



جامعة أبو بكر بلقايد - تلمسان

Université Abou Bakr Belkaïd de Tlemcen

Faculté de Technologie

Département de Génie Biomédical

Laboratoire de Recherche de Génie Biomédical

MEMOIRE DE PROJET DE FIN D'ETUDES

Pour obtenir le Diplôme de

MASTER en Génie Biomédical

Spécialité : Informatique Biomédicale

Présenté par : **BENYAMINA Zohra et HARTANI Abdelhakim**

**Réalisation d'une application de reconnaissance
d'alphabet des sourds muets**

Soutenu le 14 juin 2015 devant le Jury

Mme	BENCHAIB Yasmine	MCB	Université de Tlemcen	Président
M.	AMMAR Mohammed	MCB	Université boumardess	Encadreur
Melle	FEROUI Amel	MCB	Université de Tlemcen	Examineur

Année universitaire 2014-2015

Remerciements :

Nous remercions Allah le tout puissant de nous avoir donné le courage et la volonté de mener un terme à ce présent travail.

Nous tenons tout d'abord, d'adresser nos sincères remerciements et notre profonde gratitude à notre Encadreur Monsieur Ammar Mohammed pour nous accepter de nous encadrés, pour l'aide qu'il nous a apportée et pour la confiance qu'il nous a accordés.

Nous voudrions également adresser nos plus vifs remerciements à chacun des membres du jury : Mme BENCHAIB Yasmine, pour nos avoir fait l'honneur de présider nos jurys de mémoire, Melle FEROU Amel, pour avoir accepté la charge d'examiner cette mémoire. Merci pour l'intérêt qu'ils y ont porté et pour leurs remarques pertinentes.

Nous remercierons l'ensemble des enseignant du département de génie biomédical pour les efforts qu'ils fournissent pour notre réussite dans nos études plus particulièrement le professeur Mr Chikh M.A, Mr Abdrahim, Mr Behadada, Monsieur Karaai.....

Nous remercierons également nos camarades avec qui nous avons partagés de bons moments pendant ces cinq années et pour l'aide et pour l'ambiance qui ont régnés.

Nous tenons à remercier tous ceux qui nous ont aidés d'une manière ou d'une autre à élaborer ce modeste travail.

Dédicace :

Mes pensées vont vers ma grand-mère Alhadja qui m'a toujours donnés amour, tendresse et affection. Ta prière et ta bénédiction m'ont été d'un grand secours pour mener à bien mes études. Puisse Dieu, le tout puissant, te préserver et t'accorder santé, longue vie et bonheur.

A mes très chers parents, la source de mes efforts, de joie et de bonheur, et mon soutien moral et matériel qui s'ont toujours sacrifié pour me voir réussir. Que dieu vous préserve une longue vie heureuse.

A mon frère Abdelkader qui m'est toujours aidé, et qui m'est accompagnaient durant mon chemin d'études.

A ma sœur Sara qui était toujours à mes côtés, merci pour tes conseils, aides, et encouragements.

A ma sœur Bakhta et son marie Abdelhak et sa petite princesse hindo que j'adore.

A ma nièce Hafsa, ma compagnon fidèle dans tous les moments les plus délicats de cette vie mystérieuse.

Aux personnes dont j'ai bien aimé la présence dans ce jour, à tous mes frères Mohamed, Hamid, Billal, Walid et Yasser et mes sœurs Houria, Nacira, Aicha, Cherifa, mes belles sœurs Hafidha, Rabia, Samia, Fatma, Soumia et Amel, mes bons frères et mes nièces Hafsa,....et Amir.

A toute ma grande famille, Tous ceux que j'aime, qui m'aiment.

Je dédie également ce travail à tous ceux qui m'ont apporté leur savoir et contribué à ma formation. Qu'ils trouvent ici l'expression de ma gratitude et de mon profond respect.

A mes aimables amis, collègues d'étude, et sœurs de cœur, toi Farida, Halima, Chifaa, Zahra, Zineb, Keltoum, Samiha, Hanane et Houda. Être votre amie est un plaisir et une joie énergisante.

A tous ceux qui, un jour, ont pensé à moi, les plus beaux mots ne sauraient exprimer mes sentiments et mon profond respect envers vous, que dieu me vous gardez dans ma vie, nchallah.

Je tiens énormément à remercier mon binôme Hartani Abdelhakim, je te souhaite un avenir plein de bonheur, de prospérité et de réussite, merci.

A tous ceux que j'aime, Je dédie le fruit de mon projet de fin d'études.

Zahira 

Dédicaces

A ma mère,

A la plus merveilleuse des mères. J'espère réaliser, en ce jour, l'un de tes rêves. Aucun mot ne saurait exprimer mon respect, ma considération et l'amour que je te porte. Ta présence constante à mes côtés, tes encouragements et tes prières m'ont été d'une aide précieuse et m'ont permis d'atteindre le but désiré. Puisse Dieu le tout puissant te donner santé et longue vie afin que je puisse te combler à mon tour...

A mon père,

Aucune dédicace ne saurait exprimer à sa juste valeur tout l'amour, le respect, l'attachement et la reconnaissance que je te porte. Tu m'as enseigné la droiture, le respect et la conscience du devoir. Ce travail est le fruit de tous tes sacrifices, tes encouragements, ton désir de me voir arriver et ton soutien permanent durant ce long parcours. Puisse Dieu, le tout puissant, te procurer santé, bonheur et longue vie...

J'espère

Avoir répondu aux espoirs que vous avez fondés en moi. Je vous rends hommage ici par ce modeste travail en guise de ma reconnaissance éternelle à jamais... Que Dieu, le Tout Puissant vous garde et vous procure santé, bonheur et longue vie.

Toute ma famille,

Grands et petits qui attendent ce travail avec impatience et surtout mes frères amis abdelouaheb sid ahemd sarah aicha ,mon binôme benyamina zahira Trouvez ici l'expression de ma profonde affection et mon respect.

Mes professeurs,

Qui m'ont enseigné depuis ma première scolarité. Trouvez ici l'expression de mes respects et mon éternelle reconnaissance. A tous mes amis surtout djamel otmane krimo, collègues et toute la promotion master en génie biomédicale Que tous vos rêves soient exaucés et la réussite comble votre vie.

M .Hartani Abdelhakim

Résumé

Ce projet de fin d'étude a pour objectif de réaliser une application de reconnaissance d'alphabet des sourds muets en utilisant java script Android.

Pour la réalisation de notre application nous avons utilisé l'algorithme Sift qui se base sur le calcul des points d'intérêt qui sont invariants à plusieurs transformations au changement d'intensité, de mise à l'échelle et de rotation, ce qui fait de lui un descripteur très robuste.

Notre application comporte deux étapes principales : une étape qui permet d'extraire les caractéristiques d'une image et de calculer ses descripteurs et une étape consiste à mettre en place une procédure de mise en correspondance.

Mots clés : sourds muets, Android, Sift, points d'intérêt, descripteurs, mise en correspondance.

TABLE DES MATIERES

Remerciements.....	i
Dédicace.....	ii
Table des matières.....	iv
Table des figures.....	vii
liste de tableaux.....	iv
Glossaire.....	xi
Introduction générale.....	11

CHAPITRE I. Les sourds muets

1. Introduction.....	13
2. Qui sont les sourds muets ?	13
3. Comment avoir une surdité ?.....	14
4. Histoire de la langue des signes	14
5. La langue des signes	15
6. Les différences entre la langue des signes et la langue orale	15
7. Quelques caractéristiques de la Langue des Signes	16
8. Composition d'un signe.....	16
8.1. La configuration	16
8.2. Le mouvement.....	17
8.3. L'emplacement.....	18
8.4. L'orientation.....	18
8.5. La mimique faciale.....	19
9. Différentes classes de signes.....	19
10. Interaction entre les deux mains	20
11. Enrichissement de la langue des signes	21
12. La Langue des Signes Française (LSF)	21
13. L'alphabet de la langue des signes française	22
14. La langue des signes est une langue universelle	23
15. Conseils et attitudes	24
16. Conclusion	25

CHAPITRE II. Les méthodes de reconnaissance des alphabets

1.	Introduction.....	27
2.	La reconnaissance des gestes.....	27
2.1.	Type de reconnaissance de gestes	27
2.1.1.	La reconnaissance statique	28
2.1.2.	La reconnaissance dynamique.....	28
2.2.	Définition des gestes	28
2.3.	Les fonctions du geste de la main.....	28
2.3.1.	La fonction ergotique	28
2.3.2.	La fonction épistémique	29
2.3.3.	La fonction sémiotique	29
2.4.	Les gestes de la main.....	29
2.5.	Le problème de la reconnaissance des gestes	29
3.	Les méthodes de reconnaissance des gestes.....	30
3.1.	La reconnaissance en utilisant des points caractéristiques spatio-temporelle (STIP)	30
3.2.	Reconnaissance de gestes dynamiques de la main	30
3.3.	La reconnaissance de gestes en utilisant un système multi-caméras.....	31
3.4.	La reconnaissance en utilisant les gants	32
3.5.	La reconnaissance de la langue des signes en utilisant des sous-unités linguistiques.....	32
3.6.	Reconnaissance des gestes par le suivi 3D de la main	33
3.7.	Le système de traduction de langue des singes en utilisant Kinect	34
4.	Conclusion	35

CHAPITRE III. Sift et résultats et implémentations

Partie I :SIFT.....		36
1.	Introduction	36
2.	Définition	36
3.	L'algorithme SIFT	36
3.1.	Calcul des points d'intérêts et des descripteurs.....	37
3.1.1.	Construction de l'espace des échelles (la pyramide)	37

3.1.2.	Détection des extrema locaux dans les DoG	38
3.1.3.	Localisation précise des points d'intérêt et élimination des points non pertinents	39
3.1.4.	Affectation d'orientation aux points d'intérêt	39
3.1.5.	Calcul des descripteurs	40
3.2.	Mise en correspondance	40
4.	Les techniques utilisées dans la recherche de correspondance.....	41
5.	Pour Quoi SIFT ?	41
Partie II : Implémentation et résultats		43
1.	Introduction	43
2.	La conception du système	44
2.1.	Création de la base des signes	44
2.2.	Reconnaissance de l'alphabet.....	44
3.	Acquisition d'image	44
4.	Calcul des points d'intérêts et des descripteurs	45
5.	Base de données	45
6.	Environnement de développement	46
6.1.	Android.....	46
6.2.	Architecture matérielle	46
7.	Présentation des interfaces de notre système.....	46
7.1.	L'interface principale	46
7.2.	Fenêtre des paramètres	47
7.3.	Fenêtre des singes	48
7.4.	Fenêtre de la caméra	49
7.5.	Fenêtre d'explorer	50
8.	Expérimentation et résultats	52
9.	Conclusion	56

Table des figures

Figure I.1: Le signe [boire].....	17
Figure I.2: les configurations "c", "boule" et "pince" , applicable au verbe [boire].....	17
Figure I.3: le signe [voir].....	17
Figure I.4: le signe [opérer] effectué sur le buste.....	18
Figure I.5: Le signe [je t'envoie].....	18
Figure I.6: le signe [impossible] à gauche et [détester] à droite.....	19
Figure I.7: La phrase [une caisse pleine de pomme].....	20
Figure I.8: Les signes [TOMBER DE] et [TABLE].....	21
Figure I.9: Exemple d'une phrase en LSF composé des signes "je donne un gâteau au garçon qui est à ma droite" [Garçon] "index vertical ", [gâteau] et [donner].....	22
Figure I.10: Alphabets de la langue des signes.....	23
Figure II.1: Geste Statique.....	28
Figure II.2: Geste dynamique.....	28
Figure II.3: Exemple de détection des STIP sur la séquence de geste « Salut ».....	30
Figure II.4: (a) Image binaire, (b) Image de distances, (c) le squelette, (d) Squelette.....	31
Figure II.5: Construction de l'histogramme de gradients orientés.....	31
Figure II.6: Exemples des gants : (a) le 5DT Data Glove 5 Ultra, et (b) le CyberGlove.....	32
Figure II.7: Aperçu sur les 3 types de sous-unités et les 2 différences entre les 2 niveaux de classifieur.....	33
Figure II.8: Suivi 3D de la main en fonction du temps.....	34
Figure II.9. : Les fonctions principales du système.....	34
Figure II.10: Les modes (a) et (b) du système.....	35
Figure III.1: Étapes de calcul des points d'intérêts et des descripteurs	37
Figure III.2: Une pyramide de gaussienne.....	38
Figure III.3 : Structure de détection des extrema.....	38

Figure III.4: Détection d'extrema locaux.....	39
Figure III.5: Construction de l'histogramme des orientations.....	39
Figure III.6: Construction d'un descripteur SIFT ; vecteur de taille 128 (8x16).....	40
Figure III.7. Schéma global du système.....	43
Figure.III.8. Interface d'acquisition.....	44
Figure.III.9. trois images d'alphabet A.....	45
Figure III.10. Interface principale.....	47
Figure. III.11. Fenêtre des paramètres.....	47
Figure. III.12. fenêtre des Singes.....	49
Figure .III.13. Fenêtre du camera	50
Figure .III.14. Fenêtres Explorer.....	51
Figure III.15. Graphe de résultat de matching entre les alphabets.....	53
Figure III.16. Graphe de résultat de matching entre les alphabets.....	54
Figure III.17. Graphe de résultat de matching entre trois images de même alphabet.....	55

liste des tableaux

Tableau III.1: techniques utilisées dans la recherche de correspondance.....	41
Tableau III.2: étude comparative entre SIFT, SURF et PCA-SIFT.....	42
Tableau III.3: Résultats des modifications des paramètres.....	52
Tableau III.4: Résultat de matching entre les alphabets.....	53
Tableau III.5: Résultat de matching entre les alphabets.....	54
Tableau III.6: Résultat de matching entre trois cas du même alphabet.....	55

Glossaire

LS : Langue des Signes.

LSF : langue des signes français.

ASL : langue des signes Américain.

STIP : Points d'intérêt spatio-temporel.

SVM : Support Vector Machine.

HOG : Histogramme de gradient orienté.

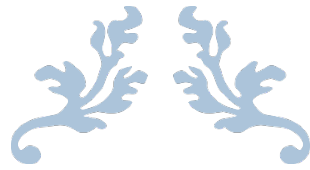
DTW : Déformation Temporelle Dynamique.

K-NN : plus proches voisins.

SIFT : Scale Invariant Feature Transform (transformation de caractéristiques invariante à l'échelle).

PI : Points d'intérêts.

DOG : différence de Gaussien.



INTRODUCTION GENERALE



Introduction générale :

Au cours de notre vie, que nous entendant ou non, nous utilisons des signes pour transmettre notre message à l'autre. Par exemple, un bébé apprend vite à pointer le doigt vers l'objet qu'il désire. Souvent, lorsque nous ne parlons pas la langue d'une personne étrangère, nous utilisons des signes pour essayer de le faire comprendre.

La Langue des Signes est le moyen d'expression utilisé par les communautés des sourds ou malentendants pour communiquer entre eux et constitue de ce fait la forme la plus évoluée en communication gestuelle.

Le système de reconnaissance des gestes de la langue des signes est l'une des recherches les plus en développement aujourd'hui. Des nouvelles techniques ont été développées récemment.

Dans ce mémoire, nous proposons une application pour la reconnaissance des alphabets des sourds muets basée sur l'algorithme SIFT.

Le reste du manuscrit sera organisé comme suit :

Dans le premier chapitre, nous présentons la communauté des sourds muets et la langue des signes et ses caractéristiques, la différence entre elle et la langue orale, les compositions d'un signe et ses différentes classes gestuel par la main, ses fonctionnalités, ses types ainsi quelque conseils et attitude pour faciliter la communication entre un sourd et un entendant.

Dans le deuxième chapitre nous avons abordé la reconnaissance des gestes et ses types, les gestes de la main et ses fonctions ainsi les problèmes rencontrés et les différentes méthodes de reconnaissance des gestes proposées dans la littérature.

Dans le troisième chapitre, nous décrivons l'algorithme SIFT que nous avons utilisé dans notre application, ses paramètres, les étapes et l'architecture générale pour la réalisation de cette dernière.

En fin, nous terminons par une conclusion générale qui résume, l'essentiel de notre travail et bien sûr les perspectives envisagées pour l'amélioration de notre application.



Chapitre I



Sourds muets et la langue des signes

« Mon langage est aussi valable que le vôtre, plus valable même, parce que je peux vous communiquer en une image, une idée plus élaborée que vous pouvez le faire en cinquante mots ». Les Enfants du Silence (pièce de théâtre, 1986).

1. Introduction :

Le langage mimique des sourds muets constitue un mode de communication très particulier. Tandis que toutes les langues humaines, quelle que soit leur diversité pour le linguiste, sont orales, le langage mimique est uniquement visuel. Cela signifie non seulement qu'il est perçu sans aucun appel à l'ouïe, mais qu'il utilise une matière tout à fait originale (gestes, attitudes, mouvements...).

L'emploi des gestes est évidemment chose courante dans les relations interindividuelles. Mais dans le langage mimique des sourds muets il s'agit de gestes élaborés qui permettent des échanges complexes, ceux que requiert la communication entre individus adultes et civilisés. Il est donc loin, bien que son but soit essentiellement pratique, d'une gesticulation élémentaire, pragmatique ou affective. [1]

La Langue des Signes est le moyen d'expression utilisé par les communautés de sourds ou malentendants pour communiquer entre eux. La Langue des Signes est une véritable langue à part entière (avec un lexique, une syntaxe...) et constitue de ce fait la forme la plus évoluée en communication gestuelle. [2]

2. Qui sont les sourds muets ?

D'après la définition du dictionnaire un sourd est qui ne perçoit pas ou qui perçoit difficilement les sons et un muet est qui n'a pas ou plus l'usage de la parole. En effet un sourd-muet est personne privée de l'ouïe et de la parole. [3]

Au niveau physique, les organes qui sont touchés sont bien distincts et en aucun cas liés.

- La surdit  implique une perte partielle ou totale de l'audition et touche donc l'oreille (externe, moyenne et/ ou interne).
- La mutit , d'un point de vue m dical, d signe l'absence de cordes vocales ou l'incapacit  de la personne   les utiliser.

Sauf des exceptions rare, les cordes vocales des personnes sourdes fonctionnent parfaitement, mais leur perception des sons d ficiante influence leur capacit    les utiliser,   vocaliser.

En fonction de leurs capacités audio-phonologiques, il leur est donc plus ou moins difficile de prononcer des mots puisqu'elles ne s'entendent pas.[4]

3. Comment avoir une surdité ?

La surdité est un état pathologique caractérisé par une perte partielle ou totale du sens de l'ouïe. La surdité dite de perception (définitive et incurable) peut être liée à :

- un problème génétique (héréditaire)
- une maladie in-utéro ou une maladie infantile
- un accident (traumatisme sonore ou crânien)

Ce déficit n'est pas forcément total. Un sourd peut parfois entendre les fréquences basses, il ressent les vibrations (musique, coup sur le sol, quand on tape dans les mains etc...). [5]

4. Histoire de la langue des signes :

Chaque culture a son histoire. Le monde des entendant, c'est la naissance de l'écriture, et, le monde des sourds, c'est la naissance de la langue des signes. Pour chacun d'eux c'est un évènement sacré.

Cette histoire elle rappelle que les sourds se sont battus et se battent encore pour faire respecter cette langue, leur culture, ses mythes et ses valeurs, et, continuent leur combat à ce jour.

Les premières tentatives connues est de Charles-Michel de l'Epée qui naît le 24 novembre 1712 à Versailles. Après des études de droit, il s'oriente vers l'église. Vers 1760, grâce à la rencontre de deux jumelles sourdes qui communiquent entre elles en langue des signes, il pense pouvoir instruire religieusement les sourds et leur apprendre à lire et à écrire. Il crée alors la première école gratuite permettant aux sourds de milieux défavorisés de recevoir une instruction.

Il est le premier à baser son enseignement sur des gestes qui viennent des sourds eux-mêmes. En organisant des séances publiques d'apprentissage, il montre au reste du monde que les sourds peuvent être instruits et considérés comme des Hommes, des citoyens, et que les signes peuvent être une langue et peuvent exprimer la pensée humaine autant qu'une langue orale. Il fait de l'instruction des sourds une vraie vocation, donnant tout ce qu'il a pour eux. Le

fait qu'il ait regroupé des élèves sourds a favorisé la communication entre eux, ce qui a permis le développement et le perfectionnement de la Langue des Signes Française (LSF). [6]

5. La langue des signes :

La langue des signes est une langue vivante et complexe. Comme toute langue, elle évolue. Elle possède une pleine capacité d'expression et d'abstraction. Elle permet donc la transmission de savoirs, d'où sa légitimité éducative. Elle possède sa propre syntaxe, qui est intimement liée à la perception visuelle, puisque cette langue répond à une logique visuelle et non auditive. La grammaire de la langue des signes n'est pas identique à celle de la langue vocale. Ainsi, la place des mots dans la phrase n'est pas la même.

Les signes sont basés sur l'utilisation des mains, du vue et de l'espace : les configurations des mains, leur orientation, leur emplacement et leur mouvement forment des signes qui équivalent à des mots, disposés devant soi comme sur une scène de théâtre. La disposition de ces signes, ainsi que la direction du regard, permettent de visualiser les relations (actif, passif...), le temps (signes tournés vers l'arrière pour le passé, vers l'avant pour le futur). Le visage et le mouvement des épaules servent aussi à exprimer les nuances du discours. [7]

6. Les différences entre la langue des signes et la langue orale :

Des données différencient radicalement les langues des signes des langues orales ; Une étude réalisée sur des signes isolés de l'ASL (American Sign Language) et des mots de l'américain parlé [8] a montré que :

- Pour un même concept, le temps d'émission moyen d'un signe isolé est approximativement deux fois plus long que celui d'un mot.
- À contenu et quantité d'information égale, le temps d'émission d'un discours en langue des signes et en américain parlé est approximativement le même.
- Le système auditif humain est adapté à la discrimination temporelle tandis que le système visuel est adapté à la discrimination spatiale [9].

- Le signal analysé par l'œil correspond directement un mouvement des articulateurs (les mains, les bras...), tandis que l'oreille n'analyse que l'effet sonore produit par le mouvement des articulateurs (les cordes vocales).

Selon Christian Cuxac [10] ces différences conditionnent le fonctionnement des langues des signes. Les niveaux syntaxiques et sémantiques sont structurés de façon à rattraper le temps perdu en ce qui concerne l'émission de signes isolés de tout contexte. Cela se fait au moyen d'une spatialisation des rapports syntaxiques ainsi que d'une utilisation fréquente d'informations simultanées.

7. Quelques caractéristiques de la Langue des Signes :

La Langue des Signes est une véritable langue, a un lexique, une syntaxe. L'expression de phrases en LS ne se réduit pas aux gestes produits par les deux mains, c'est le corps tout entier qui peut être mis à contribution pour exprimer une phrase. On peut distinguer trois principales parties qui interviennent : les mains, la tête et le buste.

Dans cette étude on se restreint uniquement aux mains, c'est pourquoi les autres parties du corps ne seront pas abordées par la suite. Chaque phrase est constituée d'une suite de gestes des mains que l'on appelle *signes* et qui sont agencés suivant une syntaxe régie par une logique spatiale et temporelle. [18]

8. Composition d'un signe :

Chaque geste d'une main peut être décomposé en cinq paramètres qui sont indépendants et peuvent être aussi bien dynamiques qu'invariants durant l'émission du signe

8.1. La configuration :

La configuration correspond à la forme de la main définie par les doigts et la paume (figure I.1). [2]

Par exemple selon la configuration utilisée, une action est appliquée à un instrument différent. La figure 1.1 illustre le signe [boire] générique. En modifiant uniquement la forme de

la main. On peut faire référence à des formes de récipient différentes. Par exemple on peut exprimer le fait de boire dans un verre (configuration "c " avec une seule main), un bol (configuration "boule" avec les deux mains), ou encore une tasse (configuration "pince" avec une seule main) (figure I.2)



Figure I.1. : Le signe [boire]



Figure I.2. : les configurations "c", "boule" et "pince" , applicable au verbe [boire].

8.2. Le mouvement :

Le mouvement de la main représente une action exercée par ou sur l'objet. De plus , la dynamique du mouvement permet de différencier les aspects du verbe. [10]



Figure I.3. : le signe [voir].

8.3. L'emplacement :

L'emplacement de la main représente l'endroit où l'action est effectuée ou un endroit générique. [10]

Par exemple, un verbe comme [opérer] possède un emplacement générique, mais peut être réalisé à d'autres endroits du corps, pour indiquer une opération spécifique : opérer du ventre, de l'œil, du cœur... (Figure I.4)



Figure I.4. : le signe [opérer] effectué sur le buste

8.4. L'orientation :

L'orientation permet de conjuguer certains verbes : ou de préciser l'orientation d'objets. [10]

En langue des signes, il existe deux classes de verbes : les verbes directionnels et les verbes non directionnels. Les verbes directionnels se conjuguent dans l'espace.

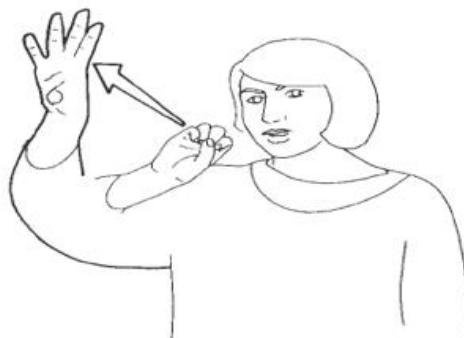


Figure I.5. : Le signe [je t'envoie].

8.5. La mimique faciale :

Permet d'exprimer le mode du discours (interrogatif, négatif...) ou joue le rôle des compléments de manière (plaisir, colère, envie, gêne...). Ces compléments de manière permettent aussi d'exprimer le point de vue du signeur sur le contenu de l'énoncé. [11]

Certain signes ne se différencient que par la mimique faciale, comme par exemple les signes [impossible] et [détester] (Figure I.6).

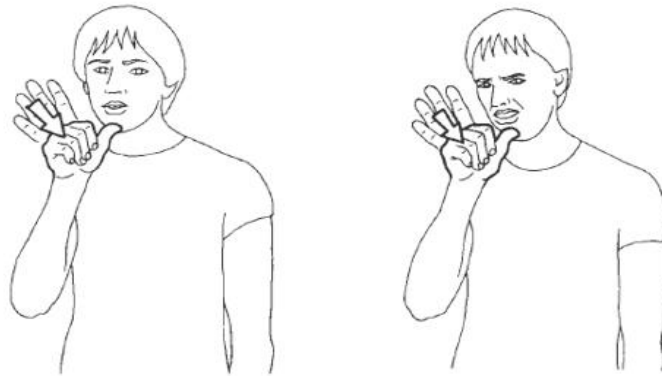


Figure I.6. : le signe [impossible] à gauche et [détester] à droite.

9. Différentes classes des signes :

En LS on peut distinguer différentes classes des signes, chacune correspondant à un usage particulier.

- L'ensemble des signes dits "*standards*" correspond à des mots (nom, verbe, adjectif...) ayant un sens bien établi. Ce sont les signes qui ont été le plus étudiés en reconnaissance des gestes et il existe des systèmes permettant d'effectuer la reconnaissance de manière fiable sur des vocabulaires de taille limitée. [12]
- **des *spécificateurs de forme et de taille*** permettent de décrire un objet, un animal, une scène. C'est à l'aide de la forme des mains, de leur orientation et de leur mouvement que le signeur décrit une forme et les dimensions d'un objet.
- **des *classificateurs*** sont similaires aux spécificateurs dans le sens où ils représentent également un objet (ou personne, animal...) et donc la forme de la main est en rapport

avec celle de l'objet ou de sa fonction. Mais ils ont un tout autre rôle car ils servent en quelque sorte de prénom. Lorsqu'un objet a été cité dans une phrase à l'aide d'un signe du vocabulaire standard, un classificateur peut ensuite être utilisé pour représenter cet objet dans le reste de la phrase. On peut avec ce classificateur préciser la position de l'objet ou décrire une trajectoire qu'il a empruntée.

Dans la figure (I.7), on peut voir une phrase qui regroupe les trois classes de signe. Le premier signe est un **spécificateur** de forme et de taille, il décrit une caisse en symbolisant ses côtés. Le signe suivant est le signe **standard** [POMME]. Enfin, la troisième étape montre des signes qui sont des **classificateurs** symbolisant des pommes que l'on place à différents endroits pour représenter le tas de pomme que contient la caisse.[18]



Figure I.7. : La phrase [une caisse pleine de pomme].

10. Interaction entre les deux mains :

Un signe peut aussi bien faire intervenir une main que deux mains et les deux mains ont différentes façons d'interagir. Lorsque les deux mains sont impliquées dans un signe, deux cas se présentent :

- Dans le premier cas on voit apparaître un rôle pour chaque main. Une main est dite **dominante** et a pour rôle de décrire "l'action", tandis que l'autre main qui est appelée main **dominée** sert de référence à cette action [2]. Par exemple, avec le signe [TOMBER DE] (figure I.8), on peut voir la main dominante (main droite dans la figure) qui mime l'action de tomber d'un support (référence donnée par la main dominée). En général la main dominante se déplace au cours du geste tandis que la main dominée reste statique.

- Dans le deuxième cas, les deux mains sont complètement synchronisées : leurs paramètres sont identiques ou symétriques. Par exemple, avec le signe [TABLE] (figure 1.8), les deux mains ont une configuration et une orientation identiques et des mouvements de même trajectoire mais de sens opposés. [18]



Figure 1.8. : Les signes [TOMBER DE] et [TABLE]

11.Enrichissement de la langue des signes :

La langue des signes, telle qu'elle est pratiquée au quotidien dans la communauté sourde s'est enrichie depuis sa reconnaissance à ce jour. Des néologismes fleurissent chaque jour pour évoquer de nouveaux concepts. C'est la preuve que la langue est bien vivante et que ses usagers sont soucieux de pouvoir parler de tout ce qui les intéresse, que ce soit dans le secteur médical, social, juridique, politique, culturel, haute technologie...[13]

12.La Langue des Signes Française (LSF) :

La LSF est une langue vivante. Sa structure est complexe. Elle utilise le corps et l'espace environnant comme supports signifiants. Comme toute autre langue des signes, ses constructions grammaticales et lexicales sont précises et inévitables.

En effet, la LSF n'a pas seulement une propriété figurative et descriptive. Elle autorise l'abstraction dans différents registres tels que l'humour, la poésie, le chant, l'argot, etc.

En LSF, on peut discuter de philosophie, de psychologie, d'art, de religion et de bien d'autres domaines à haut pouvoir d'abstraction.[14]



Figure I.9. : Exemple d'une phrase en LSF composé des signes
"je donne un gâteau au garçon qui est à ma droite"
[Garçon] "index vertical ", [gâteau] et [donner]

Le temps d'action s'exprime horizontalement par les signes :

- Le présent s'indique directement près de corps
- La future s'annonce devant le corps avant de faire les signes
- Le passé s'annonce derrière le corps avant de faire les signes

13.L'alphabet de la langue des signes française :

La langue des signes se compose de gestuelles qui représentent un mot entier ou une phrase, et d'un alphabet dactylogique (le correspondant signé de l'alphabet latin).

Cet alphabet est constitué de signes dont les configurations gestuelles de la main représentent les lettres de l'alphabet. Les sourds l'utilisent pour épeler les mots qui n'ont pas d'équivalent en L.S.F. (comme les néologismes, les noms-propres peu utilisés), ou dans certains signes reprenant la lettre dactylogique initiale du mot en français écrit. [15]

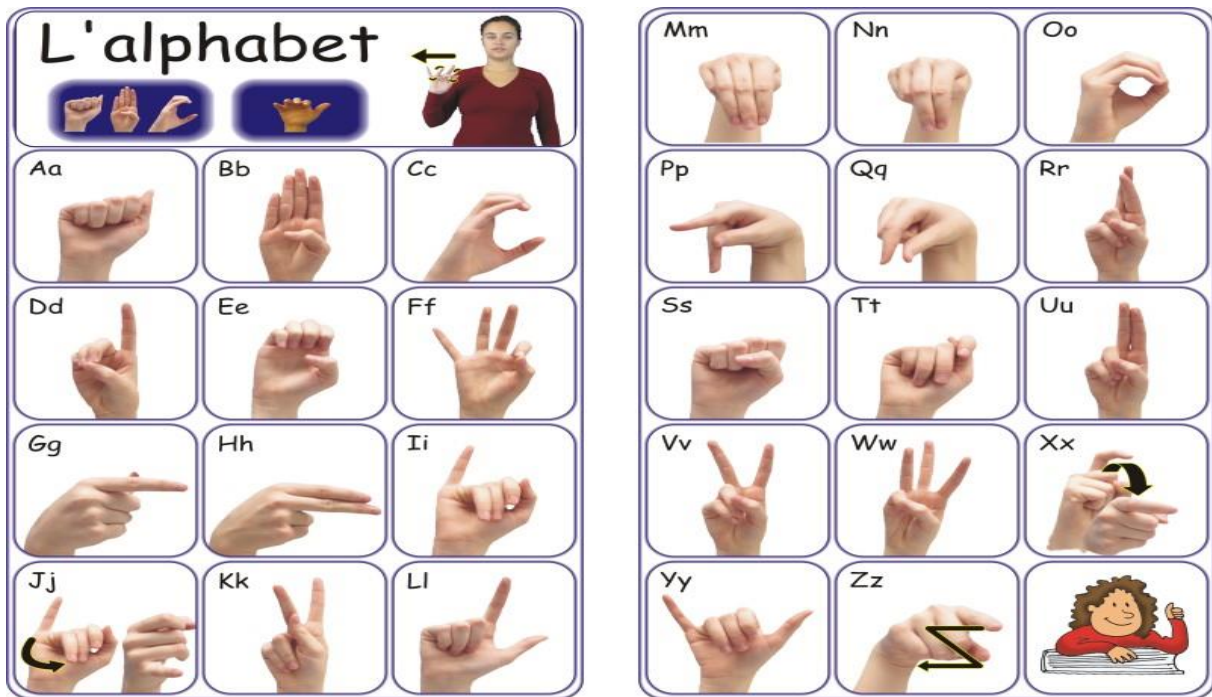


Figure I.10. : Alphabets de la langue des signes.

14. La langue des signes est une langue universelle :

« Toutes les langues des signes sont plus proches entre elles que ne le sont les langues vocales ». (Delaporte, 2002).

Comme l'énonce **Yves Delaporte** dans la citation ci-dessus les langues des signes ont beaucoup plus de similitudes entre elles que ne l'ont les langues vocales. Un sourd chinois mettra par exemple beaucoup moins de temps à comprendre et à communiquer avec un de ses amis sourds français que ne le feront deux personnes entendantes française et chinoise.

Cependant, chaque pays possède sa propre langue avec ses règles syntaxiques, lexicales, ses exceptions...

En 1850, **Ferdinand Berthier**, professeur sourd de l'Institut National des Jeunes Sourds de Paris, proclamait : « Pendant des siècles, des lettrés de tous pays ont cherché à découvrir une langue universelle et ont échoué. Mais elle existe autour de nous, et c'est la langue des signes. » Cette prise de conscience du caractère international de ce qui fut longtemps désigné sous le terme de langage mimique. [16]

Le mystère est percé : les sourds de différents pays peuvent communiquer ensemble sans problème, après un court temps d'adaptation où chacun prend connaissance du lexique de son interlocuteur. Dans les congrès internationaux de sourds, les échanges informels ainsi que les conférences se passent même d'interprètes ! Autrement dit, entre les sourds du monde entier, les barrières linguistiques n'existent pas. [14]

Il existe par exemple la langue des signes américaine (ASL), la langue des signes britannique (BSL), la langue des signes belge (langue des signes de Belgique Francophone, LSF), la langue des signes québécoise (LSQ), etc.

15. Conseils et attitudes :

La surdit  est un handicap de communication et d'acc s   l'information. Les personnes sourdes :

- ✓ compensent par une concentration accrue, ce qui peut les fatiguer.
- ✓ sont susceptibles d' tre bruyantes car elles sont inconscientes du bruit qu'elles g n rent.
- ✓ perdent plus facilement l' quilibre (probl me rencontr  chez certains sourds, d    l'oreille interne).
- ✓ communiquent dans la langue des signes de leur pays et utilisent aussi la lecture sur les l vres.
- ✓ ont besoin d'une bonne visibilit  ( clairage, distance de communication).
- ✓ pour appeler un sourd, si crier est inutile, il existe 8 m thodes efficaces, comme par exemple faire clignoter la lumi re d'une pi ce.

Voici quelques attitudes cl s   adopter en pr sence d'une personne sourde :

- ✓ Pr sentez-vous le visage d gag ,  clair  (attention au contre-jour).
- ✓ Parlez lentement en articulant (sans exc s) avec des mots simples.
- ✓ Accompagnez vos paroles de gestes simples.
- ✓ Pensez   utiliser papier et crayon.
- ✓ Assurez-vous que l'information est bien comprise (acquiescement) et attirez l'attention.
- ✓ Accompagner les signaux sonores par des signaux lumineux.

- ✓ Doubler les messages oraux (comme les annonces dans les magasins) par des textes écrits. [14]

16. Conclusion :

La langue des signes va s'imposer progressivement. Son essor et l'accès à l'enseignement permettent alors aux sourds d'accéder à de vrais métiers et de se regrouper en associations.

C'est aussi un potentiel de vie : mieux que tout autre mode de communication, la langue des signes permet à la personne sourde un développement harmonieux de ses facultés cognitives et psychologiques. Apprise dès la petite enfance, elle permettra l'accès naturel aux différents apprentissages, y compris celui de la langue française versus écrit, dans le respect des capacités cognitives individuelles, c'est l'élément clef qui ouvre les portes du monde et de la connaissance aux sourds, qui leur permet d'être des Hommes parmi les Hommes.

Dans le prochain chapitre on va présenter les différentes techniques qui sont déjà développées dans le domaine de la reconnaissance des gestes et plus précisément dans la reconnaissance de la langue des signes.



Chapitre II



L'état de l'art

(Les méthodes de reconnaissance des gestes)

1. Introduction :

Lorsqu'on ne peut pas transmettre un message, on utilise les gestes comme moyen de communication complémentaire. Il existe différentes communications gestuelles suivant les besoins d'expressivité. Les plus simples sont les gestes de commande qui ne se combinent pas entre eux ou avec un autre canal de communication, les gestes co-verbaux ou les mimiques sont un peu plus complexes et sont complémentaires vis à vis de la parole. Au plus haut degré de complexité se situent les gestes des Langues des Signes qui se combinent pour former des phrases et ce que nous intéressent dans notre mémoire.

Dans la langue des signes, chaque geste a déjà attribué ce qui signifie, et des règles strictes de contexte et de la grammaire peuvent être appliquées à faire de la reconnaissance traitable.

Ce chapitre sera consacré à l'étude des méthodes de reconnaissance des gestes proposées dans la littérature et plus précisément les gestes de la langue des signes.

2. La reconnaissance des gestes :

La reconnaissance des gestes désigne l'ensemble des opérations permettant d'analyser une scène à savoir la capture des gestes (par exemple à l'aide d'une caméra ou d'un gant dotés de capteurs), la segmentation, l'évaluation des poses et l'interprétation. [19]

Des applications d'analyse et de surveillance peuvent profiter des méthodes que proposent les scientifiques dans ce domaine.

2.1. Type de reconnaissance de gestes :

On peut distinguer deux types de reconnaissance : la reconnaissance statique ou dynamique, c'est-à-dire respectivement basée sur une ou plusieurs frames. [19]

2.1.1. La reconnaissance statique : utilise les données spatiales, un frame après l'autre. Les approches statiques comparent des informations préenregistrées avec l'image courante.

2.1.2. La reconnaissance dynamique : utilise des caractéristiques temporelles dans la tâche de reconnaissance. Les données utilisées peuvent être de bas ou de haut niveau. La reconnaissance basse niveau est basée sur des données spatio-temporelles sans trop de traitements. Les méthodes de plus haut niveau sont basées sur les données de l'estimation de pose.



Figure II.1.Geste Statique



Figure II.2.Geste dynamique

2.2. Définition des gestes :

Les gestes sont un des moyens de communication les plus riches que l'être humain possède. Ils permettent d'agir sur le monde physique, et servent aussi à communiquer. De plus, le geste permet à la fois d'émettre des informations, mais aussi d'en recevoir. [20]

2.3. Les fonctions du geste de la main :

Cadoz [21] définit trois fonctions principales de la main :

2.3.1. La fonction ergotique : la main joue le rôle d'organe moteur et agit sur le monde physique pour le transformer. Elle applique aux objets des forces, pour les déplacer ou les déformer.

2.3.2. La fonction épistémique : la main joue le rôle d'organe de perception. Le sens du toucher donne des informations sur la forme, l'orientation, la distance, la grandeur, le poids, la température, les mouvements des objets, etc.

2.3.3. La fonction sémiotique : la main joue le rôle d'organe d'expression pour l'émission d'informations visuelles. Cela comprend le geste co-verbal, qui accompagne la parole, ou les gestes permettant une communication basique lorsqu'on ne peut pas utiliser la parole, comme les gestes utilisés dans le langage des signes qu'ils sont basés sur des principes de linguistique et qu'ils permettent de combiner des gestes et des signes pour former des structures grammaticales utiles à la conversation. [20] Ce type de geste que nous intéressent dans notre thèse et ce que nous avons déjà cité auparavant.

2.4. Les gestes de la main :

Les gestes de la main sont un canal de communication naturel et intuitif chez l'homme pour interagir avec son environnement. Ils servent à manipuler des objets, à renforcer la parole, ou à communiquer dans un environnement bruyant. Ils peuvent aussi représenter un langage à part entière avec la langue des signes. Les gestes peuvent avoir une signification différente suivant la langue ou la culture : les langues des signes en particulier sont spécifiques à chaque langue. [20]

2.5. Le problème de la reconnaissance des gestes :

Le principal problème de la reconnaissance des gestes reste la segmentation et le traçage d'objets d'intérêt dans une scène réelle (problème quasiment inexistant en utilisant des gants) et l'interprétation des gestes dans un flux d'information généralement très bruyant. [19]

Et plus précisément d'explicitier certains des problèmes rencontrés lorsque l'on cherche à concevoir un système de reconnaissance des gestes de la langue des signes et de proposer des solutions adaptées. Les trois aspects traités ici concernent la simultanéité d'informations véhiculées par les gestes des mains, la synchronisation éventuelle entre les deux mains et le fait que différentes classes de signes peuvent se rencontrer dans une phrase en langue des signes.

3. Les méthodes de reconnaissance des gestes :

3.1. La reconnaissance en utilisant des points caractéristiques spatio-temporelle (STIP) :

Consiste à utiliser des points caractéristiques spatio-temporels (STIP) afin de décrire localement les mouvements par des histogrammes de flux optique qui sont ensuite classés par une machine à support vecteurs (SVM). [22]



Figure II.3. : Exemple de détection des STIP sur la séquence de geste « Salut ».

3.2. Reconnaissance de gestes dynamiques de la main :

Le geste dynamique est caractérisé par les signatures statiques de début et fin de geste (histogramme des orientations du gradient) et par la signature dynamique (superposition des squelettes de chacune des images composant la séquence). La reconnaissance du geste se fait par une mesure de similitude entre ces signatures et des signatures d'un alphabet connu.

La méthode proposée a été testée sur une quarantaine de gestes pris parmi un alphabet de 10 gestes simples. Moyennant l'utilisation de plusieurs séquences d'apprentissage pour un même geste, le taux de reconnaissance obtenu atteint les 100%. [23]

Les différentes étapes sont présentées dans la figure suivante :

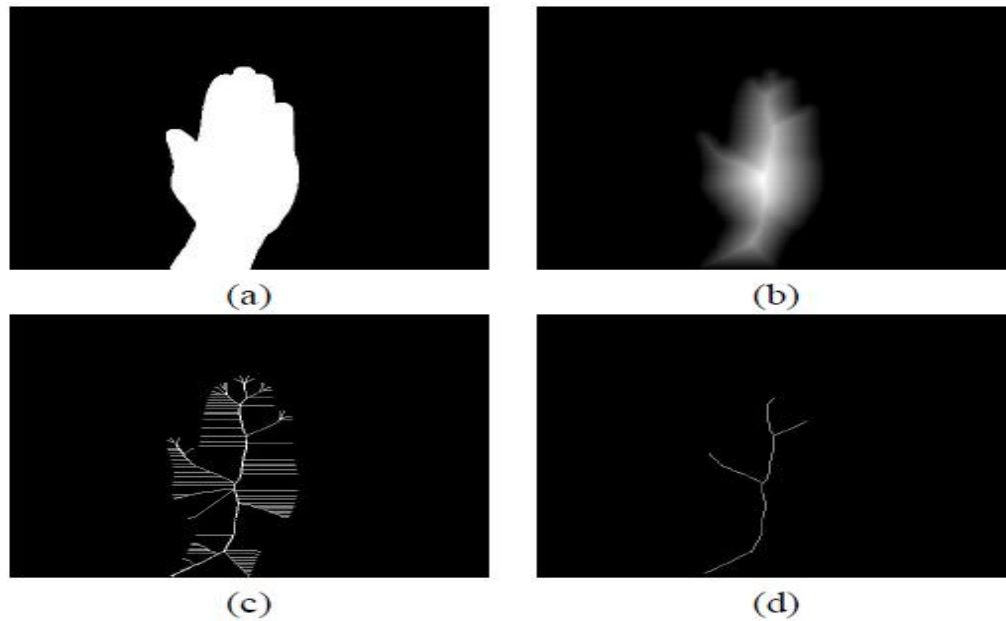


Figure II.4 : (a) Image binaire, (b) Image de distances, (c) le squelette, (d) Squelette simplifié.

3.3. La reconnaissance de gestes en utilisant un système multi-caméras :

Ils détectent les points d'intérêts dans la scène afin d'extraire les régions caractéristiques et les décrivent par des histogrammes de gradient orienté HOG. Ensuite, ils effectuent un suivi temporel de ces descripteurs et classent les gestes en utilisant l'algorithme K-means à partir d'une base de gestes d'apprentissage. La reconnaissance est réalisée par un algorithme de plus proches voisins (K-NN). [24]

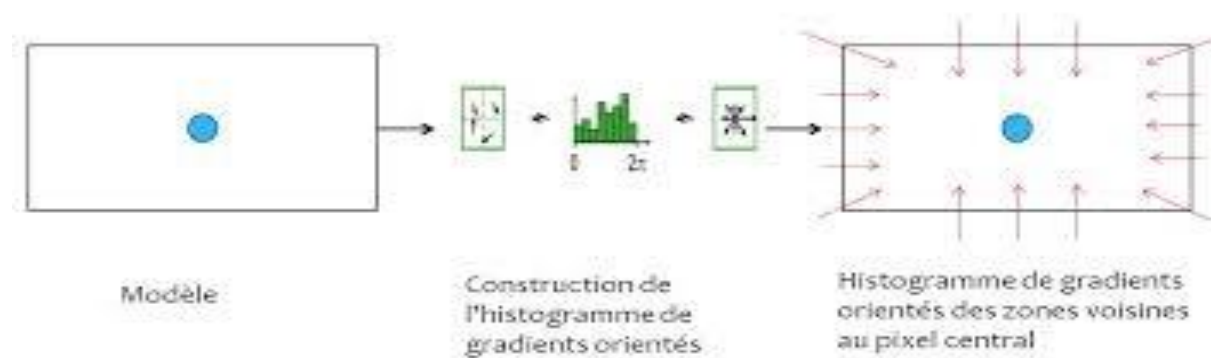


Figure II.5 : Construction de l'histogramme de gradients orientés.

3.4. La reconnaissance en utilisant les gants :

On générale les gants sont munis de capteurs au niveau du gant qui fournissant la position de la main et les angles des articulations des doigts et toutes les informations sur la configuration de la main directement.

Les gants sont utilisés de longue date pour la reconnaissance de la langue des signes, car ils fournissent les positions précises et fiables des articulations de la main. Malheureusement, ces gants ont un coût élevé et sont encombrants, leur utilisation est contraignante pour l'utilisateur. [20]



Figure II.6. : Exemples des gants : (a) le *5DT Data Glove 5 Ultra*, et (b) le *CyberGlove*.

3.5. La reconnaissance de la langue des signes en utilisant des sous-unités linguistiques :

La reconnaissance de la langue des signes en utilisant des sous-unités linguistiques présente trois types de sous-unités ; Ces qui apprendre à partir des données d'apparence ainsi que ceux déduit à partir suivi 2D ou 3D des données. Ces sous-unités sont ensuite combinées à l'aide d'un classifieur au niveau de signe ; deux approches sont présentées :

1. La première utilise des modèles de Markov pour coder les changements temporels entre des sous-unités.
2. Le deuxième utilise le patron Boosting pour appliquer le descripteur discriminatif de sélection en même temps le codage de l'information temporelle. Cette approche est plus robuste au bruit et la performance s'améliore ainsi dans les tests indépendants de

signataire, l'amélioration des résultats de la 54% atteint par le Chaînes de Markov à 76%. [25]

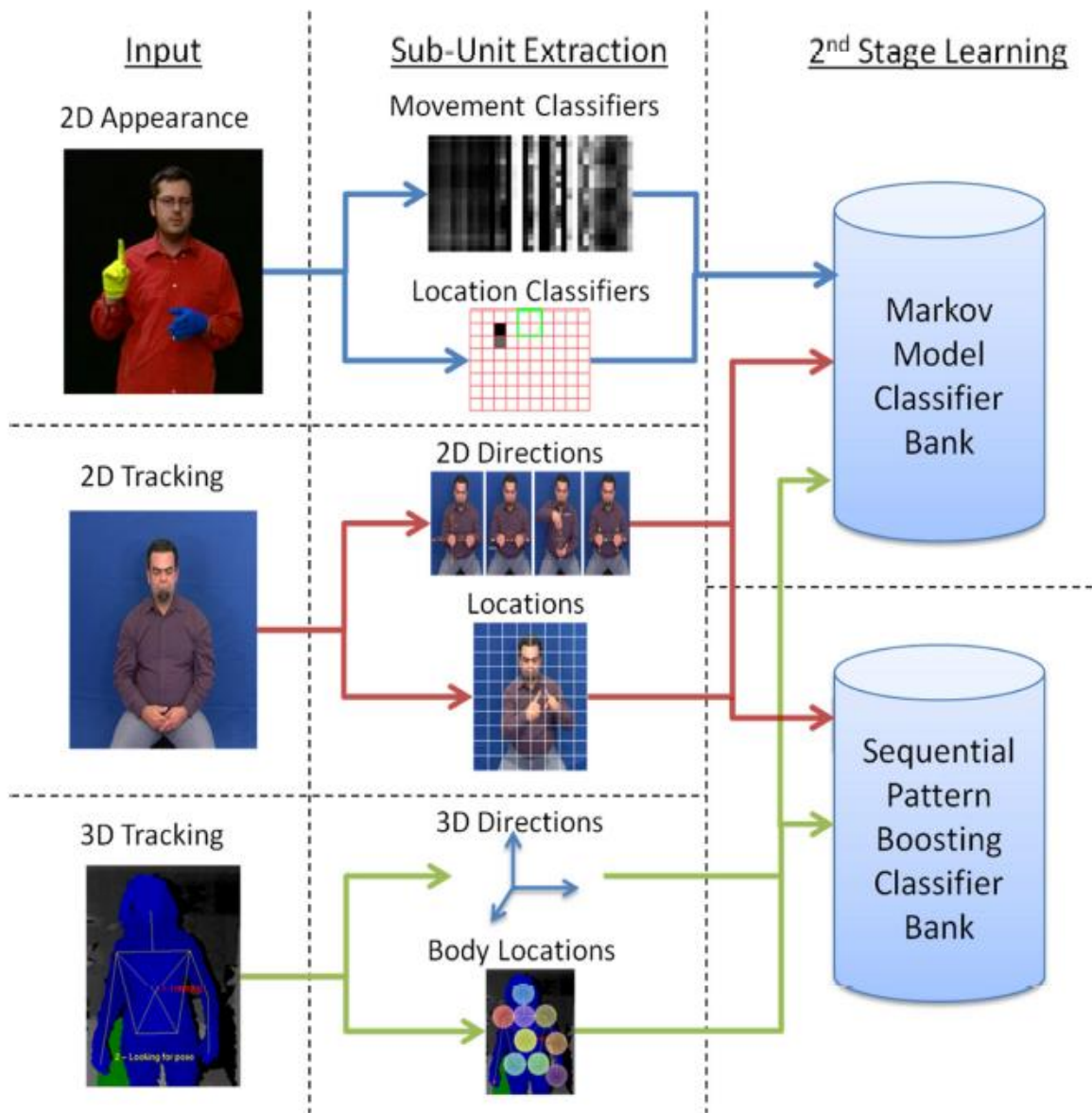


Figure II.7 : Aperçu sur les 3 types de sous-unités et les 2 différences entre les 2 niveaux de classifieur.

3.6. Reconnaissance des gestes par le suivi 3D de la main :

Cette approche est basée sur la classification des trajectoires 3D de la main acquise à l'aide d'un capteur Kinect. Chaque geste effectué est représenté par les coordonnées 3D de la main en fonction du temps. La reconnaissance est effectuée par le calcul de la déformation temporelle dynamique (DTW). [22]

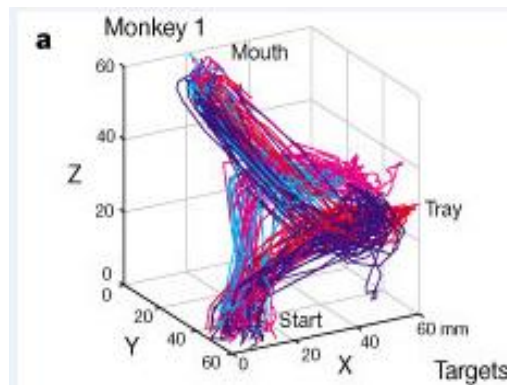


Figure II.8. Suivi 3D de la main en fonction du temps.

3.7. Le système de traduction de langue des signes en utilisant Kinect :

Ce système de traduction de la langue du signe utilise des images acquises par Kinect, est incorporé pour faciliter la communication entre des sourds et des personnes normales.

Le principal rôle de ce système est donné dans les figures II.7 et II.8 et il est composé de 2 modes ; **Mode de traduction (a)** lequel traduit les signes de la langue des signes à un texte ou parole et **Mode de communication (b)** lequel une personne normale peut communiquer avec signataire pendant avatar. [26]

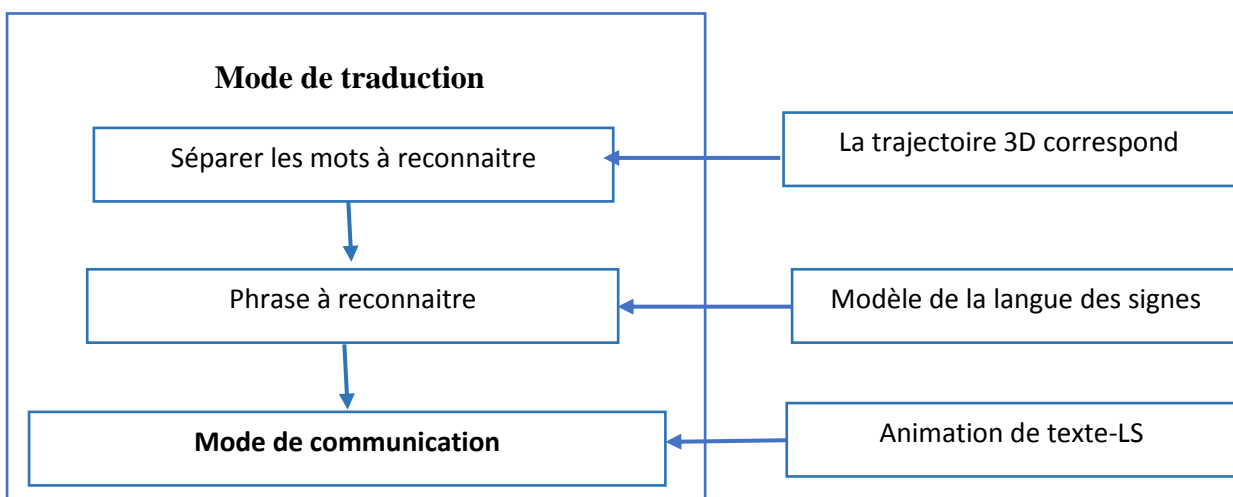
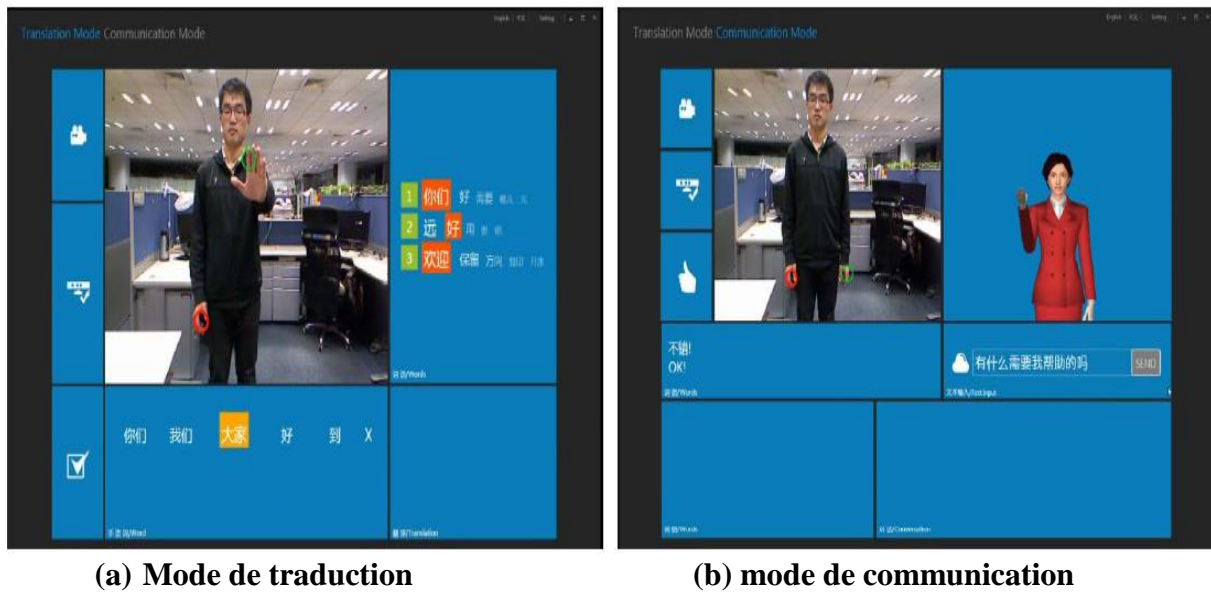


Figure II.9. : Les fonctions principales du système.



(a) Mode de traduction

(b) mode de communication

Figure II.10 : Les modes (a) et (b) du système

4. Conclusion :

Nous avons présentés dans ce chapitre les différentes approches de la reconnaissance de gestes. La littérature étudiée nous a permis de dégager les principaux défis de la reconnaissance de geste, de présenter les outils et les méthodes proposées.

Dans le chapitre qui suit on va présenter la méthode que nous avons choisi pour notre système de reconnaissance d'alphabet des sourds muets ainsi leur architecture générale et l'environnement de développement.



Chapitre III

SIFT Implémentation et résultats

(SCALE-INVARIANT FEATURE
TRANSFORM)

Partie I : SIFT

Scale-invariant feature transform

1. Introduction :

Un processus d'extraction de caractéristiques est un processus qui consiste à décrire une image, une séquence, ou un objet de l'image par plusieurs attributs dans un espace multidimensionnel. Ce processus d'extraction fournit la donnée d'entrée à un système de reconnaissance.

Dans la littérature, il existe plusieurs descripteurs de forme. Parmi ces descripteurs, nous avons choisi l'algorithme SIFT qu'on va détailler dans la première partie de ce chapitre. La deuxième partie sera consacrée à l'implémentation et la réalisation de l'application logicielle.

2. Définition :

Scale-invariant feature transform (**SIFT**), que l'on peut traduire par "transformation de caractéristiques invariante à l'échelle", est un algorithme utilisé dans le domaine de la vision assistée par ordinateur pour détecter et identifier les éléments similaires entre différentes images numériques (éléments de paysages, objets, personnes...), ceci ayant pour objectif de décrire des zones d'intérêts dans une image. Il a été développé en 1999 par le chercheur David Lowe, et le propriétaire du brevet est l'Université de la Colombie-Britannique.

L'idée générale de SIFT est donc de trouver des points d'intérêts qui sont invariants à plusieurs transformations au changement d'intensité, de mise à l'échelle et de rotation, ce qui fait de lui un descripteur très robuste.

3. L'algorithme de SIFT :

L'algorithme SIFT se divise en 2 grandes étapes :

- **La première étape** permet d'extraire les caractéristiques d'un objet et de calculer ses descripteurs, c'est-à-dire, de détecter les caractéristiques qui sont les plus susceptibles de représenter cet objet, de le définir et de le discriminer par rapport aux autres.
- **La seconde étape** consiste à mettre en place une procédure de mise en correspondance (« matching »).

3.1. Calcul des points d'intérêts et des descripteurs :

C'est La transformation d'une image en ensemble de vecteurs descripteurs et se compose en 5 étapes :

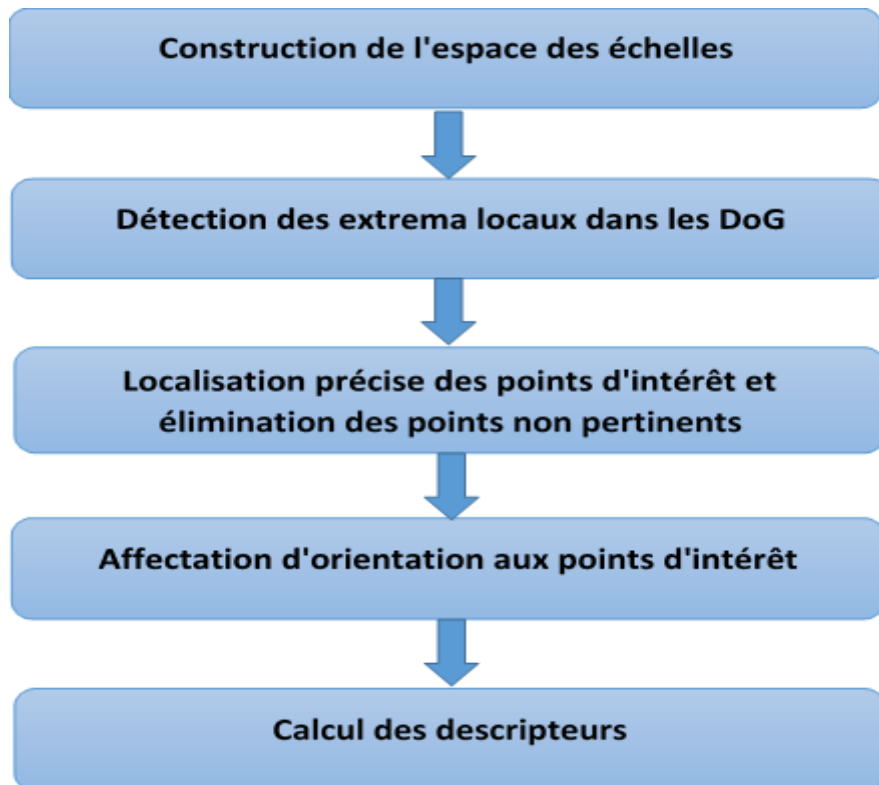


Figure III.1. Étapes de calcul des points d'intérêts et des descripteurs

3.1.1. Construction de l'espace des échelles (la pyramide) :

Consiste en un lissage d'image sur plusieurs échelles et à des dimensions différentes et permet de justifier l'invariance des descripteurs par rapport à la mise à l'échelle.

Une pyramide de gaussienne utilisé permet de représenter des données sur plusieurs échelles.

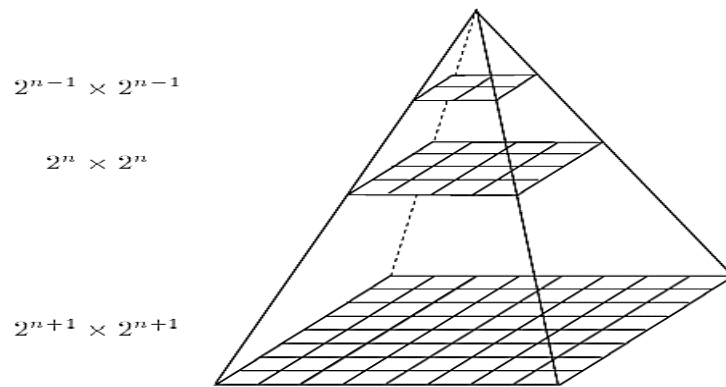


Figure III.2. Une pyramide de gaussienne.

3.1.2. Détection des extrema locaux dans les DoG :

L'idée est de calculer une différence de gaussiennes (DoG), entre deux images consécutives d'une même octave dans la pyramide de gaussiennes pour obtenir une pyramide de DoG ; On souhaite détecter les extrema locaux dans les contours.

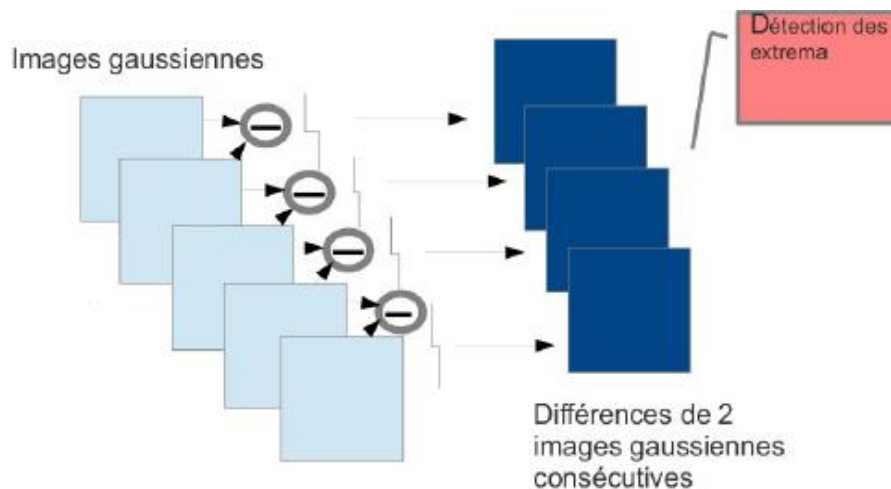


Figure III.3. Structure de détection des extrema

Cette étape permet de détecter des points d'intérêts sur plusieurs échelles et sur différents niveaux de résolution. Les points d'intérêts recherchés constituent les extrema locaux des images des DoG à travers les différentes échelles. Chaque pixel des images des DoG est alors comparé à ses 26 voisins ; 8 voisins dans la même échelle et 9 voisins sur les deux échelles voisines qui l'encadrent. Si le pixel est un extremum local, c'est-à-dire s'il est supérieur ou inférieur à ses voisins alors il est sélectionné en tant que point candidat pour la suite.

Détection d'extrema locaux: le pixel marqué X est comparé à ses 26 voisins sur 3 échelles mais dans la même octave.

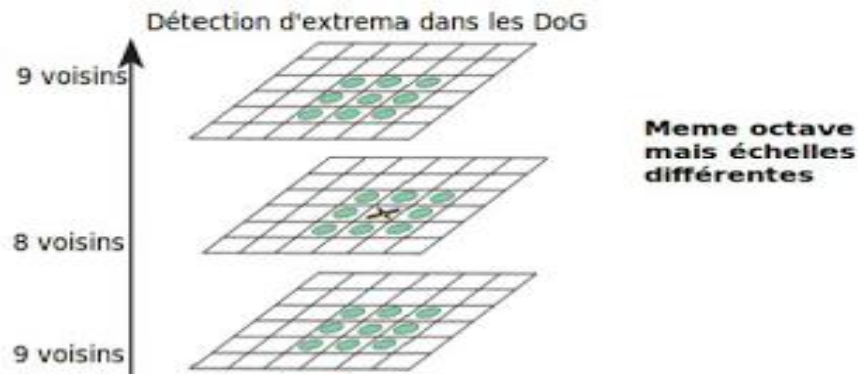


Figure III.4. Détection d'extrema locaux

3.1.3. Localisation précise des points d'intérêt et élimination des points non pertinents :

Après la détection des extremums dans l'espace des échelles, l'algorithme élimine les points de faible contraste, puis les points situés sur les arêtes. C'est la détection des coordonnées exactes des points d'intérêt grâce à une interpolation.

3.1.4. Affectation d'orientation aux points d'intérêt :

Calcul des histogrammes d'orientations en fonction du voisinage. Permet de justifier l'invariance des descripteurs par rapport à la rotation. Il s'agit à présent d'attribuer à chaque PI sélectionné une ou des orientations en utilisant la direction des gradients des voisins directs de ce point. Pour cela on parcourt tous les pixels de toutes les images gaussiennes à toutes les octaves et on leur affecte une orientation.

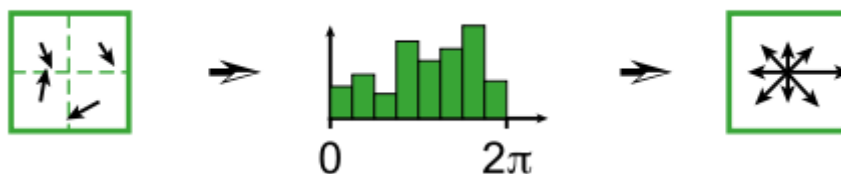


Figure III.5. Construction de l'histogramme des orientations.

3.1.5. Calcul des descripteurs :

Génération des vecteurs descripteurs associés à chaque point d'intérêt. On considère toujours autour du point-clé, une région de 16×16 pixels, subdivisée en 4×4 zones de 4×4 pixels chacune. Sur chaque zone est calculé un histogramme des orientations comportant 8 intervalles.

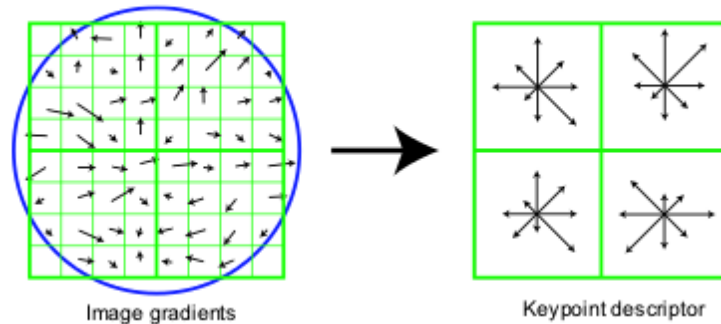


Figure III.6. Construction d'un descripteur SIFT ; vecteur de taille 128 (8×16).

3.2. Mise en correspondance :

Grâce au calcul des descripteurs des points d'intérêt des deux images, l'algorithme SIFT peut trouver des paires de points d'intérêt communs aux deux images par comparaison de leurs descripteurs (qui sont stables par rotation, changement d'échelle, variation de la luminosité...).

Il repose sur le calcul des distances euclidiennes dans l'espace des descripteurs (espace de dimension 128). La méthode se base sur la construction d'un *k-d tree* et sur l'emploi l'algorithme *Best Bin First*.

L'analyse des correspondances des points trouvées pour décider si l'image requête est similaire à des images de la base. Cette analyse est faite par recherche d'une transformation affine entre l'image requête et les images de la base par le calcul de distance euclidienne.

4. Les techniques utilisées dans la recherche de correspondance :

Étape	Techniques utilisées	Avantages
Extraction des points-clés	Pyramide de gradients, Différence de gaussiennes (DoG), Interpolation	Précision, stabilité, invariance aux modifications d'échelle et à la rotation
Calcul des descripteurs	Génération des vecteurs descripteurs, Histogramme des orientations de l'image	Stabilité relative aux transformations affines et à la luminosité
Recherche des correspondances avec les descripteurs de l'image requête	La distance euclidienne, Arbre kd, Plus proche voisin approximatif (<i>Best Bin First</i>)	Efficacité, Rapidité

Table III.1 : techniques utilisées dans la recherche de correspondance

5. Pour Quoi SIFT ?

Des études comparatives de descripteurs locaux ont été présentées dans [29]. Les auteurs ont présenté une évaluation comparative de plusieurs descripteurs locaux couplés à divers détecteurs des zones d'intérêts. L'intérêt de ces études est de comparer des couples détecteur-descripteur, et de conclure sur leur efficacité pour une application de recherche de similarités entre images.

D'après cette étude [29], le descripteur SIFT reste la référence en matière de détection de points d'intérêt. Il combine les DoG qui sont invariants à la translation, rotation et changement d'échelle avec un descripteur basé sur les distributions d'orientations de gradient qui, de plus, est robuste aux changements d'illumination.

En résumé, les descripteurs basés sur l'algorithme SIFT sont les plus performants. Pour le détecteur de région d'intérêt, et la recherche des similarités entre des images et pour cela nous choisissons d'utiliser celui également pour obtenir des meilleurs résultats.

Celle [29] met en évidence les avantages et inconvénients de SIFT, PCA-SIFT et SURF grâce au tableau suivant :

MÉTHODE	TEMPS	ÉCHELLE	ROTATION	FLOU	ILLUMINATION	AFFINE
SIFT	normal	meilleur	meilleur	meilleur	normal	bon
SURF	meilleur	bon	normal	bon	meilleur	bon
PCA-SIFT	bon	normal	bon	normal	bon	bon

Table.III.2 étude comparative entre SIFT, SURF et PCA-SIFT

Partie II : Implémentation et résultats

1. Introduction :

Dans ce chapitre nous allons présenter les méthodes et les moyens utiliser pour la réalisation et le développement de notre application.

Voici ci-dessous une représentation globale de notre système.

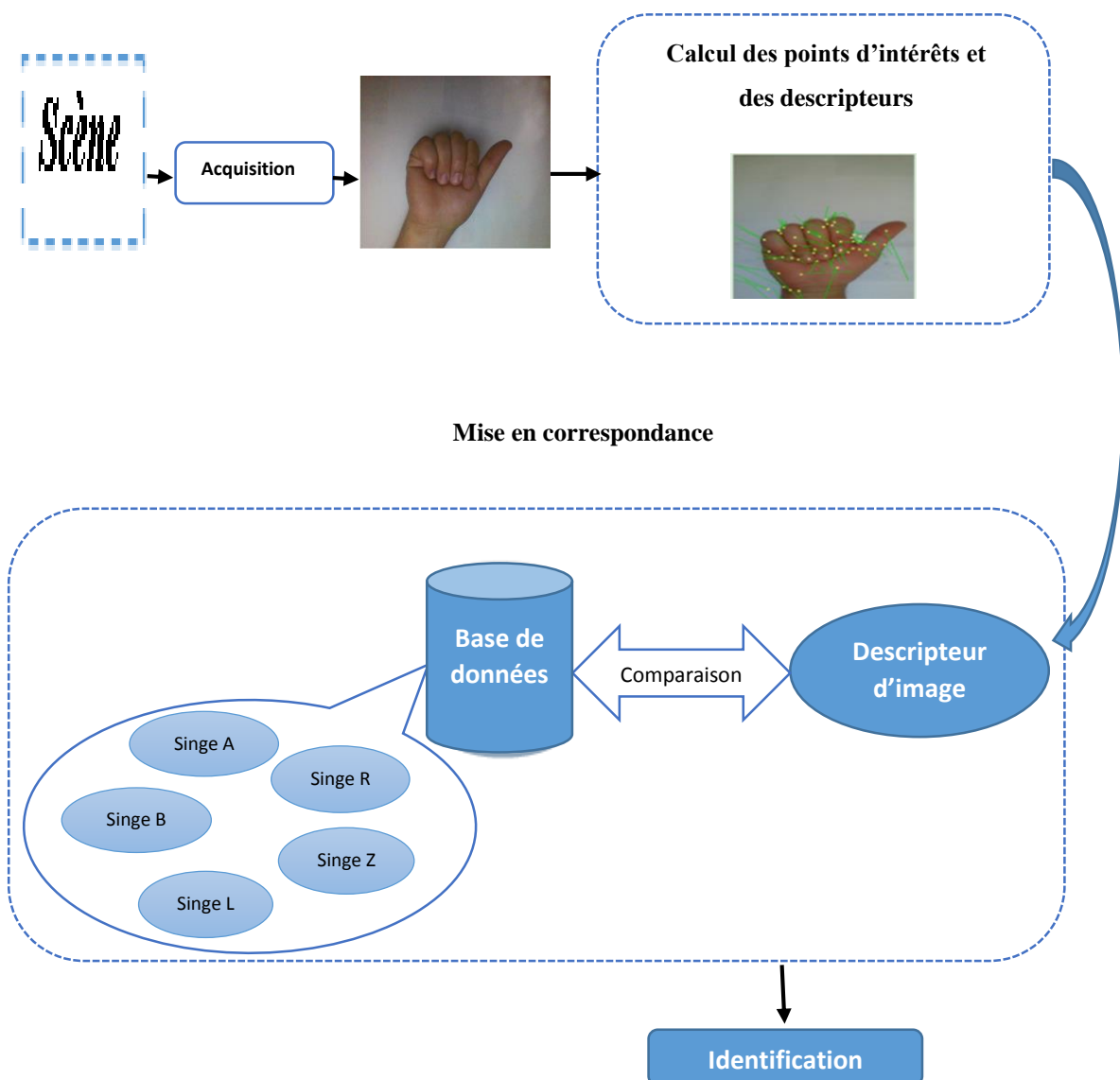


Figure III.7. Schéma global du système.

2. La conception du système :

La conception générale de notre application se devise on 2 étapes principales.

2.1. Création de la base des signes :

Pour cette étape, nous avons créé une base des images des alphabets. Pour chaque alphabet nous avons calculé les descripteurs correspondants.

2.2. Reconnaissance de l'alphabet

Nous sélectionnons une image requête acquise par une webcam puis on compare cette image de signe avec les descripteurs de toutes les alphabets de la base.

L'alphabet est identifié par le nombre des descripteurs similaire le plus élevés.

S'il y a plusieurs alphabets qui ont des descripteurs similaires aux descripteurs d'une image requête, notre application prend les trois premiers alphabets par ordre décroissant comme réponse. Si elle ne trouve pas des descripteurs similaires elle affiche un point d'interrogation « ? ».

3. Acquisition d'image :

Pour notre système nous avons utilisé un capteur de vision de type webcam qui fournit des images couleur de 120 x 160 de résolution et de 480*640 de dimensions , avec une fréquence de 15 images/seconde, le choix de cette webcam de faible résolution est due au fait, que nous nous intéressons à la forme de la main et non pas à ses détails et pour diminuer le temps de calcule des descripteurs.

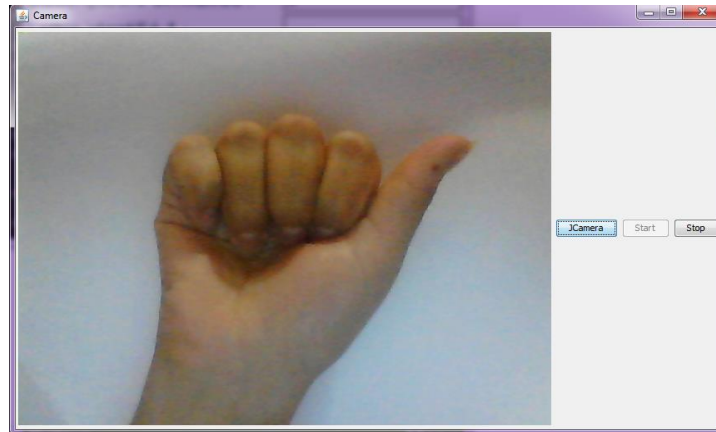


Figure.III.8. Interface d'acquisition.

4. Calcul des points d'intérêts et des descripteurs

La phase de calcul des points d'intérêts et des descripteurs est primordiale, elle consiste à extraire à partir de l'image des informations qui décrivent son contenu.

5. Base de données :

La première étape dans notre système est la création de la base des données qui est composé des images des singes attribué à aux alphabets de A jusqu'à Z.

Pour chaque alphabet on a choisi 3 cas différents (voir figure III.10).



Figure.III.9. trois images d'alphabet A.

6. Environnement de développement :

6.1. Android :

Android est très segmentée en termes de performances, d'architectures résolutions. Les processeurs des appareils sous Android peuvent être des ARM, des x86 d'INTEL.

L'appareil utilisé pour la majorité des tests était un Nexus 4, choisi pour sa prise en charge d'Android 4.3 et qui côté performance se place dans la moyenne des téléphone récent sous Android haut de gamme.

Le développement d'applications se fait en Java SE via l'utilisation du SDK Android. Du code natif peut être écrit en C++ et interfacé au code Java.

6.2. Architecture matérielle :

Nous avons développé notre application sur un PC ayant les caractéristiques suivantes :

- Processeur : intel core i5 2.40 GHZ.
- RAM: 8Go.
- Système d'exploitation: Windows 7.
- Webcam: intégré.

7. Présentation des interfaces de notre système :

7.1. L'interface principale :

Cette interface permet d'accéder à différentes fenêtres (Paramètres, Singes, Camera, Explorer) et de quitter notre système.

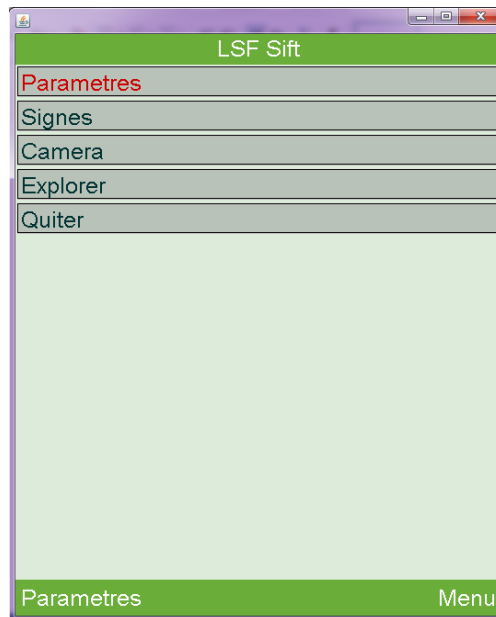


Figure III.10. Interface principale.

7.2. Fenêtre des paramètres :

Cette fenêtre sert à modifier les paramètres de l'algorithme Sift pour rechercher les meilleurs paramètres qui nous donne des bons résultats.

Nous avons travaillé avec les paramètres suggérés par **G.Lowe** [30] comme paramètres par default.

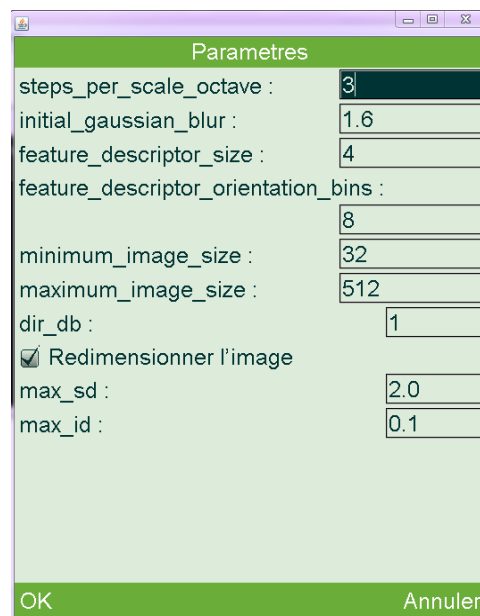


Figure. III.11. Fenêtre des paramètres.

Voici l'explication du des paramètres [30] que nous avons utilisé :

➤ **Étapes par octave échelle (steps per scale octave) :**

Les point-clés candidats sont extraites à toutes les échelles entre la taille maximale et minimale de l'image. Cet espace d'échelle est représenté dans chaque octave couvrant un nombre fixe d'étapes discrètes à partir d'échelle σ_0 à $2 \sigma_0$. $\sigma_0 = 3$ et pas plus de 10.

➤ **Flou gaussien initial (initial gaussian blur) :**

La localisation précise des points-clés nécessite un lissage initial de l'image. Si les images sont déjà floues, nous pourrions réduire le flou initial σ_0 légèrement pour obtenir plus de points-clés. La valeur par défaut $\sigma_0 = 1.6$ px.

➤ **La taille caractéristique de descripteur (feature descriptor size) :**

Le descripteur de Sift se compose de $n \times n$ histogrammes de gradient, chacun d'un bloc de 4×4 px. $n = 4$.

➤ **Paramètres d'orientation de descripteur (feature descriptor orientation bins) :**

C'est le nombre de bin d'orientation b par 4×4 px bloc. la valeur par défaut $b = 8$.

➤ **Taille de l'image minimale (minimum image size) :**

L'espace échelle s'arrête si la taille de l'octave serait inférieure à la taille minimale de l'image.

➤ **Taille d'image maximale (maximum image size) :**

L'espace d'échelle se commence avec la première octave inférieure ou égale à la taille d'image maximale.

7.3. Fenêtre des singes :

Cette fenêtre représente les singes des alphabets.

En premier temps, la valeur des descripteurs est nulle (avant le calcul des descripteurs), puis on calcule les descripteurs pour les trois cas de chaque alphabet et les affiches.

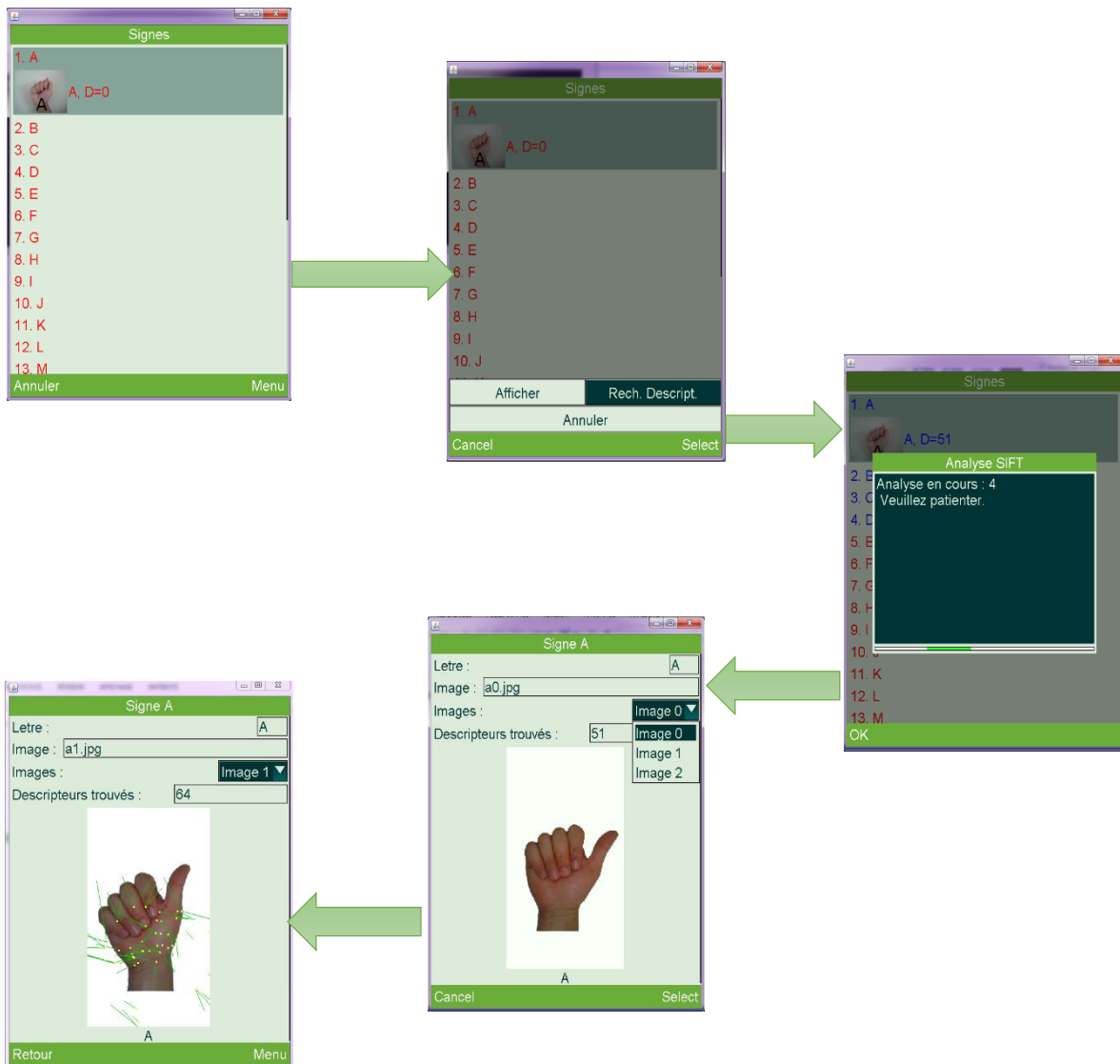


Figure. III.12. Fenêtre des Singes.

7.4. Fenêtre de la caméra :

Cette fenêtre nous permet d'accéder à la webcam et d'acquérir une image d'un signe.

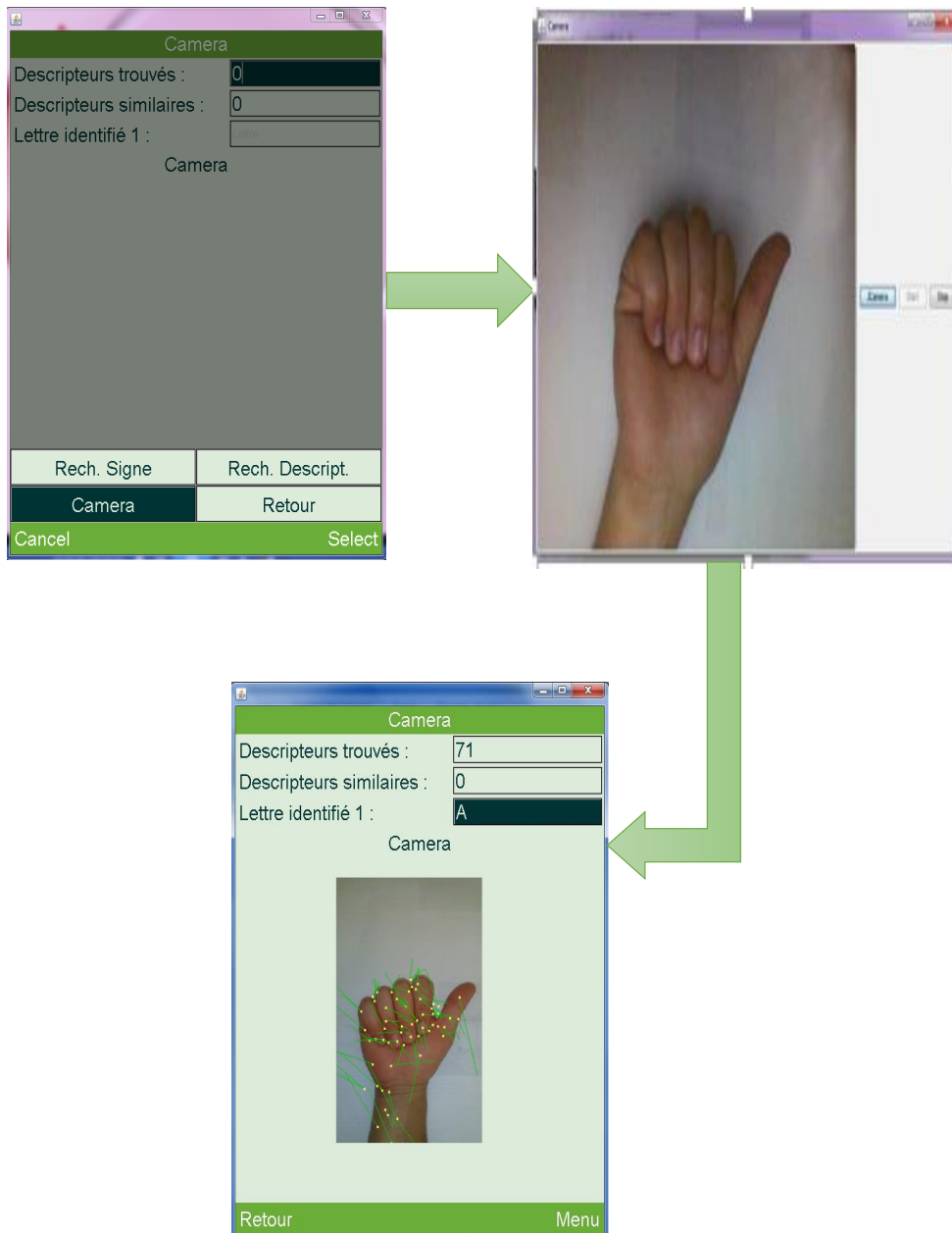


Figure .III.13. Fenêtre du camera

7.5. Fenêtre d'explorer :

Cette fenêtre nous permet d'accéder à notre base de teste qui contient des images des alphabets numérotés au hasard. Nous choisissons par la suite une image et on calcule leurs descripteurs puis on les compare avec les descripteurs de la base pour trouver l'alphabet correspondant.

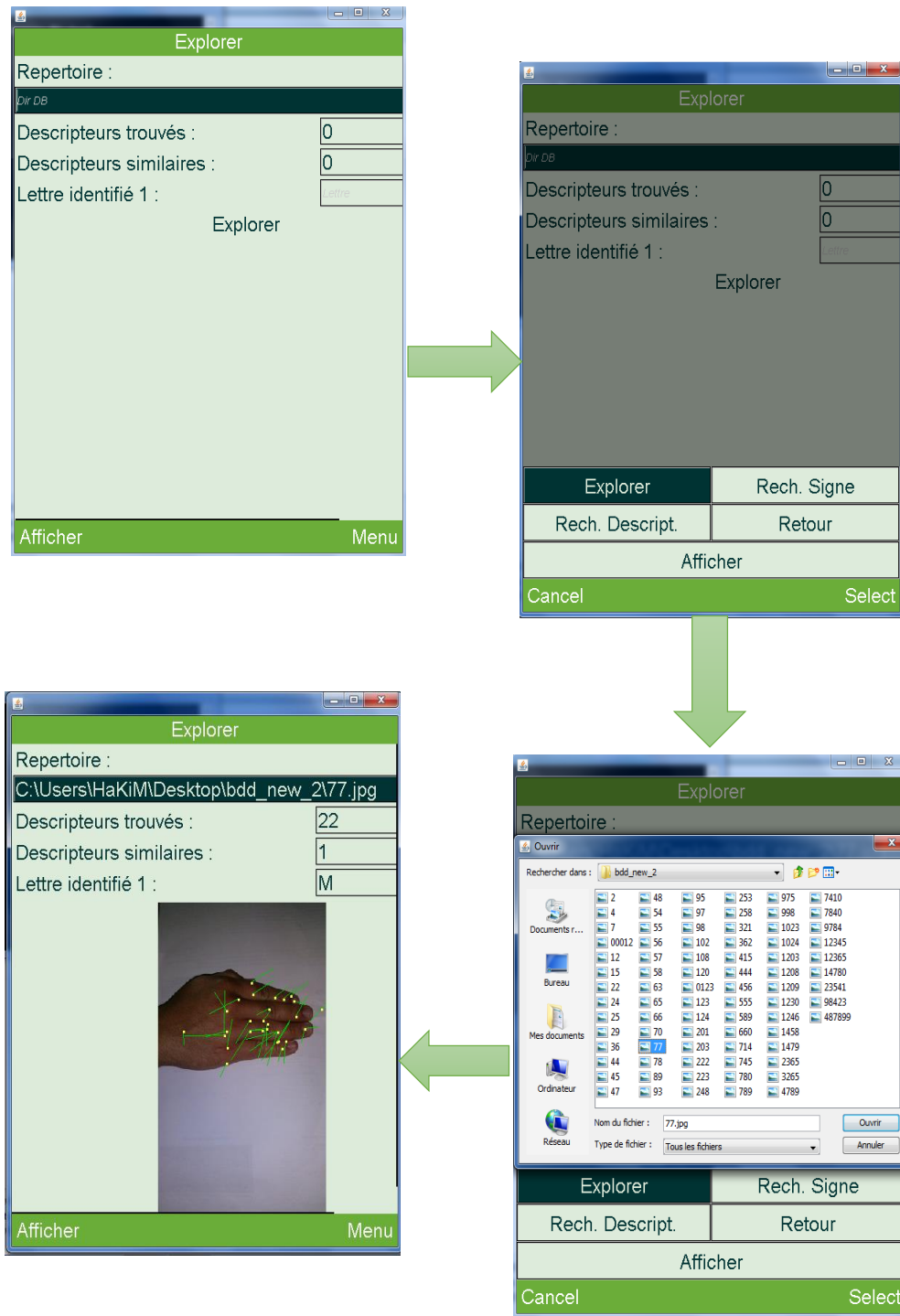


Figure .III.14. Fenêtres Explorier.

8. Expérimentation et résultats :

Afin de fixer les paramètres optimaux de l'algorithme de SIFT nous avons modifié chaque paramètre de l'algorithme. Cette expérimentation a été réalisé sur des alphabets qui semblent presque similaire a des autre. Pour cela nous avons sélectionné 3 images :

- Image 01 (singe C qui est semblable a singe O)
- Image 02 (singe U qui est semblable a singe R)
- Image 03 (singe K qui est semblable a singe V)

Les résultats obtenus sont présentés dans le tableau III.3

Paramètres Images	Steps per scale octave	Initial gaussian blur	Feature descriptor size	Feature descriptor orientation bins	Max sd	Max id	Confusion
Image 01	8.0	2.5	5.0	9.0	3.0	0.8	oui
	7.0	2.0	4.5	8.7	2.8	0.5	non
	5.0	1.0	4.2	8.5	2.4	0.2	non
	3.0	1.6	4.0	8.0	2.0	0.1	non
Image 02	8.0	2.3	5.0	9.0	3.0	0.8	oui
	6.0	2.0	4.8	8.5	2.7	0.5	oui
	4.0	1.8	4.3	8.2	2.2	0.2	non
	3.0	1.6	4.0	8.0	2.0	0.1	non
Image 03	6.0	2.2	5.0	9.0	3.0	0.8	oui
	5.0	2.0	4.8	8.7	2.8	0.5	oui
	4.0	1.8	4.5	8.5	2.5	0.2	oui
	3.0	1.6	4.0	8.0	2.0	0.1	non

Table.III.3.Résultats des modifications des paramètres

Afin de montrer la robustesse de l'algorithme SIFT, nous avons chargé des images de notre base sur le site web IPOL JORNAL [31].

Nous avons obtenu les résultats présentés dans les tableaux suivants (voir tableau III.3 et III.4).

	A	B	C	D	E	F	G	H	I	J	K	L
A	148	0	0	0	0	0	0	0	0	0	0	0
B	0	79	0	0	0	0	0	0	0	0	0	0
C	0	0	89	0	0	0	0	0	0	0	0	0
D	0	0	0	107	0	0	0	0	0	0	0	0
E	0	0	0	0	110	0	0	0	0	0	0	0
F	0	0	0	0	0	95	0	0	0	0	0	0
G	0	0	0	0	0	0	75	00	0	0	0	0
H	0	0	0	0	0	0	00	78	0	0	0	0
I	0	0	0	0	0	0	0	0	95	0	0	0
J	0	0	0	0	0	0	0	0	0	87	0	0
K	0	0	0	0	0	0	0	0	0	0	81	0
L	0	0	0	0	0	0	0	0	0	0	0	90

Table III.4. : Résultat de matching entre les alphabets

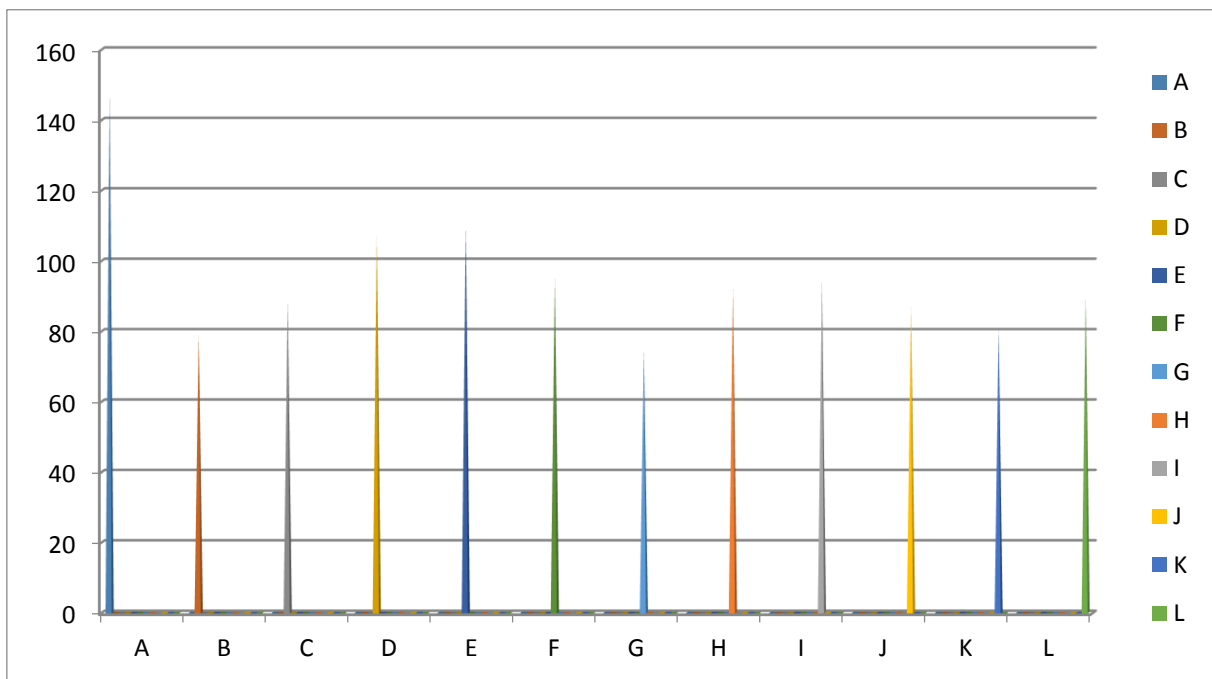


Figure III.15. Graphe de résultat de matching entre les alphabets.

	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
M	74	12	0	0	0	0	0	0	0	0	0	0	0	0
N	12	59	0	0	0	0	0	0	0	0	0	0	0	0
O	0	0	107	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	84	0	0	0	0	0	0	0	0	0	0
Q	0	0	0	0	103	0	0	0	0	0	0	0	0	0
R	0	0	0	0	0	124	0	0	08	0	0	0	0	0
S	0	0	0	0	0	0	106	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	115	0	0	0	0	0	0
U	0	0	0	0	0	08	0	0	124	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	135	0	0	0	0
W	0	0	0	0	0	0	0	0	0	0	177	0	0	0
X	0	0	0	0	0	0	0	0	0	0	0	139	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	115	0
Z	0	0	0	0	0	0	0	0	0	0	0	0	0	103

Table III.5. Résultat de matching entre les alphabets

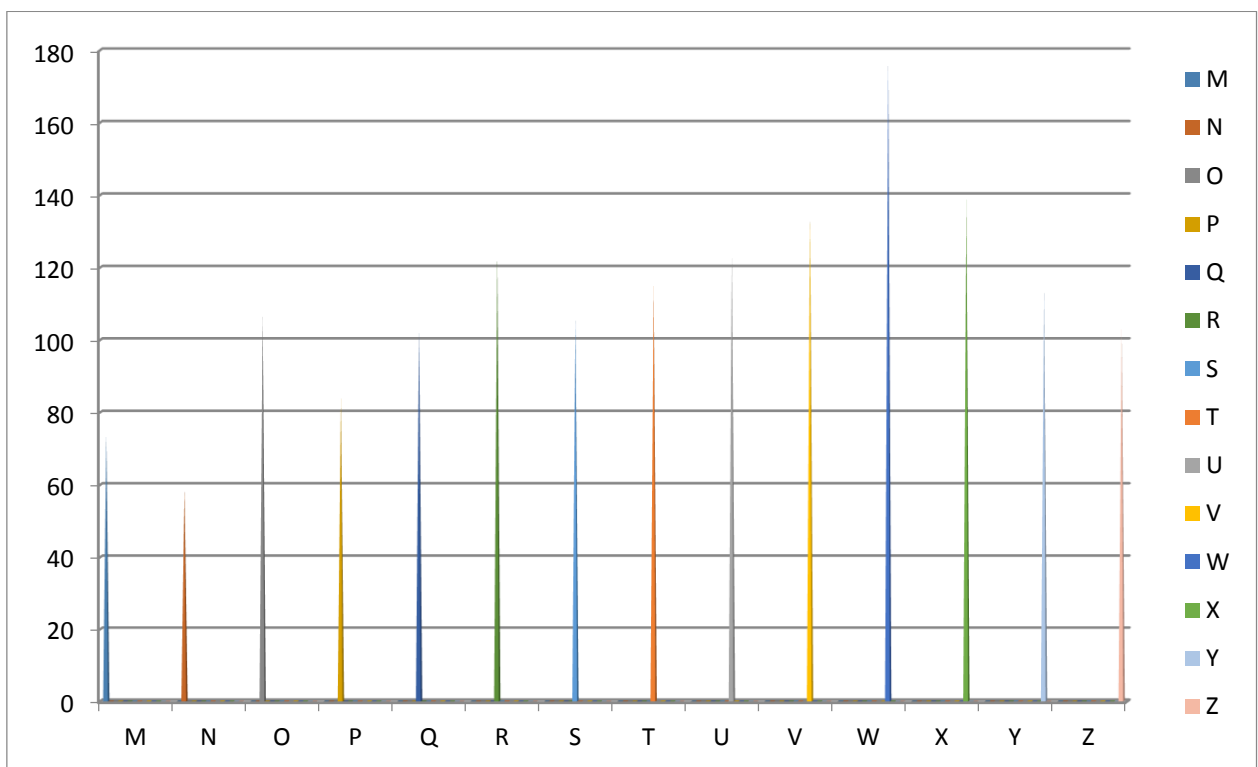


Figure III.16. Graphe de résultat de matching entre les alphabets.

Nous avons aussi comparé trois cas du même alphabet. Les résultats sont représentés dans le tableau 6.

	A0	A1	A2
A0	59	10	17
A1	9.0	60	18
A2	19	18	48

Table III.6. : Résultat de matching entre trois cas du même alphabet.

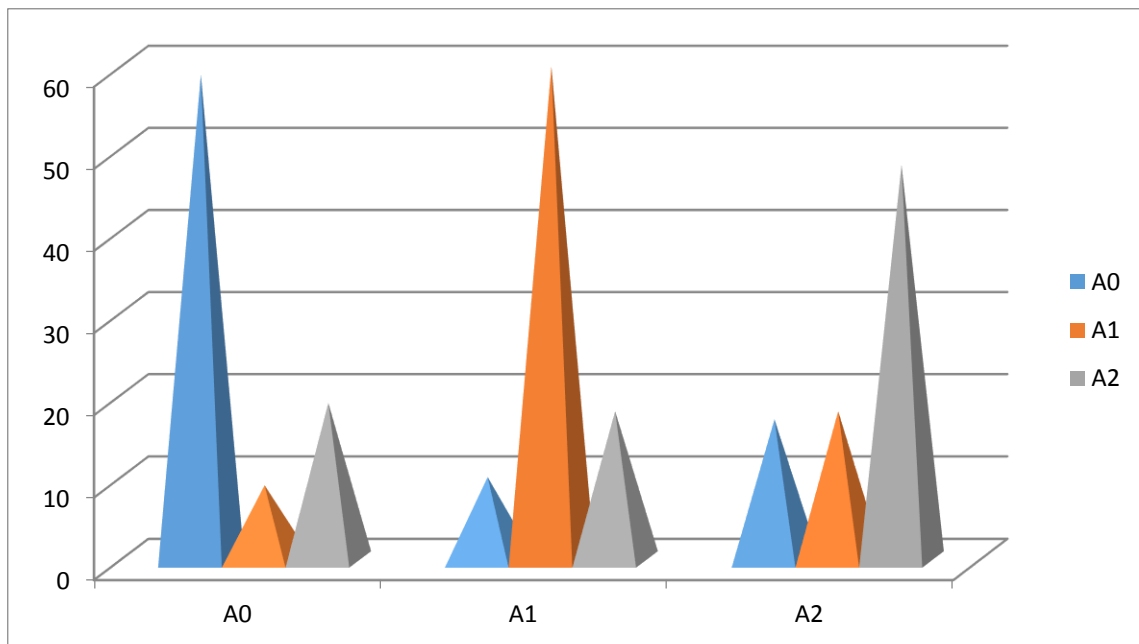


Figure III.17. Graphe de résultat de matching entre trois images de même alphabet.

D'après les résultats du tableau **III.3** ; les paramètres idéals qui donnent des bons résultats de matching et qui peuvent différenciés entre des alphabets similaires sont : de 03 pour Steps per scale octave, 1.6 pour initial gaussian blur, 04 pour feature descripteur size et 08 pour feature descripteur orientation bins et qui conforment la suggérassions de G.lowe

Nous avons calculé par la suite le nombre des points similaires entre chaque singe d'alphabet et les singes des autres alphabets.

D'après les résultats que nous avons obtenue dans les tableaux (**III.4** et **III.5**) nous pouvons dire que :

- la comparaison entre l'alphabet et elle-même permet de donnée le nombre de correspondances le plus élevé.
- La comparaison entre des alphabets semblent similaire permet de donnée un nombre de descripteurs similaires non négligeable.
- la comparaison entre deux alphabets différents permet de donnée un nombre de descripteur presque nul.

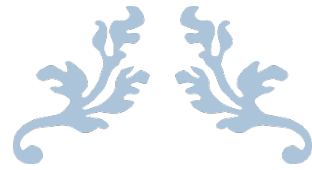
9. Conclusion :

L'algorithme SIFT vient combler en grande partie les limites des méthodes d'extraction de points d'intérêts déjà développées avant lui par Harris, et plus tard par Mohr et Schmid. En effet, il a contribué à l'amélioration des techniques d'extraction d'information dans une image en apportant un algorithme robuste. Il est invariant à la translation, la rotation et changement d'échelle.

Pour toutes ces raisons, nous avons choisi l'algorithme SIFT pour réaliser notre application de reconnaissance de l'alphabet des sourds muets.

La méthode Sift qu'est adapté à notre système, pour détecter et identifier les éléments similaires entre différentes images et de trouver des points-clés qui sont invariants au changement d'intensité, de l'échelle et de rotation.

Nous avons réalisé notre application de reconnaissance de l'alphabet des sourds muets, sous le Platform java script Android pour faciliter sa portabilité sur les smartes phones.



CONCLUSION GENERALE



Conclusion générale :

Nous avons proposé à travers ce projet une application de reconnaissance des alphabets des sourds muets basée sur l'utilisation du descripteur de forme SIFT.

Nous avons exposé dans le premier chapitre, le problème des sourds muets. Cette catégorie utilise les gestes de la main afin de communiquer avec le reste de la société. Dans le deuxième chapitre, nous avons présenté une étude sur les différentes techniques et systèmes déjà proposés dans la littérature.

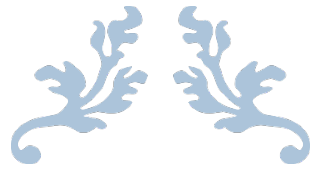
Dans le dernier chapitre, nous avons présenté en détail l'approche SIFT qui permet de détecter les points d'intérêts d'une image qui sont invariants au changement d'intensité, de mise à l'échelle et de rotation, ce qui fait de lui un descripteur très robuste.

Notre application de reconnaissance des alphabets des sourds muets est basée sur deux étapes principales :

- ✓ la première étape consiste à calculer les descripteurs de chaque image de singe d'alphabet
- ✓ la deuxième étape consiste à comparer les descripteurs de l'alphabet à reconnaître avec tous les descripteurs des alphabets de A jusqu'à Z. L'alphabet est identifié par le nombre de points similaires le plus élevé.

L'application a été développée sous Android et a été aussi chargée sur un smartphone pour faciliter l'utilisation de cette dernière.

Les résultats obtenus sont très convainquant mais comme dans tout projet, il reste néanmoins des améliorations concernant les performances du système, en particulier pour les alphabets qui semblent avoir une forme presque similaire.



REFERENCES



Bibliographie :

- [1] Oléron Pierre. «Études sur le langage mimique des sourds-muets. I. Les procédés d'expression ».1952 vol. 52, n°1. pp. 47-81.
- [2] Cuxac C. « La langue des Signes Française (LSF), les voies de l'iconicité, Faits de Langues 15-16 » (2000).
- [3] Larousse de Poche 2006, Paris
- [4] APEDAF asbl et Anne-Charlotte Prévot. « 5 bonnes raisons pour ne plus utiliser sourd-muet».
- [5] Paul Bert « 100 Voix, Pisourd », Dossier Pédagogique, out 2013.
- [6] Magazine de l'Art Pi. Le magazine artistique des Sourds et de la Langue des Signes SÉRIE 2012 (www.art-pi.fr).
- [7] Dr Denise Busquet. « La surdit  de l'enfant Guide pratique   l'usage des parents » Octobre 2005.
- [8] Bellugi U et Klima E.(1979). «the signs of language ».(Harvard University Press),Cambridge.
- [9] Mc Neill D. « Hand and Mind: What Gestures Reveal about Thought » universit  Chocago Press, 1992.
- [10] Cuxac C «Autour de la langue des signes » vol.10 Universit  de Paris 1983.
- [11] Annelies Braffort « Reconnaissance et compr hension de gestes, application   la langue des signes » th se de doctorat de universit  de paris-XI, juin 1996.
- [12] Heinz H., Bauer B. (2000), « Video-Based Continuous Sign Language Recognition Using Statistical Methods », Actes de *IEEE International Conference on Automatic Face and Gesture Recognition*.
- [13] Christiane Fournier « Que recouvre aujourd'hui le terme de bilinguisme L.S.F./Fran ais » Connaissances surdit s mars 2005.

[14] http://www2.univ-paris8.fr/ingenierie-cognition/master-handi/etudiant/projets/site_sam/sourds/info_sourds.htm

[15] Aliyah Morgenstern «Introduction à la langue des signes française, La place du Sourd et de sa langue en France »

[16] Nathalie Monteillard «La langue des signes internationale », Acquisition et interaction en langue étrangère, 15 | 2001, 97-115, Université Paris VIII.

[17] Descours Pauline «L'histoire de la langue des signes française et de la langue des signes brésilienne : quelles influences pour les populations sourdes ? » thèse master 1 université stendhal juin 2011.

[18] Bossard Bruno « Problèmes posés par la reconnaissance de gestes en Langue des Signes» Université Paris XI, juin 2002.

[19] Julien Thomet «Une vue d'ensemble de la reconnaissance de gestes » Université de Fribourg.

[20] Simon Conseil « Suivi tridimensionnel de la main et reconnaissance de gestes pour les interfaces homme-machine », thèse de doctorat, Université Paul Cézanne Aix-Marseille III, mars 2008.

[21] Cadoz C. « le geste canal de communication homme/machine – la communication instrumentale » technique et science informatique, vol.13, 1994.

[22] Maher Mkhinini et Patrick Horain «Reconnaissance de gestes : approches 2D et 3D», Institut Mines-Télécom/Télécom Sud Paris.

[23] Bogdan ionescu, Didier coquin, Patrick lambert « de gestes dynamiques de la main», Laboratoire d'Informatique, Systèmes, Traitement de l'Information et de la Connaissance, France.

[24] N. Dalal et B. Triggs « Histograms of oriented gradients for human detection », Computer Vision and Pattern Recognition, 2005.Computer Society Conference on, vol. 1, 2005.

- [25] Helen Cooper, Eng-Jon Ong, Nicolas Pugeault et Richard Bowden « Sign Language Recognition using Sub-Units », Centre for Vision Speech and Signal Processing, University of Surrey, 2012.
- [26] Xiujuan Chai, Guang Li, Yushun Lin, Zhihao Xu, Yili Tang, Xilin Chen «Sign Language Recognition and Translation with Kinect », Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS Beijing, China.
- [28] David G. Lowe, « Object recognition from local scale-invariant features », Proceedings of the International Conference on Computer Vision, vol. 2, 1999.
- [29] Juan L., Gwun O., « A comparison of sift, pca-sift and surf », International Journal of Image Processing (IJIP), (2009)
- [30] David G. Lowe « Distinctive Image Features from Scale-Invariant Keypoints » , Proceedings of the International Conference on Computer Vision,2004.mo
- [31] demo.ipol.im/demo/my_affine_sift/