

Application pour les sourds et les muets

Arwa Ezzeddine

Département physique , Institut national des sciences appliquées et de technologie, Tunisie

auteurs correspondant : Ezzeddine Arwa ;
Gastli Mouna ; Mighri Hadhami ; Jannene Dalia

Extrait : **POUR** interagir les uns avec les autres, la communication joue un rôle vital dans la vie humaine. Les personnes ayant une difficulté spécifique , en particulier les personnes sourdes et muettes, ne peuvent pas communiquer avec les personnes normales. C'est pour cela ,les langages des signes (gestes spéciaux) sont inventés pour permettre à ces personnes de se communiquer. Différentes méthodes sont utilisées pour reconnaître le signe (mouvement des gestes) et le convertir en texte et en voix. Pour ce document , un examen est fait des techniques existantes. Un système de reconnaissance des gestes et de la voix ,basé sur LabView a été développé pour identifier le geste de la main et pour convertir la voix en signe afin d'aider les sourds et les muets à interagir plus facilement avec un ordinateur et avec autrui .
mots clés : Reconnaissance Optique des Caractère(OCR), Traitement d'image,Langage de Signe,LABView,virtual instrument

Date de présentation : 29/03/2020

Note : 17.5 , Mention très bien

I. Extrait :

Depuis la révolution tunisienne, les sourds et les muets font preuve qu'ils sont capables d'acquérir ce que les personnes normaux peuvent faire. Mais leurs difficultés spécifiques restent intacte et cette différence dresse des barrières qui empêchent une communication efficace entre eux.

Selon les données et les statistiques sur les handicapés en Tunisie, 465 personnes bénéficiaires de cartes de handicapés inscrits auprès du ministère des affaires sociales et la répartition de la prévalence des handicaps se présente comme suit: oculaire : 10,7% et auditif 11.8%.

Comme les gestes sont des mouvements corporels expressifs et significatifs pour la communication, les personnes handicapées utilisent différentes formes de mains et de mouvements pour transmettre le message. La reconnaissance des gestes est l'interprétation mathématique d'un mouvement humain par des dispositifs informatiques. La reconnaissance des gestes de la main est une méthode dans laquelle les gestes produits par l'utilisateur sont décodés par le système. Ce système aide les personnes muettes et sourdes à entrer leurs données en utilisant les gestes de la main et les données entrées sont converties en l'alphabet et la voix respectif . Le système de reconnaissance des gestes de la main fournit une solution réalisable pour contrôler diverses applications domestiques, industrielles et biomédicales. Récemment, le développement des technologies d'assistance pour les personnes handicapées s'est considérablement accéléré, ce qui a permis d'améliorer les systèmes traditionnels. De même, l'utilisation croissante de l'ordinateur dans le travail et les loisirs a conduit au développement d'applications de manipulation associées au PC, utilisant principalement des interfaces graphiques. L'augmentation des performances des ordinateurs personnels et la réduction de leur coût ont rendu les systèmes basés sur le PC réalisables. Il faut une technologie qui convertit le langage des signes en alphabet et en voix pour que les sourds et les muets puissent communiquer avec des personnes normales.

II. Introduction :

De nos jours, le traitement d'images fait partie des technologies en plein essor. Il constitue également un domaine de recherche essentiel pour les disciplines de l'ingénierie et de la technologie. Nous avons développé et mis en œuvre de manière innovante des applications de traitement d'images, principalement basées sur la boîte à outils de vision LabVIEW et IMAQ.

La boîte à outils de vision IMAQ présente un ensemble complet de fonctions de traitement et d'acquisition d'images numériques qui améliorent l'efficacité des projets et réduisent l'effort de programmation des

utilisateurs en obtenant de meilleurs résultats dans des délais plus courts. Par conséquent, la boîte à outils IMAQ de LabVIEW est un outil intéressant à analyser en détail et à travers ce chapitre, il sera présenté différentes théories sur le traitement numérique des images et différentes applications dans le domaine de l'acquisition et de la transformation d'images.

III. Littéraire :

Le système de la connaissance des geste de la main et la voix est dédié pour les sourds et les muets . Comme pour n'importe quel système de reconnaissance , la phase de la collection des données reste la plus importante partie du système. Pour ce fait , plusieurs technologies sont utilisés notons Vision based approach, and OCR approach :

1- Vision based approach :

L'image est un outil nécessaire pour l'interaction Homme-Machine sans avoir besoin de matériel supplémentaire. Il existe de nombreuses techniques utilisées pour la détection après quelques opérations de prétraitement d'image, ces méthodes peuvent être divisées en deux parties : les approches basées sur l'apparence et les approches basées sur un modèle 3D .

2- OCR approach :

La reconnaissance optique des caractères est la plus récente approche utilisée pour la vision de l'ordinateur . Elle permet la détection des caractères selon plusieurs technologies de traitement d'image puis la conversion en audio.

IV. Présentation LabView :

Ni Labview est un environnement de développement professionnel pour la programmation graphique dans un langage dit "G". Ce logiciel est utilisé dans le monde entier pour créer des applications de test de matériel et pour contrôler précisément le flux de données lorsqu'il est crucial. Il est également populaire dans les systèmes d'exploitation en temps réel. Comme les outils de base fournis par l'environnement ne suffisent peut-être pas pour un usage professionnel, Ni Labview apporte également un ensemble de bibliothèques supplémentaires pour les calculs avancés. Pour les applications de vision artificielle et d'imagerie scientifique, la National Instruments Corporation a développé une bibliothèque de vision Imaq (également appelée "Drivers for A Hardware"), qui fait partie du module de développement de la vision. Pour faciliter le traitement des images, le logiciel d'acquisition d'images Ni et les pilotes mentionnés ont été utilisés. Dans Labview, une interface utilisateur est construite avec un ensemble d'outils et d'objets. L'interface utilisateur est connue sous le nom de "panneau avant". Le code est utilisé pour la représentation graphique des fonctions de contrôle des objets du panneau avant. Le diagramme de blocs contient ce code. Le bloc-diagramme ressemble à l'organigramme. Le programme Labview est appelé "instruments virtuels" ou "Vis" parce que leur apparence et leur fonctionnement imitent ceux des instruments physiques, tels que les oscilloscopes et les multimètres, etc.

Le Vi contient les trois éléments suivants :

I) Le panneau avant sert d'interface utilisateur

II) Le schéma fonctionnel qui contient le code source graphique définissant la fonctionnalité du Vi

III) L'icône et le panneau de connexion qui identifient le Vi afin qu'il puisse être utilisé dans un autre Vi.

Un Vi dans un autre Vi est appelé Sub Vi. Un sous-vi correspond à un sous-programme en langage de programmation textuelle.

V. Processus :

Le projet dispose de 2 grande parties qui a fait recourt à deux fonctionnalités que le logiciel Labview présente: une concerne la vision "vision and motion" et l'autre "Connectivity" pour configurer le son.

Traitement d'image:

dédié pour ceux qui utilisent la langue des signes pour communiquer. Le code décrypte le message émis par l'interlocuteur et affiche ce qu'il veut dire sur l'écran afin de permettre au récepteur de bien comprendre l'intention.

Reconnaissance Optique des Caractères (OCR):

dédié pour les non voyants et ceux qui ont un dysfonctionnement visuel. Le code lit ce qu'en lui présente sous forme d'écriture sur papier : lettre, journal, note... et le génère sous forme de son.

Les objectifs fonctionnels sont présentés dans le graph suivant :

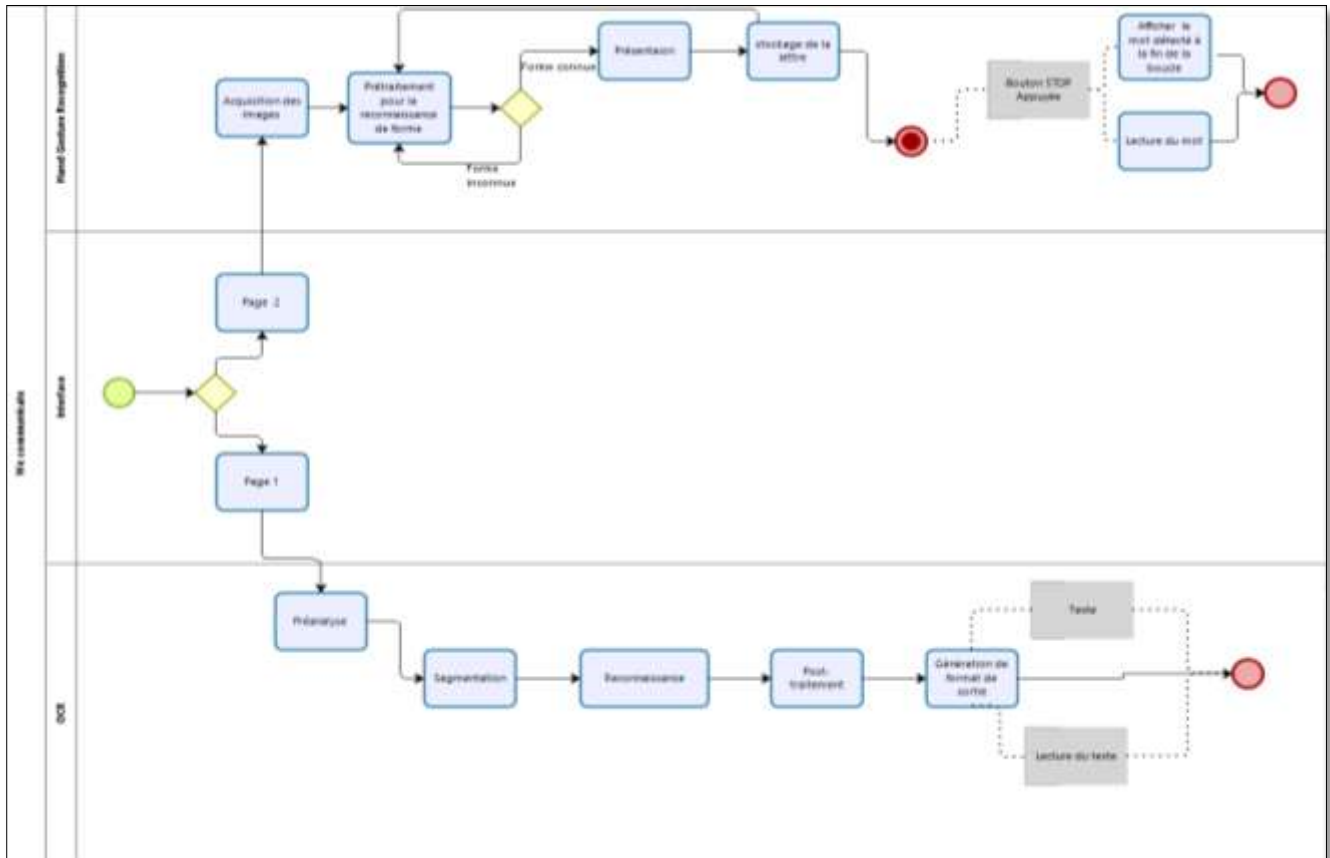


Fig1 : le processus du traitement de système

VI. La reconnaissance optique des caractère (OCR) :

Cette technique consiste à afficher une image numérique contenant du texte au caméra de l'ordinateur et à partir de l'image montrée, il sera capable de transformer ces pixels à des lettres. Généralement les chercheurs classe la reconnaissance optique des caractère en 5 étapes après l'acquisition de l'image de la caméra. préanalyse de l'image, Segmentation, Reconnaissance, Post traitement, Génération du format de sortie comme illustré sur le diagramme de flux suivant :

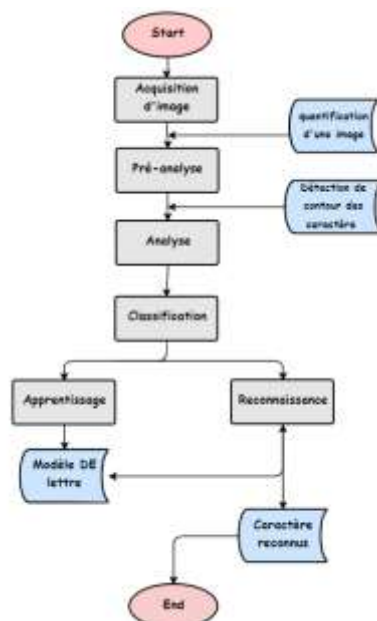


Fig2 : diagramme de flux de l'OCR

La 1ère étape :

Préanalyse de l'image: le but étant est d'améliorer la qualité de l'image : redressement d'images inclinées ou déformées, corrections de contraste, passage en noir et blanc, détection de contours...

La 2ème étape :

Segmentation ou analyse : permet d'isoler dans l'image les lignes de texte et les caractères. Cette phase peut aussi détecter le texte souligné, les cadres, les images. Une sorte de "listing" des éléments de votre fichier initial .

La 3 -ème étape :

Reconnaissance : les caractères sont comparés à une forme connue qui s'approche au plus possible des caractères selon une distance ou une vraisemblance appelée "likelihood". Le mot sera cadré en vert .

La 4ème étape :

Post traitement : permet l' utilisation de dictionnaires linguistiques et syntaxiques pour réduire le nombre d'erreurs de reconnaissance : systèmes à base de règles, ou méthodes statistiques basées sur des dictionnaires de mots, de syllabes, de N-grammes (séquences de caractères ou de mots).

La 5ème étape :

génération du format de sortie : le texte a été reconnu, il s'agit maintenant de le générer dans un format permettant à l'utilisateur de l'utiliser. Le format .txt pour du texte brut qui ne tiendra pas compte de la mise en page, ou alors le format (.doc) ou (.pdf) pour des technologies plus développées tenant compte de la mise en page, des cadres, des paragraphes ...

finallement, le mot sera affiché.

VII. Implémentation LabView :

Pour la partie OCR , on a utilisé le module “Connectivity” pour générer le son .

Plusieurs fonctions figurent dans ce module qui permette la réalisation du code complet.

Pour l'acquisition de l'image : IMAQ Grab VI configure qui va assurer l'acquisition de l'image du caméra du webcam.

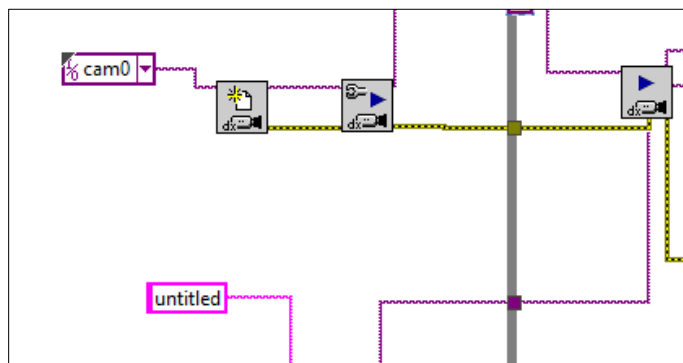


fig3 : l'acquisition de l'image

Après le traitement décrit dans la partie précédente , la détection se fait et on aura une lecture audio du mot lit .

dans notre cas , on a donnée un exemple de 4 mots seulement : 'AIDEZMOI', 'BONJOUR', 'SALUT', 'MERC'

les fonctions utilisés sont illustré dans le code suivant :

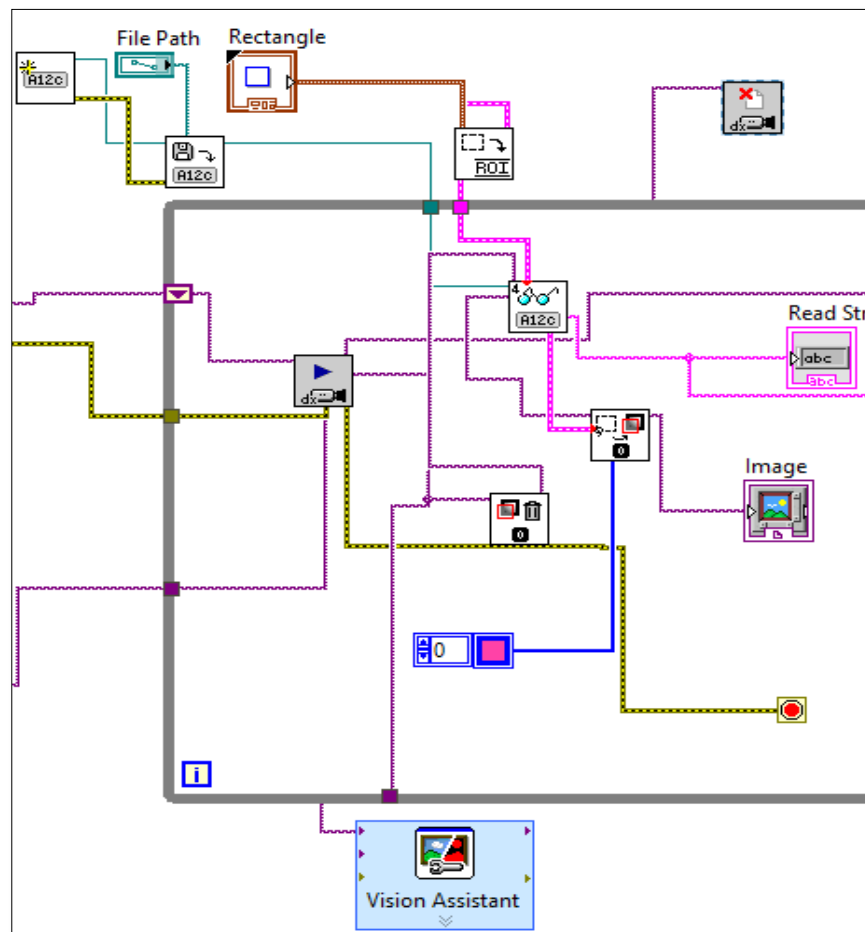


Fig4 : traitement OCR

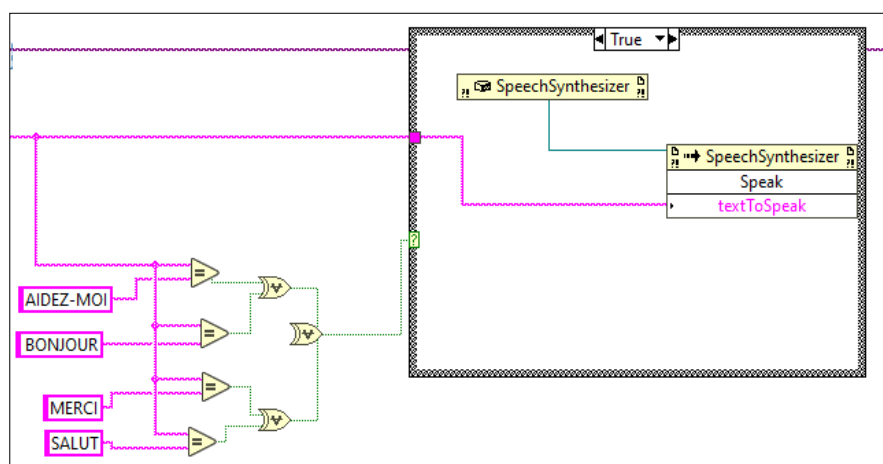


fig5 : La présentation audio du mot

La partie graphique est présentée dans le front panel comme ci-dessous. Il est à noter que la partie graphique représente l'interface utilisateur GUI qui montre le résultat.

Le rectangle vert représente le mot à détecter et qui sera par la suite écrit dans la partie 'Read String'.

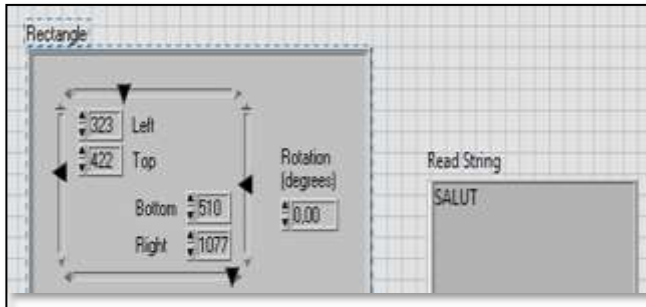


fig6 : dimension du figure

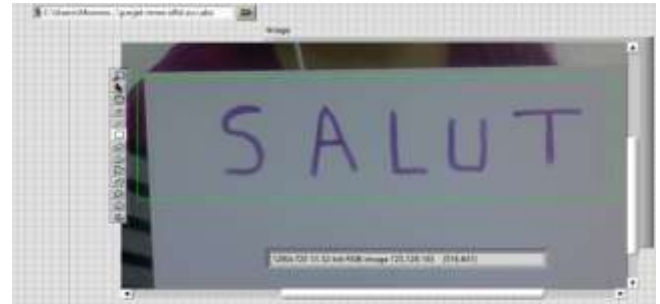


fig7 : le mot lu à partir de la caméra

VIII. Reconnaissance des gestes de la main :

Le processus de reconnaissance des gestes est principalement divisé en trois parties successives:

- l'acquisition des images
- le prétraitement avec la correspondance des formes
- la présentation des résultats

L'algorithme utilisé pour le prétraitement est défini en 4 étapes :

a) Extraction d'un plan unique d'une image en couleur .

b) Érosion : Réduit la luminosité des pixels qui sont entourés par des voisins avec une intensité plus faible. Il est utile pour éliminer les petits bruits blancs.

c) Dilatation: Augmente la luminosité des pixels entourés par des voisins avec une intensité plus élevée. (Une dilatation a l'effet inverse d'une érosion: Il augmente la zone blanche dans l'image ou la taille de l'objet au premier.)

d) test.

Sur l'image traitée, on applique l'algorithme de détection des formes (Grayscale Value Pyramid).

Lorsqu'un geste est détecté, on superpose une boîte de délimitation autour de celui-ci et on visualise la correspondance score à côté.

IX. Implémentation LabVIEW :

Chaque partie sera discutée dans cette section. Pour ceci on a utilisé la fonctionnalité 'vision and motion' qui nécessite l'installation de deux modules supplémentaires: **vision acquisition software** et **vision development module** . Ce module contient toutes les fonctions nécessaires pour faire la conversion de l'image en langage de signe compréhensible par les muets .

Pour l'acquisition de l'image , on doit tout d'abord configurer les images par donner leurs chemins spécifique , puis l'ouverture du caméra permettra d'acquérir l'image à comparer .

Le code de cette partie est illustré comme suit :

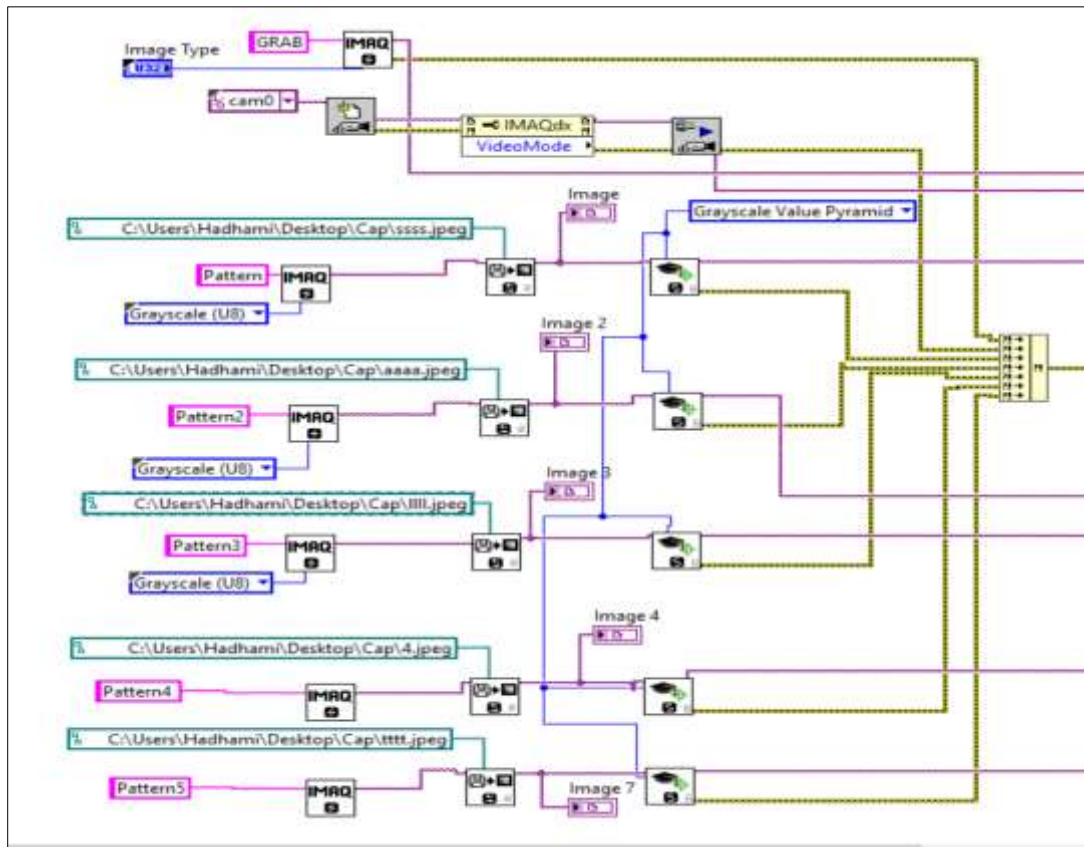


fig8 : acquisition de l'image pour la reconnaissance des gestes de le main

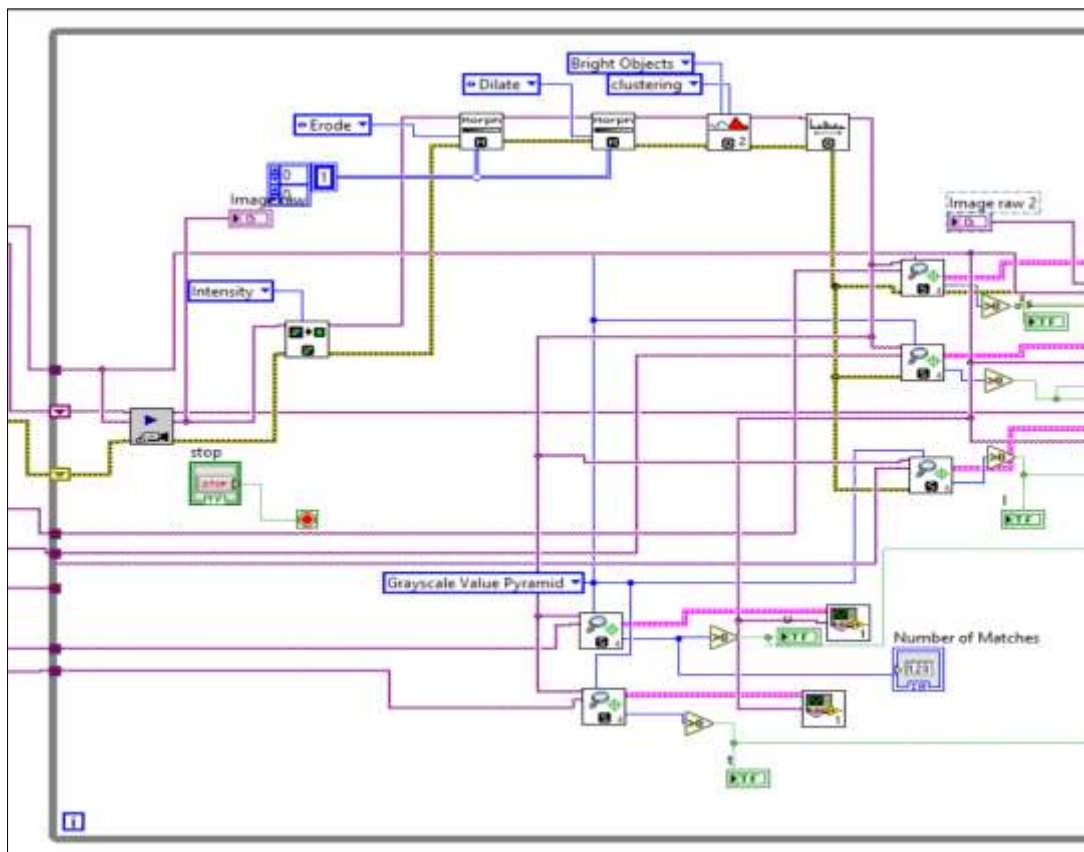


fig9 : traitement de la reconnaissance des gestes

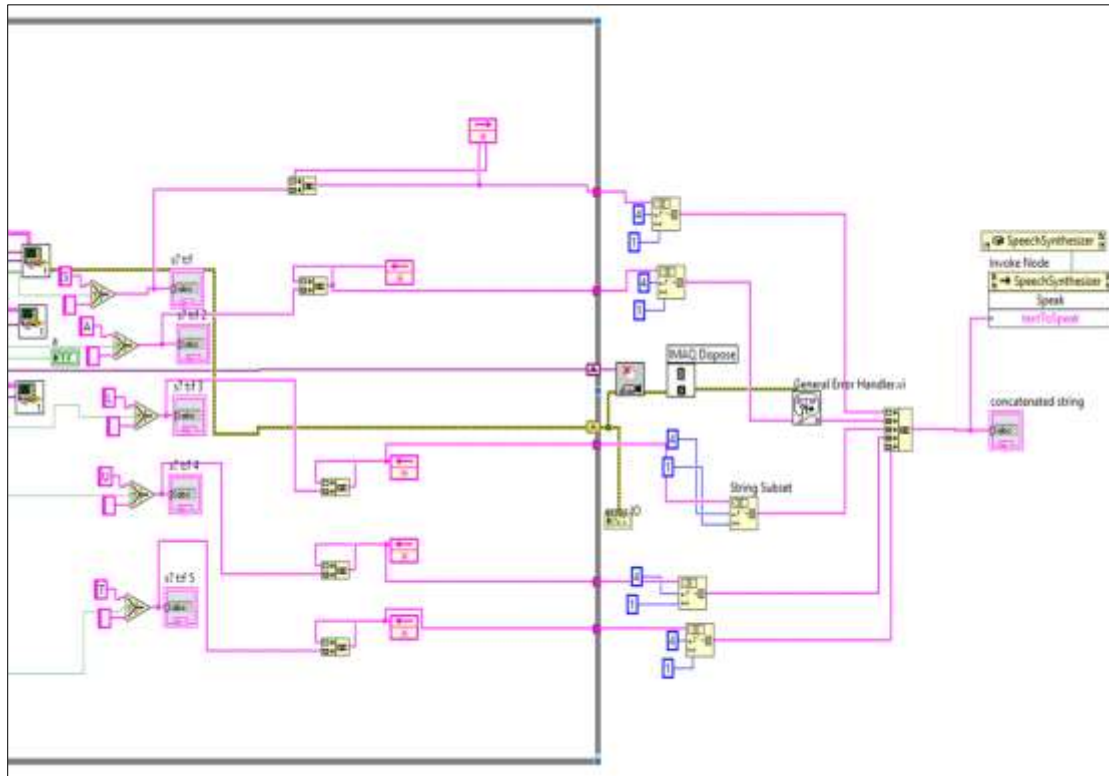


fig10 : la présentation

Pour la partie présentation , un SUBVI est utilisé sous le nom de hi.vi :

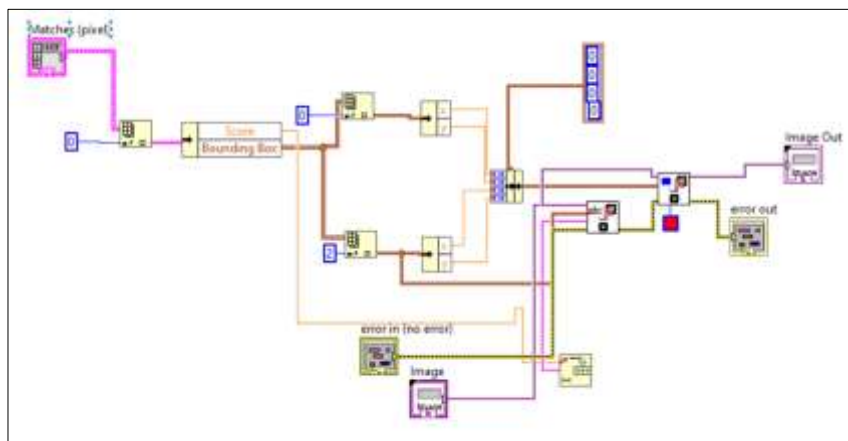


fig11 : le subVi hi.vi

Ce diagramme est un SubVI utilisé pour superposer un rectangle sur un motif détecté et lui attribuer un score correspondant .Le rectangle et la boîte de score suivent un geste lorsqu'il se traduit en coordonnées x et y. Les coordonnées sont extraites de la concordance des modèles VI . Lorsqu'un compteur de détection de motif est supérieur à 0, il fixe la valeur booléenne de l'indicateur de modèle correspondant à 1 (la couleur de l'indicateur passe du jaune au vert). L'image brute et l'image traitée sont toutes deux affichées. En cas d'erreur, la boucle est terminée et le groupe d'erreurs de sortie (état d'erreur, code et source) est présenté à l'utilisateur.

Pour l'interface utilisateur , où la première personne peut voir le résultat , on a just fait capture de 2 tests de signes . Voilà le résultat :

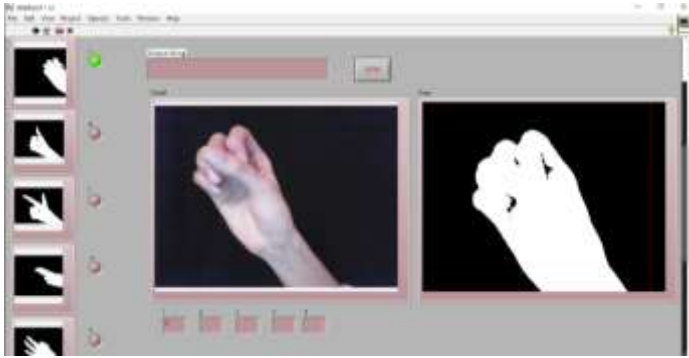


fig12 : lecture de la lettre S

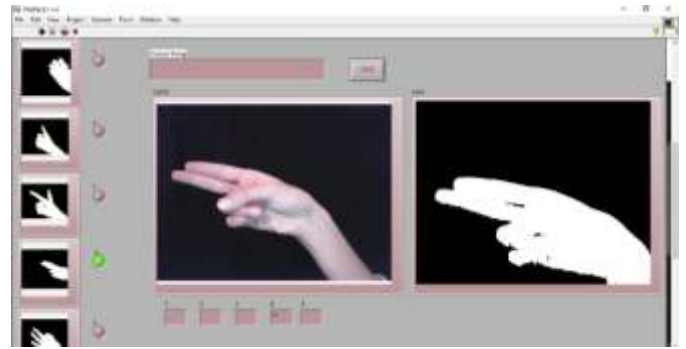


fig13 : lecture de la lettre A

Cette interface est un peu compliqué et présente plusieurs parties :

*Gauche :

Modèles de gestes et indicateurs de reconnaissance: Le modèle de chaque geste est affiché afin d'informer un utilisateur sur des modèles correspondants . En outre, les indicateurs à diode prennent l'un des deux états de couleur possibles :

- jaune : le geste n'a pas été détecté.
- vert : le geste a été détecté.

*Milieu :

Affichage des images brutes et traitées:

- Cam0: affiche une image brute, non traitée provenant de la caméra intégré au PC
- Cam: affiche une image traitée c prête à être recherchée parmi le modèle.

Si un modèle est trouvé, le programme dessine une boîte dite "correspondance de modèles", avec une valeur de score qui représente l'exactitude de la correspondance (cela se produit lorsqu'un dépasse un certain seuil).

*Droite :

Paramètres d'acquisition et un indicateur d'erreur: l'utilisateur peut démarrer/arrêter l'acquisition d'images, sélectionner le mode seuil ou le type d'image saisie. De plus, le champ "error out" indique si une erreur survenue lors du calcul de l'algorithme ou non.

X. Le système Global :

Pour une meilleure lisibilité , une interface qui regroupe les 2 parties est présenté d'une manière organisé où chaque partie est présenté dans une page comme illustre la figure suivante :



fig13 : l'interface global du système

XI. Conclusion :

En guise de conclusion , nous avons exploité le logiciel **LabVIEW** hors de ses dimensions communes pour servir un besoin social humanitaire visant l'inclusion des personnes en difficulté de communication dans les différents domaines de la vie, y compris leur inclusion en industrie. Ce logiciel fournit différents modules facilitant l'acquisition et l'extraction d'information à partir d'images et vidéos en temps réel. Nous avons réussi à distinguer 5 gestes de l'alphabet de la langue des signes par le programme et de décrypter plusieurs mots comme ' salut' et de le prononcer.

Même si nous avons éliminé le problème de l'éclairage et de l'arrière-plan en plaçant un noir dans la deuxième partie, le problème des différentes formes des mains reste irrésolu . Pour ceci, il faut appliquer une correspondance géométrique des formes. Cet algorithme permet de faire correspondre un motif à une image mise à l'échelle. Néanmoins, son implication nécessiterait une du temps supplémentaire pour le développement du programme.