is '1' for a pin configured as input and is driven low by an external device, the data will be written back with a '0'.
- 2: If this instruction is executed on the TMR0 register (and, where applicable, d = 1), the prescaler will be cleared if assigned to the Timer0 module.
- 3: If Program Counter (PC) is modified, or a conditional test is true, the instruction requires two cycles. The second cycle is executed as a NOP.

Table 2 jeu d'instructions du pic16f877

V.7 Schéma électronique de la carte de commande

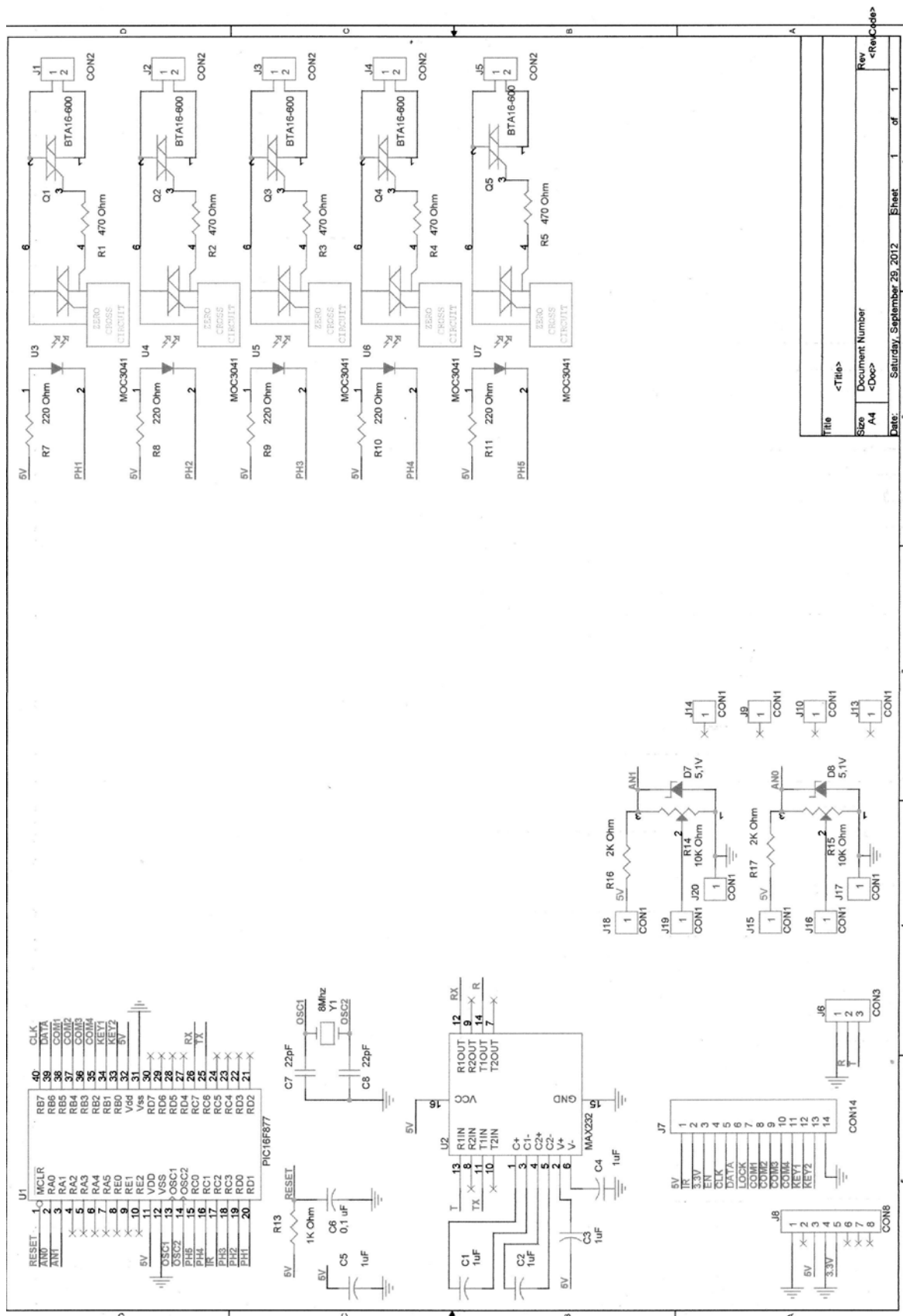


Figure 1.19 Schéma de la carte PC de commande

V.8 Circuit imprimé de la carte de commande

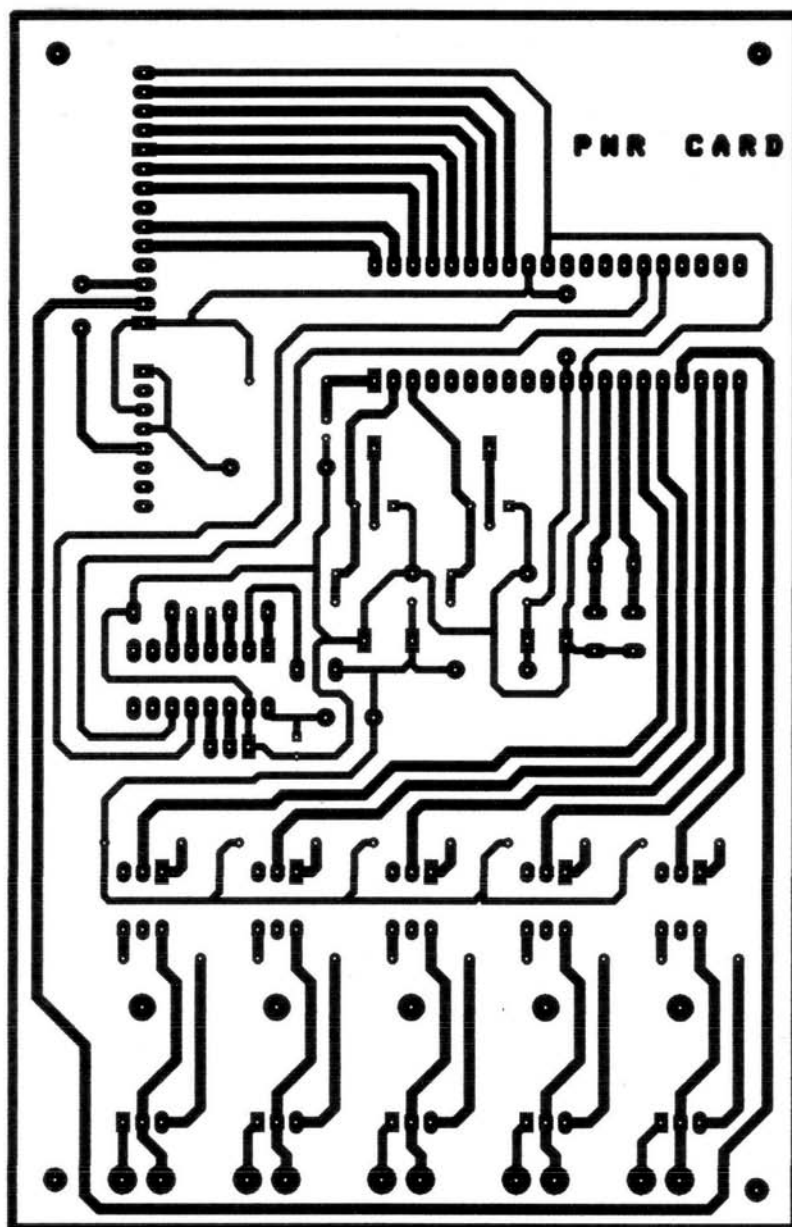


Figure 1.20 Circuit imprimé de la carte pc de commande



Figure 1.21 dispositif de commande

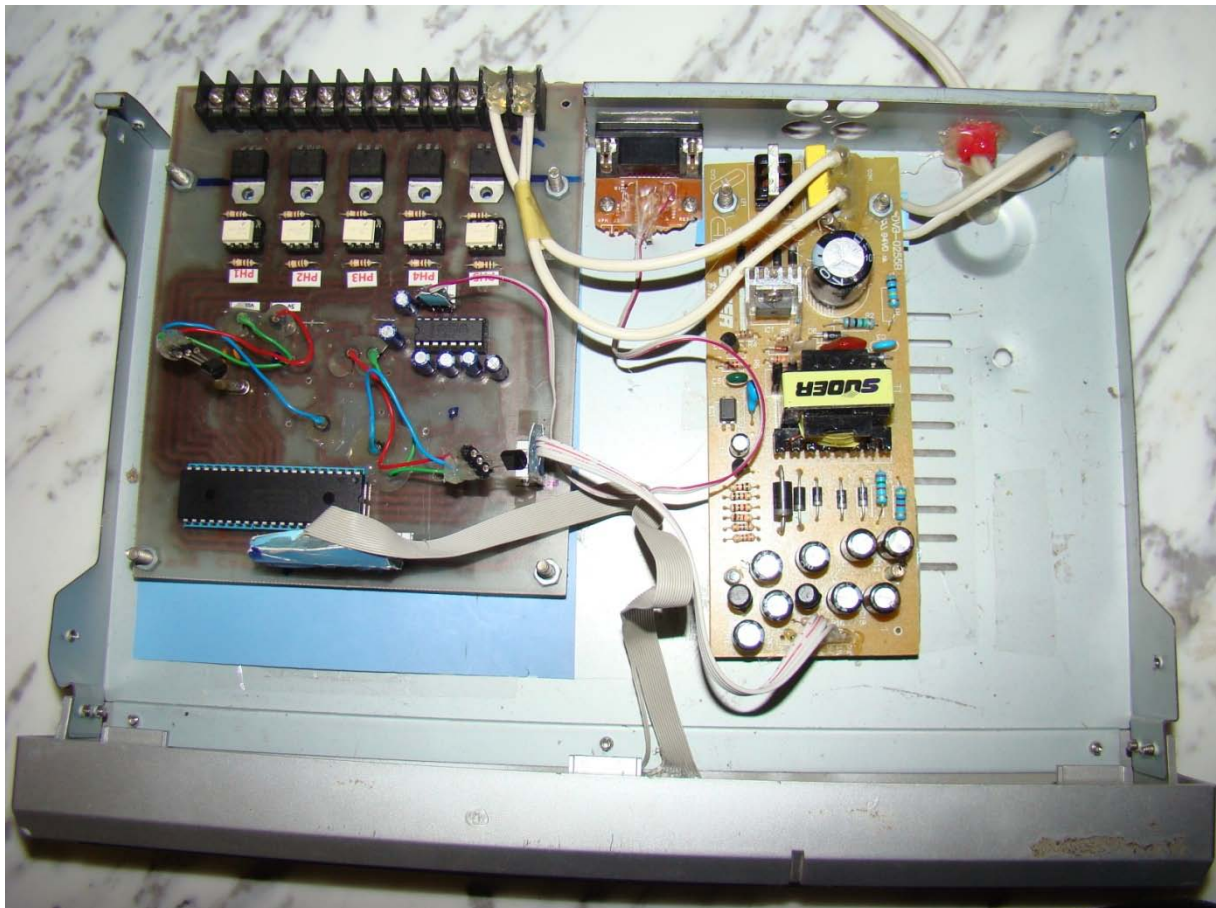


Figure 1.22 Dispositif de commande sans couvercle

V.9 Schéma de la carte d'alimentation

Nous avons utilisé dans notre application une alimentation à découpage, dont la régulation est assurée par des composants électrotechnique utilisés en commutation . Ce mode de fonctionnement diffère de celui des alimentations linéaires dans lesquelles les composants électrotechniques sont utilisés en mode linéaire. Pour pallier les inconvénients des alimentations linéaires : poids élevé et faible rendement, nous avons utilisé une alimentation utilisée dans les cartes des récepteurs satellite. La figure ci-dessous montre son schéma électrique.

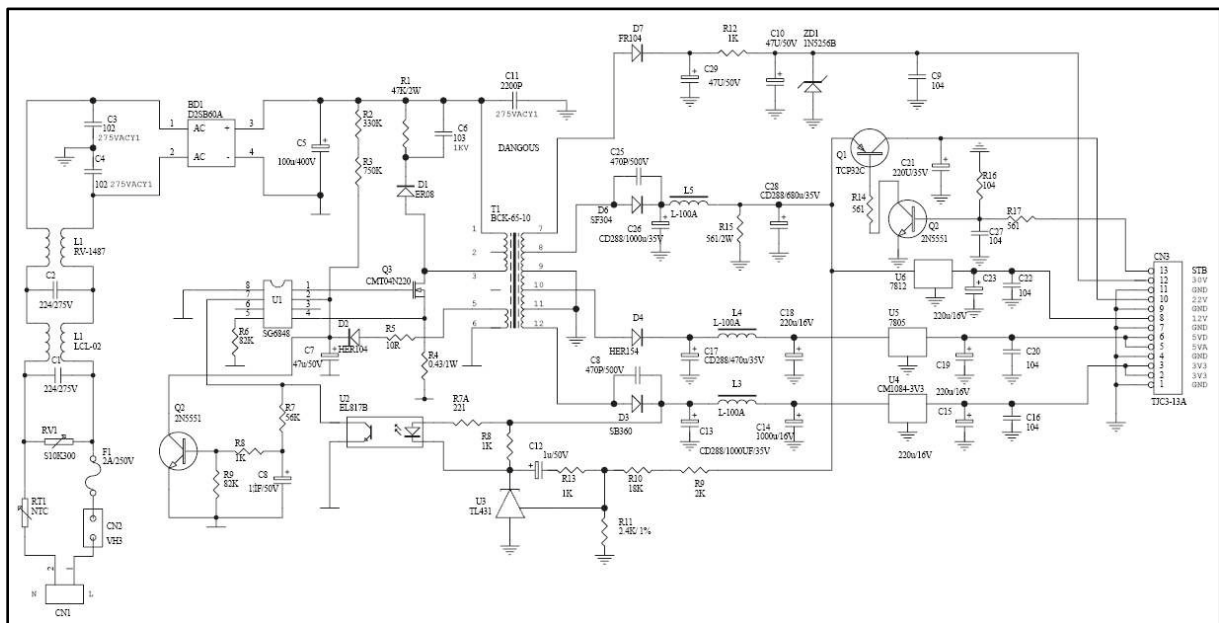


Figure 1.23 Schéma électrique de la carte d'alimentation

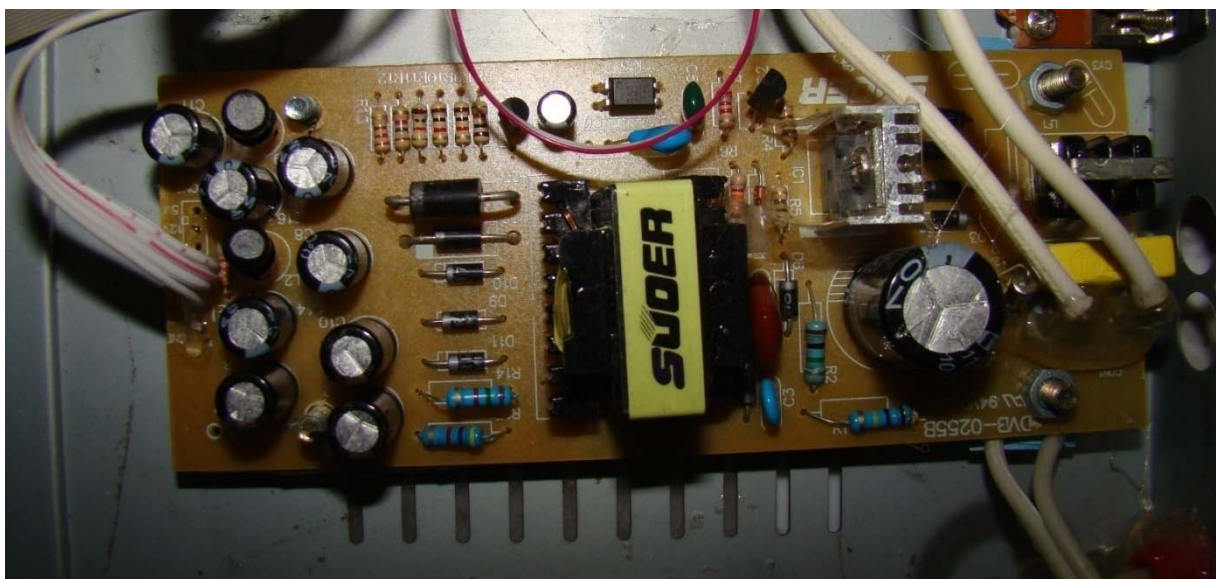


Figure 1.24 Carte d'alimentation

Conclusion

Le travail que nous avons entrepris dans le présent projet, vise à apporter des solutions techniques pour assurer un maximum d'aide et de confort au non-voyant, il permet de simplifier les tâches quotidiennes effectuées par le non-voyant dans son domicile, il peut aussi être utilisé par les handicapés moteurs.

Cette carte de commande, communique en temps réel avec le logiciel de reconnaissance vocale installé sur un ordinateur fonctionnant sur windows XP ou version plus récente, ce logiciel de reconnaissance vocale réalisé par notre équipe est menu d'une interface graphique moderne, ergonomique et interactive, conçue pour répondre aux exigences du non voyant.

Cette interface permet la communication directe de l'utilisateur avec la machine, qui à son tour, communique avec la carte électronique, cette dernière commande les équipements à distance.

Les objectifs visés au début de ce travail ont été atteints, nous avons pu réaliser un système d'aide au non-voyant par commande vocale, interactif est fonctionnel à 100%.

Le prototype réalisé peut être reproduit et commercialisé par des industriels. Notre réalisation essentiellement basée sur la conception d'une carte de commande programmable à base de microprocesseur, ce qui la rend plus maniable et sa prise en main est plus facile.

Bibliographie

- [1] M.C. Amara Korba, D. Messadeg, R. Djemili, H. Bourouba. "Robust Speech Recognition Using Perceptual Wavelet Denoising and Mel-frequency Product Spectrum Cepstral Coefficient Features", *Informatica Journal*, Vol. 32, No 3, pp. 283-288, 2008.
- [2] N.Q. Trung; P.T. Nghia, "The perceptual wavelet feature for noise robust Vietnamese speech recognition", *Communications and Electronics ICCE2008*, Vol. 2, pp. 258-261, 2008.
- [3] Shi-Huang Chen, Hsin-Te Wu, Yukon Chang, T. K. Truong, "Robust voice activity detection using perceptual wavelet-packet transform and Teager energy operator", *Pattern Recognition Letters*, vol. 28, pp. 1327-1332, 2007.
- [4] M.T. Johnson, X. Yuan and Y. Ren, "Speech signal enhancement through adaptive wavelet thresholding", *Speech Communication, Elsevier*, Vol. 49, pp. 123-133, 2007.
- [5] T. Haci, E. Ergun, "Speech Enhancement based on undecimated wavelet packet-perceptual filter-banks and MMSE-STSA estimation in various noise environments". *Elsevier, Digital Signal Processing*, 2007.
- [6] A. Saeed, M. T. Manzuri, D. Roodhollah, "An improved wavelet-based speech enhancement by using speech signal features", *Elsevier, Computer and Electrical Engineering* Vol. 32, pp. 411-424. 2006.
- [7] B. kotnik, Z. Kacic, "A noise robust feature extraction algorithm using joint wavelet packet subband decomposition and AR modelling of speech signals", *Elsevier, Signal Processing*, Vol. 87, pp. 1202-1223, 2006.
- [8] Yu Shao, Chip-Hong Chang, "A versatile speech enhancement system based on perceptual wavelet denoising", *ISCAS 2005*, Vol. 2, pp. 864-867, 2005.
- [9] Shao, Y. Chang, C.-H, "A versatile speech enhancement system based on perceptual wavelet denoising", *IEEE international Symposium on circuits and systems*, Vol. 2, pp. 864-867, 2005.
- [10] C.C. lin, S.H. Chen, T.K. troung and Y. Chang, "Audio classification and categorization based on wavelets and support vector machine". *IEEE transaction on speech and audio processing*, vol.13, pages 644-651, 2005.
- [11] D. Dimitriadis, P. Maragos and A. Potamios, "Auditory teager energy cestrum coefficients for robust speech recognition". In *European Conference On speech communication and Technology*, pages 3013-3016, 2005.
- [12] Y. Guermeur, A. Eliseef and D. Zelus, "A Comparative study of multi-class classifiers. Applied stochastic model in business and industry", Vol. 21, 2005.
- [13] V. Wan & J. Carmichael, "Polynomial dynamic time Warping kernel support vector machines for dysarthric speech recognition with sparse training data". In *INTERSPEECH*, 2005.
- [14] V. Wan & S. Renais, "Speaker verification using sequence discriminant support vector machines". *IEEE Transaction on Speech and Audio Processing*, Vol. 13, 2005.
- [15] M. Phothisonothai, P. Kumhom, and K. Chamnongthai, "Single-Channel Noise Reduction for Multiple Background Noises Using Perceptual Wavelet Packet Transform and Fuzzy Logic", *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 8, No. 6, pp. 613-620, 2004.
- [16] S.H. Chen, J. Wang, "Speech Enhancement Using Perceptual Wavelet Packet Decomposition and Teager Energy Operator", *Springer, The Journal of VLSI Signal Processing*, Vol. 36, No. 2, pp. 125-139, 2004.

- [17] M. Deviren, "Revising speech recognition systems: dynamic bayesian networks and new computational paradigms". Phd thesis, Université Henri Poincaré, Nancy, France, 2004.
- [18] Z. Donglai and K. K. Paliwal, "Product of power spectrum and group delay function for speech recognition", Proc. ICASSP, pp. 125-128, 2004.
- [19] O. Farooq and S. Datta, "Wavelet-based Denoising for Robust Feature Extraction for Speech Recognition", electronics letters, Vol. 39, No 1, pp. 163-165, 2003.
- [20] B. kotnik, Z. Kacic and B. Horvat, "The usage of wavelet packet transformation in automatic noisy speech recognition systems", Proceeding EROCON 2003, pp. 131-134, 2003.
- [21] H.A. Murthy and V. Gadde, "The Modified Group Delay Function and Its Application to Phoneme Recognition", Proc. ICASSP, vol. 1, pp. 68-71, 2003.
- [22] J. Ajmera, I. McCowan & H. Bourlard. "Speech/Music Discrimination using Entropy and Dynamism Features in a HMM Classification". *Frameuiork; Speech Communication*, vol. 40, pp. 351-363, 2003.
- [23] M. Deviren and K. Daoudi, "Frequency filtering or wavelet filtering". In joint Intl. Conf. on Artificial Neural Networks and Neural Information Processing, ICANN/ICONIP, 2003.
- [24] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals". *IEEE transaction on speech and audio processing*, Vol. 10, No. 5, pp. 293-302, 2002.
- [25] I. J. Kim, S.I. Yang and Kwon, "Speech enhancement using adaptive wavelet shrinkage". In *ISIE-2001*, vol. 1, pp. 501-504, 2001.
- [26] R. Gemello, D. Albesano, L. Moisa and R. De Mori, "Integration of fixed and Multiple resolution analysis in a speech recognition system". In *ICASSP-01*, 2001.
- [27] O. Farooq and S. Datta, "Robust features for speech recognition based on admissible wavelet packets", *Electronics letters*, Vol. 37, No 5, pp. 1554-1556, 2001.
- [28] N. Gowda and Z. Tufekci, "Mel-scaled discrete wavelet coefficients for speech recognition", *Proc. Int. Conf. on acoustic, Speech and signal processing*, Vol. 3, pp. 1351-1354, Istanbul, Turkey, 2000.
- [29] S. Mallat. "Une exploration des signaux en ondelettes". Editions de l'Ecole polytechnique, 2000.
- [30] R. Sarikaya and J.H.L. Hansen. "High resolution speech feature parameterization for monophone-based stressed speech recognition". *IEEE. Signal processing letters*, vol. 7, No 7, pp. 182-185, 2000.
- [31] S. Saha, "Image compression from DCT to wavelets", *ACM grossroads*, Vol. 6, No. 3, pp. 644-651, 2000.
- [32] A. Ganapathiraju & J. Picone, "Hybrid SVM/HMM Architectures for Speech Recognition". In *Neural Information Processing Systems*, 2000.
- [33] F. Jabloun and A. Enis Cetin, "The teager energy based feature parameters for robust speech recognition in car noise", In *ICASSP 99*, 1999.
- [34] S. Mallat, "A wavelet tour of signal of signal processing". Academic press, 1998.
- [35] I. Pinter, "Perceptual wavelet-representation of speech signals and its application to speech enhancement", *Computer speech & language*, Vol. 10, No. 1, pp. 1-22, 1996.
- [36] Yifan Gong, "Speech recognition in noisy environments: a survey". *Speech communication*, Vol. 16, pp. 261-291, 1995.

- [37] Ephraim, "Gain-adapted Hidden Markov Models for Recognition of Clean and Noisy Speech". IEEE Trans. Signal Processing, Vol. 40, pp. 1303-1316, 1992.
- [38] F. Jabloun and A. Enis Cetin, "The teager energy based feature parameters for robust speech recognition in car noise", In ICASSP 99, 1999.
- [39] R. Vergin and D. O'Shaughnessy, "Generalized Mel frequency cepstral coefficients for large-vocabulary speaker-independent continuous speech recognition", IEEE Trans. Speech, Audio Process., vol. 7, no. 5, pp. 525-532, Sept. 1999.
- [40] S. Mallat, "A wavelet tour of signal of signal processing". Academic press, 1998.
- [41] V.N. Vapnik, "Statistical learning theory", John Wiley and Son, Inc N.Y. 1998.
- [42] L.R Rabiner and B. Juang, "A Tutorial on hidden Markov Models and select application in speech recognition", Proceedings of IEEE, Vol. 77, No. 2, pp. 257-285, 1989.
- [43] B. Moore, "An introduction to the psychology of hearing", Academic Press, 1997.
- [44] J. S. Bridle, "Optimization and search in speech and language processing". Survey of the state of the art in human language technology, pp. 423-428, 1995.
- [45] O. Cappé, J. Laroche et E. Moulines, "Regularized estimation of cepstrum envelope from discrete frequency points", IEEE ASSP Work-shop on application of signal processing to audio and acoustic (1995), p. 213-216. 1995.
- [46] L. Rabiner, B. H. Juang, "Fundamentals of Speech Recognition". Prentice Hall Signal Processing Series, 1993.
- [47] J.W. Picone, "Signal modeling techniques in speech recognition", Proc. IEEE, Vol. 81, No. 9, pp. 1215-1247, 1993.
- [48] L. Rabiner and B.-H. Juang, "Fundamentals of speech recognition", Prentice-Hall, 1993.
- [49] I. Daubechies, "Ten lectures on wavelets", Society for industrial and applied Mathematics, 1992.
- [50] B. Yegnanarayana and H.A. Murthy, "Significance of Group Delay Functions in Spectrum Estimation", IEEE Trans. Signal Processing, Vol. 40, pp. 2281-2289, 1992.
- [51] B. Boser, I. Guyon and V. Vapnik, "A training algorithm for optimal margin classifiers". In COLT'92, pp. 144-152, 1992.
- [52] H. Hermansky, N. Morgan, A. Bayya et P. Kohn, "Compensation for the effect of the communication channel in auditory-like analysis of speech (RASTA-PLP)". Proceedings of the European Conference on Speech Communication and Technology, pp 1367-1370, 1991.
- [53] H. Hermansky, N. Morgan, A. Bayya et P. Kohn. "RASTA-PLP speech analysis". Rapport technique TR-91-069, 6 pp, International Computer Science Institute, Berkeley (CA, États-Unis), 1991.
- [54] J.-P. Haton, J.-M. Pierrel, G. Pérennou, J. Caelen et J.-L. Gauvain. "Reconnaissance automatique de la parole", 239 p, Collection AFCET - Dunod informatique, Dunod, 1991.
- [55] T. Galas et X. Rodet, "Generalized functional approximation for source-filter system modeling", Proc. of Eurospeech 1991, pp. 1085-1088. 1991.

- [56] H. Bourlard & C. Welckens. "Links between Markov models and multilayer perceptrons". In Trans. PAMI, Vol. 12, pp. 1167-1178, 1990.
- [57] H. Hermansky, "Perceptual linear predictive (plp) analysis of speech", Journal of Acoustical Society of America, Vol. 87, pp. 1738-1752, 1990.
- [58] T. Galas et X. Rodet, "An improved cepstral method for deconvolution of source-filter systems with discrete spectra : application to musical sound signals", In ICMC, 1990.
- [59] K. Hornik, M. Stinchcombe and H. White. "Multilayer feedforward networks are universal approximators". Neural Networks, vol. 2, 1989.
- [60] L.R. Rabiner, J.G. Wilpon & F.K. Soong, "High performance connected digit recognition using hidden Markov models", IEEE Trans. Acoust., Speech, Signal Processing, Vol. 37, No. 8, pp. 1214-1225, 1989.
- [61] J.J. Hopfield. "Learning algorithms and probability distributions in feed-forward networks". In Nat. Acad. Sci, pp. 8429-8433, 1987.
- [62] Y. Lecun, "Une procédure d'apprentissage pour réseaux à seuil asymétrique". In proc. Cognitiva, pp. 599-604, 1985.
- [63] V.N. Vapnik, "Estimation of dependence based on empirical data", Springer-Verlag, N.Y, 1982.
- [64] C. S. Myers et L. R. Rabiner, "Connected digit recognition using a level building DTW algorithm". IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 29, pp 351-363, 1981.
- [65] S.B. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences", IEEE Trans. Acoustics, Speech, Signal Processing, Vol. 28, No. 4, pp. 357-366, 1980.
- [66] H. Sakoe, "Two level DP-matching a dynamic programming based pattern matching algorithm for connected word recognition". IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. 27, pp. 588-595, 1979.
- [67] J.-P. Zerling, "Articulation et coarticulation dans les groupes occlusive-voyelle en français". Thèse de doctorat de 3ème cycle, Université de Nancy 2, Nancy (France), 1979.
- [68] L.R. Rabiner and R.W. Schafer, "Digital processing of speech signals", Prentice-Hall, 1978.
- [69] H. Sakoe et S. Chiba, "Dynamic programming algorithms optimization for spoken word recognition". IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 26, No. 1, pp. 43-49, 1978.
- [70] F. Itakura, "Minimum production residual principle applied to speech recognition". IEEE Transaction on Acoustics, Speech and Signal Processing, vol. 23, pp 67-72, 1975.
- [71] J.M. Makhoul, "Linear prediction: a tutorial review", Proc. IEEE, Vol. 63, No. 4, pp. 561-579, 1975.
- [72] A.V. Oppenheim and R.W. Schafer, "Digital Signal Processing", Englewood Cliffs, NJ: Prentice-Hall, 1975.
- [73] A. V. Aho, J. E. Hopcroft et J. D. Ullman, "The design and analysis of computer algorithms, chapitre 7 : The fast Fourier transform and its application", pp. 251-276, 1974.
- [74] H. Sakoe et S. Chiba, "A dynamic programming approach to continuous speech recognition", Proceedings of the 7th International Conference on Acoustics, article 20C-13, 6 pp, 1971.

- [75]** J.W. Cooley & J.W. Tukey, "An algorithm for machine calculation of complex Fourier series", Math.Comput, vol. 19, pp. 297-301, 1965.
- [76]** F. Rosenblatt, "Principles of neurodynamics", Spartan Books, 1962.
- [77]** R. E. Bellman, "On a routing problem", Quaterly Journal of Applied Mathematics, vol. 16, pp. 87-90, 1958.
- [78]** R. E. Bellman. "Dynamic Programming", Princeton University Press, 1957.
- [79]** W. Koenig, H.K. Duhn & L.Y. Lacy, "Thesound spectrograph", J. Acoust. Soc. Am, Vol. 18, pp. 19- 49, 1946.
- [80]** D.Gabor, "Theory of communication", J. IEEE, Vol. 93, pp. 429-457, 1946.
- [81]** Calliope, "La parole et son traitement automatique", livre, Collection technique et scientifique des télécommunications, CNET - ENST, Masson.
- [82]** B. Carneno, A. Drygajlo, "Perceptual speech coding and enhancement using frame-synchronized fast wavelet-packet transform algorithms". IEEE Trans. Signal Process, Vol. 47, No 6, pp. 1622-1635, 1999.
- [83]** P. Srinivasan, L.H. Jamieson, "High quality audio compression using an adaptive wavelet decomposition and psychoacoustic modelling". IEEE Trans. Signal Process, Vol. 46, No 4, pp. 1085-1093, 1998.
- [84]** D. L. Donoho, "De-noising by Soft-thresholding", IEEE Trans. Inform Theory, Vol. 41, No. 3, pp. 613-627, 1995.

Information financière

Quantité	produits
01	Station à souder Weller WSD 81
01	Oscilloscope numérique UT2025C
03	Multimètre fluke 15b
01	EasyPIC6 Development System
01	Programmateurs Superpro

Table 3. Nomenclature des dépenses